HAL
open science

Données numériques et prise en compte de l'apprenant dans les environnements informatiques pour l'apprentissage humain

Julien Broisin, Eric Sanchez, Amel Yessad, Françoise Chenevotot

To cite this version:

Julien Broisin, Eric Sanchez, Amel Yessad, Françoise Chenevotot. Données numériques et prise en compte de l'apprenant dans les environnements informatiques pour l'apprentissage humain. 9ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2019), pp.1-149, 2019, 978-2-901384-01-4. hal-03274910

HAL Id: hal-03274910

https://hal.science/hal-03274910v1

Submitted on 30 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Actes de la 9ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain

Données numériques et prise en compte de l'apprenant dans les environnements informatiques pour l'apprentissage humain

Édités par
Julien Broisin, Eric Sanchez, Amel Yessad, Françoise Chenevotot

Du 04 au 07 Juin 2019
Sorbonne Université, LIP6
Paris, France

La conférence EIAH 2019 a été organisée par le laboratoire LIP6 de l'Université Sorbonne Université (Paris, France), sous l'égide de l'ATIEF (Association des Technologies de l'Information pour l'Éducation et la Formation)

ISBN : 978-2-901384-01-4
Composition : Julien Broisin, Eric Sanchez, Amel Yessad

Introduction aux actes

La conférence Environnements Informatiques pour l'Apprentissage Humain (EIAH) est une conférence internationale francophone bi-annuelle, placée sous l'égide de l'ATIEF (Association des Technologies de l'Information pour l'Éducation et la Formation). L'objectif de la conférence est de permettre à la communauté scientifique francophone et pluridisciplinaire de se rencontrer pour présenter et confronter ses travaux traitant des questions de recherche sur l'éducation en général, et sur le numérique au service de l'éducation en particulier. La 9ème édition de la conférence EIAH (EIAH'2019) s'est tenue à Paris du 4 au 7 juin 2019 sur le campus Pierre et Marie Curie (Jussieu). Elle a été organisée et portée par Sorbonne Université. Elle succède à la conférence EIAH'2017 organisée à Strasbourg en 2017.

Pour cette nouvelle édition, nous avons souhaité aborder la question des **données numériques et la prise en compte de l'apprenant dans les EIAH**. En effet, la collecte automatique de données et leur traitement est aujourd'hui un chantier en plein essor. Dans le champ de recherche des EIAH, ces travaux relèvent de l'analytique de l'apprentissage (*learning analytics*) ou de la fouille de données éducatives (*educational data mining*). Le développement des méthodes statistiques et plus particulièrement celles qui concernent l'intelligence artificielle semblent offrir de nouvelles perspectives à la science des données (*data sciences*) pour l'éducation et la formation. L'une de ces perspectives constitue la thématique scientifique de la conférence EIAH'2019 : la question des données numériques et la manière dont elles peuvent être exploitées pour favoriser la prise en compte de l'apprenant dans les environnements numériques dédiés à l'éducation.

La conférence EIAH'2019 a mobilisé les communautés de chercheurs, praticiens, ingénieurs pédagogiques, entrepreneurs et décisionnaires dans une discussion conjointe mêlant sciences, technologies, pratiques et politiques au service des données numériques pour la prise en compte de l'apprenant. En particulier, l'objectif de la conférence EIAH'2019 était de dresser un état de la recherche francophone sur les approches, outils et méthodes permettant de collecter, stocker et traiter les données d'apprentissage résultant des interactions entre les utilisateurs et les systèmes numériques éducatifs, mais aussi de constituer un panorama des avancées technologiques s'appuyant sur la fouille et l'analyse de ces données pour améliorer l'expérience utilisateur, tant du point de vue des interfaces homme-machine que des scénarios d'apprentissage et des systèmes intelligents proposés aux apprenants. La conférence a ainsi tenté d'adresser les challenges d'aujourd'hui résultant des données d'apprentissage et plus largement des données numériques :

- Dans quelle mesure et comment les données d'apprentissage permettent-elles d'élaborer un modèle de l'apprenant ? Quelles nouvelles connaissances sur les processus d'apprentissage peuvent être extraites à partir de ces données ?

- Dans quelle mesure ces connaissances conduisent-elles à l'élaboration de modèles prédictifs ? Comment peuvent-elles être utilisées pour concevoir des dispositifs adaptatifs ou des dispositifs permettant le suivi de l'apprenant ? Comment peuvent-elles être utilisées pour améliorer l'ergonomie des interfaces ?

- Comment prendre en compte les questions éthiques soulevées par la collecte, le traitement et la réutilisation de ces données ? Quels sont les impacts sur les EIAH des nouvelles lois de régulation du traitement des données numériques ?

Pour tenter d'apporter des éléments de réponse à la formalisation d'un paradigme méthodologique et théorique autour de ces questions dont la liste n'est pas exhaustive, les contributions ont abordé

ces questions sous la forme de présentations de travaux théoriques ou empiriques en informatique ou en sciences humaines (sciences de l'éducation, psychologie, didactique, ergonomie, sciences cognitives, etc.).

Cette nouvelle édition a été marquée par le succès des ateliers présidés par Sandra Nogry (Université Paris 8, France) et Claudine Piau-Toffolon (Le Mans Université, France). Ainsi, une journée complète a été consacrée à la tenue de dix ateliers dont un sur l'enseignement de l'informatique qui a été organisé conjointement avec l'International Federation for Information Processing (IFIP).

Dans le cadre d'EIAH 2019 nous avons également souhaité consacrer une place importante aux jeunes chercheurs avec une session des doctorants organisée par Sébastien Jolivet (LIG, Université Grenoble-Alpes), Jean-Charles Marty (LIRIS, Université Savoie Mont Blanc) et Nathalie Guin (LIRIS, Université de Lyon). Cette session a regroupé onze jeunes chercheurs qui ont pu échanger autour des avancées de leurs travaux de thèse.

La communauté EIAH a également réussi à nouer des relations avec le secteur économique et la Direction du Numérique pour l'Éducation du Ministère de l'Éducation Nationale. Ainsi, une session EdTech organisée par Jean-Marc Labat (LIP6, Sorbonne Univesrité) a permis à des chercheurs, des formateurs et des chercheurs d'échanger sur les questions vives qui mobilisent la communauté EIAH.

La session consacrée aux démonstrations a permis de mobiliser 8 présentations d'outils développés au sein de communautés.

Cette année, nous avons souhaité organiser une table ronde à la mémoire de François Villemonteix décédé brusquement cette année. Nous remercions les personnes qui ont œuvré à l'organisation de cette table ronde.

Emmanuel Sander (Université de Genève) et Kaśka Porayska-Pomsta (University College London, Knowledge Lab) étaient les deux conférenciers invités de cette dition 2019.

EIAH 2019 en chiffres

L'importance d'une conférence s'apprécie avant tout sur des critères qualitatifs mais nous revenons dans le Tableau 1 sur quelques chiffres qui, selon nous, attestent du dynamisme de la communauté EIAH.

La conférence EIAH 2019 aura bénéficié de l'envoi de 80 résumés dont 74 se sont concrétisés sous la forme de 50 propositions d'articles longs, 16 propositions d'articles courts, et 18 propositions de posters. Ces soumissions reflètent bien la pluridisciplinarité de notre domaine de recherche, avec un bon équilibre entre sciences humaines et informatique, ainsi qu'un nombre important de papiers rédigés par des équipes pluridisciplinaires.

Le processus de sélection qui a été mis en place a mobilisé un comité de programme international constitué de 86 chercheurs de France, Suisse, Belgique, Québec, Sénégal, Algérie, Allemagne, Luxembourg. Ce processus a consisté à évaluer, en double aveugle et par 3 relecteurs, les soumissions proposées, puis à offrir la possibilité aux auteurs de répondre aux remarques faites par les experts (phase de rebuttal). Des experts seniors (meta-relecteurs) ont animé une discussion scientifique qui a permis de prendre une décision finale pour chacune des soumissions.

Sur l'ensemble des propositions, 33 (45%) ont été acceptées dans la catégorie (article long, article court ou poster) initialement sélectionnée par les auteurs. 25 propositions (34%) ont été écartées sans possibilité de révision. Pour 11 articles longs, il a été proposé de les présenter sous un format court. Ce sont donc 49 propositions (66%) sur les 74 initiales qui composent les actes de la conférence.

	Soumis	Papiers longs acceptés	Papiers longs acceptés (%)	Papiers courts acceptés	Papiers courts acceptés (%)	Posters acceptés	Posters acceptés (%)	Rejetés	Rejetés (%)
Papiers longs	50	21	42%	11	22%	5	10%	13	26%
Papiers courts	16	0	0%	6	38%	2	13%	8	50
Posters	8	0	0%	0	0%	5	63%	3	38%
Total	**74**	**21**		**17**		**12**		**24**	**32**%

Table 1: Statistiques de l'édition 2019 de la conférence EIAH

Remerciements

Il nous faut ici remercier chaleureusement le comité d'organisation de cette nouvelle édition qui était présidé par Françoise Chenevotot du Laboratoire de didactique André Revuz, France, et Amel Yessad du LIP6, Sorbonne Université, France.

Nous remercions également tout particulièrement les membres du comité de programme qui, à travers les expertises des papiers qui leur ont été soumis, sont les principaux architectes de la qualité de cette conférence.

Julien Broisin, Université Toulouse 3 Paul Sabatier, France
Eric Sanchez, Université de Fribourg, Suisse
Présidents du comité de programme EIAH 2019

Comités de la conférence

Présidents du comité de programme

Julien Broisin	IRIT, Université Toulouse 3 Paul Sabatier, France
Eric Sanchez	Université de Fribourg, Suisse

Comité de programme

Marie-Hélène Abel	Heudiasyc, Université de Technologie de Compiègne
Franck Amadieu	CLLE-LTC, Université de Toulouse
Georges-Louis Baron	Sorbonne Paris Cité, Université Paris Descartes
Monique Baron	LIP6, UPMC Sorbonne Universités
Mireille Betrancourt	TECFA, Université de Genève
Jacques Beziat	Université de Caen
François Bouchet	LIP6, Sorbonne Université
Denis Bouhineau	Université Grenoble Alpes
Yolaine Bourda	LRI, CentraleSupélec
Jean-François Bourdet	Université du Maine
Nabila Bousbia	LMCS, (Ecole nationale Supérieure d'Informatique)
Amel Bouzeghoub	SAMOVAR, Télécom SudParis
Eric Bruillard	EDA
Pierre-André Caron	CIREL, Université de Lille1
Thibault Carron	LIP6, Université Paris 6
Jean-François Ceci	Laboratoire Passages - Université de Pau et des Pays de l'Adour
Bernadette Charlier	Université de Fribourg
Christophe Choquet	LIUM, Le Mans Universit´
Colin De La Higuera	Universit de Nantes
Bruno De Lièvre	Université de Mons
Nicolas Delestre	LITIS, Normandie Université, INSA Rouen
Christian Depover	Université de Mons
Michel Desmarais	Ecole Polytechnique de Montréal
Christophe Després	LIUM, Le Mans Université
Philippe Dessus	LaRAC, Université Grenoble Alpes
Fahima Djelil	Université de Haute Alsace
Hassina El Kechai	TECHNE, Université de poitiers
Catherine Faron Zucker	Université Nice Sophia Antipolis
Serge Garlatti	IMT Atlantique
Sebastien George	LIUM, Le Mans Université
Celso Gonçalves	Universidade Estadual de Minas Gerais
Monique Grandbastien	LORIA, Université de Lorraine
Viviane Gueraud	LIG, Université Grenoble Alpes
Nathalie Guin	LIRIS, Université de Lyon
Nathalie Huet	CLLE-LTC, Université Toulouse 2
Sébastien Iksal	LIUM, Le Mans Université

Dominique Jaccard	University of Applied Sciences Western Switzerland
Stéphanie Jean-Daubias	LIRIS, Université de Lyon
Celine Joiron	MIS, Université de Picardie
Jean-Marc Labat	LIP6, Sorbonne Université
Yacine Lafifi	LabSTIC, Université du 8 Mai 1945
Pierre Laforcade	LIUM, Le Mans Université
Mona Laroussi	IFEF
Elise Lavoué	LIRIS, Université Jean Moulin Lyon 3
Dominique Leclet	Université de Picardie Jules Verne
Marie Lefevre	LIRIS, Université Lyon 1
Anne Lejeune	LIG
Dominique Lenne	Heudiasyc, Université de Technologie de Compiègne
Pascal Leroux	CREN
Paul Libbrecht	German Institute for International Educational Research (DIPF)
Domitile Lourdeaux	CNRS
Vanda Luengo	LIP6, Sorbonne Université
Stéphanie Mailles-Viard Metz	Praxiling, Université de Montpellier
Iza Marfisi-Schottman	LIUM, Le Mans Université
Pascal Marquet	Université de Strasbourg
Jean-Charles Marty	LIRIS
Agathe Merceron	Beuth University of Applied Sciences
Florian Meyer	Université de Sherbrooke
Christine Michel	LIRIS, Université de Lyon
Gaëlle Molinari	TECFA, Université de Genève
Mathieu Muratet	LIP6, Sorbonne Université
Roger Nkambou	Université du Québec À Montréal (UQAM)
Thierry Nodenot	LIUPPA, IUT de Bayonne Pays Basque
Sandra Nogry	Laboratoire Paragraphe, Université Cergy-Pontoise
Lahcen Oubahssi	LIUM, Le Mans Université
Daniel Peraya	TECFA, Université de Genève
Claudine Piau-Toffolon	LIUM, Université du Maine
Marianne Poumay	LabSET, Université de Liège
Chrysta Pélissier	Praxiling, Université de Montpellier 3
Mar Pérez-Sanagustín	IRIT, Université Toulouse 3 Paul Sabatier, PUC
Christophe Reffay	Université de Franche-Comté
Robert Reuter	AES, University of Luxembourg
Nicolas Roland	Université libre de Bruxelles
Margarida Romero	Université Côte d'Azur
Daniel Schneider	TECFA, Université de Genève
Karim Sehaba	LIRIS, Université Lumière Lyon 2
Hassina Seridi	LabGED, Badji Mokhtar University
Franck Silvestre	IRIT, Université Toulouse 1 Capitole
Denise Sutter Widmer	TECFA, Université de Genève
Nicolas Szilas	TECFA, Université de Genève
Pierre Tchounikine	Université of Grenoble
Gaëtan Temperman	Université de Mons

Marc Trestini	LISEC
Lucile Vadcard	LaRAC, Université Grenoble Alpes
Rémi Venant	LIUM, Le Mans Université
Amel Yessad	LIP6, Sorbonne Université

Présidents du comité d'organisation

Amel Yessad	LIP6, Sorbonne Université
Françoise Chenevotot	LDAR, ESPE Lille Nord de France

Membres du comité d'organisation

Mathieu Muratet	LIP6, INS-HEA
Jean-Marc Labat	LIP6, Sorbonne Université
Vanda Luengo	LIP6, Sorbonne Université
François Bouchet	LIP6, Sorbonne Université
Odette Auzende	LIP6, Sorbonne Université
Matthieu Cisel	EDA, Université Paris Descartes
Iryna Nikolayeva	LIP6, Sorbonne Université
Mathieu Vermeulen	IMT Lille Douai

Table des matières

Jeux numériques

Éléments de réflexion autour de la conception du jouet pour un jeu sérieux **1**
Bertrand Marne

Analyser la disparition de la gratuité dans les MOOC au prisme des buts d'accomplissement **13**
Matthieu Cisel

Instrumentation des pratiques d'accompagnement par des formateurs et des conseillers en lien avec l'usage d'un serious game **25**
Pierre-André Caron

Analyse d'usage

L'Orthodyssée des Gram - Stratégies d'étayage et verbalisations du raisonnement métalinguistique lors d'interactions logopédiste - enfant autour d'un jeu grammatical en ligne **37**
Mireille Rodi, Kostanca Cuko and Thierry Geoffre

LabNbook, plateforme numérique support des pédagogies actives et collaboratives en sciences expérimentales **49**
Cédric d'Ham, Claire Wajeman, Isabelle Girault and Patricia Marzin-Janvier

Étude empirique sur la ludicisation de classes au Brésil et en Suisse avec Classcraft **61**
Guillaume Bonvin, Eric Sanchez and Gonçalves Celso

Analytique de l'apprentissage et tableaux de bord

Evaluer les apprentissages : comparaison de deux techniques de diagnostic de connaissance . **73**
Marie Sacksick

Détection Automatique d'Interactions entre Pairs dans les Jeux Sérieux Multi-Joueurs . . . **79**
Mathieu Guinebert, Amel Yessad, Mathieu Muratet and Vanda Luengo

Moodleboard: un tableau de bord dynamique et interactif pour les ingénieurs pédagogiques . **85**
Véronique Sebastien, Didier Sebastien and Dominique Gay

Tableaux de Bord d'aide au Suivi pour la Formation Professionnelle en Ligne: Étude d'Indicateurs et Lignes Directrices **91**
Mohamed Mouaici, Laurence Vignollet, Mael Etienne and Christine Galez

Dispositifs en ligne et hybrides

La (dé-)synchronisation des transitions dans un processus d'évaluation formative exécuté à distance : impact sur l'engagement des étudiants **97**
Jean-Francois Parmentier and Franck Silvestre

Former à la classe inversée en contexte de visioconférence : retours sur deux itérations d'une recherche orientée par la conception **109**
Vincent Barré, Florian Meyer, Matthieu Petit, Nathalie Lefebvre and Cécile Gandon

Comparaison de questions posées et votées en ligne dans le cadre d'une classe inversée **121**
Fatima Harrak, François Bouchet and Vanda Luengo

Analyse d'usage et ingénierie des EIAH

Processus d'évaluation longitudinale (PEL) d'une plateforme pédagogique (LMS) : le cas de LabNbook **133**
Nadine Mandran, Patricia Marzin-Janvier, Maëlle Planche, Aous Karoui and Isabelle Girault

Contribution de différents outils de mesure à l'évaluation des usages d'une plateforme numérique par un processus longitudinal : cas du travail à distance **139**
Maëlle Planche, Nadine Mandran, Isabelle Girault, Patricia Marzin-Janvier, Cedric d'Ham and Claire Wajeman

Modèles et interactions pour générer des événements dynamiques horodatés: le cas de la formation professionnelle à l'organisation des soins infirmiers **145**
Catherine Pons Lelardeux, Galaup Michel, Arthur Ramolet, Catherine Mercadier, Lagarrigue Pierre and Hervé Pingaud

Élaboration d'une méthodologie d'instrumentation pédagogique en contexte universitaire . . **151**
Vincent Bettenfeld, Christophe Choquet, Crétin-Pirolli Raphaëlle and Claudine Piau-Toffolon

Ingénierie des EIAH

Dispositif d'évaluation pour un certificat d'aptitudes cliniques : quelles mesures pour quelles compétences ? **157**
Olivier Catteau, Odile Beyne-Rauzy, Anne Mayere and Nicolas Savy

Génération de scénarios adaptés dans un jeu d'apprentissage selon une approche dirigée par les modèles **169**
Pierre Laforcade and Youness Laghouaouta

Réalité augmentée pour l'apprentissage conceptuel en sciences : quels principes de conception pour les EIAH ? Cas du dispositif DEAPE LEARN en électromagnétisme **181**
Julien Da Costa, Nicolas Szilas and Andreas Mueller

Modélisation de l'apprenant et feedbacks

Étayer dynamiquement une activité de conception expérimentale en biologie : premiers résultats d'expérimentations **193**
Catherine Bonnat and Anne Lejeune

Français langue de scolarisation : Réflexions sur les référentiels de compétences et l'adaptive-learning **199**
Thierry Geoffre and Nathalie Colombier

Modèle d'Exercices et Parcours d'Apprentissage Prenant en Compte le Raisonnement de l'Élève en Mathématiques au Collège **205**
Elann Lesnes-Cuisiniez and Brigitte Grugeon-Allys

Apprentissage de l'informatique

Reconnaissance de Motifs et Répétitions : Introduction à la Pensée Informatique **211**
Yvan Peter, Marielle Léonard and Yann Secq

Objet Tangible ou Simulation Numérique : Deux Situations Équivalentes pour l'Apprentissage de la Programmation ? **223**
Grégoire Fessard, Patrick Wang and Ilaria Renna

Soutien à l'apprentissage de la programmation : conception et évaluation d'un indicateur sémantique **235**
Julien Broisin and Clément Hérouard

Ingénierie des EIAH

Etudier des objets intermédiaires pour analyser les moments-clés de la conception d'un EIAH **247**
Matthieu Cisel

Muse-Logique: Un Système Tutoriel Intelligent pour l'apprentissage du raisonnement logique **259**
Roger Nkambou and Robert Serge

Modèle et outil pour assister la scénarisation des activités pédagogiques orientées RV **271**
Oussema Mahdi, Lahcen Oubahssi, Claudine Piau-Toffolon and Sébastien Iksal

Analyse d'usage

Ingénierie d'un EIAH doté d'étayages adaptatifs **283**
Anne Lejeune, Catherine Bonnat, Viviane Gueraud and Patricia Marzin-Janvier

"Marketplaces" de l'apprentissage en ligne : appréhender les stratégies des instructeurs . . . **289**
David Pontalier and Matthieu Cisel

Écriture numérique à distance : interventions de co-écriture dans un environnement dynamique **295**
Christine Blanchard Rodrigues and Marion Andriamasinoro

Outils en réseau et interactions entre pairs dans la réalisation d'activités collectives à distance **301**
Sonia Androwkha

Adoption des EIAH

TGRIS : Outil de simulation pour la gestion des émotions dans la formation à l'entretien des conseillers pédagogiques **307**
Iza Marfisi-Schottman, Isabelle Vinatier and Elisabetta Bevacqua

Impact de la personnalisation d'une interface graphique sur l'intention d'adoption d'un EIAH **319**
Zeller Arnaud and Mohib Najoua

L'acceptabilité des tablettes tactiles en classe : étude de l'évolution des perceptions d'élèves du secondaire **331**
Julie Mulet, Cécile Van de Leemput, Sandrine Rome, Jan Van der Linden and Franck Amadieu

Métadonnées

Un schéma OWL pour la description d'éléments de formation **343**
Daniel Antelme, Nicolas Delestre and Nicolas Malandain

Nodefr-1 : Métadonnées pédagogiques et Web des données **349**
Yolaine Bourda, Nicolas Delestre, Romuald Verrier, Erwan Le Gall, Yohan Colmant, Marie Peterlongo and Rosa Maria Gomez

Posters

Des jeunes passeurs de pratiques : étude de l'appropriation à partir d'une analyse des données numériques **355**
Laëtitia Pierrot

Experimentation du passage d'une activite debranchee au programme **359**
Adam Michel, Patrice Frison and Daoud Moncef

Un Karaoké pour Entraîner la Prosodie en Lecture **363**
Erika Godde, Gérard Bailly and Marie-Line Bosse

Valorisation des données pour l'amélioration des plateformes d'hébergement de MOOC . . . **367**
Frédérick Bigrat, Mélody Laurent and Cécile Méadel

Apport d'un outil d'intelligence artificielle à l'enseignement-apprentissage des langues **371**
Grégory Miras, Marie Lefevre, Najib Arbach, Louis Rapilly and Théo Dumarski

Évaluation de la satisfaction en formation pour adulte **375**
Laurie Acensio, Frédéric Hoogstoel and Luigi Lancieri

Conception d'un outil d'aide à la compréhension écrite pour les apprenants de français langue étrangère **379**
Yuming Zhai, Gabriel Illouz and Anne Vilnat

Soutien à la régulation cognitive et émotionnelle de l'apprentissage à distance : Intégration d'outils au design pédagogique d'un cours hybride **383**

Bernies Maud and Molinari Gaëlle

La collaboration dans une recherche orientée par la conception : l'exemple de PLAY **387**
Elsa Paukovics

Interactions en jeu dans les situations d'apprentissage instrumenté en Arts Visuels **391**
Ana-Vanêssa Lucena

Prototype de feedback visuel des productions écrites d'apprenants francophones de l'anglais sous Moodle **395**
Nicolas Ballier, Thomas Gaillat and Erin Pacquetet

Session
Jeux numériques

Éléments de réflexion autour de la conception du jouet pour un jeu sérieux

Bertrand Marne[1]

[1] ICAR UMR 5191, Université Lumière Lyon 2, France
bertrand.marne@ens-lyon.fr

Résumé. Les jeux sérieux destinés à l'apprentissage peuvent proposer plus qu'un supplément de motivation aux EIAH. Certains auteurs avancent que ceux fondés sur des jouets en transforment également les aspects pédagogiques. Le jouet est alors défini comme l'artefact « avec » lequel joue plutôt que celui « auquel » on joue. Ce sont les qualités exploratoires et expérimentales suggérées par « avec » que ces auteurs proposent d'exploiter. Dans la perspective de questionner la place du jouet dans un jeu sérieux, nous commençons par interroger leur conception. Dans cet article, nous proposons trois études de terrain menées en situation écologique. Elles visent à identifier des patrons de conception pour l'intégration de jouets dans les jeux sérieux. Nous discutons les résultats des trois études et proposons des fondations de patrons de conception pour la construction d'une simulation, et pour la conception des interactions ludiques avec elle.

Mots-clés. Jeu sérieux, ingénierie, jouet, micro-monde, patrons de conception

Abstract. Serious learning games can provide more than just an extra engagement to ITS. Some authors argue that those based on toys also transform their pedagogical aspects. The toy is then defined as an artifact you play "with" rather than the artifact you play "at". The exploratory and experimental qualities suggested by "with" are pointed out by these authors. In order to question serious game based on toys, we began with the investigation of their design. In this article, we propose three field studies conducted in an ecological situation. They aim to identify design patterns for the integration of toys into serious games. We discuss the results of the three studies and propose design pattern foundations for the design of a simulation, and for the design of playful interactions with it.

Keywords. Serious Game, design, toy, microworld, design patterns

1 Introduction

Dans les nombreuses acceptions des jeux sérieux (JS) à vocation d'apprentissage, nous choisissons de focaliser cette étude sur les JS cherchant à obtenir une motivation intrinsèque des apprenants-joueurs plutôt que les JS ayant recours aux mécanismes de gamification. Parmi ceux-ci, nous visons plus précisément les JS qui ne cherchent pas seulement à augmenter la motivation et l'engagement à utiliser un EIAH. Mais, ceux dans lesquels les interactions sont utilisées pour proposer de l'exploration et de

l'expérimentation afin de tenter de proposer un apprentissage plus constructiviste que transmissif aux apprenants-joueurs. Ryan et al., en s'appuyant explicitement sur les travaux de Schell [18], et dans une réinterprétation implicite du concept de micro-mondes [14], proposent de centrer la conception de cette catégorie de JS sur la construction d'un jouet systémique [16]. Nos travaux s'inscrivent dans cette problématique générale qui interroge les apports de la notion de jouet dans les méthodes de conception des JS.

Dans cet article nous décrivons nos recherches, constituées de trois études de terrain, pour extraire les fondations de patrons de conception [1, 12] de JS articulés sur ces jouets systémiques. Nous commençons par présenter les travaux antérieurs qui nous ont servi de base pour définir les concepts de notre problématique. Dans une seconde partie, nous présentons la méthodologie de recherche que nous avons choisie et les biais qu'elle provoque pour mener les trois études que nous décrivons et discutons dans la troisième partie. Dans la conclusion, nous faisons le bilan de ces travaux et évoquons nos futures recherches sur ces questions.

2 Le jouet comme fondement de jeux sérieux constructivistes

Certains auteurs comme Bogost et al., Ryan et al. considèrent que les JS pour l'apprentissage ne doivent pas viser en premier lieu la motivation, mais un apprentissage expérientiel [2, 4, 16]. En effet, ils montrent que les JS de ce type proposent des expériences d'apprentissage plus constructivistes, qui par ailleurs induisent une motivation intrinsèque. Alors que les autres, souvent articulés sur les principes de gamification, apportent plutôt une motivation extrinsèque [2]. Or, la gamification peut aussi avoir des effets négatifs sur l'apprentissage, comme la baisse de la motivation intrinsèque [6] ou la course à la récompense [5].

Les travaux de Ryan et al. [16] en utilisant la « *Procedual Rethoric* »[1] proposée par Bogost [4] et le « *Deep Conceptual Learning* »[2] proposé par Gee [7] invitent à fonder la conception d'un JS sur l'utilisation d'un jouet grâce à la « *Lens of the Toy* » de Schell [18]. Selon Ryan et al., un bon jouet est un système complexe qui propose de nombreuses affordances ludiques et qui sollicite les capacités cognitives de reconnaissance de structures, de raisonnement stratégique et de résolution de problèmes [16]. En cela, ils enrichissent la notion de jouet proposée par Schell. Ce dernier définit le jouet comme « le jouet est artefact *avec* lequel on joue ». Et non l'artefact *auquel* on joue, qui est lui plutôt le jeu : l'objet-jeu se décrit alors comme l'ensemble de règles (obstacles et buts) qui encadrent l'utilisation du jouet [18].

Ryan et al., sans l'expliciter dans leurs travaux, retrouvent des notions proches des micro-mondes, puisqu'ils considèrent qu'un jouet, au cœur d'un JS éducatif, doit être un modèle concret des systèmes qui dirigent le domaine à enseigner [14, 16]. Et ainsi, ils décrivent le jouet comme un outil qui invite, par l'action de jeu, à l'exploration des systèmes, à la reconnaissance de ses structures et à l'expérimentation pour essayer de contrôler, puis de maîtriser ces systèmes. C'est très proche des micro-mondes, ces EIAH qui mettent les utilisateurs dans une situation d'autonomie afin de développer des connaissances par la découverte, la conception d'hypothèses et l'expérimentation de celles-ci [14]. Ainsi, lorsqu'un apprenant-joueur manipule un micro-monde ou un

1 Capacité d'un processus interactif à porter un message
2 Un apprentissage profond est émergent et se fait intrinsèquement par l'activité de jeu

jouet selon Ryan et al., il cherche à comprendre les règles et les équilibres existant dans les systèmes que le jouet présente, et ce de deux manières. Soit en l'explorant : il provoque des causes et en observe les conséquences pour en induire des règles. Soit en expérimentant : il cherche à obtenir des conséquences en multipliant les possibilités de causes pour vérifier une ou plusieurs règles, ses hypothèses, par abduction [15].

Pour construire ces JS fondés sur des jouets/micro-mondes, Ryan et al. proposent, sans la vérifier expérimentalement, une méthode qui ressemble beaucoup à certains de nos travaux (légèrement antérieurs) : les 6 facettes de la conception des jeux sérieux [10, 11]. Les 6 facettes forment un modèle formel de la conception des JS qui a été évalué expérimentalement [11] et qui a été conçu pour faciliter la collaboration entre les parties prenantes de cette conception. La méthode de Ryan et al. se décompose en 5 étapes chronologiques (contrairement aux 6 facettes) :

1. Identifier les mécaniques et dynamiques impliquées dans les systèmes du domaine que l'on cherche à enseigner. Cette étape correspond de façon remarquable à deux facettes. La facette 1 : la définition des *Objectifs pédagogiques* qui correspond à la modélisation du domaine à enseigner et la facette 2 : la *Simulation du domaine* qui décrit les modèles formels des systèmes à l'œuvre dans le domaine à enseigner.
2. Présenter le système de façon à faciliter la reconnaissance de ses structures sous-jacentes. Ce qui rejoint les concepts décrits dans la facette 2 de *Simulation du domaine*.
3. Fournir un outil pour que l'apprenant-joueur ait un contrôle incarné et ludique. Soit un outil à la fois simple et aux affordances multiples pour une interaction fine et ludique avec le système. Il s'agit exactement de ce que propose la facette 3 : *Interactions avec la simulation* qui concerne la conception de l'interface ludique avec la simulation (facette 2) proposée à l'apprenant-joueur.
4. Ajouter des buts pour créer des étapes dans la découverte du système par les apprenants-joueurs. C'est en correspondance avec la facette 4 : *Problèmes et progression* qui concerne les défis et les obstacles à proposer aux apprenants-joueurs afin de les faire progresser à la fois dans le jeu et dans les compétences recherchées.
5. Fournir le nécessaire pour permettre aux apprenants-joueurs de partager leurs nouvelles compétences. Ce qui ressemble à la facette 6 : *Conditions d'utilisation* qui concerne notamment l'élaboration des conditions d'institutionnalisation des apprentissages.

Assez logiquement, compte tenu de leurs propositions en opposition aux techniques de gamification, les étapes proposées par Ryan et al. ne reprennent pas les contenus de la facette 5 : *Décorum* correspondant à la conception des éléments de gamification d'un JS.

En reprenant la définition du jouet donnée dans la « *Lens of the Toy* » de Schell [18], nous avons constaté que parmi ces 5 étapes, seules les trois premières sont centrées sur l'élaboration du jouet. Les deux suivantes sont centrées sur la construction de l'objet-jeu.

Nos travaux ont pour objectif de questionner les JS fondés sur des jouets systémiques. Et notamment, d'explorer la possibilité que ces JS, comme les micro-mondes [14], permettent aux apprenants-joueurs la découverte par l'exploration, la compréhension et la modélisation des systèmes qui régissent les domaines enseignés par des raisonnements inductifs. Mais aussi, l'expérimentation pour vérifier et conforter les modèles hypothétiques par des raisonnements abductifs.

Dans cette perspective, nous avons commencé nos travaux par interroger la conception des JS de cette catégorie. Notamment, nous avons cherché à mettre à l'épreuve sur le terrain les propositions méthodologiques de Ryan et al. et tenté d'extraire des éléments de patrons de conception [1, 12] de ces situations.

3 Propositions méthodologiques

Pour proposer et mettre à l'épreuve des méthodes de conception de JS fondés sur des jouets, nous avons choisi de nous inscrire dans une démarche de recherche écologique proche de la recherche-action. Ce choix est à la fois guidé par des contraintes pratiques et des choix épistémologiques.

À propos de ces derniers, nous cherchons à extraire des éléments de patrons de conception pour aider à la fabrication des JS centrés sur des jouets. Dans le domaine des JS, les méthodes récentes d'extraction de patrons de conception s'intéressent à leur extraction à partir de JS déjà existants [9]. Ces méthodes d'extraction permettent d'identifier les bons résultats et d'en inférer des patrons de conception, mais ne permettent malheureusement pas d'identifier les obstacles et problèmes survenus lors de la conception, et les moyens de les résoudre. C'est pourquoi, en complément, nous avons choisi de travailler sur une extraction de patrons *durant* le processus de conception lui-même. Ceci implique une approche très écologique, comme celles de la recherche-action, en suivant les concepteurs aussi souvent que possible.

De plus, comme les JS fondés sur les jouets sont rares, nous avons aussi choisi d'être partie prenante, pour aider les concepteurs que nous suivions en leur apportant, quand c'était nécessaire, des éléments méthodologiques élaborés à partir des expériences antérieures et de l'état de l'art. Ainsi, nous avons choisi de nous inscrire dans une épistémologie proche de la recherche collaborative orientée par la conception (ROC) [17], elle-même inspirée de la Design Based Research [19].

Ce choix nous a également permis de surmonter l'une des principales contraintes pratiques de ce type de recherche : trouver un terrain d'étude dans lequel la recherche-action peut être suffisamment longue pour permettre d'observer et d'agir sur des situations variées, sans que le risque économique pesant sur les concepteurs soit grand. Dans ce contexte proche de la ROC [17], nous avons choisi le cadre pédagogique d'un enseignement universitaire sur les JS.

Bien que ces choix méthodologiques aient rendu possible une étude d'une certaine ampleur par sa durée et son nombre de participants, nous avons dû tenir compte d'un certain nombre de biais. Par exemple, le fait que nous soyons partie prenante, ou le fait que ce public qui découvre les JS est peut-être assez différent du public que nous pourrions cibler avec des patrons de conception.

Pour prendre en compte ces biais, au-delà de l'analyse des JS produits [9], de celle des prototypes conçus à chaque itération de la production, et des observations directes des méthodes employées, nous avons aussi collecté des réponses à des questionnaires et à des entretiens proposés à trois occasions : avant le début de la conception et des apports méthodologiques, pendant la conception, et après la production du JS final et son évaluation scolaire. Lors de toutes ces observations, notre collecte d'informations a été guidée par plusieurs objectifs. Nous avons cherché à savoir quelles étaient les préconceptions des étudiants sur les méthodes de conception de JS et comment et pourquoi ces méthodes ont pu changer au cours du projet, afin d'avoir pour chacun

des participants une idée des connaissances antérieures, des choix et des sources d'information qui ont participé à l'évolution de celles-ci au cours des travaux.

Nous avons travaillé en trois étapes et avec plusieurs groupes d'étudiants. Progressivement, nous avons amélioré les aides méthodologiques fournies en nous appuyant sur leurs difficultés à appréhender et à mettre en œuvre la notion de jouet.

4 Résultats de l'étude de terrain et discussion

Cette étude de terrain s'est déroulée sur une période de 3 ans (une étape par an) auprès de 104 étudiants ayant développé 23 prototypes de JS par groupes de 3 à 12 personnes. 58 de ces étudiants étaient en Master 2 (M2) de journalisme dans un cursus axé sur les pratiques numériques journalistiques. Les 46 autres étudiants sont en Diplôme Universitaire (DU) de level design (niveau bac+3).

Les étudiants ont travaillé à la conception et réalisation de prototypes dans le cadre de leur formation aux JS en général, et aux newsgames [3] en particulier pour les étudiants journalistes. Ces projets sont menés de façon agile, sur un semestre avec entre 12 et 24 h de cours méthodologiques auxquels s'ajoutent de l'accompagnement.

Les prototypes de JS issus de ces travaux restent pour la plupart assez rudimentaires, aussi bien dans leurs aspects pédagogiques, que ludiques, voire du mélange de ces deux aspects. L'objectif n'était pas de faire concevoir un JS idéal, mais en produisant un premier prototype de JS, de permettre aux étudiants de se familiariser avec ce type de média. Ainsi, nous ne décrivons ni discutons des qualités ludiques ou pédagogiques des projets portés par ces étudiants. Nous nous limitons à décrire et discuter la présence, la conception et l'intégration de jouets dans les JS produits.

Pour travailler, les étudiants bénéficiaient d'un bagage méthodologique et logiciel fourni lors des premiers cours. Ces contenus ont varié entre les étapes, en fonction des résultats précédents. Les trois études sont présentées dans les sous-sections suivantes.

4.1 Première étude : le jouet présenté comme une des méthodes de conception

La première étude menée a réuni 26 étudiants de M2 de journalisme, regroupés dans 6 projets de JS (dont un groupe de 9, pour un projet de plus grande ampleur) dont ils ont choisi le sujet librement.

Les étudiants ont bénéficié d'une introduction aux JS fondée sur *Journalism at Play* de Bogost [3] et mettant en avant les newsgames comme des médias systémiques et expérientiels. Puis, la méthodologie des 6 facettes de la conception des jeux sérieux et leurs patrons de conception [10] leur a été expliquée parce qu'elle décompose la conception de façon intelligible, en particulier pour des novices. Au cours de ces apports méthodologiques, la notion de jouet a été définie et présentée comme une bonne pratique pour combiner les facettes de la simulation du domaine et de l'interface ludique avec cette simulation.

Pour accompagner la réalisation technique, des outils auteurs ont été proposés avec une grande liberté pour les choisir. Nous les classons en plusieurs catégories : un outil de conception (Legadee [8]), des outils fabrication de narrations interactives

(Twinery[3] et Klynt[4]), un outil de réalisation de jeux de rôles (RPG Maker[5]), un outil auteur de JS point & clic (eAdventure [13]), des outils de réalisation de jeux sans programmation (GDevelop[6] et Construct[7]).

Les 6 travaux qui en résultent sont synthétisés dans le Tableau 1. La colonne *Genre* décrit à quel genre vidéoludique appartient le JS, la colonne *Ambition* décrit le but que les étudiants voulaient atteindre avec. La colonne *Principales mécaniques* décrit les interactions et les moyens ludiques principaux mis en œuvre dans le JS. Enfin, la colonne *Méthode d'apprentissage visée* décrit comment les étudiants projetaient que se fasse l'institutionnalisation des savoirs dans leurs JS.

Leurs travaux n'ont pour la plupart pas dépassé le stade de prototypes et à ce titre n'ont pu être évalués que lors de playtests[8] rudimentaires. Au cours de ceux-ci, institutionnalisation des avoirs liés à l'utilisation du JS n'a pas pu être mesurée.

Tableau 1. Synthèse des JS produits lors de l'étude 1

Nom du JS	Genre	Ambition	Outils auteur	Principales mécaniques	Méthode d'apprentissage visée
ISIS the End[9]	Narration interactive	Faire comprendre les mécanismes de la radicalisation	Klynt	À chaque étape un choix détermine l'étape suivante	Exploration + texte de remédiation quand il y a erreur
Citizens of Nowhere	Narration interactive	Découvrir les épreuves endurées par les réfugiés	Twinery	À chaque étape un choix détermine l'étape suivante. Des choix font varier des jauges déclenchant des événements	Exploration + expérimentation des hypothèses
Dealopolis	Narration interactive	Comprendre l'articulation drogue-corruption	RPG Maker	Collecte d'indices dans une carte. Puis QCM.	Exploration (indices, réponses au QCM) + textes de remédiation
Femmes à travers le monde	Associations d'objets	Sensibilisation aux risques d'agression pour les femmes	Construct	Association de choix en fonction de la situation	Le joueur perd toujours : Exploration + texte de remédiation
Reportage Ebola	Narration interactive	Comprendre les conditions du journalisme africain	Twinery	À chaque étape un choix détermine l'étape suivante	Exploration + textes de remédiation
Vélo en ville	Narration interactive	Sensibiliser aux dangers de la circulation pour les cyclistes	Twinery	À chaque étape un choix détermine l'étape suivante	Exploration + textes de remédiation

Notre première observation est que les projets des étudiants sont quasiment tous des projets de narration interactive, dans lesquelles les interactions se résument le plus souvent à faire des choix qui orientent la suite de l'histoire (des choix souvent binaires et non cornéliens : entre bon et mauvais choix). Les étudiants avaient prévu que l'institutionnalisation des savoirs se ferait soit en faisant un mauvais choix, lisant

3 Twinery : https://twinery.org/ (consulté le 26/04/2019)
4 Klynt : http://www.klynt.net/ (consulté le 26/04/2019)
5 RPG Maker : https://www.rpgmakerweb.com/ (consulté le 26/04/2019)
6 Gdevelop : https://gdevelop-app.com/ (consulté le 26/04/2019)
7 Construct : https://www.scirra.com/ (consulté le 26/04/2019)
8 Technique d'évaluation des jeux très employée dans l'industrie des jeux vidéo.
9 Isis the End est de plus grande ampleur (projet de l'année des étudiants). Il est d'ailleurs disponible en ligne : http://www.isistheend.com/ (consulté le 26/04/2019)

les textes de remédiation, et en recommençant ; soit en lisant des « indices » fournis avant les choix. Un projet se distingue : *Citizens of Nowhere* parce qu'en plus de la narration interactive, il ajoute un système de jauges (santé, argent) qui varie en fonction des choix et qui peut ré-aiguiller la narration. Tardivement dans la conception, les étudiants avaient voulu mettre en place une petite simulation au cœur de leur JS très narratif. Leur souhait était d'utiliser la notion de jouet systémique pour provoquer, en plus de l'exploration, un apprentissage par l'expérimentation. Nous avons d'ailleurs retenu l'idée de *jauge* pour modéliser les variables à l'œuvre dans la simulation des systèmes, car d'après nos entretiens finaux, il a semblé parlant pour l'ensemble des étudiants.

À l'exception du groupe de *Citizens of Nowhere*, nous avons constaté que les étudiants n'avaient pas fait le pas de côté suggéré par Bogost [3]. Il propose d'éviter les méthodes journalistiques narratives pour fournir des contenus plus expérientiels et systémiques, dans lesquels l'apprentissage se fait par des combinaisons d'exploration et d'expérimentation. Du point de vue méthodologique, nous n'avions donc pas réussi à proposer des moyens qui permettaient aux étudiants d'envisager le jouet comme solution valide pour construire des expériences avec ce type d'apprentissage, plutôt que des JS fondés sur des récits journalistiques.

Nous pensons également que le choix des outils auteurs les plus simples à prendre en main (Twinery, Klynt et RPG Maker) par les étudiants a joué un rôle important dans le choix des contenus puisqu'ils sont conçus pour les narrations interactives. Nous avons utilisé ces observations pour notre seconde étude.

4.2 Seconde étude : le jouet placé au centre des méthodes de conception

Elle s'est déroulée d'une part auprès de 32 étudiants en M2 de journalisme (regroupés dans 8 projets de JS) et d'autre part auprès de 23 étudiants en DU de level design (regroupés dans 5 projets de JS).

Pour les journalistes, la mise en œuvre a été fondée sur les mêmes principes que l'étude précédente avec les changements suivants : le concept de jouet a été placé au centre des enseignements méthodologiques, sans pour autant en proposer une méthode de conception détaillée. Son articulation avec les *systèmes* décrits par Bogost [3] est réifiée grâce au principe des *jauges*, retenu à la précédente itération de l'étude. Elles permettent de modéliser les variables systémiques dont le jouet peut faire varier les valeurs. De plus, afin d'éviter l'influence trop forte des outils auteurs dans le genre de jeu, il est demandé, d'une part, de faire le premier des deux playtests avec un prototype non numérique (jeu physique, jeu de plateau, etc.). D'autre part, pour le second prototype, les outils auteurs les plus versatiles (GDevelop et Construct) sont proposés en priorité. Les travaux de ces groupes sont présentés dans Tableau 2.

Les travaux présentés dans le Tableau 3 sont ceux des groupes d'étudiants level designers pour lesquels nous avons introduit plusieurs différences avec les journalistes. À la place des travaux de Bogost sur les newsgames, un travail de réflexion a été mené sur *The Lens of the Toy* de Schell avec la recherche de bons exemples de jouets dans les jeux vidéo populaires. De plus, le domaine a été imposé : les étudiants devaient concevoir un JS de niveau collège sur le système immunitaire.

Auprès des étudiants journalistes (Tableau 2), nous observons que la proposition de JS a été plus diversifiée que lors de l'étude précédente. Il y a encore deux JS narratifs. Mais cette fois, cherchant à intégrer un jouet en fin de conception, ces deux groupes

ont ajouté un système de jauges incrémentées/décrémentées en fonction des choix. À la manière des tests dans les magazines estivaux, elles classent l'apprenant-joueur dans une catégorie à la fin du JS. D'autres JS conçus lors de cette étude n'ont pas d'aspect systémique. Le jeu de l'oie, le casse brique et le QCM humoristique ne proposent pas d'expérimentation aux apprenants-joueurs, seulement la lecture d'informations. Dans *La Catalogne* et *Make Hummus not Walls* les étudiantes ont pourtant intégré un « jouet » (au sens de Schell), mais totalement décorrélé des aspects systémiques du domaine à découvrir (donc, pas au sens de Ryan et al.). Seuls les JS *Enfance dans le monde* et *Jeu de l'Amazonie* proposent aux joueurs de comprendre comment fonctionnent des systèmes en leur permettant d'explorer, puis expérimenter. Pour le premier des JS, cet aspect systémique a été introduit par un système de jauges (santé, éducation, plaisir, etc.), pour le second par une petite simulation qui est interrogée à chaque action du joueur.

Tableau 2. Synthèse des JS produits lors de l'étude 2 avec des étudiants journalistes

Nom du JS	Genre	Ambition	Outils de développement	Principales mécaniques	Méthode d'apprentissage visée
Enfance dans le monde	Jeu d'adresse	Sensibiliser aux difficultés des enfants des pays pauvres	GDevelop	Se déplacer en évitant de se faire attraper	Exploration + expérimentation
Jeu de l'Amazonie	Stratégie au tour par tour	Comprendre les équilibres écologiques et économiques de l'agriculture en Amazonie	GDevelop	Tour par tour : choix (déforestation, plantation, jachère, récolte, etc.) en fonction de ses ressources	Exploration + expérimentation
La Catalogne	Jeu d'action	Comprendre les enjeux de l'indépendance de la Catalogne	Construct	Se déplacer avec agilité en collectant des informations	Lecture des informations
Make Hummus not Walls	Casse brique	Connaître des faits sur la Barrière de séparation d'Israël	GDevelop	Récupérer des messages par agilité	Lecture des informations
Make Twitter Great Again	Suite de QCM humoristiques	Faire connaître les tweets du Président des USA	Klynt	À chaque étape un QCM à remporter ou perdre	Exploration + textes de remédiation
Rentabilise Neymar	Narration interactive	Comprendre les transferts et le management des joueurs de football	Twinery	À chaque étape un choix détermine l'étape suivante + Système de jauges → ajouté à la fin	Exploration + texte de remédiation + expérimentation
Garde la pêche	Narration interactive	Apprendre à avoir une alimentation équilibrée à faible impact écologique	Twinery	À chaque étape un choix détermine l'étape suivante + Système de jauges → ajouté à la fin	Exploration + texte de remédiation + expérimentation
Un toit sur Lyon	Jeu de l'oie	Apprendre les démarches pour se loger	GDevelop	Déplacement sur un plateau + Questions à chaque étape	Exploration + textes de remédiation

Quatre des huit JS produits sont non systémiques, même si deux d'entre eux contiennent des « jouets » (pour coller aux demandes de l'enseignant). Dans les deux JS narratifs, les étudiants sont parvenus à contourner les contraintes des outils auteurs (avec des pions réels, à comptabiliser à la fin de la partie) pour introduire un aspect systémique fondé sur des jauges. Les deux autres JS s'approchent plus encore de la notion de jouet selon Ryan et al., en intégrant un système avec lequel il est possible de

jouer pour en expérimenter le fonctionnement et les limites. Dans ces deux JS, les playtests ont montré que des apprenants-joueurs exploraient et expérimentaient les jouets-systèmes sans nécessairement suivre les buts des JS.

Nous avons constaté que les outils auteurs les plus versatiles étaient très difficiles à pendre en main par des étudiants peu familiers des outils numériques. C'est ce qui a conduit certains groupes, après avoir fait un JS non numérique systémique, à se résoudre à réaliser un JS numérique non systémique, conçu avec des outils plus simples à appréhender comme Twinery ou Klynt.

Au contraire, les étudiants en level design (Tableau 3) étaient pour la plupart familiers du numérique et de la programmation et se sont facilement approprié l'outil auteur Construct à partir duquel tous les projets ont été menés.

Tableau 3. Synthèse des JS produits l'étude 2 avec des étudiants level designers.

Nom du JS	Genre	Outils de développement	Principales mécaniques	Méthode d'apprentissage visée
Deficience	Tower defence	Construct	Association de défenses à des vagues de microbes	Exploration + expérimentation
Heal and Help	Jeu de plateau	-	Choisir des défenses pour contrer des événements infectieux	Exploration + expérimentation
Immune Defence	Tower defence	Construct	Association de défenses à des vagues de microbe	informations textuelles + entraînement
Immunis	Jeu de tir	Construct	Poursuivre et détruire les microbes ciblés par anticorps	informations textuelles + entraînement
Infiltrhex	Jeu de stratégie au tour par tour	Construct	À chaque tour, choisir des actions pour les microbes afin de traverser les défenses	Exploration + expérimentation

Nous observons que les JS réalisés appartiennent à des genres assez proches : le genre *tower defense* est assez proche du genre *jeu de stratégie au tour par tour*. Nous distinguons néanmoins deux familles de choix de méthodes institutionnalisation des savoirs, indépendantes des genres : d'une part, les JS dans lesquels il est prévu que l'apprentissage se fasse par l'apport d'informations (tutoriels) et son renforcement est prévu par l'entraînement. D'autre part, ceux dans lesquels il est prévu que l'apprentissage se fasse en explorant, puis en expérimentant différentes solutions.

Nous avons constaté que l'ensemble des JS proposés sont systémiques, et nous pensons que cela est lié à trois causes : en premier lieu, la nature du sujet imposé aux étudiants (*système* immunitaire). En second lieu, l'absence de la barrière technique de la maîtrise de l'outil auteur. Et enfin en dernier lieu, grâce à la culture vidéoludique des étudiants, associée au travail d'inventaire qui a été fait de systèmes et des jouets dans les jeux vidéo populaires. Cependant, seuls les JS qui proposent de l'exploration semblent être vraiment structurés autour d'un jouet au sens de Ryan et al. En effet, ils sont les seuls à proposer des artefacts que les candidats aux playtests ont pris plaisir à manipuler indépendamment des instructions et des buts donnés dans les JS.

Nous avons tiré 3 observations principales de l'ensemble de cette étude 2 :

- L'introduction du concept de jauge semble faciliter la conception d'artefacts systémiques (Bogost [3]), mais pas forcément de jouets (Ryan et al. [16]).
- Le manque de maîtrise des outils auteurs versatiles (Construct, Gdevelop) est un obstacle à l'implémentation de jouets dans les prototypes.
- Les sujets de JS qui paraissent plus systémiques facilitent la création de jouets. Donc, il nous faut proposer une méthode d'identification des aspects systémiques dans tous les sujets (étapes 1 et 2 de Ryan et al. [16]).

4.3 Troisième étude : avec une méthode de conception détaillée du jouet

Cette troisième étude a impliqué 23 étudiants de DU de level design et 11 étudiants en M2 de journalisme, cette fois répartis équitablement sur 3 projets de JS avec un thème imposé : L'Europe politique. Nous avons été convaincus d'instaurer ce mélange d'étudiants par le changement de stratégie d'étudiants journalistes : ils avaient d'abord conçu des prototypes de JS non numériques, plutôt systémiques. Puis, finalement, ils avaient décidé de les abandonner face aux problèmes d'implémentation dans les outils auteurs (Construct, Gdevelop).

Comme nous avions (volontairement) imposé un thème qui semble peu systémique, nous avions aussi fourni aux étudiants une méthode détaillée de conception de jouet, fondée d'une part sur Ryan et al. [16] articulée avec les 6 facettes [11]. Et d'autre part sur l'observation des difficultés rencontrées lors des deux premières études. Elle décrit le travail qui doit être opéré dans les facettes de la *Simulation du domaine* et des *Interactions avec la simulation*.

Le Tableau 4 contient le texte de l'une des fiches méthodologiques de synthèse fournies aux étudiants. Le Tableau 5 synthétise les JS produits lors de cette étude 3.

Tableau 4. Récapitulatif de la méthode de conception des jouets proposée aux étudiants

Facette n°2 : Concevoir la simulation	Facette n°3 : Construire les interactions

1. **Choisir** un sujet systémique
2. Décrire et **modéliser** les systèmes
  1. Identifier les **paramètres** du système → **jauges**
  2. Identifier les **éventements** qui font varier les paramètres
  3. Identifier comment les paramètres **interagissent** entre eux
  4. Identifier les **règles** qui régissent le système
  5. **Simplifier** en se concentrant sur ce que l'on veut faire comprendre aux utilisateurs/ices
3. **Programmer** la simulation des systèmes
  ○ Système à base d'équations
  ○ Système à base de règles
  ○ Ou les deux
4. **Tester** la simulation pour vérifier...
  ○ Sa fiabilité
  ○ Sa simplicité
	1. Identifier comment le **hasard** peut jouer dans le système
	2. Imaginer une **métaphore**
	  1. Imaginer plusieurs **interactions** potentielles pour chaque paramètre
	  2. Imaginer plusieurs **feedbacks** possibles pour chaque changement du système
	  3. **Coordonner** interactions et feedbacks
	  4. Imaginer un environnement et des objets pour « **emballer** » les interactions et les feedbacks
	3. **Vérifier** que les interactions+feedbacks permettent...
	  1. D'explorer le jouet pour essayer de comprendre le système : **explorer** les règles du système
	  2. D'**expérimenter** les règles que les utilisateurs/ices ont comprises
	4. **Tester :** est-il intéressant de manipuler le jouet, sans but et sans contrainte

Tableau 5. Synthèse des JS produits lors de l'étude 3

Nom du JS	Genre	Ambition	Outils de développement	Principales mécaniques	Méthode d'apprentissage visée
Enter the EU	Simulation	Comprendre le processus d'entrée d'un pays dans l'UE	Construct	Choix binaires à chaque tour pour influencer des jauges	Exploration + expérimentation
Élections européennes	Succession de QCM	Comprendre la constitution de listes électorales	Unity 3D[10]	Choisir ou non les cartes qui sont présentées.	Exploration + textes de remédiation
Procédure législative	Simulation	Comprendre l'élaboration des lois dans l'UE	Construct	Choix binaires à chaque tour pour influencer des jauges	Exploration + expérimentation

10 Unity 3D : https://unity3d.com/ (consulté le 26/04/2019)

Nous avons fait deux observations principales. D'une part, les outils auteurs n'ont pas été un obstacle technique : les groupes étaient grands et peuplés de suffisamment d'étudiants connaisseurs.

D'autre part, nous avons observé que pour deux des trois projets, les étudiants se sont emparés des éléments méthodologiques pour identifier ce qui faisait système dans leurs sujets afin de concevoir des simulations autour de jauges. Pour ces deux JS, l'interface choisie est celle de *Reigns*[11] qui propose un moyen simple et facilement ludique d'interfacer une simulation multifactorielle. Les étudiants ont pu constater à l'occasion des playtests que des participants exploraient et expérimentaient en *jouant* avec la simulation pour inférer la compréhension des systèmes, parfois sans poursuivre les buts du JS, c'est-à-dire comme pour des jouets au sens de Ryan et al.

Si nous avons constaté que les 3 premières étapes de la méthode de Ryan et al. associées aux 6 facettes semblent donner de bons résultats, il nous faut néanmoins poursuivre nos études. Principalement en vérifiant, d'une part, quels sont les éléments méthodologiques fournis qui sont les plus utiles. D'autre part, en vérifiant qu'ils ne mènent pas toujours à des interfaces de type *Reigns* ou si ce résultat est un biais issu des échanges entre les étudiants. Ces travaux complémentaires devraient nous permettre, autour de la base méthodologique que nous venons d'établir, de construire des patrons de conception structurés pour concevoir des JS articulés sur des jouets systémiques.

5 Conclusion

Dans la perspective de fournir des Jeux Sérieux (JS) permettant un apprentissage par l'exploration et l'expérimentation grâce à des jouets/micro-mondes très systémiques, nous avons choisi de mener trois études de terrain en situation écologique. Elles sont inspirées de la recherche collaborative orientée par la conception, et menées auprès d'une centaine d'étudiants ayant participé à la conception de 23 prototypes de JS. Les objectifs de ces études étaient, d'une part, de mettre à l'épreuve la méthode de conception proposée par Ryan et al. [16] et, d'autre part, d'extraire des éléments de patrons de conception pour les JS centrés sur des jouets/micro-mondes.

Nous avons observé que pour que les participants appréhendent correctement la notion du jouet systémique, comme un micro-monde, il est nécessaire d'aller au-delà de la définition et des éléments méthodologiques proposés par Schell [18] dans *The Lens of the Toy*. Ainsi, nous montrons que les 3 étapes de conception de jouets proposées par Ryan et al. [16] associées aux 6 facettes de la conception des jeux sérieux [10] ont aidé les participants à intégrer et concevoir leurs JS autour de jouets systémiques.

Nous proposons donc (Tableau 4) des principes fondateurs pour des patrons de conceptions associés aux facettes de *Simulation du domaine* et d'*Interactions avec la simulation*. Ces fondations devront être enrichies, raffinées et formalisées pour constituer des patrons complets.

Nous avons aussi observé l'influence (négative) de certains outils auteurs dans la conception et l'implémentation de jouets systémiques dans les JS. Pour éviter cet effet négatif, nous envisageons de rendre formels les futurs patrons de conception afin de faciliter leur implémentation dans les outils auteurs.

11 Reigns : https://www.devolverdigital.com/games/view/reigns (consulté le 26/04/2019)

Remerciements. L'auteur remercie le LABEX ASLAN (ANR-10-LABX-0081) de l'Université de Lyon pour son soutien financier dans le cadre du programme « Investissements d'Avenir » (ANR-11-IDEX-0007) de l'État français géré par l'Agence Nationale de la Recherche (ANR).

Bibliographie

1. Alexander, C. et al.: A pattern language. Oxford University Press US (1977).
2. Bogost, I.: Exploitationware. In: Colby, R. et al. (eds.) Rhetoric/Composition/Play through Video Games: Reshaping Theory and Practice of Writing. pp. 139–147 Palgrave Macmillan US, New York (2013).
3. Bogost, I. et al.: Newsgames: Journalism at play. Mit Press, Cambridge, MA (2012).
4. Bogost, I.: Persuasive Games: The Expressive Power of Videogames. MIT Press, Cambridge, UK (2007).
5. Callan, R.C. et al.: How to avoid the dark side of gamification: Ten business scenarios and their unintended consequences. In: Gamification in education and business. pp. 553–568 Springer (2015).
6. Deci, E.L. et al.: A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol. Bull. 125, 6, 627 (1999).
7. Gee, J.P.: Deep learning properties of good digital games: How far can they go? Serious Games Mech. Eff. 67–82 (2009).
8. Marfisi-Schottman, I.: Méthodologie, modèles et outils pour la conception de Learning Games. INSA de Lyon (2012).
9. Marfisi-Schottman, I., Piau-Toffolon, C.: Extraire et réutiliser des patrons de conception à partir de Learning Games existants. Presented at the Atelier méthodologies de conception collaboratives des EIAH de la Conférence EIAH 2015 June 2 (2015).
10. Marne, B. et al.: Articuler motivation et apprentissage grâce aux facettes du jeu sérieux. In: Actes de la conférence EIAH 2011. pp. 69–80 Université de Mons, Mons, Belgique (2011).
11. Marne, B.: Modèles et outils pour la conception de jeux sérieux : une approche meta-design. Université Pierre et Marie Curie (UPMC) (2014).
12. Meszaros, G., Doble, J.: A pattern language for pattern writing. In: Pattern Languages of Program Design-3. pp. 529–574 Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (1997).
13. Moreno-Ger, P. et al.: The <e-Game> Project: Facilitating the Development of Educational Adventure Games. In: Proceedings of the IADIS International Conference on Cognition and exploratory learning in digital age (CELDA 2005). , Porto, Portugal (2005).
14. Papert, S.: Microworlds: transforming education. In: Artificial intelligence and education. pp. 79–94 (1987).
15. Peirce, C.S.: Minute Logic: Chapter I. Intended Characters of this Treatise. MS [R] 425, (1902).
16. Ryan, M. et al.: Deep Learning Games through the Lens of the Toy. In: Meaningful Play 2012. pp. 1–29 , East Lansing, USA (2012).
17. Sanchez, É., Monod-Ansaldi, R.: Recherche collaborative orientée par la conception. Un paradigme méthodologique pour prendre en compte la complexité des situations d'enseignement-apprentissage. Educ. Didact. 9, 2, 73–94 (2015).
18. Schell, J.: The Art of Game Design: A book of lenses. Morgan Kaufmann (2008).
19. Wang, F., Hannafin, M.J.: Design-based research and technology-enhanced learning environments. Educ. Technol. Res. Dev. 53, 4, 5–23 (2005).

Analyser la disparition de la gratuité dans les MOOC au prisme des buts d'accomplissement

Matthieu Cisel

Université Paris-Descartes, Laboratoire EDA,
45 rue des Saint-Pères,75006 Paris, France
matthieucisel@gmail.com

Résumé. Les MOOC, à leur apogée, délivraient gratuitement des certificats et attestations de réussite par centaines de milliers chaque année. La gratuité a désormais presque disparu. Nous nous proposons de chercher à mieux identifier les conséquences de cette évolution via une enquête qualitative menée auprès d'une quarantaine d'apprenants sur leur intérêt vis-à-vis du certificat. L'analyse est menée au prisme de la théorie des buts d'accomplissement. Nous montrons que malgré l'existence d'apprenants autodirigés hostiles à ces attestations, l'intérêt pour le certificat n'est pas incompatible avec le fait de suivre le cours dans une logique de buts de maîtrise – comprendre avec la volonté de maîtriser des acquis plus qu'avec celle de montrer des performances via l'obtention d'un certificat. Nous dressons ici une typologie des arguments mobilisés pour légitimer un intérêt, ou à l'inverse un certain désintérêt pour les attestations de réussite.

Mots-clés. MOOC, motivation, certificats, persistance

Abstract. At their peak, each year MOOCs would deliver free statements of accomplishment by the hundreds of thousands. Such free certificates have almost completely disappeared. Based on the achievement goal theory, we try to assess the consequences of such an evolution through qualitative interviews. Some learners show little interest for those certificates, notably because of their little value on the job market. Nevertheless, mastery goals are not incompatible with the interest for the certificate. We here build a typology of arguments used by learners to legitimize their interest, or on the contrary their lack of interest for the free statements of accomplishment that are delivered in MOOCs.

Keywords. MOOC, motivation, certificates, retention

1 Introduction

1.1 Une disparition progressive des certificats gratuits

Les plates-formes américaines qui dominaient et dominent toujours la scène mondiale des MOOC ont largement renoncé à la gratuité de ces cours en ligne [9,25]. On peut certes généralement encore accéder aux ressources des cours sans payer, principalement les vidéos pédagogiques. Néanmoins, l'obtention gratuite d'attestations de réussite, ou de certificats comme nous les appellerons ici, n'est plus d'actualité – en 2018, la plate-forme française FUN fait figure d'exception en la

matière. Cette mutation constitue à certains égards un retour au modèle traditionnel de formation à distance dont les ressources pédagogiques seraient librement accessibles, et dont l'Open University a constitué l'un des pionniers [28].

Il est difficile d'évaluer l'impact, sur l'utilisation de ces cours, de cette disparition progressive de la gratuité de la certification. Nous ne disposons pas de chiffres précis, faute de la mise à disposition généralisée des traces par les plates-formes qui ont changé de politique en matière de certification – le caractère stratégique de ces données y est sans doute pour beaucoup. Néanmoins, une meilleure compréhension du rapport des participants au certificat peut sans doute aider à appréhender les conséquences de ces évolutions impulsées par les plates-formes.

Les nombreuses enquêtes quantitatives sur les motivations des apprenants suggèrent que ceux-ci s'inscrivent avant tout selon une logique de buts de maîtrise [8,10], c'est-à-dire avec l'intention de gagner en compétences ou en connaissances plus qu'avec celle d'obtenir des certificats qu'ils pourraient faire valoir. Ce faisant, nous retrouvons des résultats connus de longue date ; dès les années 1960, les travaux américains sur l'éducation non formelle [20], comme les travaux français en sociologie des loisirs [12], mettaient au jour des tendances similaires.

Si les enquêtes diffusées dans les MOOC révèlent un certain nombre de régularités empiriques sur ces problématiques, elles ne permettent néanmoins pas de comprendre les mécanismes sous-jacents. La question initiale qui a motivé cette recherche peut être formulée dans un premier temps avec le vocabulaire du sens commun : quels sont les arguments de ceux qui déclarent porter un intérêt au certificat, et quels sont les arguments des apprenants qui affirment leur désintérêt pour le certificat ?

Le travail de problématisation qui a suivi la formulation de ces questions nous a amené à mobiliser la théorie des buts d'accomplissement [14, 27], dont la section suivante permet d'exposer les grandes lignes.

1.2 Théorie des buts d'accomplissement et notions complémentaires

La théorie des buts d'accomplissement a été impulsée par les travaux de Dweck [14] et Nicholls [27]. Les deux concepts phare sont les buts de maîtrise et de performance. Les buts de maîtrise, ou buts d'apprentissage, représentent le cas où « l'objectif de la personne est d'apprendre, de développer de nouvelles habiletés, de comprendre ou de maîtriser quelque chose de nouveau. Être compétent signifie apprendre, comprendre, progresser » [7]. Les buts de performance sont « des buts où la comparaison sociale joue un rôle déterminant. Centré sur la valorisation de soi, l'objectif est d'obtenir des jugements favorables sur sa compétence. Avoir appris ou progressé ne suffit pas pour se sentir compétent. Il importe avant tout de savoir où l'on se situe par rapport aux autres » (ibid).

Selon Elliot et Dweck [15] les « buts qu'un individu poursuit créent un cadre pour interpréter et réagir aux événements qui se présentent ». La théorie des buts d'accomplissement va s'élargir pour introduire deux buts supplémentaires : l'évitement de la performance, et l'évitement de la maîtrise. La question de l'évitement de la performance repose sur le fait que « Chercher un jugement favorable n'a pas la même signification et n'implique pas les mêmes conduites que chercher à éviter un jugement défavorable ». Elliot et Harackiewicz [16] scindent les buts de performance en deux buts, « l'un orienté vers la démonstration de sa propre compétence (performance-approach goal : rechercher d'un jugement favorable de la

part d'autrui), l'autre but consistant à éviter la démonstration de son incompétence (performance-avoidance goal : évitement d'un jugement défavorable de la part d'autrui) » [7].

Cette distinction permet de considérer la tendance à l'évitement de l'échec comme un but à part entière. La scission ayant touché les buts de performance affectera de la même manière les buts de maîtrise pour conduire au modèle à quatre buts actuellement utilisé à travers la naissance du but intitulé « évitement de la maîtrise » (mastery-avoidance goal). Le terme « évitement » correspond au fait que « l'objectif est d'éviter de perdre, ou le moins possible, le niveau de compétence atteint antérieurement […], éviter de perdre la main aussi bien dans des activités professionnelles que de loisir » [17,18]. Si notre enquête a permis de mettre à jour de tels buts d'évitement, nous ne nous attarderons pas sur la question dans cette contribution.

Le dernier point que nous aborderons réside dans le fait que les apprenants pouvaient poursuivre plusieurs buts de manière simultanée, et qu'il était dès lors plus pertinent de parler de profils de buts [31]. Il existe une hiérarchie dans les buts, considération qui nous amènera à prendre un certain nombre de précautions terminologiques. Nous dirons d'un participant qu'il s'inscrit « avant tout » dans une logique de buts de maîtrise, si les buts de performance sont secondaires par rapport aux buts de maîtrise. Cette considération est centrale à notre propos, dans la mesure où elle permet d'expliquer aisément pourquoi l'intérêt pour le certificat n'est pas incompatible avec les buts de maîtrise. Elle ne permet pas en revanche d'appréhender la diversité des configurations possibles et des arguments mobilisés, diversité que le travail qualitatif que nous avons mené vise à mettre au jour.

En sus des concepts issus de cette théorie, nous utiliserons le concept d'apprenant autodirigé, qui s'appuie notamment sur le concept d'auto-régulation (Zimmerman, 1990). La théorie de l'apprentissage autodirigé tire son origine notamment dans les ravaux de Tough et de Knowles [23]. Aussi reprendrons-nous avec Jezegou [19] la définition qu'en donne ce dernier : « Un processus dans lequel les individus prennent l'initiative, avec ou sans l'aide d'autrui, de déterminer leurs besoins de formation, de recenser les ressources humaines et matérielles nécessaires à la formation, de sélectionner et de mettre en œuvre les stratégies de formation adéquates, d'évaluer les résultats de leur formation » [23].

Les instruments mobilisés dans de multiples recherches montrent que le caractère particulièrement autodirigé de la plupart des apprenants de MOOC [22]. Cette considération nous a amené à définir avec Carré [6] le concept d'apprenant autodirigé : « L'apprenant autodirigé est à la fois fortement engagé dans son propre projet (autodétermination), armé de techniques et de ressources cognitives, matérielles et humaines dont il est capable de réguler les usages en fonction de ses propres objectifs (autorégulation), le tout étant fortement soutenu et dynamisé par un sentiment affirmé de son efficacité personnelle à apprendre ».

1.3 Reproblématisation

A la lueur des différents concepts présentés jusqu'à présent, nous pouvons désormais préciser davantage notre problématique et fournir quelques résultats partiels.

Comment buts d'accomplissement et rapport au certificat s'articulent-ils pour les utilisateurs de MOOC dont le profil est dominé par les buts de maîtrise ?

Ainsi, si les enquêtes suggèrent que le profil des participants penche plutôt du côté des buts de maîtrise, elles confirment également la théorie dans le sens où des buts de performance sont également poursuivis, même de manière secondaire : la plupart des répondants déclarent porter un certain intérêt au certificat, qui pourrait jouer un rôle central dans la motivation des participants pour aller jusqu'au terme de la formation, même lorsqu'il est gratuit, peu reconnu, et qu'il n'a pas vocation à être utilisé dans un contexte professionnel.

Le lien entre intérêt pour le certificat et buts d'accomplissement de soi est encore mal compris, et les instruments mobilisés à ce jour dans les enquêtes ne permettent pas d'explorer la question en détail. La question du rôle du certificat, et plus généralement des diplômes, dans la rétention des apprenants a été étudiée dès les années 1970 a donné lieu à des modèles célèbres avec pour objet les études universitaires [29]. Ils ont par la suite été adaptés à la formation à distance et à la formation d'adultes. Ces travaux suggèrent que la disparition d'un paramètre comme le certificat rend la persistance plus sensible aux événements extérieurs que peuvent représenter les responsabilités familiales ou professionnelles des apprenants. Néanmoins, ces recherches ne permettent pas de prendre en compte le fait que le MOOC est une formation à distance particulière, en ce sens qu'il n'y a le plus souvent pas de reconnaissance formelle des certificats ni sur le plan de la formation académique ni sur celui de la formation tout au long de la vie.

2 Méthodologie

Cette enquête est fondée sur une quarantaine d'entretiens semi-directifs avec des apprenants de MOOC. Plusieurs stratégies ont été employées pour sélectionner nos sujets : diffusion d'une annonce via la lettre d'information du MOOC que nous organisions, le *MOOC Monter un MOOC de A à Z,* et via les réseaux sociaux de plusieurs MOOC. Le message de quelques lignes expliquait le contexte de l'entretien, sa durée approximative, et invitait à envoyer un courriel à une adresse dédiée, en se présentant rapidement.

Au cours de la phase qualitative de ce travail, quarante-et-un entretiens ont été menés, entre septembre 2014 et juin 2015. Bien que les entretiens aient été réalisés il y a plus de quatre ans, il nous semble que les arguments sont toujours d'actualité, et qu'ils éclairent de manière pertinente les évolutions récentes des politiques de certification des plates-formes. La population d'enquête, à majorité féminine (24 femmes pour 17 hommes) était en moyenne âgée de 41 (± 11) ans. La majorité des enquêtés disposaient au moins d'un master, mais six n'avaient que le bac ou un diplôme inférieur. Le nombre de MOOC suivis par ces apprenants était particulièrement élevé, avec 5,8 MOOC suivis en moyenne (± 7,5).

Le canevas d'entretien comprenait, en sus des variables sociodémographiques, des questions sur les motivations pour s'inscrire, sur les comportements observables au sein des MOOC suivis (nature et nombre d'actions réalisées, etc.), et sur les comportements qui n'étaient pas observables par les traces d'interaction. Les entretiens étaient réalisés par l'outil de visioconférence *Skype* ou par téléphone, et

enregistrés via une application dédiée. Une fois retranscrits, les entretiens étaient analysés selon le principe de l'analyse thématique [2]

Ces thématiques peuvent être rangées en trois grandes catégories. Dans une première catégorie, nous regroupions les éléments relatifs à la motivation pour s'inscrire et pour persister au sein de la formation. Dans une seconde catégorie, nous regroupions les perspectives des participants relatives aux différentes caractéristiques du dispositif (activités prescrites, interactions avec les autres participants, etc.). Dans une troisième catégorie, nous rangions les thématiques relatives aux actions réalisées par les participants qui ne pouvaient être capturées par l'analyse des traces d'interaction.

Au cours des entretiens, nous avons mobilisé des items permettant de dresser le profil de buts des enquêtés, en nous focalisant sur le dernier MOOC qu'ils avaient suivi, le profil de buts pouvant varier selon le type de cours. Nous commencions l'entretien par une question ouverte : Quelles sont les raisons qui vous ont poussé à vous inscrire à ce MOOC ? Lorsque plusieurs motifs concurrents étaient présentés, cas le plus fréquent, nous demandions à l'enquêté de les hiérarchiser, via la question suivante : Parmi les différents motifs que vous avez cités, quel est celui qui vous semble plus déterminant dans votre inscription ? Pour identifier l'existence éventuelle de buts d'évitement parmi les apprenants où dominait la logique de buts de maîtrise, nous avons rajouté une question liée à la connaissance des thématiques abordée : Dans quelle mesure maîtrisez-vous les thématiques abordées dans ce MOOC ? Nous la doublions d'une question de relance le cas échéant : Dans quelle mesure cherchez-vous à éviter de « perdre la main » sur des sujets sur lesquels vous vous êtes déjà penchés ? Enfin, pour tous les participants ayant obtenu un certificat gratuit, nous posions les questions suivantes : Quelle valeur le certificat du MOOC revêt-il à vos yeux ? Qu'en avez-vous fait une fois que vous l'avez obtenu ?

Dans la mesure où nous avons par ailleurs réalisé sur la question du rapport au certificat une enquête quantitative [10] ayant réuni plus de 8000 réponses, il ne nous est pas apparu pertinent de mener une analyse quantitative du travail qualitatif présenté ici. Nos enquêtés ne sont pas nécessairement représentatifs des apprenants de MOOC, et, notamment si l'on s'intéresse au lien entre profils de motifs d'inscription et persistance dans le MOOC, le croisement entre enquêtes et traces d'interaction [21] représente une approche vraisemblablement plus intéressante.

3 Résultats

Nous commencerons par présenter les arguments des participants qui déclarent porter peu d'intérêt au certificat du MOOC. Nous ferons la distinction entre ceux qui affirment que ce désintérêt découle de sa faible valeur sur le marché de l'emploi, et ceux qui, parce qu'ils estiment être des apprenants autodirigés, considèrent qu'ils n'ont pas besoin de cette source de motivation extrinsèque. Nous nous pencherons ensuite sur le cas de ceux qui, tout en s'inscrivant dans une logique de but de maîtrise, affirment que le certificat possède une certaine valeur à leurs yeux. Dès lors, pour ces participants, sa suppression pourrait avoir des conséquences sur leur persistance au sein du cours.

3.1 Comprendre le désintérêt pour le certificat

3.1.1 De la faible valeur du certificat

Le désintérêt pour le certificat est une position le plus souvent soutenue par des participants s'inscrivant dans une logique de buts de maîtrise, et ne voyant pas d'utilité au certificat du fait de sa faible valence, au sens de sa faible valeur instrumentale. Un certain nombre d'éléments peuvent être avancés par l'enquêté pour justifier cette position. Le premier est d'ordre professionnel, notamment lorsque le participant n'envisage pas de mobilité en termes d'emploi. Alternativement, l'enquêté peut envisager une mobilité professionnelle mais considérer que le certificat n'a pas suffisamment de valeur sur le marché de l'emploi pour lui être d'une quelconque utilité. Ce cas de figure est illustré par une étudiante de Master en *Français Langue Etrangère* ayant suivi un MOOC de manière optionnelle dans le cadre de son cursus.

Je me destine vers l'enseignement en France donc je sais que ça ne va pas être quelque chose qui va me rapporter quoi que ce soit dans ma vie professionnelle. L'enseignement en France étant ce qu'il est, tout ce qui ne sort pas, n'étant pas un vrai diplôme ne sert à rien. (E31)

Elle explique ne pas attacher une grande importance au certificat dans la mesure où celui-ci n'est pas reconnu sur le marché du travail français. Enfin, certains participants déclarent qu'ils accorderaient davantage de valeur au certificat si ce dernier était plus reconnu sur le marché de l'emploi.

3.1.2 La revendication d'un apprentissage autodirigé

Le second type d'argument est avancé par les apprenants qui se déclarent autodirigés [6]. Pour un certain nombre de participants, la démarche d'autoformation qu'ils suivent est antinomique avec la logique de buts de performance qu'implique la recherche du certificat. Ce cas correspond aussi bien au participant qui estime que son apprentissage ne doit être guidé que par sa seule satisfaction personnelle, qu'à celui du participant autodirigé, ayant une idée précise de ce qu'il veut apprendre à travers le dispositif. Nous illustrerons ce dernier cas par la figure du développeur Web auto-entrepreneur, utilisant des dizaines de MOOC, et ayant créé des cartes mentales pour identifier les lacunes qu'il doit combler par le MOOC. Il déclare être opposé au principe même de la délivrance de certificats dans le contexte d'un MOOC.

Pour moi, ça fait pas partie de la philosophie des MOOC, la certification. Ça peut être bien parce que t'as beaucoup de gens qui recherchent ça, mais je pense que c'est des gens qui ont beaucoup de choses à régler. Tu vois, c'est personnel, y a un besoin de reconnaissance, de diplôme, et autre. Quelque part si tu utilises les MOOC pour apprendre des choses, tu n'as pas besoin d'une validation […] Si t'as envie d'un diplôme, pour moi tu fais une formation diplômante. (E33)

En un sens, il estime qu'il doit assumer la responsabilité du processus d'évaluation de l'apprentissage. Faible valence du certificat et revendication d'une forte propension à l'autodirection constituent les deux principaux arguments avancés par les tenants d'un discours hostile aux certificats de MOOC. Ce discours n'est pas tenu par l'ensemble des apprenants déclarant s'inscrire avant tout dans une logique de buts de maîtrise, aussi reviendrons-nous maintenant sur les arguments de ces derniers, afin

de mieux appréhender pourquoi la logique de maîtrise peut être compatible avec un certain intérêt pour le certificat.

3.2 Buts de maîtrise et intérêt pour le certificat

Nous pouvons identifier au moins trois arguments avancés par les enquêtés déclarant porter de l'intérêt au certificat, sans que cet intérêt soit de nature professionnelle. En premier lieu, l'apprenant peut vouloir valoriser des apprentissages informels, notamment auprès de cercles d'amis ou de proches. En second lieu, le certificat peut être considéré comme nécessaire au maintien de la motivation, bien qu'il n'ait pas vocation à être utilisé par la suite. En dernier lieu, il peut servir à valider que le contenu du cours est bien maîtrisé, l'intérêt pour le certificat découlant alors précisément de la logique de buts de maîtrise dont se revendiquent ces apprenants.

3.2.1 Valoriser des apprentissages en dehors du contexte professionnel

Certains participants affichent sur les réseaux sociaux les certificats obtenus, et ce bien qu'ils n'envisagent pas de les utiliser dans un contexte professionnel. Ils ne s'inscrivent pas pour autant avant tout dans une logique de buts de performance, dans la mesure où cette valorisation reste secondaire par rapport aux buts de maîtrise. Nous illustrerons cette position par les propos d'une mère au foyer ; bien qu'elle n'envisage pas de reprendre son activité professionnelle à court terme, elle affiche sur les réseaux sociaux les certificats qu'elle obtient pour montrer à autrui qu'elle est active intellectuellement malgré son statut de mère au foyer. Les certificats de MOOC lui permettent ainsi de valoriser ses apprentissages, chose qu'elle ne peut faire avec un livre. Face à la question de la valeur qu'elle accorde au certificat, elle répond comme suit :

Ben moi ça a une valeur qui est, je les ai mis sur LinkedIn parce que ça a une valeur aussi par rapport au fait d'être à la maison. Je ne sais pas si ça se passe comme ça en France, mais en Suisse, on a un peu l'impression avec mon mari, c'est assez drôle, quand on croise des gens, on lui pose la question ce qu'il fait, mais moi on me pose pas la question parce que je suis à la maison donc sûrement que je fais rien. Quelque part ils ont une valeur, mais le fait de dire, bah ouais, je suis à la maison, actuellement je ne suis pas employée mais je ne fais pas rien. […] Je suis sûr qu'il y a des mamans qui lisent des livres, qui font plein de choses. L'avantage c'est que les MOOC ça permet d'une certaine manière de les valoriser. (E7)

Précisons que d'autres enquêtés déclarent afficher les certificats obtenus sur les réseaux sociaux, sur leurs blogs et leurs sites personnels. Certains affichent ainsi des collections de plusieurs dizaines de certificat dans des domaines variés. Néanmoins, il est fréquent chez les enquêtés qui s'intéressent au certificat de déclarer ne pas désirer montrer à autrui les certificats obtenus. Ils peuvent alors légitimer leur intérêt par le rôle que joue le certificat dans le maintien de la motivation.

3.2.2 Une aide au maintien de la motivation

Un certain nombre d'enquêtés déclarent en jouant un rôle de récompense, même symbolique, le certificat aide à maintenir la motivation tout au long de la formation, ou apporte une forme de satisfaction a posteriori. Cette situation est illustrée par les propos de ce consultant inscrit à de nombreux MOOC.

Souvent le réflexe c'est de dire t'es allé jusqu'ici, il y a quand même la certification, donc c'est une manière de ne pas se démotiver, c'est une manière de lutter contre l'abandon. (E1)

Le cas est également illustré par ce cadre d'entreprise qui « collectionne » les certificats de MOOC et fait le parallèle avec les « bons points ».

Je pense que c'est vrai, je préfère quand même avoir un petit papier, je reconnais, et puis ça fait plaisir, à titre personnel aussi. On s'amuse avec Sophie, parce qu'on a plein de certificats maintenant, donc on est très contents. Bon, bah c'est un petit, c'est un bon point qu'on nous donne, on est réalistes, mais un petit bon point avec l'université du Maryland, un petit certificat j'imagine que Sophie aura ça un jour, de Harvard ou autre chose, bon bah, c'est fun aussi. (E27)

L'enquêté souligne ici l'aspect ludique que revêt leur obtention, lui et sa femme jouant à les amasser.

3.2.3 Valider que les contenus sont effectivement maîtrisés

Dernier cas de figure, le participant cherche à travers l'obtention du certificat une validation de l'institution lui permettant de se rassurer quant au fait qu'il maîtrise le contenu du cours. Le cas est notamment illustré par cette ingénieure de recherche en agronomie qui suit, entre autres, des cours de statistiques. Elle utilise fréquemment la discipline dans le cadre de son activité professionnelle, sans pour autant y avoir été formée. L'obtention de certificats de MOOC de statistiques la conforte dans le fait qu'elle utilise correctement les statistiques.

Pour moi pour les certificats de statistiques, oui ça a de l'importance, parce que ça me conforte que j'ai bien, même si j'ai pas suivi des enseignements pendant ma formation initiale sur ces contenus-là, je maîtrise quand même le sujet. Ça me valide un peu ce que je fais. (E18)

Bien qu'il suive le cours dans une logique de buts de maîtrise, le participant peut chercher soit à valider des compétences ou des connaissances liées au MOOC, soit chercher à valider des connaissances qu'il aura acquises de manière informelle, en amont du cours, comme c'est le cas pour cette ingénieure avec les statistiques. Nous retrouvons ici le cas de la légitimation des apprentissages informels, dont il reste à déterminer l'impact sur la motivation intrinsèque des participants. Nous rejoignons ici le débat sur l'effet des récompenses extrinsèques matérielles, comme ici les attestations et certificats, pour les activités où l'apprenant est déjà motivé intrinsèquement [11]. Ces travaux ont montré chez l'enfant que dans ce type de configuration, l'obtention de la récompense, et le plaisir associé finissent par l'emporter sur la raison intrinsèque de s'engager dans l'activité. Reste à déterminer dans quelle mesure une telle dynamique peut se retrouver dans le cas d'adultes apprenant en ligne.

4 Discussion

4.1 Des apprenants autodirigés intéressés par le certificat

La fondatrice de la célèbre plate-forme américaine Coursera, affirmait dans un article polémique sur les faibles taux de certification [24] que les utilisateurs de MOOC

étaient pour la plupart autodirigés. Dès lors, le fait de s'inscrire dans une logique de buts de maîtrise les rendait selon elle relativement insensibles à la délivrance du certificat, minimisant par conséquent l'impact de la disparition des certificats gratuits. L'analyse des entretiens que nous avons exposée ici nous invite à nuancer un tel point de vue, en suggérant que l'apprentissage autodirigé et la logique de buts de maîtrise ne sont pas incompatibles avec l'intérêt pour le certificat, conformément à ce que l'on pourrait attendre au regard de la théorie des buts d'accomplissement, et plus précisément des écrits sur les profils de buts [31].

Certes, nous avons vu le cas de figure où l'apprenant autodirigé maintient une posture incompatible avec l'intérêt pour le certificat, au nom d'une logique de buts de maîtrise dont il se revendique[1]. Ces apprenants sont attachés à leur mode d'autoformation, qui s'apparente à ce que l'on qualifie parfois dans la littérature francophone d'autoformation intégrale [13] – il n'y a aucune médiation de la part d'enseignants, au point de refuser de suivre la logique d'un dispositif de formation. Ils préfèrent sélectionner dans le MOOC les ressources correspondant aux connaissances et aux compétences qu'ils souhaitent acquérir, même si cela signifie renoncer à l'éventuel certificat. Ce type d'apprenant autodirigé est vraisemblablement en effet peu affecté par la disparition des certificats gratuits.

Néanmoins, on peut aussi émettre l'hypothèse selon laquelle les apprenants autodirigés s'accommodent globalement du retour de la logique d'hétéroformation qui sous-tend le suivi d'un MOOC, y compris ceux qui déclarent porter peu d'intérêt au certificat du fait de sa faible valeur. Il est vraisemblable que nombre d'apprenants fortement autodirigés se soumettent complètement au contrôle pédagogique du dispositif [26], en suivant dans une large mesure les instructions de l'équipe pédagogique. C'est un phénomène qui correspond à la trajectoire inverse du *saut autodidactique* de Bouchard [3]. Alors que dans le cas du saut autodidactique, l'apprenant fait le choix de mener par lui-même un projet d'apprentissage de manière autonome malgré l'existence d'alternatives, c'est le phénomène contraire dont nous sommes vraisemblablement témoins avec les MOOC.

La plupart des rôles généralement dévolus à l'apprenant autodirigé se formant en ligne sont récupérés dans le cadre d'un MOOC par l'équipe pédagogique à l'origine du dispositif : de la sélection des ressources à la fixation des objectifs en passant par le cadencement de l'essentiel du projet d'apprentissage éventuel. Notre analyse des entretiens nous amène à avancer l'hypothèse selon laquelle la majorité des apprenants autodirigés, y compris ceux qui sont habitués à s'autoformer via différents supports, livres, ressources éducatives libres, blogs, sans la médiation d'une équipe pédagogique, sont favorables au transfert du contrôle pédagogique. Selon cette hypothèse, l'autoformation intégrale représente pour eux un mode d'autoformation coûteux en temps et en ressources, qu'ils n'adoptent que forcés par l'absence d'alternative compatible avec leurs contraintes.

Les recherches d'Arendt et Shelton [1] montrent que l'un des principaux inconvénients des *Open Coursewares* aux yeux de leurs utilisateurs réside dans l'absence de certification associée à l'apprentissage. Les MOOC constituent pour eux une solution à l'absence d'organisation de ces apprentissages informels, et le

[1] Nous sommes dans la configuration conflictuelle que constitue la rencontre d'un fort contrôle pédagogique et d'un fort contrôle psychologique [26].

certificat, par les activités qu'il suppose, permet de structurer la démarche. Bien qu'ils n'envisagent pas de le valoriser, l'existence du certificat jouerait sur le maintien de la motivation tout au long de la formation, c'est-à-dire la *persistance* de la motivation au sens de Vallerand et Thill [30], y compris chez les participants n'envisageant pas de l'utiliser dans un contexte professionnel. Il permet d'une certaine manière de faciliter la fixation d'un but défini. Il aide au maintien de la persistance de la motivation, mais aussi à la *direction* de cette motivation. En d'autres termes, il aide le participant à orienter ses actions vers un objectif déterminé.

4.2 Validation externe

Nous avons ensuite vu que le certificat pouvait servir précisément du fait de la logique de buts de maîtrise dont se revendiquent les participants. On peut alors distinguer plusieurs situations selon le niveau de connaissances sur la thématique du MOOC. Pour la plupart des participants, qui débutent ou n'ont que quelques notions sur la thématique, ce certificat vient valider des connaissances ou des compétences nouvellement acquises, leur valider le fait qu'ils ont bien compris le contenu du cours. Pour les participants possédant une certaine connaissance du sujet – notamment du fait de leur activité professionnelle, il vient possiblement légitimer des connaissances acquises de manière informelle, ou s'assurer que les connaissances n'ont pas été perdues, selon alors une logique de buts d'évitement.

Nous émettons, à la suite de ces entretiens, l'hypothèse selon laquelle ils sont conscients des lacunes de leur propre processus d'autoformation. C'est l'idée que colportent les propos de l'ingénieure en agronomie cherchant à se rassurer quant à ses compétences en statistiques. Le participant peut estimer que les apprentissages informels sont incomplets car insuffisamment structurés ; dans la mesure où ils s'inscrivent dans une logique de buts de maîtrise, ils cherchent une évaluation externe de leurs acquis, composante qui fait défaut à de nombreuses formes d'autoformation, et ce sans que cette évaluation externe ne s'inscrive dans une logique de buts de performance.

Dans un cas comme dans l'autre, le certificat joue un rôle dans le maintien au sein de la formation. Cette considération nous amène à la notion d'autoformation d'accompagnement proposée par Dumazedier [13] ; il désigne les pratiques visant à « compléter, corriger l'action éducative limitée ou illusoire de l'institution » ; elle peut concerner des personnes très instruites. La formation initiale ne permet pas d'acquérir l'ensemble des connaissances ou compétences requises à l'exercice de l'activité professionnelle.

5 Conclusion

Les certificats gratuits jouent vraisemblablement un rôle essentiel dans le maintien de la motivation des apprenants, y compris pour ceux qui s'inscrivent avant tout selon une logique de buts de maîtrise. Selon toute vraisemblance, leur disparition, qu'elle soit le choix d'une équipe pédagogique ou qu'elle soit imposée par les évolutions des plates-formes, est susceptible de faire chuter considérablement les statistiques de rétention des apprenants, même pour ceux qui s'inscrivent avant tout selon une logique de buts de maîtrise. Il nous faut dès lors nuancer les propos de ceux qui

affirment que, parce que les enquêtes montrent que sont minoritaires les apprenants s'inscrivant avant tout pour obtenir un certificat valorisable professionnellement, la disparition des certificats n'aura que peu d'impact.

Avec la disparition progressive de la gratuité, il est vraisemblable qu'il ne reste plus qu'un noyau réduit d'apprenants pour acheter les certificats payants, et réaliser les activités afférentes. Qui sont-ils et quels sont les buts qu'ils poursuivent ? L'une des principales limites de cette enquête est de s'être cantonnée aux apprenants inscrits avant tout dans une logique de buts de maîtrise. Il serait nécessaire de prolonger nos questionnements en étudiant davantage la population des apprenants s'inscrivant avant tout dans une logique de buts de performance d'une part, et préciser le rapport des apprenants aux certificats payants d'autre part, indépendamment de leur profil de buts [31]. On pourra également prolonger les travaux croisant les enquêtes relatives aux motivations pour s'inscrire avec des traces d'interaction collectées par les plateformes [21], pour tâcher d'établir un lien entre profil de buts, rapport au certificat, et expérience au sein des MOOC.

Se pose enfin la question du lien entre le rapport au certificat et les caractéristiques sociodémographiques des apprenants, notamment. L'enquête fondée sur un questionnaire que nous avons diffusé dans douze MOOC [10] a montré que dans les participants expriment un intérêt plus marqué pour les certificats gratuits dans les pays à faible développement humain, et ce même chez les apprenants déjà diplômés. On peut émettre l'hypothèse selon laquelle des certificats de MOOC d'institutions considérées comme prestigieuses par l'apprenant peuvent être perçus comme un signal d'intérêt dans un marché de l'emploi compétitif. Ce premier constat issu d'enquêtes quantitatives gagnerait à être complété par des approches qualitatives plus susceptibles de mettre en évidence les mécanismes à l'œuvre.

6 Références

1. Arendt, A. M., et Shelton, B. E. (2009). Incentives and Disincentives for the Use of OpenCourseWare. *The International Review of Research in Open and Distributed Learning*, *10*(5).
2. Beaud, S. et Weber, F. (2010). *Guide de l'enquête de terrain.* Paris : Armand Colin.
3. Bouchard, P. (1994). *Pourquoi apprendre seul : les déterminants du choix éducatif chez des professionnels autodidactes.* (Thèse de doctorat non publiée) Université de Montréal. Québec. Canada.
4. Boutinet, J.-P. et Collectif. (2009). *L'ABC de la VAE*. Toulouse : Erès.
5. Carré, P. (2001). *De la motivation à la formation.* Paris : L'Harmattan.
6. Carré, P. (2003). La double dimension de l'apprentissage autodirigé, Contribution à une théorie du sujet social apprenant. *Revue Canadienne pour l'étude de l'éducation des adultes*, 17(1).
7. Carré, P. et Fenouillet, F. (2009). *Traité de psychologie de la motivation.* Paris : Dunod.
8. Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D. et Emanuel, E. (2013). The MOOC phenomenon: Who takes massive open online courses and why? Consulté le 5 juin 2017 sur http ://papers.ssrn.com/sol3/papers.cfm?abstract_id=2350964
9. Cisel, M. (2017). Les MOOC français à l'épreuve de la viabilité économique. *Revue Française des Sciences de l'Information et de la Communication. 12(1).*
10. Cisel, M. (2018). Le MOOC vu comme un projet d'apprentissage. *Éducation et Formation, e-307, 7*.

11. Deci, E. L., Koestner, R. et Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*, *125*(6), 627.
12. Dumazedier J. (1962). *Vers une civilisation du loisir ?* Paris : Seuil.
13. Dumazedier, J. (2002). *Penser l'autoformation. Société d'aujourd'hui et pratiques d'autoformation.* Paris : Chronique Sociale.
14. Dweck, C. S. (1986). Motivational processes affecting learning. *American psychologist*, *41(10),* 1040.
15. Elliot, E. S. et Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, *54*(1), 5–12.
16. Elliot, A. J., et Harackiewicz, J. M. (1994). Goal setting, achievement orientation, and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *66*(5), 968–980.
17. Elliot, A. J. et Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of personality and social psychology*, 72(1), 218.
18. Elliot, A. J. et McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519.
19. Jézégou, A. (2009). Apprentissage autodirigé et formation à distance. *Distances et savoirs*, *6*(3), 343–364.
20. Johnstone, J. W. C. et Rivera, R. J. (1965). *Volunteers for Learning: A Study of the Educational Pursuits of American Adults.* Chicago : Aldine Publishing Company.
21. Kizilcec, R. F. et Schneider, E. (2015). Motivation as a Lens to Understand Online Learners. *ACM Transactions on Computer-Human Interaction* (TOCHI), 22(2).
22. Kizilcec, R. F., Pérez-Sanagustín, M. et Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers et Education*, *104*, 18–33.
23. Knowles, M. S. (1975). *Self-directed Learning: A Guide for Learners and Teachers*. New York : Association Press.
24. Koller, D., Ng, A., Do, C. et Chen, Z. (2013). Retention and intention in massive open online courses. *EDUCAUSE Review*. Consulté le 5 juin 2017 à https://er.educause.edu/articles/2013/6/retention-and-intention-in-massive-open-online-courses
25. Koller, D. (2015) An update on assessments, grades, and certification. Consulté le 5 juin 2017 à : http://coursera.tumblr.com/post/131520811622/an-update-on-assessments-grades-and
26. Long, H. (1989). *Self-directed learning: emerging theory and practice*. University of Oklahoma : Norman.
27. Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91*(3), 328–346.
28. Rosell-Aguilar, F. (2013). Delivering unprecedented access to learning through podcasting as OER, but who's listening? A profile of the external iTunes U user. *Computers et Education*, *67*, 121– 129.
29. Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, *45*(1), 89–125.
30. Vallerand, R. et Thill, E. (1993). *Introduction à la psychologie de la motivation*. Laval, Québec : Editions Etudes Vivantes.
31. Wormington, S. V., et Linnenbrink-Garcia, L. (2017). A new look at multiple goal pursuit: The promise of a person-centered approach. *Educational Psychology Review*, *29*(3), 407-445.

Instrumentation des pratiques d'accompagnement par des formateurs et des conseillers en lien avec l'usage d'un Serious Game

Pierre-André Caron

Univ. Lille, EA 4354 – CIREL – Centre Interuniversitaire de Recherche en Education de Lille, F-59000 Lille, France
pierre-andre.caron@univ-lille.fr

Résumé. Cette communication rend compte de l'instrumentation des pratiques d'accompagnement menée par des formateurs et des conseillers en lien avec l'usage d'un dispositif de type Serious Game qui a pour but de favoriser chez les apprenants la compréhension et l'identification de compétences transversales. Nous mobilisons pour cette recherche le cadre théorique du modèle d'acceptation de la technologie amendé par nos travaux. Une première enquête, quantitative par questionnaire, nous permet de justifier la décomposition proposée par notre modèle, d'explorer les différentes modalités de l'utilité perçue par les usagers du dispositif de formation et de montrer, dans le contexte étudié, comment l'utilisabilité et l'injonction discriminent son usage réel.

Mots-clés. Serious Game, formateur, injonction, dispositif, modèle d'acceptation de la technologie

Abstract. This paper reports on the results obtained concerning the instrumentation of support practices by trainers and counsellors in connection with the use of a Serious Game-type system designed to encourage learners to understand and identify transversal skills. We mobilize the theoretical framework of the technology acceptance model amended by our work in the context of introduction and use of a Serious Game by trainers and counsellors. An initial quantitative questionnaire survey explores the different modalities of perceived utility, justifies the decomposition proposed by our model, and shows, in the context studied, how usability and injunction discriminate the device used.

Keywords. Serious Game, trainer, device, technology acceptance model, professionalization process.

1 Introduction

Cette communication rend compte d'un recherche portant sur la nature de l'instrumentation des pratiques d'accompagnement par des formateurs et des conseillers en lien avec l'usage d'un dispositif de formation dans lequel s'inscrit un Serious Game. Dans un premier temps, nous proposons d'articuler les problèmes spécifiques, rencontrés par les formateurs et les conseillers interrogés, aux questions de recherche que nous nous posons. Ces problèmes concernent d'une part l'appropriation de la notion de compétence transversale et d'autre part l'appréhension du dispositif numérique SkillPass, les questions qui sous-tendent notre recherche prennent alors la forme suivante :

• Quelles sont les perceptions des professionnels utilisant SkillPass ?

• Quels intérêts trouvent-ils à son usage, sur le plan de l'utilité et de l'utilisabilité du dispositif SkillPass ?

• Quels sont les rapports qu'ils entretiennent avec l'injonction institutionnelle et collégiale qui leur est faite d'intégrer les outils numériques à leurs pratiques d'accompagnement ?

Dans un deuxième temps, nous réalisons un état de la recherche et nous proposons de mobiliser le cadre théorique du modèle d'acceptation de la technologie (TAM))[1] amendé par nos travaux. Nous formulons ainsi la problématique initiale de notre recherche : *« Appréhender, au regard de la théorie de l'acceptation des technologies, les perceptions et motivations des formateurs utilisant le dispositif SkillPass »*. Notre recherche s'inscrit dans une démarche déductive, pour autant, certains résultats propres aux rôles particuliers que jouent l'injonction collégiale et l'injonction institutionnelle nécessitent pour être interprétés de compléter, d'une part les référents théoriques, d'autre part le cadre d'interprétation et d'analyse. Cet amendement a pour conséquence une problématique de recherche consolidée portant sur: *« La compréhension de l'impact du dispositif SkillPass sur les pratiques et la professionnalisation des formateurs et des conseillers et la perception que les formateurs et les conseillers ont de l'utilité réelle du dispositif SkillPass quant à ses finalités exprimées »,* dans la partie « perspectives » de notre communication, la mise en place d'une démarche inductive est alors évoquée.

Dans un troisième temps, nous proposons de présenter les démarches d'enquêtes qui ont été réalisées : la première, dont nous rendons compte dans cette communication, est quantitative par questionnaire, elle permet d'explorer les différentes modalités de l'utilité perçue et de justifier la décomposition proposée par notre modèle. La seconde démarche est qualitative sur la base d'entretien, bien que déjà réalisée, cette démarche inductive n'est évoquée que dans la partie perspectives de notre communication, elle mêle analyse thématique de contenu et analyse lexicale. Elle permet de mettre en évidence, chez les formateurs et les conseillers, le lien qu'ils établissent entre le dispositif SkillPass et la formation à l'identification des compétences transversales, elle permet de comprendre comment le dispositif SkillPass agit sur les pratiques professionnelles et comment plus généralement les actions d'accompagnement instrumentées s'inscrivent pour les formateurs et les conseillers dans une démarche de professionnalisation.

2 Problématique et cadre d'analyse

L'arrivée d'une nouvelle technologie dans un cadre professionnel signifie souvent, pour les acteurs de terrain, l'arrivée d'un changement dans la manière dont ils exercent leur profession ; comme tout évènement saillant, ce changement peut alors être instigateur d'une posture réflexive, qui permet à l'acteur de s'interroger sur l'évolution de ses pratiques. Le contexte de notre recherche, l'introduction d'un Serious Game dans des pratiques d'accompagnement et de formation, est donc riche de cet évènement qui invite formateurs et conseillers à une introspection professionnelle. Le dispositif que nous étudions est composé d'actions de formations impliquant l'usage d'un Serious Game (SkillPass) qui a pour but de favoriser chez les apprenants la compréhension et l'identification de compétences transversales. C'est, pour notre recherche, l'occasion de tenter de capturer, en ce moment privilégié, le sens que les formateurs et conseillers (utilisateurs du dispositif) donnent à l'intégration (puis à l'usage) des technologies dans le cours d'une activité d'accompagnement et de formation.

Notre recherche adresse deux problèmes rencontrés par les formateurs dans l'usage du dispositif proposé. Le premier concerne la finalité du dispositif, c'est dire la formation de jeunes adultes à l'identification de compétences transversales. Cette notion polymorphe est en soit difficile et parfois mal maîtrisée par les formateurs eux même[2]. Le second problème concerne l'appréhension du dispositif numérique pour répondre à cette finalité. Le dispositif SkillPass ne se limite pas à l'usage d'un Serious Game, c'est un dispositif riche qui instruit une triple alternance : la première entre les temps de jeu et les temps d'échange, la deuxième entre le lieu virtuel du jeu et le lieu réel de la salle où sont accompagnés les jeunes, la troisième entre le rôle de l'accompagnant que le jeune tient dans le jeu et le rôle d'accompagné qu'il tient lors des séances de mise en posture réflexive. Pour définir les liens que les formateurs et conseillers tissent entre ces deux problèmes nous proposons d'interroger dans ce contexte, l'intention d'usage et l'acceptation qu'ils ont du dispositif SkillPass.

Dans la littérature scientifique, l'intention d'usage est souvent corrélée à la notion d'acceptation au travers des notions d'utilité et de facilité d'utilisation. En s'inspirant du modèle attente-valeur de l'action raisonné conçu par Fishbein & Ajzen [3], Davis [4] propose le cadre théorique du modèle de l'acceptation des technologies, l'auteur postule que l'intention d'usage d'une technologie est un bon prédicteur de son usage et qu'elle dépend principalement de deux facteurs, la perception l'utilité et la perception de la facilité d'utilisation (Fig. 1).

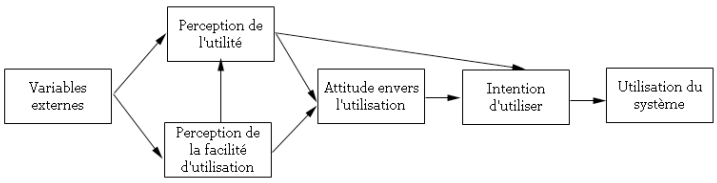


Fig. 1. Modèle du TAM selon Davis [4]

Pour notre recherche, mobiliser ce cadre théorique se justifie pour plusieurs raisons, d'une part c'est un cadre validé internationalement par des méta-recherche [5], [6], d'autre part c'est un cadre souvent mobilisé en formation [7], enfin, sa simplicité permet une mise en œuvre et une adaptation assez simple dans notre contexte. Depuis les travaux de Davis, de nombreux autres travaux ont été menés pour compléter le modèle prédictif [8], [1], [9], [10]. Or une des difficultés d'appliquer ce modèle, est que, si de nombreuses études ont été menées dans des contextes de formation [11], la plupart concerne l'intention usage des apprenants et très peu concernent l'intention d'usage des enseignants ou des formateurs. Seules huit études ont ainsi été recensées pour ce public par Nucci [12] et toutes ont été menées sur des enseignants en position de formés. Dans le domaine des Serious Games, le modèle TAM a été récemment complété [13] pour y intégrer comme nous le faisons la dimension injonction et le concept d'innovation personnelle, mais toujours pour des utilisateurs en position de formé. Notre contexte de recherche se distingue donc de ces recherches car il concerne des formateurs, public pour lequel nous postulons que la nature indissociable des tâches d'enseignement et d'apprentissage le différencie des autres publics. Autrement dit, nous faisons l'hypothèse que, vis à vis de la théorie de l'acceptation des technologies, notre public est spécifique, dans la mesure où ce que le formateur imagine de l'usage des apprenants pourrait influer sur l'appréhension qu'il a du dispositif.

Nous postulons (Fig. 2) que, pour notre public composés de formateurs et de conseillers, la nature duale de l'activité ciblée « Formation **et** Apprentissage », influe sur la perception d'utilité qu'ils ont du dispositif. Pour prendre en compte cet élément nous distinguons « l'utilité pour le formateur » de « l'utilité pour le jeune».

L'utilité pour le jeune (UTE) concerne la projection que fait le formateur de l'utilité de l'outil pour son public. Par exemple, le formateur peut penser qu'avec le dispositif SkillPass, les jeunes comprennent mieux comment identifier les compétences qu'ils ont construites.

L'utilité propre concerne l'estimation que le formateur se fait de l'utilité de l'outil pour lui même. Dans notre modèle cette utilité se décompose en trois sous dimensions :

- L'utilité pour moi même (UPM), c'est, indépendamment du contexte du dispositif SkillPass, ce en quoi son usage favorise la compréhension de la notion de compétence et les bénéfices espérés pour la formation et l'accompagnement. Cette dimension que nous postulons, se rapproche des dernières applications du TAM au contexte des Serious Games [13].
- L'utilité générale (UTG), cette utilité concerne tout les bénéfices professionnels (autres que pour former / accompagner / ou se familiariser à la notion de compétence) que le formateur ou le conseiller peut avoir de l'usage du dispositif.
- L'utilité pour former aux compétences (UPF), c'est l'utilité principalement ciblé par le dispositif, ce en quoi le dispositif facilite l'action des conseillers et des formateurs lorsqu'ils accompagnement des jeunes à l'identification des compétences qu'ils ont développées.

Notre modèle est complété par deux dimensions :

L'utilisabilité (UTB) concerne la facilité pour le formateur à utiliser le dispositif. L'utilisabilité est corrélée avec d'autres concepts tels que l'ergonomie, l'efficacité, la satisfaction.

L'injonction (UTI) concerne la perception qui pousse les formateurs à accepter le dispositif plus pour des raisons d'appartenance à un organisme de formation, un environnement collégial plutôt que par convictions pédagogiques. Cette injonction est une forme de contextualisation du concept d'acceptation sociale présente dans le modèle dont nous nous inspirons [1] et repris par les derniers travaux appliquant le TAM aux Serious Games [13].

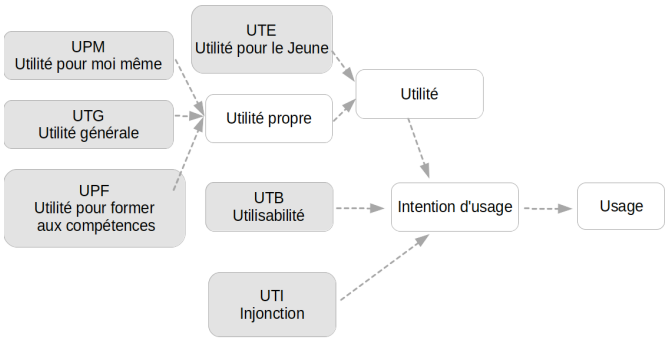


Fig. 2. adaptation par l'auteur du modèle TAM pour le projet LudoErgoSum, 2018.

Pour mener à bien cette recherche nous mobilisons donc le cadre théorique de l'acceptation de la technologie que nous complétons pour sa dimension utilité et que nous précisons dans sa dimension acceptation sociale. Notre objectif initial de recherche consiste, à ce stade, à interroger, au regard de la théorie de l'acceptation des technologies, les perceptions des formateurs utilisant le dispositif SkillPass.

3 L'étude menée

Sur le plan méthodologique, deux démarches complémentaires ont été menées. Notre communication rend compte de la première qui est de type quantitative par questionnaire, elle permet d'explorer les différentes modalités de l'utilité perçue, de discriminer les différents profils d'utilisateurs et de quantifier l'influence de l'injonction dans la mise en œuvre du dispositif. Une seconde démarche, qualitative sur la base d'entretien, est abordée dans les perspectives de nos travaux. Cette démarche mêle deux analyses complémentaires : une analyse thématique de contenu et une analyse lexicographique. Ces deux analyses explorent chez les formateurs et les conseillers les perceptions qu'ils ont du dispositif SkillPass, et la manière dont le dispositif participe à identification des compétences transversales chez les apprenants.
L'étude quantitative, a pris la forme d'un questionnaire général de 133 questions dont 34 concernaient notre recherche LudoErgoSum. Ces 34 questions ont été regroupées en 6 dimensions correspondant au modèle théorique projeté [14] : UTF (utilité pour

former), UTG (utilité générale), UPM (utilité pour moi même), UTE (utilité pour le jeune), UTB (utilisabilité), UTI (injonction). 128 formateurs et conseillers ont été contactés, 88 ont répondu au questionnaire, 63 ont répondu entièrement, parmi ces 63 répondants certains ont suivi une formation mais n'ont eu qu'un usage occasionnel du dispositif, seuls 32 répondants ont eu un usage conséquent du dispositif SkillPass indépendamment de l'action de formation. L'étude porte donc, selon le type de questions, sur les 63 ou sur les 32 répondants. Une première analyse descriptive, portant sur les 63 répondants, s'appuie sur les indicateurs de position et de dispersion (Fig. 3), les réponses sont exprimées sous forme de likert de 5.

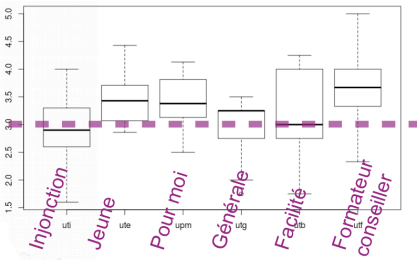


Fig. 3. Boites à moustache résumant les indicateurs de position et de dispersion

Ces indicateurs montre que l'appréhension du dispositif SkillPass est dans toutes ses dimensions globalement positive, toutes les moyennes étant supérieures à 3. Ils précisent la grande variabilité de la perception d'utilisabilité, c'est à dire la variabilité de la facilité d'usage du dispositif SkillPass. C'est cette dimension qui montre la plus grande amplitude. Concernant les différentes dimensions de l'utilité, l'usage du dispositif est plébiscité.

- l'utilité du dispositif pour moi même (UPM)
- l'utilité du dispositif pour le former (UTF),
- l'utilité du dispositif postulé par le formateur pour le jeune (UTE).

Un deuxième analyse porte sur les réponses des 32 individus ayant continué d'utiliser le dispositif, cette analyse est réalisée en composantes principales (ACP). Dans un premier temps (Fig.4) nous calculons la contribution de chaque dimension aux axes proposés par l'ACP par la méthode des cosinus carrés.

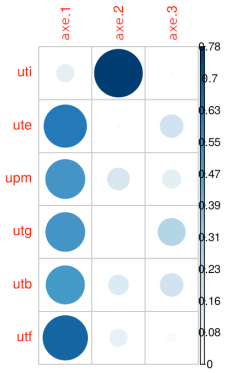


Fig. 4. Contribution de chaque dimension aux deux axes de l'ACP

Nous voyons que toutes les dimensions autres que l'injonction contribuent au premier axe de l'ACP, le deuxième axe est quand à lui principalement caractérisée par la dimension injonction. L'inertie des axes factoriels suggère le nombre judicieux de composantes principales à étudier : les 2 premiers axes de l' ACP expriment 68.7% de l'inertie totale du jeu de données (Fig.5) ; cela signifie que 68.7% de la variabilité totale du nuage des individus et des variables est représentée dans ce plan. C'est un pourcentage assez important, et le premier plan représente donc convenablement la variabilité contenue dans une grande part du jeu de données actif.

Dans un deuxième temps (Fig. 5) nous affichons donc le graphique 2d de l'ACP concernant les deux premiers axes. L'analyse en composante principale (ACP) a pour but d'identifier les dimensions qui rendent le mieux compte de la variabilité des résultats. Cette analyse structure les perceptions des utilisateurs selon deux axes. Le premier concerne les différentes dimensions de l'utilité (en particulier utilité pour moi même et utilité pour le formateur), qui comme on peut s'y attendre sont pour les formateurs de bons discriminants à large spectre de l'utilité du dispositif SkillPass. Les distinctions que nous faisons dans cette recherche des différentes formes de l'utilité ne sont pas clairement illustrées par cette ACP (les vecteurs UTG, UTE et UTF sont presque colinéaires). Le second axe concerne quant à lui la perception de l'injonction, l'analyse nous apprend qu'il peut expliquer 20 % de la variabilité et que c'est un facteur qui est peu corrélé avec la perception d'utilité (les vecteurs sont perpendiculaires). Comme vu précédemment ces deux dimensions, utilité et injonction, suffisent à expliquer près de 70 % des perceptions des usagers.

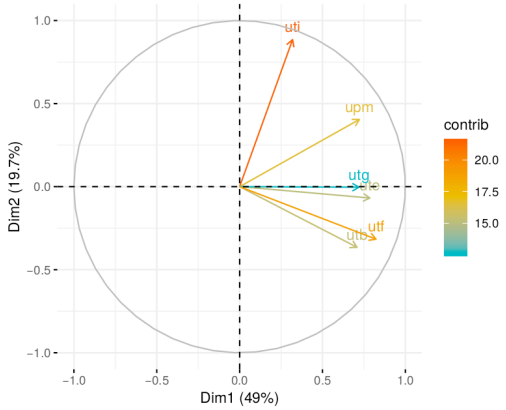


Fig. 5. Analyse en composantes principales

On remarque que deux vecteurs se distinguent également : l'utilité pour moi même et l'utilité pour former, la première dimension renvoie à des aspects de l'utilité dans le but de se professionnaliser, la seconde aux activités de formation. C'est un des résultats de cette étude que de montrer que dans l'usage de SkillPass la perception de l'utilité en terme de professionnalisation constitue un champ relativement indépendant à explorer, c'est ce que nous proposons dans les perspectives de notre recherche. Enfin nous voyons que les vecteurs de l'utilité pour former et de l'utilisabilité sont presque colinéaires, c'est un résultat déjà constaté du modèle d'acceptation des technologies qui montre que la perception d'utilisabilité influe sur la perception d'utilité cf (Fig. 1).

A ce stade de notre recherche nous convoquons les travaux [15], [16] et [17] qui proposent de classer les usagers d'une innovation selon 4 catégories : Objecteurs, Pragmatiques de la continuité, Pragmatiques du Changement, Passionnés. Les travaux de Moore [15] postulent qu'il existe un gouffre entre les passionnées et les 3 autres catégories, et que l'adoption d'une innovation par les passionnés n'induit pas l'adoption par les trois autres catégories.

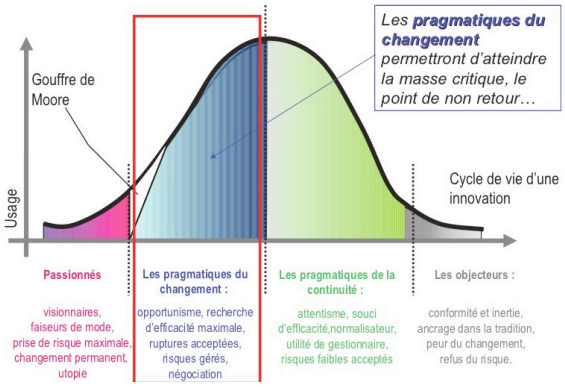


Fig. 6. Classement en quatre catégories [17]

Sur la base des réponses au questionnaire, nous proposons de mettre en œuvre une Classification Hiérarchique sur Composantes Principales (HCPC), cette classification a pour but de mettre en évidence des profils des répondants et de comprendre la dynamique de l'acception des technologies selon les différents profils.

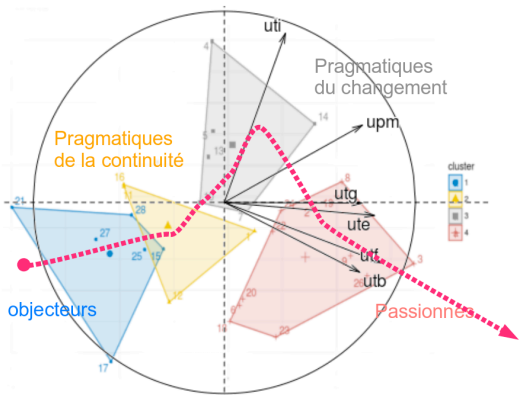


Fig. 7. Classification Hiérarchique sur Composantes Principales

L'HCPC que nous réalisons met en évidence quatre profils dont d'ordonnancement suit globalement la dimension horizontale de l'utilité. Notre recherche identifie (Fig. 7) en bleu les objecteurs, puis en jaune les pragmatiques de la continuité, en gris les pragmatiques du changement, en rose les passionnés. Si l'axe principal est bien l'axe horizontal, nous remarquons que les dimensions injonction et utilisabilité jouent un rôle certes secondaire mais loin d'être neutre. La première discrimine les trois premiers profils, la seconde discrimine le groupe des passionnés. L'hypothèse de Moore qui postule un différentiel motivationnel entre les passionnés et les autres catégories est ici clairement illustrée.

Pour les objecteurs, les pragmatiques de la continuité et les pragmatiques du changement c'est l'injonction qui détermine l'intention d'usage, pour les passionnés c'est l'utilisabilité qui est déterminante.

4 Conclusion et perspectives

En conclusion, nous souhaitons énoncer une première série de résultats : la perception du Serious Game SkillPass par le formateur ou le conseiller est globalement positive, bien qu'il y ait une grande variabilité de la perception d'utilisabilité. La perception de l'utilité n'est pas principalement reliée à l'usage du dispositif SkillPass pour former dans le domaine des compétences, elle est reliée de façon plus générale à tous les aspects de l'utilité et donc aussi par les conséquences qu'induisent ces aspects en terme de professionnalisation. De façon plus précise, il est possible de distinguer deux éléments influençant la perception d'utilité : pour les objecteurs, les pragmatiques de la continuité, les pragmatiques du changement, l'injonction est un élément de discrimination important, pour ces profils la perception de l'utilité du dispositif, et donc son usage, est directement corrélée à l'injonction perçue. Les utilisateurs qui sont les moins sensibles à l'injonction, quelle soient institutionnelle ou collégiale, sont aussi ceux qui trouvent le moins utile le Serious Game et le dispositif dans lequel il s'inscrit. Les utilisateurs les plus sensibles à l'injonction sont plus enclin à le trouver utile. L'utilisabilité du dispositif n'est vraiment discriminante que pour les utilisateurs les plus passionnés. L'analyse en composante principale des données recueillies par l'enquête quantitative met ainsi en évidence une dynamique au travers deux résultats (Fig8) :

Fig. 8. Dynamique de l'adoption du dispositif SkillPass

1- La **perception de l'utilité générale** du dispositif dans tous ses aspects comme discriminant principal de l'intention d'usage, 2- La dynamique entretenue par **l'injonction** (institutionnelle, collégiale), et **l'utilisabilité** du dispositif, ce dernier facteur discriminant le groupe le plus actif.

Au titre des perspectives, nous abordons l'histoire de la recherche dans laquelle s'inscrit notre communication. Cette histoire ne peut pas, faute de place être relatée avec autant de détails dans cette communication, pour autant, nous souhaitons en décrire les enjeux et les principaux résultats. A la suite des travaux dont nous venons de rendre compte, nous avons produit deux autres analyses à partir d'une enquête menée sous forme d'entretiens qui abordent les thèmes relatifs aux pratiques émergentes d'accompagnement. En terme de volume, cette enquête porte sur des extraits de 12 entretiens auprès de formateurs et de conseillers, dont la retranscription

correspond à un texte d'environ 130 000 mots. La première analyse menée est thématique selon la méthode de L'Écuyer [18], la seconde est une analyse lexicale utilisant le logiciel Iramuteq [19]. La première explore le lien entre l'usage du dispositif SkillPass et l'acquisition de nouveaux gestes professionnels, elle montre qu'au de là de la simple animation de séances d'accompagnement de jeunes, il existe des enjeux individuels, collectifs et institutionnels à l'usage du dispositif. La deuxième analyse précise ces enjeux : le rôle ambigu du dispositif à la fois simple instrument mis à disposition des formateurs mais aussi catalyseur d'une injonction à l'action qui ré-interroge l'ensemble des actions professionnelles.

Au terme de cette communication, nous souhaitons évoquer les différentes démarches de recherche que nous avons mises en œuvre. Sur le plan méthodologique, la succession de méthodes déductives et inductives, le recours à des analyses quantitatives, qualitatives et mixtes, se révèlent opérantes pour aborder ce problème multi-niveau ; elles nous permettent d'analyser assez finement l'impact de l'introduction du dispositif, tant au niveau micro de l'action, qu'au niveau meso de son organisation ainsi qu'au niveau macro de l'institution et des stratégies de professionnalisation. A l'issue de ces recherches, se pose alors la question des nouveaux possibles. Or la recherche menée montre comment l'usage du dispositif induit pour tous ses acteurs une injonction à l'action, cette injonction semble en partie suscitée par l'articulation des différents espaces mis à disposition des usagers du dispositif : espace **virtuel** du jeu *versus* espace **réel** de la salle. D'où notre conclusion sous forme de question : En quoi cette articulation de différents espaces permet elle, pour le jeune, la mise perspective des expériences vécues et la formalisation de compétences, et en quoi permet elle, pour le formateur, le développement et le partage de nouveaux gestes professionnels propices à une forme de professionnalisation ?

Reference

1. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS Q. 425–478 (2003).
2. Baudouin, J.-M.: La compétence et le thème de l'activité: vers une nouvelle conceptualisation didactique de la formation. De Boeck Supérieur (2002).
3. Fishbein, M., Ajzen, I.: Belief, attitude, intention and behavior: an introduction to theory and research. (1975).
4. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Q. 13, 319–340 (1989).
5. King, W.R., He, J.: A meta-analysis of the technology acceptance model. Inf. Manage. 43, 740–755 (2006).
6. Yousafzai, S.Y., Foxall, G.R., Pallister, J.G.: Technology acceptance: a meta−analysis of the TAM: Part 1. J. Model. Manag. 2, 251–280 (2007).
7. Holden, H., Rada, R.: Understanding the influence of perceived usability and technology self-efficacy on teachers' technology acceptance. J. Res. Technol. Educ. 43, 343–367 (2011).
8. Tricot, A., Plegat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G., Morcillo, A.: Utilité, utilisabilité, acceptabilité: interpréter les relations entre trois dimensions de l'évaluation des EIAH. In: Environnements Informatiques pour l'Apprentissage Humain 2003. pp. 391–402. ATIEF; INRP (2003).

9. Venkatesh, V., Bala, H.: Technology acceptance model 3 and a research agenda on interventions. Decis. Sci. 39, 273–315 (2008).
10. Février, F.: Vers un modèle intégrateur « expérience-acceptation » : rôle des affects et de caractéristiques personnelles et contextuelles dans la détermination des intentions d'usage d'un environnement numérique de travail, (2011).
11. Luethi, J.: Perception des étudiants quant à l'utilité et l'utilisabilité d'un portfolio numérique: étude de cas à l'Université de Genève, (2012).
12. Nucci-Finke, C.: Les enseignants et le e-learning - Facteurs d'adoption ou de rejet du e-learning, dans un contexte de formation des enseignants, (2015).
13. Malaquias, R.F., Malaquias, F.F.O., Hwang, Y.: Understanding technology acceptance features in learning through a serious game. Comput. Hum. Behav. 87, 395–402 (2018).
14. Caron, P.-A.: Echelle TAM_INJ, http://refa.univ-lille.fr.
15. Moore, G.A.: Crossing the chasm. Harper Business Essentials (1991).
16. Rogers, E.M.: Diffusion of Innovations: Modifications of a Model for Telecommunications. In: Stoetzer, M.-W. and Mahler, A. (eds.) Die Diffusion von Innovationen in der Telekommunikation. pp. 25–38. Springer, Berlin, Heidelberg (1995).
17. Mallein, P., Peyrin, J.: Conception Assistée par l'Usage pour les Technologies, l'Innovation et le Changement». M Dubois Al Étude Socio-Cogn. Usages Multiméd. Rapp. Final. 52–57 (1998).
18. L'Écuyer, R.: L'analyse de contenu: notion et étapes. Méthodes Rech. Qual. 49–65 (1987).
19. Ratinaud, P., Déjean, S.: IRaMuTeQ: implémentation de la méthode ALCESTE d'analyse de texte dans un logiciel libre. Modélisation Appliquée Aux Sci. Hum. Soc. MASHS. 8–9 (2009).

Session
Analyse d'usage

L'Orthodyssée des Gram
Stratégies d'étayage et verbalisations du raisonnement métalinguistique lors d'interactions logopédiste - enfant autour d'un jeu grammatical en ligne

Mireille Rodi, Thierry Geoffre, Kostanca Cuko

Haute École Pédagogique Fribourg, UR DidaLang, 36 rue de Morat,
CH-1700 Fribourg, Suisse
rodim@edufr.ch
geoffret@edufr.ch
cukok@edufr.ch

Résumé. *L'Orthodyssée des Gram* est une application en ligne qui utilise un algorithme de « calcul » de phrases grammaticalement correctes. Actuellement, quatre jeux existent, permettant d'avoir différentes entrées de sensibilisation (2ème primaire) puis d'approfondissement (3ème à 6ème primaire) des notions de constituants de la phrase et de chaines syntagmatiques. Notre communication présente l'analyse des interactions logopédiste-enfant lors de son utilisation. Il s'agit de vérifier en quoi l'étayage de l'adulte peut soutenir la verbalisation des raisonnements métalinguistiques ; avec la perspective d'encourager ensuite les enseignants à susciter de tels échanges au niveau de groupes de besoin dans des classes ordinaires. Cet article présente le protocole d'expérimentation, la grille d'analyse et les premiers résultats issus de l'analyse de ce corpus.

Mots-clés. Application numérique, raisonnement métalinguistique, orthographe, grammaire, stratégies d'étayage.

Abstract. *L'Orthodyssée des Gram* is an online application that uses an algorithm to "calculate" grammatically correct sentences. At this time, four games are available, providing different ways of increasing awareness (2nd primary) and then improving (3rd to 6th primary) the notions of sentence components and syntagmatic chains. Our paper presents the analysis of speech therapist-child interactions during its use. The aim is to verify how adult scaffolding can support the verbalization of metalinguistic thinking, with the aim of encouraging teachers to initiate such exchanges with need groups in ordinary classes. This paper presents the experimental protocol, the analysis framework and the first results derived from the analysis of this corpus.

Keywords. Numerical application, metalinguistic thinking, spelling, grammar, scaffolding strategies.

1 Introduction

L'Orthodyssée des Gram est une application en ligne accessible à l'adresse : https://www.lafamillegram.ch/#/ (Geoffre, Hofer & Cochard, 2018). La proposition initiale des « tirettes » (Sève & Ambroise, 2009) a été convertie en application ludique soutenue par un algorithme de « calcul » de phrases grammaticalement correctes. Trois activités se sont greffées à cette première initiative, permettant d'avoir différentes entrées de sensibilisation (2e primaire) puis d'approfondissement (3e à 6e primaire) des notions de constituants de la phrase, de classes grammaticales, de fonctions et de chaines syntagmatiques. Les recherches sur l'utilisation du numérique dans l'apprentissage du langage écrit mettent rarement en exergue le fonctionnement métalinguistique de l'enfant (Cuko, Geoffre & Alvarez, 2019) ainsi que les liens entre ce raisonnement et les interventions étayantes d'un professionnel de l'enseignement ou de la santé (Marty, 2005). Notre étude, ancrée dans une approche interactionniste des processus d'acquisition/apprentissage du langage, se veut donc novatrice dans cette perspective. Elle a débuté à l'automne 2018 dans un contexte spécifique, l'interaction logopédiste[1] - enfant, avant une extension de l'expérimentation dans des classes.

2 Ancrage théorique

2.1 Ressources numériques et enseignement/apprentissage du langage écrit

À l'heure actuelle, diverses ressources numériques spécifiques sont utilisées ou développées dans le domaine de l'enseignement/apprentissage du français. Sont visés des objectifs précis[2] :

- en compréhension ;
- en production textuelle ;
- sur le plan phonologique;
- dans les domaines syntaxique et morphosyntaxique ;
- en orthographe lexicale et phonétique.

Suivant les intentions pédagogiques, l'outil numérique endosse des formes originales et ajustées au contenu. Bien qu'elles concourent à mieux appréhender l'incidence des ressources numériques sur la motivation et la participation des élèves, leurs compétences à planifier des tâches et leurs capacités d'analyse des informations, ces recherches questionnent rarement le bénéfice des TIC sur les processus d'acquisition/apprentissage des élèves dans le domaine du langage écrit et ne renseignent pas sur leurs éventuels progrès (Cuko, Geoffre & Alvarez, 2019). De même, peu d'informations sont fournies sur les démarches explicites des professionnels pour guider les enfants/élèves à exploiter au mieux cet outil en vue de construire des

[1] Par commodité d'usage dans ce manuscrit, nous utilisons le féminin pour désigner la fonction de logopédiste/orthophoniste

[2] Pour plus de détails, se référer à l'article de Cuko, Geoffre & Alvarez, 2019.

compétences langagières précises. C'est ce que propose notre étude centrée conjointement sur les stratégies d'étayage des adultes et les processus d'acquisition / apprentissage des enfants.

2.2 Stratégies d'étayage et processus d'acquisition/apprentissage du langage

Indissociable de la perspective interactionniste, le concept d'étayage (Vygotski (1934/1997 ; Bruner, 1983) est déterminé par une asymétrie de compétences entre les partenaires d'une dyade. Grâce à un processus *d'échafaudage («scaffolding process»,* Vygotski, 1934/1997), un enfant parvient non seulement à dépasser une éventuelle difficulté dans la résolution d'une tâche[3], mais également à se saisir d'une possible opportunité d'apprentissage (Rodi, 2018). Co-construites au sein de *formats* (Bruner, 1983)[4], ces structures interactives, dont le déroulement est invariant, offrent des conditions optimales pour l'actualisation de processus d'apprentissage, car elles exploitent la zone proximale de développement de l'enfant (Vygotski, 1934/1997). Sur le plan linguistique, les verbalisations de l'adulte sont donc directement subordonnées à celles de l'enfant. À ce niveau, l'étayage linguistique peut être défini comme une co-élaboration de productions verbales, ce processus étant ou non la conséquence d'une requête, d'une prévention ou d'une correction (Hudelot, 2005). Au-delà d'un simple apport d'informations, ces interventions sont une somme de conduites polysémiques guidant l'enfant sur différents plans, à la fois linguistique, cognitif, et relationnel. Les modalités d'étayage (étayage de la tâche ou étayage linguistique, de Weck & Salazar-Orvig, 2019)[5] vont différer en fonction de l'intention et du genre de tâche. Ainsi est-il possible de distinguer :

a) le *cadrage* (Hudelot, 2005) ou *régulation* (de Weck & Salazar-Orvig, 2019), basé sur le déroulement même de l'activité ;
b) le *guidage* (op.cit.), dont l'objectif est de faire participer l'enfant à la tâche en fonction du rôle attendu. L'adulte produit alors des *instructions*, *des initiations* (Hudelot, 2005) ou des *relances* (Hudelot, 2005 ; de Weck & Salazar-Orvig, 2019) de l'échange, qui peuvent prendre la forme de questionnements anticipatifs ou proactifs ou encore d'injonctions (Hudelot, 2005) ;
c) la prise de parole *à la place* de l'enfant avec une induction précise de son discours, par exemple en amorçant l'énoncé (ébauche verbale) ou en le reformulant (op. cit) ;
d) *l'évaluation* ou *ratification* à posteriori des énoncés de l'enfant (op. cit).

Appliquée à des moyens de médiation numériques, la modalité d'étayage sur la tâche implique de manière cooccurrente, et spécifique, un soutien technique lié à la manipulation de l'outil numérique (Tsourapi, Komis, & Baron, 2018). Néanmoins, ce soutien à des verbalisations techniques ne doit pas être supérieur aux régulations

[3] 6 fonctions sont définies pour l'étayage d'une tâche : l'enrôlement, la réduction des degrés de liberté, le maintien de l'intérêt sur la tâche et de l'attention sur les éléments pertinents de celle-ci, le contrôle de la frustration et la présentation de modèles de solutions (Wood, Bruner & Ross, 1976)

[4] A savoir des « *échanges habituels qui fournissent un cadre pour l'interprétation concrète de l'intention de communication entre mère et enfant* » (Bruner, 1983, p.171)

[5] Étude concernant des enfants de 5 à 7 ans en interaction avec leur mère.

linguistiques, qui démontreraient, dans ce cas, une manipulation trop coûteuse de l'outil informatique.

Cet ancrage théorique justifie notre procédure d'analyse des stratégies d'étayage produites dans les dyades logopédiste – enfant lors de la manipulation de *L'Orthodyssée des Gram* ».

3 Méthodologie

3.1 Objectif et hypothèses

L'objectif général de l'application « *L'Orthodyssée des Gram* » vise le développement d'un raisonnement métalinguistique, au-delà de la seule plausibilité sémantique : l'enfant est amené à réfléchir sur les liens entre les mots variables de la phrase (chaines d'accord) et à tester la grammaticalité des phrases potentielles. C'est à ce moment-là qu'il peut être amené à verbaliser sa démarche s'il travaille accompagné d'un adulte. Ainsi, en plus de tester l'utilisation de l'application et son éventuelle capacité à faire évoluer la réflexion des enfants/élèves, nous souhaitons vérifier deux hypothèses. D'une part, que les activités langagières proposées sont un vecteur pertinent pour accéder aux processus d'acquisition/apprentissage de l'enfant/élève, d'autre part qu'il est envisageable d'extraire des stratégies d'étayage spécifiques des professionnels pour le guider et exploiter efficacement le moyen de médiation. Nous supposons qu'un format (Bruner, 1983) caractéristique d'échanges entre les stratégies d'étayage et les raisonnements métalinguistiques de l'enfant va se profiler.

3.2 Récolte des données

Une première phase d'expérimentation est proposée à 3 logopédistes travaillant dans un cabinet indépendant avec des patients présentant des troubles du développement du langage écrit (6A03.0 / 6A03.1: Developmental learning disorder with impairment in reading and written expression[6]). Une seconde phase, en cours d'élaboration, est prévue en milieu scolaire. Concernant la première phase, sujet de cet article, nous avons recueilli nos données par le biais des enregistrements d'interactions logopédiste-patient[7], au moyen du logiciel Apowersoft®[8]. Ce logiciel permet d'enregistrer simultanément l'audio des verbalisations des interactants et la vidéo des actions de l'enfant à l'écran. Après avoir requis un consentement parental, l'application « *L'Orthodyssée des Gram* » est soumise à des patients, âgés entre 7 et 13 ans (4 au total), une fois par semaine, durant 10 séances. L'activité des « *Tirettes* » est proposée en premier lieu mais la logopédiste peut ensuite sélectionner un autre mode de jeu selon ses objectifs. L'enfant effectue un seul parcours de questions par séance et par mode de jeu. Quant à la logopédiste, elle étaye l'enfant selon ses pratiques thérapeutiques, pour

[6] Classification Internationale des Maladies : CIM11 (ICD-11 version 2018).
[7] En l'absence d'un expérimentateur afin de ne pas perturber la dynamique des échanges.
[8] Disponible à l'adresse : https://www.apowersoft.fr/apowersoft-enregistreur-decran-gratuit.

atteindre les objectifs qu'elle a déterminés. Le choix du niveau scolaire de l'activité dépend de l'évaluation des compétences orthographiques de chaque enfant (niveau propre). Il est choisi par la logopédiste.

3.3 Analyse

Nous disposons à ce jour d'un corpus de 40 séances. Ces séances sont transcrites au moyen de l'application ExpresScribe® et intégrées dans une grille. Notre méthode d'analyse est issue d'une démarche qualitative ascendante : à partir de maintes lectures du corpus transcrit, nous identifions des critères nous permettant de caractériser les stratégies d'étayage et les justifications de choix morphographiques effectués par les enfants. Une grille d'analyse réunit ces critères afin de classer les données et d'en obtenir une représentation simplifiée. Les stratégies d'étayage et les verbalisations métalinguistiques des enfants sont segmentées en interventions[9] afin de parvenir à les catégoriser précisément. Ces catégories (résumées dans le tableau 1) servent à décrire la construction des échanges et leur fonctionnement dans chaque interaction.

Tableau 1 Catégories d'analyse des stratégies d'étayage

Niv 1	Modalités	REGA	REGM	RAT
Niv 2	Forme	Consignes **CONS**	Consignes **CONS**	Evaluations **EVAL**
		Relances **REL**	Relances **REL**	Acceptations **A**
		Action non verbale **ACT**		
Niv 3	Niveau linguistique	Injonctions **INJ**		
		Commentaires **COM**		
		Demandes de clarification **DCLA**		
		Demandes de confirmation **DCONF**		
		Ebauches **EB**		
		Reformulations **REF**		
		Réponse **R**		

Trois niveaux d'analyse (Tableau 1) qualifient les stratégies d'étayage : la modalité, la forme, ainsi que le niveau linguistique.

a) La *modalité de l'étayage* permet de préciser si les stratégies concernent une régulation de l'activité (REGA, tableau 1), à savoir le cadrage et le déroulement de la tâche, une régulation métalinguistique (REGM, tableau 1), centrée sur le raisonnement morphographique qu'exige l'activité langagière ou encore une ratification (RAT, tableau 1), qui permet de clore un échange (voir exemple 2, L17).

b) La *forme de l'étayage* dépend de la modalité. Ainsi, les régulations (REGA ; REGM) sont réalisées principalement sous forme de consignes (CONS : instructions, explications sur le fonctionnement de la tâche) ou de relances (REL : sollicitations pour poursuivre l'activité ou approfondir le raisonnement) (exemples 1 et 2) ; alors que pour les ratifications (RAT), il peut s'agir d'une acceptation (A, tableau 1) de la proposition de l'enfant ou d'une évaluation (EVAL) de celle-ci, principalement sous forme d'un feed-back positif (exemple 3). Nous illustrons les formes d'étayage (REGM et REGA)

[9] Une intervention est un énoncé possédant une même valeur fonctionnelle (valeur illocutoire/genre discursif). Deux interventions différentes d'un même locuteur sont séparées par une pause, et/ou possèdent une valeur fonctionnelle différente (Bernicot, 1992).

à l'aide de trois exemples Le premier exemple se situe au début d'une séance de la dyade Noé – logopédiste A.

Exemple de forme d'étayage 1[10] : REGM

TDP	LOC	INTERVENTIONS	Modalité	Forme
1	L	alors tu vois i(l) y a une phrase écrite /	REGM	CONS
1	L	pis à chaque fois on a plusieurs possibilités /	REGM	CONS
1	L	on peut changer le mot qu(i) est là / pour mettre celui-là / ou celui-là / pis là on a aussi trois mots /	REGM	CONS
2	N	mhm !		
3	L	pis on doit essayer de faire le plus de phrases possibles	REGM	CONS
4	N	ok !		
10	N	le maitre écrit les devoirs au tableau		
11	L	est-ce que c'est une phrase qui marche ?	REGM	REL
12	N	ouais		

La logopédiste introduit le jeu (TDP 1 à 4). Pour ce faire, elle produit une série d'interventions de type consignes visant à préciser le déroulement de l'activité langagière (L1 ; L3). Quelques tours de parole plus tard (TDP 10 – 12), suite au choix que l'enfant verbalise (N10), la logopédiste produit une demande de confirmation (L11), qui a une fonction de relance (REL) sur le plan métalinguistique (de type : « *la phrase est-elle grammaticalement acceptable ?* »), puisqu'elle incite l'enfant à effectuer un raisonnement, afin de confirmer ou non sa sélection.

Dans l'exemple 2, l'enfant vient de terminer une tâche de l'activité.

Exemple de forme d'étayage 2 : REGA

TDP	LOC	INTERVENTIONS	Modalité	Forme
17	L	hé ! ça dit / bravo / tu as réussi ! /	RAT	EVAL
17	L	voilà ! / pis main(te)nant / tu peux chercher une autre phrase qu'on pourrait faire	REGA	REL
18	N	écr / le maitre (pointe chaque mot) / (tente de sélectionner une tirette)		
19	L	pour bouger / ouais / tu peux appuyer /	REGA	CONS
19	L	pis tu dois rEster appuyer / pis monter un peu comme ça	REGA	CONS

Après avoir lu les mots de l'item (N18), Noé tente avec peine de sélectionner une tirette. L'intervention de la logopédiste (L19 : CONS) a pour fonction d'expliquer la manière de s'y prendre pour parvenir à effectuer la sélection. A noter que les régulations de l'activité peuvent aléatoirement être réalisées par une action non verbale (ACT) (*sélectionne un item à la place de l'enfant*, par exemple).

L'exemple 3 présente un échange ayant lieu lorsque Noé vient de terminer avec succès une étape de l'activité des Tirettes.

[10] Légende : TDP : tour de parole V : verbe GN : groupe nominal
LOC : locuteur (**N** = **Noé** – prénom d'emprunt / **L** = **logopédiste**).

Exemple de forme d'étayage 3 : RAT

TDP	LOC	INTERVENTIONS	Modalité	Forme
60	N	(clique sur « vu » / phrase en vert)		
61	L	ouais tu l'as eu ! /	RAT	A
61	L	c'est bon /	RAT	EVAL
61	L	t(u) as trouvé la solution !	RAT	A

La logopédiste produit 3 interventions (L61) visant à ratifier sa proposition. D'une part, elle valide la réponse (première et troisième intervention), d'autre part, elle l'évalue positivement (deuxième intervention).

c) Le *niveau linguistique* de l'étayage permet de qualifier l'intention communicative du locuteur. Les interventions des logopédistes sont catégorisées selon 7 types d'intentions (tableau 1) : des demandes de confirmation (DCONF) et de clarification (DCLA), des injonctions (INJ), visant à faire agir l'enfant, des commentaires (COM) sur les actions des locuteurs, des ébauches (EB), visant à amorcer l'intervention de l'interlocuteur, des reformulations (REF), qui peuvent agir comme ratifications ou comme expansions d'une intervention de l'enfant, et des réponses (R) à une demande de ce dernier.

Ces catégories, ainsi que celles que nous avons définies pour analyser les raisonnements métalinguistiques des enfants, nous permettent de réaliser une segmentation du corpus en « *séquences* » d'acquisition/apprentissage métalinguistique[11] (SAAM), à savoir une suite d'échanges activant des processus acquisitionnels de type justifications métagraphiques, initiée par un des locuteurs (une demande, par exemple) et concernant un objet d'apprentissage langagier. La fin d'une séquence est généralement marquée par une ratification. Les séquences sont dénombrées sur la totalité du corpus. Grâce à cette segmentation, la dernière étape de notre analyse vise la définition de profils des processus acquisitionnels ayant lieu dans les interactions logopédiste – enfant (un *format* de séquence) et l'établissement de liens entre 1) les stratégies d'étayage des logopédistes, 2) les objectifs des activités et 3) les compétences actualisées par les enfants lors d'activités langagières en cours d'utilisation de l'application *l'Orthodyssée des Gram*.

4 Exemples d'analyse d'une séquence d'acquisition/apprentissage métalinguistique (SAAM - grille simplifiée)

L'exemple 4 présente la catégorisation des stratégies d'étayage d'une logopédiste (L) à partir des tirettes « la voiture roule / les voitures roulent ». C'est la SAAM14 de la première séance concernant Noé (N)[12].

[11] A noter que la démarche est en cours au moment de la rédaction de cet article.

[12] Noé est âgé de 10;03 ans et présente un trouble développemental du langage écrit (CIM11 6A03.0 / 6A03.1) ainsi qu'un déficit d'attention (CIM 11 6A05.0 Attention deficit hyperactivity disorder, predominantly inattentive presentation). Il est scolarisé en 5H.

Exemple 4 : Catégorisation des stratégies d'étayage. Dyade Noé – Logopédiste A.

TDP	LOC	INTERVENTIONS	Modalité	Forme	Niveau linguistique	SAAM
154	N	la voiture rou_ / (pointeur sur la tirette « V»)				14
155	L	alors qu'est-c(e) que tu voudrais mettre d'autre ?	REGM	REL	DCLA	14
156	N	(sélectionne l'item « roulent ») [rulã]				14
157	L	est-c(e) que ça te dit que(l)qu(e) chose quand c'est « ent » à la fin ? /	REGM	REL	DCONF	14
157	L	quand est-ce qu'on met « ent » ? /	REGM	REL	DCLA	14
157	L	toi / est-c(e) que ça te dit que(l)qu(e) chose ?	REGM	REL	DCONF	14
158	N	(actionne la tirette GN, sélectionne l'item « les voitures ») / attends / la voiture / roule (chuchote) / heu / les voitures (chuchote) / heu il y en a plusieurs?				14
159	L	exactement !	RAT	A	COM	14
160	N	bon on teste ça				14
161	L	on teste ça	RAT	A	REF	14
162	N	(clique sur « vu ») / (phrase en vert)				14
163	L	bravo ::!	RAT	EVAL	COM	14
164	N	bon / c'est bon !				14

Suite à la présentation de l'item « la voiture roule / les voitures roulent », Noé effectue dans un premier temps un choix erroné (156N : la voiture / roulent). La logopédiste intervient avec une régulation métalinguistique (157L : REGM), qui agit comme une relance en 3 temps, afin de centrer l'attention de l'enfant sur le sens de la flexion verbale *-ent*. Dans un premier temps, elle produit une demande de confirmation qui exige une réponse peu coûteuse cognitivement (réponse par oui ou non) centrée sur la flexion verbale (l'enfant a lu [rulã]). L'enfant ne prenant pas son tour de parole, elle poursuit avec une demande de clarification, suscitant une réflexion de type explicative, en lien avec une règle morphosyntaxique enseignée dans le contexte scolaire, puis termine l'échange par une demande de confirmation, centrée sur une relance du processus de réflexion. L'incidence de ces 3 interventions est visible au tour de parole suivant (158N) : l'enfant s'engage dans une démarche de raisonnement métalinguistique, par une verbalisation chuchotée, qui peut évoquer une forme de langage égocentrique (Vygotsky, 1934/1997) dont la fonction est de planifier sa réflexion. Il propose une justification sémantique (*il y en a plusieurs*), dont il demande confirmation à l'adulte qui ratifie la proposition (159L : RAT) par un commentaire affirmatif. La séquence se termine par une ratification (163L : RAT) sous forme de feed-back positif qui trouve écho dans l'intervention suivante de Noé.

L'exemple 5 est issu de la première séance de la dyade Élise[13] – Logopédiste B, à partir de la tirette « des feuilles colorées servent à bricoler ».

[13] Élise (E ; prénom d'emprunt) est âgée de 12;02 et scolarisée en 8H. Elle présente un trouble développemental du langage écrit (CIM11 6A03.0 / 6A03.1)

Exemple 5: Catégorisation des stratégies d'étayage. Dyade Élise – Logopédiste B

TDP	LOC	INTERVENTIONS	Modalité	Forme	Niveau linguistique	SAAM
164	E	des feuilles colorées servent à bricoler (actionne les tirettes et sélectionne « des » « feuilles » « colorées » « servent »)				10
165	L	ok /	RAT	A	COM	10
165	L	pourquoi t(u) as choisi ça ?	REGM	REL	DCLA	10
166	E	parce que c'est masculin pluriel				10
167	L	héhé pourquoi c'est masculin pluriel ?	REGM	REL	DCLA	10
167	L	où / c'est masculin ?	REGM	REL	DCLA	10
168	E	non féminin				10
169	L	pourquoi c'est féminin ?	REGM	REL	DCLA	10
170	E	parc(e) que / la feuille colorée				10
171	L	on peut dire la feuille colorée / héhé / donc c'est féminin	RAT	A	REF	10
172	E	singulier / pluriel				10
173	L	ok /	RAT	A	COM	10
173	L	qu'est-ce que t(u) as mis au féminin pluriel ?	REGM	REL	DCLA	
174	E	ben des feuilles colorées / le groupe nominal				10
175	L	oui /	RAT	A	COM	10
175	L	parc(e) que c'est le groupe nominal /	RAT	A	REF	10
175	L	ok / des feuilles colorées	RAT	A	REF	10
176	E	et « servent » / le verbe / vue qu'on avait dit que « des » ou « les » c'est comme [il] avec « s » ou elle avec « s »				10
177	L	oui ça montre que c'est au pluriel / et là du coup c'est féminin /	RAT	A	COM	10
177	L	c'est elles ou c'est ils ?	REGM	REL	DCONF	
178	E	elles				10
179	L	héhé / pourquoi / comment tu sais ?	REGM	REL	DCLA	10
180	E	parc(e) que on parlait avant que c'était au féminin / du coup tout doit être féminin				10
181	L	héhé exactement /	RAT	EVAL	COM	10
181	L	ok /?	REGM	REL	DCONF	10
181	L	donc c'est elle avec « s » (4 sec)	REGM	CONS	COM	10
181	L	donc du coup est-ce que la terminaison de verbe est correcte?	REGM	REL	DCONF	
182	E	oui				10
183	L	héhé / très bien	RAT	EVAL	COM	10
184	E	(valide la réponse)				10

Élise sélectionne les items adéquats (164E). Bien que la réponse soit appropriée, la logopédiste effectue une relance, sous forme de demande de clarification (165L : REGM) incitant un raisonnement métalinguistique chez l'enfant. La réponse d'Élise (166E) fait référence à un niveau de réflexion sur le genre et le nombre (niveau attendu), mais elle se trompe sur le genre du groupe nominal analysé, ce qui suscite une nouvelle relance (167L : REGM), à savoir deux demandes de clarification, la deuxième précisant la problématique (le genre). L'enfant peut ainsi reformuler sa réponse selon l'attente de la logopédiste, qui, néanmoins, relance un processus de réflexion métalinguistique explicatif (169L : REGM - DCLA). Élise argumente en précisant le déterminant qui définit le genre (170E), ce que la logopédiste ratifie par une reformulation (171L : RAT). Or l'enfant complète, en hésitant, sa réponse, par la définition du nombre du groupe nominal (172E). Cette hésitation est probablement à la source de la nouvelle demande de clarification que fait la logopédiste (173L : REGM - REL) qui semble vouloir s'assurer ici de la compréhension de la notion de genre et de nombre. La réponse appropriée d'Élise est ratifiée par 2 reformulations (175L : REGM). De sa propre initiative, l'enfant poursuit son raisonnement sur l'accord du verbe (176E), réflexion validée par un commentaire explicatif (177L : RAT). L'adulte attend une confirmation de la compréhension de cette notion grammaticale par deux relances : une demande de confirmation (177L : REGM) et une demande de clarification (178L : REGM) qui exhorte l'enfant à argumenter plus avant (180E). La réponse satisfait les attentes de la logopédiste qui ratifie la proposition par un commentaire évaluatif positif (181L : RAT) et Élise peut valider sa réponse (184E).

5 Discussion

Les deux séquences analysées mettent en exergue la construction des processus d'acquisition/apprentissage activés dans les interactions logopédiste – enfant. Chaque locuteur, en fonction de son rôle et de son statut, contribue à la mise en place de ce processus. Précisons que, dans ce contexte, les échanges présupposent l'existence d'un contrat, qualifié de « *pédago-thérapeutique* » (Rodi, 2018), qui engage l'adulte dans un rôle d'étayage des productions, offrant à l'enfant la possibilité d'actualiser ses compétences langagières. Ainsi, même si l'activité de l'enfant est dans la cible des attentes de la logopédiste, celle-ci ne se contente pas de ratifier la réponse. A maintes reprises, dans les 2 exemples analysés, des relances sont proposées à l'enfant pour lui suggérer un raisonnement métalinguistique permettant d'argumenter ses choix morphographiques. Ces régulations métalinguistiques respectent la zone proximale de développement, c'est à dire qu'elles sont adaptées à son niveau de compétences (les demandes adressées à Élise sont ainsi plus exigeantes sur le plan du raisonnement morphographique que celles proposées à Noé). Nous notons ici que la redondance des phrases proposées dans les différents jeux de l'application[14] diminue l'effet de double tâche lié à la lecture et permet à l'enfant/élève de se focaliser davantage sur une réflexion métalinguistique. Il est également intéressant de souligner que les interventions des logopédistes ne réfutent pas les choix des enfants. Probablement que les attentes ne sont pas définies en termes de résultats, de « juste » ou « faux », puisque c'est une réflexion liée aux représentations morphographiques qui est demandée au travers de la médiation du jeu informatique. Il est parallèlement possible que la logopédiste, en fonction de son statut de thérapeute, garante d'une alliance thérapeutique, effectue un travail de figuration (Goffman, 1973), en évitant de mettre en danger la face de l'enfant (un refus de sa proposition peut engendrer une frustration). Par ailleurs, nous identifions systématiquement, à la fin des séquences, des ratifications sous forme de feed-back positif. Ces ratifications ont pour fonction d'une part de clore la séquence, d'autre part de valider le travail de l'enfant. Nous interprétons tous ces indicateurs comme une manifestation du contrat pédago-thérapeutique (Rodi, 2018).

De son côté, l'enfant tente de s'approprier les propositions de son interlocuteur et de les mettre à profit. Les relances des logopédistes trouvent réponses mais, plus encore, elles suscitent un processus de réflexion qui se généralise au cours de l'interaction (sans demandes de son interlocuteur, Élise poursuit son raisonnement, par exemple). Ainsi, au sein de ces échanges, ont lieu diverses manifestations de co-influences, de co-constructions enclines à servir de support à l'apprentissage. Ces blocs d'échanges, que nous avons définis comme *séquences d'acquisition/apprentissage métalinguistiques*, sont identifiés tout au long des interactions. Nos analyses des verbalisations des enfants nous confirment fréquemment une évolution quantitative (nombre de SAAM) et qualitative (niveau des justifications), soulignant un processus d'acquisition/ apprentissage effectif.

La récurrence de ces échanges relève d'une structure organisée de comportements conversationnels, telles qu'on les retrouve dans les formats (Bruner, 1983). L'enfant peut progressivement anticiper les attentes du professionnel et s'engager dans une réflexion métalinguistique sans une intervention spécifique de ce dernier, ce qui

[14] Les mêmes phrases sont utilisées pour structurer les quatre jeux de l'application.

confirme le potentiel acquisitionnel de l'activité langagière proposée, ce qui sera l'objet d'un futur article.

6 Conclusion

Nous avons présenté une première étude des possibles bénéfices de l'utilisation de l'application *L'Orthodyssée des Gram* dans une interaction dyadique logopédiste-enfant. Cette première partie de l'étude vise à acquérir des données empiriques d'une part sur les verbalisations métalinguistiques des enfants, d'autre part sur les stratégies d'étayage des logopédistes. Comme nous l'avons constaté dans l'analyse des deux exemples, la grille développée dans ce travail permet d'extraire ce qui nous semble être des formats d'échanges co-construits par les interventions successives des interlocuteurs. Si l'on s'intéresse aux possibilités de différenciation, on constate qu'il est possible de sélectionner un niveau[15] en fonction des compétences de l'enfant/élève, ce qui permet au professionnel d'adapter les stratégies d'étayage à la zone proximale de l'enfant/élève et d'optimiser ainsi l'effet d'acquisition/apprentissage ; d'où l'intérêt d'une étude détaillée des verbalisations de chaque interlocuteur qui, à terme, permettra de dégager des profils d'interaction favorables à l'émergence de séquences d'acquisition/apprentissage métalinguistique.

À terme, notre projet vise une transposition de cette démarche en classe afin que les enseignant-e-s puissent susciter de tels échanges au niveau de groupes de besoin dans des classes ordinaires et d'en étudier les effets. Il est effectivement possible d'imaginer divers usages de cet outil, en situation individuelle d'apprentissage, tel que présentée dans cet article, mais également dans une forme collective, par exemple au tableau interactif (TBI) lors d'une phase de négociation graphique avec l'ensemble de la classe, ou en groupe. Lors de la transposition de cet outil numérique dans le contexte de la classe, nous examinerons également l'intérêt ou non d'introduire une fonction visant la construction d'une banque de phrases créées par des didacticiens (corpus de base) ou mutualisées par des enseignants (corpus dynamique).

Nous envisageons donc *l'Orthodyssée des Gram* comme une ressource numérique autonome, utilisable en individuel dans la classe ou avec un groupe d'élèves, en fonction d'objectifs précis, ou avec un patient en thérapie logopédique, ce dernier point étant l'objet du présent article. Nous pensons également son intégration à un environnement plus large, de type plateforme d'enseignement/apprentissage du français langue de scolarisation. L'application, en tant que support dynamique, engage l'enfant/élève dans une réflexion métalinguistique propice à une interaction plus fonctionnelle que les outils traditionnels. Cette interaction relève potentiellement d'une négociation graphique étayée par la logopédiste ou encouragée par un-e enseignant-e avec un groupe d'élèves, allant dans le sens de la négociation graphique comme exercice de classe propice aux apprentissages (Sautot & Geoffre, 2017).

[15] Ce niveau est dépendant des structures syntaxiques employées dans les phrases proposées et des flexions qui apparaissent sur les tirettes. Les phrases utilisées dans cette expérimentation ont été construites en suivant la progression des structures syntaxiques proposée par Sautot & Geoffre (2019) selon les degrés d'extension des chaines d'accord (Geoffre, 2013).

Références

1. Bernicot, J. (1992). *Les actes de langage chez l'enfant*. Presses Universitaires de France. Consulté le 10.01.19 à l'adresse https://www.cairn.info/les-actes-de-langage-chez-l-enfant--9782130447528.htm
2. Bruner, J. (1983). *Le développement de l'enfant* (8e éd.). Paris : Presses Universitaires de France - PUF.
3. Cuko, K., Geoffre, T., Alvarez, L. & Centre de Recherche sur l'Enseignement/Apprentissage par les Technologies numériques (C·R·E/A·TE). (2019). Les ressources numériques pour faciliter l'enseignement/apprentissage du français langue 1 : une revue systématique et complétée de la littérature pour orienter les recherches et les développements futurs. *Revues systématiques du C·R·E/A·TE*, 1. http://doc.rero.ch/record/324244
4. de Weck, G., & Salazar-Orvig, A. (2019). L'apport des études de corpus à l'analyse de l'étayage. *Corpus*, (19). Consulté le 10.01.19 à l'adresse http://journals.openedition.org/corpus/4173
5. Geoffre, T. (2013). *Vers le contrôle orthographique au cycle 3 de l'école primaire : étude psycholinguistique et propositions didactiques* (Thèse de doctorat). Université Stendhal-Grenoble 3, Grenoble. En ligne : https://www.researchgate.net/publication/323607223
6. Geoffre, T., Hofer, D. & Cochard, B. (2018). *L'Orthodyssée des Gram*. Consulté le 10.01.2019 à l'adresse https://www.lafamillegram.ch/#/
7. Hudelot, C. (2005). Pratiques interlocutives en crèche et en maternelle. Réflexions à partir de la description d'un poster par des petits groupes d'enfants de 2 à 3 ans. SRED, 11 : Scolariser la petite enfance ? *Actes du deuxième colloque «Constructivisme et éducation»* vol. *I*, *342*, 352.
8. Marty, N. (2005). *Informatique et nouvelles pratiques d'écriture*. Paris: Nathan.
9. Rodi, M. (2011). Interactions logopédiste – enfant : Comment se construisent des échanges potentiellement acquisitionnels dans une situation de jeu symbolique ? *Glossa*, *110*, 49-68.
10. Rodi, M. (2018). Les interactions logopédiste-enfant comme lieu d'apprentissage de capacités langagières. L'exemple des séquences potentiellement acquisitionnelles. *SHS Web of Conferences*, *46*, 10010. https://doi.org/10.1051/shsconf/20184610010
11. Sautot, J.-P. & Geoffre, T. (2017). De quoi la négociation graphique est-elle l'exercice? *Repères*, (56), 73-90. https://journals.openedition.org/reperes/1196
12. Sautot J.-P. & Geoffre T. (2019). Typologie de phrases pour l'étude de la langue ou la négociation orthographique. In *Scolagram*.
    En ligne : https://scolagram.u-cergy.fr/index.php/chapter-2/production-et-reception-du-metalinguistique/53-typologie-de-phrases-pour-l-orthographe
13. Sève, P., & Ambroise, C. (2009). Images, ciseaux, tirettes… Un exemple de bricolage didactique au CE1 autour des relations nom / verbe. *Repères*, (39), 103-123. https://doi.org/10.4000/reperes.373
14. Tsourapi, C., Komis, V., & Baron, G.-L. (2018). L'étayage des enseignants de l'école maternelle au cours des activités de programmation avec le logiciel ScratchJr. *In*: G. Parriaux, J-Ph. Pellet, G.-L. Baron, E. Bruillard & V. Komis (Eds.)*, De 0 à 1 ou l'heure de l'informatique à l'école Actes du colloque Didapro 7 – DidaSTIC*. Peter Lang. https://doi.org/10.3726/b13387/34
15. Vygotski, L.-S. (1934/1997). *Pensée et langage* (4e édition). Paris : La Dispute.
16. Wood, D., Bruner, J.-S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry and allied disciplines*, *17*, 100.

LabNbook, plateforme numérique support des pédagogies actives et collaboratives en sciences expérimentales

Cédric d'Ham[1], Claire Wajeman[1], Isabelle Girault[1], Patricia Marzin-Janvier[2]

[1] Université Grenoble Alpes, CNRS, Grenoble INP, LIG, F-38000 Grenoble, France
cedric.dham@univ-grenoble-alpes.fr,
[2] Université de Bretagne Occidentale, CREAD, F-29200 Brest, France

Résumé. Cette communication présente LabNbook, une plateforme numérique dans laquelle les apprenants créent et partagent des cahiers de laboratoire et/ou des rapports scientifique au sein d'un espace de travail structuré et supervisé par leurs enseignants. Nous décrivons les outils de production et de communication intégrés à la plateforme ainsi que les éléments didactiques et épistémologiques qui justifient leur conception. LabNbook a été diffusé à grande échelle dans deux établissements d'enseignement supérieur et nous décrivons quelques aspects de l'utilisation de la plateforme par les étudiants et enseignants.

Mots-clés. Plateforme numérique, pédagogies actives et collaboratives, sciences expérimentales, cahier de laboratoire, rapport scientifique

Abstract. This paper describes LabNbook, a learning management system dedicated to experimental sciences. In LabNbook, the learners create and share laboratory notebooks and/or scientific reports. The teachers structure and monitor the workspaces of their students. We describe the production and communication tools integrated in the environment and we justify their conception with epistemological and didactical elements. LabNbook has been deployed in a large-scale pilot study in higher education. We describe some aspects of its current usage by students and teachers.

Keywords. Online environment, active and collaborative pedagogies, experimental sciences, laboratory notebook, scientific report

1 Introduction

Dans le rapport *Les innovations pédagogiques numériques et la transformation des établissements d'enseignement supérieur* [1] adressé en juin 2018 à la ministre de l'enseignement supérieur, les auteurs écrivent : « La transformation pédagogique des premier cycles universitaires et la transformation numérique des établissements sont liées et susceptibles de se renforcer mutuellement : les IPN [Innovations Pédagogiques Numériques] sont en effet de nature à faciliter […] la formation des étudiants inscrits dans les premiers cycles de l'enseignement supérieur, en particulier en licence ». On peut cependant s'interroger sur l'utilisation des outils numériques en formation initiale dans l'enseignement supérieur. L'usage des plateformes pédagogiques génériques LMS (« Learning Management Systems ») telles Moodle, Chamilo, etc. s'est effectivement généralisé, mais de tels systèmes servent

généralement plus à administrer les cours et à échanger des ressources qu'à soutenir des pédagogies innovantes. Dans une étude menée par l'association EDUCAUSE [2], il est indiqué que, si les étudiants sont satisfaits de leur usage des LMS, les fonctionnalités les moins appréciées sont celles censées soutenir la *collaboration sur des projets* et la *participation à des groupes de travail*. A l'heure où les pédagogies actives et collaboratives sont désignées comme des méthodes d'enseignement à privilégier [3], quels sont les outils numériques disponibles pour les soutenir ?

Notre champ d'étude est l'apprentissage des sciences expérimentales : physique, chimie, biologie et géologie. Dans ces disciplines, il existe de multiples exerciseurs et simulateurs pour travailler sur des concepts particuliers (*cf.,* par exemple, le site PhET). Ces ressources peuvent être utilisées au sein d'une séquence pédagogique, mais ne soutiennent pas le déroulement de la séquence ni les interactions entre étudiants et enseignants. Pour ces objectifs, des environnements ont été produits par des équipes de recherche, principalement dans le cadre de l'apprentissage par investigation. Citons par exemple WISE [4], SCY [5] ou Go-Lab [6]. Ces environnements complexes répondent aux besoins des pédagogies pour lesquelles ils ont été conçus en intégrant différentes ressources et outils, tels que :

- des contenus de cours (textes, vidéos),
- des simulations,
- des données issues d'expériences (enregistrées ou pilotées à distance),
- des outils pour produire des inscriptions scientifiques (traitements de données, schémas, textes),
- des outils de communication synchrone ou asynchrone,
- des outils de régulation de l'activité (séquencement et partage de tâches notamment).

La contrepartie de ces environnements riches est qu'ils sont souvent complexes à configurer et se conforment à un modèle pédagogique prédéfini. Dans le cas où l'enseignant souhaite construire ses contenus d'apprentissage et choisir sa pédagogie, il utilise préférentiellement un environnement générique comme Google Drive ou Windows 360. Ces environnements n'ont pas été pensés à des fins pédagogiques : ils ne différencient pas les rôles étudiant / enseignant et ils n'intègrent pas de vue agrégée pour suivre le travail de l'ensemble d'une classe.

Un nouveau type d'environnements de travail apparait dans les cursus d'enseignement de certaines universités : les « Electronic Lab Notebooks ». Ces environnements, pour la plupart commerciaux, sont conçus à destination des laboratoires de recherche et développement. Ils sont utilisés par les chercheurs pour consigner et partager leurs résultats au sein de leur équipe. Si ces environnements sont plus adaptés au travail scientifique que Google Drive par exemple, ils n'ont pas non plus été initialement conçus pour un usage pédagogique et leur diffusion dans le monde de l'enseignement reste actuellement peu étendue.

L'objectif de cette communication est de présenter LabNbook, une nouvelle plateforme numérique pour l'apprentissage des sciences, et de justifier les choix de conception par des éléments épistémologiques et didactiques. LabNbook est accessible en ligne et permet aux apprenants (niveau lycée ou enseignement supérieur) de créer et partager un cahier de laboratoire et/ou un rapport scientifique. Cette plateforme a été conçue pour que les enseignants puissent aisément mettre en

œuvre les contenus et la pédagogie de leur choix : travaux pratiques, apprentissage par investigation, apprentissage par problème, pédagogie de projet…
Dans la suite de cet article, nous décrirons les outils constitutifs de la plateforme : outils de production, de collaboration, de communication et de suivi. Dans un deuxième temps nous ferons état des usages actuels de la plateforme qui a été déployée au sein de la COMUE Grenoble Alpes depuis 16 mois.

2 Des outils pour produire et écrire des contenus scientifiques

Dans *La vie de laboratoire*, Latour et Woolgar décrivent les chercheurs en science comme des personnes dont l'activité principale est de produire et échanger des « inscriptions littéraires » [7]. LabNbook est construit autour de cette idée : la communication écrite occupe une place centrale dans l'activité du scientifique et dans la construction de nouvelles connaissances. Deux processus sont intervenus dans la conception des outils de production de la plateforme : un processus de transposition didactique [8] de l'activité des chercheurs en sciences expérimentales a permis de caractériser les types d'écrits produits par les apprenants dans leurs apprentissages actifs ; un processus d'ingénierie informatique a permis de concevoir des outils utiles à la production de ces écrits.

2.1 Un processus de transposition didactique

Pour déterminer les types d'écrits utiles aux étudiants, nous nous sommes appuyés sur une transposition pour l'enseignement de la démarche expérimentale du scientifique. Ces dernières années, la démarche d'investigation a été la plus citée comme support de l'enseignement des sciences [9]. Cependant, LabNbook n'est pas dédié à une pédagogie particulière et nous avons préféré utiliser une description plus classique de la démarche expérimentale [10]. Nous avons fait évoluer cette description en lui donnant une structure itérative et en la faisant se développer dans les deux mondes des théories & modèles et objets & événements proposés par Tiberghien dans ses travaux sur la modélisation en sciences [11]. Notre transposition de la démarche expérimentale pour l'enseignement consiste en une séquence de sept types de tâches pouvant être assignés à l'étudiant (*cf.* Fig.1). Nous les décrivons succinctement :

- *Définir le problème :* le problème est pris avec une acceptation large : il peut correspondre soit à une investigation (par exemple déterminer quelle est la condition pour qu'un objet flotte) soit à un problème technique (par exemple déterminer la concentration d'une solution par un dosage) ; il peut être issu de réflexions théoriques ou de l'observation d'un fait singulier.
- *Proposer des hypothèses* : l'hypothèse, au sens large, peut être une relation entre grandeurs (cas de l'investigation par exemple) ou un résultat numérique anticipé (cas d'un problème technique par exemple).
- *Concevoir l'expérience* : ce type de tâche a pour objet de produire un protocole expérimental qui est un document ayant un statut bien précis en sciences expérimentales [12].

- *Manipuler et collecter des données* : les données peuvent être des phénomènes observés ou des mesures.
- *Traiter les données* : ce type de tâche correspond à appliquer des transformations mathématiques aux données numériques ou des traitements aux images collectées ; c'est ici qu'ont lieu les tâches de modélisation.
- *Interpréter les résultats* : ce type de tâche est central dans la démarche car c'est ici que la plupart des actions de contrôle sont effectuées.
- *Diffuser les résultats* : la diffusion peut se faire à destination de l'enseignant et/ou des pairs.

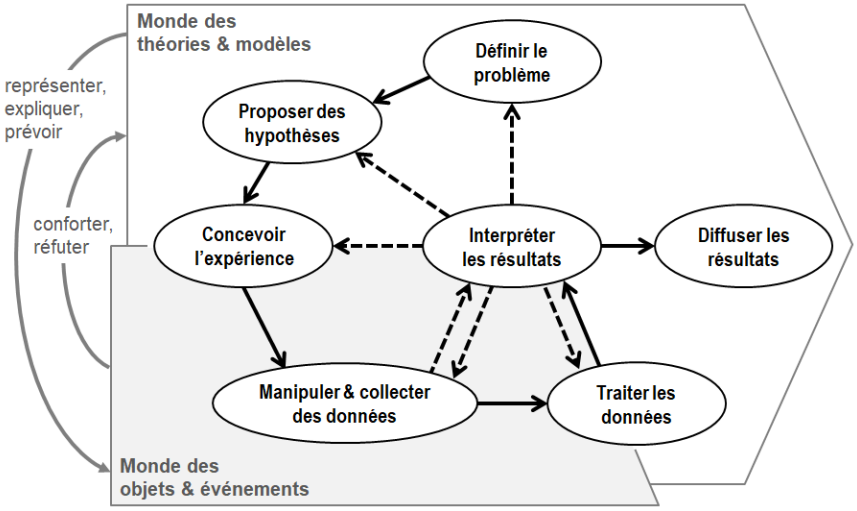


Fig.1. Une transposition en enseignement de la démarche du scientifique : le cycle expérimental

Les types de tâches de la démarche sont ici présentés séquentiellement mais cela ne signifie pas que la démarche expérimentale est linéaire. Les chemins en pointillés sur la figure correspondent aux différents retours que l'étudiant peut faire suite à des contrôles de l'activité. De plus, cette présentation ne signifie pas que dans une séquence d'enseignement, tous les types de tâches sont à la charge de l'étudiant : c'est l'enseignant qui, selon la pédagogie choisie, décide des tâches dévolues à l'étudiant. Finalement, nous avons incorporé dans le schéma l'idée que la démarche expérimentale met nécessairement en jeu des éléments de deux mondes : le monde des théories & modèles et le monde des objets & événements [11]. En situant les étapes du cycle expérimental dans ces deux mondes, trois types de tâches se trouvent placés à l'interface des deux mondes : *Concevoir l'expérience*, *Traiter les données* et *Interpréter les résultats*. Nous pensons que ces moments de la démarche sont centraux dans la construction de sens par les étudiants car ils les obligent à articuler des concepts ou relations issus du monde des théories & modèles avec des matériels, phénomènes ou mesures issus du monde des objets & événements.

2.2 Un processus d'ingénierie informatique

Le cycle expérimental que nous avons présenté n'est pas linéaire et l'enseignant peut choisir de confier uniquement certains types de tâches aux étudiants. La plateforme LabNbook ne doit donc pas contraindre le déroulement du travail selon une séquence prédéterminée de tâches. En ceci, notre plateforme se distingue de la plupart des environnements d'investigation : les tâches assignées aux étudiants et leur séquence sont à l'initiative de l'enseignant. Dans LabNbook, la définition d'une activité pour les étudiants débute nécessairement par un travail de l'enseignant au cours duquel il définit la « mission » qu'il souhaite proposer aux étudiants. Cette mission présente les tâches assignées aux étudiants et leur séquence à travers la structuration de l'espace de travail dans lequel les étudiants produisent leurs écrits. L'espace de travail est scindé par l'enseignant en parties, chacune contenant des consignes de travail et donnant accès à des outils de production au choix de l'enseignant. La figure 2 montre un exemple d'espace de travail étudiant pour une des missions de LabNbook. Selon les missions, l'espace de travail des étudiants peut correspondre à un cahier de laboratoire ou à un rapport scientifique, voire à un intermédiaire fréquemment rencontré dans l'enseignement supérieur, le compte-rendu de travaux pratiques.

Fig.2. Exemple de structuration en 4 parties de l'espace de travail des étudiants dans LabNbook

La plateforme intègre quatre outils de production avec lesquels les étudiants éditent les documents – les « LabDocs » – qui alimentent l'espace de travail : textes, dessins, protocoles expérimentaux et tableaux de données avec graphiques. Les outils de production ont été choisis en accord avec notre transposition de la démarche expérimentale. Le traitement de texte (TinyMCE) et l'outil de dessin (Zwibbler) sont des outils génériques et peuvent servir pour tout type de tâche. Dans le traitement de texte nous avons incorporé un éditeur d'équation (EpsilonWriter) ; l'outil de dessin intègre des bibliothèques d'images utiles pour les schémas scientifiques comme les montages de chimie ou d'électricité. Dans le cycle

expérimental, nous avons identifié trois types de tâches comme étant cruciaux de par leur position intermédiaire entre le monde des théories & modèles et le monde des objets & événements. Pour deux de ces types de tâches, nous avons développé deux outils pédagogiquement innovants que nous présentons plus en détail.

L'outil d'écriture de protocole expérimental correspond au type de tâche *Concevoir l'expérience*. La conception de cet outil s'appuie sur des recherches qui montrent l'intérêt de ce type de tâches pour l'apprentissage [13], sa complexité [14] et l'utilité d'un étayage par un outil numérique [15]. Dans LabNbook, cet outil est un éditeur de texte structuré où l'étudiant décrit l'expérience qu'il projette de réaliser. Pour relier la conception de l'expérience aux étapes précédentes, l'outil propose par défaut de rappeler (1) la question de recherche ou l'objectif de l'expérimentation et (2) les hypothèses ou les résultats attendus. Ces deux items correspondent aux types de tâches *Définir le problème* et *Proposer des hypothèses*. Ensuite l'étudiant décrit (3) le principe de la manipulation, (4) la liste du matériel et (5) le mode opératoire. Si l'enseignant estime qu'un étayage est nécessaire, il peut proposer une liste de matériels sélectionnables ainsi qu'un ensemble d'actions prédéfinies à partir desquelles l'étudiant compose son mode opératoire. Cet outil, hautement configurable, aide donc les étudiants à concevoir des protocoles expérimentaux grâce à la structuration qu'il impose et aux étayages définis par l'enseignant [16].

L'outil fourni pour gérer des données et créer des graphiques XY correspond aux deux types de tâches *Manipuler et collecter les données* et *Traiter les données*. L'outil ressemble aux tableurs habituels et, s'il comporte moins de fonctionnalités, il en intègre d'autres spécifiques aux sciences expérimentales et ayant une fonction pédagogique. L'étudiant peut notamment :

- qualifier certaines données comme étant « aberrantes » : ces données restent visibles mais ne sont pas prises en compte dans les calculs ;
- faire des calculs avec des noms de variables explicites : par exemple, un calcul de vitesse peut s'écrire « vitesse = distance/temps » ;
- afficher des données sur un graphique XY avec des échelles linéaires ou log ;
- afficher sur ces données des barres d'incertitude en X et/ou Y ;
- modéliser les données expérimentales par des fonctions mathématiques paramétrées : les paramètres sont reconnus comme tels par le système et l'étudiant modifie leurs valeurs pour ajuster visuellement le modèle mathématique aux points expérimentaux du graphique ; il est aussi possible de contrôler l'ajustement à partir de la valeur affichée de l'écart-type des résidus.

La fonctionnalité de modélisation est particulièrement appréciée des enseignants car elle permet aux étudiants d'appréhender la signification des paramètres du modèle lors d'un travail de modélisation expérimentale. Un exemple est donné dans la figure 3 : les données expérimentales validées sont représentées par des cercles et les données aberrantes par des croix ; deux fonctions linéaires modélisent les enveloppes de la courbe expérimentale ; une troisième fonction ajuste les données expérimentales par un modèle approprié ; sous le graphique sont affichés les choix de l'étudiant (fonctions mathématiques et valeurs des paramètres) ainsi que l'écart-type des résidus.

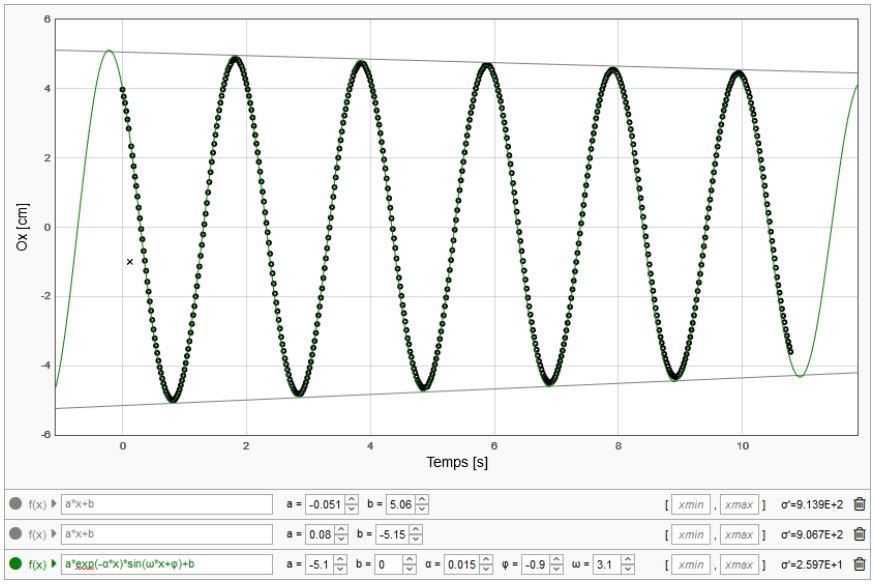


Fig.3. Utilisation de l'outil tableur-grapheur pour produire une modélisation mathématique des oscillations amorties pseudo-périodiques d'un pendule

Le dernier type de tâche du cycle expérimental concerne la diffusion du travail réalisé. L'étudiant peut sélectionner les documents (les « LabDocs ») de son espace de travail qui composent la version finale de son travail. Actuellement, seule la diffusion du travail des étudiants à destination de l'enseignant est possible. La diffusion du travail vers les pairs est une fonctionnalité qui sera prochainement introduite dans la plateforme.

Trois types de tâches de notre cycle expérimental n'ont pas d'outil spécifique dans LabNbook : *Définir le problème*, *Proposer des hypothèses* et *Interpréter les résultats*. Même si les deux premiers types de tâches sont partiellement couverts par l'outil d'écriture de protocole, les outils de texte et de dessin peuvent être utilisés efficacement pour ces trois types de tâches. Nous n'avons pas intégré d'outil support à la définition d'hypothèses comme on peut en trouver dans d'autres environnements [17]. En effet, un tel outil est adapté à l'investigation scientifique avec simulations, mais il ne nous apparaît pas suffisamment versatile pour notre plateforme qui se veut ouverte à diverses approches pédagogiques.

3 Des outils de collaboration, de communication et de suivi

Dans cette partie, nous décrivons les outils de la plateforme qui n'ont pas un caractère disciplinaire et qui ne sont donc pas issus d'une transposition didactique.

3.1 Outils supports de la collaboration entre étudiants

L'importance de l'interaction sociale dans le travail scientifique est un point qui ne semble pas remis en cause, tant dans les laboratoires de recherche [7] que dans les activités d'apprentissage [18] et notamment dans celles qui mettent en œuvre des activités de résolution de problèmes [19].
Dans LabNbook, l'espace de travail est partagé par l'équipe et tous les étudiants ont les mêmes droits sur les contenus. Ainsi les LabDocs écrits peuvent naturellement être le produit d'un travail collaboratif. L'aspect « synchrone » de la co-écriture n'est que partiellement pris en charge : quand un étudiant modifie un LabDoc, celui-ci reste visible pour les collaborateurs mais il est verrouillé en modification. Le LabDoc est libéré lors de sa validation. Les modifications apportées sont toutefois répercutées en synchrone sur l'interface des autres étudiants de l'équipe.
La plateforme offre un support aux échanges entre pairs par l'intermédiaire de trois outils spécifiques. Le premier outil est un porte-document partagé. Il permet aux étudiants d'échanger des ressources (URL ou documents téléchargés) qu'ils estiment utiles pour leur mission. Le second outil est une messagerie interne à la plateforme. Par rapport à une messagerie externe, cette messagerie comporte deux avantages : (i) elle indique le contexte de travail (mission) dans lequel les messages sont créés et (ii) les étudiants peuvent l'utiliser pour échanger leurs LabDocs entre équipes (option activable ou non par l'enseignant). Le troisième outil de communication est un fil de discussion disponible pour chaque LabDoc créé dans LabNbook. Ceci permet aux étudiants d'utiliser des espaces de discussion contextualisés par rapport à leur travail de production scientifique.

3.2 Outils supports des rétroactions des enseignants vers les étudiants

Un des objectifs de LabNbook est que les enseignants puissent suivre continuellement la progression des étudiants afin de pouvoir faire des retours sur leur travail. Pour cela nous avons mis en place trois outils.
Le premier outil mis à disposition des enseignants est un tableau de bord qui synthétise de la progression des étudiants dans leurs missions. Ce tableau, encore peu fourni, contient deux indicateurs généraux : la date de dernière modification faite par un étudiant dans l'espace de travail et le nombre de LabDocs produits par l'équipe. Nous allons prochainement intégrer de nouveaux indicateurs dans le tableau de bord grâce aux traces d'activité des étudiants recueillies dans la plateforme. Des algorithmes d'analyse vont permettre d'évaluer finement l'avancement des étudiants, leur degré de collaboration au sein de l'équipe et leur façon de prendre en compte les annotations fournies par les enseignants sur leur travail (voir ci-dessous).
Un deuxième outil permet aux enseignants de visualiser à tout moment l'espace de travail des étudiants. Certaines informations privées ne sont toutefois pas accessibles aux enseignants : les messages, les fils de discussion et les LabDocs auxquels les étudiants ont attribué un statut « caché ». L'espace de travail contient ainsi des espaces privés où les étudiants peuvent s'exprimer hors du regard de l'enseignant.
Le dernier outil permet à l'enseignant de faire des rétroactions aux étudiants. Dans l'espace de travail des étudiants, l'enseignant peut surligner des textes et ajouter des

annotations qui s'affichent sous forme de bulles. Nous avons ajouté une mémoire à cet outil afin que les enseignants puissent réutiliser leurs annotations précédentes écrites dans un contexte similaire.

4 Usages de la plateforme

La création de LabNbook a débuté en 2010. Pendant plusieurs années, la plateforme est restée un outil de recherche accessible à un nombre restreint d'utilisateurs de lycée [20] et de premier cycle universitaire [21]. En 2017, nous avons estimé que la plateforme était suffisamment stable pour être diffusée à plus large échelle et elle a été proposée au sein de la COMUE Grenoble Alpes. Vingt unités d'enseignement ont adopté la plateforme et les utilisations ont rapidement augmenté à la rentrée 2017. Ainsi, de cette date à la fin de l'année 2018 (16 mois), nous dénombrons 180 enseignants ayant encadré 2590 étudiants sur 160 missions différentes. Ces missions ont donné lieu à la création de 4870 espaces de travails partagés par autant d'équipes. Comme nous l'avions anticipé, les pédagogies implémentées par les enseignants dans la plateforme sont très diverses : LabNbook est utilisé pour que les étudiants écrivent leurs comptes-rendus dans le cadre de TP classiques, mais aussi comme support à des séquences ambitieuses d'apprentissage par projet [22].
Une étude en cours évalue l'impact de l'utilisation de LabNbook sur l'évolution des stratégies pédagogiques des enseignants. Les résultats de cette étude ne sont pas donnés ici. Nous nous limitons à présenter quelques chiffres illustrant l'utilisation de la plateforme par les étudiants. Cette utilisation ne traduit pas l'acceptabilité de la plateforme par les utilisateurs. Elle indique simplement les outils identifiés comme utiles par les étudiants (outils choisis) ou par les enseignants (outils prescrits).
Concernant l'utilisation des outils de production, nous nous attendions à ce que les outils les plus facilement identifiables et les plus versatiles soient les plus utilisés, soit : le traitement de texte ; l'outil de dessin ; le tableur-grapheur ; l'éditeur de protocoles. Nous avons compté, pour chaque outil, le nombre de LabDocs utilisés par les étudiants à leur initiative ou à l'initiative d'un enseignant (*cf.* table 1). Cette distinction vient du fait que les LabDocs peuvent être créés *ex-nihilo* par un étudiant ou être incorporés par l'enseignant dans l'espace de travail pour être ensuite complétés par les étudiants.

Table 1. Utilisation des outils de production : nombre de LabDocs modifiés par les étudiants de septembre 2017 à décembre 2018

	Nombre de LabDocs utilisés à l'initiative	
Outil de production de	d'un enseignant	d'un étudiant
textes	11 967 (51%)	16 894 (50%)
tableaux et graphiques	5 481 (23%)	7 102 (21%)
dessins	3 260 (14%)	6 209 (18%)
protocoles expérimentaux	2 909 (12%)	3 596 (11%)

Comme attendu, l'outil le plus utilisé (la moitié des usages) est le traitement de texte. C'est, en effet, le seul outil utilisable pour tous les types de tâches du cycle

expérimental. L'outil tableur-grapheur est le second en termes d'usage. Il est effectivement utilisé dans les missions de physique et chimie où l'activité de modélisation a souvent une place importante. Il n'est pas étonnant que l'outil d'écriture de protocole soit le moins utilisé : cet outil n'a pas d'équivalent et, comme cela est observé dans les innovations, son utilité n'est probablement pas reconnue d'emblée par les utilisateurs qui ont tendance à privilégier des outils familiers tel que le traitement de texte. Par ailleurs, la conception d'expérimentations est une activité complexe pour les étudiants et elle n'est pas mise en œuvre dans toutes les missions. Le fait que cet outil représente plus de 10% des utilisations, à l'initiative des enseignants ou des étudiants, est un résultat tout de même remarquable.
L'utilisation des outils de communication est donnée dans la table 2. Nous avons compté les ressources échangées grâce au porte-document, les messages postés dans la messagerie interne, les commentaires postés dans les fils de discussion contextualisés et les annotations faites par les enseignants. Concernant les messages, nous avons indiqué ceux postés par un étudiant ou par un enseignant. Si possible, nous avons indiqué le nombre d'équipes d'étudiants concernées par les communications (pour la messagerie, ce chiffre n'est pas disponible car les messages sont liés à un utilisateur qui peut appartenir à plusieurs équipes).

Table 2. Utilisation des outils de communication : nombre d'échanges effectués entre les utilisateurs de septembre 2017 à décembre 2018

Communication par échange de	Nombre d'échanges	Nombre d'équipes concernées
ressources externes (postées par un étudiant)	550	128
messages postés par un enseignant	68	-
messages postés par un étudiant	1 954	-
commentaires (postés par étudiant)	1 712	432
annotations (postées par un enseignant)	17 074	1 830

En termes comptables, on voit que peu de ressources externes sont échangées par les étudiants. Il est possible qu'ils utilisent d'autres moyens d'échange ou qu'ils consultent peu de documents en dehors de ceux fournis par les enseignants. La messagerie n'est quasiment pas utilisée par les enseignants. Pour les étudiants, la messagerie est utilisée autant que les commentaires des fils de discussion attachés aux documents produits. Ce résultat est étonnant : nous avions imaginé que les commentaires, plus contextualisés, auraient été plus utilisés que les messages et ce d'autant plus que l'enseignant peut décider de ne pas donner accès à la messagerie dans une mission. L'outil de communication le plus utilisé est de loin l'outil d'annotation qui permet aux enseignants de faire des retours sur les productions des étudiants : plus de 17000 annotations ont été produites par les enseignants en 16 mois.
Une caractéristique commune aux différents outils de communication est que leur utilisation concerne un pourcentage relativement faible d'équipes d'étudiant. Par exemple, le partage de ressources externes ne concerne que 128 équipes sur les 3837 équipes de plus de 1 étudiant : 3,3% des équipes partagent donc en moyenne 4,3 documents et 96,7% des équipes n'utilisent pas cet outil. De la même façon, 11,3% des équipes échangent en moyenne 4,0 commentaires et 88,7% des équipes n'échangent pas de commentaires. Pour les annotations, la discrimination est moins

importante puisque 37,6% (1830/4870) des équipes reçoivent en moyenne 9,3 annotations et 62,4% des équipes n'en reçoivent pas. On voit ainsi que l'utilisation des outils de communication est très différente selon les équipes. Il est probable que, suivant les situations pédagogiques mises en place, l'utilité de ces outils soit très différente. Par exemple, il est évident que les outils de communication n'ont pas d'utilité lorsque le travail sur la plateforme est intégralement réalisé en présentiel ! Une étude plus approfondie sur l'utilisation de la plateforme comme outil collaboratif sera réalisée à partir de l'analyse des traces d'activité des utilisateurs.

5 Conclusion

LabNbook est une plateforme numérique de support aux pédagogies actives. A notre connaissance, elle n'a pas d'équivalent actuel. Elle permet à des apprenants de créer et partager un cahier de laboratoire et/ou un rapport scientifique dans un espace de travail structuré et supervisé par leurs enseignants. Nous avons présenté les divers outils intégrés à la plateforme et les justifications épistémologiques et didactiques qui ont conduit à leur conception. Une première diffusion de LabNbook a été menée à grande échelle dans un périmètre restreint et nous étudions actuellement l'impact de l'utilisation de la plateforme sur les stratégies pédagogiques des enseignants.
LabNbook est développée sous licence libre. Nous souhaitons promouvoir son utilisation hors de notre COMUE par différentes actions : partager le code, connecter la plateforme au LMS Moodle et introduire un outil de traduction des interfaces. Nous envisageons aussi LabNbook comme un outil de recherche. Son acceptation par les enseignants et l'utilisation active par les étudiants a permis d'enregistrer, sur la dernière année, deux millions de traces d'utilisation en sus des productions et des communications des utilisateurs. Ces données, qui peuvent être partagées, sont, à coup sûr, un matériau intéressant pour la recherche dans le domaine des EIAH.

Remerciements
Nous remercions Nicolas Balacheff pour sa relecture critique du texte.
Le développement et le déploiement de LabNbook sont possibles grâce aux soutiens accordés par les projets IDEX-Formation de la COMUE Grenoble Alpes et Appel à Manifestation d'Intérêt « Transformation pédagogique et numérique » du Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

Références

1. Dulbecco, P., Beer, M.-C., Delpech de Saint-Guilhem, J., Dubourg-Lavroff, S., Pimmel, E.: Les innovations pédagogiques numériques et la transformation des établissements d'enseignement supérieur. Les rapports de l'I.G.A.E.N.R. 2018–049, (2018).
2. Pomerantz, J., Brown, M., Brooks, D.C.: Foundations for a Next Generation Digital Learning Environment: Faculty, Students, and the LMS. EDUCAUSE Center for Analysis and Research, Louisville, CO (2018).

3. Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., Wenderoth, M.P.: Active learning increases student performance in science, engineering, and mathematics. PNAS. 111, 8410–8415 (2014).
4. Slotta, J.D., Linn, M.C.: WISE Science: Web-Based Inquiry in the Classroom. Teachers College Press, New York, NY (2009).
5. de Jong, T., van Joolingen, W.R. *et. al.*: Learning by creating and exchanging objects: The SCY experience. British Journal of Educational Technology. 41, 909–921 (2010).
6. de Jong, T., Sotiriou, S., Gillet, D.: Innovations in STEM education: the Go-Lab federation of online labs. Smart Learning Environments. 1, (2014).
7. Latour, B., Woolgar, S.: La vie de laboratoire - La production des fait scientifiques. La Découverte, Paris (1979).
8. Chevallard, Y.: La transposition didactique - Du savoir savant au savoir enseigné. La Pensée Sauvage, Grenoble (1991).
9. Pedaste, M., Mäeots, M., Siiman, L.A., de Jong, T., van Riesen, S.A.N., Kamp, E.T., Manoli, C.C., Zacharia, Z.C., Tsourlidaki, E.: Phases of inquiry-based learning: Definitions and the inquiry cycle. Educational Research Review. 14, 47–61 (2015).
10. Develay, M.: Sur la méthode expérimentale. Aster. 8, 3–15 (1989).
11. Tiberghien, A., Vince, J., Gaidioz, P.: Design-based Research: Case of a teaching sequence on mechanics. International Journal of Science Education. 31, 2275–2314 (2009).
12. Girault, I., d'Ham, C., Ney, M., Sanchez, E., Wajeman, C.: Characterizing the Experimental Procedure in Science Laboratories: A preliminary step towards students experimental design. International Journal of Science Education. 34, 825–854 (2011).
13. Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R., Hmelo-Silver, C.E.: Design and Reflection Help Students Develop Scientific Abilities: Learning in Introductory Physics Laboratories. Journal of the Learning Sciences. 19, 54–98 (2010).
14. Berthet, A., Girault, I., d'Ham, C.: Difficultés d'élèves pour élaborer un protocole expérimental - Un exemple en classe de terminale S. Le Bup. 109, 1395–1408 (2015).
15. Girault, I., d'Ham, C.: Scaffolding a Complex Task of Experimental Design in Chemistry with a Computer Environment. Journal of Science Education and Technology. 23, 514–526 (2014).
16. Bonnat, C., Marzin-Janvier, P., Girault, I., d'Ham, C.: Modélisation didactique pour la conception d'étayages dans un EIAH : exemple d'une activité de conception expérimentale en biologie. STICEF. 25 (2019).
17. Hovardas, T., Xenofontos, N.A., Zacharia, Z.C.: Using Virtual Labs in an Inquiry Context: The Effect of a Hypothesis Formulation Tool and an Experiment Design Tool on Students' Learning. In: Optimizing STEM Education With Advanced ICTs and Simulations. pp. 58–83. IGI Global (2017).
18. Vygotsky, L.: Pensée et langage. Messidor éditions sociales, Paris (1934).
19. Wismath, S.L., Orr, D.: Collaborative Learning in Problem Solving: A Case Study in Metacognitive Learning. The Canadian Journal for the Scholarship of Teaching and Learning. 6 (2015).
20. Bonnat, C., Marzin-Janvier, P., Girault, I.: Analyse des conceptions d'élèves sur le vivant dans une situation de conception expérimentale avec un environnement informatique. Recherche en Didactique des Sciences et des Technologies. A paraitre (2019).
21. Wajeman, C., Girault, I., d'Ham, C., Marzin, P.: Students' reflection on experimental design during an innovative teaching sequence with LabBook. In: Lavonen, J., Juuti, K., Lampiselkä, J., Uitto, A., and Hahl, K. (eds.) Electronic Proceedings of the ESERA 2015 Conference. pp. 744–755. University of Helsinki, Helsinki, Finland (2016).
22. Hoffmann, C., Girault, I., Kahane, C., d'Ham, C., Planche, M.: Utilisation d'une plateforme numérique dans un dispositif d'apprentissage par problèmes (APP). In: Question de Pédagogies dans l'Enseignement Supérieur (QPES). Brest, France (2019).

Étude empirique sur la ludicisation de classes au Brésil et en Suisse avec Classcraft

Guillaume Bonvin[1], Eric Sanchez[1], Celso Gonçalves[2]

[1] Université de Fribourg, 20 Avenue de l'Europe, 1700 Fribourg, Suisse
guillaume.bonvin@unifr.ch, eric.sanchez@unifr.ch
[2] Universidade do Estado de Minas Gerais, Rue Ver. Geraldo Moisés da Silva
38.302-192 Ituiutaba, Minas Gerais, Brésil
celsosbgoncalves@yahoo.com

Résumé. Ce papier porte sur une étude empirique conduite dans des écoles secondaires, en Suisse et au Brésil et qui concerne la gestion de classe avec Classcraft, un jeu numérique de gestion de classe. Les travaux qui ont été menés visent à comprendre la manière dont le comportement des élèves est influencé par le jeu. Ils s'appuient sur un modèle qui situe l'expérience de jeu dans un contexte qui comprend la situation de jeu, le travail de l'enseignant qui intègre le jeu et la culture organisationnelle de l'institution dans laquelle ce dernier exerce. La méthodologie mixte qui a été adoptée consiste dans la collecte de traces numériques d'interaction complétées avec des données collectées dans le cadre d'entretiens avec les enseignants et leurs élèves. Les résultats obtenus montrent que, pour comprendre les effets du jeu sur les élèves, il est nécessaire de prendre en compte la manière dont l'enseignant s'est approprié le jeu et l'influence de l'institution sur cette appropriation.

Mots-clés. Gamification, ludicisation, gestion de classe, Classcraft, Comportement des élèves

Abstract. This paper concerns an empirical study about classroom management using Classcraft, a digital game dedicated to classroom, conducted in secondary schools in Switzerland and Brazil. The study aims to understand how students' behaviour is influenced by the game and rely on a model that situates the game experience in a context that includes the game; the teacher's task, which integrates the game; and the organisational culture of the institution in which the teacher exercises his or her profession. The mixed methodology used in the studies is founded on the collection of digital interaction activity trace data, complemented by data collected during interviews with teachers and their students. The research results indicate that, in order to understand the effects of the game on students, it is necessary to consider the way the teacher appropriates the game and the influence of the institution over this appropriation.

Keywords. Gamification, ludicization, Classroom management, Classcraft, Students' behavior

1 Introduction

La gamification, parfois traduite par ludification en français, est généralement définie comme l'emploi d'éléments ou de mécaniques ludiques dans le cadre d'une situation réputée non ludique comme une situation d'apprentissage [1]. Des définitions plus récentes tentent de prendre en compte le caractère subjectif du jeu [2]. Elles conduisent à considérer que la question de la conversion d'une situation d'apprentissage en jeu relève d'un processus de ludicisation [3].

Nous avons ici pour objectif de discuter les effets de la ludicisation sur le comportement d'élèves de l'enseignement secondaire. L'analyse que nous souhaitons mener prend en compte la dimension subjective du jeu, les décisions prises par l'enseignant et la culture organisationnelle de l'institution au sein de laquelle il est mis en place. Cette analyse s'appuie sur des travaux empiriques qui ont permis d'expérimenter Classcraft, une plateforme de jeu dédiée à la ludicisation de la gestion de classe [4]. Menées en Suisse et au Brésil, ces expérimentations ont permis de recueillir des données d'observation dans des classes d'enseignants du secondaire qui collaborent avec les chercheurs.

Dans une première partie, nous revenons sur les travaux qui se sont attachés à définir le concept de gamification et le modèle que nous avons élaboré. Nous décrivons, dans une seconde partie, l'étude empirique qui a été menée ainsi que les observations effectuées et des données collectées. L'influence du jeu sur les élèves sera analysée dans la troisième partie de ce papier.

2 Cadre conceptuel et problématique

La gamification est originellement définie comme l'usage d'éléments ou de mécaniques ludiques dans le cadre d'une situation réputée ne relevant pas du jeu. Les éléments de jeu sont ici des composants avec lesquels le joueur interagit tels que des récompenses sous forme de points ou de badges, de tableaux de scores. D'autres éléments sont relatifs au game design, des éléments choisis par le créateur du jeu comme un temps limité ou des niveaux de jeu. Les mécaniques ludiques sont les méthodes mobilisées par le joueur pour jouer. Pour les jeux de plateau, les mécaniques usuelles sont par exemple le lancer d'un dé ou le déplacement de pions [5]. Dans le champ de l'éducation, la gamification est employée pour obtenir l'engagement des apprenants dans les activités d'apprentissage ou des comportements souhaités. Ce dernier objectif relève de la gestion de classe. Définie comme l'ensemble des « pratiques éducatives auxquelles l'enseignant a recours afin d'établir, de maintenir et, au besoin de restaurer dans la classe des conditions propices au développement des compétences des élèves » [6], la gestion de classe est souvent un grand défi que l'enseignant rencontre dans son activité professionnelle. Différentes applications dédiées à la gamification de la gestion de classe sont aujourd'hui disponibles. Par exemple, The Carrot Rewards for schools[1] propose des systèmes de récompense qui permettent à l'enseignant d'accorder des points à leurs élèves selon

[1] www.carrotrewards.co.uk

qu'ils respectent, ou pas, les règles de vie de la classe. ClassBadges[2] est une autre plateforme numérique qui permet d'attribuer des badges aux élèves qui réussissent certains objectifs prédéfinis. Enfin, Class Dojo[3] est destiné à envoyer des "feed-back positifs" aux élèves qui démontrent des capacités attendues telles que "travailler de manière assidue" ou "être respectueux avec les autres". L'efficacité de cette dernière application a pu être démontrée dans le cas d'une étude empirique, en particulier du point de vue de l'amélioration de l'engagement des élèves [7]. D'autres études ont également rapporté des améliorations en ce qui concerne une autorégulation des comportements et une diminution de la fréquence des comportements inappropriés [8]. Les explications avancées pour expliquer ces effets mettent généralement en avant le système de récompense du jeu. Par exemple, Rocha Seixas et col. [9] ont montré que le nombre de récompenses attribuées est corrélé avec des performances comportementales plus élevées que la moyenne ce qui les conduits à conclure sur les effets positifs de ces récompenses.

Ainsi, la conception de ces applications s'appuie sur des éléments de jeu qui relèvent de choix de game design prenant la forme de points ou de badges et étant censés influencer la manière dont un élève se comporte. Pourtant, la notion d'élément de jeu est problématique. Aucun élément ne peut être considéré comme étant intrinsèquement ludique et, s'appuyer sur une définition de la gamification qui postule le contraire, c'est adopter un point de vue qui nie le caractère subjectif du jeu [10]. Des éléments ou des mécaniques ludiques, tels que les récompenses, ne sont pas par essence ludiques. Ils peuvent être utilisés dans d'autres types de situations. Ainsi il est difficile de considérer qu'il est possible de « faire un jeu » comme le suggère l'étymologie du terme gamification.

Des définitions plus récentes de la gamification tentent de prendre en compte le fait qu'un jeu, au sens de situation de jeu (play) et non pas de l'artefact (game), résulte avant tout d'une interprétation de celui qui joue. Gamifier consisterait alors à introduire dans une situation des affordances qui s'appuient sur des principes du game design et qui permettent, au sujet impliqué dans cette situation de la considérer comme un jeu. Le concept d'affordance perçue renvoie aux propriétés qu'un sujet peut attribuer à un objet [11]. Une telle définition de la gamification permet alors de prendre en compte le caractère subjectif du jeu, ce que Jacques Henriot [12] résume en écrivant que « le jeu n'est pas dans la chose, mais dans l'usage qu'on en fait ». Ce point de vue permet également d'insister sur le caractère performatif du jeu. Le jeu-situation (play) émerge d'une rencontre entre un jeu-objet (game) et un sujet qui devient joueur parce qu'il accepte d'entrer dans un « cercle magique » [13] où les actions qu'il réalise prennent un sens différent. Gamifier ce n'est donc pas « faire un jeu » (littéralement gami-fier), mais faire en sorte qu'une situation ordinaire puisse être interprétée comme un jeu. C'est pourquoi, en nous appuyant sur Genvo [10], nous avons proposé que le terme ludicisation soit utilisé à la place de gamification [4]. Le suffixe "icisation" permet en effet de souligner l'idée d'une transformation et le préfixe "ludus" indique la nature de cette transformation. Ludiciser une situation ordinaire c'est donc y introduire des affordances perçues qui s'appuient sur des principes du game design dans le but de permettre à un sujet d'interpréter cette

[2] www.classbadges.com
[3] www.classdojo.com

situation comme un jeu. Dans la suite, nous illustrons ce concept à partir de l'analyse de Classcraft.

L'adoption de cette approche nous conduit à considérer que la ludicisation de la gestion de classe dépend de la manière dont l'enseignant conçoit et anime la situation ludique et, d'autre part, de la culture organisationnelle du système éducatif. C'est ce qui est décrit dans le modèle ci-dessous (fig. 1), adapté de Margolinas [14] et de ses travaux sur la structuration du milieu.

Fig. 1. Modèle de ludicisation inspiré de Margolinas [14]

Ce modèle permet de prendre en compte que la situation de classe ludicisée dépend de décisions prises par l'enseignant lorsqu'il conçoit la situation. La conception est elle-même influencée par la culture organisationnelle de l'institution dans laquelle il exerce et qui influence ses choix. À travers ce modèle nous faisons également l'hypothèse que l'influence de chacun des niveaux peut aussi s'exercer dans l'autre sens. La situation peut amener l'enseignant à réviser ses choix de conception et lui-même peut introduire des changements dans le fonctionnement de son institution.

Ce modèle constitue le cadre conceptuel que nous avons élaboré pour analyser la ludicisation de la gestion de classe avec Classcraft. Cela nous conduit à nous interroger sur la situation de classe, sur les effets de l'usage du jeu sur le comportement des élèves. Est-ce que les comportements escomptés par les enseignants sont bien présents ? Un autre volet de notre problématique porte sur l'interprétation de ces effets. En particulier, en lien avec notre modèle, nous nous intéressons au rôle de l'enseignant ainsi que celui de la culture organisationnelle du système éducatif dans lequel cet enseignant exerce. Nous formulons ainsi l'hypothèse que les effets du jeu dépendent de la manière dont l'enseignant le met en place et que les choix de l'enseignant dépendent d'attentes et de contraintes institutionnelles.

3 Une étude empirique

3.1 Classcraft, un jeu de gestion de classe

Classcraft [4] est un jeu numérique dédié à la gestion de classe. C'est une plateforme numérique accessible via un ordinateur, une tablette ou un ordiphone. Classcraft

présente quelques similarités avec les applications précédemment décrites. Les joueurs sont récompensés ou pénalisés en fonction de leur comportement en classe. Les comportements attendus sont décrits par les règles du jeu. Les joueurs endossent des rôles tels que mage, guerrier ou guérisseur qui leur confèrent des pouvoirs spécifiques. Des règles sont implémentées par défaut dans la plateforme. Elles peuvent être modifiées par l'enseignant pour les adapter à ses objectifs ou aux règles qui prévalent dans l'établissement scolaire où il exerce. Par exemple, un élève pourra perdre 10 points de vie (HP) s'il oublie d'apporter son matériel scolaire en classe ou pourra se voir attribuer 10 points d'expérience (XP) s'il aide un camarade en difficulté. Les points reçus peuvent être utilisés pour acheter des pouvoirs qui ont des effets sur ce que l'élève est autorisé à faire en classe. Par exemple, le pouvoir « invisibilité » permet d'arriver 5 minutes en retard sans encourir une punition. Les élèves jouent en équipe et les comportements d'entraide et de collaboration sont encouragés. Certains pouvoirs sont en effet des pouvoirs collaboratifs qui permettent d'aider un coéquipier en difficulté. Le pouvoir d'invisibilité peut ainsi être utilisé pour éviter une punition à un camarade qui arrive en retard. Inversement, certaines sanctions attribuées par l'enseignant à un élève impactent également les membres de son équipe. D'autres fonctionnalités de Classcraft permettent de gérer et d'animer des activités d'apprentissage. Par exemple, le sonomètre peut être utilisé pour limiter le niveau de bruit lors des travaux de groupes. Si ce niveau dépasse un seuil déterminé par l'enseignant, tous les joueurs sont pénalisés. La "bataille de boss" consiste dans un questionnaire qui est conçu par l'enseignant et pour lequel chaque réponse exacte débattue au sein d'une équipe est une victoire sur un adversaire virtuel dans le jeu. Cette victoire permet de gagner des pièces d'or (Coins) qui peuvent être utilisées pour personnaliser son avatar.

Les différentes équipes ne sont généralement pas en compétition, mais chaque élève peut consulter un tableau de bord sur lequel sont indiquées des informations telles que le niveau atteint, les points disponibles ou les pouvoirs dont il dispose. Le tableau de bord affiche également son avatar qui dépend du rôle qu'il a choisi. L'enseignant est donc un maître du jeu qui choisit les règles, s'assure qu'elles sont respectées, définit des défis et récompense ou sanctionne.

Classcraft ne consiste donc pas dans la juxtaposition d'éléments ou de mécaniques ludiques. Ces éléments sont au contraire intégrés pour constituer une métaphore cohérente. La classe devient une bataille et le but du jeu consiste à survivre aux aléas du combat et à acquérir de nouveaux pouvoirs. Cette métaphore est soutenue par l'iconographie qui s'inspire de celle du jeu World of Warcraft. Elle est également soutenue par le lexique employé. Ainsi, un élève qui a perdu tous ses points de vie est « tombé au combat ». La métaphore résulte aussi des rôles qui sont attribués aux différents personnages. Le guérisseur peut sauver un camarade en difficulté. La métaphore déployée dans l'univers Classcraft vise à ce qu'un élève donne une autre signification aux comportements attendus dans la classe. Ainsi, obtenir des points, c'est permettre à son avatar de gagner des pouvoirs qui sont le gage de sa « survie » et celle de son équipe.

Comme tout champ de bataille, Classcraft est également sujet à des aléas qui peuvent se révéler bénéfiques ou négatifs. Pour animer son cours, l'enseignant peut lancer l'outil "Les Cavaliers de l'île de Vay". De manière aléatoire, l'ensemble de joueurs est pénalisé ou au contraire gratifié de points supplémentaires, de manière

complètement arbitraire. L'ensemble de ces éléments font de Classcraft une plateforme dédiée à la ludicisation de la gestion de classe. Pour les élèves, les actions effectuées dans la classe ne sont pas modifiées en tant que telles. Il s'agit d'adopter le comportement attendu d'un élève de l'enseignement secondaire. Les raisons qui motivent ce comportement ne sont pas non plus différentes de la situation usuelle qui prévaut dans une école. Ce comportement est encadré par des règles qui sont ici des règles du jeu semblables aux règles de vie de la classe. Ce qui est modifié par l'application c'est le sens donné au comportement dans la classe. Il ne s'agit plus d'adopter un comportement attendu par l'enseignant, mais plutôt un comportement qui permettra de franchir les niveaux du jeu, d'acquérir des pouvoirs et de survivre au sein d'une bataille pour laquelle l'entraide entre équipiers est cruciale. Autrement dit, et pour reprendre la terminologie de Goffman [15], ludiciser c'est rendre possible un changement de cadre d'interprétation de la situation. La classe, en tant qu'expérience vécue par les élèves, constitue un cadre primaire régi par des règles sociales qui relèvent de la gestion de classe usuelle. Le jeu Classcraft constitue un cadre transformé. Il permet que s'établissent, entre les joueurs et entre les joueurs et l'enseignant, en tant que maître du jeu, des interactions qui relèvent d'un second degré. Ainsi une pénalité attribuée à un élève n'est plus une sanction liée au cadre scolaire, mais une pénalité de jeu, se traduit par une perte de points. Cette transformation est rendue possible par les affordances perçues du jeu qui permettent à un élève de devenir joueur, c'est-à-dire d'attribuer une signification nouvelle aux interactions. Selon ce point de vue, la subjectivité du jeu est centrale.

3.2 Contexte de l'expérimentation

À travers cette étude descriptive, nous voulons comprendre comment le comportement des élèves est influencé par le jeu. Nous faisons l'hypothèse que le jeu influence le comportement des élèves et que son orchestration est influencée par l'enseignant, la culture organisationnelle et les élèves eux-mêmes. À cet effet, nous conduisons des expérimentations du jeu dans deux écoles pilotes situées respectivement dans le Canton de Fribourg (codée ECF), en Suisse et à Uberlandia au Brésil (codée EUB). Le tableau ci-dessous (table 1) résume les caractéristiques des expérimentations menées. Dans la suite, les élèves seront codés à l'aide d'une lettre à laquelle nous adjoignons le code de son école.

Table 1. Caractéristiques des classes dans lesquelles les expérimentations sont menées

	Fribourg (ECF)	**Uberlândia (EUB)**
Classe		
Nombre d'élèves	23 élèves	35 élèves
Âge	14-15 ans	10-12 ans

Rôles	12 filles	15 filles
	5 Mages	2 Mages
	2 Guerrières	5 Guerrières
	5 Guérisseuses	8 Guérisseuses
	11 garçons	20 garçons
	5 Mages	12 Mages
	4 Guerriers	4 Guerriers
	2 Guérisseurs	4 Guérisseurs

3.3 Données collectées

Les données collectées proviennent de deux sources. Elles consistent dans la collecte et l'analyse des traces numériques d'interaction qui nous permettent de mettre en place une méthodologie que nous qualifions d'analytique du jeu ou playing analytics [16] en référence à l'analytique de l'apprentissage. Ces traces se présentent sous la forme d'obsels, c'est-à-dire sous la forme d'événements de jeu horodatés qui intègrent des informations permettant de leur donner du sens. Par exemple, le joueur dont l'identifiant est "bGjjjjF6Jk7nQurh3" a, le 18 janvier 2017 à 8h10, perdu 10 points parce qu'il a oublié d'apporter son matériel en classe. Le 12 avril 2018, 1911 obsels avaient été collectés. Ces obsels sont visualisés avec la plateforme kTBS4LA [17] qui permet une sélection des données et certains traitements statistiques élémentaires.

Des observations de classe et des entretiens semi-directifs (de type focus groups) sont conduits avec les élèves et les enseignants. Les observations de classe font l'objet de l'utilisation d'une grille qui permet de consigner ce que disent et ce que font les enseignants et leurs élèves. Sont également notés tous les événements relatifs au jeu. Les observations réalisées sont horodatées. L'observateur saisit également des mots clefs ainsi que des commentaires sur la séance à laquelle il a participé. Les entretiens sont conduits en prenant comme référence les différents différents niveaux de notre modèle d'analyse : jeu et comportement des élèves, décisions et interventions de l'enseignant et contexte institutionnel dans lequel le jeu est mis en place. Pour alimenter ce dernier point, des entretiens ont également été menés avec les responsables des établissements concernés.

4 Résultats et discussion

4.1 Le jeu et son influence sur le comportement des élèves

Les données que nous avons collectées permettent d'attester de l'influence du jeu sur le comportement des élèves qui va dans le sens des attentes des enseignants. Dans un entretien, l'enseignant de l'UEB indique qu'il a réussi à obtenir que les élèves utilisent le dictionnaire pour certaines tâches. Les élèvent relèvent qu'ils ont pris conscience de l'intérêt de travailler de manière plus collaborative et de demander l'avis des autres membres de leur équipe avant de fournir une réponse. Ils

mentionnent que le jeu a facilité l'intégration d'élèves arrivés en cours d'année scolaire. De son côté, l'enseignant relève que certains travaux demandés pour la préparation des cours qui n'étaient généralement pas effectués par les élèves le sont aujourd'hui de manière beaucoup plus systématique.

Des résultats similaires sont rapportés par les enseignants de l'ECF qui indiquent que les élèves se sont approprié des règles de vie de classe qu'ils avaient du mal à intégrer jusqu'alors. Ceci est confirmé par certains élèves lors des entretiens : "*Je pense que grâce au jeu, on a pris conscience qu'on était responsable de notre comportement en classe*" (M-ECF). J-ECF mentionne que les élèves de son équipe se sont entraidés pour échapper aux conséquences négatives, pour toute l'équipe, des échecs de l'un d'entre eux. Lors d'un entretien, E-ECF explique que le jeu permet de "*faire passer des règles de manière positive. Normalement, elles passent de manières négatives. Si on ne les respecte pas, ils nous punissent. Tandis que là, si on respecte les règles on est récompensé*".

Néanmoins, il est possible d'identifier des élèves plus réfractaires au jeu. Par exemple, A-ECF dont le comportement est jugé "perturbateur" par ses enseignants et qui elle-même reconnaît la justesse de ce constat, n'a pas pour autant modifié son comportement. Elle ne rejette néanmoins pas totalement le jeu, car elle en souligne l'intérêt pour "*présenter les règles de manière positive*" et y participe en recherchant des erreurs dans les énoncés des enseignants pour gagner des points.

Le jeu se traduit donc par des effets différenciés selon les élèves. Les traces recueillies montrent que la manière de jouer dépend du genre. Le nombre d'obsels recueillis chez les filles est inférieur à celui des garçons et, en ce qui concerne les pouvoirs utilisés, les filles, contrairement aux garçons, utilisent de manière plus fréquente des pouvoirs qui permettent d'aider un membre de leur équipe (pouvoirs collaboratifs) plutôt qu'elles-mêmes (pouvoirs individuels). Néanmoins, un modèle statistique de l'utilisation des pouvoirs nous a permis d'identifier que la variable explicative de la probabilité d'utiliser tel ou tel pouvoir n'est pas le genre, mais le rôle joué dans le jeu. Ainsi, les pouvoirs collaboratifs sont plus utilisés par les filles parce qu'elles choisissent majoritairement le rôle de "guérisseuse" et que ce rôle se manifeste par la capacité à utiliser de tels pouvoirs.

Les données collectées confirment que le jeu permet, dans la majorité des cas et des élèves, d'obtenir des effets positifs en ce qui concerne la gestion de classe. Néanmoins, les résultats obtenus varient suivant les élèves et les contextes. Par ailleurs, l'expérience de jeu (*play*) ne semble pas homogène. Cela traduit que la manière dont le jeu est joué et son influence dépend d'autres facteurs que les propriétés intrinsèques du jeu (*game*) telles que les éléments et les mécaniques ludiques. Ces facteurs sont à rechercher dans la manière dont le jeu est introduit dans la classe, ce que nous discutons dans le paragraphe suivant.

4.2 Implémentation et orchestration du jeu

À l'ECF, les interactions entre le chercheur et l'enseignant qui le premier a intégré Classcraft dans son établissement ont été grandement facilitées du fait que ce dernier

était très enthousiaste à utiliser Classcraft pour la gestion de sa classe. Quatre autres enseignants avaient également manifesté leur intérêt, mais considéraient la mise en place du jeu complexe et astreignante. Ils ont finalement rejoint le projet après quelques semaines, convaincus par le discours de l'enseignant qui a initié le projet. Ce dernier a en effet mentionné auprès de ses collègues les bénéfices qu'il retirait de cette expérience, tels qu'un plus grand respect des règles de vie de la classe et une augmentation de l'implication des élèves. Dans un premier temps, cet enseignant avait souhaité découvrir le jeu par lui-même et ce n'est que plus tard qu'il a fait appel au chercheur, mais d'abord pour bénéficier de son expertise sur le jeu. Au final, c'est une collaboration entre 5 enseignants et le chercheur en charge du projet qui s'est mise en place. La figure 2 indique, au cours du temps, le nombre d'obsels collectés pour les situations animées par les différents enseignants (différentes couleurs). Le niveau d'activité, évalué en termes d'événements de jeu, se poursuit sur l'ensemble de l'année scolaire pour l'enseignant qui le premier a été impliqué dans le projet (couleur bleue). L'activité mesurée pour les autres enseignants montre leur implication plus tardive, progressive et moins importante. Les semaines pour lesquelles le nombre d'événements est moins important correspondent aux vacances scolaires.

Fig. 2. Nombre d'obsels collectés selon les enseignants de l'ECF

Le jeu a été présenté de manière à ce que la situation soit bien interprétée comme un jeu par les élèves. Il a aussi été adapté en prenant en compte leur intérêt pour des "pouvoirs" attrayants. Par exemple, les élèves qui franchissent le dernier niveau du jeu se voient gratifiés d'un "pouvoir" qui leur permet d'échanger avec un de leur camarade de classe pendant cinq minutes à la fin d'une évaluation sommative. C'est un "pouvoir" très convoité qui renforce l'intérêt à gravir les différents niveaux du jeu.

Finalement, le jeu a pris une place considérable pour la classe, car ce sont au total 5 enseignants qui l'ont utilisé. Néanmoins, *Classcraft* n'est pas présent dans toutes les séances et les enseignants alternent des séances avec et sans le jeu.

À l'EUB, un seul enseignant est concerné et, malgré le manque d'équipement audiovisuel dont il dispose, il s'est approprié le jeu et l'a intégré dans ses pratiques. Les élèves utilisent leurs téléphones. Les données collectées montrent que le jeu a été mis en place de manière très différente. À l'ECF les obsels qui concernent l'ajout ou le retrait de points XP ou HP concernent respectivement 34.4% et 15.9% du total contre 21.4% et 2.7% à l'EUB. A contrario, les obsels qui concernent les *Coins* représentent 20.8% à l'ECF contre 52.4% à l'EUB. Ces données montrent que l'enseignant de l'EUB utilise principalement le système des *Coins* ce qui signifie qu'il utilise Classcraft pour animer les activités de Classe (sous la forme de défis qui permettent de remporter des *Coins*). Les aspects du jeu qui concernent le système de

sanctions et récompense pour amener les élèves à adopter certains comportements en classe (XP et HP) sont moins utilisés que chez ses collègues suisses.

Ces données sont confirmées par les observations de classes et les entretiens. Contrairement à l'expérience brésilienne, les enseignants suisses se sont principalement cantonnés à l'utilisation des outils de gestion de classe de type sanction-récompense. L'enseignant principal voulait *"se concentrer sur la maîtrise technique du jeu"*. Cela a été réalisé au détriment de l'animation et cela a été relevé par certains élèves: *"Pour que ce soit plus un jeu, il faut un côté plus participatif. Je ne sais pas comment dire. Je trouve qu'actuellement ce n'est pas un jeu, car ça manque de vie, ça ne change rien" (G-ECF)*. À l'ECF, l'expérience ludique a donc été limitée.

De plus, des élèves ont profité d'un certain manque de coordination entre les enseignants et de certaines lacunes dans l'appropriation du jeu (application insuffisante des sentences) pour bénéficier d'avantages indus (accumulation de XP sans perdre des HP). Cela a été révélé durant le focus group et totalement assumé par la grande majorité des élèves: *"Nous avons profité du système."*; *"Pour moi c'est au prof de se renseigner.[...] On ne va pas faire son métier"* (V-ECF).

L'appropriation du jeu par les enseignants est donc très différente dans les deux contextes étudiés. Elle se manifeste par des usages différents qui se traduisent par des expériences de jeu différentes chez les élèves.

4.3 Culture organisationnelle des institutions et mise en place du jeu

À Fribourg, toutes les remarques faites à un élève pour son comportement durant le temps de classe sont consignées dans le carnet de correspondance (agenda scolaire) et doivent être signées par l'autorité parentale. Un contrôle précis est donc effectué par le corps enseignant et cette démarche s'inscrit dans le règlement de l'établissement. Durant le temps consacré à l'expérimentation du jeu, les enseignants n'ont pas voulu utiliser la plateforme comme moyen de communication avec les parents de leurs élèves. Classcraft a eu comme conséquences, que les sanctions (pertes de points dans le jeu) n'étaient plus communiquées systématiquement aux parents, excepté "les morts au combat" qui amenaient les élèves à de véritables sanctions. Cela a été perçu par les élèves comme un avantage majeur et largement partagé durant le focus group: *"Classcraft met une meilleure ambiance à la maison, les parents ne sont pas fâchés à cause des remarques"; "Classcraft atténue les remarques qu'on pourrait se choper. Tout part dans le jeu et du coup les parents ne sont au courant de rien [rire général]*" D-ECF. Il faut relever que les enseignants n'avaient pas identifié cela lors de la mise en place du jeu.

Du point de vue des institutions, il faut souligner que, à l'ECF, l'expérimentation est soutenue par le directeur de l'établissement scolaire. En effet, c'est après un échange avec le directeur de cet établissement que la décision a été prise d'expérimenter le jeu. Des enseignants ont été sollicités par le directeur et une séance d'information a été organisée. C'est au cours de cette séance que des enseignants se sont déclarés intéressés. Nous considérons cet appui institutionnel important, car, au-

delà des facilités administratives qu'il a permises, il légitime de travail effectué par les enseignants. Le directeur de l'établissement scolaire a également joué un rôle important pour la mise en place de la recherche elle-même en se positionnant favorablement auprès du département de l'instruction publique fribourgeois. Ce soutien a facilité les démarches nécessaires pour la conduite de l'expérimentation et le recueil des données.

Le contexte brésilien est tout différent. L'enseignant exerce dans une école militaire et ne bénéficie pas, *a priori*, d'un soutien de sa hiérarchie pour expérimenter des approches éducatives innovantes. Lui-même se décrit comme un enseignant qui souhaite se positionner en opposant aux méthodes qui prévalent dans son établissement et le jeu lui permet d'affirmer sa singularité. Sa hiérarchie a reçu des retours positifs des parents dont certains ont aidé leur enfant à prendre en main la plateforme à la maison et ont apprécié que son niveau d'engagement dans les activités scolaires ait augmenté. L'enseignant bénéficie ainsi d'un soutien indirect qui, dans une certaine mesure, légitime l'utilisation du jeu.

Ainsi, les observations réalisées permettent de dresser un tableau paradoxal de l'utilisation du jeu. C'est l'institution la plus facilitatrice qui obtient, *in fine*, les résultats les plus faibles en ce qui concerne l'appropriation du jeu par les enseignants. Ce paradoxe est levé si on prend en compte les motivations personnelles des enseignants et leur attitude vis-à-vis de l'institution.

5 Conclusion

Les travaux que nous avons menés dans le cadre de cette étude montrent que l'introduction d'un même jeu peut se traduire, selon les élèves, les enseignants et les institutions considérées par des situations très différentes dont les effets sur les élèves ne sont pas moins différents. À l'ECF Classcraft est utilisé comme un outil de gestion de classe principalement orienté vers le contrôle des comportements des élèves. À l'UEB, Classcraft est plutôt utilisé pour l'animation de la classe et l'obtention d'un engagement des élèves dans les activités d'apprentissage. Ainsi, les effets observés sont différents suivant les classes, mais des variations interindividuelles au sein d'une même classe sont également observables. Les raisons de ces variations sont en lien avec son caractère subjectif. Elles peuvent en effet être attribuées à la manière dont le jeu est perçu par les élèves, par les enseignants, mais également par l'institution.

La compréhension du processus de ludicisation des situations d'apprentissage est donc un défi. Une analyse réductrice qui ne prendrait en compte que les caractéristiques du jeu ne pourrait pas permettre de relever ce défi. En effet, les résultats que nous avons obtenus attestent de la complexité des interactions qui se jouent entre différents niveaux qui sont la situation de classe, la situation de conception et la culture organisationnelle de l'institution. Cette compréhension implique donc de mettre en place une approche systémique qui consiste dans la caractérisation de l'influence du jeu sur les élèves, la mise en lien de ces effets avec la manière dont l'enseignant intègre le jeu à ses pratiques et l'identification de la manière dont l'institution influence elle-même cette intégration.

Remerciements : Nous remercions la *Leading House for the Latin American Region* (Université de Saint Gall) qui a financé ces travaux et les enseignants et les directions des établissements qui ont permis la mise en place des expérimentations.

Références

1. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: Du game design au gamefulness : définir la gamification. Sciences du jeu 2014, Questionner les mises en forme ludiques du web: gamification, ludification et ludicisation (2014)
2. Seaborn, K., Fels, D.: Gamification in theory and action: A survey. International Journal of Human-Computer Studies (2015) 14-31
3. Genvo, S.: Penser les phénomènes de "ludicisation" du numérique. Pour une théorie de la jouabilité. Revue des sciences sociales (2011) 68-77
4. Sanchez, E., Young, S., Jouneau-Sion, C.: Classcraft: from gamification to ludicization of classroom management. Education and Information Technologies, Vol. 20 (2016) 497-513
5. Lieberoth, A.: Shallow gamification: Testing psychological effects of framing an activity as a game. Games and Culture, Vol.10, (2015) 229-248
6. Nault, T., Fijalkow, J.: Introduction à la gestion de la classe : D'hier à demain. Revue des sciences de l'éducation, Vol. 25 (1999) 451-466
7. Maclean-Blevins, A., Muilenburg, L.: Using Class Dojo to Support Student Self-regulation. In: EdMedia 2013-World Conference on Educational Media and Technology. Edited by Couros, A., Irvine, V., Victoria, V.: Association for the Advancement of Computing in Education (AACE) Canada (2013)
8. Dadakhodjaeva, K.: The good behavior game: Effects on and maintenance of behavior in middle-school classrooms using class dojo (2017)
9. da Rocha Seixas, L., Gomes, A., de Melo Filho, IJ.: Effectiveness of gamification in the engagement of students. Computers in Human Behavior, Vol. 58 (2016) 48-63
10. Genvo, S.: Penser les phénomènes de ludicisation à partir de Jacques Henriot. Sciences du jeu (2013)
11. Gibson, J.: The theory of affordances. In: Perceiving, Acting, and Knowing: Toward an Ecological Psychology. Edited by Shaw Bransford, R., Hillsdale, J.: Erlbaum Associates (1977)
12. Henriot, J.: Le jeu. Presses Universitaires de France Paris (1969)
13. Huizinga, J.: Homo ludens: essai sur la fonction sociale du jeu Gallimard Paris (1951)
14. Margolinas, C.: La structuration du milieu et ses apports dans l'analyse a posteriori des situations. In: Les débats de didactique des mathématiques. Edited by Margolinas, C.: La Pensée sauvage éditions Grenoble (1995) 89-102
15. Goffman, E.: Les cadres de l'expérience. Editions de Minuit Paris (1991)
16. Sanchez, E., Mandran, N.: Exploring Competition and Collaboration Behaviors in Game-Based Learning with Playing Analytics. In Lavoué, E., Drachsler, H., Verbert, K., Broisin, J., Pérez-Sanagustín M. : Lecture Notes in Computer Science, Vol. 10474 Cham: Springer (2017) 467-472
17. Casado, R., Guin, N., Champin, P., Lefevre, M.: kTBS4LA : une plateforme d'analyse de traces fondée sur une modélisation sémantique des traces. Atelier ORHEE Rendez-Vous, Font-Romeau France (2017)

Session
Analytique de l'apprentissage et tableaux de bord

Evaluer les apprentissages : comparaison de deux techniques de diagnostic de connaissance

Marie Sacksick[1]

[1] Université Paris 8, CHArt – France
& Universidad de Cordoba, KDIS – Espagne
& Domoscio – France
marie.sacksick@ntymail.com

Résumé. L'item Response Theory (IRT) et le Bayesian Knowledge Tracing (BKT) sont deux modèles d'évaluation de la maîtrise d'une connaissance par un apprenant. Ces deux modèles sont généralement comparés sur la précision de la prédiction d'une bonne réponse d'un apprenant à une question. Nous montrons dans cet article que ces deux modèles s'adressent à des situations distinctes pour lesquelles chacun d'eux peut être utilisé, et nous émettons l'hypothèse que cette hiérarchisation par la précision évolue en fonction de la quantité de réponses de l'apprenant.

Mots-clés. Théorie de la Réponse à l'Item, Bayesian Knowledge Tracing, Système Tuteur Intelligent, comparaison de modèles, modèle prédictif, modèle explicatif.

Abstract. Item Response Theory (IRT) and Bayesian Knowledge Tracing (BKT) are two evaluation models of the mastery degree of a concept by a learner. In general, these two models are compared based on the prediction prevision of the correct answer of a learner to a question. We show in this article that these models tackle different situations for which each of them can be used, and we hypothesize that this hierarchy by the precision tends to evolve according to the amount of questions the learner answered to.

Keywords. Item Response Theory, Bayesian Knowledge Tracing, Intelligent Tutoring System, model comparison, predictive model, explicative model.

1 Introduction

Lors de la dernière décade, le domaine des systèmes de tuteur informatisé (ITS, Intelligent Tutoring Systems) est entré dans une phase d'opérationnalisation et d'industrialisation, avec des entreprises comme Knewton aux Etats-Unis ou Domoscio en France. Dans un ITS, pour fournir une aide adaptée qui favorise l'apprentissage, il est nécessaire d'estimer la maîtrise d'une connaissance par un apprenant. Le choix de la technique estimant la maîtrise d'une connaissance est sujet à discussion aussi bien dans le milieu de la recherche que dans le milieu industriel.
Cette maîtrise est généralement traduite par la probabilité d'obtenir une réponse correcte à une question, cette probabilité pouvant dépendre de la question elle-même.

Ce besoin d'estimation se retrouve dans d'autres outils en support à l'apprentissage humain comme dans le système RiARiT [1].
Lorsque l'on a besoin d'estimer la maîtrise d'une connaissance par un apprenant, il faut sélectionner parmi les techniques de diagnostic qui permettent une telle estimation. Pour cela, il est courant de choisir une technique pour sa précision, c'est-à-dire sa capacité à prédire correctement si un apprenant va répondre correctement ou non à une question. Nous souhaitons souligner dans cet article d'autres critères et des spécificités aux situations, et comparer les plus connus des modèles que sont l'Item Response Theory (IRT, théorie de la réponse à l'item) et le Bayesian Knowledge Tracing (BKT). Il en existe également d'autres comme le Performance Factor Analysis [2] ou encore le Knowledge Tracing Machine [3].
Dans un premier temps nous présenterons plus avant les modèles IRT et BKT, avant de les mettre en regard dans la troisième partie.

2 Spécificités des modèles IRT et BKT

Dans cette partie, nous présentons les modèles IRT et BKT, en particulier les idées et les hypothèses sur lesquelles ils reposent.

2.1 IRT et modèles associés

L'Item Response Theory (IRT) repose sur un modèle probabiliste qui suppose une relation entre la validité de la réponse de l'apprenant à une question d'un côté et les caractéristiques de la question et la maîtrise du concept par l'apprenant de l'autre.
Etant donnés un apprenant A_j et une question Q_i, la probabilité de succès de l'apprenant à la question s'écrit :

$$P_i\left(obs = correct \middle| A_j\right) = g(A_j, Q_i) \quad \textbf{(1)}$$

Dans le modèle 2-PL (2 Parameters Logistic model), la fonction g peut s'écrire :

$$g\left(A_j, Q_i\right) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}} \quad \textbf{(2)}$$

Où θ_j est le niveau de maîtrise de la connaissance sur lequel repose la question par l'apprenant ; a_i est le niveau de discrimination de la question ; b_i représente le niveau de difficulté de la question.
Disposant des données de N_j apprenants ayant répondu à N_i questions, il est possible d'estimer la valeur de ces paramètres grâce à la minimisation de la vraisemblance ; une fonction basée sur la probabilité décrite en (1).
Le modèle de l'IRT repose sur plusieurs hypothèses [4], [5], notamment celle de la « stabilité de l'apprenant », qui suppose que le niveau de maîtrise θ est identique à tous les moments où l'apprenant répond aux questions.
La fonction g présentée dans (2) est une des plus connues, mais il existe d'autres possibilités comme le modèle de Rasch qui considère que $a_i = 1$. Il est également

possible d'étendre le modèle afin de considérer que la question dépend de plusieurs concepts ; la probabilité de répondre correctement dépend alors de plusieurs maîtrises [4].

2.2 BKT et modèles associés

Le Bayesian Knowledge Tracing (BKT) est un modèle basé sur les chaînes de Markov cachées dans lequel on trouve deux états : un état de maîtrise de la connaissance, et un état de non maîtrise. Partant du principe qu'il est impossible de savoir dans lequel de ces deux états se trouve l'apprenant, il est en revanche possible d'avoir des observations, que sont les réponses à des questions, qui donnent des indications sur l'état de l'apprenant. Lorsqu'un apprenant réussit une question, il peut être soit dans l'état de maîtrise, soit avoir deviné (*probabilité « guess »*) et être en réalité dans l'état de non maîtrise. Lorsqu'un apprenant répond incorrectement à une question, il peut effectivement ne pas connaître, ou bien s'être trompé pour des motifs comme la fatigue (*probabilité « slip »*). Répondre à une question offre à l'apprenant l'opportunité de passer de l'état « d'ignorance » à « connaissance » (*probabilité de transition*). Il est également possible que l'apprenant maîtrise déjà le concept (*probabilité initiale* avant le début du test et de ses réponses aux questions). On pourra retrouver toutes les équations associées dans [6].
Plusieurs variantes du BKT ont été proposées afin d'assouplir le modèle, notamment dans [7] où les paramètres de « *slip* » et « *guess* » sont contextualisés, ou dans [8] qui cherche à déterminer le degré de difficulté d'une question et à le prendre en compte dans le calcul des états.

3 Comparaison des deux modèles

Nous considérons dans la suite la situation suivante : dans le cadre d'un outil d'apprentissage adaptatif, nous cherchons à déterminer quelle méthode serait la plus adaptée entre l'IRT et le BKT. Cet outil d'apprentissage adaptatif recommande des contenus pédagogiques aux apprenants en fonction de leurs objectifs, de leur niveau de maîtrise des connaissances, et de leur parcours sur la plateforme d'apprentissage, et ce jusqu'à ce que ces objectifs pédagogiques soient atteints.

3.1 Fiabilité des modèles

Généralement, les modèles IRT et BKT sont comparés comme des modèles classiques d'apprentissage automatique ; autrement dit, pour plusieurs ensembles de données une méthode de validation croisée est utilisée (entraînement sur une partie des données pour tester sur l'autre partie des données), et la performance est comparée en termes de RMSE (Root Mean Squared Error) et d'AUC (Area Under the Curve).
Selon les variantes des modèles qui sont choisies, et selon les bases de données sur lesquelles elles sont testées, les résultats sont évidemment différents. Ainsi par exemple dans l'article de Minn et al. [9], l'IRT est bien meilleur en termes d'AUC

mais le BKT est un peu au-dessus en termes de RMSE. Ces performances dépendent des jeux et de données et du contexte d'utilisation de l'EIAH ; c'est la raison pour laquelle Lallé et al. ont développé un système d'assistance à la conception de techniques de diagnostic des connaissances qui permet de comparer plus facilement diverses techniques [10]. Néanmoins, il est admis que l'IRT a une capacité de prédiction généralement supérieure à celle du BKT [11].
Dans la situation pratique considérée, dans la mesure où le domaine et la population cibles ne sont pas précisés, il serait plus intéressant de choisir d'utiliser l'IRT sur ce critère de précision.

3.2 Esprit des modèles

Pour autant, on ne peut se contenter d'une comparaison de performance de précision sur une base de données. En effet, ces modèles peuvent être comparés selon d'autres critères d'utilisation, comme nous le montrons dans la suite.

Situation d'apprentissage. Une des hypothèses de l'IRT est que le niveau de connaissance de l'apprenant est fixe ; elle est questionnable puisqu'on peut se demander à quel moment un apprenant n'apprend plus, hors du moment où il maîtrise parfaitement le sujet. En dehors de cette remarque, on considère généralement que la connaissance de l'apprenant peut être considérée comme fixe lorsqu'il a déjà étudié et travaillé le sujet, qu'il est dans une situation d'évaluation, et lorsque les questions sont présentées à l'apprenant dans une fenêtre temporelle réduite. A l'inverse, dans le modèle du BKT la situation d'apprentissage supposée est celle d'une évolution d'un état de non maîtrise vers un état de maîtrise de la connaissance. Il est donc plus adapté dans une situation où l'apprenant débute et découvre le domaine.
Dans le cas pratique que nous considérons ici, il semble plus adapté d'utiliser le BKT.

Réalisme des modèles. Lorsque l'IRT est utilisé dans un outil durant une situation d'apprentissage, le modèle perd une grande partie de son réalisme et de son aspect explicatif (en opposition à son aspect prédictif). En effet, dans le cas de l'IRT, l'ordre des réponses de l'apprenant n'a pas d'importance : c'est à rebours de cette notion de progression et d'évolution qu'on a dans l'apprentissage.
Au contraire, le BKT inclut cette notion d'évolution de la connaissance. Dans un cadre évaluatif, comme une évaluation finale, il semble peu adapté : la probabilité de transition est alors théoriquement basse et beaucoup de réponses seront nécessaires pour faire évoluer les apprenants d'un état à l'autre. Il est donc tout à fait hors de propos de l'utiliser dans le cas d'une évaluation.
Cette tension entre un modèle « hautement complexe » qui a une bonne capacité de prédiction et un modèle « hautement structuré » qui a une bonne interprétabilité est mise en avant dans [12]. C'est également ce qu'il se joue ici entre l'IRT qui est plus prédictif qu'explicatif, et le BKT qui est plus explicatif que prédictif.

3.3 Conclusion du cas d'application

Dans le cas d'application considéré, nous cherchons à vérifier qu'un objectif pédagogique est atteint, autrement dit que le degré de maîtrise d'une connaissance est passée au-dessus d'un seuil. Il est donc nécessaire que l'estimation de cette maîtrise puisse être faite à la fois en posant le moins de questions possibles à l'apprenant (pour ne pas l'ennuyer) et à la fois avec le plus de certitude possible.
Il semble y avoir une contradiction entre le fait que l'IRT suppose une maîtrise fixe de l'apprenant, et le fait que l'IRT puisse prédire avec précision la justesse ou non de la réponse au moment où cette maîtrise varie le plus, autrement dit au moment des premières interactions avec la connaissance.
Dans des travaux futurs nous nous proposons donc d'étudier non pas uniquement la précision de ces deux algorithmes, mais l'évolution de cette précision dans la situation où l'apprenant découvre la connaissance : tend-on à avoir effectivement au début une prédominance du BKT qui s'efface petit à petit vers une prédominance de l'IRT ?
Si tel est le cas, il sera ensuite possible de combiner ces deux modèles en les utilisant en parallèle : il sera possible de choisir le meilleur en fonction du contexte, comme il est courant de faire lorsque l'on souhaite faire de la prédiction.

4 Conclusion

Dans cet article, nous avons comparé selon plusieurs angles les modèles de BKT et d'IRT ; ils pourraient également être comparés selon d'autres angles comme la problématique du départ à froid ou encore l'homogénéité de la population cible. Ces modèles reposent donc sur des hypothèses différentes qui ne les destinent pas au même usage. En cela, il peut ne pas être pertinent de les comparer sur des données similaires, ou uniquement sur des données qui pourraient « donner l'avantage » à l'un plutôt qu'à l'autre, car respectant mieux les hypothèses du modèle. Par ailleurs, lorsque l'on cherche à savoir lequel de ces modèles serait le plus pertinent à utiliser dans un outil lors de son développement, il est intéressant de regarder plus avant l'adéquation entre l'utilisation qui sera faite de l'outil et les hypothèses des modèles, afin de s'assurer que le calcul des paramètres donne des valeurs cohérentes.
Cette mise en miroir entre l'IRT et le BKT permet d'émettre l'hypothèse que le premier modèle est effectivement performant lorsque beaucoup de questions sont posées à l'apprenant, puisque la compréhension se stabilise autour d'un point petit à petit, là où le BKT est plus performant d'un point de vue de la prédiction sur les premières questions, et donc plus performant en général lorsque peu de questions sont posées à l'apprenant.
Nous nous proposons de vérifier cette hypothèse sur des ensembles de données dans le cas d'examens et dans le cas de plateformes d'apprentissage dans le cadre d'un travail futur.

Remerciements. Je souhaite remercier Matthieu Cisel pour son aide attentive lors de l'évolution de cette idée, mes directeurs de thèse Sebastian Ventura et Charles Tijus

pour leur accompagnement, ainsi que Benoit Praly et Simon Lemerle pour leurs relectures précieuses.

References

[1] B. Clement, D. Roy, P.-Y. Oudeyer, et M. Lopes, « Multi-armed bandits for intelligent tutoring systems », *J. Educ. Data Min.*, vol. 7, nº 2, p. 20-48, 2015.

[2] P. I. Pavlik Jr, H. Cen, et K. R. Koedinger, « Performance Factors Analysis–A New Alternative to Knowledge Tracing. », *Online Submiss.*, 2009.

[3] J.-J. Vie et H. Kashima, « Knowledge Tracing Machines: Factorization Machines for Knowledge Tracing », *ArXiv181103388 Cs Stat*, nov. 2018.

[4] M. D. Reckase, *Multidimensional Item Response Theory*. New York, NY: Springer New York, 2009.

[5] G. L. Thorpe et A. Favia, « Data Analysis Using Item Response Theory Methodology: An Introduction to Selected Programs and Applications. », 2012.

[6] M. V. Yudelson, K. R. Koedinger, et G. J. Gordon, « Individualized bayesian knowledge tracing models », in *International Conference on Artificial Intelligence in Education*, 2013, p. 171–180.

[7] R. S. J. d. Baker, A. T. Corbett, et V. Aleven, « More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing », in *Intelligent Tutoring Systems*, vol. 5091, B. P. Woolf, E. Aïmeur, R. Nkambou, et S. Lajoie, Éd. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, p. 406-415.

[8] Z. A. Pardos, N. T. Heffernan, C. Ruiz, et J. Beck, « Effective Skill Assessment Using Expectation Maximization in a Multi Network Temporal Bayesian Network », in *Proceedings of the The Young Researchers Track at the 9th International Conference on Intelligent Tutoring Systems*, 2008.

[9] S. Minn, Y. Yu, M. C. Desmarais, F. Zhu, et J. J. Vie, « Deep Knowledge Tracing and Dynamic Student Classification for Knowledge Tracing », in *arXiv:1809.08713 [cs]*, 2018.

[10] S. Lallé, V. Luengo, et N. Guin, « Assistance à la conception de techniques de diagnostic des connaissances », présenté à Environnements Informatiques pour l'Apprentissage Humain, Toulouse, 2013, p. 12.

[11] Z. A. Pardos et N. T. Heffernan, « KT-IDEM: introducing item difficulty to the knowledge tracing model », in *International Conference on User Modeling, Adaptation, and Personalization*, 2011, p. 243–254.

[12] M. Khajah, R. V. Lindsey, et M. C. Mozer, « How deep is knowledge tracing? », *ArXiv160402416 Cs*, mars 2016.

Détection Automatique d'Interactions entre Pairs dans des Jeux Sérieux Multi-Joueurs

Mathieu Guinebert[1], Amel Yessad[1], Mathieu Muratet[1,2] and Vanda Luengo[1]

[1] Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, LIP6, F-75005 Paris, France

[2] INS HEA, 58-60 Avenue des Landes, 92150 Suresnes, France

Résumé. Dans les jeux sérieux multi-joueurs (JSMJ), les apprenants interagissent entre eux à travers des activités de jeu. Cet article pose les bases d'une détection automatique des interactions entre pairs pouvant émerger d'un scénario de JSMJ. Cette détection permettrait aux enseignants et concepteurs de JSMJ de construire des scénarios adaptés à leurs besoins favorisant certaines interactions. Nous présentons une approche d'extraction de propriétés d'interactions de scénarios de JSMJ. Nous avons évalué notre approche en comparant les propriétés extraites du scénario à celles extraites des traces obtenues auprès de 114 étudiants. Les résultats de cette évaluation montrent que (1) l'algorithme d'extraction couvre bien l'ensemble des traces des apprenants et que (2) les propriétés d'interactions calculées à partir du scénario sont de mêmes types et ont des répartitions similaires à celles extraites des traces.

Mots-clés. Interactions entre pairs, Jeux sérieux Multi-joueurs, Analyse de Scénario

Abstract. In Multi-player learning games (MPLG), learners interact with each other through game activities. This article throws the basis of automatic detection of peer interactions that could emerge in MPLG scenario. This detection should allow teacher and designer to construct scenarios adapted to their needs, promoting some interactions. We present an algorithm extracting interaction features in MPLG scenarios. We evaluated this algorithm by comparing the properties extracted from the scenario to the ones extracted from the traces obtained from 114 students. The results of this evaluation show that (1) the extraction algorithm covers well the learners' traces and that (2) the calculated features from the scenario are of the same type with similar distribution as the one extracted from the traced paths.

Keywords. Peer Interactions, Multi-Player Learning Games, Scenario Analysis

1 Introduction

Les jeux sérieux multi-joueurs suscitent un intérêt grandissant chez les enseignants en leur permettant de mettre en place des scénarios où les apprenants peuvent interagir les uns avec les autres [1]. Plusieurs travaux de recherche ont déjà montré l'intérêt des interactions entre pairs dans les JSMJ sur la motivation et l'engagement des apprenants [2]. Dans cet article, nous utiliserons le terme interaction pour désigner ces interactions entre pairs. Nous ne nous intéresserons pas à l'interaction des apprenants avec le système.

Les interactions sont des phénomènes sociaux particulièrement complexes composées d'éléments verbaux, non-verbaux et sociaux [3] les rendant extrêmement difficiles à détecter automatiquement. Les interactions sont souvent mal-définies, ambigües [4] et leurs définitions manquent de consensus. La confusion fréquente entre collaboration et coopération en est un parfait exemple [5]. Pour cette raison, notre travail de recherche s'est orienté vers la définition de propriétés de bas niveau issues de la description des activités du JSMJ. Ces propriétés objectivables permettent aux enseignants ou aux concepteurs de décrire des interactions de haut niveau telles que la coopération ou la compétition. Ces descriptions sont alors exploitées pour détecter les interactions entre pairs.

A travers ce travail, nous cherchons à répondre à deux questions : (1) Comment décrire formellement des interactions entre pairs ? (QR1) et (2) Peut-on déterminer, en amont de la réalisation d'un scénario par les apprenants, les interactions entre pairs au sein d'un scénario JSMJ ? (QR2). Pour répondre à ces deux questions, nous avons mené une expérimentation visant à comparer les interactions détectées dans les traces de joueurs avec celles obtenues en analysant la structure du scénario.

2 Détecter des Interactions entre Pairs dans des Scénarios de JSMJ

2.1 Modélisation des Interactions

La revue de la littérature a montré qu'il n'existe pas de définitions consensuelles pour les interactions (par ex. la collaboration). Ce constat nous a amené à définir une interaction comme une combinaison logique de propriétés de plus bas niveau [6]. Nous pourrions par exemple définir la collaboration ainsi : plusieurs joueurs d'une même équipe acquièrent les mêmes connaissances à l'issue de l'activité de jeu sans être en conflit (l'acquisition d'une ressource par un joueur n'empêche pas les autres joueurs de l'équipe d'atteindre leurs objectifs). En termes de propriétés, nous définissons la collaboration en se servant des deux propriétés suivantes : (1) construction commune de connaissances (SCK) et (2) conflit intra-équipe (IC). La collaboration est donc décrite ainsi : $SCK \wedge \neg IC$.

Tout utilisateur peut donc proposer sa propre combinaison logique de propriétés pour désigner une interaction. Ce choix est justifié par le besoin de rendre notre système plus flexible et adaptable aux différents contextes et acteurs.

En participant à une activité (telle que décrite dans [7]), chaque apprenant est associé à un rôle prototypique (i.e. un comportement prédéfini associé à l'activité). Nous considérons que les interactions dans une activité sont dépendantes des objectifs des joueurs. Ainsi, nous associons une *goal function (GF)* au couple "rôle prototypique"/"activité". Cette fonction associe l'ensemble des ressources et compétences visées par un rôle dans une activité. Cette fonction, définie par l'enseignant ou le concepteur, sert à caractériser les objectifs d'un joueur dans le cadre d'une activité. Si l'apprenant réussit l'objectif, la fonction GF renvoie vraie.

La description d'un rôle est donnée par des fonctions de consommation et de productions de ressources. A la fin de chaque activité, l'apprenant peut atteindre ou pas ses objectifs. Nous considérons qu'un objectif est atteint si l'ensemble des ressources produites et consommées par le joueur, ainsi que l'ensemble des compétences travaillées par ce dernier, correspondent à celles définies par la fonction GF. Si aucune fonction GF n'est associée à un rôle, ce dernier est dit « passif », sinon il est « actif ». En effet, un joueur peut perdre ou obtenir des ressources dans une activité du fait d'actions d'autres joueurs et ce indépendamment de sa volonté. Celui-ci, bien qu'ayant été impacté par l'activité, n'avait pas d'objectifs dans celle-ci.

Le système que nous avons développé détecte 5 propriétés différentes définies comme suit (ces propriétés sont inspirées des travaux de Kumpulainen et Mika [4] et leur définition est détaillée dans [6]). Individualiste (Ind) : un apprenant est individualiste s'il est l'unique joueur actif dans une même activité. Conflit Externe (EC) : un ou plusieurs apprenants d'une même équipe n'atteignent pas leurs objectifs au contraire des joueurs d'une autre équipe dans la même activité. Conflit Interne (IC) : Un ou plusieurs apprenants n'atteignent par leurs objectifs au contraire de membres de la même équipe dans une même activité. Concertation (CA) : les apprenants d'une même équipe sont actifs et ne sont pas en conflit interne. Construction Commune des Connaissances (SCK) : les apprenants d'une même équipe sont actifs, ne sont pas en conflit interne et obtiennent tous les mêmes compétences à l'issue de l'activité.

Ces propriétés s'appuient sur les éléments de la modélisation d'une activité : les rôles prototypiques, les ressources produites et consommées, les compétences travaillées, les objectifs des rôles, les contraintes sur la composition des équipes, etc.

2.2 Extraction des séquences d'un scénario de JSMJ

La décomposition de nos interactions en propriétés nous permet de proposer une réponse à la question QR1. Grâce à cette décomposition, il devient possible de détecter les interactions ayant lieu au sein d'une activité de JSMJ en cherchant à identifier les propriétés la composant. Lorsque les joueurs jouent, ils enchainent les différentes activités. Chaque partie de jeu correspond donc à une séquence chronologique d'activités. L'enchaînement des activités est lié aux ressources à disposition des joueurs. Dans notre approche, l'ordre des activités n'a aucune influence sur le type ou le nombre des propriétés détectées dans une séquence puisque nous nous intéressons uniquement à la répartition (ou fonction de distribution cumulative) des propriétés dans un scénario. Si deux séquences utilisent les mêmes activités mais ordonnées différemment alors les mêmes types et nombres d'interactions seront identifiés.

Répondre à la question QR2 revient d'abord à trouver les différentes séquences d'activités que les apprenants peuvent utiliser pour atteindre la fin du jeu et ensuite à appliquer notre modèle de propriétés sur les activités composant les différentes séquences d'un scénario.

En raison du nombre élevé de séquences d'activités possibles dans un scénario, nous travaillons uniquement sur un sous-ensemble des séquences qui doit préserver l'ensemble des propriétés d'interactions pouvant émerger du scénario. Nous avons donc développé un algorithme se basant sur la description des activités pour extraire, d'un scénario de JSMJ, un sous ensemble (noté SA) des séquences valides pour atteindre la fin d'un scénario. Une seule combinaison d'activités est retenue pour des séquences ayant les mêmes activités répétées (nous ne pouvons avoir la séquence b->a dans SA si la séquence a->b y est déjà).

3 Expérimentations

Nous avons développé deux scénarios d'un jeu sérieux multi-joueurs que nous avons fait tester à 114 étudiants de deuxième année de licence. Grâce à ces tests, nous avons obtenus 132 séquences à deux joueurs sur le premier scénario (noté TD2) et 62 séquences à quatre joueurs sur le deuxième scénario (noté TD4).

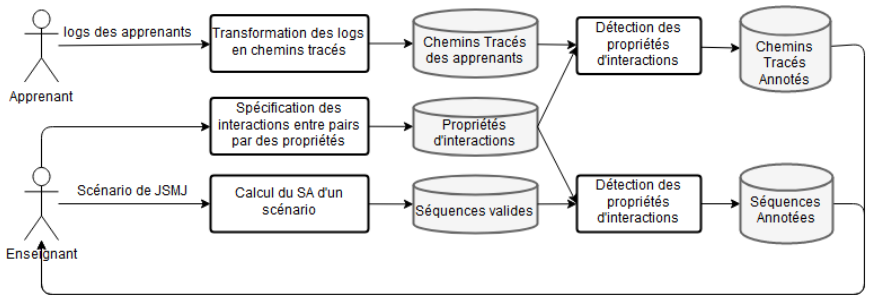


Fig. 1. Framework pour la détection automatique des interactions multi-joueurs dans un scénario JSMJ et des chemins tracés

Nous qualifions de chemin tracé (CT), l'ensemble des activités utilisées dans une séquence par les joueurs. La Figure 1 permet de visualiser le fonctionnement global de notre système dans le cadre de cette expérimentation.

D'abord, nous avons détecté les propriétés d'interactions d'une part dans les CT et d'autre part dans l'ensemble SA calculé, et ceci pour les deux scénarios TD2 et TD4. Ensuite, nous avons analysé ces propriétés en comparant leurs répartitions dans SA et CT (voir Table 1).

Le premier résultat que nous avons obtenu est que l'ensemble SA inclut tous les chemins tracés des joueurs et que seulement 13% des séquences SA du TD2 (7% des séquences SA pour le TD4) sont réellement réalisées par les joueurs. Ce résultat montre que l'ensemble SA est bien plus exhaustif que les chemins tracés réalisés par les joueurs.

Table 1. Les pourcentages de répartition des propriétés d'interactions dans les séquences du SA, dans les séquences utilisées du SA et dans les CT des joueurs

Propriétés (%)	Séquences du TD2			Séquences du TD4		
	SA	Séquences utilisées	CT	SA	Séquences utilisées	CT
EC	0	0	0	47.32	48.4	48.02
IC	20	17.34	15.14	10.53	10.56	10.96
Ind	20	16	22.69	10.53	6.15	4.24
CA	40	41.33	36.08	21.08	21.11	21.26
SCK	20	25.33	26.09	10.53	13.78	15.52

Nous avons appliqué le test du khi2 sur les résultats de la Table 1 pour vérifier l'indépendance des répartitions de propriétés, celle du SA et celle des CT des joueurs. L'idée étant alors la suivante : les apprenants sont libres de leurs actions et peuvent réutiliser plusieurs fois une même séquence du SA (comme indiqué plus haut). Ainsi, même si CT et SA utilisent la même modélisation, la répartition de leurs propriétés n'est pas obligatoirement similaire ou dépendante. Or, les résultats du test montrent clairement une dépendance entre les deux répartitions, et ceci dans le cas des deux scénario TD2 et TD4. Ce résultat montre, dans le cas de cette étude, que les propriétés extraites à partir des séquences du SA sont similaires (en type mais aussi en distribution) aux propriétés que nous pouvons extraire des CT des joueurs. L'analyse, à priori, d'un scénario JSMJ permet donc de fournir des informations pertinentes sur les interactions. Ce résultat est important car il permet aux enseignants et aux concepteurs de scénarios de détecter les interactions de leurs scénarios, en amont de leur utilisation par des apprenants, et donc de pouvoir les modifier/corriger si ces interactions ne répondent pas à leurs besoins. Toutefois, ces résultats sont propres au contexte de l'étude présentée ici, d'autres études devraient être menées dans d'autres contextes pour les valider.

4 Conclusion

Dans ce papier, nous avons travaillé sur le problème de la détection automatique d'interactions entre pairs dans les JSMJ. Pour répondre à nos deux questions de recherche, nous avons proposé deux contributions.

La première, consiste à décomposer les interactions en propriétés détectables au sein d'activités multi-joueurs. L'avantage de ces propriétés étant de permettre à un enseignant ou un concepteur de définir lui-même sa vision d'une interaction. Les définitions des interactions faisant consensus, cet aspect s'avère intéressant.

La deuxième a été de fournir un algorithme en mesure d'extraire les séquences d'activités que les joueurs peuvent emprunter pour finir le jeu. L'application de notre modèle de propriétés d'interactions sur ces séquences permet donc d'obtenir les propriétés d'interactions issues du scénario étudié.

Les résultats que nous avons obtenus sont encourageant et semblent indiquer que notre méthode est capable de répondre à notre problème. Toutefois ceux-ci étant issus d'une unique expérimentation, ils ne sont pas encore suffisants pour être généralisés.

Remerciements

Nous voudrions remercier la chaire EIAH de Sorbonne-Universités pour avoir financé ce travail et les étudiants ayant participé à l'étude expérimentale.

Références

1. Turkay, Selen, et al. "Toward understanding the potential of games for learning: Learning theory, game design characteristics, and situating video games in classrooms." Computers in the Schools 31.1-2 (2014): 2-22.
2. Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An investigation of the interrelationships between motivation, engagement, and complex problem solving in game-based learning. Educational technology & society, 17(1), 42-53.
3. Kumpulainen, Kristiina, and Mika Mutanen. "The situated dynamics of peer group interaction: An introduction to an analytic framework." Learning and instruction 9.5 (1999): 449-473.
4. Kreijns, K., Kirschner, P.A., Jochems, W.: Identifying the Pitfalls for Social Interaction in Computer-supported Collaborative Learning Environments: A Review of the Research. In: Computers in Human Behaviour, Vol. 19 Issue 3 (2003) 335–353
5. Kozar, O. (2010). Towards Better Group Work: Seeing the Difference between Cooperation and Collaboration. In English Teaching Forum (Vol. 48, No. 2, pp. 16-23).
6. Guinebert, M., Yessad, A., Muratet, M., & Luengo, V. (2018, October). Analysis of Peer Interaction Features in Multi-Player Learning Game: Semi-Automatic Approach and Proof of Concept. In ECGBL 2018 12th European Conference on Game-Based Learning (p. 162). Academic Conferences and publishing limited.
7. Guinebert, M., Yessad, A., Muratet, M., & Luengo, V. (2017, September). An ontology for describing scenarios of multi-players learning games: toward an automatic detection of group interactions. In European Conference on Technology Enhanced Learning (pp. 410-415). Springer, Cham.

Moodleboard : un tableau de bord dynamique et interactif pour les ingénieurs pédagogiques

Véronique Sébastien[1], Didier Sébastien[2], Dominique Gay[3]

[1] Institut Indianocéanique du Numérique (Université de La Réunion),
[2] Direction des Usages du Numérique (Université de La Réunion),
[3] Laboratoire d'Informatique et de Mathématiques (Université de La Réunion),
15 avenue René Cassin, 97400 Saint Denis, REUNION (FRANCE)
veronique.sebastien/didier.sebastien/dominique.gay@univ-reunion.fr

Résumé. Cet article présente Moodleboard, un tableau de bord dynamique et interactif d'aide à la décision destiné aux ingénieurs pédagogiques et aux administrateurs fonctionnels de plateformes pédagogiques Moodle. Ce tableau de bord présente différents indicateurs pouvant être explorés de manière interactive, avec plusieurs finalités : recueillir des statistiques d'usages de la plateforme à différents niveaux de granularité, mais aussi modéliser les différents types de cours numériques, les classifier, repérer les détournements d'usage, ou encore les innovations mises en place par les équipes pédagogiques et administratives des établissements. Nous exposons dans cet article la méthodologie de conception de Moodleboard fondé sur des entrepôts de données massives, des algorithmes de fouille de données et des outils de visualisation interactifs. Nous décrivons également les premiers résultats mis en évidence sur les cours numériques de l'Université de La Réunion.

Mots-clés. Tableau de bord, aide à la décision, fouille de données, plateforme pédagogique, clustering.

Abstract. This article deals with Moodleboard, a dynamic and interactive dashboard which can be used as a decision support tool by pedagogical engineers and Moodle platforms administrators. This dashboard presents various indicators which can be explored in an interactive way to achieve different goals: obtain statistics about the platform usage with different granularities, but also model digital courses, classify them, detect misuses or innovations implemented by teaching or administrative teams in an institution. In this article, we describe the design of Moodleboard, based on datawarehouses, datamining algorithms and datavizualisation environments. We also give the first results highlighted by Moodleboard on the pedagogical platform of Reunion Island University.

Keywords. Dashboard, decision support, datamining, pedagogical platform, clustering

1 Introduction et contexte

A l'ère des MOOCs et des réseaux sociaux, les universités proposent aujourd'hui de nombreux services administratifs et pédagogiques en ligne à leurs étudiants et personnels. Cependant, la mise en place de ces nouvelles technologies ne s'accompagne que rarement d'un retour fiable sur leurs usages et sur leurs apports réels à la communauté. Dans ce contexte, l'Université de La Réunion (UR) s'est doté d'un Observatoire des Usages du Numérique (OBSUN) en 2015. L'OBSUN a notamment mis en place des entrepôts de données permettant le recueil et l'analyse des traces d'utilisation des différents services numériques déployés à l'UR.

Parallèlement à ce contexte, la plateforme Moodle [1] de l'UR a vu son usage démultiplié ces dernières années. Elle est ainsi passée de 250 cours en 2015 à 1400 en 2018. Cependant, les contenus et les usages peuvent grandement varier d'un cours à l'autre. Peu d'enseignants ont suivi une formation à l'usage de cet outil et/ou sollicitent l'accompagnement de la Direction des Usages du Numérique (DUN). Les attentes en la matière se focalisent plutôt autour d'un accompagnement individuel et ponctuel du fait du peu de disponibilité des enseignants-chercheurs, malgré un intérêt manifeste pour ces outils.

Les ingénieurs pédagogiques et administrateurs fonctionnels de la plateforme, bien que validant chaque demande de création de cours, ont donc une visibilité restreinte sur la manière dont la plateforme est utilisée par les enseignants. La détection d'initiatives innovantes ou de cours non utilisés reste difficile. Par ailleurs, ces derniers sont régulièrement sollicités par les décideurs afin d'obtenir des indicateurs précis à différentes échelles. L'accès à ces indicateurs doit être simple et disponible à tout moment et en tout lieu.

Face à ces constats récurrents dans les établissements proposant des plateformes pédagogiques en ligne, l'OBSUN a développé Moodleboard, un tableau de bord dynamique et interactif, en partenariat avec la DUN et le Laboratoire d'Informatique et de Mathématiques (LIM) de l'UR. Dans la partie suivante, nous présentons tout d'abord d'autres travaux sur les tableaux de bord d'observation des activités d'apprentissage en ligne. Puis nous décrivons la méthodologie de conception de Moodleboard et ses principaux indicateurs et fonctionnalités. Dans une troisième partie, nous analysons les résultats d'un tel outil à l'UR. Enfin, nous discutons des perspectives de développement de notre outil et des possibilités de modélisation qu'il ouvre.

2 Les tableaux de bord pour l'analyse de l'apprentissage et de l'enseignement

De nombreux outils d'observation et d'analyse des activités pédagogiques portées par le numérique existent. Chacun se dédie à des plateformes d'apprentissage et à des activités spécifiques, même si l'on retrouve souvent des indicateurs-types similaires : suivi des connexions des apprenants, scores obtenus, détection des décrocheurs, etc.

Les chercheurs du projet HUBBLE ont publié en 2017 un rapport sur les modèles et processus de génération de ces tableaux de bord [2]. Nous reprenons de cette étude la définition de la notion de tableau de bord d'apprentissage qui y est donnée, elle-

même issue des travaux de Schwendimann et al. [3] : « un tableau de bord pour l'analyse de l'apprentissage est un affichage unique agrégeant différents indicateurs sur les apprenants, les processus d'apprentissage et les contextes d'apprentissage au travers d'une ou de plusieurs visualisations ». On adjoint à cette définition l'activité d'enseignement, qui a naturellement un impact sur les processus d'apprentissage.

En ce qui concerne spécifiquement les outils dédiés à la plateforme Moodle, on constate que l'offre est assez importante bien que plutôt uniforme. Le plugin « Dashboard » [4] par exemple permet de générer de nombreux graphiques à partir de requêtes sur la base de données de Moodle. Cependant, l'utilisation de cet outil reste réservée à des informaticiens spécialistes de l'architecture de Moodle. L'outil Intelliboard [5] est beaucoup plus accessible et visuellement attractif. Il permet d'avoir des statistiques en temps réel sur l'utilisation de la plateforme pédagogique, mais aussi de suivre dans le détail la progression des apprenants sur un cours donné. Des personnalisations sont également possibles quant à la disposition des différents graphiques et à l'envoi de notifications en cas de détection de profils « décrocheurs ». Cependant, le module Intelliboard est payant, et requiert d'autoriser la société à accéder aux données de la plateforme de l'établissement. L'Open University Hellénique a également développé un outil très abouti pour suivre la progression des apprenants sur Moodle [6] avec notamment une détection des ressources non exploitées par les apprenants, et la visualisation des interactions au sein des forums de discussion. L'intérêt de cet outil réside également dans la possibilité pour l'enseignant de réagir face à un constat directement depuis le tableau de bord via l'envoi de messages ciblés par exemple.

L'ensemble des outils présentés permettent d'avoir une vision très nette de l'engagement d'un étudiant sur une telle plateforme. Cependant, cet engagement est grandement conditionné par la nature et la qualité des ressources et activités qui y sont proposées. En effet, peut-on réellement mesurer l'efficacité d'un apprentissage sur la base de statistiques de téléchargement d'une ressource pédagogique ? Il convient donc d'avoir une vision plus précise des ressources et activités proposées aux étudiants. C'est pourquoi nous proposons dans ce qui suit une méthodologie et un outil d'observation intuitive des contenus d'une plateforme pédagogique.

3 Moodleboard : de l'entrepôt de données aux indicateurs

Moodleboard est un tableau de bord interactif et dynamique permettant de visualiser des données massives sur les cours d'une plateforme pédagogique. La **Fig. 1** présente une capture d'écran de l'outil. Les caractéristiques observables sont le nombre de connexions aux cours depuis le début de l'année scolaire, le nombre d'inscrits, de sections, d'enseignants, de ressources téléchargeables, de messages dans les forums, de devoirs à rendre, de tests, de wikis, d'événements et de glossaires à la date d'observation.

Cliquer sur les différents graphiques permet de filtrer les données et d'explorer ainsi des catégories spécifiques de cours : il est alors possible par exemple de visualiser les caractéristiques de l'ensemble des cours de L2 en Sciences, Technologies et Santé en deux clics. Une barre de recherche permet également d'isoler certains cours par

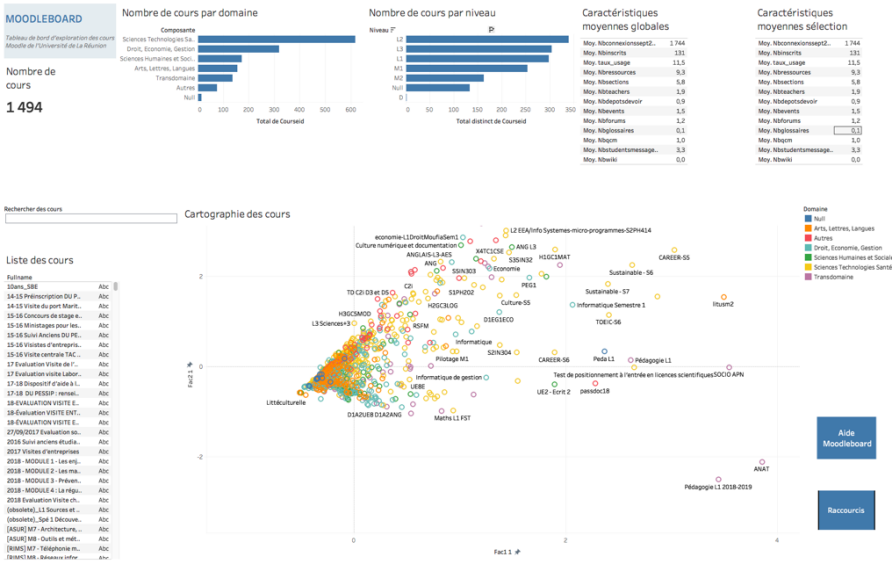


Fig. 1. Capture d'écran de l'outil Moodleboard : statistiques générales et cartographie des cours

mots-clés. On pourra ainsi comparer entre eux l'ensemble des cours d'anglais de la plateforme, indépendamment des filières ou niveaux concernés. Les tableaux « caractéristiques moyennes globales » et « caractéristiques moyennes sélection » permettent d'opérer des comparaisons entre un sous-groupe de cours et l'ensemble des cours de la plateforme. Il est également possible de ne sélectionner qu'un seul cours pour consulter ses caractéristiques propres.

La cartographie des cours représente un positionnement de ces derniers dans un espace à deux dimensions. Ce positionnement est réalisé en fonction des différentes caractéristiques citées ci-dessus à l'aide d'une Analyse en Composantes Principales (ACP). Dans notre exemple, appliqué au cas de l'Université de La Réunion, les deux premiers axes de l'ACP expliquent 44% de la variance totale (voir partie 4). Des cours proches sur la carte possèdent donc des caractéristiques similaires. Des points isolés représentent des cours sortant de la norme selon les critères observés. Les cours ayant une composante 1 assez forte (plutôt à droite sur la cartographie) ont une participation assez importante dans leurs forums et un nombre de connexions élevé. Les cours ayant une composante 2 assez forte (plutôt en haut sur la cartographie) ont un faible nombre d'inscrits et de connexions mais un nombre de ressources et d'activités plus élevé.

Moodleboard s'appuie sur un entrepôt de données Big Data Hadoop dans lequel sont enregistrés quotidiennement les indicateurs de chacun des cours. Les requêtes sont exécutées chaque nuit afin d'extraire les caractéristiques des cours à partir de la base de données de Moodle. Divers traitements sont appliqués aux données, notamment afin de calculer les nouvelles coordonnées des cours via l'ACP. Le tableau de bord est connecté à l'entrepôt de données et donc rafraichi automatiquement dès que de nouvelles données y sont intégrées.

Cet entrepôt permet également d'étudier le processus de construction d'un cours en ligne dans le temps. Sur certains cours, on constate ainsi l'augmentation du nombre d'étudiants inscrits d'année en année, mais également l'augmentation du nombre de ressources et d'activités, traduisant l'enrichissement progressif du cours en ligne par l'enseignant. Un taux d'usage de chaque cours est également calculé directement à partir de l'outil de tableau de bord. Ce taux permet de classer les cours du plus exploité par ses apprenants au moins exploité. Il est calculé en divisant le nombre de connexions réalisées par des étudiants du cours depuis le début de l'année scolaire (cette date est redéfinie à chaque nouvelle année scolaire dans le script d'extraction des données) par le nombre d'inscrits au cours.

Des matrices de nuages de points entre variables permettent de visualiser les liens pouvant exister entre différentes caractéristiques des cours. Des liens peuvent ainsi être identifiés entre la taille de groupes d'apprenants et la participation dans les forums. L'ACP réalisée sert également de base à un partitionnement (clustering) des cours. L'algorithme utilisé est le k-means avec un nombre optimal de partitions défini à l'aide du critère de Calinski-Harabasz. Ce partitionnement permet d'identifier d'éventuels groupes homogènes de cours sur la base des caractéristiques observées, et de répondre à des hypothèses du type : « les enseignants d'un même domaine ou d'une même discipline construisent-ils des cours de la même manière ? ».
La partie suivante présente quelques résultats mis en évidence à l'Université de La Réunion.

4 Étude de cas : résultats à l'Université de La Réunion

La principale plateforme Moodle de l'Université de La Réunion compte 1494 cours en 2019. Les visualisations générées par Moodleboard démontrent un usage très homogène au regard des caractéristiques observées, à quelques exceptions près. Les partitionnements réalisés ne permettent pas de retrouver des domaines ou des niveaux d'études particuliers, ce qui démontre que ces paramètres n'ont pas d'influence sur la façon de construire des cours sur la plateforme.

Cependant, notre outil met très nettement en évidence les cours sortant de l'usage classique « dépôt de document ». Ces cours, qui représentent 9% de la population observée, appartiennent à quatre clusters de petite taille identifiés par l'algorithme k-means. Ils se caractérisent par un taux d'usage entre 4 fois et 30 fois plus élevé, et des activités plus nombreuses, notamment au niveau du nombre de tests proposés, de l'activité des forums et du dépôt de devoirs. Si les effectifs de ces cours ne sont pas systématiquement plus élevés que la moyenne, on constate en revanche que le nombre d'enseignants l'est (entre 3 et 5 enseignants). Là encore, le domaine d'étude ne semble pas être un facteur discriminant l'appartenance ou non à ce groupe de cours « hors norme », bien que les cours en sciences soient globalement surreprésentés sur la plateforme pédagogique.

La cartographie des cours permet également d'identifier très facilement les cours « exceptionnels » car présentant des valeurs remarquables sur les variables retenues. A titre d'exemple, le cours de Physiologie de la L1 STAPS (Sciences et Techniques des Activités Physiques et Sportives) présente plus de 1000 inscrits, plus de 500 000

connexions depuis le mois de septembre 2018, et surtout plus de 1000 messages d'étudiants sur les forums. Il conviendrait de se pencher sur les raisons d'une telle utilisation intensive par les étudiants et d'adapter l'accompagnement pédagogique et technique proposé aux usagers de ce cours.

Moodleboard facilite également la détection des cours consacrés à la formation aux outils numériques. En effet, ces derniers ont des caractéristiques propres à les éloigner de la masse des cours classiques. Le cours « bac à sable » par exemple est évidemment peu représentatif car de futurs formateurs s'y entraînent à créer des activités. Mais il n'a plus été modifié depuis 2016 et pourrait donc faire l'objet d'un archivage.

5 Conclusion et perspectives

Nous avons décrit dans cet article la méthodologie de conception et les premiers résultats de notre outil Moodleboard. Ce tableau de bord interactif et dynamique s'appuie sur des entrepôts de données, des algorithmes de fouille de données, et des outils de visualisation de données (dataviz) afin de fournir des indicateurs utiles aux ingénieurs pédagogiques et administrateurs fonctionnels de plateformes Moodle. L'interactivité proposée par Moodleboard permet à ces derniers de sélectionner eux-mêmes l'échelle d'observation permettant de répondre à leurs besoins.

De nombreuses perspectives sont envisagées pour cet outil, depuis l'analyse des mots-clés et descriptions des cours, jusqu'à l'exploitation des résultats des tests proposés sur la plateforme : sont-ils plutôt utilisés à titre de positionnement initial, d'entrainement, ou d'évaluation ? la discipline étudiée a-t-elle un impact sur la réussite à ces tests ? Il serait également intéressant de comparer les cartographies obtenues d'un établissement à l'autre, et de mesurer l'impact des politiques incitatives en matière de numérique sur l'usage des plateformes pédagogiques. Ces travaux pourraient aboutir à une modélisation des cours numériques en fonction des types de formations et de leurs caractéristiques (présentiel, distanciel, hybride, continue ou initiale) à des fins d'adaptabilité des cours et des accompagnements pouvant être proposés.

Remerciements. Les auteurs remercient Eloi Le Renard pour ses Travaux d'Étude et de Recherche qui ont permis d'initier ce projet dans le cadre de son Master 1 d'Informatique.

Références

1. https://moodle.org/, visité le 27/01/2019
2. Dabbebi, I., Gilliot, J.: Modèles et processus de generation de tableaux de bord. Livrables L5.3 et L5.4 du projet ANR HUBBLELEARN (2017)
3. Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., & Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies* , *10* (1), 30-41.
4. https://moodle.org/plugins/block_dashboard, visité le 27/01/2019
5. https://www.intelliboard.net/, visité le 27/01/2019
6. Gkontzis, A. F., Karachristos, C. V., Lazarinis, F., Stavropoulos, E. C., & Verykios, V. S. (2017). Assessing Student Performance by Learning Analytics Dashboards. *Διεθνές Συνέδριο για την Ανοικτή & εξ Αποστάσεως Εκπαίδευση*, *9*(1A), 101-115.

Tableaux de Bord d'Aide au Suivi pour la Formation Professionnelle en Ligne : Étude d'Indicateurs et Lignes Directrices

Mohamed Mouaici[1,2], Laurence Vignollet[1], Chrisitine Galez[1], and Mael Etienne[2]

[1] Université Savoie Mont Blanc, France
{mohamed.mouaici,laurence.vignollet,christine.galez}@univ-smb.fr
[2] Société Logipro, France
{mohamed.mouaici,mael.etienne}@logipro.com

Résumé. Nous présentons ici une étude expérimentale qui vise à évaluer la pertinence de six indicateurs de l'activité d'apprenants dans un contexte de formation professionnelle en ligne, choisis parmi ceux communément référencés dans la littérature. L'étude a été réalisée sur un échantillon de données incluant 84 activités e-learning suivies par 214 apprenants. Nous démontrons que le choix d'indicateurs de suivi peut être fortement impacté par le contexte d'application et ses contraintes intrinsèques, par le type de formation, par le scénario pédagogique, ainsi que par les objectifs de la formation. Ensuite, nous donnons quelques lignes directrices qui découlent de notre étude, et qui pourraient faciliter la conception d'un tableau de bord d'aide au suivi, en termes d'indicateurs, dans le cadre de la formation professionnelle en ligne.

Keywords: Étude Expérimentale, Indicateurs, Lignes Directrices, Suivi d'apprentissage, Formation Professionnelle en Ligne.

Abstract. In this paper, we present an experimental study that aims to evaluate the relevance of six indicators selected from those commonly referenced in the literature, in a context of online professional training. The study was conducted on a set of data including 84 e-learning activities followed by 214 learners. We show that the choice of indicators can be impacted by their context of application and its intrinsic constraints, by the type of training, by the pedagogical scenario, as well as by the objectives of the training. Then, we give some guidelines that come out of our study, and that could facilitate the design of a learning analytics dashboard, in terms of indicators, as part of online professional training.

Keywords: Experimental Study, Indicators, Guidelines, Learning Monitoring, Online Professional Training.

1 Introduction

Nos travaux de recherche s'inscrivent dans le domaine des "Learning Analytics Dashbords (LAD)" dédiés à la formation professionnelle en ligne. Nous tra-

vaillons en collaboration avec un organisme de formation (OF) qui dispense divers types de formations professionnelles dans différents secteurs d'activités. L'OF dispose aussi de sa propre plateforme d'apprentissage (LMS) où les apprenants peuvent consulter les différentes activités e-learning auxquelles ils sont inscrits. Une activité e-learning peut être composée de plusieurs types de contenu multimédia : vidéos, textes, quiz, etc. Actuellement, l'objectif de l'OF est de se doter d'un LAD efficace i.e. qui lui permettrait de mieux suivre les apprenants avec la volonté d'accroitre leur réussite et leur satisfaction, mais aussi d'améliorer la qualité des formations proposées.

Afin de répondre à ce besoin, nous avons adopté une approche centrée utilisateur en incluant les parties prenantes tout au long du processus de conception [1]. Les parties prenantes impliquées dans nos travaux sont les ingénieurs pédagogiques de l'OF dont le rôle est de concevoir les formations professionnelles. De plus, certains d'entre eux ont la responsabilité de suivre les apprenants et d'aider ceux qui sont en difficulté et qui peuvent donc échouer. Pour cela, ils ont besoin d'un LAD incluant un ensemble d'indicateurs leur permettant de comprendre le comportement des apprenants vis-à-vis de la réussite/l'échec dans les formations professionnelles en ligne. Avoir ce type de suivi leur permettrait non seulement d'aider les apprenants à mieux réussir mais aussi d'améliorer les formations s'ils constatent qu'une activité représente un point de blocage pour la majorité des apprenants (par exemple un concept mal expliqué).

Suite à plusieurs séances de Focus Group[1], nous avons constaté que les besoins en termes de LAD sont différents pour les ingénieurs pédagogiques en fonction de leurs objectifs de suivi. Ce même constat d'absence de consensus dans la définition des LAD a été observé lors de notre étude de l'état de l'art[2,3,4]. De plus, nous avons constaté que les solutions de type LAD dédiées à la formation professionnelle sont très peu explorées dans la littérature[5,6,7]. En effet, le peu de travaux qui existent s'intéresse plutôt à l'aspect informel de la formation professionnelle : apprentissage par la pratique sur le lieu de travail, interaction avec les collègues ou avec son réseau professionnel[6,7].

Ainsi, la contribution principale de cet article réside dans l'étude de la pertinence de six indicateurs régulièrement mentionnés dans la littérature, vis-à-vis de la réussite/l'échec des apprenants. Les résultats de cette étude nous conduisent à la proposition de plusieurs lignes directrices qui pourront faciliter la conception d'un LAD dédié à la formation professionnelle en ligne.

2 Étude Expérimentale

Notre étude expérimentale consiste donc à tester la pertinence de six indicateurs communément référencés dans la littérature[2,3,4,6], dans notre contexte de la formation professionnelle. L'étude de la pertinence signifie que l'on cherche s'il existe un lien entre la réussite/l'échec de l'apprenant et les indicateurs choisis qui permettrait d'expliquer la réussite/l'échec. Autrement dit, si nous disposons d'un tableau de bord avec les six indicateurs, serait-on capable de comprendre objectivement la raison de la réussite/de l'échec de l'apprenant. Pour ce faire,

un ensemble de données assez conséquent (84 activités et 214 apprenants) a été utilisé dans notre étude afin de garantir la validité de nos résultats.

La rareté d'études pour la formation professionnelle en ligne nous a amenés à choisir ces indicateurs en nous basant sur des travaux ayant des similarités avec notre contexte : l'apprentissage en ligne[2,3,4,6]. Ces six indicateurs sont :

Temps total passé par activité (I1) : durée totale que l'apprenant a passée sur l'ensemble des contenus d'une activité e-learning.

Nombre de visualisations de contenus (I2) : nombre total de fois où l'apprenant a visualisé le contenu d'une activité.

Nombre de jours actifs sur une activité (I3) : nombre total de jours où l'apprenant a été actif (il a fait une action sur l'activité : visualisation d'un contenu par exemple) sur une activité e-learning.

Nombre de jours actifs sur la plateforme (I4) : nombre total de jours où l'apprenant a été actif (il fait une action sur la plateforme : connexion à l'espace apprenant, ou sur l'activité) sur la plateforme d'apprentissage.

Temps moyen passé par visualisation d'un contenu (I5) : temps moyen passé par un apprenant à chaque visualisation de contenu.

Temps moyen par jours actifs (I6) : moyenne de temps passé par jour par l'apprenant sur une activité e-learning.

Par la suite, notre démarche expérimentale pour les six indicateurs a été appliquée sur un échantillon de données construit à partir des traces d'apprentissage issues du LMS de l'OF (pour les 84 activités suivies par 214 apprenants). Notre processus d'analyse de ces indicateurs est établi en deux étapes :

1. **Étude de corrélation entre les indicateurs et la réussite des apprenants :** dans un premier temps, nous avons privilégié la réussite/l'échec comme objectif de suivi, en partant des résultats obtenus par les apprenants aux quiz d'évaluation. Nous avons cherché à observer le comportement des apprenants en analysant l'intensité de la liaison qui existe entre les indicateurs et la réussite des apprenants. Pour cela, nous avons calculé le facteur de corrélation entre les indicateurs et la réussite des apprenants (Tableau 1).
2. **Examen des résultats avec les ingénieurs pédagogiques :** nous avons procédé à un examen critique de nos résultats avec les ingénieurs pédagogiques qui ont plus de connaissances sur la nature des formations et les paramètres qui pourraient influencer la pertinence de ces indicateurs et notamment sur les différents scénarios pédagogiques utilisés.

Le tableau 1 illustre les facteurs de corrélation que nous avons obtenus entre les indicateurs et la réussite des apprenants. Ainsi, sur les 84 activités, nous rapportons la valeur minimum, maximum et la valeur moyenne pour le facteur de corrélation obtenu avec chaque indicateur.

Les valeurs moyennes des facteurs de corrélation que l'on observe dans le tableau 1 démontrent que la liaison entre les indicateurs étudiés et la réussite des apprenants est très faible. En effet, le facteur de corrélation moyen est très proche de zéro pour les six indicateurs. Cela indique que la réussite des apprenants n'est

Table 1. Facteurs de corrélation entre les indicateurs étudiés et la réussite des apprenants. FC(Ii,R) : Facteur de Corrélation entre l'indicateur i et la **R**éussite.

Corrélations	Minimum	Maximum	Moyenne
FC(I1,R)	-.639	.73	**.010**
FC(I2,R)	-.592	.693	**.084**
FC(I3,R)	-.063	.735	**.050**
FC(I4,R)	-.624	.789	**.068**
FC(I5,R)	-.558	.711	**-.008**
FC(I6,R)	-.66	.724	**.008**

pas directement liée à ces indicateurs. De plus, le facteur de corrélation varie de -63.9% à 78.9% sur l'ensemble des activités. Nous avons alors observé en détail la distribution de notre échantillon par rapport à la réussite sur ces activités pour l'ensemble des indicateurs. Les Fig. 1 et Fig.2 illustrent un exemple de la distribution des notes obtenues par les apprenants (donc la réussite ou l'échec) en fonction des indicateurs I1 et I2 pour une même activité représentative (le facteur de corrélation est de 71% pour I1 et de 25 % pour I2, et avec un nombre d'apprenants important : 71 apprenants). Sur ces deux figures, la ligne horizontale désigne la note minimale requise pour réussir, la ligne verticale correspond à la moyenne de l'échantillon par rapport à l'indicateur concerné sur cette activité, et les points désignent la note obtenue par les apprenants sur l'activité. Les résultats détaillés et les données de notre étude sont accessibles via ce lien [1].

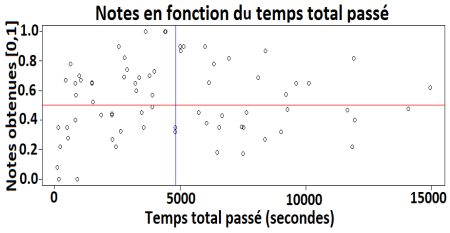


Figure 1. Distribution des notes en fonction du temps total passé par activité (I1).

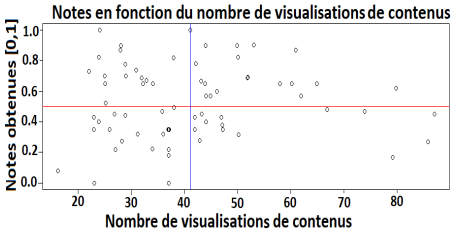


Figure 2. Distribution des notes en fonction du nombre de visualisations (I2).

Les résultats observés sur cette activité démontrent la difficulté de trouver une tendance claire entre la réussite et ces deux indicateurs I1 et I2. Nous observons par exemple que certains apprenants passent beaucoup de temps contrairement à d'autres pour réussir alors que d'autres passent moins de temps que la moyenne (la ligne verticale) en ayant de meilleurs résultats. Ce constat est valable aussi pour les autres indicateurs sur l'ensemble des 84 activités. Ceci doit amener, dans le contexte de la formation professionnelle en ligne, à être prudent quant à l'utilisation d'indicateurs pourtant proposés et validés dans la littérature et à étudier de manière approfondie les facteurs d'influence de ces

1. https ://owncloud.logipro.com/owncloud/s/Nx7dQAXsywKpE4

indicateurs même pour des formations paraissant assez similaires en première approche (formations en ligne). L'importance de ces indicateurs dans la littérature nous a poussés à approfondir avec nos parties prenantes notre analyse afin de mieux comprendre et expliquer les résultats que nous avons obtenus et de proposer ainsi des lignes directrices qui permettront de faciliter la conception d'une solution d'aide au suivi efficace dans le cadre de la formation professionnelle.

3 Analyses des Résultats et Lignes Directrices

Les analyses que nous avons réalisées avec les ingénieurs pédagogiques, via des séances de travail [1], nous ont permis de confirmer notre constat précédent concernant la pertinence des indicateurs étudiés pour le suivi de la réussite. De plus, nous avons pu mettre en évidence certaines particularités liées au contexte de la formation professionnelle et qui permettraient d'expliquer la corrélation très faible entre les indicateurs étudiés et la réussite des apprenants. Nous avons identifié deux grands types de formations : **les formations validées par le résultat des évaluations** et **les formations validées par le temps passé** où la durée minimum exigée est en général fixée par le financeur de la formation. Dans ce deuxième cas de figure, la présence d'évaluations dans les activités e-learning n'a pas forcément d'enjeux aux yeux de l'apprenant car elles ne sont pas pénalisantes pour sa réussite. Par conséquent, tous les apprenants passent à peu de choses près une même durée totale à se former. Ceci pourrait bien justifier l'absence de corrélation entre le temps passé (I1) et la réussite des apprenants. Cette même contrainte peut aussi expliquer la liaison faible entre les cinq autres indicateurs et la réussite des apprenants, car, par exemple, un nombre de visualisations ou un nombre de jours actifs important, peut ne pas être significatif si l'apprenant se limite seulement à la durée qu'on lui impose, ce qui est très fréquent dans la formation professionnelle.

Grâce à cette analyse, ainsi qu'aux échanges réguliers avec les ingénieurs pédagogiques, nous avons pu définir des lignes directrices pour la conception d'un LAD dans un contexte de formation professionnelle. Il s'agirait donc :

- de prendre en compte **les particularités et les contraintes administratives liées à la formation professionnelle**. En effet, nous avons relevé via notre étude que des contraintes administratives propres à la formation professionnelle font que la réussite de l'apprenant est sujette à des critères différents en fonction du type de formation. La réussite aux formations avec évaluations obligatoires est déterminée en fonction des notes obtenues à ces évaluations, tandis que la réussite aux formations validées par le temps passé est avérée si la durée totale passée par l'apprenant est au moins égale à la durée minimale exigée par le financeur de formation. De ce fait, le besoin en termes d'indicateurs, pour le suivi de la réussite, peut changer en fonction du type de formation et des contraintes propres au contexte.
- de **permettre aux utilisateurs de personnaliser leur tableau de bord** : les contraintes précédentes mettent en évidence l'importance d'avoir une solution personnalisable que les utilisateurs finaux pourraient adapter à leurs

différents besoins de suivi pour chaque type de formation. Cette nécessité est aussi apparue au fil des réunions de concertation, chaque utilisateur exprimant finalement des attentes différentes quant aux indicateurs de suivi.

- de pouvoir **utiliser des indicateurs spécifiques pour prévenir les risques liés à la formation professionnelle** et notamment les risques psychosociaux. Ce type de risque est lié à la surcharge qu'un apprenant peut subir en travaillant tout en suivant sa formation[8]. En plus des effets sur la santé des individus, les risques psychosociaux ont un impact sur les performances et sur la réussite de l'apprenant[8]. Avoir des indicateurs dédiés à ces risques peut être utile pour comprendre les difficultés des apprenants dans la formation professionnelle[8].

4 Conclusion et Perspectives

Nous avons présenté dans cet article une étude expérimentale de six indicateurs issus de la littérature afin d'évaluer leur pertinence dans notre contexte de la formation professionnelle en ligne. Ces indicateurs ont été testés sur un échantillon de données représentatif en étudiant la corrélation entre chaque indicateur et la réussite des apprenants aux évaluations. Les résultats démontrent qu'il est très difficile de trouver un lien clair entre la réussite aux évaluations et les indicateurs étudiés. Par la suite, nous avons expliqué les raisons qui ont conduit à ces résultats et qui se résument à des particularités et à des contraintes propres à la formation professionnelle. Ceci nous a amenés à proposer des lignes directrices qui pourront être utiles pour la conception d'un LAD dans le contexte de la formation professionnelle en ligne. Dans nos travaux futurs, nous nous investirons dans la conception d'un LAD à partir d'indicateurs pertinents pour chaque type de formation et sur un mécanisme permettant de personnaliser son tableau de bord d'aide au suivi.

Références

1. Abras, C., Maloney-Krichmar, D., Preece, J. : User-centered design. Citeseer (2004)
2. Verbert, K., Govaerts, S., Duval, E., Santos, J., Assche, F., Parra, G., K. : Learning dashboards : an overview and future research opportunities. Springer-Verlag (2014)
3. Schwendimann, B., Rodriguez-Triana, M., Vozniuk, A., Prieto, L., Boroujeni, M., Holzer, A., Gillet, D., Dillenbourg, P. : Perceiving learning at a glance : A systematic literature review of learning dashboard research. In : IEEE(2017)
4. Charleer, S., Klerkx, J., Duval, E., De Laet, T., Verbert, K. : Creating effective learning analytics dashboards : Lessons learnt. In : ECTEL, Springer (2016)
5. Littlejohn, A. : Learning and Work : Professional Learning Analytics. SoLAR (2017)
6. Ruiz-Calleja, A., Prieto, L., Ley, T., Rodríguez-Triana, M. : Learning analytics for professional and workplace learning : A Literature Review. In : ECTEL (2017)
7. Littlejohn, A., Margaryan, A. : Technology-enhanced professional learning. In : Handbook of research in professional and practice-based learning, Springer (2014)
8. Marc, J. : Le recours aux formations à distance (e. learning) dans la formation professionnelle des salariés. Présentation, influence sur les acteurs et éléments de vigilance. In : Institut National de Recherche et de Sécurité (INRS) (2014)

Session
Dispositifs en ligne et hybrides

La (dé-)synchronisation des transitions dans un processus d'évaluation formative exécuté à distance : impact sur l'engagement des étudiants

Jean-François Parmentier[1] and Franck Silvestre[2]

[1] Université de Toulouse, UPS, IRES, F-31400 Toulouse
jf.parmentier@gmail.com
[2] Université Toulouse 1 Capitole,
Institut de Recherche en Informatique de Toulouse (IRIT)
118 Route de Narbonne, F-31062 Toulouse CEDEX
franck.silvestre@irit.fr

Résumé. Le processus en N Phases implanté dans la plateforme Tsaap-Notes permet une mise en oeuvre dans un contexte de cours en face à face d'une variante de l'instruction par les pairs. Or son utilisation dans un enseignement hybride, mêlant face à face et distance, a mis en évidence une limitation forte : un nombre significatif d'étudiants ne mènent pas le processus jusqu'à son terme. Cet article présente son adaptation pour une utilisation dans des enseignements hybrides ou à distances. L'adaptation principale consistant à ne plus synchroniser la transition entre les phases est inspirée de la plateforme DALITE. Les résultats de l'expérience menée pour mettre à l'épreuve ce nouveau processus démontrent clairement l'augmentation du nombre d'étudiants réalisant les N phases de bout en bout.

Mots-clé: Instruction par les pairs · expérimentation · pédagogie active · évaluation formative · évaluation par les pairs· enseignement à distance · enseignement hybride · analyse des données d'apprentissage · Tsaap-Notes · DALITE

Abstract. The N Phases process implemented in the Tsaap-Notes platform provides an implementation in a face-to-face course context of a variant of peer instruction. However, its use in hybrid teaching, combining face-to-face and distance, has highlighted a strong limitation : a significant number of students do not complete the process. This article presents its adaptation for use in hybrid or distance learning. The main adaptation consisting in no longer synchronizing the transition between phases is inspired by the DALITE platform. The results of the experiment conducted to test this new process clearly demonstrate the increase in the number of students performing the N phases from start to finish.

Keywords: Peer Instruction · experiment · active learning · formative assessment · peer assessment· distance learning · blended learning · learning analytics · Tsaap-Notes · DALITE

1 Introduction

Parmi les différentes pratiques de pédagogies actives, le processus d'évaluation formative baptisé « Instruction par les Pairs » (IP) concerne les enseignements dispensés en face à face dans des amphithéâtres pouvant accueillir plusieurs centaines d'étudiants. L'IP a été introduit par Mazur pour la première fois en 1991 durant des cours de physique pour des étudiants de licence [8]. L'approche consiste, pour une même question fermée (principalement des questions à choix exclusif), à sonder les étudiants deux fois : une première fois après une phase de réflexion individuelle, puis une deuxième fois après que les étudiants aient débattu avec leurs voisins sur le bien-fondé de la réponse qu'ils ont fournie à l'issue du premier sondage. Les résultats du deuxième sondage révèlent alors un plus grand nombre de bonnes réponses. Différentes expérimentations ont démontré les bénéfices de cette approche en termes d'apprentissage par rapport à une approche classique purement transmissive [8,6,14]. Devant les résultats probants apportés par l'IP tant sur l'engagement des étudiants que sur leurs résultats d'apprentissage, beaucoup d'universités se sont lancées dans l'utilisation des systèmes de votes individuels (tels des boitiers), en appliquant ou non l'IP [7,10,9].

La plateforme Web Tsaap-Notes [12] permet l'orchestration de séquences d'évaluation formative structurées de manière similaire à celles de l'IP. L'application, conçue pour une utilisation durant les cours en face à face, introduit l'argumentation écrite par les étudiants, la confrontation de points de vue anonymisée et l'évaluation par les pairs. Le processus supporté par Tsaap-Notes, a été baptisé processus en N Phases [13] ; les expérimentations menées en présentiel dans différents cours ont montré des résultats semblables à ceux obtenus en instruction par les pairs traditionnelle : fort taux d'engagement des étudiants et améliorations des bonnes réponses à l'issue du deuxième vote [12].

Parmi les enseignants utilisateurs de Tsaap-Notes, certains ont souhaité exploiter l'outil en dehors de leur enseignement en présentiel. Les premiers retours informels convergeaient vers le constat suivant : un nombre significatif d'étudiants ne prenaient pas part à la deuxième phase du processus, celle d'évaluation par les pairs. En conséquence, le nombre d'évaluations des arguments écrits était faible sur la plupart des contributions.

Ces retours nous ont conduits à nous interroger sur la pertinence de l'utilisation du processus en N Phases à distance et à poser la question de recherche suivante : sous quelle(s) condition(s) le processus N Phases conserve-t-il son caractère engageant dans un contexte d'activité menée en dehors du temps de classe ?

L'étude de l'application DALITE, conçue pour la mise en œuvre à distance du processus d'IP, nous a conduits à émettre deux hypothèses :

1. l'implantation d'un processus à transitions non synchronisées, tel celui existant dans DALITE permettant aux étudiants de participer aux différentes phases en une seule connexion à l'application, réduira de manière significative le taux de perte d'étudiants entre la première et la deuxième phase ;

2. l'augmentation du nombre d'étudiants à la deuxième phase permettra, quasi mécaniquement, d'augmenter le nombre d'évaluations de chaque contribution.

Afin de valider ces hypothèses, nous avons adapté le processus en N Phase et modifié Tsaap-Notes en conséquence puis mis en place un dispositif expérimental permettant de comparer les résultats obtenus entre l'utilisation de la version initiale du processus et l'utilisation de la version adaptée.

L'article est composé de cinq parties. Une revue bibliographique des logiciels permettant d'orchestrer les processus d'instruction par les pairs et dérivés est présentée dans la section 2. Le processus en N Phases adapté pour une utilisation hors classe est présenté dans la section 3. L'expérimentation et les résultats obtenus pour mettre à l'épreuve nos hypothèses sont détaillés respectivement dans les sections 4 et 5. Enfin, la section 6 conclut l'article en présentant les perspectives ouvertes par les résultats obtenus.

2 Logiciels pour l'orchestration des processus IP et dérivés

Une revue bibliographique a permis d'identifier l'existence de deux logiciels permettant l'orchestration des processus d'IP et dérivés : Tsaap-Notes et DALITE. Les deux plateformes sont présentées de manière détaillée avant d'être comparées.

2.1 Tsaap-Notes

La plateforme Tsaap-Notes [11] a été conçue pour une utilisation pédagogique en face à face. Son objectif initial est de proposer une amélioration de l'instruction par les pairs sur deux points :

— L'alignement pédagogique : dans sa forme initiale, l'instruction par les pairs utilise le débat oral. Les étudiants ne sont alors pas invités à formaliser par écrit leur raisonnement. Or, durant les évaluations sommatives, il est très fréquemment demandé aux étudiants d'expliquer par écrit les résultats qu'ils obtiennent. L'approche IP n'est donc pas optimale en termes d'alignement constructiviste [3] : l'activité proposée durant le cours pour atteindre l'objectif d'apprentissage n'est pas de même nature que celle proposée pour évaluer de manière sommative le niveau d'atteinte de l'objectif d'apprentissage.

— La connaissance des raisonnements des étudiants : lors de l'instruction par les pairs à l'aide de dispositifs de vote simples, l'enseignant ne dispose comme information que de l'histogramme des réponses. Or ce qui est attendu des étudiants n'est pas seulement de voter pour la bonne réponse, mais aussi d'être capable de justifier ce choix. Ainsi connaitre les raisonnements des étudiants qui les ont conduits aux différents choix de réponses est crucial afin de pouvoir les corriger. En particulier, il est essentiel de détecter les faux positifs, c'est-à-dire les raisonnements faux

qui conduisent au bon choix de réponse, pour que les étudiants concernés ne repartent pas avec une fausse information sur la validation de leurs acquis. En général, dans la mise en œuvre de l'IP, l'enseignant tente d'obtenir cette connaissance en se déplaçant aléatoirement dans la classe en en écoutant les débats entre élèves, ou en demandant à quelques étudiants d'expliquer publiquement leurs raisonnements. Or ces pratiques ont à la fois délicates à mener et peu représentatives.

Le processus en N Phases, dans sa forme générique, se déroule en 3 phases successives déclenchées par l'enseignant. Les transitions entre les phases sont donc synchronisées pour tous les étudiants. Les trois phases se déclinent de la manière suivante :

1. Chaque étudiant prend connaissance individuellement de la question et des choix proposés, puis effectue un choix et écrit une argumentation.
2. On propose alors à chaque étudiant d'évaluer jusqu'à trois explications proposées par ses pairs. Cette évaluation s'effectue en donnant une note de 1 à 5 (sous forme d'étoiles). Ils peuvent ensuite changer leur choix de réponse s'ils ont été convaincus par une des explications alternatives proposées (Fig. 1) [13].
3. La dernière phase propose à la fois une synthèse des votes sous forme d'histogramme, ainsi qu'un affichage des justifications les mieux notées. L'enseignant prend alors connaissance de quelques explications parmi les mieux notées et effectue des commentaires sur leur pertinence.

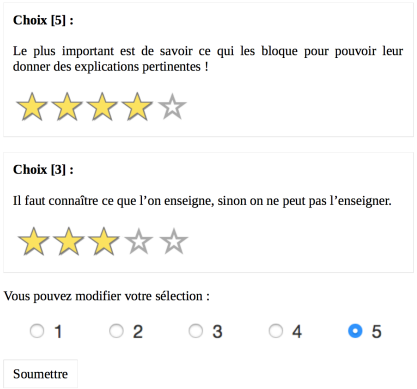


Figure 1. Phase 2 de Tsaap-Notes.

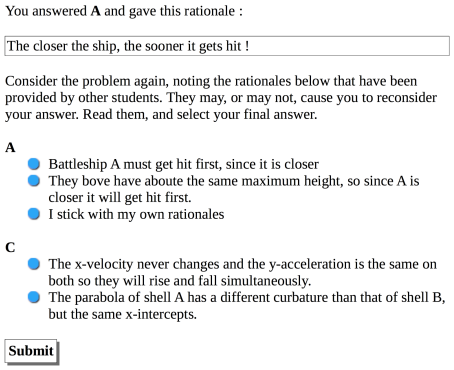


Figure 2. Phase 2 de DALITE.

Dans la pratique, il est fréquent d'observer un raisonnement incorrect parmi les explications les mieux notées qui correspondent cependant au choix de la réponse correcte. Ainsi le procédé permet de détecter des faux positifs. De plus, l'enseignant peut consulter les explications les mieux notées correspondant à des choix de réponses incorrectes. La phase d'évaluation par les pairs donne

accès à l'enseignant à la connaissance des raisonnements « séduisants », donc représentatifs des conceptions erronées, sans que celui-ci n'ait besoin de lire l'ensemble des contributions écrites.

Les expérimentations menées dans différents cours en face à face ont montré des résultats semblables à ceux obtenus avec l'IP : fort taux d'engagement des étudiants et améliorations des bonnes réponses à l'issue du deuxième vote [12].

Ce processus a été étendu à l'usage de questions ouvertes. Dans ce cas, l'étudiant n'a pas de choix à sélectionner et ne fait qu'indiquer son raisonnement. La dernière phase n'affiche pas d'histogrammes. La validation s'effectue par les commentaires oraux de l'enseignant sur les différentes réponses qui ont été les mieux évaluées.

2.2 DALITE

DALITE est une application Web conçue pour exécuter à distance le processus d'Instruction par les Pairs [5]. Elle est donc destinée à un usage hors classe, soit en complément d'un enseignement présentiel [4,5], soit dans le cadre d'une formation 100% en ligne comme un MOOC [2].

La phase de débat, initialement réalisée à l'oral, a été adaptée à l'écrit. La plateforme a donc été conçue pour inciter les étudiants à s'expliquer par écrit à travers des justifications de QCM, à comparer ces justifications à celles de leurs pairs et à réfléchir à la qualité des justifications proposées. Dans sa version la plus récente [2], répondre à une question pour un étudiant se déroule en 3 phases successives :

1. Visualisation de la question, des choix, et écriture d'une argumentation.
2. Pour deux des choix de réponses, on propose à l'étudiant 3 ou 4 justifications écrites par d'autres étudiants. Ils doivent alors lire les explications puis sélectionner le choix et l'explication qu'ils pensent être les meilleurs. Ils peuvent décider de conserver leur propre explication (Fig. 2). L'un des deux choix proposés correspond toujours à la réponse correcte [4]. L'autre choix correspond soit à la réponse qu'a choisie l'étudiant (dans le cas où celui-ci a répondu incorrectement), soit à une réponse sélectionnée en amont par l'enseignant et correspondant habituellement à la réponse incorrecte la plus fréquente.
3. La dernière phase leur résume ce qui vient de se passer : leur choix et explication précédents, leur nouveau choix et leur nouvelle explication ainsi que l'explication donnée par un expert, mais sans leur indiquer à quel choix cela correspond [4,5,1].

De manière alternative, la phase 2 peut aussi se dérouler de manière séquentielle en proposant des choix de réponses un par un, et en demandant à l'étudiant de voter « pouce levé » ou « pouce baissé » pour chacun des choix [2].

Ainsi, les élèves participent à toutes les phases du processus dès leur première connexion à l'application. Les transitions entre les phases ne sont donc pas synchronisées d'un étudiant à l'autre. L'enseignant peut alors leur demander de compléter les activités en dehors des cours (en complément d'un enseignement

présentiel), ou au cours d'une formation en ligne, dans le cadre d'un MOOC par exemple.

Lors de la création d'une question, l'enseignant doit fournir [1] :

— la question sous forme QCM en indiquant la bonne réponse,
— une explication justifiant ce choix,
— le « deuxième meilleur choix » qui sera proposé aux étudiants ayant voté pour la bonne réponse.

L'outil a été utilisé dans différents cours de mécanique [5,1] pour des étudiants âgés de 17 à 19 ans. Les analyses ont montré qu'utiliser DALITE comme outil supplémentaire à un cours traditionnel conduit à une amélioration des gains conceptuels mesurés par un test standardisé de mécanique [4,5]. Les gains sont comparables à l'utilisation de l'instruction par les pairs en classe. Ainsi, DALITE semble pouvoir être utilisé comme une alternative au Peer Instruction en classe.

2.3 Comparaison des processus

Tsaap-Notes et DALITE proposent des fonctionnalités globalement similaires : réponse à des questions QCM sous forme d'écrit (phase 1 identique), confrontation de point de vue sous forme de visualisation des choix et argumentations des autres participants, changement de vote possible, et affichage final des histogrammes. Il y a cependant des différences significatives :

— **Le caractère synchronisé des transitions entre les phases :** DALITE fonctionne en transitions de phases non synchronisées : lorsqu'un étudiant se connecte, il effectue immédiatement les différentes phases indépendamment de ce que font les autres étudiants. Au contraire, Tsaap-Notes fonctionne en transition de phases synchronisées : les différentes phases sont déclenchées l'une après l'autre par l'enseignant.
— **Sélection vs Évaluations des contributions :** dans DALITE l'étudiant sélectionne 1 explication (qui peut-être la sienne) qui correspond à son second vote. Dans Tsaap-Notes l'étudiant doit évaluer chacune des contributions qu'on lui présente et peut changer ou non son vote en conséquence.
— **Informations pour l'enseignant :** dans DALITE, seuls les histogrammes des votes sont affichés à l'enseignant. Tsaap-Notes propose de plus l'affichage des explications les mieux notées.

Les différences constatées s'expliquent assez naturellement en comparant les objectifs de conception initiaux. L'utilisation de Tsaap-Notes en cours en face à face nécessite une maitrise du temps passé sur chaque phase et donc une gestion synchronisée des cadences des étudiants. L'objectif d'échanges entre enseignant et étudiants sur les argumentations écrites explique une présentation des résultats incorporant les contributions des étudiants avec leur note moyenne issue de l'évaluation par les pairs, là où DALITE s'intéresse d'abord au changement d'avis entre les deux votes.

La non-synchronisation des transitions proposée par DALITE a directement inspiré la conception du processus en N Phases adapté à une utilisation à distance. Les principales caractéristiques du nouveau processus sont présentées dans la section suivante.

3 Processus en N Phases à transitions non synchronisées

3.1 Principes généraux

Le nouveau processus permet à l'étudiant de réaliser l'ensemble de ses tâches lors de sa première visite sur l'application. L'étudiant est donc invité à répondre à la question posée en fournissant sa contribution textuelle ; après soumission de sa réponse, le système lui propose immédiatement d'évaluer un lot de contributions. L'étudiant peut modifier sa réponse si la confrontation de point de vue induite l'y pousse. À l'issue de cette étape, le processus se termine de deux manières différentes selon que l'activité se déroule dans un contexte hybride ou entièrement à distance.

Dans le contexte d'un enseignement délivré entièrement à distance, l'étudiant accède immédiatement au feedback proposé par la plateforme : son score à la question, la réponse et l'explication attendues, la note moyenne obtenue par sa contribution, les argumentations des étudiants les mieux notées ayant répondu correctement. Dans le contexte d'une utilisation hybride, l'étudiant accèdera au feedback sur son travail après que l'enseignant ait décidé de publier les résultats sur la question. Cette publication est en général effectuée pendant un cours en présentiel.

3.2 Garantir l'accroissement du nombre d'évaluateurs

Le processus présenté dans la section 2.1 dispose de l'ensemble des contributions des étudiants pour effectuer l'affectation entre les contributions et les étudiants évaluateurs. Dans ce cas de figure, il est possible de générer le plan d'affectation à partir d'une analyse de l'ensemble des réponses. Un des premiers algorithmes utilisés par Tsaap-Notes pour réaliser le plan d'affectation est décrit de manière détaillée dans [13]. Dans le cas de notre nouveau processus, les affectations doivent être réalisées au fil de l'eau à partir des contributions disponibles à l'instant où l'étudiant accède à la plateforme. L'algorithme utilisé consiste à proposer à l'étudiant courant un lot d'au plus 5 contributions alternatives correspondant aux contributions qui ont été les moins évaluées par les pairs à cet instant. Cette approche garantit qu'à l'issue de l'activité, chaque contribution a été évaluée par un nombre maximum d'évaluateurs possible.

Cependant, l'algorithme pose un problème de démarrage à froid, car il ne fonctionne qu'en présence de contributions fournies par les pairs. Or, le ou les premiers étudiants connectés à la plateforme ne disposent pas de telles contributions. La solution imaginée pour régler ce problème est le suivant : l'enseignant, quand il prépare sa question, dispose d'une interface lui permettant de fournir la réponse attendue à la question ainsi qu'autant de réponses factices imaginées pour cette question qu'il le souhaite. Ces réponses fournies par l'enseignant sont alors utilisées par l'algorithme d'affectation pour le ou les premiers étudiants se connectant à la plateforme tant que des contributions issues des pairs ne sont pas encore disponibles.

4 Dispositif expérimental

Afin de mettre à l'épreuve les deux hypothèses formulées dans l'introduction, une expérimentation a été menée sur deux ans. La première année proposait à des étudiants d'utiliser Tsaap-Notes dans sa version initiale. La deuxième année, une version incluant le processus à transitions non synchronisées a été testée sur un autre groupe d'étudiants aux caractéristiques similaires.

La comparaison des deux processus a été effectuée en analysant les données d'apprentissage recueillies au regard de deux indicateurs :

1. Combien d'étudiants participent à la phase d'évaluation des contributions ? L'objectif étant de maximiser la confrontation de points de vue afin d'améliorer l'apprentissage des étudiants.
2. Par combien d'étudiants est évaluée chacune des explications ? L'objectif étant de garantir la pertinence des contributions qui sont sélectionnées pour permettre à l'enseignant de remédier aux conceptions erronées.

4.1 Population et organisation

Les deux expérimentations ont eu lieu au cours d'un enseignement d'initiation à l'algorithmique à destination des étudiants de L1 à l'Université Paul Sabatier, en 2016-2017 (groupe 1) puis en 2017-2018 (groupe 2), aux premiers semestres universitaires.

Les thématiques abordées et l'organisation du cours étaient similaires. Chaque semaine était composée d'un cours en amphithéâtre (CM) en classe entière d'une durée de 1h ainsi que de deux heures de travaux pratiques sur machines en petits groupes (quart de classe). Le nombre d'étudiants présents aux cours magistraux était similaire entre les deux années (en moyennes 73 et 69).

L'année 2016-2017, le cours en amphithéâtre avait lieu chaque vendredi et les étudiants devaient ensuite effectuer un Devoir Maison (DM) utilisant Tsaap-Notes. Avant le lundi soir, ils devaient répondre à trois questions ouvertes (phase 1 de contribution), puis ils devaient, entre le mardi et le prochain cours, évaluer les contributions de trois autres étudiants pour chacune des questions (phase 2 d'évaluation). Enfin, au début de chaque cours suivant, l'enseignant faisait des commentaires oraux sur certaines réponses des étudiants (phase 3 de présentation des résultats). Ce cycle a eu lieu 3 fois, conduisant donc à 3 DM.

Les mêmes questions ont été posées durant l'année 2017-2018. Les cadences étant cette fois-ci non synchronisées, ils avaient toute la semaine pour compléter leurs réponses et évaluer immédiatement après 3 autres contributions pour chacune des questions. Au début du cours suivant, de manière similaire à l'année précédente, l'enseignant effectuait des commentaires oraux sur quelques contributions et quelques évaluations.

Les questions posées portaient sur des aspects métacognitifs (réflexion sur leur propre apprentissage). Une question type et une réponse typique d'étudiant sont présentées dans le tableau 1.

Table 1. Exemple de question et de réponse d'étudiant.

Question	Quel était le point le plus important pour vous dans le cours de vendredi dernier (thème du 18 novembre) ? Expliquez pourquoi. Soyez précis et détaillé.
Réponse typique	"Le plus important est la notion d'optimisation des programmes pour qu'ils renvoient un résultat en moins de temps possible et avec un cout en mémoire plus faible. En cas de boucle, chercher à faire le moins de tours possible."

Deux types de données ont été analysées : le nombre d'étudiants complétant chacune des phases (écriture puis évaluation) pour chacune des questions de chacun des DM, puis le nombre moyen d'évaluations reçues par une contribution.

5 Résultats

Un nombre important d'étudiants hors de la population cible ayant effectué le DM1 de l'année 2016-2017, l'analyse des données s'est effectuée uniquement sur les DM 2 et 3 pour chacune des années. Au total 6 questions ont été posées. En moyenne 53 étudiants ont répondu à chacune des questions pour l'année 2016 et 51 pour l'année 2017, conduisant à un total de 624 contributions écrites.

La figure 3 présente les résultats pour chacun des DM. Les valeurs indiquées représentent la moyenne des contributions sur les 3 questions du DM correspondant. Par exemple, pour le DM2 de l'année 2016-2017, 60 étudiants ont répondu à la question 1, 59 à la question 2 et 64 à la question 3, conduisant à une moyenne de 61. Les valeurs pour chacune des questions étant relativement similaires au sein d'un même DM, le choix a été fait d'effectuer des moyennes par DM. Au contraire, on constate une diminution du nombre de participants entre le DM2 et le DM3, pour chacune des années. Celle-ci est de 25% pour l'année 2016 (interval de confiance à 95% : [14%, 37%]) et de 23% pour l'année 2017 ([13%, 36%]), l'écart entre les deux résultats étant non statistiquement significatif (p-value = 1 avec un test exact de Fisher). La proportion de participants aux DM par rapport au nombre moyen d'étudiants en cours est aussi similaire entre les deux années (83% pour le DM2 et 64% pour le DM3). Ces deux résultats confirment l'hypothèse d'une population similaire entre les deux années.

Une différence majeure entre les deux années est cependant observée dans le nombre de participants à la phase d'évaluation des contributions. Ainsi, on observe une perte moyenne de 21% (interval de confiance à 95% : [14%, 30%]) entre les deux phases durant l'année 2016 et une perte nulle ([0%, 4%]) durant l'année 2017. Cette différence est statistiquement très significative ($p < 8 \cdot 10^{-8}$) (calcul sur le tableau 2). Les deux populations étant supposées identiques, cette différence de comportement est attribuée au résultat de la modification du processus.

Durant l'année 2016-2017, des interviews ponctuelles et informelles semblent montrer une acceptation et un intérêt de la part des étudiants à évaluer le travail

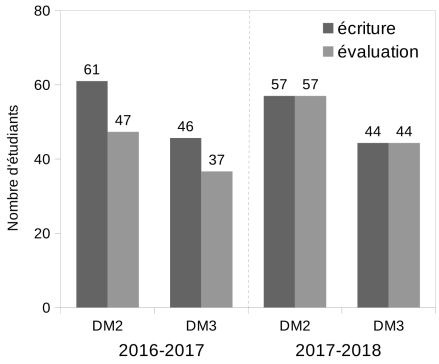


Figure 3. Nombre d'étudiants ayant participé aux phases d'écriture et d'évaluation pour chacun des DM. Les valeurs sont les moyennes sur les trois questions du DM.

Table 2. Nombre moyen d'étudiants participants aux différentes phases. Les participants sont ajoutés entre les différents DM (ex : 101 = 57 + 44).

	écriture et évaluation	**écriture seule**
2016-2017	84	23
2017-2018	101	0

des autres étudiants. Cependant les étudiants déclarent avoir des difficultés à organiser leur travail au cours de la semaine, expliquant un oubli régulier de participer à la phase d'évaluation. Ces déclarations semblent confirmer que la différence de comportement entre les deux années s'explique majoritairement par le changement de processus.

La perte d'étudiants observée en 2016-2017 diminue l'impact pédagogique du dispositif : moins d'étudiants participent à la phase d'évaluation, diminuant ainsi la qualité de l'évaluation moyenne de chaque contribution. En théorie, comme chaque élève doit évaluer 3 contributions, chacune de celles-ci est évaluée 3 fois, la moyenne obtenue donnant un indicateur relativement correct de la qualité de la contribution. Or en raison de la perte d'étudiants entre les deux phases, le nombre d'évaluateurs est plus faible que le nombre de contributeurs, conduisant à une moyenne de 2.2 évaluations par contribution durant l'année 2016-2017. Celle-ci est à 2.9 pour l'année 2017-2018. La moyenne du nombre d'évaluations est donc en nette augmentation et converge vers la moyenne attendue de 3. Le résultat est inférieur à 3 en raison de la présence des fausses contributions créées par l'enseignant et du nombre plus faible d'évaluations des 3 dernières contributions étudiantes. Les résultats sont synthétisés dans le tableau 3.

Table 3. Synthèse des données et des résultats

Année universitaire	2016-2017	2017-2018
Présence moyenne en cours d'amphi	73	69
Transitions	synchronisées	dissociées
Participation moyenne (écriture)	53	51
Participation moyenne (évaluation)	42	51
Perte moyenne	21 %	0 %
Nombre moyen d'évaluations par contribution	2.3	2.9

6 Conclusion et perspective

Le processus d'instruction par les pairs et ses dérivés assistés par la technologie sont reconnus pour les bénéfices qu'ils apportent aux étudiants durant leur apprentissage. En particulier, le processus en N Phases implanté dans Tsaap-Notes, favorise l'alignement entre les activités proposées en cours et celles proposées lors des évaluations sommatives et permet à l'enseignant d'accéder aux raisonnements des étudiants. Cependant, le processus en N Phases, comme l'IP, a été conçu pour une mise en œuvre durant les cours en face à face. Lors des premières expérimentations portant sur l'utilisation de l'application Tsaap-Notes dans un contexte d'enseignement à distance, l'analyse des données d'apprentissage a révélé deux dysfonctionnements : une perte non négligeable de participants à l'activité entre la première et la deuxième phase du processus ainsi qu'un nombre faible d'évaluateurs par contributions à l'issue de la phase de confrontation de point de vue et d'évaluation par les pairs. Ces résultats nous ont incités à concevoir et à implanter dans Tsaap-Notes une adaptation du processus en N Phases intégrant la non-synchronisation des transitions entre les phases telle qu'elle est implantée dans la plateforme DALITE. L'analyse des données d'apprentissage à l'issue des expérimentations faisant usage de la nouvelle version du processus, a mis en évidence la disparition de la perte de participants entre la première et la deuxième phase et l'augmentation significative du nombre moyen d'évaluateurs par contribution.

Bien que le nouveau processus apporte une réponse satisfaisante pour l'utilisation de Tsaap-Notes dans un contexte d'enseignement à distance, de nouvelles problématiques émergent. En effet, afin d'augmenter le nombre d'évaluateurs pour chaque contribution, l'algorithme initial utilisé par Tsaap-Notes a été modifié pour maximiser le nombre d'évaluateurs par contribution. En contrepartie, aucun autre critère n'est utilisé pour effectuer l'association entre contributions et évaluateurs. À titre d'exemple, la version initiale de l'algorithme s'appuyait sur l'indice de confiance des réponses pour planifier les appariements. Par ailleurs, le processus ne règle pas le problème de la "coupure à chaud" : à ce jour, le processus ne permet pas d'évaluer la contribution du dernier étudiant participant à l'activité. Ces problèmes guident aujourd'hui notre réflexion sur la conception de nouveaux processus.

Références

1. Bhatnagar, S., Desmarais, M., Whittaker, C., Lasry, N., Dugdale, M., Charles, E.S. : An analysis of peer-submitted and peer-reviewed answer rationales, in an asynchronous Peer Instruction based learning environment. In : International Educational Data Mining. Madrid, Spain (Jun 2015)
2. Bhatnagar, S., Lasry, N., Desmarais, M., Charles, E. : DALITE : Asynchronous Peer Instruction for MOOCs. In : European Conference on Technology Enhanced Learning. Lyon, France (Sep 2016)
3. Biggs, J. : Enhancing teaching through constructive alignment. Higher education **32**(3), 347–364 (1996)
4. Charles, E.S., Lasry, N., Lenton, K., Whittaker, C., Dugdale, M., Bhatnagar, S. : Les réseaux conceptuels collectifs en enseignement et en apprentissage : l'usage de TIC pour relier la science scolaire avec la réalité extérieure (2014)
5. Charles, E.S., Lasry, N., Whittaker, C., Dugdale, M., Lenton, K., Bhatnagar, S., Guillemette, J. : Beyond and within classroom walls : Designing principled pedagogical tools for students and faculty uptake. In : Computer Supported Collaborative Learning (2015)
6. Crouch, C.H., Mazur, E. : Peer instruction : Ten years of experience and results. American journal of physics **69**(9), 970–977 (2001)
7. Fagen, A.P., Crouch, C.H., Mazur, E. : Peer Instruction : Results from a Range of Classrooms. The Physics Teacher **40**(4), 206–209 (Apr 2002). https://doi.org/10.1119/1.1474140
8. Mazur, E. : Peer instruction (1997)
9. Parmentier, J.F., Lamine, B., Bonnafé, S. : Changer les conceptions en mécanique des étudiants en L1 à l'Université. In : M01 Mini symp. Formation et pédagogie. AFM, Association Française de Mécanique, Lyon (2015)
10. Rudolph, A.L., Lamine, B., Joyce, M., Vignolles, H., Consiglio, D. : Introduction of interactive learning into French university physics classrooms. Physical Review Special Topics - Physics Education Research **10**(1), 010103 (Jan 2014). https://doi.org/10.1103/PhysRevSTPER.10.010103
11. Silvestre, F. : Conception et mise en oeuvre d'un système d'évaluation formative pour les cours en face à face dans l'enseignement supérieur. Thèse de doctorat, Université Paul Sabatier, Toulouse, France (novembre 2015), (Soutenance le 25/11/2015)
12. Silvestre, F., Vidal, P., Broisin, J. : Reflexive learning, socio-cognitive conflict and peer-assessment to improve the quality of feedbacks in online tests. In : Design for Teaching and Learning in a Networked World, pp. 339–351. Springer (2015)
13. Silvestre, F., Vidal, P., Broisin, J. : Un nouveau processus d'évaluation pour améliorer la qualité des feedbacks dans les tests en ligne. Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation **24**(spécial), pp–181 (2016)
14. Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N., Su, T.T. : Why Peer Discussion Improves Student Performance on In-Class Concept Questions. Science **323**(5910), 122–124 (Jan 2009). https://doi.org/10.1126/science.1165919

Former à la classe inversée en contexte de visioconférence : retours sur deux itérations d'une recherche orientée par la conception

Vincent Barré[1], Florian Meyer[2], Matthieu Petit[2], Nathalie Lefebvre[2], Cécile Gandon[3]

[1] LIUM – EA 4023, Le Mans Université, 72 085 Le Mans Cedex 9, France
vincent.barre@univ-lemans.fr

[2] Université de Sherbrooke, 2500 boul. de l'Université, Sherbrooke, QC J1K 2R1, Canada
{florian.meyer, matthieu.petit, nathalie.lefebvre}@usherbrooke.ca

[3] Université de Bretagne Sud, 27 rue A. Guillemot, 56000 Lorient, France
cecile.gandon@univ-ubs.fr

Résumé. Dans notre projet de recherche collaborative orientée par la conception, une formation sur la classe inversée en contexte de visioconférence a été donnée à deux reprises à des enseignants et conseillers pédagogiques. Le design pédagogique de cette formation se structurait dans un cadre misant sur la création d'une communauté d'apprentissage, la présence à distance et le vécu en tant qu'apprenant d'une expérience de classe inversée en contexte de visioconférence. Suite à la première itération, des constats empiriques, tant du point de vue de l'apprentissage, que de l'activité des formateurs ou des caractéristiques technopédagogiques mises en œuvre, ont motivé l'élaboration d'un premier travail de recherche visant à décrire la présence suscitée au sein de cette communauté, selon le modèle de la présence à distance [11]. Ce travail a notamment permis de mettre en évidence une certaine carence en matière de présence sociocognitive et de soulever des questions sur la présence pédagogique.

Mots-clés. Recherche collaborative orientée par la conception, Téléprésence, Classe inversée, Présence à distance, Analyse de forum

Abstract. As part of a collaborative design-based research project, flipped classroom training in a videoconferencing context was given twice to teachers. The pedagogical design of this training was structured in a framework based on the creation of a learning community, the learning presence and the real experience as a learner of a flipped classroom experience in a videoconferencing context. Following the first iteration, empirical findings, from the point of view of learning, as well as the activity of the instructors or the technopedagogical features implemented, led to the development of a first research project aimed at describing the presence created within this community, according to the learning presence model [11]. This work helped to highlight a certain lack of socio-cognitive presence and to raise questions about the pedagogical presence.

Keywords. Design-based research, Telepresence, Flipped classroom, Learning presence, Chat analysis

1 Introduction

Dans le cadre du projet TOPIC (*Téléprésence comme OPportunité pour l'Innovation et la Conception pédagogiques*), notre équipe franco-québécoise de chercheurs et conseillers pédagogiques s'est engagée dans une démarche de recherche collaborative orientée par la conception pédagogique (RCOC) [1], [2] autour d'un double besoin de (i) comprendre comment faire évoluer des dispositifs de formation initiale et continue impliquant visioconférence et technologies liées à la téléprésence et (ii) construire un répertoire de connaissances rigoureuses et critiques autour des approches pédagogiques adaptées, des effets sur l'apprentissage des étudiants, ou encore des conditions facilitant et enrichissant le travail enseignant dans des contextes de formation exploitant de telles infrastructures. Dans notre contexte, les technologies liées à la téléprésence participent à ce que les sens des utilisateurs soient soumis à des stimuli qui donnent l'impression d'être sur un (ou plusieurs) site distant via une haute qualité d'image et de son.

Peu de travaux abordent ces aspects, en particulier sur les approches et stratégies pédagogiques favorisant l'apprentissage [3], [4], [5], [6], [7] dans un tel contexte, alors que les enseignants, les formateurs et les chercheurs ont besoin de clés de compréhension et d'outils permettant d'y faire évoluer leurs dispositifs de formation[7].

C'est dans une démarche collaborative impliquant deux conseillères pédagogiques et deux chercheurs (ci-après « les formateurs ») que nous avons conçu, implanté, vécu et analysé la formation « *Inverser sa classe en contexte de visioconférence* ». La dimension collaborative de cette RCOC s'est nourrie de la complémentarité des profils des formateurs et de leur expérience en matière de classe inversée (CI), et s'est déployée à toutes les phases du projet : conception, implémentation, déroulement et analyse de la formation, même si l'implication de tous n'a pas été la même sur chaque phase.

Considérant le praticien en formation comme étant un acteur engagé et réflexif [8], les participants ont été placés en situation de classe inversée, les amenant à porter un regard réflexif sur celle-ci du point de vue de l'apprenant. La formation a été jouée deux fois, déroulant deux itérations du cycle de conception / recherche [1], où participants et formateurs étaient situés sur trois campus au Québec et en France. Il convenait donc de favoriser la création d'une communauté d'apprentissage en ligne [9], [10] au sein de laquelle les participants auraient l'occasion d'interagir librement et aisément afin de confronter leurs réflexions, leurs compréhensions et leurs solutions au regard des cibles d'apprentissage visées. Assurer une présence à distance [11], [12] suffisante, en particulier lors des moments asynchrones de la formation, était dès lors crucial [7].

Les données collectées et le ressenti des formateurs apportent de premières pistes de réponses à la question de recherche consistant à savoir en quoi les interactions durant la formation ont favorisé, ou non, une expérience éducative riche. Ont-elles permis, en particulier, une présence à distance [11] suffisante ? Et si oui, selon quelles dimensions ? Ce questionnement visait à la fois la réingénierie de situations de formations en contexte de téléprésence, et à comprendre comment mieux les concevoir.

Afin de respecter la logique de RCOC, cet article présente le modèle exploité pour concevoir la formation, son cadre théorique et son organisation. Sur cette base, deux itérations vécues sont présentées : leur design technopédagogique, une problématique émergente, un cadre théorique et méthodologique, les premiers résultats et leur interprétation. Une discussion et les perspectives de recherche concluent l'article.

2 Conception de la situation d'apprentissage

L'entrée principale d'un processus de RCOC passe par la conception de la situation d'apprentissage étudiée [2]. Dans notre cas, les formateurs l'ont conçue à l'aide de la méthode d'ingénierie pédagogique MISA [13] et l'ont offerte au printemps 2017 [14], puis à l'automne 2017. Une troisième itération s'est déroulée au printemps 2019.

Conçue pour des enseignants / formateurs souhaitant utiliser une approche de CI dans un contexte où les temps « présentiels » se déroulent au sein d'un réseau de salles connectées en visioconférence, la formation visait à développer les éléments de compétences suivants :

1. Planifier une situation de CI tenant compte des caractéristiques, conditions et contraintes d'une situation d'enseignement / apprentissage en salle de visio ;
2. Scénariser le déroulement d'activités d'une situation d'enseignement / apprentissage de type CI en salle de visio ;
3. Identifier les ressources nécessaires à produire pour une situation d'enseignement / apprentissage de type CI en exploitant des outils adaptés ;
4. Identifier l'échéancier pour leur projet et l'appropriation des ressources technopédagogiques nécessaires à la production du matériel pédagogique.

Les objectifs de la formation visaient l'acquisition de connaissances et d'habiletés sur le plan de la scénarisation d'une unité d'apprentissage en CI. Pour que ce soit réellement significatif pour les participants [21], il leur était proposé de se mettre en action et de scénariser une séquence d'enseignement dont le choix devait être relié à leur contexte d'enseignement et leur servir ultérieurement dans le cadre d'un cours.

Une des thématiques de la formation étant la CI, il était important pour les formateurs de s'accorder sur une définition commune : « *la classe inversée est fondée sur l'idée d'une double topicité (en classe et hors classe) [15], [16], celle-ci articulant un ensemble d'activités d'apprentissage impliquant une redistribution des rôles au sein de ces espaces d'apprentissage* » [17] et que donc « *la classe inversée doit être considérée comme un dispositif hybride de formation qui combine le mode présentiel et le mode virtuel asynchrone* » [17]. Par conséquent, afin de permettre aux participants d'appréhender cette complexité dans un temps contraint, l'équipe a fait le choix de les placer eux-mêmes dans un dispositif de type CI, leur permettant d'en découvrir les aspects importants et de développer des compétences professionnelles appropriées, tout en ayant l'occasion de vivre l'expérience [18] du point de vue de l'apprenant.

Le modèle de conception de la formation est détaillé dans [14]. Elle « *s'est structurée selon un type de CI que Lebrun [19] définit comme une combinaison d'activités de préparation et d'activités de découverte et d'appropriation à la maison et d'activités de réalisation et d'échange en classe. Ce type de CI (type 3) est, selon Lebrun [19], une hybridation de deux types qu'il qualifie de prototypiques : le type 1 où les apprenants sont en posture plutôt passive, écoutent la leçon à la maison et réalisent en classe les exercices conçus par l'enseignant et le type 2 où les élèves, à la maison, collectent les informations et développent les activités à réaliser en classe afin de construire les savoirs qu'ils ont identifiés.* »

Notre intention étant de permettre aux apprenants de porter un regard réflexif sur la formation, nous avons conçu un dispositif inspiré de la CI qui débute après une

première phase de mise en place et de création de communautés. Ainsi, cinq temps favorisant la double topicité de l'apprentissage [17] ont été vécus sur deux semaines par les participants (Fig. 1). Les temps présentiels étaient virtuels et synchrones, c.-à-d. qu'ils se vivaient dans un contexte de visioconférence afin de couvrir les deux thématiques de la formation et de permettre aux participants d'être engagés « *dans une diversité de situations (...) qui leur fasse prendre conscience des différents moyens de "donner" et de "solliciter" les signes de la présence* » [20].

Les activités ont été scénarisées de manière à soutenir les participants dans la définition de leur propre projet de CI et la planification, l'identification et la préparation à la conception des ressources et des moyens technopédagogiques nécessaires.

Un temps complémentaire de retour d'expérience aurait également été intéressant (mais non mis en œuvre) après une durée suffisante pour que les formés puissent eux-mêmes transformer un enseignement et l'expérimenter auprès de leurs étudiants.

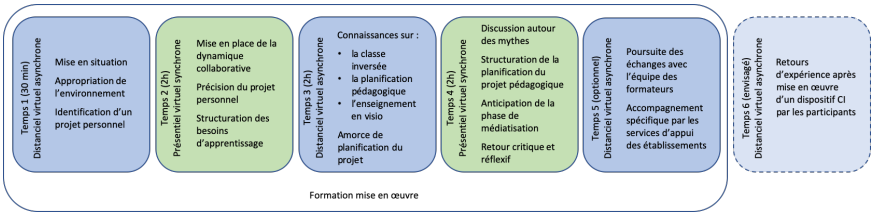


Fig. 1. Structuration globale de la formation

Dans notre contexte mêlant présentiel et distanciel dans un environnement virtuel, la question des outils technopédagogiques était cruciale en vue de proposer un espace de collaboration ouvert et simple d'accès. Dans cette optique, la formation misait sur des technologies simples, largement diffusées et relativement fermées : *Google Drive*, *Slack* et *Socrative*. Le courriel servant pour la mise en place de la formation.

3 Première itération du cycle de conception / recherche

La formation a impliqué six participants et trois formateurs sur deux sites (deux participants et deux formateurs sur l'un, quatre et un sur l'autre) ainsi qu'un participant et un formateur isolés. Cette disposition ne relevait pas d'un choix expérimental particulier, mais du fonctionnement normal d'une formation multisites où formateurs et formés se répartissent sur les sites connectés les plus proches de leur lieu d'attache. Un participant n'a participé qu'aux temps 1 et 2, les six autres ont suivi toute la formation. Il s'agissait de six enseignants et enseignants-chercheurs, ainsi qu'un conseiller pédagogique. Cette première itération constituant un test, nous avions explicitement souhaité limiter le nombre de participants tout en gardant une configuration représentative du fonctionnement attendu en régime normal.

Les documents de cours étaient déposés avant chaque séance sur *Google Drive*, tandis qu'un canal public dans *Slack* permettait les contributions (dépôt de productions ou interactions écrites entre participants et/ou formateurs). Un canal privé de *Slack* permettait aux formateurs de se synchroniser.

Durant l'ensemble de la formation, 97 contributions publiques ont été recensées sur *Slack*, majoritairement sur les temps 2 et 4 présentiels et sur le temps 3 asynchrone. *Slack* fut peu mobilisé en amont et en aval de la formation (temps 1 et 5) alors qu'il constituait le seul moyen de communication. Parmi ces 97 contributions, 14 fichiers ont été déposés par les participants (11 en lien avec les activités scénarisées et 3 pour illustrer leurs propos), tandis que 2 fichiers ont été déposés par les formateurs.

3,1 Analyse

Dès la fin de la formation, de premiers constats empiriques ont soulevé des questions et amorcé le travail d'analyse : les participants avaient-ils pleinement bénéficié des situations pédagogiques planifiées ? L'expérience éducative [9], [11], [12] avait-elle été aussi riche que souhaité ? En accord avec l'approche RCOC, il fallait mener une analyse précise et structurée visant à explorer en quoi les interactions durant la formation ont favorisé, ou non, l'expérience éducative. Répondre à cette question visait non seulement à offrir des pistes de réingénierie rigoureuses, mais également à comprendre l'apprentissage d'enseignants / formateurs en contexte de visioconférence.

Slack constituait l'unique canal de communication durant les temps 1, 3 et 5 asynchrones et était explicitement prévu par le scénario pédagogique comme lieu d'échange pour certaines activités présentielles des temps 2 et 4. Il était donc naturel d'analyser en premier lieu les contributions qui y ont été déposées afin de questionner la présence à distance au sein de l'espace numérique de communication et l'émergence, ou non, d'une communauté d'apprentissage en ligne.

Une analyse préalable montre que le temps 1 d'appropriation asynchrone ne fut effectué que par trois participants sur les sept, avec pour seules interactions des réponses de bienvenue des formateurs. Le second temps asynchrone (temps 3) s'est quasi exclusivement tenu en autonomie (sauf pour un participant). Les retours attendus via *Slack* ont été effectués tardivement : un la veille du temps 4 et quatre dans les 2 h le précédant, sur six participants actifs. Au niveau des interactions entre participants et formateurs, les échanges dans *Slack* se sont uniquement produits la veille du temps 4, pour commenter les fichiers déposés par deux participants (4 jours plus tôt et la veille pour l'un et le jour même pour l'autre). Ces commentaires n'ont conduit à des interactions qu'avec le premier des deux participants. Durant les temps 2 et 4 synchrones, seuls les formateurs ont réagi au dépôt d'une contribution par les participants (sauf pour un cas d'interpellation directe entre deux participants au temps 4). Certaines activités prévues dans le scénario n'ont cependant pas généré d'échanges dans *Slack*, même si des échanges oraux ont eu lieu. Ces échanges n'ont pas été analysés lors de cette itération de construction des outils, mais l'ont été lors des suivantes.

Une façon de répondre à notre questionnement consista alors à analyser les contributions perçues via *Slack* selon les trois dimensions de la présence à distance [11] : la « *présence sociocognitive* » qui existe lorsque les apprenants collaborent (et donc confrontent, coconstruisent, etc.) afin de résoudre un problème soulevé par une situation complexe, la « *présence socioaffective* » qui prend forme lorsque les apprenants créent un climat social équilibré et agréable, et la « *présence pédagogique* » qui plonge les participants dans un cadre pédagogique structuré, modéré et animé pour réaliser les activités proposées.

Pour ce faire, nous avons construit notre propre classification, pragmatique et liée à nos besoins, sur la base des attentes exprimées dans le scénario et de nos observations, tout en nous plaçant dans la perspective d'un raffinement itératif lors des prochaines formations. Sept catégories ont ainsi émergé, couvrant l'ensemble des contributions : encouragement / motivation, présentation, discussion pédagogique (sans rapport direct avec la séquence) ou informelle, dépôt d'une contribution, réaction à une contribution, demande de précisions / compléments et apport de précisions / compléments. Chacune de ces catégories a été rapprochée d'un type de présence à distance (pouvant différer selon la nature de l'auteur de la contribution) et nous avons utilisé cette typologie pour caractériser les 97 contributions publiques de *Slack* [14].

Les données collectées permettent un premier éclairage sur les interactions qui se sont produites durant la formation. Mais ont-elles favorisé, ou non, une expérience éducative riche ? Constatons, tout d'abord, la faiblesse des interactions écrites entre participants. Lors des temps synchrones, les interactions se sont majoritairement faites oralement, celles sur *Slack* étant généralement à l'initiative des formateurs. Une exception notable : *Slack* a parfois été utilisé par les participants pour faire des « retours » sans interrompre l'orateur. La présence de plusieurs formateurs permettait alors une réponse immédiate et/ou d'apporter des compléments ou des précisions, éventuellement sous forme de ressources complémentaires. Les participants se sont néanmoins peu approprié *Slack* ou la visioconférence pour interagir entre eux.

Lors du temps asynchrone, le scénario suggérait *Slack* pour partager les questions et susciter un débat entre les participants : ce ne fut pas le cas, ce que le dépôt tardif des productions a sans doute renforcé. Ce comportement avait été anticipé comme possible, et fut analysé avec les participants durant le temps 4.

Bien que le scénario pédagogique ait été conçu pour amener les participants à aborder une « situation problématique » (la CI) en collaborant à distance, cette collaboration s'est principalement faite par l'intermédiaire des formateurs. Les interactions sociales de confrontation de points de vue, d'ajustement mutuel, de négociation et de délibération entre participants, et la *présence sociocognitive* qui en résulte [11], ont été largement intermédiées par les formateurs, laissant peu de place aux interactions directes entre participants. L'absence d'interactions directes (parmi les 30 contributions typées « présence sociocognitive », 23 sont liées au dépôt initial d'une contribution) limitant fortement l'apparition de collaboration contradictoire [11].

La collaboration à distance reste une difficulté que le formateur peut aider à dépasser en jouant un rôle de « facilitateur » au travers de ses interactions avec les apprenants. Ce rôle de facilitateur, qui correspond à la *présence pédagogique* [11], peut se décliner sur trois facettes : la coordination, l'animation et la modération. L'analyse des contributions nous éclaire sur la présence pédagogique des formateurs et montre une carence sur ses facettes coordination (liée au scénario pédagogique) et animation durant les temps asynchrones. La fonction animation semble mieux tenue durant les temps synchrones (avec un rôle de personne-ressource clair), mais la fonction coordination reste limitée (le cadre mis en place n'a pas favorisé les échanges entre participants dans l'espace numérique) ou trop directive, plaçant les formateurs en position d'intermédiaires, ne favorisant pas les échanges entre participants.

Peu d'interactions, dans *Slack*, correspondent à une composante *socioaffective* de la présence. Cette faiblesse est cohérente avec le peu d'interactions entre participants, mais ne préjuge pas nécessairement d'une carence de ce type de présence (même si les

liens affectifs entre les sites distants restent à confirmer). Le fait que le groupe soit constitué d'universitaires engagés volontairement dans une démarche de formation favorisait en effet certainement le climat égalitaire, détendu et cordial observé.

Cette première itération montre donc une faiblesse du côté présence sociocognitive et des interactions limitées. Ainsi, à la question soulevée à cette étape de recherche, il est possible de répondre que l'expérience éducative est moins riche que ce qui était visé par les formateurs. Il convenait donc d'en tenir compte dans l'itération suivante.

4 Seconde itération du cycle de conception / recherche

Sur la base des premiers constats, une nouvelle itération s'est déroulée à l'automne 2017. Les objectifs de la formation et son organisation technopédagogique sont restés inchangés, tandis que le design pédagogique a été adapté. Nous avions exclu de modifier l'organisation technopédagogique afin de ne pas modifier tous les paramètres en même temps, d'autant qu'aucun participant n'avait éprouvé de difficulté particulière liée aux outils. Des modifications ont, par contre, été apportées au niveau des ressources. En effet, plusieurs participants avaient souhaité des informations, références ou exemples complémentaires qui ont donc été mis à disposition en amont de chacun des temps. L'intention était ici d'augmenter la présence pédagogique [11] et permettre aux participants d'accroitre leur perception de la contrôlabilité [21], tout en facilitant l'acquisition de connaissances nécessaires au développement des compétences ciblées [22], favorisant ainsi une présence sociocognitive plus riche.

L'organisation des temps de travail a été précisée et affinée afin de mieux montrer l'importance de l'investissement dans chacun d'eux (et en particulier dans le temps 3 asynchrone) [21]. La durée des activités des temps 2 et 4 a été ajustée afin de laisser plus de temps pour échanger, confronter les perceptions et travailler au développement des projets individuels (présence sociocognitive). Il était, par exemple, proposé aux participants que les questions soient prioritairement posées dans *Slack* pour permettre aux formateurs d'y répondre de manière asynchrone (présence pédagogique).

Un investissement plus réactif des formateurs dans *Slack* a été instauré afin que tous les participants aient rapidement des réponses, des rétroactions ou des suggestions (présence pédagogique) et puissent ainsi rester engagés et motivés (grâce à une meilleure perception de la valeur de l'activité).

Cette seconde itération a impliqué six participants dans deux salles (deux dans l'une, sans présence de formateurs, et quatre dans l'autre) et quatre formateurs dans trois salles (un dans la salle avec les quatre participants, deux dans une salle, un dans une autre). Tous les participants ont suivi la formation dans son ensemble et étaient enseignants, enseignants-chercheurs, doctorants et conseiller pédagogique. Contrairement à la première itération, la volonté n'était pas de limiter le nombre de participants, même si le volume est resté le même (probablement à cause d'une période non-optimale pour mobiliser des enseignants-chercheurs et surtout d'une insuffisante communication autour de la formation).

Durant cette seconde itération, 127 contributions publiques sur *Slack* ont été relevées. La répartition temporelle de ces messages (tableau 1) montre une plus forte utilisation de *Slack* lors des temps 1 et 3 à distance que lors de l'itération 1.

Tableau 1. Répartition des messages échangés (*Slack*)

	Temps 1		Temps 2		Temps 3		Temps 4		Temps 5		Total
Itération 1	8	8%	27	28%	32	33%	26	27%	4	4%	97
Itération 2	15	12%	29	23%	51	40%	17	13%	15	12%	127

Parmi ces 127 contributions, 24 fichiers ont été déposés par les participants, 8 par les formateurs. Notons que le temps 3 asynchrone a été mobilisé par plusieurs participants pour retravailler et redéposer les productions réalisées lors du temps 2.

Au-delà des contributions sur *Slack*, l'analyse de l'enregistrement vidéo met en évidence 175 interventions orales lors du temps 2 (121 des formateurs et 54 des participants) et 125 lors du temps 4 (78 des formateurs et 47 des participants).

Dès la fin de la formation, à l'instar de la première itération, les formateurs ont fait des constats empiriques avec le sentiment que les participants étaient, dans l'ensemble, satisfaits et avaient réalisé des apprentissages relatifs aux compétences ciblées. Les changements apportés ne semblaient donc pas, à priori, avoir eu d'effet négatif et il semblait que la participation et l'engagement étaient mieux répartis et plus significatifs.

Exploitant la même méthodologie et le même cadre théorique, les données de cette seconde itération ont été analysées dans la perspective d'une double boucle d'amélioration : tant au niveau de la formation, que des outils d'analyse.

Au niveau de la formation, contrairement à l'itération 1, les temps 1 et 3 furent nettement plus riches. Le temps 1 fut investi par les six participants, tandis que cinq se sont impliqués tout au long de la semaine du temps 3 en partageant régulièrement leurs productions sur la plateforme *Slack* (même si un pic est observé la veille du temps 4). Dans les deux cas, les interactions se sont cependant limitées à un dialogue avec les formateurs (sans interactions entre participants sur *Slack*, mais en tenant des échanges suivis avec les formateurs). Cet investissement supérieur est à mettre en relation avec la réingénierie de la formation visant à donner plus de sens à l'investissement dans les temps asynchrones et avec la réactivité accrue des formateurs dans *Slack* (tous les messages des participants ont reçu des réponses dans les 24 h et majoritairement dans l'heure). Il convient également de noter que les échanges se sont poursuivis avec un participant lors du temps 5 post-formation (15 messages sur *Slack* en cinq jours).

Au niveau des outils d'analyse, la classification des contributions ayant émergé lors de la première itération [14] a été affinée grâce au corpus plus important de contributions (sur la base de deux itérations) et à une vision plus structurée du besoin d'observation (analyse de la présence à distance via les contributions recensées). L'idée était de pouvoir caractériser chaque contribution par un ensemble de marqueurs permettant d'en analyser la contribution à telle ou telle dimension de la présence à distance, tout en permettant également une analyse par rapport au scénario pédagogique. Concernant la présence à distance perçue, cela amène à un ensemble de cinq métadonnées ayant émergé de l'analyse des traces : l'initiateur (participant ou formateur), le récepteur (participant, formateur ou les deux), sa sémantique (message de service, présentation, encouragement / motivation, échange pédagogique en lien avec l'activité scénarisée, autres échanges), sa nature (message technique, dépôt d'une production — fichier ou message — prévue dans le scénario, question / demande, réponse à une question / demande, commentaire, apport de ressource complémentaire) et son type (message ou ressource).

Concernant l'analyse des contributions au regard du scénario pédagogique, deux autres métadonnées sont calculées : l'activité du scénario concernée par la contribution et le temps durant lequel elle se produit (qui n'est pas nécessairement celui anticipé).

Les deux itérations ont été analysées avec cette classification revue, mettant en évidence 29 types de contributions dans *Slack* pour la première itération et 22 pour la seconde (soit 37 types observés au moins une fois sur les 360 combinaisons de marqueurs possibles). Cette classification fut également utilisée pour analyser les interventions orales de la seconde itération, mettant en évidence 32 types de contributions (dont 11 observés à la fois dans *Slack* et oralement).

Les types observés ont chacun été rapportés à une dimension de la présence à distance, sur la base de regroupements liés à la sémantique de la contribution et à son initiateur (tableau 2) afin de rester cohérent avec l'analyse effectuée sur la première itération, tout en apportant un niveau plus fin de compréhension.

Tableau 2. Identification des types de présence

Sémantique	Initiateur	Présence	Nb types regroupés
Message de service	Tous	Socioaffective	2
Présentation	Tous	Socioaffective	3
Encouragement / motivation	Participant	Socioaffective	5
Encouragement / motivation	Formateur	Pédagogique	6
Échange pédagogique en lien avec l'activité scénarisée	Participant	Sociocognitive	10
Échange pédagogique en lien avec l'activité scénarisée	Formateur	Pédagogique	5
Autres échanges	Tous	Socioaffective	7

La réingénierie du scénario visait, en particulier, à mieux impliquer les participants dans le temps 3. Cela passait par de nouvelles consignes, rappelées avec insistance lors du temps 2, et un investissement plus réactif des formateurs lors du temps 3. L'analyse détaillée de ce temps 3 (tableau 3), où *Slack* représentait l'unique canal de communication, montre, en particulier, une augmentation des échanges pédagogiques (passant de 13 à 18 pour les participants et de 8 à 24 pour les formateurs).

Une analyse plus fine, basée sur les cinq marqueurs, montre plus particulièrement une forte augmentation des contributions de nature « commentaire » (passant de 7 à 24) alors que les contributions d'autres natures sont restées globalement stables. L'analyse de ces commentaires montre une progression tant au niveau des participants (passant de 1 à 8) que des formateurs (passant de 6 à 16). Au niveau de la sémantique associée, on observe que les échanges pédagogiques sont ceux qui ont le plus progressé (passant de 0 à 3 pour les participants et de 1 à 11 pour les formateurs).

L'analyse des interventions orales durant les temps 2 et 4 montre comme attendu une forte proportion d'échanges pédagogiques de nature « commentaire » (66 faits par les formateurs et 11 par les participants), mais également de nature « question / demande » pour les formateurs (30 interventions) et « réponse à une question / demande » pour les participants (29 interventions).

Le scénario de la seconde itération suggérait que les participants utilisent *Slack* pour poser leurs questions reliées aux différentes notions et contenus abordés. L'analyse des contributions écrites montre que ce ne fut pas le cas, celles de nature « question » initiées dans *Slack* par les participants restant faibles (deux pour la première itération,

une pour la seconde). L'analyse des contributions orales, quant à elle, montre 17 contributions de nature « question » effectuées par les participants, majoritairement suite à une réponse faite par un formateur (échange suivi initié par un formateur).

Afin de répondre aux attentes identifiées lors de la première itération, des ressources complémentaires avaient été mises à disposition des participants. Bien qu'elles aient été mobilisées, les besoins ont perduré et les formateurs ont déposé 2, 5 et 7 ressources complémentaires lors des temps 2, 3 et 4 de la seconde itération, contre 4, 0 et 6 lors de la première. L'investissement supérieur de l'équipe de formateurs lors du temps 3 asynchrone est aussi visible sur cet aspect, les ressources complémentaires apportées renforçant la composante animation de la présence pédagogique.

Cet investissement supérieur, qui s'est traduit par des productions plus nombreuses et plus abouties, ne s'est pas retrouvé au niveau du travail collaboratif entre participants qui reste largement intermédié par les formateurs, tant sur *Slack* qu'oralement (seulement 2 échanges oraux suivis entre participants, l'un au temps 2 l'autre au temps 4, lors d'activités où ils devaient exprimer leur ressenti sur la formation).

De ce point de vue, l'amélioration de la présence sociocognitive [11] reste en deçà des attentes, mais le faible nombre de participants était sans doute un facteur limitant. Si la présence socioaffective reste bonne pour les mêmes raisons que précédemment, la présence pédagogique est renforcée dans ses composantes animation et modération, tout en restant en retrait sur sa composante coordination.

Tableau 3. Analyse des contributions (focus sur le temps 3 distanciel asynchrone)

Initiateur	Récepteur	Sémantique	Nature	Type	Iter 1	Iter 2
Participant	Participant	Encouragement	Commentaire	message		1
Participant	Formateur	Encouragement	Réponse	message	1	
Participant	Formateur	Encouragement	Commentaire	message	1	
Participant	Les deux	Encouragement	Commentaire	message		1
Participant	Les deux	Échange péda.	Dépôt	message		4
Participant	Formateur	Échange péda.	Dépôt	ressource	1	
Participant	Les deux	Échange péda.	Dépôt	ressource	11	9
Participant	Formateur	Échange péda.	Réponse	message	1	2
Participant	Formateur	Échange péda.	Commentaire	message		3
Participant	Formateur	Autres	Commentaire	message		3
Formateur	Participant	Encouragement	Question	message	1	
Formateur	Participant	Encouragement	Réponse	message	3	
Formateur	Participant	Encouragement	Commentaire	message	4	3
Formateur	Participant	Échange péda.	Question	message	1	4
Formateur	Participant	Échange péda.	Réponse	message	2	4
Formateur	Participant	Échange péda.	Commentaire	message	1	11
Formateur	Participant	Échange péda.	Apport compl.	message	2	
Formateur	Participant	Échange péda.	Apport compl.	ressource	2	5
Formateur	Participant	Autres	Commentaire	message	1	2

5 Conclusion et perspectives

A l'issue des deux premières itérations de la formation, un constat s'impose : la présence sociocognitive semble faible et les interactions entre participants étaient très limitées. Cela provient-il du contexte numérique contraignant et d'une relative inadéquation des outils de communication, du faible nombre de participants ou cela témoigne-t-il d'un engagement cognitif faible ou superficiel des participants malgré des choix pédagogiques des formateurs visant à permettre des perceptions positives de l'activité et de sa contrôlabilité [21] ?

Dans une prochaine itération, il pourrait être envisagé de miser plus sur une démarche structurée tel le cycle de l'apprentissage expérientiel de Kolb [18] et dépasser la phase d'expérimentation pour introduire des phases d'analyse, de généralisation et de transfert. Il pourrait également être envisagé de créer des activités de résolution de problèmes en collaboration afin de susciter les « *transactions entre les apprenants lors de la pratique d'enquête* » [11] et favoriser la présence sociocognitive.

D'autre part le faible niveau d'interaction entre les participants dans *Slack* nous interroge. Il peut, bien évidemment, provenir du faible nombre de participants ne réunissant pas la masse critique pour entretenir des échanges suivis. L'analyse de la troisième itération qui s'est déroulée au printemps 2019 avec 22 participants sur quatre sites nous permettra une compréhension plus fine de ce point (les interactions semblent effectivement plus riches). Une autre hypothèse pourrait aussi être que les apports de connaissances étaient liés à trop d'éléments complexes différents tels qu'enseigner en contexte de visioconférence et enseigner en CI. Une telle hypothèse méritera un travail de recherche plus fin, relatif notamment à l'identification des apprentissages réels des participants ou aux outils numériques exploités et leurs impacts sur les choix et actions des participants. Peut-être serait-il aussi pertinent de distinguer les notions théoriques et pratiques reliées au contexte de visioconférence, afin de miser avant tout sur l'expérience vécue de ce contexte (afin d'y sensibiliser et familiariser les participants), et en faire un objet de réflexion et d'apprentissage distinct tel qu'expérimenté lors de la troisième itération de la formation.

Bibliographie

1. Sanchez, E., Monod-Ansaldi, R.: Recherche collaborative orientée par la conception. Education et didactique 9(2), 73-94 (2015)
2. Wang, F., Hannafin, M. J.: Design-based research and technology-enhanced learning environments. Educational Technology Research and Development 53(4), 5-23 (2005)
3. Alhlak, B. A., Ramakrisnan, P., Hameed, Z. S., Mohsenu, H. R.: Video conference: integrated tool for identifying CSF in education development in UiTM. Procedia-Social and Behavioral Sciences 67, 102-113 (2012)
4. BECTA: What the research says about video conferencing in teaching and learning. https://webarchive.nationalarchives.gov.uk/20130402162028/https:/www.education.gov.uk/publications/eOrderingDownload/15007.pdf
5. Dahl, A.: Visioconférence en éducation : exploite-t-on son potentiel pédagogique? DistanceS 11(1), 1-16 (2009)

6. Lawson, T., Comber, C., Gage, J., Cullum-Hanshaw, A.: Images of the future for education? Videoconferencing: a literature review. Technology, Pedagogy and Education 19(3), 295-314 (2010)
7. Macedo-Rouet, M.: La visioconférence dans l'enseignement : ses usages et effets sur la distance de transaction. Distances et savoirs 7(1), 65-91 (2009)
8. Paquay, L.: Continuité et avancées dans la recherche sur la formation des enseignants. Les cahiers de recherche du GIRSEF 90, 2-38 (2012)
9. Garrison, D. R., Anderson, T., Archer, W.: Critical inquiry in a text-based environment: computer conferencing in higher education. The Internet and higher education 2, 87-105 (2000)
10. Shea, P., Bidjerano, T.: Learning presence as a moderator in the community of inquiry model. Computers & Education 59(2), 316-326 (2012)
11. Jézégou, A.: La présence en e-learning : modèle théorique et perspectives de recherche. La revue internationale de l'apprentissage en ligne et de l'enseignement à distance 26(1) (2012)
12. Kawachi, P.: Unwrapping presence. Distances et savoirs 9(4), 591-609 (2011)
13. Paquette, G.: L'ingénierie pédagogique. PUQ, Sainte-Foy (2002)
14. Meyer, F., Barré, V., Lefebvre, N., Gandon, C.: Développer des compétences pour enseigner en classe inversée dans un contexte de visioconférence: un projet de recherche orientée par la conception. In: Le numérique en éducation, PUQ, pp.181-210 (2019)
15. Bishop, J., Verleger, M.: The flipped classroom: A survey of the research. In : ASEE national conference proceedings, Atlanta, GA, vol. 30(9), pp.1-18 (2013)
16. Lage, M. J., Platt, G. J., Treglia, M.: Inverting the classroom: A gateway to creating an inclusive learning environment., 30-43 (2000)
17. Nizet, I., Galiano, O., Meyer, F.: Vers un cadrage théorique pour comprendre la classe inversée. In : La classe inversée ou comment rendre mon enseignement intéressant. De Boeck, Bruxelles (2016)
18. Kolb, D. A.: Experiential learning. Prentice-Hall, New Jersey (1984)
19. Lebrun, M.: Les classes inversées. paysage ouvert et visite guidée. Available at: http://lebrunremy.be/WordPress/?p=791
20. Jacquinot-Delaunay, G.: Chapitre 6. Absence et présence dans la médiation pédagogique ou comment faire circuler les signes de la présence. In : Pratiquer les TICE : Former les enseignants et les formateurs à de nouveaux usages. De Boeck, Louvain-la-Neuve (2002) 103-113
21. Viau, R.: Savoir motiver les étudiants. In : Se former à la pédagogie de l'enseignement supérieur. Association québécoise de pédagogie collégiale, collection PERFORMA, Montréal (2014) 235-250
22. Bachy, S.: Un modèle-outil pour représenter le savoir technopédagogique disciplinaire des enseignants. Revue internationale de pédagogie de l'enseignement supérieur 30(2) (2014)

Comparaison de questions posées et votées en ligne dans le cadre d'une classe inversée

Fatima Harrak, François Bouchet, Vanda Luengo

Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, LIP6, F-75005 Paris, France
fatima.harrak@lip6.fr
francois.bouchet@lip6.fr
vanda.luengo@lip6.fr

Résumé. Les questions des élèves sont utiles pour leur apprentissage et l'adaptation pédagogique des enseignants. Nous étudions ici les questions posées en ligne par des étudiants de première année de médecine, utilisées par les professeurs pour préparer des sessions de questions-réponses. Comme les étudiants peuvent aussi voter sur les questions posées, on peut s'interroger sur la valeur pédagogique du vote : a-t-il le même impact en termes d'apprentissage, et les étudiants votent-ils sur des questions similaires à celles qu'ils posent ? Pour répondre à ces questions, nous avons développé un schéma de codage de la nature des questions, puis conçu un annotateur automatique que nous avons appliqué à l'ensemble du corpus. La comparaison votants vs. non-votants révèle que les votants réussissent mieux, mais peu de différences apparaissent entre la nature des questions posées et celles votées pour les étudiants qui font les deux. Ce résultat confirme la valeur du vote comme alternative à la formulation de question pour les étudiants sachant déjà formuler leurs propres questions.

Mots-clés. Question d'élève, vote d'élève, profil d'apprenant, schéma de codage, classe inversée

Abstract. Students' questions are useful for their learning experience as well as to help teachers to adapt their pedagogy. We study here a corpus of questions asked online by 1st year medical students, used by their professors to prepare their questions and answers sessions. As students can also vote on questions already asked, we can wonder about the pedagogical value of such votes: do they help learning as much, and are students voting on questions similar in nature to the ones they ask? To answer, we have developed a coding scheme of questions according to their nature and built an automatic annotator to tag the whole corpus. Comparing voters and non-voters reveals the former perform better than the latter, and there are little differences between questions asked and voted for students who do both. This result confirms the value of voting as an alternative to asking one's own questions for students who know how to ask questions.

Keywords. Student's question, student's vote, student's profile, coding scheme, blended learning

1 Introduction

Les questions des élèves jouent un rôle important dans le processus d'apprentissage, non seulement pour aider les élèves à mieux apprendre [1], mais aussi pour l'enseignant afin de déterminer ce qui a été compris et d'adapter sa pédagogie en conséquence. Les environnements en ligne et EIAH peuvent éliminer de nombreux obstacles qui empêchent les élèves de poser des questions en classe [2]. Nous nous intéressons ici à une formation hybride dans laquelle les étudiants doivent poser chaque semaine des questions à partir de supports de cours étudiés à distance avant le cours (selon une approche de classe inversée), pour aider les enseignants à préparer leurs séances de questions-réponses en présentiel. Cependant, compte tenu du volume de questions posées, les enseignants n'ont souvent pas assez de temps pour répondre à chaque question et doivent donc sélectionner celles auxquelles ils vont répondre. Pour limiter le nombre de questions et les aider dans ce choix, ils encouragent les étudiants à voter sur les questions déjà posées avant d'en poser de nouvelles. D'un point de vue pédagogique, cela suppose que les étudiants lisent les questions des autres, ce qui peut également avoir un impact positif en les forçant à s'interroger sur leur propre compréhension des points abordés par leurs camarades. Mais on peut aussi penser qu'un vote n'est pas exactement équivalent à une question. En effet, dans le cadre théorique Active-Constructive-Interactive (ACI) proposé par Chi [3], le fait de poser une question est une activité de nature « constructive » (mobilisant des processus cognitifs tels que la recherche de lacunes dans ses connaissances et la restructuration de celles-ci), plus à même d'aider l'apprentissage qu'une activité « active » comme le fait de voter (mettant uniquement en jeu une recherche dans ses connaissances pour savoir si on saurait ou non répondre à cette question). Face à cette contradiction sur la valeur d'un vote, nous avons conduit des analyses afin d'explorer comment les votes sont associés à la réussite des étudiants et aux questions qu'ils posent. Plus précisément, notre objectif était de répondre aux deux questions de recherche suivantes : (QR1) Quel est l'impact du vote sur la performance d'un étudiant, notamment en comparant ceux qui posent ou non également des questions ? (QR2) Y a-t-il des différences dans la nature des questions posées et votées par les étudiants ?

Pour traiter ces questions de recherche, nous avons défini un schéma de codage adapté aux questions des étudiants et un système d'annotation automatique pour annoter l'ensemble du corpus de questions dont nous disposions. Dans la suite de cet article, nous proposons dans la section 2 un état de l'art sur les différentes taxonomies de questions. Nous présentons dans la section 3 le contexte de l'étude et les données utilisées, avant de décrire dans la section 4 la méthodologie de catégorisation de questions et d'annotation automatique. Enfin, nous présentons dans la section 5 les résultats des analyses effectuées pour répondre aux deux questions de recherche et nous concluons avec quelques perspectives et limites de ce travail en section 6.

2 Etat de l'art

Les chercheurs ont étudié le comportement de questionnement des élèves dans divers contextes éducatifs, tels que la classe [4], le tutorat [5] et les environnements

d'apprentissage en ligne [6]. Scardamalia et Bereiter [7] font la distinction entre les questions fondées sur le texte et celles fondées sur le savoir (ces dernières ayant un potentiel éducatif plus fort). Bien que cette distinction soit pertinente, n'ayant pas accès aux transcriptions des vidéos et diapositives avec lesquelles les étudiants ont interagi avant de poser leurs questions, il était difficile d'identifier ceci automatiquement. D'autres chercheurs ont proposé une typologie de questions distinguant celles pouvant faire l'objet d'une investigation scientifique (*e.g.* comparaison, cause à effet, prédiction, exploration) des autres [4]. Bien que notre but ne soit pas d'encourager à poser un certain type de questions, cette distinction pourrait s'appliquer à notre travail, mais est difficile à réaliser sans experts du domaine. Graesser et Person [5] ont pour leur part élaboré une taxonomie de questions posées pendant les séances de tutorat, utilisée pour la génération automatique de questions. Bien que leur taxonomie puisse être pertinente ici, certaines catégories comprenaient des « questions de raisonnement approfondi » de haute qualité, associées à des modèles de raisonnement difficiles à identifier automatiquement. Enfin, des recherches récentes [8] ont utilisé une version réduite de la taxonomie de Bloom [9] pour établir un lien entre rétroaction pratique et performance de l'apprenant en matière d'évaluation. Cette taxonomie, en raison de son origine, tend toutefois à être plus appropriée aux questions de l'enseignant qu'à celles des élèves.

L'analyse des questions d'apprentissage a été utilisée à des fins très diverses afin d'améliorer l'efficacité de l'enseignement et l'apprentissage des élèves. Ainsi, Graesser et Person [5] ont étudié la relation entre les questions posées et la réussite, constatant une corrélation positive entre la réussite des élèves et la « qualité » de leurs questions (une fois acquise une certaine expérience en tutorat) et une corrélation négative entre la réussite des élèves et le nombre total de questions. Les recherches de Chin [10, 11], ont montré que les questions se situent sur un continuum allant de « superficielles» à « profondes » et que le type de questions que posent les élèves dépend de la façon dont ils abordent leurs tâches d'apprentissage. En se basant sur ces études, on pourrait conclure que poser des questions « superficielles » est généralement peu utile.

Les votes des élèves ont également été étudiés pour analyser le comportement des élèves dans les forums en ligne. Bihani et al. [12] ont utilisé le nombre de votes sur les questions et réponses des étudiants et les réponses de l'enseignant pour révéler les paires de questions/réponses pertinentes pour les futurs cours. Zeng et al. [13] ont également utilisé le nombre de votes pour détecter la confusion au sein de messages de forum de discussion, constatant que les messages contenant de la confusion sont considérés comme importants et reçoivent plus de votes. Jiang et al [14] ont analysé les étudiants considérés comme des "influenceurs" (utilisateurs dont les messages génèrent beaucoup de réponses dans les forums d'un MOOC). Ces influenceurs ont des résultats plus faibles et reçoivent moins de votes que les utilisateurs actifs (ceux qui postent régulièrement sur le forum). De même, Wong et al. [15] ont analysé les votes (positifs et négatifs) sur les messages et les commentaires des utilisateurs actifs. Contrairement à [14], ils ont constaté que les utilisateurs actifs sont aussi des utilisateurs influents qui apportent généralement une contribution positive aux discussions du forum du MOOC. Les votes des élèves ont donc surtout été utilisés pour analyser le comportement des élèves, mais la nature des questions votées n'a apparemment pas encore été explorée.

Dans l'ensemble, les taxonomies proposées jusqu'à présent dépendent essentiellement du contexte, et nous avons décidé de définir un nouveau schéma de codage utilisant une approche fondée sur les données. Dans cet article, nous nous intéressons donc à l'analyse de la nature des questions posées et votées par les étudiants.

3 Contexte

Nous avons considéré l'ensemble des questions posées par des étudiants de 1ère année de médecine et pharmacie de l'université Joseph-Fourier de Grenoble en 2012-2013. 1608 étudiants étaient inscrits cette année-là, bien que tous n'aient pas posé de questions. La Faculté de médecine dispose d'un système de formation hybride pour ses étudiants de 1ère année (PACES). Chaque semestre se termine par un concours (en janvier et mai) éliminatoire (seule une partie des étudiants qui ont échoué au concours est autorisée à repasser l'année une seule fois, les autres devant se réorienter). Chaque cours est composé de deux à six séquences de 4 semaines (*cf.* Figure 1).

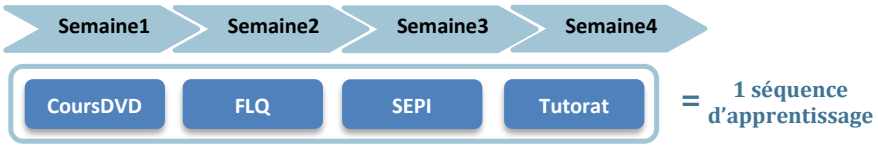


Fig. 1. Les quatre activités d'une séquence d'apprentissage sur quatre semaines

Dans chaque séquence, la première semaine consiste à étudier le cours sur DVD-ROM ou sur le site Medatice (diapositives + vidéos du professeur). La deuxième semaine est consacrée à la Formulation en ligne des questions (FLQ) pour les enseignants : ces questions concernent exclusivement les cours multimédias étudiés la semaine précédente. Les élèves peuvent voir les questions posées par les autres élèves de leur groupe et voter pour celles auxquelles ils veulent aussi une réponse, mais ils ne peuvent pas y répondre. En fin de semaine, les questions sont envoyées aux enseignants intervenant la troisième semaine, qui les utilisent pour structurer leurs sessions d'enseignement interactives en classe. Au cours de ces sessions, l'enseignant répond à certaines questions posées en ligne par les étudiants. La quatrième semaine est consacrée à des séances de tutorat afin de tester les connaissances acquises lors de la séquence de formation par le biais d'un auto-test utilisant des questions à choix multiples (QCM), qui sont ensuite corrigées par un professeur auxiliaire. Il y a deux séances de tutorat de 2 heures/semaine. Chaque étudiant peut vérifier individuellement ses notes et son classement par rapport à l'ensemble de la promotion, et il est nécessaire s'être connecté à la plateforme de questions pour pouvoir consulter ses notes. Ces résultats de tutorat permettent aux élèves d'évaluer la qualité de leur apprentissage.

Pour chacun des 13 cours, nous avons donc 2 à 6 ensembles de questions (un par séquence) posées par 429 étudiants (6457 questions au total) et votées par 672 étudiants (10951 votes) pendant la deuxième semaine de chaque séquence. La répartition des questions est inégale (*cf.* Table 1), avec plus de questions au 1er semestre, notamment car certains étudiants sont obligés d'arrêter à la fin de celui-ci. On note que seul un élève

sur quatre a posé au moins une question, ce qui peut s'expliquer en partie par l'encouragement à voter au lieu de poser des questions (pour forcer à lire les questions des autres mais aussi pour réduire le nombre de questions reçues par courriel).

Table 1. Distribution des questions posées par cours

BCH	BPH	HBD	BCE	ANT	PHS	SSH	ICM	MAT	Spec.
19%	17%	15%	11%	10%	9%	8%	6%	3%	1%

BCH = Biochimie, BPH = Biophysique, HBD = Histoire et biologie du développement, BCE = Biologie cellulaire, ANT = Anatomie, PHS = Physiologie, SSH = Santé, société, humanité, ICM = Initiation à la connaissance du médicament, MAT = Mathématique, Spécialité = Pharmacie, Odontologie, Maïeutique

4 Catégorisation et annotation de questions

4.1 Schéma de codage

Afin d'identifier la nature des questions posées par les étudiants, nous avons décidé de développer notre propre schéma de codage. Nous avons travaillé sur un échantillon de 800 questions (12% du corpus) issues de deux cours (BCH et HBD), considérés par l'équipe pédagogique comme étant parmi les plus difficiles et ayant suscité le plus de questions (*cf.* Table 1). Cet échantillon a été divisé en 4 sous-échantillons de 200 questions pour appliquer 4 étapes successives de catégorisation.

(1) L'**étape de découverte** consistait à regrouper empiriquement des phrases ayant des similitudes pour en extraire des concepts significatifs. Bien que les enseignants demandent aux élèves de poser des questions simples (c.-à-d. d'éviter des questions comme « Pourriez-vous expliquer à nouveau X ? De plus, Y n'était pas clair »), 40% des questions pouvaient être divisées en plusieurs questions indépendantes. Une fois les phrases segmentées en questions dites simples, nous avons regroupé celles dont la structure (ex : « qu'est-ce que X ? » et « qu'est-ce que Y ? ») et la sémantique (ex : « qu'est-ce que X ? » et « pourriez-vous définir X ? ») semblent similaires. Des groupes de questions ont ensuite reçu des "étiquettes" (ex : « définition d'un concept »), puis nous avons identifié les exclusions mutuelles entre étiquettes (ex : une question simple ne peut pas être à la fois une vérification et une demande de ré-explication). Cela nous a conduit à définir le concept de « dimensions », ensembles d'étiquettes de types de questions similaires mais mutuellement exclusives. Chacune de ces étiquettes individuelles ("vérification", "ré-explication"...) sont des valeurs pouvant être associées à une dimension. Chaque question simple peut alors être associée à une annotation dans ce schéma de codage en choisissant une valeur par dimension. Une annotation de question peut donc être vue comme un vecteur de 4 valeurs (ex : "Pourriez-vous réexpliquer la différence entre un composé ionisable et un partiellement ionisable ?" est représenté par le vecteur [Ree,0,Lie,0], cf. Table 2).

(2) L'**étape de consolidation** consistait à annoter le deuxième sous-échantillon pour valider les dimensions et les valeurs précédemment identifiées. Cela a conduit à divers ajustements des dimensions pour s'assurer qu'elles étaient bien indépendantes les unes

des autres (par exemple l'ajout de la valeur "correction" dans Dim2, non identifiée précédemment). Parallèlement, les dimensions identifiées ont été revues et validées par un professeur expert enseignant dans le cadre de PACES, qui a estimé que les catégories étaient potentiellement pertinentes pour analyser les questions des étudiants.

(3) Dans l'**étape de validation**, nous avons effectué une double annotation pour valider l'ensemble de nos catégories sur le troisième sous-échantillon de 200 phrases. Ensuite, deux annotateurs humains ont utilisé comme référence unique le schéma de codage créé à la fin de l'étape précédente pour annoter chaque segment (238 au total). A l'issue de l'étape précédente, trois dimensions avaient été identifiées : Dim1 (relative au type de question), Dim2 (relative à la modalité d'explication), Dim4 (facultative, annotée uniquement si la question est une vérification, relative à la nature de ce qui est vérifié). La dimension appelée "Dim3" plus loin n'existait pas encore à cette étape. Les annotateurs humains ont fait deux annotations distinctes et indépendantes sur chaque dimension, et leur accord a été évalué à l'aide du Kappa de Cohen (K1=0,72, K2=0,62 où K1 et K2 correspondent respectivement au Kappa de Dim1 et Dim2). Pour Dim4, en raison de son caractère facultatif, les deux annotateurs n'ont pas nécessairement annoté les mêmes questions (annotateur 1 : 82 questions ; annotateur 2 : 68 questions) : sur les 68 en commun, le kappa valait 0,66. Puis ils se sont rencontrés pour discuter et résoudre les désaccords, ce qui a conduit à un affinement final des catégories (par exemple, séparation des catégories Dim1 et Dim4, ajout de la catégorie Dim3). Finalement, tout l'échantillon (600 phrases) a été ré-annoté sur les 4 dimensions pour tenir compte des changements et fournir une référence à laquelle comparer l'annotation automatique. Cette version finale du schéma de codage est présentée dans la Table 2.

Table 2. Schéma de codage créé à partir de l'annotation manuelle

Dim1	**Type de question**	**Description**
Ree	Ré-expliquer / redéfinir	Demander de revenir sur un concept déjà expliqué
App	Approfondir un concept	Approfondir une connaissance, clarifier une ambiguïté ou demander plus de détails pour mieux comprendre
Ver	Validation / vérification	Vérifier ou valider une hypothèse
Dim2	**Modalité d'explication / Sujet de question**	**Description**
Exe	Exemple	Exemple d'application (cours/exercice)
Sch	Schéma	Schéma d'application ou explication sur ce dernier
Cor	Correction	Correction d'un exercice en cours/examen
Dim3	**Type d'explication**	**Description**
Def	Définir	Définir un concept ou un terme
Man	Manière (comment ?)	La manière comment procéder
Rai	Raison (pourquoi ?)	Demander la raison
Rol	Rôles (utilité ?)	Demander l'utilité / fonction
Lie	Lien entre concepts	Vérifier le lien entre deux concepts, le définir
Dim4	**Type de vérification (facultatif)**	**Description**
Err	Erreur / contradiction	Détecter une erreur/contradiction en cours ou dans l'explication de l'enseignant
Con	Connaissances du cours	Vérifier une connaissance
Exa	Examen	Vérifier une connaissance attendue à l'examen

(4) Finalement, dans l'**étape d'évaluation**, le dernier sous-échantillon de 200 segments a été annoté manuellement par les deux annotateurs experts (avec un kappa accru de 0,83 sur Dim1, 0,76 sur Dim2 et 0,47 sur Dim3). Ce sous-échantillon, non utilisé pour l'entrainement de l'annotateur automatique, a été utilisé pour son test.

4.2 Annotation automatique

Afin d'annoter l'ensemble de questions posées par les étudiants (et à terme pour l'utiliser en ligne pour analyser les questions collectées), un outil semi-automatique à base de règles et de mots clés pondérés manuellement a été utilisé antérieurement pour segmenter et annoter les questions automatiquement. Bien qu'efficace sur les questions qu'il annote (kappa élevé), une proportion importante de questions ne sont pas annotées par cet outil [10], ce qui nous amène à utiliser ici une annotation entièrement automatisée basée sur des techniques d'apprentissage automatique sur le corpus des questions (*cf.* Figure 2 pour un résumé des différentes étapes décrites ci-dessous).

La première étape a consisté à transformer les 923 segments annotés manuellement en vecteurs de mots. Tout d'abord, nous avons utilisé la version française de WordNet [17], base de données lexicale reliant des concepts sémantiques entre eux dans une ontologie selon une variété de relations sémantiques (tels que synonymes et hyperonymes). L'objectif est de ramener différentes expressions synonymes à une même expression dans les questions (ex : pour la valeur « Rai » dans dim 3, les mots synonymes « cause », « raison » et « motif » sont remplacés dans le texte par « pourquoi »), pour diminuer la diversité lexicale et renforcer certaines expressions pour le traitement qui suit. Nous avons effectué par la suite un ensemble de pré-traitements classiques sur le corpus de 923 segments : tokenisation, stemmatisation, suppression de ponctuation et de stopwords (mots creux non porteurs de sens), etc. Puis, nous avons extrait tous les unigrammes et bigrammes, avec une approche de type sac de mots, et compté leurs occurrences dans l'échantillon de 600 questions (723 segments) de l'étape de validation. Chaque segment est représenté par un vecteur de mots (nombre d'occurrences de chaque unigramme/bigramme (n-grammes) possible sur chaque segment). Le nombre de n-grammes étant très important par rapport au nombre de segments, nous avons réduit celui-ci pour conserver les mots-clés les plus importants et les plus significatifs en utilisant une technique de sélection d'attributs (suppression des n-grammes les moins fréquents et corrélés).

La deuxième étape a consisté à entraîner un classifieur pour automatiquement annoter chaque valeur de dimension (par ex. « réexpliquer »). Nous avons essayé 6 techniques de classification différentes tels qu'implémentées dans RapidMiner (Modèle linéaire généralisé, Gradient Boosted Trees, Arbre de décision, K-NN, Règle d'induction et Naïve Bayes) sur chaque dimension séparément (les dimensions étant conçues comme indépendantes). Chaque classifieur est entraîné en prenant en entrée un ensemble de vecteurs de mots représentant les 723 segments de l'ensemble d'entraînement, et l'étiquette à prédire est la valeur associée manuellement au segment dans cette dimension. Le modèle est ensuite évalué sur un échantillon indépendant de 200 segments sans étiquettes, afin d'assurer une bonne estimation de la performance

sur des données non vues. Enfin, nous avons calculé les valeurs Kappa entre les valeurs prédites par le classifieur et les valeurs correspondantes trouvées par l'annotation manuelle. Les meilleurs résultats ont été obtenus par l'algorithme Gradient Boosted Trees avec un Kappa moyen sur chaque dimension de 0,70 (*cf.* Table 3) – une valeur suffisamment élevée pour appliquer l'annotation automatique au corpus complet.

Table 3. Kappas obtenus des différentes techniques de classification utilisées

Dimension	Modèle linéaire généralisé (GLM)	Gradient Boosted Trees (GBT)	Naive Bayes	K-NN (k=2)	Arbre décision (C4.5)	Règle induction
Dim1	0.68	**0.70**	0.29	0.57	0.36	0.70
Dim2	0.17	0.77	0.10	0.43	**0.79**	0.37
Dim3	**0.69**	0.63	0.37	0.61	0.63	0.58
Dim4	0.62	**0.66**	0.38	0.60	0.13	0.66

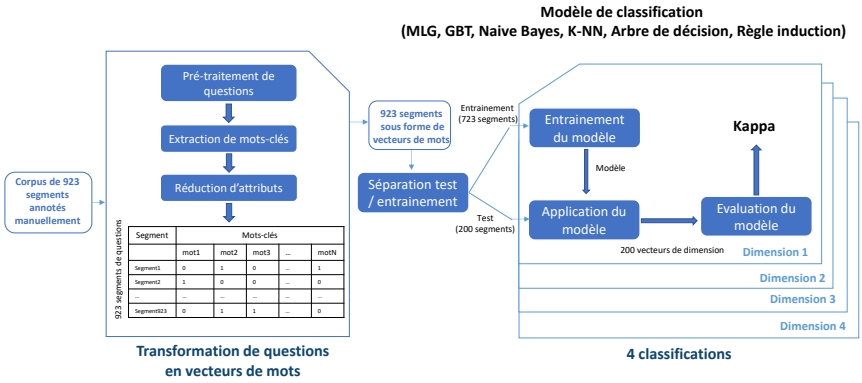


Fig. 2. Processus d'annotation à base d'apprentissage automatique

5 Différences de nature entre questions posées et votées

A l'issue de la section précédente, nous disposons donc d'un corpus complet de 6457 questions annoté automatiquement. Nous nous sommes concentrés sur quatre cours ayant généré le plus de questions (*cf.* Table 1) et considérés par les enseignants comme les plus difficiles : BCH, HBD, BCE et ANT. Les trois premiers ont lieu au premier semestre et ANT au second semestre. Nous n'avons pas fusionné les questions des différents cours, car des études antérieures [16, 18] sur ces données ont montré des différences significatives entre les cours. De plus, considérer les cours séparément permet de vérifier si des tendances similaires apparaissent d'un cours à l'autre.

Pour répondre aux questions de recherche, nous avons distingué 4 sous-populations sur chacun des cours considérés en fonction de l'activité des étudiants en distinguant : Q pour les étudiants « ayant posé au moins une question » (NQ sinon), et V pour les étudiants « ayant voté sur au moins une question » (NV sinon). En croisant les deux, cela donne donc les 4 sous-populations suivantes : QV, QNV, NQV et NQNV.

5.1 Impact du vote sur la performance

Méthode. Pour examiner la QR1, pour évaluer la performance, nous avons considéré pour chaque étudiant et sur chaque cours : (1) la note moyenne obtenue sur les QCM du cours (NotMoy, sur 20), qui peut donner une mesure à court terme de l'impact des questions posées, et la note finale obtenue au concours à cette matière (NotFin, sur 20), qui peut donner une mesure de l'effet à plus long terme.
Nous avons ensuite fait des comparaisons 2 à 2 entre QV et QNV (pour mesurer l'impact du vote chez les étudiants posant des questions) et entre NQV et NQNV (pour l'impact du vote chez les étudiants ne posant pas de questions), sur chacun des 4 cours, pour les deux variables. Pour ces 2 variables, nous avons utilisé des tests Mann-Whitney U au lieu de t-tests (les distributions ne suivant pas une loi normale). Nous avons effectué 2 fois 7 tests (2 comparaisons de population avec 4 cours et 2 variables à chaque fois, sauf NotFin, manquante pour BCE) et les seuils de significativité ont été corrigés par la méthode de Holm–Šidák pour éviter les erreurs de type I.

Résultats. En ce qui concerne les QV par rapport à QNV, 1 seul résultat significatif a été obtenu sur les cours du premier semestre (BCH, HBD, BCE) où parmi les étudiants ayant posé des questions, ceux n'ayant pas voté ont eu une meilleure note finale que les étudiants ayant voté (pour HBD). En revanche, pour ANT, en dépit d'effectifs plus réduits, les étudiants qui ont posé des questions et votés ont mieux réussi que ceux n'ayant fait que poser des questions, aussi bien aux QCM du cours (*cf.* Table 4) qu'au concours final (*cf.* Table 5).
En ce qui concerne les NQV par rapport à NQNV, 3 résultats significatifs ont également été obtenus (sur 7 tests) : pour BCH et HBD, les étudiants n'ayant pas posé de questions et n'ayant pas non plus voté ont eu une meilleure note sur les QCM que ceux ayant uniquement voté. Ce résultat se retrouve également au niveau du concours final pour HBD. Aucune différence n'a été observée au cours du deuxième semestre (ANT). En résumé, lorsqu'une différence a été observée, les élèves qui ont voté (sans poser de question) ont obtenu des résultats inférieurs, tant dans le cours que dans l'ensemble.

Table 4. Caractérisation des votants en termes de Note Moyenne (NotM) pour les étudiants qui posent des questions (QV vs. QNV) et ceux qui n'en posent pas (NQV vs. NQNV) et statistiques descriptives (quartiles et médiane) des 4 sous-populations sur chacun des 4 cours

	QV vs. QNV	NQV vs. NQNV	QV				QNV				NQV				NQNV			
	p	*p*	*N*	Q1	Md	Q3	*N*	Q1	Md	Q3	*N*	Q1	Md	Q3	*N*	Q1	Md	Q3
BCH	.049	**.001***	181	6.50	8.50	11.50	59	7.64	9.67	12.50	217	5.60	8.00	10.17	980	5.50	8.00	10.87
HBD	.176	**.000***	154	7.69	10.12	13.75	53	8.25	11.67	14.00	252	6.00	8.67	11.27	956	6.00	9.25	12.50
BCE	.128	.039	83	7.47	10.40	12.55	47	6.40	9.00	11.70	117	6.00	8.00	10.80	1133	5.00	7.67	10.75
ANT	**.001***	.005	42	11.20	13.80	15.35	46	6.43	10.20	13.35	23	8.57	10.00	12.70	968	5.00	8.00	11.80

* significatif avec $p < .05$ après correction de Holm-Šidák (.000 signifie $p < .001$)

Table 5. Idem Table 4 pour la Note Finale (NotF)

	QV vs. QNV	NQV vs. NQNV	QV				QNV				NQV				NQNV			
	p	*p*	*N*	Q1	Md	Q3	*N*	Q1	Md	Q3	*N*	Q1	Md	Q3	*N*	Q1	Md	Q3
BCH	.080	.015	168	5.25	8.25	11.50	56	6.31	10.38	13.31	197	4.25	6.50	9.75	880	3.75	6.75	10.50
HBD	**.002***	**.000***	144	6.50	9.25	12.12	44	8.19	11.50	13.06	229	4.50	7.25	10.50	886	4.00	7.75	10.94
BCE	-	-	83	N/A	N/A	N/A	49	N/A	N/A	N/A	118	N/A	N/A	N/A	1368	N/A	N/A	N/A
ANT	**.000***	.706	42	13	15.38	16.19	45	7.50	11.50	14.25	22	8.13	11.88	14.69	1116	3.5	7.00	12.00

* significatif avec $p < .05$ après correction de Holm-Šidák (.000 signifie $p < .001$)

5.2 Comparaison de nature de questions posées et votées

Méthode. Pour comparer la nature des questions que posent les étudiants à la nature des questions sur lesquelles ils votent, nous avons dû nous concentrer sur la population des étudiants qui font les deux (QV). Pour ces étudiants, sur chacun des 4 cours, nous avons considéré toutes les questions sur lesquelles ils ont voté pour calculer la proportion de chaque type de question votée dans chaque dimension. Par exemple, si dans BCH, un élève a voté sur une question de ré-explication et une autre de vérification (étiquetées [Ree,0,Sch,0] et [Ver,0,0,Con]), sur la dimension 1, il aurait voté à 50 % sur des questions de valeur "Ree" (ré-explication) et à 50% de ses questions de valeur "Ver" (validation). Ces proportions sont codées entre 0 et 1, de sorte que pour chaque étudiant, sur chaque cours, on obtient un *vecteur de vote* composé de 14 (3+3+5+5+3) valeurs comprises entre 0 et 1, et en suivant la même approche pour les questions posées, on peut également obtenir un *vecteur de question* posée de 14 valeurs.

Une fois le pré-traitement effectué, la comparaison des questions votées aux questions posées consistait à comparer pour chaque cours, pour chaque valeur d'une dimension (ex : la valeur "Ree" de la dimension 1), la distribution de la proportion des questions posées par les étudiants et celle de questions votées dans cette dimension. En d'autres termes, comparer deux distributions (non distribuées normalement) entre 0 et 1 pour la même population d'élèves, ce qui a été fait en effectuant 14 tests de Wilcoxon, en utilisant comme en 5.1 la méthode Holm-Šidák pour ajuster la valeur p critique.

Résultat. Un seul test sur 56 a révélé un résultat statistiquement significatif (*cf.* Table 6), répondant ainsi négativement à la QR2.

Table 6. Différences entre voter et poser une question selon la nature des questions

	Ree	App	Ver	Exe	Sch	Cor	Def	Man	Rai	Rol	Lie	Err	Con	Exa
BCH	.176	.427	.11	.31	.411	.017	.92	.000*	.236	.352	.295	.514	.259	.078
HBD	.382	.851	.717	.755	.809	.225	.093	.728	.007	.285	.003	.043	.941	.706
BCE	.476	.067	.015	.515	.723	.929	.89	.652	.797	.681	.118	.51	.686	.033
ANT	.826	.087	.551	.204	.795	.18	.076	.485	.212	.198	.039	.611	.691	.701

* significatif avec $p < .05$ après correction de Holm-Šidák (.000 signifie $p < .001$)

6 Discussion et conclusion

La comparaison des questions posées et votées par les étudiants révèle donc que, dans notre cas, **le vote ne semble pas être un bon substitut au fait de poser des questions**. En effet, quand un effet est mesuré, il est plutôt négatif pour les étudiants ayant voté, surtout lorsqu'ils n'ont pas posé de questions par ailleurs. Cette tendance semble cependant s'inverser plus tard dans l'année, où le fait de voter en complément de poser des questions entraîne de meilleurs résultats, tant aux QCM de cours qu'à l'examen final. Dans notre contexte, il semble donc que des activités « actives » (au sens de Chi – ici, voter) complémentaires à des activités « constructives » (ici, poser des questions) soient plus efficaces que des activités « constructives » seules, mais qu'une activité « active » seule soit plus négative en termes d'apprentissage que de la passivité. La différence entre le premier et le second semestre pourrait être liée au fait que les étudiants les plus en difficulté ont été forcés à arrêter à la fin du premier semestre, ainsi qu'à la baisse générale d'activité sur la plateforme : les étudiants qui continuent à y participer sont donc probablement plus motivés que la moyenne.
Le second résultat est que **pour les étudiants qui votent et posent des questions**, c'est-à-dire les plus impliqués et ceux qui réussissent le mieux (notamment au second semestre), il n'y a globalement **pas de différence de nature entre les questions sur lesquelles ils votent et celles qu'ils posent eux-mêmes**. Ce résultat est intéressant car il suggère une interprétation possible du résultat précédent, à savoir que les votes de ces étudiants correspondent à des questions qu'ils se posent vraiment. Il est possible que des étudiants se contentant de voter ne font pas l'effort de formuler leurs propres questions, et que s'ils le faisaient, elles seraient d'une nature différente.

Ce travail est exploratoire et présente donc plusieurs limites : même si tous les étudiants se connectent à la plateforme de questions, nous n'avons pas accès à des logs permettant de savoir s'ils ont vraiment lu les autres questions posées. Les étudiants qui se connectent en premier n'ont également pas de questions sur lesquelles ils peuvent voter, sauf s'ils se reconnectent par la suite pour voir les nouvelles questions. Une expérience davantage contrôlée dans laquelle les étudiants doivent poser des questions et/ou voter sur des questions précédemment posées après avoir vu une vidéo, pour pouvoir passer à la suite, permettrait de vérifier les interprétations précédentes. Néanmoins le travail préalable réalisé ici rend désormais possible ce type d'expérience. Dans notre contexte, ces résultats suggèrent qu'encourager les étudiants à formuler leurs questions, plutôt que de se contenter de voter sur les questions des autres, serait une stratégie positive pour l'apprentissage. En effet, il est possible que pour certains étudiants, voter donne le sentiment de faire ce qui est attendu d'eux, sans pour autant développer les stratégies métacognitives mises en jeu lorsqu'on se pose ses propres questions (identifier les concepts clés, tester sa compréhension, résumer ce qui a été appris…). Cela pourrait être fait en encourageant les étudiants à poser une question avant de pouvoir consulter celles des autres. Du point de vue des enseignants, cela signifie qu'il est d'autant plus critique de leur proposer une visualisation plus efficace que le « mur de questions » actuel, ce qui est une des perspectives de ce travail.
En résumé, voter est une bonne stratégie pour les étudiants sachant déjà formuler leurs propres questions. En revanche, pour ceux en difficulté, cela peut retarder la prise de conscience de leurs lacunes et leur capacité à les combler activement.

Remerciements : Nous remercions Pierre Gillois de nous avoir fourni les données.

References

1. Sullins, J., McNamara, D., Acuff, S., Neely, D., Hildebrand, E., Stewart, G., Hu, X.: Are you asking the right questions: The use of animated agents to teach learners to become better question askers. In International Florida Artificial Intelligence Research Society Conf., FLAIRS (2015).
2. Otero, J., Graesser, A.C.: PREG: Elements of a model of question asking. Cognition and instruction. 19, 143–175 (2001).
3. Chi, M.T.H.: Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science. 1, 73–105 (2009).
4. Chin, C., Kayalvizhi, G.: Posing Problems for Open Investigations: What questions do pupils ask? Research in Science & Technological Education. 20, 269–287 (2002).
5. Graesser, A.C., Person, N.K.: Question asking during tutoring. American educational research journal. 31, 104–137 (1994).
6. Li, H., Duan, Y., Clewley, D.N., Morgan, B., Graesser, A.C., Shaffer, D.W., Saucerman, J.: Question asking during collaborative problem solving in an online game environment. In: International Conference on Intelligent Tutoring Systems. pp. 617–618. Springer (2014).
7. Scardamalia, M., Bereiter, C.: Text-Based and Knowledge Based Questioning by Children. Cognition and Instruction. 9, 177–199 (1992). https://doi.org/10.1207/s1532690xci0903_1.
8. Supraja, S., Hartman, K., Tatinati, S., Khong, A.W.: Toward the Automatic Labeling of Course Questions for Ensuring their Alignment with Learning Outcomes. In: Proc. of The 10th International Conf. on Educational Data Mining. pp. 56–63. Xiangen Hu, Tiffany Barnes, Arnon Hershkovitz, Luc Paquette, Wuhan, China (2017).
9. Bloom, B.S., Engelhart, M.B.: Furst, EJ. Hill, WH, & Krathwohl, DR Taxonomy of educational objectives. The classification of educational goals. Handbook I: Cognitive domain. New York: Longmans Green. (1956).
10. Chin, C., Brown, D.E.: Student-generated questions: A meaningful aspect of learning in science. International Journal of Science Education. 24, 521–549 (2002).
11. Chin, C., Osborne, J.: Students' questions: a potential resource for teaching and learning science. Studies in science education. 44, 1–39 (2008).
12. Bihani, A., Ullman, J.D., Paepcke, A.: FAQtor: Automatic FAQ generation using online forums. In: International Conference on Educational Data Mining. pp. 529–532. Kristy Elizabeth Boyer, Michael Yudelson (2018).
13. Zeng, Z., Chaturvedi, S., Bhat, S.: Learner Affect Through the Looking Glass: Characterization and Detection of Confusion in Online Courses. In: International Conference on Educational Data Mining. pp. 272–277 (2017).
14. Jiang, Z., Zhang, Y., Liu, C., Li, X.: Influence Analysis by Heterogeneous Network in MOOC Forums: What Can We Discover?. In: Int. Conf. on Educ. Data Mining (2015).
15. Wong, J.-S., Pursel, B., Divinsky, A., Jansen, B.J.: An analysis of MOOC discussion forum interactions from the most active users. In: International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction. pp. 452–457. Springer (2015).
16. Harrak, F., Bouchet, F., Luengo, V., Gillois, P.: Profiling Students from Their Questions in a Blended Learning Environment. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. pp. 102–110. ACM, New York, NY, USA (2018).
17. Sagot, B., Fišer, D.: Building a free French wordnet from multilingual resources. In: OntoLex (2008).
18. Harrak, F., Bouchet, F., Luengo, V.: From Students' Questions to Students' Profiles in a Blended Learning Environment. 1. 6, 54—84 (2019).

Session
Analyse d'usage et ingénierie des EIAH

Processus d'évaluation longitudinale (PEL) d'une plateforme pédagogique (LMS) : le cas de LabNbook

Nadine Mandran[1], Patricia Marzin[2],[1], Maelle Planche[1], Aous Karoui[1], Isabelle Girault[1]

[1] Univ. Grenoble Alpes, CNRS, Grenoble INP*, LIG, F-38000 Grenoble France*
[2] Univ Brest, Univ Rennes, CREAD, F-29200 Brest, France

Résumé. Cet article présente d'une part les concepts pour construire un processus d'évaluation longitudinale (PEL) d'une plateforme d'apprentissage et d'autre part l'organisation de ce PEL. L'objectif de cette évaluation est d'inscrire ce PEL au niveau de la recherche, dans une posture constructiviste pragmatique pour travailler de manière itérative et incrémentale à partir du terrain. Le PEL est construit sur la base des outils de suivi de projet, en utilisant une démarche qualité qui garantit une traçabilité de la production des données et des résultats. Le cas d'étude choisi est LabNbook (https://labnbook.fr), qui est une plateforme pour l'apprentissage des sciences expérimentales.

Mots-clés. Evaluation, longitudinal, traçabilité, processus, qualité

Abstract. This article presents on the one hand the concepts to build a longitudinal evaluation process (LEP) of a learning platform in a research context and on the other hand the organization of this LEP. Evaluations' objectives are threefold: at the research level, at the technical level and at the societal level. At the research level, this LEP is inscribed in a pragmatic constructivism posture to work iteratively and incrementally from the field. At the level of project monitoring, the LEP is built on the basis of the tools of the quality approach which can guarantee a traceability of the production of data and results. The case study chosen is the LabNbook platform to learn experimental sciences in a context of higher education.

Keywords. Evaluation, longitudinal, traceability, process, quality

Introduction

Cet article présente des concepts et des outils pour concevoir un processus d'évaluation longitudinale (PEL) d'une plateforme d'apprentissage dans un contexte de recherche. L'évaluation d'une plateforme de support à l'activité d'apprentissage pour des apprenants ou des enseignants en situation réelle est complexe. Des facteurs techniques, humains, sociétaux peuvent influencer les résultats de cette évaluation [1].

Notre contribution se propose de répondre à la question suivante : quel processus d'évaluation peut-on mettre en place, en fonction des utilisateurs, de leurs rôles, du

* Institute of Engineering Univ. Grenoble Alpes

contexte d'enseignement sur plusieurs années et de nos questions de recherche ? Ce processus a été construit à partir des problématiques de recherche des didacticiens et avec l'objectif de vérifier que les objectifs du projet à évaluer ont été atteints. Pour qualifier ce processus d'évaluation longitudinale, nous définirons un cadre de recherche pour conduire ces évaluations, dans un état de l'art présenté ci-dessous.

2- Etat de l'art

Depuis l'apparition des premières plateformes de type e-learning, leurs méthodes d'évaluation ont évolué au fil des années. Ainsi, les approches utilisées au début, tenaient compte essentiellement de l'utilisabilité, en vérifiant le respect des normes d'utilisabilité théoriques ainsi que des heuristiques de Nielsen [2]. Certaines approches consistaient à se conformer à une check-list de normes ou de principes [3], sans pour autant tenir compte du contexte d'apprentissage, raison pour laquelle ces approches sont considérées comme réductrices. Dans un deuxième temps, il y a eu plus de synergie entre les recherches sur l'utilisabilité et celles sur les EIAH. Ainsi, des méthodes comme « l'apprentissage avec les heuristiques du logiciel » ont vu le jour [4]. Il s'agit d'une approche qui consiste à établir une liste de normes adaptées au contexte d'apprentissage. Similairement, Mehlenbacher et al. [5] mettent en place un ensemble de 21 heuristiques d'utilisabilité pour évaluer les plateformes d'apprentissage et les expériences des concepteurs. D'autres approches ont également été proposées, à l'image de l'évaluation systématique des plateformes e-learning (eLSE) [6]. Celle-ci est dérivée du SUE (Systematic Usability Evaluation) conçu à l'origine pour évaluer des systèmes hypermédias. En 2011, Granic et Cukisic [7] proposent une évaluation qui associe des tests-utilisateurs (relatifs à l'utilisabilité et à l'apprentissage) à des évaluations ergonomiques de la part d'experts en IHM et en EIAH. Aussi, dans une méta-étude publiée en 2014, sur 30 plateformes d'investigation, Donelly et al. mettent en avant une liste de critères pour identifier des plateformes d'investigation efficaces [8]. Parmi ces critères, les fonctionnalités d'implémentation fournies aux enseignants, l'impact sur l'apprentissage, le développement d'autonomie et de pratiques d'apprentissage métacognitives, et la mise en œuvre de la collaboration. Bien que ces critères puissent être spécifiques aux plateformes d'investigation, ils sont issus d'une comparaison des impacts d'apprentissage entre les plateformes analysées, négligeant tout de même les aspects d'utilisabilité.

Dans notre revue de l'état de l'art, nous n'avons pas trouvé de travaux relatant des processus d'évaluation, ni d'évaluations qui couvrent à la fois des objectifs de recherche et des objectifs opérationnels. Notre contribution propose de prendre en charge ces aspects.

3- Cas d'étude : le projet et la plateforme LabNbook

L'objectif de ce projet est d'évaluer le modèle de structuration de l'activité proposé et supporté par la plateforme LabNbook. Il s'agit également de pérenniser les

pratiques innovantes induites et de donner des recommandations pour des institutions. LabNbook est une plateforme pédagogique (Learning Management System) pensée pour des étudiants et des élèves. Elle leur propose un espace de travail partagé qui peut être utilisé comme cahier de laboratoire numérique et/ou comme support de rapports. LabNbook fournit des outils afin que des équipes pédagogiques structurent des situations et suivent la réalisation des activités par les étudiants. LabNbook a été testé en classe et a évolué à partir des résultats obtenus [9]. Dans le cadre du projet Idex-formation, LabNbook est actuellement utilisée par 25 enseignants et 1400 étudiants.

4- Processus d'évaluation longitudinale pour LabNbook (PEL)

4.1 Cadre scientifique et posture épistémologique

Cette recherche produit d'une part de la connaissance scientifique et d'autre part des outils pour accompagner l'activité humaine (p.ex., une interface, un modèle, une plateforme). Ces deux productions sont totalement entrelacées et interdépendantes. Pour construire la connaissance, toutes les sources d'observations et les méthodes de production des données sont mobilisées. Le but de la connaissance est d'élaborer des modèles intelligibles de l'expérience humaine, offrant des repères adaptés et viables. Ainsi, la connaissance prend la forme de modèles génériques activables. Nous retiendrons le Design Based Research (DBR), et la méthode THEDRE [10] (Traçability Human Experiment Design Research). L'objectif du DBR est de faire avancer de manière conjointe la recherche et les pratiques. Il définit trois niveaux de mesures : 1- le niveau d'adoption de la technologie par les enseignants : intérêt de l'enseignant pour la technologie, adaptation du domaine enseigné à la technologie, amélioration des pratiques de l'enseignant, intérêt de l'enseignant dans l'expérimentation; 2- le niveau d'institutionnalisation et la continuité de l'utilisation du dispositif d'enseignement sans technologie : coordination avec les décideurs, enthousiasme des élèves, apprentissage des étudiants et enthousiasme des enseignants ; 3- la continuité de l'utilisation du dispositif avec des technologies : l'enseignant structure son enseignement avec le dispositif, ensuite il utilise le dispositif pour sa classe. Le DBR correspond à notre besoin de prendre en compte le terrain et de construire une évaluation avec ces trois types de mesures. Nous retenons les trois niveaux de mesures du DBR, et le processus de THEDRE, ses outils de la démarche qualité (DQ) et la démarche centrée utilisateur (DCU) pour cadrer nos évaluations.

4.2 Concepts pour construire le processus

4.2.1-Acteurs et Rôles

Notre processus implique quatre acteurs pour l'évaluation de la plateforme et la production de la connaissance scientifique : le chercheur en didactique/EIAH, le développeur, le méthodologue, l'utilisateur. Nous distinguons quatre types

d'utilisateurs : 1- le responsable d'UE qui coordonne l'équipe enseignante qui utilise LabNbook; 2- les enseignants impliqués dans les UE. 3- les étudiants « témoins » choisis pour leur implication lors de l'utilisation de LabNbook, qui participent à des interviews pour connaitre leur avis et pour construire le questionnaire baromètre étudiants. 4- les étudiants utilisateurs de LabNbook qui sont mobilisés pour répondre aux questionnaires baromètres étudiants.

4.2.2. - Temporalité de l'évaluation

Notre objectif étant d'étudier un changement des pratiques au fil du temps, nous avons besoin de trois moments d'observation : avant, pendant et après l'utilisation de la plateforme.

4.2.3.- Méthodes de production de données et outils de mesures des SHS

Pour l'évaluation de LabNbook avec les utilisateurs, nous utilisons deux types de méthodes de production de données issues des sciences humaines et sociales (SHS). La démarche qualitative [11] : elle permet de comprendre et d'identifier des phénomènes. Le nombre d'utilisateurs consultés est faible (entre 6 et 20 personnes). En revanche, la variété des profils est importante pour observer le maximum de cas. Les outils de production des données sont des observations in-situ, des interviews ou des focus-groups. Les outils d'analyse de données sont de l'analyse thématique et de l'analyse du discours.

La démarche quantitative [11] : elle permet de quantifier des faits ou des avis. La sélection des utilisateurs repose soit sur une méthode d'échantillonnage (i.e. quotas, aléatoire) soit sur un recensement. Dans ce cas l'ensemble des utilisateurs est consulté. Les outils de production des données sont des questionnaires ou des traces d'activités. Les outils d'analyse des données sont des analyses statistiques ou des visualisations.

Les outils de mesures sont les dispositifs qui permettent le recueil des données. Dans notre PEL, pour la démarche qualitative nous utilisons les outils de mesures suivants : 1/Entretien semi-directif : technique d'investigation pour interroger en face à face une personne, à l'aide d'un guide d'entretien ; 2/Carnet de bord : outil qui va enregistrer chronologiquement les différents événements survenus lors de l'utilisation d'une application [12] ; 3/ Observation de l'activité : outil pour étudier les comportements de l'utilisateur par l'observation de ce dernier en situation réelle, dans un contexte d'utilisation, de travail [13]. Dans le cadre de LabNbook, il s'agit d'étudier les missions proposées aux étudiants par les enseignants. L'objectif est de comprendre comment ils ont construit leur enseignement sur la plateforme. Pour la démarche quantitative, nous utilisons : 1/Questionnaire baromètre : outil qui produit des données à partir des réponses des utilisateurs ou des consommateurs (attentes, satisfaction, etc) à intervalles de temps réguliers. Il va permettre de réaliser des comparaisons et donc de percevoir une amélioration, une stagnation ou une détérioration de la situation [13]. Ce questionnaire contient des mesures standard de satisfaction, d'accord et il intègre des questions pour mesurer si les enseignants mettent en place des pratiques de « Student-Centered Learning » [14] 2/Les traces d'activité : outil qui capture les actions réalisées par l'utilisateur avec une application

informatique. Elles portent une dimension temporelle car la capture enregistre la date et l'heure de l'action.

4.3. Processus d'évaluation longitudinale (PEL)

4.3.1 Organisation du processus

Notre évaluation est longitudinale, elle va se répéter au cours du temps. Par exemple pour connaître le changement de pratiques des enseignants nous mesurons les pratiques avant l'utilisation (T0) de LabNbook et après l'utilisation (T+1, etc). Pour représenter ce processus d'évaluation longitudinal, nous proposons une cartographie (Fig. 1). Pour chaque type d'utilisateur, pour chaque période d'utilisation de la plateforme nous indiquons les objectifs expérimentaux et/ou les outils de mesures (seule la cartographie des objectifs est présentée). Nous proposons des questionnaires baromètres pour les enseignants. Ce questionnaire leur est administré avant la première utilisation de LabNbook pour connaître leur perception a priori et à la fin de chaque semestre. L'objectif est de suivre l'évolution de leurs pratiques dans le temps. Un autre questionnaire est utilisé avec les étudiants.

4.3.2 Mise en œuvre et résultats du PEL sur LabNbook

Le PEL actuel a nécessité plusieurs versions et il est actuellement en cours d'utilisation pour l'évaluation de LabNbook. L'ensemble des données recueillies, a permis de faire évoluer la plateforme LabNbook et de rendre des comptes aux financeurs. Nous avons réalisé 21 entretiens, recueillis 91 questionnaires baromètres enseignants et 600 questionnaires étudiants. Au niveau conceptuel, les retours d'utilisation de la plateforme ont permis d'identifier un problème de compréhension des concepts de cahier de laboratoire ou de rapports d'étudiants.

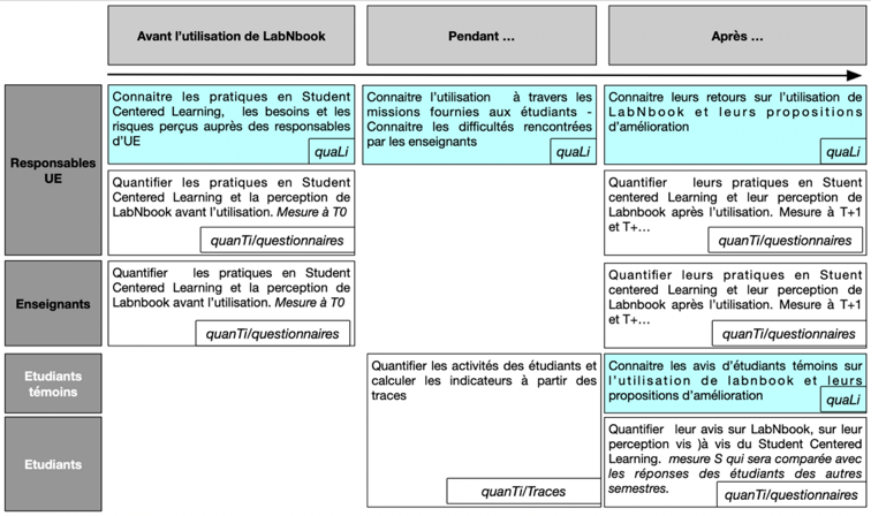


Fig. 1. Processus d'évaluation longitudinale de la plateforme LabNbook

5 Conclusions, limites et perspectives

La structuration en temporalité/acteurs (Fig. 1) offre une vision globale et structurante de l'évaluation longitudinale. Ce processus a été construit sur les principes du DBR : évaluer selon trois niveaux à partir du terrain en utilisant des outils de mesures qui se complètent pour répondre aux questions de recherche. Une difficulté rencontrée est le maintien du nombre de personnes enquêtées d'une année à l'autre. Il faut pour cela envisager des moyens de motivations des étudiants mais aussi des enseignants pour répondre aux questionnaires. Des outils d'analyse systématiques doivent également être proposés pour compléter ce PEL.

Remerciements. Idex-Formation de la COMUE Grenoble-Alpes et « Appels à Manifestations d'Intérêt » du Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

Références

[1] F. Jambon, N. Mandran, B. Meillon, and C. Perrot, 'Évaluation des systèmes mobiles et ubiquitaires : proposition de méthodologie et retours d'expérience', in Ergo-IA'08 'L'humain au cœur des systèmes et de leur développement', Biarritz, 2008, pp. 107–116.
[2] J. Nielsen, 'Usability methods', Usability Inspection Methods, John Wiley, New York, pp. 178–249, 1994.
[3] O. Parlangeli, E. Marchigiani, and S. Bagnara, 'Multimedia systems in distance education: effects of usability on learning', Interacting with computers, vol. 12, no. 1, pp. 37–49, 1999.
[4] D. Squires and J. Preece, 'Predicting quality in educational software', Interacting with computers, vol. 11, no. 5, pp. 467–483, 1999.
[5] B. Mehlenbacher et al., 'Usable e-learning: A conceptual model for evaluation and design', in Proceedings of HCI International, 2005, vol. 2005, p. 11th.
[6] R. Lanzilotti, C. Ardito, M. F. Costabile, and A. De Angeli, 'eLSE methodology: A systematic approach to the e-learning systems evaluation.', Journal of Educational Technology & Society, vol. 9, no. 4, 2006.
[7] A. Granić and M. Ćukušić, 'Usability Testing and Expert Inspections Complemented by Educational Evaluation: A Case Study of an e-Learning Platform', Journal of Educational Technology & Society, vol. 14, no. 2, pp. 107–123, 2011.
[8] D. F. Donnelly, M. C. Linn, and S. Ludvigsen, 'Impacts and characteristics of computer-based science inquiry learning environments for precollege students', Review of Educational Research, vol. 84, no. 4, pp. 572–608, 2014.
[9] C. Bonnat, P. Marzin, and I. Girault, 'Analyse des conceptions d'élèves sur le vivant, dans une situation de conception expérimentale avec un environnement informatique.' Recherche en Didactique des sciences et des technologies, 2018.
[10] N. Mandran and S. Dupuy-Chessa, 'THEDRE: a Traceable Process for High Quality in Human Centred Computer Science Research', 26th ISD, Larnaca-Chypre, 2017.
[11] J. W. Creswell, Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications, 2013.
[12] R. Bernhaupt, 'Usability and User Experience Evaluation Methods', Mass Customization for Personalized Communication Environments: Integrating Human Factors: Integrating Human Factors, p. 232, 2009.
[13] J. Copans, L'enquête ethnologique de terrain (Ethnological field investigation). Armand Colin, 2011.
[14] M. J. Hannafin and K. M. Hannafin, 'Cognition and student-centered, web- based learning: Issues and implications for research and theory', in Learning and instruction in the digital age, Springer, 2010, pp. 11–23.

Contribution de différents outils de mesure à l'évaluation des usages d'une plateforme numérique par un processus longitudinal : cas du travail à distance

Maëlle Planche[1], Isabelle Girault[1], Nadine Mandran[1], Patricia Marzin[2], Cédric d'Ham[1], Claire Wajeman[1]

[1]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, F-38000 Grenoble France [1]Univ Brest, [2]Univ Rennes, CREAD, F-29200 Brest, France

Résumé. Cet article explore la contribution d'un processus d'évaluation longitudinal élaboré pour comprendre les usages d'une plateforme pédagogique spécifiquement conçue pour l'enseignement scientifique : la plateforme LabNbook. En prenant comme exemple le travail à distance, nous présentons quatre outils de mesures mobilisés en vue d'évaluer les pratiques des enseignants et leurs évolutions dans le temps. Ainsi, nous montrons que ces outils s'avèrent complémentaires, à la fois pour le développement des connaissances sur les pédagogies actives et pour l'amélioration des performances de la plateforme.

Mots-clés. Evaluation longitudinale ; Méthodologie ; LabNbook ; Plateforme LMS ; Pédagogie active

Abstract. This article explores the contribution of a longitudinal evaluation process developed to understand the uses of a LMS specifically designed for scientific teaching: the LabNbook platform. Through a case study, we highlight the contribution of several measurement tools to the evaluation of teachers practice and their evolutions when using LabNbook. This article shows how each of these tools turns out to be essential to jointly develop the knowledge about Student Centered Learning and the improvement of the platform performances.

Keywords. Longitudinal evaluation; Methodology; LabNbook; LMS platform; Active Learning

1 Introduction

Dans l'enseignement scientifique, une place importante est donnée aux « pédagogies actives » qui favorisent l'autonomie des étudiants et leur implication dans le processus d'apprentissage. Des études rigoureuses ont montré la supériorité de ces pédagogies, que ce soit en termes d'apprentissage conceptuel [1] ou d'acquisition de nouvelles compétences scientifiques [2]. La mise en place de ces pédagogies implique des tâches nouvelles pour les étudiants comme pour les enseignants. Les étudiants doivent devenir autonomes dans la production de solutions ; ils travaillent généralement en collaboration lors de séquences impliquant une activité en présentiel et/ou à distance, synchrone ou asynchrone. Les enseignants doivent suivre, structurer et orienter le travail des étudiants tout en évitant d'imposer une méthode de résolution préétablie. Toutes ces tâches accroissent la complexité des séquences pédagogiques et font émerger le besoin d'un support adapté. C'est dans ce contexte de transformation de l'enseignement scientifique que la plateforme LabNbook a vu le jour. Elle est le fruit de recherches sur la conception d'environnements informatiques supports aux

pédagogies actives en sciences menées par les enseignants-chercheurs en didactique des sciences expérimentales et intègre des résultats de leurs travaux en didactique des sciences. En 2016, LabNbook a obtenu un financement Idex-Formation de la COMUE Université Grenoble Alpes, dans le but de promouvoir les pédagogies actives au sein de ses établissements, de pérenniser l'utilisation de la plateforme et sa dissémination au niveau national et international. Ce projet comporte un volet d'évaluation conséquent dont l'objectif principal est de comprendre les usages et l'appropriation de la plateforme par des enseignants et des étudiants en sciences expérimentales. L'évaluation est conduite dans le temps, c'est pourquoi nous utilisons le terme de longitudinale. Cet article explore la contribution de ce processus et vise à montrer comment chaque outil de mesure développé participe à la compréhension des transformations des pratiques enseignantes et étudiantes avec l'utilisation de LabNbook. Nous allons nous centrer sur un des indicateurs permettant de mesurer cette transformation : la place du travail à distance dans les enseignements supportés par la plateforme.

2 Le volet scientifique du projet LabNbook

Dans cette partie, nous commençons par présenter l'état de l'art, la problématique didactique générale ainsi que les différentes questions de recherches qui sous-tendent le volet scientifique du projet. Nous présenterons ensuite la plateforme LabNbook, puis nous terminerons cette partie en présentant le processus d'évaluation longitudinal (PEL) élaboré à la fois dans une optique de recherche en didactique, de développement continu et de déploiement de la plateforme.

2.1 Contexte, état de l'art, problématique et questions de recherche

De l'idée de « pédagogies actives » mentionnée dans l'introduction, nous retenons « l'apprentissage centré sur l'apprenant » [3]. À partir des travaux de Freeman et al. [1] et de Hannafin [3] ainsi que d'un glossaire [4], nous proposons des indicateurs permettant d'évaluer le degré de centration sur l'apprenant des pédagogies mises en œuvre dans les situations d'enseignements étudiées : 1/ Déléguer les modalités de résolution de problèmes aux étudiants ; 2/ Expliciter les objectifs d'apprentissage et les compétences visées ; 3/ Donner aux élèves la possibilité d'évaluer leurs apprentissages en cours d'activité ; 4/ Proposer une organisation qui permette de travailler en dehors de la classe traditionnelle ; 5/ Encourager le travail collaboratif. Ces pédagogies peuvent être supportées par des environnements informatiques, tels que LabNbook, appelés « Student Centered Learning Environnements » (SCLE) [5].

La problématique didactique générale qui sous-tend le volet scientifique du projet, vise à étudier la transformation des pratiques enseignantes et étudiantes médiées par LabNbook. Notre recherche explore principalement deux aspects des usages de la plateforme : 1/ Les facteurs qui prévalent à l'appropriation de LabNbook par les enseignants ; 2/ la transformation des pédagogies mises en place vers des pédagogies centrées sur l'apprenant, impulsée et/ou soutenue par LabNbook. Pour répondre à ces questions, plusieurs types de mesures sont effectuées, comme par exemple sur le travail collaboratif au sein de LabNbook ou encore sur l'accompagnement individuel des étudiants. Afin d'illustrer la contribution du PEL et des différents outils de mesure élaborés, nous avons choisi de nous centrer ici sur le travail à distance médié par LabNbook, c'est-à-dire l'ensemble des tâches réalisées par les étudiants sur la plateforme en dehors des heures d'enseignement, à l'université ou ailleurs. Il peut être individuel ou collectif.

2.2 Présentation de la plateforme LabNbook

LabNbook est une plateforme pédagogique en ligne spécifiquement conçue pour l'enseignement scientifique, grâce à des outils qui permettent de produire des protocoles expérimentaux et des tableaux de données. Pensée pour l'enseignement du supérieur et du secondaire, elle offre aux étudiants un espace de travail partagé pour leurs travaux collaboratifs et peut être utilisée comme cahier de laboratoire en ligne et/ou comme support d'écriture collaborative de rapports. Elle a été conçue comme support pour aider les enseignants et les apprenants à mettre en œuvre des démarches d'investigation scientifique. Sa conception s'appuie sur le constat que les tâches confiées aux élèves sont complexes et peuvent engendrer des difficultés. Il y a donc nécessité de mettre à disposition des enseignants et des étudiants des outils et des ressources qui permettent la mise en place d'étayages réduisant la complexité de la tâche et/ou favorisant l'accompagnement des étudiants dans l'accomplissement de la tâche [6]. Dans LabNbook, une « mission » désigne une activité pédagogique dévolue à des équipes d'étudiants, correspondant à un espace de travail pré-structuré par l'enseignant pour réduire la complexité de la tâche. Cet espace est composé de parties éventuellement accompagnées de consignes de travail. La plateforme offre des outils supports pour la collaboration entre étudiants (porte-document partagé, messagerie interne et fils de discussion) et des outils supports des rétroactions des enseignants vers les étudiants (tableau de bord, espace de visualisation, outil d'annotation).

2.3 Processus d'évaluation longitudinale pour LabNbook (PEL)

Nous présentons ici une synthèse des grands principes du PEL élaboré sur la base de la méthode Thèdre [7]. Un travail en étroite collaboration avec des enseignants-utilisateurs est réalisé afin d'explorer l'impact de la plateforme sur leurs pratiques et sur celles des étudiants. L'originalité de cette évaluation repose : 1/ sur sa dimension longitudinale qui implique d'intégrer l'évolution de différents types de mesures au fur et à mesure de l'utilisation de la plateforme ; 2/ sur l'association de différents points de mesures pour comprendre les attentes des enseignants, les réponses que la plateforme apporte à leurs besoins, leurs usages et plus généralement, les transformations de leurs pratiques pédagogiques médiées par LabNbook. Le PEL mobilise à la fois des mesures de type quantitatif et qualitatif réalisées sur des données de nature factuelle (*i.e.* qui portent sur les pratiques et les usages réels de la plateforme) ou de nature narrative ou déclarative (*i.e.* qui renvoient à la manière dont les utilisateurs justifient leurs usages, à leurs motivations ou leurs réticences vis-à-vis de LabNbook).

Table 1. Outils de mesure du PEL organisés par type de traitement et nature des données

	Mesure quantitative	**Mesure qualitative**
Données de nature factuelle	Traces d'activité*	Analyse des missions* Carnets de bord Débriefings
Données de nature narrative	Questionnaires baromètres* Questionnaire de satisfaction	Entretiens semi-directifs*

* correspond aux outils de mesure pris en compte dans la quatrième partie de cet article

L'évaluation longitudinale se déroule sur trois années d'enseignement (2017-2020). Selon les outils mobilisés, les mesures sont réalisées sur une ou plusieurs

périodes : avant enseignement, après enseignement lors de la première année d'enseignement, de la deuxième et de la troisième année d'enseignement. Pour la première année universitaire (2017-2018), l'évaluation a concerné 112 enseignants et 2000 étudiants utilisateurs sur les établissements de la COMUE Université Grenoble Alpes.

3 Contributions des différents outils de mesure pour la recherche sur le travail à distance avec LabNbook

Parmi l'ensemble des outils de mesure déployés dans le cadre du PEL, quatre sont mobilisés pour fournir des informations sur le travail à distance des étudiants médié par LabNbook : les entretiens semi-directifs, les questionnaires baromètres, des traces d'activité informatiques et l'analyse des missions

3.1 Le questionnaire-baromètre

Un questionnaire-baromètre (questionnaire « classique » soumis de façon répétée et à intervalles réguliers à une cohorte d'utilisateurs) est proposé aux enseignants-utilisateurs de LabNbook tout au long du processus d'évaluation afin de mesurer l'évolution de leurs attentes, de leurs usages et de leur satisfaction à mesure qu'ils utilisent et s'approprient l'environnement informatique. Les données recueillies peuvent être considérées comme de nature « narrative » et nous permettent d'opérer des traitements de type quantitatif. Nous exploitons ici les données obtenues aux vagues 1 (avant l'utilisation avec LabNbook) et 2 (après un semestre d'utilisation). Deux questions du questionnaire-baromètre, qui relèvent à la fois des attentes et des pratiques pédagogiques des enseignants, contribuent à fournir des éléments de réponse à la question du travail distant : 1/ LabNbook facilite(ra) le travail des étudiants entre les séances ; 2/ Pour cette formation, je demande aux étudiants du travail en dehors des séances.

Les résultats obtenus témoignent d'une satisfaction générale des enseignants-utilisateurs ayant répondu au questionnaire sur les apports de la plateforme en matière de travail distant. Les évolutions de ces données permettent également de pointer certaines tendances concernant « la satisfaction à l'usage », à mesure qu'ils se familiarisent avec l'outil. Ainsi, concernant la première question, les enseignants ne semblent pas déçus : après un semestre d'utilisation, 73 % d'entre eux estiment que LabNbook facilite effectivement le travail des étudiants entre les séances (vs 72% en vague 1). De même, l'intégralité des enseignants interrogés en vague 2 dit demander du travail aux étudiants en dehors des séances, contre 96 % en vague 1. Ces données offrent une vision globale des attentes des enseignants et de leur satisfaction en matière de travail distant après une première phase d'appropriation de l'outil. Mais elles demeurent limitées de par leur seule dimension « narrative » : elles nous disent peu de choses des usages réels de la plateforme et des transformations pédagogiques en lien avec son utilisation.

3.2 Apports des traces informatiques d'activité

Avec les traces numériques d'activités des enseignants et des étudiants nous pouvons explorer les évolutions de ces usages réels de la plateforme. Elles permettent principalement d'opérer des traitements de type quantitatifs et d'obtenir des données

de nature factuelle. Elles apportent donc des informations qui complètent celles obtenues avec le questionnaire-baromètre. Les traces que nous pouvons aujourd'hui mobiliser pour répondre à la question qui nous intéresse portent sur les activités réalisées par des étudiants de huit unités d'enseignements (UE) de sciences expérimentales ayant utilisé la plateforme deux années consécutives. Deux types de mesures sont susceptibles de fournir des informations sur le travail à distance médié par LabNbook :

- **Nb_total_actions :** nombre total de traces correspondant à des ajouts, duplications ou éditions de documents par tous les étudiants d'une mission
- **Nb_actions_SW**: cette mesure donne des indications sur une partie des actions faites « à distance » parmi l'ensemble des actions mesurées ci-dessus : celles réalisées par les étudiants les soirs et les weekends (SW). Il n'est pas possible, en effet, d'identifier parmi celles effectuées en journée, celles qui sont réalisées sur le temps libre des étudiants, à l'université ou ailleurs. Les mesures présentées ici sont donc certainement très en-deçà du travail distant réel.

Pour cet article, nous nous centrons sur l'évolution entre l'année 1 et 2 de la part d'actions_SW réalisées parmi l'ensemble de ces actions. Les données recueillies montrent des différences significatives entre UE dans les proportions d'actions_SW réalisées par les étudiants (de 6 % à 29 % selon les UE) et les évolutions de ces proportions entre les deux années : trois UE ont vu cette proportion augmenter et cinq présentent la tendance inverse.

Ce sont les mesures qualitatives (attentes des enseignants, analyse des missions) qui permettront le mieux d'expliquer ces différences entre UE : par exemple, les attentes des enseignants des UE dans lesquelles les actions_SW sont les moins nombreuses, pourraient davantage porter sur le travail collaboratif que sur un travail individuel de préparation du TP. A contrario, les UE dans lesquelles elles ont augmenté sont peut-être celles pour lesquelles le scénario pédagogique proposé a évolué avec l'utilisation de la plateforme, en incluant par exemple une partie de préparation du TP en amont de la séance. Si les données issues des questionnaires baromètres et des traces informatiques fournissent des éléments de cadrage indispensables à la description et à la compréhension des usages de la plateforme, elles n'ont de sens que rapportées au contexte pédagogique dans lequel elles s'inscrivent et aux situations pédagogiques mises en place.

3.3 Les entretiens semi-directifs

Des entretiens de type semi-directifs ont été réalisés avant utilisation de la plateforme avec les responsables d'UE impliqués dans le projet afin d'identifier leurs attentes et motivations à utiliser LabNbook dans le cadre de leur enseignement. En fin de projet, ils seront de nouveau interrogés pour comprendre leur satisfaction à l'issue de leur expérience et la manière dont ils ont utilisé LabNbook. Des liens seront faits avec les situations pédagogiques observées dans les missions.

Les entretiens fournissent des éléments de compréhension beaucoup plus précis, favorisant l'interprétation des données observées dans les traces et les questionnaires. Ainsi, les attentes en matière de travail à distance avec LabNbook ressorties des entretiens initiaux révèlent des objectifs pédagogiques variables d'un enseignant à l'autre : pour certains, il s'agit de responsabiliser les étudiants en les contraignant à devenir plus actifs dans leur processus d'apprentissage ; pour d'autres, il y a nécessité de « gagner du temps » en les incitant à préparer le TP en amont. Pour d'autres encore, cette question renvoie plutôt au temps de maturation nécessaire à la compréhension des phénomènes observés et plus généralement à « *la construction de sens par les étudiants* » [8].

3.4 L'analyse des missions des enseignants

Dans le cadre du PEL, nous portons une attention particulière au contenu des missions que les enseignants proposent à leurs étudiants dans LabNbook : elles fournissent des données qualitatives de nature factuelle qui complètent celles obtenues avec les entretiens (notamment les consignes et autres contenus donnés hors LabNbook). Cette analyse permettra de saisir à la fois les points communs et les différences entre les situations pédagogiques mises en place avec LabNbook chaque année.

Lorsque LabNbook est utilisé comme support de formation à la démarche expérimentale, plusieurs tâches peuvent être réalisées à distance par les étudiants : proposer des hypothèses, définir le problème, concevoir l'expérience, interpréter les résultats, diffuser et/ou traiter les résultats. À partir des consignes données par l'enseignant, il devient ainsi possible d'identifier celles réalisables à distance. Ainsi, pour chaque situation pédagogique proposée, nous regardons la nature du travail demandé aux étudiants et sa place dans le processus pédagogique : préparer le TP, refaire le travail, diffuser ses résultats, etc. À la fin de la seconde année d'expérimentation de LabNbook, ces missions seront de nouveau analysées pour comprendre si la place et la nature du travail à distance avec LabNbook ont évolué.

Conclusion

Le cas du travail à distance a permis d'illustrer la contribution du PEL et des différents outils de mesure élaborés pour évaluer les transformations pédagogiques impulsées et/ou soutenues par LabNbook. C'est l'association de ces différentes mesures et la prise en compte de leur évolution dans le temps qui permet le mieux de comprendre et d'expliquer les usages, les attentes, la satisfaction et les motivations des utilisateurs à continuer à utiliser la plateforme. Ainsi, l'analyse de l'évolution des missions, couplée aux entretiens de satisfaction en fin de projet, fourniront des informations riches et complémentaires pour expliquer les tendances observées dans l'évolution des traces d'activité et des réponses apportées aux questionnaire-baromètre.

Références

1. Freeman, S., et al.: Active learning increases student performance in science, engineering, and mathematics. PNAS 111, vol. 23 (2014) 8410–8415.
2. Etkina E. et al.: Design and Reflection Help Students Develop Scientific Abilities: Learning in Introductory Physics Laboratories. *J. of the Learning Sciences* 19, vol. 1 (2010) 54–98.
3. Hannafin, M.J.: Student-centered learning. Chapter, Encyclopedia of the Sciences of Learning. Springer US (2012) 3211–3214.
4. http://edglossary.org/student-centered-learning/
5. Land, S.M, Hannafin, M.J. & Oliver, K.M.: Student-Centered Learning Environments: Foundations, Assumptions, and Design. Theoretical Foundations of Learning Environments, Chapter 1, Taylor & Francis (2012).
6. Zacharia, Z.C. et al. .E.: Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: a literature review. Educational Technology Research and Development 63, vol. 2 (2015) 257–302.
7. Mandran, N.:. London : ISTE Editions (2018).
8. Tiberghien, A., Vince, J., Gaidioz, P.: Design-based Research: Case of a teaching sequence on mechanics. International Journal of Science Education. 31, 2275–2314 (2009).

Modèles et interactions pour générer des événements dynamiques horodatés : le cas de la formation professionnelle à l'organisation des soins infirmiers

Catherine Pons Lelardeux[1], Michel Galaup[2], Herve Pingaud[3], Arthur Ramolet[4], Catherine Mercadier[5], and Pierre Lagarrigue[6]

[1] IRIT, Université de Toulouse (France), INU Champollion, Serious Game Research Network
catherine.lelardeux@univ-jfc.fr
[2] EFTS, Université de Toulouse (France), INU Champollion, Serious Game Research Network michel.galaup@univ-jfc.fr
[3] LGC, Université de Toulouse (France), INU Champollion, Serious Game Research Network herve.pingaud@univ-jfc.fr
[4] SGRL, Université de Toulouse (France), INU Champollion, Serious Game Research Network arthur.ramolet@univ-jfc.fr
[5] Agence Régionale de Santé Occitanie, GIP E-Sante (France)
catherine.mercadier@esante-occitanie.fr
[6] ICA, University of Toulouse (France), INU Champollion, Serious Game Research Network pierre.lagarrigue@univ-jfc.fr

Résumé. Notre travail est motivé par le constat des limites pour dispenser aux élèves-infirmiers des enseignements en organisation du travail avec des méthodes pédagogiques traditionnelles. Le travail présenté dans cet article a pour objectif de permettre une confrontation à l'activité professionnelle dans un service hospitalier virtuel de 14 lits. Nous nous intéressons à la modélisation des situations dynamiques dans cet environnement sociotechnique. Nous proposons un modèle et des interactions utiles à la génération automatique d'événements dynamiques horodatés. Cela permet de recréer de manière artificielle la dynamique du contexte et les échanges pluriprofessionnels.

Mots-clé: Scénario interactif, modélisation, scénario pédagogique, jeu de simulation, environnement virtuel

Abstract. It seems to be hardly possible to train future nurses to organize their work in a real-life professional environment. Therefore, we propose to design a virtual medical unit dedicated to training which simulates the professional context from 1 to 14 patients. In this paper, we particularly focus on modeling dynamic and professional activities. We propose a model and the interactions which allow to control and launch events at a specific time period to artificially recreate the dynamic of the context and the interprofessional exchanges.

Keywords: Educational scenario, virtual environment, sociotechnical system, nursing training

1 Contexte et problématique

Aujourd'hui, les personnels infirmiers évoluent dans des contextes de travail variés et leurs pratiques professionnelles font évoluer les modalités d'organisation du travail et donc de la formation au métier. En 2017, un groupe de travail, piloté par l'Agence Régionale de Santé (ARS) en Occitanie et composé de formateurs de 11 Instituts de Formation en Soins Infirmiers (IFSI) a montré qu'un jeune diplômé est formé pour prendre en charge 5-6 patients, alors qu'il devra prendre en charge une douzaine de patients en réalité. De plus, il est souligné que même en stage, les activités afférentes aux soins, la délégation de tâches aux aide-soignants (AS) et la gestion des imprévus ne sont pas confiées à l'élève-infirmer. Ces activités représentent 50 pour cent du travail d'un Infirmier Diplômé d'Etat (IDE). Par ailleurs, il est important de noter qu'un infirmier est très fréquemment interrompu 6 à 7 interruptions/heure; Ces interruptions ont un impact sur la charge de travail. L'impossibilité de confronter les étudiants même en fin de formation à la gestion de l'activité réelle nous conduit à proposer une confrontation aux réalités professionnelles dans un environnement virtuel (EV). Cet environnement, intitulé Clinical Organizer Nurse Education (CLONE), devrait permettre de simuler le contexte professionnel, d'acquérir l'ensemble des compétences nécessaires à l'organisation du travail et de leur offrir la possibilité de situer leurs interventions au sein de l'équipe interprofessionnelle : aides-soignants, professionnels médicaux et paramédicaux. Pour cela, il est nécessaire de pouvoir contrôler l'ensemble des événements et interactions possibles avec l'ensemble des acteurs du système et d'avoir la capacité de déclencher de manière artificielle des événements horodatés tels que l'arrivée du médecin, la distribution de repas etc. L'objectif consiste à maintenir une certaine liberté d'action aux utilisateurs tout en respectant les règles métiers régissant l'activité professionnelle, la sémantique des informations métiers, les temporalités liées aux pratiques professionnelles.

2 Modélisation de scenarii interactifs non-linéaires

2.1 Travaux antérieurs et travaux connexes

Scénario scripté Lorsqu'il s'agit de transmissions de connaissances techniques individuelles, l'utilisation d'ontologies ou de graphes orientés apporte des solutions adaptées pour décrire un scénario procédural linéaire. Mais lorsque la quantité de connaissances à manipuler est importante, l'explosion combinatoire rend impossible la description de tels scenarii. Dans le contexte étudié, le volume de connaissances à manipuler est relativement important (plus de 40 actions par patient). **Logique linéaire** La modélisation de scénario non-linéaire peut s'appréhender par l'approche générative : la logique traditionnelle ou la logique linéaire [4]. La logique linéaire définit un cadre conceptuel pour spécifier des

scénarios interactifs. Elle fournit un cadre sémantique à la modélisation de la causalité, du raisonnement automatique ainsi que des mécanismes d'allocation de ressources pour une histoire [5]. Elle offre des mécanismes de déductions rigoureux et automatiques qui permettent d'assurer le contrôle de la gestion du déroulement du jeu et notamment du respect du souhait des auteurs tout en garantissant la liberté d'actions des joueurs [1]. Pour cela, le personnage-joueur est utilisé comme un agent et les actions de jeu comme des éléments de base. Ici, seules les actions de jeu sont prises en compte, mais pas la connaissance ou la partage d'information par les joueurs. En outre, la dimension temporelle n'est pas prise en compte et les modèles ne permettent pas de décrire le déclenchement d'événements indépendants ou la survenue d'interruptions horodatées. **Logique temporelle** Une autre approche consiste à utiliser la logique d'Allen. Celle-ci propose de décrire des relations entre intervalles temporels. Elle permet notamment de décrire des événements et de les positionner les uns par rapport aux autres grâce à 13 relations de base. La logique des instants de Vilain-Kautz est basée sur le même principe et permet de décrire les événements avec 3 relations binaires. Dans les deux cas, cela permet de décrire un chemin possible respectant des contraintes liées à l'apparition d'événements mais pas liées à la connaissance d'informations ou l'atteinte d'objectifs complexes (combinaison d'action de jeu et de connaissance d'informations). **Modèle Eurikat** Nous approfondissons les travaux réalisés dans le cadre de la thèse de Pons-Lelardeux [7]. Ces travaux proposent une logique et un modèle permettant d'assurer la gestion du déroulement du scénario, de respecter les contraintes métier tout en garantissant le liberté d'actions des joueurs. Classiquement, les contraintes sont basées sur des règles métiers, des attendus pédagogiques liés au référentiel de compétences. L'originalité du modèle proposé repose sur la prise en compte de la connaissance d'informations par les joueurs et de l'atteinte d'objectifs complexes, résultants de combinaisons d'objectifs unitaires basés sur la réalisation d'actions et la connaissance d'informations. Dans cet article, nous enrichissons le modèle en prenant en compte les temporalités permettant le contrôle d'événements horodatés dans la gestion du déroulement du scénario.

3 Modélisation de scenarii pédagogiques interactifs non-linéaires

3.1 Unité médicale de l'Hôpital virtuel : CLONE

Scénario narratif Dans CLONE, l'utilisateur joue le rôle d'un personnel infirmier affecté à un service de médecine générale dans lequel sont hospitalisés entre 1 et 14 patients. Il peut interagir avec l'équipe des professionnels et les patients. Il pourra en particulier s'organiser avec l'équipe des AS et leur déléguer des tâches. Les AS sont des personnages non-joueur (PNJ) pilotés par des agents autonomes. Ils réalisent les soins qui leur ont été confiés. Le scénario est découpé en 4 grandes étapes scénaristiques : (1) Transmissions : 6h30, échanges d'informations avec l'équipe de nuit; (2) Planification des activités : élaboration d'un plan de soins

et équilibrage de la charge de travail entre les membres de l'équipe; (3) Service : 7h00, Alors que le temps est figé dans les précédentes étapes, ici, le temps défile jusqu'à 13h30. L'utilisateur réalise les activités professionnelles dans l'unité médicale. Au fil du temps, des événements horodatés surviennent par exemple, 'le brancardier vient chercher le patient X pour son examen','le patient K a chuté', 'le médecin prescrit de nouveaux médicaments à A' etc. (4) Transmissions : 13h30; le quart touche à sa fin et le joueur transmet des informations à l'équipe suivante.

Étapes scénaristiques de CLONE L'apport de la didactique professionnelle [6] et l'analyse des pratiques professionnelles[2] nous a permis la compréhension de l'activité et notamment du triplet savoir-théorique, savoir-être, savoir-faire. Les scénarios composant CLONE se veulent être des images de situations professionnelles, validées par les experts, prenant en considération les dimensions pragmatiques et conceptuelles sous-jacentes aux compétences identifiées comme étant cruciales dans l'activité des infirmières. Nous essayons de prendre en compte l'ensemble des facteurs environnementaux et sociaux qui peuvent participer au développement des compétences visées. De ce travail résulte une subdivision du scénario en étapes scénaristiques; mais pour situer des actions et des informations dans l'espace temporel de l'EV, il est nécessaire d'établir une subdivision plus fine. Pour cela, un ensemble de 9 créneaux temporels sont définis et ancrés aux étapes scénaristiques. Ainsi, des actions et/ou des informations peuvent être proposées dans des créneaux horaires particuliers.

Univers graphique L'univers graphique du service hospitalier (figure 1) se décompose en 3 zones : la barre de menu (zone 1), la vue globale du service (zone 2), la mémoire virtuelle (zone 3). La barre de menu affiche (de gauche à droite) des indicateurs associés à l'avatar de l'IDE : pictogramme, taux de charge effectif et nombre de vies, des indicateurs associés à l'AS : pictogramme, taux de charge effectif, et pictogrammes des intervenants (picto grisé=personnage absent, coloré=présent à l'instant t), icône d'accès au planning, icône d'accès au(x) dilemme(s), icône d'aide). Dans la partie basse de la zone de menu, la jauge de temps est présentée sous forme de barre horizontale divisée en 7 créneaux horaires allant de 7h à 13h30. Un bouton permet d'accélérer le temps de manière artificielle et de passer du créneau courant au créneau horaire suivant.

Fig. 1. L'univers graphique de la vue du service hospitalier

3.2 Modélisation des interactions avec impact temporel

Nous nous référons à la théorie des affordances [3] et à l'approche instrumentale de Rabardel [8].

Causalité information-temporalité Deux types d'informations sont utilisées. Une information statique est une information délivrée soit suite à la réalisation d'une action soit suite à la lecture d'une information masquée et présente dans un document métier. Une information dynamique est une information délivrée par un PNJ en cours de jeu. En réalité, les informations dynamiques peuvent interrompre l'activité courante (interruption) ou bien être délivrées de manière asynchrone (notification). Nous avons fait le choix de les différencier afin de proposer des interactions spécifiques pour chacune d'entre elles. Lorsqu'un PNJ interpelle le joueur, il apparaît une notification graphique sous son avatar. Cette notification indique au joueur que le personnage souhaite lui transmettre une information. Un simple clic affiche l'information sous forme textuelle dans une fenêtre superposée à la scène. Par exemple, 'Le petit déjeuner est servi' est une notification qui apparaît lorsqu'il est 8h00 dans l'EV. 'Je viens chercher le patient pour son scanner' est une interruption, cette information est délivrée par le PNJ 'brancardier' à 11h. Elle s'impose au joueur sans préavis.

Causalité action-information et action-temporalité Nous utilisons différents types de conditions qui peuvent être combinées entre elles: (1) précondition liée à la connaissance d'une information; (2) précondition liée à un créneau temporel; (3) postcondition : délivrer une information a posteriori; (4) postcondition : réaliser une action connexe : chaînage d'actions. Une action n'est proposée au joueur que si l'ensemble de ses préconditions est respecté. Une action est définie par une durée et possède une caractéristique cible pour être localisée sur un patient. Elle peut être répétée 0 ou n fois durant la partie. Elle fait progresser la jauge de temps d'une durée égale à la durée de l'action dans l'EV. Dans certains cas, une action peut être affichée/cachée à partir d'un créneau horaire particulier (par exemple, 'aider à prendre le petit-déjeuner'). Il est inutile de proposer cette action si le petit-déjeuner n'a pas encore été servi au patient ou bien s'il a été débarrassé.

4 Résultats

L'analyse de l'activité professionnelle nous a permis de structurer les scénarios de manière à pouvoir automatiser leur création. Une bibliothèque d'une quinzaine de scénarios a été construite à partir des modèles présentés dans cet article avec le concours d'une vingtaine de formateurs. Les scénarios se composent d'une combinaison de patients virtuels dont les dossiers médicaux ont été conçus à partir de patients réels. Leur difficulté varie selon plusieurs critères : le nombre de patients, les pathologies, les imprévus etc. Un premier prototype a été expérimenté dans 5 ISFI d'Occitanie avec 574 étudiants de 3ème année. Le contexte de l'expérimentation porte sur l'utilisation de l'EV dans une séance de formation menée par un formateur. Le formateur présente brièvement le contenu de la séance aux étudiants : les étudiants effectuent une première prise en main

avec un scénario à 1 patient avant de se confronter à un scénario où n patients sont hospitalisés (où n est entre 1 et 14). Des questionnaires ont été distribués aux étudiants à l'issu des séances de formation. 61 pour cent estiment que l'EV représente bien la réalité du contexte professionnel. 78 pour cent jugent ne pas avoir pas vécu de situations d'apprentissage similaires. Les événements horodatés permettent de recréer de manière artificielle le rythme du contexte professionnel avec des événements clés d'un service hospitalier comme les repas, la visite du médecin, l'arrivée/le départ de patient du service, l'arrivée des résultats de bilans sanguins etc. Leur utilisation permet de recréer la dynamique des relations interprofessionnelles et celle liée à l'état de santé des patients.

5 Conclusion

Un des éléments clés contribuant à faire évoluer les modalités d'organisation du travail des futurs infirmiers réside dans la formation. Le modèle et les interactions proposées dans cet article, permettent de reproduire dans un EV le contexte et la dynamique des situations impliquant une grande variété de professionnels et un nombre important de patients. Le logiciel CLONE permet ainsi une confrontation aux réalités professionnelles dans un environnement entièrement contrôlé et offre enfin la possibilité de travailler des situations d'organisation, de collaboration, et de délégation de tâches. Les premiers résultats obtenus lors des expérimentations montrent que les étudiants jugent de façon très positive la représentation du contexte professionnel et l'apport des situations d'apprentissage au travers de cet outil. Nous mesurerons l'impact de cet EV au regard des compétences visées via un recueil de données qualitatives et un recueil qualitatif.

References

1. Dang, K.D.: Aide à la réalisation de systèmes de pilotage de narration interactive : validation d'un scénario basée sur un modèle en logique linéaire. phdthesis, Université de La Rochelle (Apr 2013)
2. Galaup, M., Lelardeux, C.P., Lagarrigue, P.: De l'analyse du travail à un learning game pour la formation. Éducation permanente (2017)
3. Gibson, J.J.: The Ecological Approach to the Visual Perception of Pictures. Leonardo **11**(3), 227–235 (1978)
4. Girard, J.: Linear Logic **50**(1), 1–101 (1987)
5. Martens, C., Bosser, A.G., Ferreira, J.F., Cavazza, M.: Linear logic programming for narrative generation. In: International Conference on Logic Programming and Nonmonotonic Reasoning. pp. 427–432. Springer (2013)
6. Pastré, P.: L'analyse du travail en didactique professionnelle. Revue française de pédagogie pp. 9–17 (2002)
7. Pons Lelardeux, C.: Real-time virtual collaborative environment designed for risk management training : communication and decision making. phd, Université de Toulouse, Université Toulouse III - Paul Sabatier (Oct 2017)
8. Rabardel, P.: Les hommes et les technologies; approche cognitive des instruments contemporains p. 195

Élaboration d'une méthodologie d'instrumentation pédagogique en contexte universitaire

Vincent Bettenfeld[1], Raphaëlle Crétin-Pirolli[2], Claudine Piau-Toffolon[1], Christophe Choquet[1]

[1]Laboratoire d'Informatique de l'Université du Mans
[2]Centre de Recherche en Education de Nantes
{prénom.nom}@univ-lemans.fr

Résumé. Nous proposons dans cet article les premières étapes de l'établissement d'un canevas méthodologique permettant d'assister les acteurs de l'instrumentation de situations pédagogiques à l'aide d'outils reposant sur des technologies innovantes. Afin d'atteindre cet objectif, il a été entrepris une modélisation des différentes parties prenantes ainsi que l'environnement dans lequel s'inscrit cette instrumentation, dans ses dimensions organisationnelles et matérielles. Nous présentons le contexte de la genèse de ce modèle : l'instrumentation par un environnement outillé basé sur la transcription de la parole en direct, ainsi que d'autres technologies à l'état de prototypes de recherche.

Mots-clés. Instrumentation, modélisation, situation pédagogique, prototype, transcription, analyse des besoins

Abstract. This article presents the first steps in building a methodological framework aiming to assist contributors instrumenting pedagogical situations using tools based on innovative technologies. This framework ought to be used in various possible situations. This variety is described by a model of various stakeholders and the environment where the process is taking place, considering organizational or equipment constraints. Furthermore, we present the context leading to the creation of this model: instrumentation using a toolset based on synchronous speech transcription among other technologies in the research prototype stage.

Keywords. Instrumentation, modeling, pedagogical situation, prototyping, transcription, needs analysis

1 Introduction

Pour Tchounikine, Mørch et Bannon [11], l'interaction entre éducation et technologie « est le cœur du domaine » de recherche des EIAH. Pour cette raison, l'intégration de technologies nouvelles, ou spécifiquement adaptées, dans des contextes d'enseignement afin de créer de nouvelles possibilités d'apprentissage est identifiée comme l'un des rôles du chercheur en EIAH. Les auteurs alertent néanmoins sur le fait que cette recherche d'innovation technologique ne doit pas perdre de vue la pédagogie déployée dans ces situations d'enseignement. Les présents travaux visent à présenter

les premières étapes de la modélisation du processus d'instrumentation par une technologie dont la maturité n'est pas encore complète.

2 Modélisation de Situations Instrumentées

Dans le domaine des sciences de l'éducation, certains modèles sont utilisés pour modéliser des situations pédagogiques ; plusieurs paramètres peuvent les décrire. Le modèle TPACK [5] modélise notamment les différentes compétences mobilisées par l'enseignant lors d'un cours instrumenté. Un autre modèle, ASPID (Adoption, Substitution, Progrès, Innovation, Détérioration) [4] représente plutôt la situation d'enseignement elle-même et les enjeux pédagogiques derrière l'instrumentation de celle-ci. Le modèle ASPI (Analyser, Soutenir et Piloter l'Innovation) [7] décrit le « dispositif techno-pédagogique ». Les auteurs emploient ce terme pour désigner un ensemble de variables regroupant les fonctions des acteurs de l'instrumentation, leurs caractéristiques personnelles, les relations entre eux, ainsi que les spécificités liées à l'enseignement dans le domaine concerné.

Très détaillés, le modèle PETTaL [6] ainsi que le processus d'intégration des TIC en contexte éducatif de Fiévez [2] prennent en compte l'ensemble des différents acteurs et leurs échanges, ainsi que des variables décrivant des aspects techniques et organisationnels. Le premier modèle fait une liste des variables pouvant influencer l'enseignement et l'apprentissage des sciences à l'aide de la technologie, ces variables étant regroupées en pôles thématiques. Ceux-ci sont abrégés dans l'acronyme PETTaL qui signifie : *Power factors* (contraintes et libertés des acteurs), *Environment, Teacher* (les compétences et l'entourage de l'enseignant), *Technology and Learners* (les compétences et le comportement des apprenants).

Le processus décrit par Fiévez est « itératif et constructif ». Il modélise des situations de conception continue mobilisant élèves, enseignants, directions d'établissement et cellules d'innovation pédagogique. Ces acteurs interviennent dans le processus d'intégration au cours de différentes phases : réflexion et planification pédagogique, mise en place et utilisation de la technologie en classe, ajustements préliminaires ayant lieu avant l'utilisation en fonction des premiers résultats observables, jusqu'à l'arrivée d'une intégration efficiente, et enfin évaluation de cette intégration. Ce processus étant itératif, l'évaluation finale déclenche la réflexion concernant les modifications à apporter au dispositif par la suite en renvoyant à une phase de réflexion et de planification pédagogique. Ces modèles de situations d'apprentissage instrumentés ont permis de guider en partie notre proposition méthodologique, en recouvrant toutes les étapes et toutes les facettes du processus d'instrumentation.

3 Processus d'Instrumentation de Cours Universitaires dans la Recherche en EIAH

L'objectif ciblé par les travaux présentés dans cet article est de proposer un canevas méthodologique aidant à l'instrumentation de situations pédagogiques à l'aide d'une technologie innovante. Nous limitons le contexte d'application aux situations où la

phase de décision d'adoption est déjà achevée, et en se limitant à la phase d'implémentation elle-même, sans considérer la phase de routinisation qui implique d'autres enjeux, à d'autres échelles. Afin de conseiller au mieux les responsables de cette instrumentation dans les différentes opérations à mener, les préconisations d'un canevas méthodologique doivent être reliées aux différents paramètres d'un modèle représentant spécifiquement les situations pédagogiques. Il est nécessaire de présenter des préconisations pertinentes au regard de la phase d'avancement du projet, en s'appuyant sur les phases décrites dans le modèle de Fiévez.

Certaines propriétés des modèles décrits précédemment peuvent être restreintes ou modifiées dans le cadre de l'instrumentation de cours à l'Université. La variable Environnement du modèle PETTaL possède déjà, dans le modèle d'origine, des propriétés qui permettent de décrire le cadre de l'enseignement supérieur. Néanmoins quelques modifications sont à considérer pour se positionner dans ce cadre précis – par exemple, l'identification des parents d'élèves comme parties prenantes de l'instrumentation, inappropriée ici. En revanche, les centres d'innovation pédagogique auraient tout intérêt à être modélisés comme des acteurs de ce processus, comme ils le sont dans les travaux de Fiévez, étant donné leur rôle de support dans la transformation de l'activité d'enseignement et dans l'intégration du matériel dans la salle de classe. Comme dans le modèle ASPI, nous proposons d'intégrer les chercheurs parties prenantes dans notre modèle de l'instrumentation. Le contexte de l'enseignement universitaire ainsi que le cas considéré d'une technologie en cours de développement font que leur implication a un impact sur l'évolution de l'outil intégré ainsi que sur les modalités d'instrumentation elles-mêmes.

Le processus d'instrumentation que nous visons concerne une technologie dont le niveau de maturité n'est pas celui d'un produit commercialisé/diffusé à grande échelle. En conséquence cette technologie possède un niveau de robustesse faible (une grande possibilité de fournir des erreurs), ainsi qu'une authenticité faible (son utilisation n'est pas ancrée dans des pratiques existantes). Son utilisabilité peut également être mauvaise si le dispositif est testé pour les premières fois dans ce contexte. Or, dans ce contexte de recherche, il est possible de voir la technologie arriver à maturité, ou de devoir travailler avec les chercheurs et développeurs afin de la faire évoluer. Ce degré de maturité doit donc être représenté lors de la modélisation de ces situations.

4 Étude de Cas : Instrumentation par la Technologie de Transcription de la Parole

Nos travaux s'inscrivent dans un contexte de recherche orientée par la conception. Ce type de recherche permet de générer des procédures de conception généralisables [12]. Pour cela, la démarche a été d'instrumenter des cours existants grâce à des prototypes de recherche destinés à la transcription de la parole en direct, à la proposition de ressources de manière automatique, ainsi que de la segmentation thématique automatique à partir de corpus de parole. Pour reprendre la typologie de Smith [10], la « fidélité à l'environnement » de ces instruments est faible, c'est-à-dire qu'il n'ont pas été testés dans un contexte similaire à celui d'application.

Environnement utilisé lors des cours magistraux. Cet environnement a fait l'objet d'un article décrivant sa conception, son architecture et son interface [1]. Les étudiants ont accès à la diffusion d'un flux audiovisuel montrant l'enseignant (Fig 1.a) ainsi que des supports de cours, synchronisés avec ceux projetés en classe (Fig 1.b). Lors du cours, des liens vers des ressources pertinentes sont recommandés aux étudiants (Fig 1.c). La transcription est également affichée en temps réel dans une fenêtre restreinte similaire à des sous-titres (Fig 1.d) qui peut être agrandie pour être lue plus confortablement. Les étudiants peuvent également prendre des notes dans un éditeur de texte (Fig 1.e) et les exporter à l'issue du cours.

De son côté, l'enseignant dispose d'un instrument lui permettant d'avoir un retour sur les diapositives et la vidéo visibles par les étudiants. Il peut également voir un ensemble d'indicateurs lui donnant, par exemple, la proportion d'étudiants ayant demandé plus d'explication, ainsi que les questions ouvertes posées par l'audience.

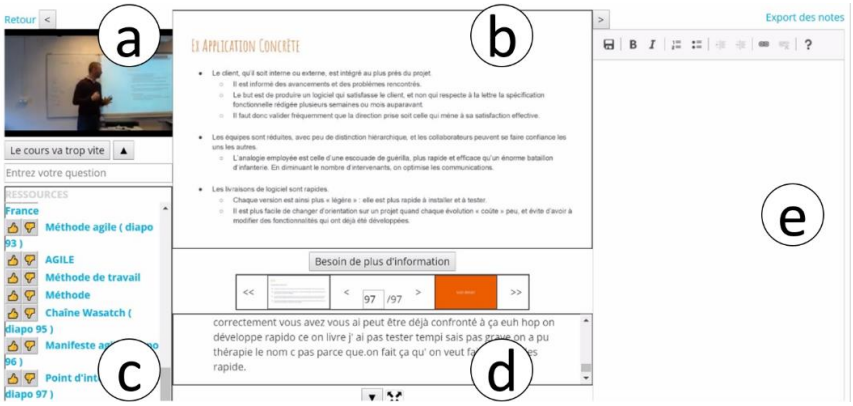


Fig. 1 Environnement proposé aux étudiants, comprenant le flux vidéo (a), les diapositives (b), la liste des ressources (c), le système de transcription (d) ainsi que l'éditeur de texte (e).

Environnement utilisé lors des travaux de groupe. Les étudiants disposent des vidéos des cours magistraux (Fig 2.b), synchronisées avec leurs transcriptions respectives (Fig 2.c). Afin de pouvoir naviguer plus facilement dans ces contenus vidéos et textuels, ils sont divisés en chapitres générés automatiquement et titrés (Fig 2.d). Cet outil permet à l'étudiant de dicter un compte-rendu, exigé par l'enseignant à la fin du module. Des indicateurs sur l'activité des groupes sont également mis à disposition des étudiants et de l'enseignant par le biais de son interface.
aa

Apports au modèle et au processus d'instrumentation en construction. Le processus d'instrumentation suivi s'est déroulé en cinq (5) phases. Dans un premier temps nous avons mené une étude des besoins (1) puis conçu un environnement et ses interfaces (2). Nous avons ensuite procédé à l'installation et l'utilisation de la technologie innovante (transcription automatique) en classe sans modifier le scénario

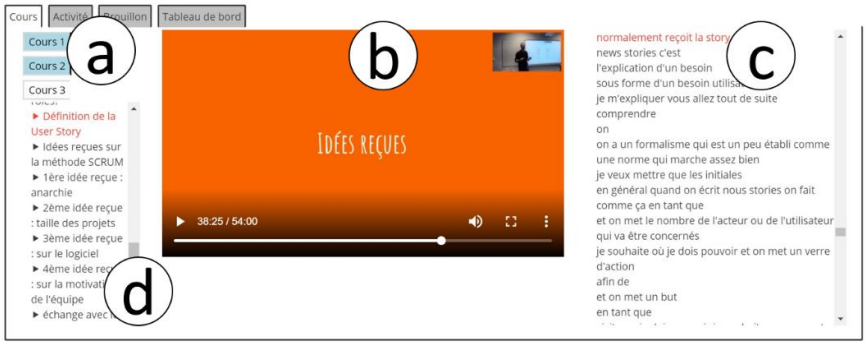


Fig. 2 Interface proposée lors des travaux pratiques. L'un des onglets (a) permet d'afficher la vidéo correspondant à un cours (b), la transcription correspondant (c), ainsi que la liste des chapitres (d).

pédagogique de l'enseignant (3). Une première session s'est tenue en mars 2018 avec l'instrumentation d'un cours en IUT, dans un contexte hybride simulé. En aval de cette expérimentation, nous (i.e. chercheurs participants du projet) avons procédé à une analyse des usages (4) : un focus group guidé par une grille d'analyse, complété par un travail d'observation et d'analyse des sessions de certains étudiants. L'ensemble des étudiants a déclaré être satisfait de la qualité de la transcription. Les possibilités offertes par l'outil se sont révélées pour la majorité d'entre eux comme trop nombreuses. Ainsi par exemple, 31,4% des actions effectuées par les étudiants étaient des manipulations de l'interface [8]. Des modifications ont été identifiées et entreprises en fonction de ces premiers résultats (5). L'analyse des usages a précédé la nouvelle itération. Trois autres expérimentations ont eu lieu depuis le début de l'année 2019. Les résultats et les analyses d'usage de ces expérimentations sont actuellement en cours de traitement et d'analyse.

Certaines technologies « transparentes » [3], c'est-à-dire celles dont l'usage est intégré dans les pratiques des acteurs de la tâche, sont à prendre en compte d'une manière plus complexe que celle avec laquelle a été menée la première itération de l'instrumentation. Par exemple, durant les deux expériences décrites, les enseignants ont utilisé leurs propres télécommandes pour faire défiler les diapositives, ce qui leur a permis de se déplacer dans l'espace de cours, et les a empêchés d'observer le feedback étudiant affiché sur leur écran. Ces instruments transparents sont en réalité des instruments qui vont avoir une incidence sur l'activité (comme la majorité des instruments au regard de la théorie de l'activité [9]). Ils doivent être modélisés comme tels pour rendre compte de la situation pédagogique.

Les expérimentations montrent que les interactions enseignant-étudiants sont impactées par la nouvelle instrumentation. Or les propriétés de ces interactions (rapidité, fréquence, importance du non-verbal) sont cruciales pour le déroulement du cours tel que prévu par les enseignants ; et certains ajustements de scénario pédagogique ont dû être improvisés pour s'y adapter. Ces informations sur les technologies transparentes peuvent être obtenues avant les expérimentations. Dans le cadre de ce projet, un cours a été filmé avant la phase d'instrumentation. Ce type de méthode permet de recueillir des informations de manière non intrusive. Dans une

situation où la captation ne serait pas possible, diriger l'entretien préalable avec l'enseignant vers le recueil des informations sur ces technologies peut permettre de malgré tout accéder à ces informations.

A l'issue des deux dernières expérimentations menées, des données qualitatives ont été recueillies par le biais de focus group et d'entretiens, menés respectivement auprès des étudiants et des enseignants participants, guidés par une grille d'entretien. Ces premiers résultats seront enrichis par l'analyse des traces d'activité enregistrées grâce au plugin et des données d'observation des comportements utilisateurs. En parallèle, le processus d'instrumentation en construction présenté dans cet article sera modélisé afin de pouvoir l'appliquer lors d'une instrumentation qui aura lieu au cours de l'année universitaire 2019-2020, dont les modalités sont encore à définir.

Références

1. Bettenfeld, V., Choquet, C., Piau-Toffolon, C. : Lecture instrumentation based on synchronous speech transcription. 18th International Conference on Advanced Learning Technologies (2018).
2. Fiévez, A., Lebrun, M., & Dumouchel, G. : L'intégration des TIC en contexte éducatif : Modèles, réalités et enjeux. Presses de l'Université du Québec (2017).
3. Graham, C.R.: Theoretical considerations for understanding technological pedagogical content knowledge (TPACK) (2011).
4. Karsenti, T : Le modèle ASPID : modéliser le processus d'adoption et d'intégration pédagogique des technologies en contexte éducatif. Teachers and Teaching (2013).
5. Koehler, M. J., Mishra, P.: What is technological pedagogical content knowledge? Contemporary Issues in Technology and Teacher Education, 9(1) (2009).
6. Mukherjee, M. M.: Technological tools for science classrooms: choosing and using for productive and sustainable teaching and learning experiences (2013).
7. Peraya, D., Jaccaz, B. : Analyser, soutenir, et piloter l'innovation : un modèle « ASPI ». Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et de l'Industrie (2004).
8. Pirolli, F., Cretin-Pirolli, R. : Apports de la transcription automatique de la parole pour l'instrumentation de séquences pédagogiques et l'enrichissement des environnements personnels d'apprentissage. Échanger Pour Apprendre en Ligne (2018).
9. Rabardel, P. : Les hommes et les technologies ; approche cognitive des instruments contemporains (1995).
10. Smith, J. D.: An alternative to technology readiness levels for non-developmental item (NDI) software. In System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on (2005).
11. Tchounikine P., Mørch A.I., Bannon L.J.: A Computer Science Perspective on Technology-Enhanced Learning Research. In: Balacheff N., Ludvigsen S., de Jong T., Lazonder A., Barnes S. (eds) Technology-Enhanced Learning. Springer, Dordrecht (2009).
12. Wang, F., Hannafin, M. J.: Design-based research and technology-enhanced learning environments. Educational technology research and development (2005)

Session
Ingénierie des EIAH

Dispositif d'évaluation pour un certificat d'aptitudes cliniques : quelles mesures pour quelles compétences ?

Olivier Catteau[1], Odile Beyne-Rauzy[2], Anne Mayère[3], Nicolas Savy[4]

[1] Université de Toulouse, IRIT UMR 5505, 118 route de Narbonne,
F-31062 Toulouse, France
olivier.catteau@irit.fr
[2] Université de Toulouse, UFR de Médecine Purpan, 37 Allées Jules Guesde,
F-31062 Toulouse Cedex 9, France
CHU de Toulouse, Médecine Interne, IUCT Oncopole, 1 avenue Irène Joliot Curie,
F-31059 Toulouse Cedex 9, France
beynerauzy.odile@iuct-oncopole.fr
[3] Université de Toulouse, CERTOP UMR 5044, 115 Route de Narbonne,
F-31077 Toulouse Cedex 9, France
anne.mayere@iut-tlse3.fr
[4] Université de Toulouse, IMT UMR 5219, 118 route de Narbonne,
F-31062 Toulouse, France
nicolas.savy@iut-tlse3.fr

Résumé. Cette communication présente les premiers résultats de l'étude de la mise en place d'un certificat d'aptitude clinique à travers la conception et la mise en œuvre d'un environnement informatique. Cet environnement informatique vise d'une part à équiper cet examen, et d'autre part à mener une démarche réflexive le concernant. Les résultats des étudiants à cet examen montrent une non redondance avec ceux des épreuves classantes typiques de ce parcours de formation. En confrontant les retours des étudiants au travers de questionnaires anonymes et l'étude du dispositif de formation lui-même, nous interrogeons l'approche qui le sous-tend concernant les compétences en général, et les compétences relationnelles en particulier.

Mots-clés. Approche par compétences, évaluation, modèles pédagogiques, collecte de données, traitement de données.

Abstract. This paper presents the first results of the study of the implementation of a clinical aptitude certificate through the design and implementation of an IT environment. This IT environment aims on the one hand to equip this examination, and on the other hand to conduct a reflexive approach concerning it. The students' results in this examination show no redundancy with those of the ranking tests typical of this training program. By comparing student feedback through anonymous questionnaires and the study of the training system itself, we question the underlying approach to skills in general and relational skills in particular.

Keywords. Competency approach, assessment, learning model, data collection, data mining.

1 Introduction

Un certificat d'aptitude clinique a été mis en place depuis 2016 à l'issue du deuxième cycle des études médicales en France (année bac+6). Ce certificat représente une innovation relative pour ce domaine, dans la mesure où il suppose de rompre avec une logique dominante de concours et d'évaluation de connaissances via des QCM, pour évaluer des compétences, en l'occurrence cliniques. Cette communication prend appui sur les premiers résultats d'une démarche initiée dans le cadre d'un programme de recherche interdisciplinaire financé par l'IReSP. A travers la conception et la mise en œuvre d'un environnement informatique, il s'agit d'une part d'équiper le dispositif d'évaluation requis pour ce certificat, d'autre part de contribuer à un retour réflexif sur le dispositif.

Nous exposerons tout d'abord le contexte d'introduction de ce nouveau certificat, ainsi que les conditions de sa mise en œuvre à l'UFR de Médecine de Toulouse Purpan. A travers l'analyse des résultats de cet examen, nous montrerons que ce certificat se distingue des autres dispositifs d'évaluation, ce qui conduit à considérer qu'il évalue des savoirs ou dispositions autres que ceux que mesurent les épreuves classantes.

Sur la base de l'analyse des résultats d'examen, qui permet d'en préciser les modalités et attendus, nous proposons une prise de recul : qu'en est-il de ce qui est ainsi évalué ? Tant les retours des étudiants via des questionnaires anonymes, que la mise en discussion des résultats intermédiaires avec ceux d'autres recherches, invitent à interroger les approches pédagogiques qui sous-tendent le dispositif, et les acceptions associées des compétences cliniques.

Enfin, nous conclurons en proposant des perspectives de recherche.

2 Conditions d'élaboration et mise en œuvre d'un dispositif d'évaluation

2.1 Contexte et Méthodologie

Dans le cadre de la réforme des études médicales, la définition d'un référentiel de compétences génériques dans le Bulletin Officiel du 16 Mai 2013 a rendu nécessaire une transformation pédagogique pour mettre en place un processus d'acquisition, de développement et de certification des compétences cliniques en médecine (1er et 2ème cycle). Le certificat de compétences cliniques (CCC) a été mis en place pour en assurer la validation [12] ; 6 rôles d'ensemble ont été spécifiées, actuellement en cours de déclinaison sous forme de compétences : « Clinicien », « Communicateur », « Acteur de Santé Publique », « Scientifique », « Responsable aux plans éthique et déontologique », « Réflexif ».

Beaucoup d'auteurs se sont penchés sur la notion de compétence. Nous utiliserons dans cet article la définition de Jacques Tardif : « Un savoir-agir complexe prenant appui sur la mobilisation et la combinaison efficaces d'une variété de ressources internes et externes à l'intérieur d'une famille de situations » [15]. Cette définition souligne le caractère complexe et situé des compétences. Or l'inscription de cette démarche par compétences dans les études médicales vient en tension avec une logique de concours, qui suppose des critères objectifs et « neutres » d'évaluation pour

l'interclassement des différents candidats. Cette logique du concours a progressivement transformé les études médicales en faisant du QCM la modalité prédominante d'évaluation, et l'organisation des enseignements. Le nouveau certificat de compétences cliniques se trouve ainsi « pris en tenaille » entre l'objectif d'évaluer des savoirs complexes, interactionnels et situés, et l'exigence de critères d'évaluation objectivables et standardisables.

Nous proposons dans le cadre de cet article de questionner les logiques en tension dans un tel dispositif d'évaluation. Nous verrons qu'une forme première de résolution de ces tensions a consisté à privilégier une approche béhavioriste, focalisée sur des comportements observables et mesurables [1]. Nous poserons ensuite des jalons vers d'autres approches susceptibles de mieux évaluer les compétences mobilisées pour affronter des situations qui réclament une réflexion sur et dans l'action [2].

Pour la mise en œuvre de cet examen, un environnement informatique a été mis en place. Son rôle est de recueillir les données d'évaluations, d'automatiser et d'accélérer leurs traitements, pour faire une remontée des résultats plus rapidement. Il permet de collecter sous forme de traces brutes un ensemble d'observés, composé à ce stade essentiellement de grilles d'évaluation. Côté *front office*, le recueil de données se fait au travers d'une application web responsive permettant aux évaluateurs de saisir leurs résultats aussi bien sur ordinateur fixe, portable ou tablette, dans des réseaux hétérogènes. Chaque clic sur une case d'évaluation fait l'objet d'une trace brute envoyée au *back office*. Une fois collectées, les traces sont modélisées dans le *back office*, puis des algorithmes de calculs sont mis en place pour produire des indicateurs [4]. Les principaux indicateurs à destination des responsables de formation concernent la réussite à l'examen, le taux de satisfaction des enseignants, le taux de satisfaction des étudiants, le taux de difficulté perçue, mais également d'autres indicateurs qui prennent appui sur une analyse statistique et docimologique des grilles d'évaluation et des résultats associés [3]. Une des finalités est de vérifier la pertinence du dispositif d'évaluation et de l'améliorer pour les futures sessions : mesures de position et de dispersion des différents résultats, analyse en composantes principales [6], corrélation avec d'autres méthodes d'évaluation [8] telles que les épreuves théoriques de 6ème année et les épreuves classantes nationales (ECN). Les principaux indicateurs à destination des étudiants relèvent de la rétro-information : réussite à l'examen, commentaires des évaluateurs, évaluation et commentaires des acteurs (patients standardisés[1]). Lorsque des indicateurs nécessitent des calculs statistiques (par exemple, coefficient de corrélation), le *back office* sollicite l'application R version 3.5.0 (2018-04-23).

Le certificat de compétences cliniques n'a pas à ce stade finalité à participer à l'interclassement des étudiants mais à vérifier la maitrise de compétences associées à la pratique clinique. Il ne s'agit pas de reproduire la logique « concours » du PACES[2] ou des ECN en favorisant des critères discriminants ; la priorité est donnée à l'homogénéité des grilles d'évaluation.

L'environnement informatique constitué permet par ailleurs, distinctement, de recueillir de façon anonyme les retours tant des enseignants que des étudiants. La

1 Patient standardisé : personne formée à simuler l'histoire d'un vrai patient et à reproduire systématiquement les signes cliniques, la personnalité, le langage corporel et les réactions émotionnelles qui auront été définis d'entrée de jeu (http://cis-ge.ch/simulation/patient-standardise/definition/)

2 PACES : Première Année Commune des Etudes de Santé

réflexivité ainsi proposée aux acteurs est un appui important pour questionner plus avant les compétences attendues et évaluées.

2.2 Expérimentation

A l'UFR de Médecine de Toulouse Purpan, le CCC a donné lieu à une nouvelle modalité d'évaluation sous forme d'un Examen Clinique Objectif Structuré (ECOS) à destination des étudiants de 6ème année.

L'examen a concerné 141 étudiants en Janvier 2016 et 134 étudiants en Février 2017 qui sont passés successivement sur 7 stations (ECG[3], Dermatologie, Thérapeutique, Annonce, Interrogatoire, Examen clinique, Synthèse). Ces stations formalisent des étapes standards de la pratique clinique. Il s'agit en particulier d'évaluer la capacité à mettre en œuvre un examen clinique, et à formuler des hypothèses de prise en charge. Un lien est établi entre plusieurs stations pour envisager un même cas sous différents angles.

L'ensemble des données collectées et analysées est résumé dans le tableau 1 ci-dessous.

Tableau 1. Effectifs collectés et analysés de 2016 à 2018

	2016	2017	2018
Station ECG	141	134	
Station ECG/Thérapeutique			135
Station Dermatologie	141	134	135
Station Thérapeutique	141	134	
Station Annonce	141	134	135
Station Annonce (retour des acteurs)		134	101
Station Interrogatoire	141	134	135
Station Interrogatoire (retour des acteurs)		125	94
Station Clinique	141	134	135
Station Examen complémentaire			135
Station Synthèse	141	134	135
Questionnaire post-évaluation A destination des étudiants		134	128
Questionnaire post-évaluation A destination des évaluateurs		30	
Résultats aux épreuves théoriques de 6ème année	131	128	124
Résultats des ECN	137	128	123

Fusion des stations ECG et Thérapeutique

Création de la station examen complémentaire

L'environnement informatique n'a été mis en place qu'en 2017 mais une analyse statistique a pu être effectuée a posteriori sur les résultats de 2016. La logistique de l'examen, organisé sur 4 demi-journées, est assez lourde avec environ 25 examinateurs

3 ECG : Electrocardiogramme

mobilisés, certains participant à plusieurs demi-journées. Chaque étudiant doit passer 7 minutes par station. Le changement de station se fait au coup de sifflet. Chaque station dispose de plusieurs « vignettes » ou sujets d'examens pour éviter la communication entre les étudiants qui ont terminé leur parcours et ceux qui ne sont pas encore passés.

L'examen a concerné 135 étudiants en Février 2018 avec une modification des stations : les thématiques ECG et Thérapeutique ont été regroupées en 1 seule station, une nouvelle station « Examen complémentaire » a été introduite à la suite de l'examen clinique (ce dernier utilisant un mannequin). C'est une autre caractéristique de ce dispositif que d'être réflexif et de supposer un environnement informatique qui équipe cette réflexivité et qui soit aisément évolutif tout en assurant des formes de comparabilité des résultats.

La communication médecin / patient est plus particulièrement considérée au cours de deux stations faisant intervenir des comédiens jouant le rôle de patients standardisés (Annonce et Interrogatoire).

Sur chaque station les examinateurs disposent d'une grille d'évaluation informatisée. Les grilles des stations ECG, Dermatologie et Synthèse sont composées de 10 à 20 critères évalués entre 1 et 6 points. Les autres grilles sont composées de 10 à 30 éléments pour lesquels l'évaluateur indique s'ils ont été effectués complètement, partiellement ou pas du tout. Le rythme de passage contraint le temps passé par les enseignants à l'évaluation. Il suppose un environnement informatique très robuste et soulève des défis quant à sa conception et son évolution pour combiner efficacité et pertinence.

Un travail exploratoire a été mené par les responsables pédagogiques pour tenter de mettre en correspondance d'une part les éléments d'évaluation et d'autre part les rôles définis dans l'annexe du BO pour décrire les compétences génériques dont l'acquisition est nécessaire pour l'obtention du certificat : Clinicien, Communicateur, Acteur de Santé Publique, Scientifique, Responsable aux plans éthique et déontologique, Réflexif [12].

En fin de parcours, l'avis des étudiants et des enseignants a été collecté sur une base volontaire et anonyme au travers de questionnaires qui leur étaient dédiés. Cette base d'information est centrale pour la recherche ici menée dans la mesure où elle permet de recueillir des avis indépendants du contexte de l'examen. Ce dernier est en effet achevé au moment où il est proposé aux étudiants de prendre quelques instants pour répondre, ce que la plupart d'entre eux font bien volontiers (100% de participation en 2017, 95% en 2018) ce moment constituant une forme de sas de sortie d'examen.

2.3 Résultats

L'environnement informatique a permis de diffuser aux candidats les résultats à l'examen en 2017 et 2018 beaucoup plus rapidement qu'en 2016 en évitant un traitement manuel et potentiellement source d'erreurs.

Les questionnaires post-évaluation démontrent un intérêt pour la démarche, tant de la part des étudiants que des enseignants-examinateurs. En 2017, le taux de satisfaction du processus d'évaluation pour les enseignants était de 100%, il était de 92% pour les étudiants (98% en 2018). A 76% les étudiants ont trouvé que l'examen était adapté par rapport à la durée et au niveau demandé (92% en 2018). En 2017, la station dite «

d'annonce » - au cours de laquelle une pathologie doit être annoncée au patient standardisé - était perçue comme étant la plus difficile par les étudiants (64% de taux de difficulté). Les compétences mises en jeu dans la relation soignant-patient, domaine encore peu enseigné au cours des études médicales, sont plus particulièrement mises à l'épreuve dans de tels contextes. La situation a évolué en 2018, la station « annonce » ayant été mieux préparée avec la mise en place d'un enseignement dédié à la relation médecin/patient ; par ailleurs les situations d'annonce ont été simplifiées, en portant sur des annonces d'examen et non directement de pathologie. Le taux de difficulté perçue par les étudiants a alors diminué à 44% alors que la station examen clinique sur mannequin a été perçue comme étant la plus difficile (49% de taux de difficulté en 2018, 44% en 2017). Des éléments d'explication peuvent tenir aux limites techniques des mannequins et au manque d'entrainement des étudiants dans ce contexte.

Le retour sur évaluation était très attendu par les étudiants. Les résultats de chaque station ont été communiqués par groupe de critères. Exemples : Savoir Etre, Diagnostic, Prise en Charge, Raisonnement… Les commentaires saisis par les évaluateurs étaient également communiqués. De plus, sur les 2 stations où intervenaient des acteurs jouant le rôle de patient standardisé (annonce et interrogatoire), une évaluation formative du candidat leur était demandée. Le résultat de cette évaluation était communiqué comme *feedback* aux étudiants mais ne rentrait pas dans le calcul de la note finale pour l'attribution du certificat de compétences cliniques. Cette évaluation pourrait toutefois devenir sommative à l'avenir. En 2017, la corrélation entre l'évaluation formative des acteurs et l'évaluation sommative des enseignants sur le savoir être des étudiants n'était que de 0.39 pour l'annonce ($p=3.4e-6$) et de 0.31 pour l'interrogatoire ($p=2.427e-4$). Il y a là un élément intéressant à investiguer plus avant pour expliciter les attendus d'une part des acteurs jouant les patients standardisés, qui ont été formés au *feedback*, et d'autre part les enseignants.

L'examen n'avait pas de visée sélective, et cela s'est confirmé par la réussite de la quasi-totalité des étudiants, seuls 1 à 2 étudiants passant chaque année à la seconde session.

Une analyse statistique plus fine des résultats, résumée dans le tableau 2, montre que chaque station a son utilité dans le dispositif. Les résultats sont calculés de la manière suivante : le score de chaque station est obtenu en faisant la moyenne pondérée de tous les éléments évalués lors du passage de l'étudiant sur la station. Les résultats sont présentés sur un barème de 20 points par station. Un élément effectué « complètement » rapporte 1 point, un élément effectué « partiellement » un demi-point et un élément non effectué ne rapporte pas de point. Le score global est obtenu sur un barème de 100 points en faisant la moyenne pondérée des 7 stations. Prenons l'exemple de la station Dermato en 2016, le coefficient de corrélation est calculé entre (1) les résultats de cette station et (2) la somme des résultats des 6 autres stations. La normalité des résultats a été vérifiée. Pour chaque calcul de coefficient de corrélation, la méthode suivante a été appliquée : (1) vérification que l'allure de l'histogramme des variables est sous forme de cloche, (2) vérification que l'allure du diagramme quantile - quantile est sous forme de droite de pente égale à 1, et (3) élimination des valeurs isolées (inférieures à $Q1 - 1.5(Q3-Q1)$ et supérieures à $Q3 + 1.5(Q3-Q1)$).

Les coefficients de corrélations calculés entre les résultats des grilles d'une station donnée et les résultats des grilles des autres stations, quand ils sont significatifs, restent

positifs et faibles (entre 0.17 et 0.24 en 2016 ; entre 0.22 et 0.36 en 2017 ; entre 0.18 et 0.29 en 2018) justifiant ainsi de l'intérêt de chaque station.

Tableau 2. Analyse statistique des résultats des différentes stations

2016

Station	Min	Q1	Med	Moy	Q3	Max	ET	Cor	p
Global	49.5	68.0	72.8	72.3	77.0	88.2	6.89		
ECG	8.0	14.0	16.0	15.2	18.0	20.0	2.88	non significatif	
Dermato	2.0	11.0	15.0	13.9	17.0	20.0	3.99	0.17 [0.00;0.33]	0.048
Thérapeu	6.0	13.0	15.0	14.6	17.0	20.0	3.26	0.19 [0.02;0.34]	0.029
Annonce	0.0	12.0	15.0	14.1	17.0	20.0	3.80	non significatif	
Interro	5.0	11.5	13.0	13.0	14.5	18.2	2.11	0.24 [0.08;0.39]	0.0039
Clinique	7.5	14.0	15.5	15.1	17.0	20.0	2.43	non significatif	
Synthèse	8.0	13.0	16.0	15.3	18.0	20.0	3.16	0.18 [0.01;0.33]	0.037

2017

Station	Min	Q1	Med	Moy	Q3	Max	ET	Cor	p
Global	49.5	68.6	73.0	72.5	77.0	85.2	6.89		
ECG	6.5	13.0	16.5	15.6	18.0	20.0	3.34	0.28 [0.11;0.43]	0.0014
Dermato	6.0	12.5	14.5	14.3	16.5	19.5	3.10	0.36 [0.21;0.51]	2.1e-05
Thérapeu	5.0	12.0	14.0	13.7	15.5	19.5	2.66	0.26 [0.10;0.42]	0.0024
Annonce	5.5	13.5	15.0	15.0	16.9	19.5	2.44	non significatif	
Interro	9.8	13.6	14.9	14.8	16.2	18.8	2.03	non significatif	
Clinique	8.0	12.5	14.0	13.9	15.5	17.5	1.95	0.35 [0.20;0.50]	3e-05
Synthèse	4.0	13.0	15.0	14.5	17.0	20.0	3.11	0.22 [0.05;0.38]	0.011

2018

Station	Min	Q1	Med	Moy	Q3	Max	ET	Cor	p
Global	48.5	67.6	72.5	71.6	76.1	89.5	6.99		
ECGThe	4.0	11.5	14.5	14.0	17.0	19.0	3.32	non significatif	
Dermato	4.0	11.5	13.5	13.5	15.5	19.0	3.04	non significatif	
Annonce	9.5	15.0	17.0	16.4	18.0	20.0	2.39	0.29 [0.12;0.44]	0.00078
Interro	8.8	13.5	15.0	14.8	16.2	19.0	1.96	non significatif	
Clinique	4.5	11.8	13.5	13.3	15.0	17.5	2.61	non significatif	
Compl	2.0	12.0	14.0	13.7	16.0	20.0	3.77	0.18 [0.01;0.34]	0.04
Synthèse	6.5	13.0	15.0	14.8	17.0	19.0	2.70	0.29 [0.12;0.44]	0.00094

La faible dispersion des résultats montre également l'homogénéité de ces derniers. Une analyse en composantes principales montre une chute d'inertie faible (39% de l'information sur les 2 premières dimensions en 2016, 47.21% en 2017 et 31.88% en 2018, cf. Figure 1). Là encore, on retrouve une homogénéité, il n'y a pas de caractère discriminant entre les stations.

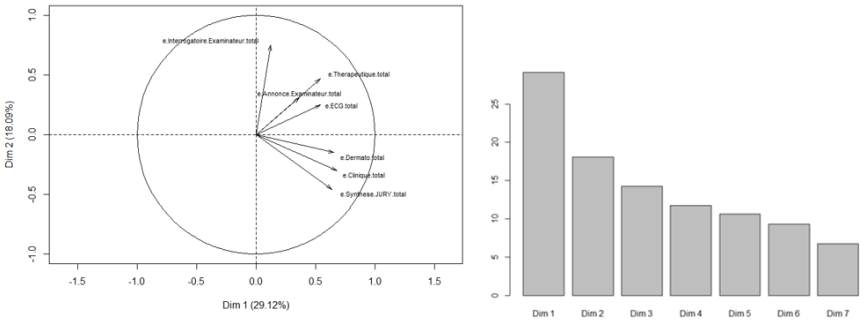


Fig. 1. Cercle des corrélations et diagramme de chute d'inertie issues de l'analyse en composantes principales des données de 2017.

Il n'y a pas non plus de redondance avec les résultats aux épreuves théoriques. En effet, comme le montre l'exemple de la figure 2, le coefficient de corrélation entre les résultats à l'ECOS et les résultats de l'UE d'enseignements théoriques de 6ème année est également significatif, positif et faible (0.40 [0.25;0.53] en 2016 ; 0.39 [0.23;0.53] en 2017, 0.25 [0.08;0.41] en 2018) justifiant de l'intérêt de l'ECOS par rapport aux modalités d'évaluation classiques existantes. La modification des stations en 2018 apporte une tendance à la diminution de la corrélation.

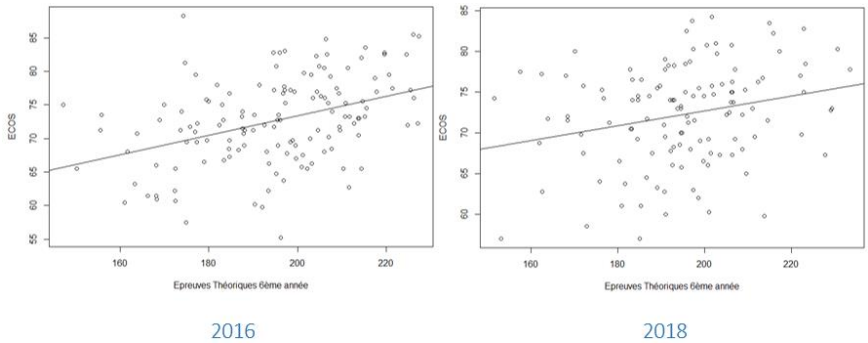


Fig. 2. Nuage de points de l'ECOS et des épreuves théoriques de 6ème année en 2016 et 2018

Enfin, il n'y a pas de redondance avec les ECN. En effet, le coefficient de corrélation entre les résultats à l'ECOS et les résultats des épreuves classantes nationales est lui aussi significatif, positif et faible (0.35 [0.19;0.49] en 2016, 0.43 [0.27;0.56] en 2017, 0.35 [0.19;0.50] en 2018). Ainsi l'examen ECOS présente une contribution spécifique, qui peut être associée à la mise en situation.

Tous ces résultats mettent en évidence l'intérêt de la démarche en ce que l'évaluation qui en résulte n'est pas redondante au regard d'autres dispositifs d'évaluation. Nous allons maintenant prendre un peu de recul au regard de ce qui peut être appréhendé comme une évaluation priorisant des critères d'évaluation objectivables et standardisables, dans une approche à dominante behavioriste.

3 Analyse critique

L'annexe du BO décrit les compétences génériques requises en les articulant autour de 7 rôles. Mais il ne donne pas d'indication sur la façon dont ces rôles et plus globalement les compétences doivent être évaluées, ici tout au long des 7 stations. Quelles conceptions sont alors retenues des attendus et des modalités de les atteindre ? En termes de ressenti, 87% des enseignants estiment que la formation actuelle permet aux étudiants d'acquérir les compétences requises. Les étudiants sont beaucoup plus circonspects quand on les interroge sur la maitrise de chacun des 7 rôles dans le cadre du questionnaire anonyme soumis après achèvement du parcours des stations d'évaluation : les taux de maitrise déclarée varient de 47% pour le rôle « d'acteur de santé publique » à 55% pour celui de « coopérateur ». On retrouve cette tendance en

2018 : les taux de maitrise déclarée varient de 50% pour le rôle « d'acteur de santé publique » à 57% pour celui de « coopérateur ». Les étudiants concernés n'avaient découvert que tardivement ces différents rôles tels qu'énoncés, auxquels ils n'avaient été sensibilisés que quelques semaines avant l'examen ; ces réponses traduisent également une difficulté de compréhension des attendus, tant les termes sont vastes et les définitions étendues.

Selon Rey, un référentiel (tel celui décrit dans le BO pour les compétences génériques) devient parole de commandement, mais ne garantit pas la faisabilité, surtout quand on aborde le caractère civique et social [13]. Rey interroge la culture du résultat susceptible d'y être associé : il y a souvent, dans nombre d'approches par compétences, confusion entre la compétence et le résultat d'une action qui s'apparente alors à une performance. A titre d'exemple, le rôle de communicateur requiert de la part de l'étudiant qu'il soit à même « d'établir une communication axée sur le patient, par le biais de la prise de décisions partagées et d'interactions efficaces fondées sur l'éthique et l'empathie ». Or il reste difficile de formuler des critères d'évaluation au-delà des formules de politesse et de civilité usuelles. Le BO laisse la mise en œuvre du référentiel sous la responsabilité des enseignants et des étudiants et fait quasi-silence sur le processus d'engendrement de ces compétences et de leur évaluation ; seules les modalités « d'évaluations en situation authentique et des auto-évaluations » sont évoquées.

Hébrard dénonce quant à lui le fait de privilégier l'apprentissage par l'action en situation au détriment des connaissances, des concepts relevant des disciplines [5]. Dans la formation d'infirmier, il constate que les grilles d'évaluations des compétences sont souvent hétérogènes et sont constituées d'une liste de preuves à vérifier ; les évaluateurs font alors face à la difficulté de cocher telle ou telle case. De plus, les dimensions affectives, conflictuelles, la question du pouvoir, les enjeux identitaires sont occultés. Dans les études médicales, les connaissances issues des sciences sociales sont en bonne part ignorées, l'enseignement étant essentiellement le fait de praticiens. Or l'approche par compétences, si elle est ramenée à une liste d'activités, ne permet pas de prendre en compte ce qui est identifié comme une caractéristique clé du travail de thérapeute, à savoir la tension intrinsèque entre d'une part un mouvement nécessaire d'objectivation du patient pour interpréter les signes éventuels de la maladie, et d'autre part une capacité à « se déprendre de sa position d'objectivation et se placer avec le patient dans une relation de vivant à vivant et de sujet à sujet » [9].

Dès lors il apparaît qu'une opération qui paraît aussi simple que de définir des listes de critères d'évaluation est en fait sous tendue par une conception plus générique de ce qu'est en l'occurrence ici un thérapeute. Une approche par compétences ramenée à une liste d'attitudes peut s'avérer très proche d'un néo-paternalisme usuel, précisément parce qu'elle n'interroge pas les dispositions requises pour être à même de mener ce double mouvement d'objectivation et de construction d'un échange effectif avec le patient.

Dans la définition de la compétence par Jacques Tardif, figure la notion de « famille de situations ». Cela interroge à deux niveaux par rapport à l'évaluation proposée : la répétition et la complexité. D'une part, cela sous-entend que la compétence ne peut être acquise que si elle a pu être mobilisée dans plusieurs situations. Quel crédit peut-on

alors donner à une certification qui ne repose que sur une situation unique (celle de l'examen) ? Une réflexion est actuellement en cours pour étendre le dispositif d'évaluation (formative et sommative) aux situations de stage de la 2ème à la 5ème année de formation. D'autre part, dans une famille de situations, il y en aura qui seront plus complexes à traiter que d'autres. On se retrouve alors face à un paradoxe : dans un souci d'homogénéité de l'examen, les « vignettes » sont construites sur des situations simples et courantes du point de vue des médecins. C'est aussi une volonté d'éviter de confronter l'étudiant à des situations de 3ème cycle.

Roegiers souligne l'importance d'établir une articulation entre le référentiel métier, le référentiel de compétences et le référentiel de formation initiale [14]. Le Boterf ajoute la nécessité d'élaborer préalablement un référentiel d'évaluation [7]. On sent bien que la volonté ministérielle d'introduire l'approche par compétences dans les études de médecine n'en est qu'au début de la démarche. La mise en œuvre de l'ECOS l'est également. Les critères d'évaluation ont été définis de manière empirique. La priorité donnée à l'opérationnel pour faire face aux défis de l'ingénierie pédagogique a conduit à mettre en partie de côté la réflexion sur les conceptions des relations soignants-patients et leurs modes d'enseignement.

Il est en cela utile d'interroger la posture béhavioriste de l'approche par compétences. Comme le souligne Boutin, le béhaviorisme est adapté pour la maitrise de comportements observables, objectivés. Dans cette approche, « le comportement visé doit être répété jusqu'à ce qu'il devienne automatique » [2]. Ainsi en est-il lors d'examens cliniques de l'identification de certains signes cliniques. Leung démontre les dérives d'une approche qui peut dans son principe de répétitivité étendue à l'évaluation tendre à l'absurdité [10] : il donne l'exemple d'un étudiant qui doit montrer au moins 4 fois qu'une compétence est mise en œuvre au cours de 7 situations filmées différentes. Meirieu dénonce l'atomisation des apprentissages globaux en une multitude de petits actes dont la recomposition peut être difficile [11].

D'autres stratégies ne sont pas à écarter : le cognitivisme peut être utile à la résolution de problèmes tandis que le constructivisme aide à affronter des situations qui réclament une réflexion sur et dans l'action [2]. Là encore il apparaît que les orientations retenues ont des liens avec l'approche retenue de l'activité thérapeutique : si les compétences cliniques reposent non seulement sur la connaissance des pathologies et des moyens d'investigation et de traitement, mais de façon imbriquée, sur « la compréhension fine de ce que peut signifier le fait d'être malade pour un individu et son entourage, ainsi que la connaissance des enjeux sociaux et institutionnels de cette expérience » [9], alors les situations de relation soignant-patient sont fondamentalement des situations qui réclament une réflexion sur et dans l'action, et une réflexion menée non sur le patient mais avec lui.

La démarche des ECOS est également répandue dans les pays anglo-saxons avec ses checklists et ses patients standardisés. Elle apporte certes l'individualisation de l'apprentissage mais revient selon certains auteurs à démontrer la performance de l'étudiant, ce potentiellement au détriment du processus sous-jacent [10]. L'approche peut être perçue comme superficielle et parfois source de démotivation. Elle encourage les étudiants à faire ce qu'on attend d'eux plutôt que de prendre le risque de développer une distance critique. A cela s'ajoute le risque pour les examinateurs de définir des

règles minimalistes de validation. L'autre risque souligné par Leung, la réduction du contenu pédagogique, est particulièrement sensible dans le contexte d'études médicales où les étudiants désertent les cours magistraux et où le primat donné aux QCM fait que toute évaluation doit se conformer aux modalités d'un QCM, quand bien même les connaissances en cause sont complexes et relèvent d'un paradigme différent de celui de *l'evidence-based medicine*, ce qui est le cas des connaissances issues des sciences sociales.

4 Conclusion et perspectives

Cette communication rend compte des premiers résultats d'une démarche innovante en termes de formation aux compétences cliniques de futurs médecins et de leur évaluation. L'environnement informatique constitué permet d'une part d'équiper l'ingénierie pédagogique, et d'autre part, distinctement, de recueillir de façon anonyme les retours tant des enseignants que des étudiants.

L'étude des résultats de cet examen de compétences cliniques montre que cet examen n'est pas redondant au regard des examens évaluant les compétences plus « théoriques » sur les maladies, les techniques diagnostiques et thérapeutiques, et les sciences dites fondamentales associées. En cela cet examen présente une contribution spécifique, dont la pleine signification reste à expliciter.

L'analyse critique proposée en complément permet d'ouvrir des pistes sur les questions que soulève cet examen. L'observation du dispositif en train de se faire permet d'interroger les paradigmes en présence. Il apparaît que, compte tenu des impératifs de calendrier, et de conceptions implicites sur les relations soignants-patients, l'approche behavioriste ait été dominante dans les premiers temps de ce dispositif.

En perspectives, la question est celle de dispositif complémentaire ou renouvelé de formation et d'évaluation qui fasse place à une approche constructiviste susceptible d'aider à affronter des situations qui réclament une réflexion sur et dans l'action [2]. Dans cette perspective, la refonte des stages et de leur évaluation pourrait être l'occasion d'une réflexion sur la mise en place de démarches pédagogique et d'évaluation accordant plus de place à cette réflexion sur et dans l'action, non seulement sur mais avec les patients. Il s'agit également de réfléchir à un dispositif permettant de mesurer la progression des compétences acquises au fil du cursus, voire au-delà si une articulation peut être établie avec la formation d'internat voire la formation continue des praticiens.

Dans ce cadre, l'EIAH doit permettre de mesurer la progression des compétences acquises en cours de formation et de contribuer à la réingénierie pédagogique, tout en étant lui-même objet d'enquête et de réflexion dans la perspective interdisciplinaire retenue.

Remerciements. Cette recherche est financée grâce au soutien de l'Institut de Recherche en Santé Publique, IReSP-17-HSR-16. Elle a bénéficié de l'aide des partenaires financeurs de l'IReSP dans le cadre de l'appel à projets général 2017 Volet Services de Santé.

J.C. Basson (CRESCO EA 7419*), E. Coeurdevey (CERTOP UMR 5044*), E. Fournales (CERTOP UMR 5044*), N. Haschar-Noé (CRESCO EA 7419*), P. Marrast (CERTOP UMR 5044*), T. Lang (INSERM UMR 1027**), M. Kelly-Irving (INSERM UMR 1027*), S. Lamy (INSERM UMR 1027*), L. Bonneville (GRICO**) et S. Grosjean (GRICO**), participent également à cette recherche dans le cadre d'une collaboration menée au sein de l'IFERISS, Institut Fédératif d'Etudes et de Recherches Interdisciplinaires Santé Société.

* Université Toulouse 3
** Université d'Ottawa

Références

1. Boutin G. : L'approche par compétences en éducation : un amalgame paradigmatique. Connexions 2004/1 n° 81 (2004) 25-41
2. Boutin G. : Le béhaviorisme et le constructivisme ou la guerre des paradigmes. Québec français 119 (2000) 37-40
3. Dessus P. : Analyser les items d'un questionnaire. http://espe-rtd-reflexpro.u-ga.fr/docs/scied-cours-qcm/fr/latest/tuto-analyse-items.html, (2018)
4. Djouad T. : Ingénierie des indicateurs d'activités à partir de traces modélisées pour un environnement informatique d'apprentissage humain. Thèse Université Claude Bernard Lyon 1 (2013)
5. Hébrard P. : L'humanité comme compétence ? Une zone d'ombre dans la professionnalisation aux métiers de l'interaction avec autrui. Les sciences de l'éducation – Pour l'Ere nouvelle 2011/2 Vol.44 (2011) 103-121
6. Husson F., Lê S., Pagès J. : Analyse de données avec R. 2ème édition, Presses Universitaires de Rennes (2016)
7. Le Boterf G. : Ingénierie et évaluation des compétences. 6ème Edition, Editions d'Organisation, Eyrolles (2011)
8. Lecoutre J.P. : Statistique et probabilités. 6ème édition, Dunod (2016)
9. Lefève C., Mino J. C.: Former de vrais thérapeutes. La place des sciences humaines et sociales dans les études de médecine, revue Etudes, n°4142, Février 2011, p. 187-198.
10. Leung W.C. : Competency based medical training : review. British Medical Journal Volume 325 (2002) 693-696
11. Meirieu P. : Si la compétence n'existait pas, il faudrait l'inventer… https://www.meirieu.com/ARTICLES/SUR%20LES%20COMPETENCES.pdf (2005)
12. Ministère de l'enseignement supérieur, de la Recherche et de l'Innovation : Etudes Médicales, Régime des études en vue du premier et du deuxième cycle, Bulletin Officiel n°20 du 16 Mai (2013)
13. Rey B. : La notion de compétence : usages et enjeux. Le français d'aujourd'hui 2015/4, n° 191 (2015) 15-24
14. Roegiers, X. : Des curricula pour la formation professionnelle initiale : la Pédagogie de l'intégration comme cadre de réflexion et d'action pour l'enseignement technique et professionnel. Bruxelles : éd. De Boeck Université (2010)
15. Tardif J. : L'évaluation des compétences : Documenter le parcours de développement. Montréal, Canada, Chenelière Education (2006)

Génération de scénarios adaptés dans un jeu d'apprentissage selon une approche dirigée par les modèles

Pierre Laforcade and Youness Laghouaouta

Laboratoire d'Informatique de Le Mans Université (LIUM)
Laval, France
prenom.nom@univ-lemans.fr

Résumé Notre intérêt se porte sur les jeux sérieux d'apprentissage nécessitant des mécanismes d'adaptation au profil de l'apprenant. Nous proposons de traiter plus précisément la génération de scénarios adaptés selon une approche d'Ingénierie Dirigée par les Modèles. Cette approche permet de guider l'analyse et la conception de la génération. Ceci est réalisé par la spécification de différents modèles selon plusieurs perspectives et dimensions. Nous avons appliqué la proposition dans le contexte du projet *Escape it !* dont le jeu sérieux a pour but de supporter le renforcement et la généralisation de compétences en performances visuelles pour des enfants autistes.

Keywords: Jeux sérieux · Adaptation · Scénario d'apprentissage · Ingénierie Dirigée par les Modèles.

Abstract In the context of learning games, we tackle the need of generating adapted learning scenarios according to elements such as the learners' progress and skills. Our proposal is a design approach relying on Model Driven Engineering techniques. It guides designers and domain experts to express, by means of models, what has to be generated, what contextualizes the generation, and what information is required from the learning game. We apply it in the *Escape it !* research project that aims to propose an "escape-room" game for helping children with autism to learn visual performance skills.

Keywords: Serious game · Adaptation · Learning scenarios · Model Driven Engineering.

1 Introduction

Les jeux d'apprentissage, *learning games*, sont des jeux sérieux visant à favoriser un apprentissage spécifique. Ils requièrent des mécanismes d'adaptation afin que tout ou partie du dispositif (éléments du jeu et/ou de l'apprentissage) puissent être adaptés dynamiquement au contexte (profil d'apprenant, profil de joueur, etc.). Dans cet article, nous nous intéressons plus particulièrement à

l'adaptation dynamique de scénarios de jeu en fonction des compétences de l'apprenant et des connaissances sur les niveaux de jeu. Ce que nous considérons comme *scénario de jeu* correspond à la spécification déclarative d'une session de jeu qui serait interprétable sans ambiguïté par le moteur du jeu : séquence ordonnée des éléments de jeu et de leurs caractéristiques. L'objectif général de notre recherche est d'identifier un cadre de conception générique qui permettrait d'aider à la conception d'un générateur de scénarios adaptés [1].

Notre travail de recherche s'inscrit dans le cadre du projet *Escape it !* présenté dans la section 2. Nous présentons ensuite brièvement des travaux connexes abordant la génération de scénarios adaptés en section 3. Ensuite, nous présentons une approche de conception exploitant des techniques de spécification issues de l'Ingénierie Dirigée par les Modèles. Cette approche est alors appliquée au contexte du projet *Escape it !* en section 4. Enfin, la section 5 présente une conclusion et des perspectives que nous abordons actuellement.

2 Contexte et motivation

L'objectif du projet est de concevoir et de développer un *learning game* sur mobile pour des enfants avec TSA (Trouble du Spectre Autistique). Le jeu servira à supporter le renforcement et la généralisation de l'apprentissage des compétences de type performances visuelles (définies dans le guide d'évaluation ABLLS-R [13]) : appariement objet-objet ou objet-image, tri, catégorisation, sériation, etc.

2.1 Vue générale du jeu d'apprentissage

Le jeu exploite un *gameplay* minimaliste inspiré des jeux de type *escape-room* : le joueur doit trouver dans une pièce (lieu de vie) des objets, parfois cachés, et les déplacer vers des zones "solutions", de manière à débloquer l'ouverture de la porte et ainsi accéder au niveau suivant. L'orientation *escape-room* a été proposée par les experts autisme impliqués dans le projet. Ils ont considéré que ce jeu pourrait être un intermédiaire intéressant pour supporter la généralisation des apprentissages entre les sessions thérapeutiques structurées et la généralisation dans l'environnement naturel de l'enfant tel que préconisé par le modèle d'intervention PRT (*Pivotal Response Training*) [9]). Le jeu sera utilisé par l'enfant, supervisé d'un adulte.

La conception du jeu a été réalisée sur la base des bonnes pratiques recensées [4][20] et à l'aide des recommandations et exigences exprimées par les experts autisme lors des séances de conception participative. L'objectif général était de réifier dans le jeu les éléments clés des approches comportementales et cognitives qui ont fait leur preuve dans la prise en charge de l'autisme [2] : utilisation du renforcement, contrôle des antécédents, guidances et estompages de guidances, façonnement et chainage. Voici quelques éléments clés de conception :

1. Cet article est une version francisée de [9] avec une focalisation portée sur la présentation de l'approche globale et non sur le processus de conception.

— Des compétences cibles : un sous-ensemble des compétences de performances visuelles définies dans [13] (celles adaptables au gameplay du jeu) ; par exemple : apparier un objet à un objet identique, trier des objets similaires, catégoriser des objets selon leurs fonctions communes, etc.
— Des durées de sessions de jeu variables : le jeu proposera des sessions de 3 à 6 niveaux, au choix de l'enfant.
— Des niveaux prenant place dans des scènes représentant des lieux de vie faisant sens pour l'enfant, groupées en thèmes.
— Des niveaux de difficulté : la difficulté correspondra à la configuration des éléments de la scène (objets visibles ou cachés, éléments de solution, positions dans la scène, etc.). Elle doit être fixée, pour chaque compétence-cible, en corrélation avec la progression des apprentissages de l'enfant.
— Des scènes variées pour encourager la généralisation des compétences : il est nécessaire que le jeu propose, pour une même compétence et un même niveau de difficulté, une grande variété de configurations de scènes de jeu.

2.2 Un besoin de génération automatique

La figure 1 illustre un exemple de scène de jeu pour la compétence B3 (appariement d'objets identiques) pour la difficulté *élémentaire*. Deux jouets "dinosaures" doivent être trouvés et placés dans le meuble et à coté de l'exemplaire rangé. Les objets utilisés pour la résolution ainsi que l'emplacement de ces objets peuvent varier : il s'agit d'un exemple de configuration générée pour une scène. Pour d'autres situations ou d'autres niveaux de difficulté, différents décors additionnels peuvent être ajoutés dans la scène, apportant ainsi de nouvelles positions possibles (pour placer les objets) ou des cachettes interactives.

Figure 1. Une exemple de configuration des éléments pour la scène *chambre*.

Dans ce contexte, un scénario de jeu correspond à une séquence ordonnée de scènes dont la configuration des différents éléments (objets, décors, cachettes, emplacements, etc.) est précisée. Ces informations (les scènes et leurs configurations) doivent être adaptées au profil de l'enfant (progression dans l'arbre des compétences, niveau de difficulté pour chaque compétence, etc.) et doivent prendre en compte les informations du jeu (l'arbre des compétences, relations entre niveau de difficulté et agencement des scènes, objets disponibles pour chaque scène, etc.). Cette adaptation est nécessaire avant chaque session de jeu.

Il serait trop coûteux de développer toutes les combinaisons possibles de configurations pour les scènes. Le jeu d'apprentissage requiert une génération dynamique de sessions de jeu adaptées au profil de chaque enfant.

3 Adaptation et génération d'activité

L'adaptation est l'action/activité d'ajuster un système en vue de le rendre adapté à quelque chose. Dans le contexte des jeux sérieux, comme de manière générale, elle peut être caractérisée par la *cible* de l'adaptation (ce qui sera adapté), par la source (à quoi sera adaptée la cible), ainsi que par les techniques employées [19]. Des états de l'art ont été réalisés pour caractériser ces différents éléments [5][21]. Pour [5], les différentes cibles ont trait à la présentation, au contrôle ou au contenu tandis que les sources concernent majoritairement le modèle utilisateur et les paramètres systèmes. Les modèles d'adaptation peuvent être implicites (dans le code) ou explicites (machines à états, règles logiques, etc.).

L'adaptation de scénarios dans les jeux sérieux est souvent abordée en considérant des scénarios de jeux narratifs et interactifs [3][16][7]. Dans le contexte du projet *Escape it !*, le concept de scénario correspond davantage à la définition générale proposée par [11] (*the global progression within a game level, its initial settings and the logical flow of events and actions that follow*) bien que nous ne sommes pas concernés par des flux d'événements ou d'actions (le jeu visé n'implique pas d'événements scriptés).

Dans [1], les auteurs ont proposé un système de génération de contenu contrôlé par des enseignants. Ces derniers sélectionnent des objets de jeu pré-créés, leur ajoutent du nouveau contenu d'apprentissage et les associent entre eux. Dans notre contexte, l'implication des experts n'est pas nécessaire pour chaque adaptation. En revanche la connaissance des objets manipulables disponibles pour chaque scène, leurs relations vis-à-vis des compétences-cibles possibles, leurs relations avec les objets-solutions semble être une base nécessaire pour générer des configurations de scènes appropriées. De telles connaissances devraient être spécifiées de manière à les rendre exploitable par l'activité d'adaptation tout en facilitant l'implication des experts dans leur expression.

Plus proche de nos préoccupations, les travaux présentés dans [15] proposent une architecture générique pour personnaliser un scénario de jeu sérieux en fonction des compétences de l'apprenant et de ses traces d'interaction [6]. L'architecture a été évaluée dans l'objectif de développer un jeu sérieux pour évaluer et réhabiliter des troubles cognitifs. Cette architecture est organisée en trois

couches : concepts du domaine, ressources pédagogiques, et ressources du jeu sérieux. En complément, leur proposition consiste à générer trois scénarios successifs (conceptuel, pédagogique et de jeu) en relation avec les couches précédentes. Pour la mise en oeuvre concrète, des techniques de représentation et manipulation d'ontologies sont utilisées. Toutefois peu d'informations sur la spécification des règles de génération sont données car leur intérêt était davantage porté sur l'analyse des traces d'interactions de l'apprenant et la mise à jour de son profil.

4 Proposition

Notre préoccupation principale concerne la génération de scénarios de jeu d'apprentissage adaptés aux profils des apprenants, tout en prenant en compte les connaissances sur le jeu et les compétences visées. Plus précisément nous avons retenu les deux problématiques et défis suivants : comment identifier et bien définir les éléments du domaine du jeu d'apprentissage (compétences, connaissances sur le jeu, éléments du modèle de l'apprenant...), ainsi que les règles de génération (défi 1) ? Comment exploiter ces informations pour diriger la génération effective des scénarios de jeu d'apprentissage (défi 2) ? Comment impliquer les experts du domaine métier dans la conception et la validation de la génération des scénarios adaptés (défi 3) ?

4.1 Cadre de référence pour les besoins de modélisation

L'Ingénierie Dirigée par les Modèles (IDM) [8] est une forme d'ingénierie générative proposant une démarche par laquelle tout ou partie d'une application informatique est générée à partir de modèles. En contrepartie, ces modèles doivent être rendus explicites et suffisamment précis pour être interprétés ou transformés par des machines. Le processus de développement peut alors être considéré comme un ensemble de transformations de modèles prenant des modèles en entrée et produisant des modèles en sortie, jusqu'à obtention d'artefacts exécutables. Le méta-modèle est alors l'entité concrétisant informatiquement le contexte de modélisation pour la conception et la manipulation des modèles. Il est l'abstraction du langage de modélisation des modèles (syntaxe, grammaire et sémantique). On considère alors que le modèle est *conforme* à son méta-modèle.

L'IDM pourrait permettre de spécifier la description du jeu, le profil de l'utilisateur, etc. en tant que modèles. L'*objet* de l'adaptation serait alors le modèle de sortie d'une transformation de modèles dont les éléments participant à sa mise en oeuvre seraient les modèles en entrée (défis 1 et 2). Ces modèles seraient exprimés à un haut niveau d'abstraction qui pourrait permettre la participation active des experts métier en interaction avec les développeurs du jeu d'apprentissage (défi 3).

4.2 Une approche 3x3x2 de modélisation

Nous proposons une approche qui permet d'appréhender la conception de la génération de scénarios adaptés. Elle s'appuie sur 3 perspectives incrémentales pour

les scénarios à générer, 3 dimensions de spécification des éléments du domaine, et 2 niveaux d'abstraction (modèle et méta-modèle) (Figure 2).

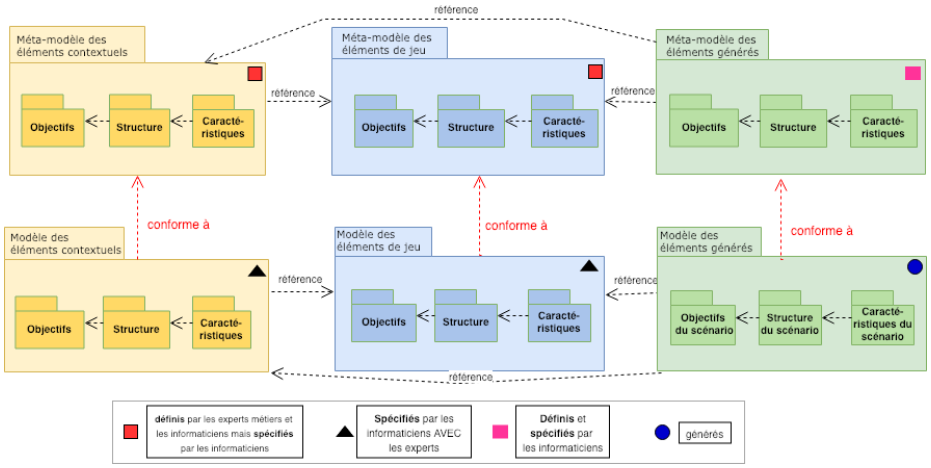


Figure 2. Approche de conception 3x3x2 proposée.

Les trois perspectives permettent de décomposer le problème de la conception de la génération en différentes étapes successives et incrémentales, chacune appréhendant un point de vue spécifique sur le scénario adapté à générer. Ces perspectives sont adaptées de la proposition de [14] :

— *Objectifs du scénario* : sélection ordonnées d'objectifs d'apprentissage, parmi ceux disponibles, en fonction de leur pertinence avec le contexte de la génération. Dans le projet *Escape it !*, cette perspective permettra de sélectionner quelles compétences visuelles sont concernées pour chacun des différents niveaux de la session de jeu, sur la base des compétences abordées, de leurs relations avec le profil de l'apprenant et en fonction de la progression concrète de l'utilisateur concerné.

— *Structure du scénario* : sélection ordonnée des exercices d'apprentissage ou des composants, à gros-grain, de jeu et d'apprentissage, en fonction des objectifs d'apprentissage fixés dans la perspective précédente. Dans *Escape it !* cela correspond, entre autres, au choix des "scènes de jeu".

— *Caractéristiques du scénario* : sélection des éléments additionnels qui participeront à préciser finement chaque élément structurel de la perspective précédente, en fonction également des objectifs fixés. Dans *Escape it !*, cela concerne la spécification détaillée de chaque scène de jeu, c-à-d toutes les informations requises (décors, objets, emplacements, etc.) par le moteur du jeu pour configurer une situation à résoudre.

Le scénario final généré sera donc composé de ces trois perspectives complémentaires. Pour chacune de ces perspectives nous proposons de considérer trois

dimensions complémentaires pour décrire les différents éléments enjeux dans la génération.

— les éléments à générer : par exemple le scénario adapté ;
— les éléments contextuels (ce qui est spécifique au contexte de la génération) : par exemple le profil de l'apprenant ;
— les éléments décrivant le jeu d'apprentissage : par exemple les objectifs concrets, les activités/exerciseurs disponibles, les éléments de configuration, etc.

Enfin, pour chaque perspective et dimension nous proposons de décrire les éléments enjeux selon deux niveaux d'abstraction :

— le niveau méta-modèle : ce niveau capture les éléments, propriétés et relations qui définissent les concepts et contraintes pour assurer la conformité des modèles ;
— le niveau modèle : ce niveau spécifie les informations concrètes nécessaires au générateur. Il est important de noter que le modèle correspondant au scénario à générer n'est pas spécifié avant la génération mais est bien produit en sortie de la génération. En revanche, le profil de l'utilisateur concerné par la génération est un modèle en entrée fournit au générateur. Les modèles décrivant le jeu d'apprentissage sont aussi des modèles en entrée. Le générateur exploite ces modèles du jeu pour l'ensemble des générations réalisées alors que le modèle de l'utilisateur peut varier.

Le processus de conception collaborative exploitant cette approche et permettant à des experts métiers d'échanger (identifier et spécifier) avec des informaticiens (réponse au défi 3) est présenté dans [10].

5 Application au projet *Escape it !*

Les prochaines sous-sections correspondent à la première perspective de l'approche (plus de détails sur les autres perspectives peuvent être obtenus dans [10]). Pour chaque dimension, nous présentons tout d'abord la description textuelle des éléments à considérer, à commencer par les éléments à générer qui dirigent l'analyse. Toutes ces informations sont issues des séances de conception collaborative que nous avons eu avec les experts autisme du projet.

5.1 Explicitation des éléments de la perspective *Objectifs*

La table 1 synthétise la description des éléments enjeux par dimension pour la perspective des objectifs. Il est intéressant de noter que l'expression des trois dimensions permet également de relever des *règles de génération et d'adaptation*. Ces règles ne peuvent pas être capturées par les trois dimensions car elles ne décrivent pas les données en entrée et en sortie de la génération mais la manière dont celle-ci devra fonctionner.

L'approche précédente est intéressante car elle permet de cadrer le périmètre des éléments à considérer pour chaque dimension. Cela permet également de mettre en évidence les règles qui seront utiles à la génération du scénario adapté.

Table 1. Éléments enjeux pour la perspective des objectifs.

Éléments descriptifs	- l'arbre des compétences visuelles, par lien de pré-requis - les 5 niveaux de difficulté (*débutant*, *élémentaire*, *intermédiaire*, *avancé*, *expert*)
Éléments contextuels	- nombre de niveaux à générer **N** (choisi par l'apprenant, entre 3 et 5) - les compétences concernées par l'enfant (certaines sont peut-être à exclure) - pour chaque compétence concernée : niveau de difficulté actuel et progression dans la difficulté courante
Éléments à générer	- les **N** compétences (ordonnées) à viser par chacun des N niveaux de la session de jeu - le niveau de difficulté pour chacune de ces compétences
Règles de génération et d'adaptation	- une compétence *éligible* pour le générateur est une compétence concernée et non déjà acquise dont les pré-requis ont au moins atteint le niveau de difficulté *intermédiaire* - le niveau de difficulté pour la compétence sélectionnée est identique à celui du profil - la sélection des **N** compétences se fait au hasard parmi les **M** éligibles ; si **M** < **N** alors la sélection des **(N-M)** restants recommencent parmi les M éligibles.

Ainsi, il est plus aisé de rejeter les informations métiers qui ne relèvent pas de la génération comme par exemple celles qui concernent davantage le temps de l'exécution du jeu (*runtime*). Par exemple, les informations régissant comment est réalisé la progression dans les niveaux de difficultés relèvent de la collecte et de la mise à jour du profil mais ne sont pas nécessaires à la génération.

5.2 Spécification des méta-modèles

Les informations collectées pour les trois dimensions peuvent alors être spécifiées (le plus formellement possible, au sens informatique du terme, c-à-d avec le moins d'ambiguïté sémantique pour un futur traitement machine). Le méta-modèle peut être considéré comme la représentation de la structure des modèles et des éléments composant ces modèles. Il permet la manipulation machine des modèles qui lui seront conformes.

La figure 3 est une représentation graphique des trois méta-modèles correspondants aux 3 dimensions de la perspective des objectifs du scénario. Pour les éléments à générer (à droite), l'*ObjectiveScenario* est composé de plusieurs *TargetSkill* (concrètement le modèle/scénario en aura autant que la valeur de *nbLevels*). Chaque *TargetSkill* référence une compétence *BxSkill* et précise la niveau de difficulté (propriété *difficultyLevel*). Les éléments de description du jeu d'apprentissage (au centre) pour cette perspective sont les compétences *BxSkill* et leurs relations *prerequisite* pour préciser l'éventuelle compétence parent d'une autre compétence. Enfin, les éléments de contexte (à gauche) décrivent le *Profile* d'apprentissage de l'enfant comme composé de plusieurs *Skill2Consider* référençant chacun une des compétences décrites et précisant ainsi pour chacun le niveau actuel de difficulté (*currentLevel*) et le statut en cours d'acquisition ou acquis de cette compétence.

5.3 Modèles

La figure 4 montre partiellement les modèles en entrée pour le générateur : la description du jeu (au centre) et un profil utilisateur (à gauche). Ces deux

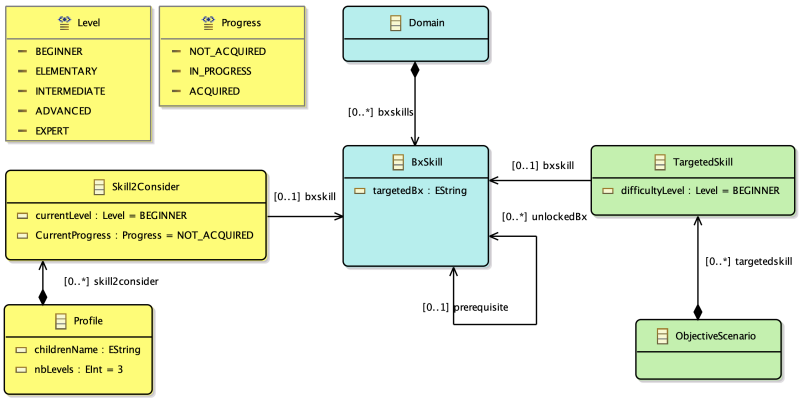


Figure 3. Représentation graphique des 3 méta-modèles correspondant aux 3 dimensions pour la perspective des *objectifs*.

modèles on été réalisés à l'aide d'un éditeur en arbre qui assure par construction la conformité des modèles à leur méta-modèle d'origine. On peut observer dans cet exemple que le jeu décrit 7 compétences (B3, B4, B9, B13, B19, B25) dont B8 (correspond au tri entre 2 catégories d'objets similaires selon [13]) a B3 pour pré-requis et est l'un des pré-requis de B13 (visibles dans la zone *Properties*). Le profil correspond ici à un enfant Tom qui souhaite 4 niveaux. Il est concerné par les 7 compétences, avec B3 et B9 acquis (non visible dans la figure), tandis que B4 (qui dépend de B3) est au niveau de difficulté *élémentaire* et les autres compétences (B8, B13 et B19) sont au niveau *débutant*.

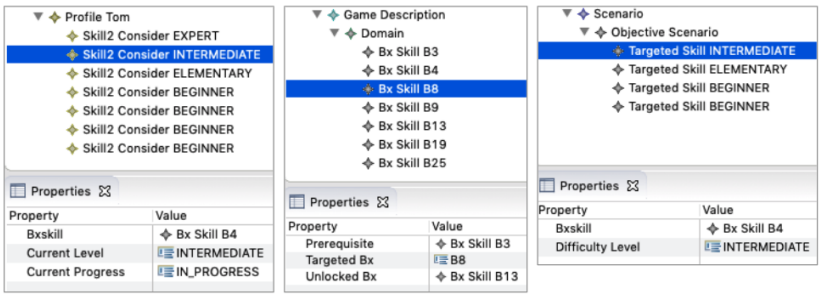


Figure 4. Exemples de modèles en entrée, un profil apprenant (à gauche), la description du jeu (au centre), et de modèle en sortie, un scénario adapté généré (à droite).

Le modèle à droite de la figure 4 est un exemple de scénario adapté qui est obtenu après génération (d'autres générations avec les mêmes modèles en entrée pourraient donner d'autres scénarios possibles). Pour cette génération, le scénario pour la perspective *objectifs* propose bien 4 compétences-cibles ordonnées :

B4 avec la difficulté *intermédiaire*, B9 (*élémentaire*), B8 (*débutant*), B8 (*débutant*). Cela est cohérent avec les règles de génération de la table 1 : seules les compétences B4-B8 (dépendantes de B3 qui est acquis) et B9 (dépendante de B4 à niveau intermédiaire) sont éligibles. Les 3 compétences ont été ensuite choisies dans un certain ordre, puis le hasard à re-sélectionner B8 pour le dernier niveau.

5.4 Mise en œuvre et validation

La génération de scénarios adaptés a été concrètement implémentée en tant que transformation de modèles, spécifiée en Java/EMF [17]. Ce framework prend en charge la manipulation des modèles en entrée et en sortie. Le développement du générateur peut donc se concentrer sur l'algorithme de génération qui doit intégrer les règles explicitées dans la table 1. Si l'on se réfère à la taxonomie des générateurs de contenu [18], notre implémentation peut être considérée *en ligne* (réalisée pendant l'exécution du jeu), *nécessaire* (le contenu généré doit être correct), *paramétrée* (le générateur connait les informations décrivant le jeu sous forme de modèle en entrée), *stochastique* (le hasard est utilisé pour départager des choix non couverts par une règle métier), *constructive* (l'algorithme du générateur ne produit jamais de contenu incohérent ou incomplet).

Le générateur a pu être testé et validé avec les experts lors des différentes sessions de conception collaborative selon les trois perspectives itératives. Des cas d'étude fictifs de profil d'enfant étaient spécifiés et utilisés pour générer plusieurs scénarios. Le générateur était alors utilisé comme *simulateur* afin de valider les règles métier de génération (certaines ont été supprimées, d'autres ajoutées ou ajustées). Cette validation, et le processus de conception entre informaticiens et experts qui exploite notre approche, sont présentés dans [10]. Cet article s'est focalisé sur la présentation de l'approche et sa validation par sa mise en application sur un cas concret de jeu sérieux. Toutefois, notre approche n'est pas seulement applicable aux jeux sérieux : elle a pu être appliquée pour la génération d'activités d'inventaires botaniques adaptées au contexte du citoyen (relevés à proximité et préférences d'inventaires du citoyen) dans le cadre du projet REVERIES.

Lorsque le développement du jeu d'apprentissage d'*Escape It !* a été réalisé (avec le moteur de jeu Unity), nous avons souhaité ne pas développer à nouveau le générateur en interne du jeu. Nous avons préféré intégrer le générateur sous forme de service Web appelé par le jeu. Cela permet de découpler le jeu des stratégies de génération qui deviennent ainsi interchangeables. Des expérimentations sont en cours afin d'améliorer l'utilisabilité du jeu et de vérifier son utilité. En ce qui concerne l'utilisation du générateur intégré, les expérimentations déjà réalisées n'ont pas révélées de situations où les scénarios proposés n'étaient pas adaptés.

6 Conclusions et perspectives

Nous avons proposé une approche de conception IDM pour la génération de scénarios adaptés dans des jeux d'apprentissage. Elle permet d'appréhender la spécification des différents modèles et méta-modèles selon trois perspectives (points

de vue incrémentaux sur les éléments à générer) et trois dimensions (les éléments à générer, les éléments décrivant le contexte de la génération, les éléments décrivant le jeu d'apprentissage). Nous avons appliqué cette approche dans le contexte du projet *Escape it !*. Cela nous a permis de concevoir un générateur de scénarios de scènes de jeu et de leurs configurations, adaptés au profil d'apprentissage des enfants. Seule la première perspective des *objectifs* du scénario est présentée dans cet article. Le générateur conçu a été concrètement développé, intégré au jeu d'apprentissage et testé avec succès.

L'approche de conception proposée semble convenir, d'après l'expérimentation réalisée dans le cadre du projet *Escape it !*, dans des contextes de jeu d'apprentissage où une seule préoccupation d'adaptation, en l'occurrence l'apprentissage, est prise en compte au niveau des éléments contextuels pour la génération. Nous travaillons actuellement dans le contexte du projet *REVERIES*[12] à une nouvelle approche qui permettrait de prendre en compte différentes préoccupations (apprentissage, ludification et inventaire citoyen).

Bien que notre proposition permette de faciliter la conception de la génération de scénarios adaptés, elle ne permet pas de faciliter la mise en oeuvre concrète des règles de génération, qui reste une tâche de programmation non négligeable. L'analyse de [5] relève que "la plupart des travaux sur les jeux sérieux n'identifient pas explicitement un modèle pour représenter les règles de l'adaptation". Nous travaillons ainsi actuellement à considérer la transformation de modèle non comme un problème algorithmique de programmation mais comme un problème de satisfaction de contraintes.

Références

1. Bielikova, M., Diveky, M., Jurnecka, P., Kajan, R., Omelina, L. : Automatic generation of adaptive, educational and multimedia computer games. vol. 2, pp. 371—-384. Signal Image Video Process. (2008)
2. Burton, L.R., McEachin, J. : A Work in Progress : Behavior Management Strategies and a Curriculum for Intensive Behavioral Treatment of Autism. New York : DRL Books (1999)
3. Delmas, G., Champagnat, R., Augeraud, M. : Plot monitoring for interactive narrative games. In : Proceedings of the International Conference on Advances in Computer Entertainment Technology. pp. 17–20. ACE '07, ACM, NY, USA (2007)
4. Ern, A. : The use of gamification and serious games within interventions for children with autism spectrum disorder (January 2014), http ://essay.utwente.nl/64780/
5. Hocine, N., Gouaich, A., Di Loreto, I., Abrouk, L. : Techniques d'adaptation dans les jeux ludiques et sérieux. Revue des Sciences et Technologies de l'Information - Série RIA : Revue d'Intelligence Artificielle, Lavoisier **25**(2), 253–280 (2011)
6. Hussaan, A.M., Sehaba, K. : Consistency verification of learner profiles in adaptive serious games. In : 11th European Conference on Technology Enhanced Learning. pp. 384–389. Springer International Publishing (2016)
7. Janssens, O., Samyn, K., Van de Walle, R., Van Hoecke, S. : Educational virtual game scenario generation for serious games. In : Proceedings of the IEEE 3rd International Conference on Serious Games and Applications for Health (2014)

8. Jézéquel, J.M., Combemale, B., Vojtisek, D. : Ingénierie Dirigée par les Modèles : des concepts à la pratique... Références sciences, Ellipses (Feb 2012)
9. Koegel LK, Ashbaugh K, K.R. : Pivotal response treatment. In : Early Intervention for Young Children with Autism Spectrum Disorder. pp. 85–112. Springer International Publishing (2016)
10. Laforcade, P., Laghouaouta, Y. : Supporting the adaptive generation of learning game scenarios with a model-driven engineering framework. In : Lifelong Technology-Enhanced Learning (ECTEL'18). pp. 151–165. Springer International Publishing, Cham (2018)
11. Lopes, R., Bidarra, R. : Adaptivity challenges in games and simulations : A survey. IEEE Transactions on Computational Intelligence and AI in Games **3**(2), 85–99 (2011). https://doi.org/10.1109/TCIAIG.2011.2152841
12. Marfisi-Schottman, I., Gicquel, P.Y., Karoui, A., George, S. : From idea to reality : Extensive and executable modeling language for mobile learning games. In : Verbert, K., Sharples, M., Klobučar, T. (eds.) Adaptive and Adaptable Learning. pp. 428–433. Springer International Publishing, Cham (2016)
13. Partington, J., Analysts, P.B. : The Assessment of Basic Language and Learning Skills-revised (the ABLLS-R) (2010)
14. Sebaha, K., Mahmood Hussaan, A. : Architecture et modèles génériques pour la génération adaptative des scénarios de jeux sérieux. application : Jeu d'évaluation et de rééducation cognitives. Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation **21**(1), 615–648 (2014)
15. Sehaba, K., Hussaan, A. : Goals : generator of adaptive learning scenarios. In : International Journal of Learning Technology. vol. 8, pp. 224–245 (10 2013)
16. Sina, S., Rosenfeld, A., Kraus, S. : Generating content for scenario-based serious-games using crowdsourcing (2014)
17. Steinberg, D., Budinsky, F., Paternostro, M., Merks, E. : EMF : Eclipse Modeling Framework 2.0. Addison-Wesley Professional, 2nd edn. (2009)
18. Togelius, J., Yannakakis, G.N., Stanley, K.O., Browne, C. : Search-based procedural content generation : A taxonomy and survey. IEEE Transactions on Computational Intelligence and AI in Games **3**(3), 172–186 (2011). https://doi.org/10.1109/TCIAIG.2011.2148116
19. Vandewaetere, M., Desmet, P., Clarebout, G. : The contribution of learner characteristics in the development of computer-based adaptive learning environments. Computers in Human Behavior **27**(1), 118 – 130 (2011). https://doi.org/https ://doi.org/10.1016/j.chb.2010.07.038
20. Zakari, H.M., Ma, M., Simmons, D. : A review of serious games for children with autism spectrum disorders (asd). In : International Conference on Serious Games Development and Applications. pp. 93–106. Springer International Publishing (2014)
21. Zniber, N., Cauvet, C. : Systèmes pédagogiques adaptatifs : état de l'art et perspectives. In : MajecSTIC 2005 : Manifestation des Jeunes Chercheurs francophones dans les domaines des STIC. pp. 300–315 (2005)

Réalité augmentée pour l'apprentissage conceptuel en sciences : quels principes de conception pour les EIAH ? Cas du dispositif DEAPE Learn en électromagnétisme

Julien Da Costa[1], Nicolas Szilas[1], Andreas Mueller[2]

[1]Université de Genève, TECFA - FPSE, 42 bd Pont-d'Arve, CH-1211 Genève 4, Suisse.
Julien.dacosta@unige.ch, Nicolas.Szilas@unige.ch
[2]Université de Genève, IUFE Sciences, 42 bd Pont-d'Arve, CH-1211 Genève 4, Suisse.
Andreas.mueller@unige.ch

Résumé. DEAPEL est un dispositif de réalité augmentée spatiale visant à soutenir la compréhension conceptuelle des phénomènes de champs électriques et magnétiques. Nous présentons ici un cadre théorique liant l'étude de la réalité augmentée aux recherches en apprentissage multimédia, changement conceptuel et didactique des sciences ainsi que ses implications pour la conception d'EIAH en réalité augmentée. Nous présenterons l'artefact issu de cette perspective et discuterons des observations issues de démonstrations publiques du dispositif.

Mots-clés. Réalité augmentée ; Apprentissage Multimédia ; changement conceptuel ; didactiques des sciences; Conception des EIAH ;

Abstract. DEAPEL is a spatial augmented reality device designed to support the conceptual understanding of electric and magnetic field phenomena. We built a theoretical framework linking the study of augmented reality to researches in multimedia learning, conceptual change and didactics. We outline design principles for TEL development with augmented reality, present the artefact and discuss preliminary observations from public demonstrations of the device.

Keywords. Augmented reality; Multimedia Learning; conceptual change; sciences didactics; TEL development;

1 Introduction

Les environnements informatiques pour l'apprentissage humain (EIAH) sont des environnements complexes, articulant des travaux liés à l'informatique et à l'apprentissage. La recherche en EIAH est largement expérimentale, théories et applications avancent ensemble, prenant en compte simultanément l'objectif didactique de l'artefact et son ingénierie [1]. La transposition des principes de psychologie cognitive et des didactiques disciplinaires au domaine des EIAH offre des lignes de conduite *permettant de dépasser le traitement ad hoc des problèmes* [2]. Nous présentons ici un dispositif d'expérimentation augmentée des phénomènes électromagnétiques fondé sur une méthodologie de type *Design Science*[1] [3].

[1]Cette approche s'appuie sur des recommandations de recherches pour développer un artefact innovant visant la résolution d'un problème identifié en apportant une contribution prescriptive

Le dispositif prend comme point de départ les recherches en didactiques des sciences indiquant que les étudiants disposent de connaissances parcellaires sur les concepts de base de l'électromagnétisme, abordent l'apprentissage des phénomènes avec des représentations alternatives aux modèles scientifiques, utilisent des définitions et des formules mathématiques routinières sans réelle compréhension conceptuelle [4, 5]. Il a notamment été mis en évidence que les élèves peuvent calculer une différence de potentiel dans des schémas complexes et abstraits, mais ne peuvent prédire ce qui se produit sur le courant, la tension, la luminosité d'une ampoule dans des circuits beaucoup plus simples utilisant une iconographie réaliste [6]. On observe que les étudiants de niveau secondaire comme universitaire ont des scores faibles sur les échelles de test conceptuel et même les étudiants ayant les meilleurs scores aux examens montrent des erreurs de compréhension sur les concepts centraux [7].

La réalité augmentée (RA), en associant/superposant en temps réel des modélisations scientifiques aux phénomènes physiques observables, est souvent présentée comme une technologie possédant un grand potentiel pour résoudre ces difficultés [8]. Toutefois peu de travaux l'ont encore investiguée sous l'angle du changement conceptuel et les principes de conception d'EIAH dans cette perspective restent limités. C'est pourquoi nous développons dans un premier temps un cadre d'étude pour l'étude de la RA sous la perspective du changement conceptuel. Nous présentons dans un second temps ces implications dans le développement de DEAPE Learn un dispositif de réalité augmentée spatiale soutenant la compréhension des phénomènes de champs électriques et magnétiques et généralisons un modèle pour le développement d'EIAH sous cette perspective. Enfin nous discuterons des observations préliminaires issues de démonstrations publiques du dispositif.

2 Cadre théorique

2.1 Réalité augmentée en contexte éducatif

La « réalité augmentée» est classiquement définie comme une sous-composante du concept de réalité mixte [9], décrivant une forme d'interaction entre l'utilisateur et la machine basée sur l'association plus ou moins prononcée d'objets réels et virtuels (numériques). Nous la présentons dans un contexte éducatif comme une interface *énactive*, permettant la superposition *sémantique*, spatiale et temporelle de multiples représentations sémiotiques à un contexte physique tangible et situé[2]. Cette vision n'est pas liée à une quelconque technologie et révèle les caractéristiques sémio-cognitives d'un EIAH avec une forte composante interface-homme-machine médiant l'interaction et la compréhension du réel (fig. 1). Malgré l'ancienneté de la RA son usage en éducation reste relativement récent et son potentiel loin d'être complètement étudié. Plusieurs méta-analyses se sont intéressées aux affordances de la RA pour l'éducation [10–12] et soulignent la difficulté de comparaison des dispositifs tant ces derniers sont divers du point de vue des technologies utilisées, des finalités et contextes d'usage.

[2] *Enactive* car centrée sur l'interaction entre cognition, action et environnement ; *sémantique* car il existe une certaine relation de sens entre visualisation numérique et le contexte.

L'utilisabilité du système est souvent une composante clé et de nombreuses études relatent que les dispositifs à l'appropriation complexe se révèlent pénalisants pour l'apprentissage [12].

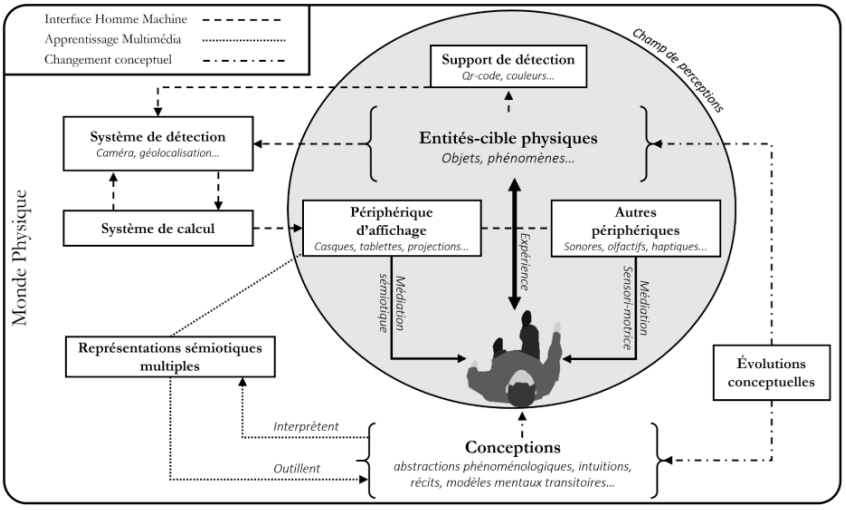


Fig. 1. Modélisation de notre cadre de référence

Dans le contexte de l'enseignement des sciences [8] la RA est principalement utilisée au sein de dispositifs de type *apprentissage par problèmes* ou *investigation*. Les résultats obtenus du point de vue de la compréhension conceptuelle sont encourageants [13] et peuvent se révéler supérieur à des dispositifs utilisant des simulations 2D ou 3D [14]. Certains travaux [15–17] constatent notamment que la manipulation de marqueurs en RA réduit le niveau d'abstraction en fournissant des référents scientifiques et spatiaux concrets pour l'apprenant, facilite la manipulation et augmente la concentration sur l'explication scientifique du phénomène étudié. L'efficacité pédagogique semble cependant fortement dépendante du design et de la scénarisation des activités, il semble ainsi nécessaire de resserrer les liens entre les théories de sciences de l'éducation et le développement d'environnement RA [18].

2.2 Changement conceptuel et modèle d'intégration de connaissance

Pour le champ de recherche en changement conceptuel [19], tout individu se représente la réalité et les représentations mentales constituent le point d'ancrage à toute construction du savoir. Ces représentations mentales (sous forme fragmentaires ou organisées) sont observables (entretiens, traces, productions) de manière récurrente chez les apprenants mis en situation de manière identique et doivent nous donner matière à penser la conception des environnements d'apprentissage.

Quel que soit le domaine, la formation, et l'âge de l'individu la formation de représentations erronées d'un point de vue scientifique fait partie du processus « normal » de construction de connaissances [20]. Le terme « erroné » n'est pas

toujours « intrinsèque » au concept mais une erreur dans sa mise en œuvre – l'apprenant activant spontanément des conceptions en réponse à certaines situations et à sa stratégie de lecture de celles-ci. Il n'y a donc pas nécessairement rupture entre conceptions naïves et expertes, certaines ne sont pas fausses (car valides dans de nombreux contextes), mais souvent insuffisamment coordonnées, priorisées et contextualisées par les étudiants. L'expression même de « changement » conceptuel peut prêter à débat lorsque de récent travaux en neurosciences tendent à démontrer que les conceptions naïves ne disparaissaient jamais du cerveau des experts, mais « cohabitent », et seraient simplement inhibées plus efficacement par ces derniers [21, 22]. Si les conflits cognitifs sont inhérents à l'acquisition de connaissances, leur résolution est un processus plus complexe, idiosyncratique et sensible au contexte. Il ne s'agit pas tant d'aider l'étudiant à *changer* ses conceptions mais à *adopter* de nouvelles conceptions en connaissant leurs domaines de validité et en reconnaissant les situations où elles s'appliquent et celles où elles ne s'appliquent pas (inhibition). Un modèle d'intégration de connaissance [23] consiste à concevoir des scénarios d'enseignement s'attachant à faire évoluer un ensemble plus ou moins disjoints d'idées sur les phénomènes à travers différents contextes sensibles et identifiables par l'apprenant. A cette fin les technologies éducatives jouent un rôle important [24] par leur capacité à représenter, simuler, étayer, « rendre visible l'invisible ». Pour autant la manière dont les étudiants interagissent et « pensent » avec les visualisations numériques, de surcroit dans un environnement de réalité physique et virtuel n'est pas neutre et est discutée ci-après.

2.3 Enjeux des visualisations en sciences

La nature des phénomènes en sciences ne nous étant pas directement perceptible et compréhensible, la recherche développe des modèles pour l'appréhender et décrit ces derniers à l'aide de représentations sémiotiques structurées et conventionnelles (textes, équations mathématiques…) ou de relations visuelles analogiques (dessins, maquettes…) [25]. Ces représentations permettent de décrire un modèle scientifique sous trois niveaux : macroscopique (ce que l'on voit, sent, expérimente), sub-microscopique (entités qui sous-tendent le niveau macroscopique et donne lieu à ce qui est perçu), symbolique (abstraction souvent mathématique du phénomène) [26].

L'importance du rôle des visualisations dans la conceptualisation est aujourd'hui bien établie [27]. Ces dernières sont des outils pour « penser » les phénomènes, faire des inférences, résoudre des problèmes [28]. De manière plus large, les visualisations sont des objets épistémiques faisant partie intégrante du processus de la science[3]. Elles ne sont toutefois pas une « panacée » car l'enseignement de la physique est complètement multimédia et apprendre demande aux étudiants de pouvoir générer, interconnecter, traduire et utiliser adéquatement plusieurs représentations sémiotiques pour résoudre des problèmes (fig. 2). La capacité de passer d'un système de représentation à un autre est très souvent le seuil critique de progrès dans la résolution de problèmes [29] et un nombre croissant de recherches en soulignent l'enjeu en enseignement des sciences [30]. Il est notamment constaté [31] que face à de multiples représentations d'un phénomène physique les étudiants : (1) n'utilisent qu'un seul type

[3] En électromagnétisme les lignes de champs en sont un exemple emblématiques aussi bien par le rôle qu'elles ont joué dans l'avancé des savoirs que par leur usage en enseignement.

de représentation (la plus familière et concrète pour eux) ; (2) ne permutent entre les représentations qu'en cas de problème avec la représentation en cours ; (3) ont besoin d'être soutenus dans leurs processus de formation de cohérence[4].

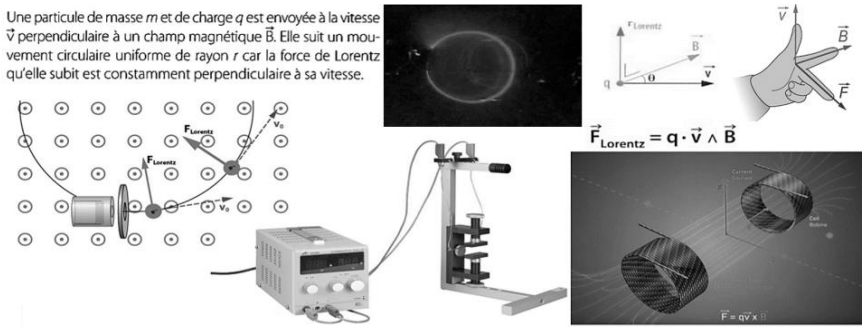


Fig. 2. Représentations de la force de Lorenz à interconnecter par un étudiant.

Plus généralement une littérature extensive dans le domaine de l'apprentissage multimédia a démontré que la confrontation des apprenants à de multiples représentations n'est pas automatiquement efficace. D'une part car les représentations ne répondent pas toujours à certains principes de conception respectant des considérations cognitives et contextuelles. D'autre part car l'apprentissage multimédia est exigeant et conséquemment implique de réduire au minimum le nombre de représentations nécessaires à l'atteinte de l'objectif d'apprentissage. Les principes de conception issues de la recherche empirique en apprentissage multimédia [32] offre sur ce sujet un cadre conceptuel bien étayé pouvant être transposé aux environnements RA.

3 Dispositif DEAPE Learn

Dans le cadre de notre problématique initiale nous avons développé un prototype de réalité augmentée spatiale permettant d'associer différentes représentations scientifiques aux phénomènes visibles des champs électriques et magnétiques. Le système offre la possibilité de manipuler les véritables objets visés par la connaissance (aimants, boussoles, circuits de courant câblés) et de projeter en temps réel (~45 images/secondes) textes, formalismes mathématiques et visualisations de manière dynamique et spatialement cohérente (précision d'environs 5 millimètres) (fig. 3). Les projections du dispositif sont basées sur des modèles mathématiques avancés afin de correspondre le plus fidèlement possible aux phénomènes observés en tout point de l'espace. La reconnaissance des objets est assurée par un système robuste de suivi de couleur et la calibration caméra-vidéoprojecteur est assurée par le kit de développement sous licence libre PapArt - Paper Augmented Reality Toolkit [34].

[4] Processus cognitif consistant à créer des connections référentielles entre divers éléments correspondants et différentes structures correspondantes au sein de multiples représentations

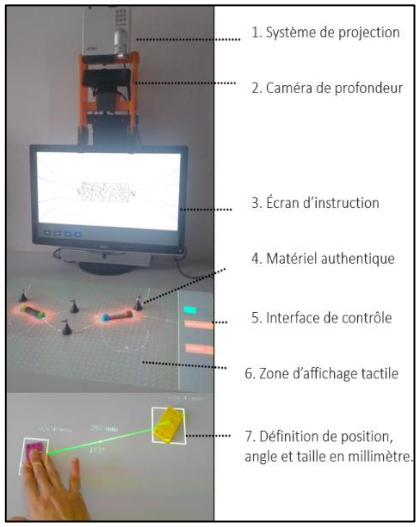


Fig. 3. Artefact technique

Afin d'assurer le guidage de l'apprenant (ou groupe de 2 ou 3 apprenants selon le scénario) le système dispense des consignes orales ou écrites sur l'écran (hors-zone de projection). Lors des manipulations expérimentales, l'élève peut permuter entre différents types de représentations soit par l'ajout d'un objet nouveau – connu du système – dans la zone de projection, soit par contacts tactiles sur la table (un contact du doigt étant reconnu comme un clic).

Principes de conception

Ce dispositif à la particularité d'intégrer dans son design notre paradigme d'étude en focalisant son intervention à quatre niveaux (fig. 4). Le modèle propose de partir des difficultés conceptuelles (A), de construire sur l'expérience sensible des phénomènes associés (B), puis de scénariser les interactions entre entités numériques et physiques (C) en contrôlant la charge cognitive associée à la manipulation du dispositif et du traitement multimédia (D).

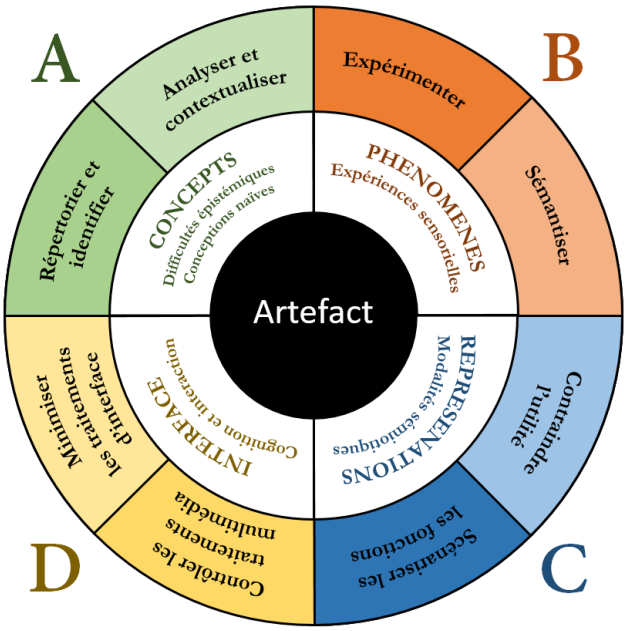


Fig. 4. Domaines d'interventions de notre paradigme d'étude

A. Partir des difficultés conceptuelles

A.1 Répertorier et identifier. Il s'agit d'identifier les difficultés conceptuelles à prendre en compte dans la conception du dispositif. Ces dernières sont bien documentées et une revue des tests standardisés, de la littérature et des bases de données didactiques suffit à construire un répertoire propre au sujet d'étude. Dans le cadre de DEAPE Learn citons notamment [35, 36] et les tests standardisés [7, 37]. Il peut toutefois se révéler nécessaire de concevoir une enquête préalable par entretiens pour identifier les difficultés conceptuelles sur des sujets moins documentés.

A.2 Analyser et contextualiser. L'étape suivante consiste à identifier les phénomènes physiques observables associés aux difficultés identifiées. Il s'agit de lier le dispositif à l'observation quotidienne, à l'expérience sensible de l'apprenant. Dans le cadre de DEAPE Learn l'attraction/répulsion des aimants, l'interaction des boussoles, la réaction de différents matériaux, le champ magnétique généré par une bobine de courants, sont autant de composantes phénoménologiques expérimentables.

B. Construire sur l'expérience sensorielle des phénomènes. La recherche suggère que les représentations physiques et virtuelles ont des capacités cognitives complémentaires pour l'apprentissage conceptuel des étudiants [38]. Sans être suffisante en elle-même, l'expérience reste centrale pour aider à dépasser certaines croyances sur les phénomènes physiques et permet de réifier certaines abstractions en les rattachant à des situations concrètes pour l'apprenant.

B.1 Expérimenter. L'expérimentation se distingue de la simulation. Cette dernière utilise un modèle computationnel qui simplifie le phénomène étudié et en contraint l'usage. Il est ainsi impossible d'en détourner la finalité (impossible d'ajouter un aimant si la simulation n'en prévoit que deux, de mettre spontanément un autre objet entre les deux aimants si le scénario ne le permet pas) ou de mettre en défaut le modèle (l'aimant est idéalisé par rapport à un véritable aimant). Dans notre approche le phénomène est authentique (donc présenté dans toute sa complexité). La manière dont l'apprenant interagit avec l'objet d'étude est beaucoup plus idiosyncratique et reflet de sa compréhension et interprétation de la situation.

B.2 Sémantiser. La relation de sens entre l'objet physique manipulé et les visualisations numériques (fig. 5, à droite) est essentielle dans cette approche et se distingue fortement des dispositifs dont l'objet manipulé (ex : QR-tag) a pour seul fonction de fournir un support à l'intégration spatiale des représentations numériques (fig. 5, à gauche).

Fig. 5. (Gauche) Affichage d'un circuit électrique en manipulant des supports de détection papier. (Droite) lien sémantique dans DEAPE Learn entre aimants authentiques et éléments virtuels

C. Scénariser l'interaction des représentations. La mise en lien de multiples représentations d'un phénomène est nécessaire à la compréhension, cependant cette interconnexion est notoirement difficile pour les étudiants. Il convient donc de réduire le nombre de représentations simultanées au minimum, de favoriser les permutations et de rendre saillant les liens entre-elles (dyna-linking) [33].

C.1 Contraindre l'utilité des représentations. La situation didactique doit contraindre l'apprenant à faire usage de représentations différentes pour résoudre un problème. Contrairement à un usage répandu il ne s'agit pas d'associer les représentations et de les animer conjointement mais plutôt de laisser l'apprenant permuter entre chaque en le guidant vers un usage approprié en fonction de la situation à analyser.

C.2 Scénariser les interactions. Si différentes représentations sont nécessaires à une compréhension effective d'un concept, c'est que ces dernières possèdent des fonctions distinctes et des relations fonctionnelles identifiables. Le modèle DeFT développé par Ainsworth [33] (Framework : Design - Function – Task) offre un cadre pour comprendre la manière dont elles s'agencent pour véhiculer du sens et peut être adapté au cadre de la RA. Les représentations numériques dispensées par le système doivent se compléter. Dans le cadre de la RA le phénomène physique étudié peut s'insérer dans ce modèle et les différentes représentations sémiotiques ont pour vocation de *compléter* l'information sur le phénomène, *contraindre* son interprétation ou *approfondir* sa compréhension (fig. 6).

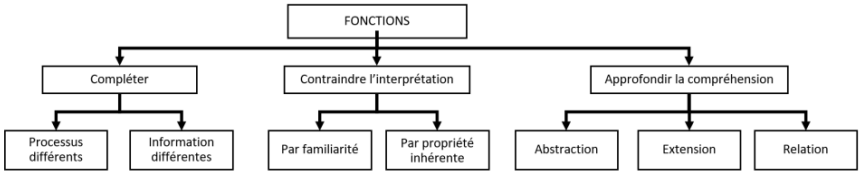


Fig. 6. Rôles fonctionnels principaux des multiples représentations d'après [33]

D. Limiter les traitements cognitifs de l'interface. Un apprentissage signifiant demande l'exploitation de ressources cognitives importantes et toutes ressources mobilisées à la compréhension et manipulation du système de RA se réalise au détriment de l'apprentissage. Nous abordons ce principe selon les deux dimensions suivantes

D.1 Minimiser les traitements d'interface. Consiste à utiliser une interface la plus transparente possible pour l'utilisateur afin de réduire au minimum les traitements cognitifs relatif à son usage. Il convient de limiter les éléments d'interface affichés simultanément et à laisser dans la mesure du possible place à l'interaction physique par manipulation d'objet concret. Pour notre prototype cela se caractérise à titre d'exemple par l'emploi de projections afin de ne pas avoir à manipuler un autre objet que le matériel d'expérience (comme une tablette ou un smartphone) ou à baser le dispositif sur un suivi de couleur pour éviter l'usage de QR-Tag intégrant le champ de perception de l'utilisateur et réduisant l'authenticité de la tâche.

D.2 Contrôler les traitements cognitifs multimédia. A partir des théories cognitives du traitement de l'information, les recherches en apprentissage multimédia ont permis d'identifier une série de principes cadres pour la conception d'environnement pédagogique. Ces principes sont référencés [32] et mis à jours en fonction de résultats d'études empiriques. Sans pouvoir les détailler ici la conception de notre prototype a fait particulièrement usage des principes de *contiguïté*, *modalité*, *cohérence*, *signalement*, *segmentation* et *pré-entrainement* définis par ce domaine. A titre d'exemple, le principe de *contiguïté spatiale et temporelle* – postulant que l'apprentissage est facilité lorsque des informations concordantes sont à proximité – est ainsi aisément exploitable en RA. Dans DEAPE Learn les visualisations de champs sont directement projetées autour des objets expérimentaux et facilitent le traitement spatial de leur mise en lien.

4 Observations préliminaires et discussion

L'interaction avec le dispositif de participants d'âge et de niveau de formation différents a pu être observée au cours de 3 démonstrations publiques de notre prototype. L'activité proposée concernait le fonctionnement d'un aimant permanent au niveau atomique (spin électronique), cristallin (domaines magnétiques) et macroscopique (orientation des pôles) ainsi que la nature du champ magnétique généré. Les participants pouvaient déplacer des aimants droits sur une surface de projection, ajouté des aiguilles aimantées et de la poudre ferromagnétique. Le dispositif permettait d'afficher l'orientation du champ magnétique terrestre, de projeter en fonction de la position et l'orientation des aimants des représentations numérique de leurs structures internes, de la position de leurs pôles, ou encore de leurs champs magnétiques sous forme de vecteurs, de lignes de champ ou d'une carte d'intensité (aperçue à droite fig. 5). L'activité donnait la possibilité d'observer le comportement d'aiguilles magnétiques ou de poudre ferromagnétique et la concordance ou non concordance avec les représentations scientifiques choisies. Ces confrontations publiques nous permettent de partager les observations suivantes :

- Le système se révéla propice à faire surgir les conceptions erronées des utilisateurs sur leur compréhension des aimants. Le fait que tous les métaux ne soient pas aimantés (a), que le magnétisme ne soient pas liés à la répartition de charges électriques positives et négatives à l'intérieur d'un aimant (b), que le champ généré par un aimant se révèle supérieur en intensité à celui de la terre (c) ou que l'orientation de la force d'attraction ne soient pas toujours dirigée de l'objet à l'aimant (d) sont des exemples de « misconceptions » identifiables dans le discours des participants.
- Les manipulations réalisées par les participants donnèrent lieu à des observations non anticipées. Cinq participants au cours des 3 présentations ont réalisé de multiples tentatives de positionnements « extrêmes » pour trouver un moyen de faire tourner une aiguille indéfiniment[5], Plusieurs participants ont également tenté sans succès de mettre en défaut le modèle en trouvant des positionnements où

[5] En contradiction avec le deuxième principe de la thermodynamique ou le fait que la force magnétique ne produise aucun travail.

l'orientation des aiguilles ne concorderait pas avec l'orientation des lignes de champs projetées numériquement. Deux participants nous interrogèrent sur la nature des « micromoteurs » qu'ils pensaient dissimulés à l'intérieur des supports des aiguilles, révélant ainsi leurs difficultés à concevoir la notion de champ magnétique en tout point de l'espace.

- L'usage de phénomènes authentiques donna lieu à divers détournements, les participants sortant spontanément des clés de voiture pour les poser sur la surface de projection ou utilisant des stylos entre deux aimants pour stopper l'attraction ou répulsion magnétique.
- La manipulation « naturelle » d'objets physiques semble avoir un effet bénéfique sur l'appropriation du dispositif par les participants mais la grande transparence de l'interface a pu perturber certains. Plusieurs utilisateurs semblent avoir eu du mal à distinguer ce qui relevait de la réalité physique du phénomène de ce qui relevait des augmentations numériques dispensées par le dispositif.
- Certaines manipulations, aisées de manière simulée à l'écran, peuvent être délicates en réalité augmentée. Par exemples les aimants s'attirent ou se repoussent et certaines configurations géométriques demandent de trouver des « astuces » pour les maintenir en place.

5 Conclusion

La présente étude contribue à la recherche sur le développement des EIAH en réalité augmentée dans lesquelles les étudiants construisent des liens entre des concepts nouveaux et leur environnement. Au niveau théorique, cette étude propose un cadre d'analyse issu des champs de recherche en apprentissage multimédia et didactique des sciences. Sans être exhaustif ce cadre éclaire l'étude de la réalité augmentée sous un angle sémio-cognitif et aborde son usage en science sous la perspective du changement conceptuel. Ce cadre comporte des implications en termes de choix de design pour la conception d'EIAH en enseignement des sciences. Nous avons développé des recommandations sur 4 niveaux de conception partant de l'identification des difficultés épistémiques du sujet d'étude jusqu'au traitement cognitif de l'interface. Ces principes basés sur des postulats de recherche offrent des lignes de conduites pour le concepteur d'EIAH devant nécessairement être arbitrées en fonction des contraintes techniques, contextuelles, institutionnelles qu'un projet d'ingénierie comporte.

La mise en œuvre de notre cadre d'analyse au sein d'un premier prototype de réalité augmentée spatiale sur les phénomènes de champs nous conforte dans sa portée pratique et les observations préliminaires nous enjoignent à étendre son développement. Des études expérimentales en cours de réalisation nous permettront prochainement de tirer des conclusions plus approfondies sur l'efficacité pédagogique de cette approche.

Références

1. Tchounikine, P., Tricot, A.: Environnements informatiques et apprentissages humains. In: Catherine Garbay, D.K. (ed.) Informatique et sciences cognitives : Influences ou

confluence ? pp. 153–186 (2011).
2. Tchounikine, P.: Quelques éléments sur la conception et l ' ingénierie des EIAH. In: 2ème assises nationales du GdR I3-Groupe de Recherche Information Interaction Intelligence. pp. 1–13 (2002).
3. Dresch, A., Lacerda, D.P., Antunes, J.A.V.: Design science research: A method for science and technology advancement. (2015).
4. Bagheri-Crosson, R., Venturini, P., Lefevre, R.: Le concept de champ magnétique et les phénomènes associés chez des étudiants du premier cycle à l'université. Dossiers des Sci. l'éducation. (2002) 23–32.
5. Chabay, R., Sherwood, B.: Restructuring the introductory electricity and magnetism course. Am. J. Phys. Vol.74. (2006) 329–336.
6. Finkelstein, N.: Learning Physics in Context: A study of student learning about electricity and magnetism. Int. J. Sci. Educ. Vol.27. Routledge (2005) 1187–1209.
7. Planinic, M.: Assessment of difficulties of some conceptual areas from electricity and magnetism using the Conceptual Survey of Electricity and Magnetism. Am. J. Phys. Vol.74. American Association of Physics Teachers (2006) 1143–1148.
8. Cheng, K.-H., Tsai, C.-C.: Affordances of Augmented Reality in Science Learning: Suggestions for Future Research. J. Sci. Educ. Technol. Vol.22. (2012) 449–462.
9. Milgram, P., Kishino, F.: Taxonomy of mixed reality visual displays. IEICE Trans. Inf. Syst. Vol.E77–D. (1994) 1321–1329.
10. Bacca, J., Baldiris, S., Fabregat, R., Graf, S., Kinshuk: Augmented Reality Trends in Education: A Systematic Review of Research and Applications. Educ. Technol. Soc. Vol.17. (2014) 133–149.
11. Radu, I.: Augmented reality in education: a meta-review and cross-media analysis. Pers. Ubiquitous Comput. Vol.18. (2014) 1533–1543.
12. Akçayır, M., Akçayır, G.: Advantages and challenges associated with augmented reality for education: A systematic review of the literature. Educ. Res. Rev. Vol.20. Elsevier (2017) 1–11.
13. Sirakaya, M., Cakmak, E.K.: The Effect of Augmented Reality Use on Achievement, Misconception and Course Engagement. Contemp. Educ. Technol. Vol.9. Contemporary Educational Technology. Faculty of Communication Sciences (2018) 297–314.
14. Yen, J.-C., Tsai, C.-H., Wu, M.: Augmented Reality in the Higher Education: Students' Science Concept Learning and Academic Achievement in Astronomy. Procedia - Soc. Behav. Sci. Vol.103. Elsevier (2013) 165–173.
15. Shelton, B.E., Stevens, R.R.: Using coordination classes to interpret conceptual change in astronomical thinking. International Society of the Learning Sciences (2004) 634–642.
16. Kerawalla, L., Luckin, R., Seljeflot, S., Woolard, A.: "Making it real": exploring the potential of augmented reality for teaching primary school science. Virtual Real. Vol.10. Springer-Verlag (2006) 163–174.
17. Fleck, S., Hachet, M., Bastien, J.M.C.: Marker-based augmented reality. In: Proceedings of the 14th International Conference on Interaction Design and Children - IDC '15. pp. 21–28. , New York, New York, USA (2015).
18. Billinghurst, M., Dünser, A.: Augmented reality in the classroom. Computer (Long. Beach. Calif). Vol.45. (2012) 56–63.
19. Vosniadou, S.: International handbook of research on conceptual change. , New York (2013).
20. Carey, S.: Science Education as Conceptual Change. J. Appl. Dev. Psychol. Vol.21. (2000) 13–19.
21. Mareschal, D.: The neuroscience of conceptual learning in science and mathematics. Curr. Opin. Behav. Sci. Vol.10. (2016) 114–118.

22. Foisy, B.: Les conceptions erronées des élèves en électricité ne disparaissent peut-être jamais de leur cerveau. (2012) 15–17.
23. Linn, M.C., Clark, D.B.: The knowledge integration perspective - connections across research and education. In: International handbook of research on conceptual change. pp. 520–538. , New York (2013).
24. Linn, M.C., Eylon, B.-S.: Science learning and instruction : taking advantage of technology to promote knowledge integration. (2011).
25. Schnotz, W., Kürschner, C.: External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. Instr. Sci. Vol.36. (2008) 175–190.
26. Gilbert, J.K.: Visualization: Theory and Practice in Science Education. Vis. Sci. Educ. Vol.3. (2008) 9–27.
27. Dori, Y.J., Belcher, J.: Learning Electromagnetism with Visualizations and Active Learning. In: Visualization in Science Education. pp. 187–216. , Dordrecht (2005).
28. Suwa, M., Tversky, B.: External Representations Contribute to the Dynamic Construction of Ideas. In: International Conference on Theory and Application. pp. 149–160 (2002).
29. Duval, R.: A cognitive analysis of problems of comprehension in a learning of mathematics, (2006).
30. Scheid, J., Müller, A., Hettmannsperger, R., Schnotz, W.: Scientific Experiments, Multiple Representations, and Their Coherence: A Task-Based Elaboration Strategy for Ray Optics. Palgrave Macmillan UK (2015) 239–252.
31. Seufert, T.: Supporting coherence formation in learning from multiple representations. Learn. Instr. Vol.13. (2003) 227–237.
32. Mayer, R.E.: The Cambridge Handbook of Multimedia Learning. (2014).
33. Ainsworth, S.: DeFT: A conceptual framework for considering learning with multiple representations. Learn. Instr. Vol.16. (2006) 183–198.
34. Laviole, J., Hachet, M.: PapARt: Interactive 3D graphics and multi-touch augmented paper for artistic creation. In: 2012 IEEE Symposium on 3D User Interfaces (3DUI). pp. 3–6 (2012).
35. Raduta, C.: General students' misconceptions related to Electricity and Magnetism. (2005).
36. Saarelainen, M., Laaksonen, A., Hirvonen, P.E.: Students' initial knowledge of electric and magnetic fields—more profound explanations and reasoning models for undesired conceptions. Eur. J. Phys. Vol.28. IOP Publishing (2007) 51–60.
37. Maloney, D.P., O'Kuma, T.L., Hieggelke, C.J., Van Heuvelen, A.: Surveying students' conceptual knowledge of electricity and magnetism. Am. J. Phys. Vol.69. (2001) S12.
38. de Jong, T., Linn, M.C., Zacharia, Z.C.: Physical and virtual laboratories in science and engineering education. Science. Vol.340. American Association for the Advancement of Science (2013) 305–8.

Session
Modélisation de l'apprenant et feedbacks

Étayer dynamiquement une activité de conception expérimentale en biologie : premiers résultats d'expérimentations

Catherine Bonnat, Anne Lejeune

Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, F-38000 Grenoble France
catherine.bonnat@univ-grenoble-alpes.fr, anne.lejeune@univ-grenoble-alpes.fr

Résumé. L'activité de conception expérimentale est au cœur des programmes de sciences en France et pose des difficultés méthodologiques et conceptuelles aux élèves. La mise en place d'étayages adaptés peut aider à pallier ces difficultés, mais reste à analyser son efficience. Pour tenter de répondre à cette problématique, nous avons conçu un EIAH doté d'étayages fixes et adaptatifs. L'apprenant y interagit avec SimDiabolo, simulation sur le thème des métabolismes cellulaires, dotée d'étayages relatifs aux concepts non objet d'apprentissage et, reçoit des rétroactions dynamiques en cas d'erreur sur un concept objet d'apprentissage. Ces rétroactions et leurs déclencheurs sont définis dans des scénarios exploités par le module élève de FORMID au fil de la progression des apprenants. Nous présentons les premiers résultats de l'analyse des traces recueillies au cours d'expérimentations en classe de seconde générale en sciences de la vie et de la terre (France).

Mots-clés. Conception expérimentale, simulation, étayage adaptatif, scénarios.

Abstract. Experimental design is one of the main activities in science programs and poses methodological and conceptual challenges for learners. Providing learners with adapted scaffolds can help to overcome these difficulties, but remains to analyze its efficiency. According to our research goal, we designed an EIAH with fixed and adaptive scaffolds. The learner interacts with SimDiabolo, simulation on the topic of cellular metabolism, which provides scaffolds related to non learning objectives, and receives dynamic feedback in case of error related to one of the learning concept. Feedbacks and their triggers are defined in the FORMID scenarios used by the FORMID runtime engine throughout the learners' progress. We present the first results of traces analysis collected during experiments in high school biology classes (France).

Keywords. Experimental design, simulation, adaptive scaffolds, scenarios.

1 Contexte scientifique et objectifs de la recherche

Cette communication présente les premiers résultats d'une recherche sur l'étayage par un EIAH de l'activité de conception expérimentale en Sciences de la Vie et de la Terre (SVT). L'activité cible de notre recherche est l'une des phases d'une démarche d'investigation dans laquelle les élèves sont amenés à concevoir et réaliser des expériences. Les enjeux de cette démarche sont doubles puisqu'ils visent, dans le

même temps, l'apprentissage d'une méthode scientifique et l'assimilation des concepts relatifs à la situation étudiée [1, 2]. Plus précisément, la conception d'un protocole expérimental permet aux élèves de faire le lien entre les activités pratiques et les concepts en jeu [3]. Cependant cette tâche reste complexe et nécessite d'être étayée. En effet, ces derniers perdent de vue le problème posé [4] suite aux difficultés méthodologiques et épistémologiques qu'ils rencontrent dans l'écriture du protocole [5]. Concevoir un étayage à partir de l'identification *a priori* de ces difficultés, peut permettre de les pallier [6]. L'étayage peut être « fixe » (i.e. systématiquement proposé à tous les élèves), ou « adaptatif » (i.e. délivré dynamiquement à l'élève sur constat d'une des difficultés préalablement recensées) [7].

Dans un travail de thèse [8], nous avions conçu et testé au lycée une situation de conception expérimentale sur le thème de la fermentation alcoolique en SVT, alors étayée par l'EIAH LabNbook. Les résultats de ce travail ont montré l'efficience des étayages fixes (pré structuration du protocole et documents ressources), mais que demeuraient des erreurs conceptuelles récurrentes dans les protocoles conçus [8, 9]. Selon certains auteurs, l'étayage adaptatif [10, 11] permet à l'apprenant d'activer ses acquis antérieurs et de mieux s'investir dans la tâche proposée [5]. Cela justifie la combinaison de plusieurs formes d'étayage et notamment de prendre en compte les erreurs tenaces en tentant de les déstabiliser par des étayages adaptatifs complémentaires.

La mise en place d'étayages adaptatifs dans un EIAH présuppose que ce dernier dispose de mécanismes de diagnostic, en temps réel, de l'activité de l'apprenant. D'autre part, analyser l'efficience des étayages construits nécessite de pouvoir retracer a posteriori la progression de chaque apprenant au fil de l'activité. Nous avons choisi pour cela de bâtir un EIAH exploitant des outils informatiques et des modèles didactiques existants. À partir des résultats des travaux de thèse, nous avons conçu SimDiabolo, une simulation web interactive et autonome sur le métabolisme fermentaire. Nous l'avons dotée d'étayages fixes et lui avons délégué la prise en charge d'étayages adaptatifs relatifs à des concepts non objet d'apprentissage (NOA). La suite FORMID [12] permet (1) de scénariser l'observation de l'activité d'un élève interagissant avec une simulation et est dotée (2) de mécanismes de délivrance dynamique de rétroactions et (3) de moyens d'observation de l'activité des élèves, en temps réel ou *a posteriori*. Pour profiter de ces fonctionnalités, nous avons rendu possible la communication entre SimDiabolo et FORMID (Figure 1). L'ayant fait, nous pouvions déléguer à différents scénarios FORMID, la prise en charge des étayages adaptatifs relatifs aux concepts objet d'apprentissage (OA) afin d'être en mesure de répondre à notre question de recherche : *Quel est l'impact des étayages proposés, sur la pertinence des protocoles produits ?*

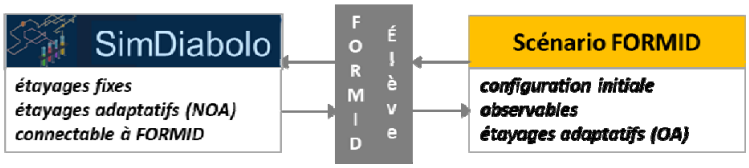


Fig. 1. Communication SimDiablo - FORMID Elève - Scénario FORMID

Dans cette communication, nous présentons l'activité proposée par SimDiabolo et sa scénarisation. Puis, après avoir décrit les expérimentations menées sur la fin 2018, nous détaillons la méthodologie d'analyse des traces recueillies et ses premiers résultats, en nous focalisant sur les erreurs relatives au concept d'anaérobie (absence de dioxygène). Ce concept, objet d'apprentissage, est à l'origine de difficultés récurrentes chez les élèves qui se traduisent par la non prise en compte des gaz présents dans le milieu et de leur dissolution [13].

2 Étayage de la situation et scénarisation de son observation

SimDiabolo propose aux élèves de lycée (15-17 ans) une activité de conception expérimentale destinée à dégager les caractéristiques du métabolisme de la fermentation alcoolique des levures (*consommation de glucose, production d'alcool et de dioxyde de carbone - réaction dépendante de la teneur du milieu en dioxygène et de sa température*). Avec SimDiabolo, l'élève rassemble le matériel nécessaire, conçoit un protocole et simule l'expérience en résultant. Simdiabolo est dotée d'étayages fixes (pré-structuration du protocole en actions paramétrées et accès à des documents ressources) et également d'étayages adaptatifs portant sur des concepts non objet d'apprentissage (ex : dispositifs de mesure).

Lors de l'analyse didactique préalable [8, 9] nous avions identifié et modélisé les erreurs fréquentes relatives aux concepts objet d'apprentissage. Les caractéristiques de ces erreurs ont été codées par des observables dans des scénarios FORMID, un observable étant associé à la rétroaction dynamiquement délivrée dans le cas de sa détection. Dans le but d'analyser l'impact de ces étayages, chaque scénario propose les mêmes observables, mais un « niveau » de rétroaction différent. De fait, nous disposons de différentes versions de l'EIAH.

Concernant le concept d'anaérobie, l'analyse didactique préalable a traduit les erreurs liées à ce concept par deux techniques potentiellement appliquées par l'élève : *contenant non bouché (solution typée S5) / contenant bouché mais partiellement rempli (solution typée S3)*. Ces techniques erronées ont été codées par des observables nommés respectivement *SoltypeS5* et *SoltypeS3*. Les rétroactions, délivrées à l'élève sur détection de l'un de ces observables sont fonction du scénario actif (Tableau 1).

Tableau 1. Scénarisation des étayages adaptatifs (cas d'un concept OA : l'anaérobie)

Contexte	Observable	Scénario	Rétroaction (si Observé)
Définition d'une solution lors de la conception du protocole : *Action* ***Prélever*** + *une ou plusieurs actions* ***Ajouter***	volume total solution créée < volume du contenant	**Scénario 1**	Aucun message
		Scénario 2 *Donner la technologie*	La fermentation alcoolique se réalise en condition d'anaérobie
		Scénario 3 *Questionner la technologie*	Quelles sont les conditions de réalisation de la fermentation alcoolique ? Consultez les documents ressources
		Scénario 4 *Donner la technique*	Supprimez les apports en dioxygène

3 Expérimentations et méthodologie d'analyse

Courant novembre 2018, nous avons testé l'EIAH dans des classes de seconde générale de trois lycées polyvalents de l'agglomération grenobloise. Les tests se sont déroulés sous le contrôle d'enseignants de SVT expérimentés, un chercheur étant présent à chaque séance. Chaque version a été expérimentée dans au moins 2 lycées différents, les séances durant 1h environ et s'intégrant dans la progression pédagogique des enseignants. 90 élèves ont ainsi participé à ces expérimentations.

À partir des traces d'observation recueillies par FORMID, nous avons filtré les observables détectés et les avons combinés en indicateurs quantitatifs et qualitatifs afin de répondre à notre question de recherche. La déclinaison de cette question relativement aux erreurs liées au concept d'anaérobie, conduisait à évaluer l'impact des rétroactions fournies à l'élève sur sa production d'un protocole expert après qu'il ait créé des solutions non bouchées ou ne remplissant pas leur contenant (Tableau 2).

Tableau 2. Sous-questions de recherches et indicateurs (cas d'un concept OA : l'anaérobie)

Sous-questions de recherche	Indicateurs résultant d'observables
A. Pertinence des solutions produites	
Est-ce que l'activité proposée aide les élèves à concevoir une solution experte (S1) ?	Nombre d'élèves ayant une solution S1. Temps et soumissions pour y parvenir.
B. Prise en charge des difficultés liées au concept d'anaérobie	
Retrouve-t-on des erreurs relatives au concept d'anaérobie ? Comment évoluent les solutions en anaérobie selon les scénarios ? Influence des documents ressource ?	Nombre de solutions en aérobie. Évolution de ces solutions. Consultation du document 1.
C. Utilisation des documents ressource et Anticipation du matériel	
Les élèves consultent-ils ces documents en amont de la conception de l'expérience ? Cela favorise-t-il le choix d'un matériel pertinent ?	Consultation des documents ressources dans la page matériel. Absence de bouchon dans la liste de matériel.

4 Premiers résultats

Pertinence des solutions produites. Nous constatons que le nombre d'élèves ayant produit au moins une solution « experte » augmente du scénario 1 au scénario 4 (Fig.2). Ce premier résultat tend à confirmer l'impact du niveau d'étayage sur les productions des élèves.

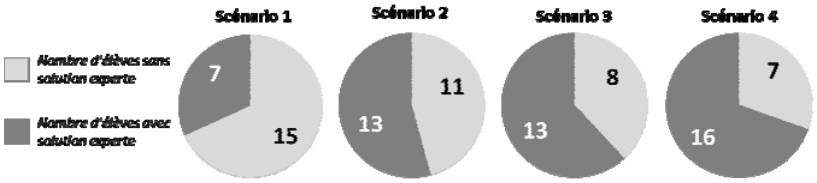


Fig. 2. Nombre d'élèves ayant produit au moins une solution « experte » selon le scénario

Nous avons également relevé le nombre important de soumissions des élèves avant qu'ils obtiennent une solution experte (moy. = 7), ce qui pourrait résulter de difficultés conceptuelles préalablement identifiées, telles que celle liée au concept d'anaérobie.

Prise en charge des difficultés liées au concept d'anaérobie. 61 élèves ont fait une erreur relative à ce concept lors de leurs premières soumissions, ce qui confirme la ténacité de cette erreur. Nous avons analysé pour ces élèves l'évolution d'une des solutions S3 ou S5 vers une solution experte en distinguant 3 cas : *pas d'évolution, évolution suite à consultation document proposé, évolution sans consultation* (Fig. 3).

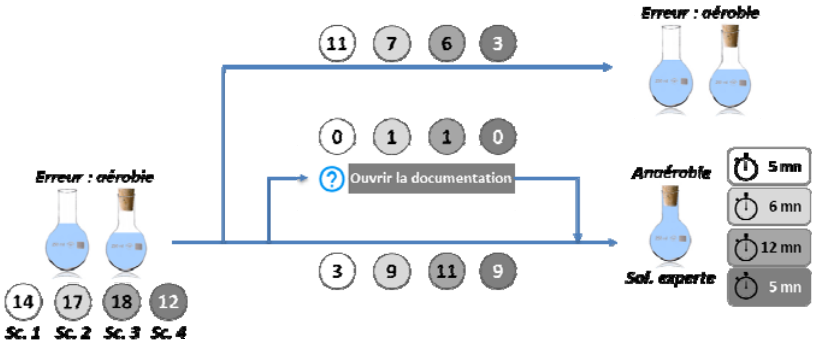


Fig. 3. Évolution, selon le scénario, des solutions en aérobie produites par les élèves

Il apparait que les rétroactions des scénarios 3 et 4 les aident davantage à modifier leurs solutions. Néanmoins, bien que la rétroaction du scénario 3 les incite à consulter un document ressource, ils ne le font pas (1 élève / 18).

Utilisation des documents ressource et anticipation du matériel. Parmi les élèves ayant fait une erreur sur le concept d'anaérobie, 51 ont consulté les documents ressources lors de la phase de collecte du matériel. Cependant, l'oubli fréquent de bouchon lors de cette phase (45 élèves), nous amène à penser que ces documents n'incitent pas assez les élèves à réfléchir sur la technique de mise en anaérobie. Le fait que l'activité nécessite de faire le lien entre la pratique et les concepts en jeu [3] peut aussi en être la cause. Parmi ces mêmes élèves, seulement 2 ont consulté la documentation lors de la conception du protocole, malgré les rétroactions qui les y incitaient. Néanmoins, la moitié d'entre eux a fini par obtenir une solution experte. Nous supposons qu'ils le doivent à la simulation des expériences conçues (rétroaction du milieu), qui serait un étayage complémentaire suffisant pour cela. Ce résultat renforce l'idée de combinaison d'étayages suggérée par certains auteurs [10].

5 Conclusion et perspectives

Ces premiers résultats tendent à prouver l'impact des niveaux de rétroactions dynamiques sur la rectification d'erreurs relatives au concept d'anaérobie, avec pour le scénario 4, davantage de solutions qui tendent vers celle qualifiée « d'experte ».

Les indicateurs - construits à partir des observables définis en amont - nous ont permis de restituer les actions des élèves et notamment celles suivant les rétroactions reçues en cas d'erreur.

Ces résultats nous encouragent à envisager des scénarios ajustant le degré d'étayage en fonction des difficultés observées. De plus, l'analyse détaillée des protocoles nous permettra d'enrichir la modélisation didactique de la situation. En effet, nous avons constaté un nombre significatif de solutions produites par les élèves qui ne correspondent à aucune de celles modélisés *a priori*. Enfin, bien que nous ne cherchions pas à évaluer les apprentissages, des expérimentations ultérieures conçues à cet effet pourraient être envisagées.

Références

1. Calmette, B., Boilevin, J.-M. (2014). Le modèle investigation-structuration et l'actualité des tensions autour des constructivismes. *Recherche en Didactique des Sciences et Technologies, 9*, 103-128.
2. Etkina, E., Karelina, A., Ruibal-Villasenor, M. (2010). Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories. *Journal of the Learning Sciences*, *19*, 54-98.
3. Coquidé, M. (2000). *Le rapport expérimental au vivant.* Mémoire d'habilitation à diriger des recherches, Université Paris-Sud, France.
4. Hmelo-Silver, C., Duncan, R. G., Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99-107.
5. Marzin-Janvier, P. (2013). *Comment donner du sens aux activités expérimentales ?* Mémoire d'habilitation à diriger des recherches. Université Joseph-Fourier - Grenoble I, France.
6. Quintana, C., Reiser, B., J., Davis, E., A., Krajcik, J., Fretz, E., & Duncan, R., G. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13*(3), 337-386.
7. Azevedo, R., Cromley, J. G., Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia ? *Contemporary Educational Psychology, 29,* 344-370.
8. Bonnat, C. (2017). *Etayage de l'activité expérimentale par un EIAH pour apprendre la notion de métabolisme cellulaire en terminale scientifique*. Thèse de doctorat, Université Grenoble-Alpes, France.
9. Bonnat, C., Marzin, P., Girault, I. (2018). Analyse des conceptions d'élèves sur le vivant, dans une situation de conception expérimentale avec un environnement informatique. *Recherche en didactique des sciences et technologies, Varia*, (sous presse).
10. McManus, T. (2000). Individualizing instruction in a web-based hypermedia learning environment: Nonlinearity, advanced organizers, and self-regulated learners. *Journal of Interactive Learning Environments, 11(3)*, 219–251.
11. Puntambekar, S., Kolodner, J. L. (2005). Toward implementing distributed scaffolding: helping students learn science from design. *Journal of Research in Science Teaching, 42*(2), 185-217.
12. Lejeune, A., Guéraud, V. (2012). Embedding observation means into the learning scenario : Authoring approach and environment for simulations-based learning. Internatinal *Conference on Advanced Learning Technologies (ICALT). Rome: IEEE, 2012*, 273-275.
13. Laugier, A., Dumon, A. (2004). L'équation de réaction: un nœud d'obstacles difficilement franchissable. *Chemistry Education Research and Practice, (5)*1, 51-68.

Français langue de scolarisation : Réflexions sur les référentiels de compétences et l'adaptive learning

Thierry Geoffre[1], Nathalie Colombier[2]

[1] HEP Fribourg, UR DidaLang, 36 rue de Morat, CH-1700 Fribourg
geoffret@edufr.ch
[2] Educlever, 59 Rue Benoît Malon, F-94250 Gentilly
nathalie.colombier@educlever.com

1 Introduction

Le développement d'un EIAH tel que la plateforme *Smart Enseigno*[1] implique celui de l'adaptive learning. Ce projet est à l'origine d'une réflexion de fond sur les référentiels de compétences parce qu'ils forment le noyau de la plateforme et de l'adaptive learning à construire. La plateforme se veut transversale au primaire et secondaire 1 (cycles 2, 3 et 4[2]), ce qui impose de construire un référentiel transversal à 9 années d'enseignement-apprentissage (désormais E/A) des domaines disciplinaires ciblés.

2 Contexte et objectifs

Les développements actuels trouvent leur origine en 2012 avec nos premiers travaux tournés vers l'individualisation et la personnalisation des ressources numériques : des QCM enrichis de degrés de choix de certitude (premier retour métacognitif sur l'activité) ; l'articulation de ressources de manière algorithmique pour créer des parcours avec activités entre elles ; le Libre Service Intelligent (LSI) qui redéfinissait l'ergonomie et la navigation en moments pédagogiques pour les apprenants.

C'est l'initiation du projet pilote de *Cartographie des* Savoirs, dans le cadre d'un consortium de recherche et développement, entre 2012 et 2014, qui nous a ensuite permis de définir une ambition bien plus vaste : la recommandation de ressources aussi précisément adaptées que possible au profil de chaque apprenant, en prenant en compte son historique de différentes manières ; l'amorce du passage à l'individualisation par l'adaptive learning. La mise en cartographie du corpus des programmes de l'Éducation nationale est alors devenue (et est toujours) une problématique centrale.

Pour ce faire, nous avons défini une approche didactique par praxéologies sous forme ontologique. La praxéologie aide alors à décrire ce qu'il « faut apprendre » : puisque les programmes scolaires sont rédigés sous forme de listes d'objectifs

[1] La plateforme *Smart Enseigno* est développée par Educlever.
[2] Pour des raisons de commodités, nous utiliserons les dénominations françaises dans cet article mais notre travail concerne à la fois les programmes français et suisse romand.

d'apprentissage, on peut les retranscrire sous forme de structures praxéologiques, c'est-à-dire de savoirs et de savoir-faire. Mais pour être utilisable par un logiciel, la praxéologie doit être modélisée par une représentation informatique aussi performante que possible. De très nombreux travaux sont consacrés à ces questions de l'organisation des bases de connaissances et des référentiels de compétences mais nous n'en développerons pas l'état de l'art dans cet article court.

Nous avons rencontré des difficultés importantes avec notre modèle utilisant les ontologies issues de la recherche fondamentale. Nous avons cherché à expliquer cet échec et notre compréhension est désormais que les problèmes survenus avec les ontologies n'étaient pas imputables au *recours* à une ontologie mais à la *modélisation* de cette ontologie. En effet, c'est toute la praxéologie d'origine qui était modélisée en ontologie, soit tout le domaine — par exemple, toute la description du domaine des mathématiques (techniques et technologies incluses). Or, nous n'avons pas besoin de la description exhaustive mais uniquement des usages nécessaires.

Pour le dire plus simplement : nous n'avons pas besoin, en tant que concepteurs d'une plateforme de contenus et de services pour l'E/A, de modéliser l'ensemble d'un domaine de connaissance (les mathématiques, le français), mais seulement les éléments structurants du point de vue de cet E/A. Nous avons alors remis le problème à plat en 2017 en collaboration avec un laboratoire d'informatique. Un nouveau point de départ a ainsi été constitué par la combinaison du socle commun des savoirs de l'Éducation Nationale et de cas d'usage (« use cases ») que nous avions définis.

À notre connaissance, la modélisation à laquelle nous avons abouti n'a jamais été publiée, et serait donc nouvelle. C'est l'architecture de ces référentiels qui fait l'objet de cet article focalisé sur le domaine de la langue (français langue de scolarisation).

3 Développement des référentiels de français

Notre cadre théorique de travail est la Théorie Anthropologique du Didactique développée par Yves Chevallard (1985 / 1991). On considère alors qu'une *tâche* t assignée à l'élève relève d'un certain *type de tâche* T, est réalisée au moyen d'une *technique* τ, justifiée par une *technologie* θ qui permet en même temps de la penser, de la produire, et qui à son tour est justifiable par une *théorie* Θ (Chaachoua, 2011). Dans notre cas, cherchant à transposer la TAD à la langue de scolarisation, la théorie est la grammaire de référence du français, au sens large, que l'on peut décomposer, dans un mouvement de transposition didactique aux différents cycles, en une « grammaire du mot » (phonologie, morphologie, sémantique), une « grammaire de la phrase » (syntaxe, morphologie, sémantique) et une « grammaire du texte » (syntaxe, genres textuels, sémantique). Cette articulation permet la construction de trois référentiels liés qui respectent la progression curriculaire et permettent de décrire l'ensemble des objets d'E/A du français des 9 années de cursus en élémentaire et secondaire 1. Ils doivent décrire dans le modèle chaque connaissance à un niveau de détail suffisamment fin pour permettre au système d'assister l'enseignant dans l'E/A, l'évaluation, la détection de prérequis manquants ou de techniques erronées possibles et connues.

3.1 Les Organisations Praxéologiques Linguistiques

Chaque référentiel regroupe donc les objets d'E/A du sous-domaine de la langue (mot, phrase, texte), que ce soit à l'écrit ou à l'oral, en réception-compréhension ou en production. Pour chaque objet d'E/A (exprimé par un groupe nominal), le référentiel prend en charge ses différentes caractéristiques linguistiques (selon les cas sémantiques, morphologiques, phonologiques, syntaxiques) afin de définir le plus exhaustivement possible les différents aspects théoriques de son E/A. Par exemple, pour la phrase, et sans être exhaustifs, on déclinera les aspects *morphosémantiques* (ponctuation, type, forme, voix), *morphosyntaxiques* (phrase simple ou complexe, constituants, fonctions des constituants, relations entre constituants), *sémantiques*.

Ces trois arborescences regroupent donc des objets d'E/A discrets, toujours verbalisés par des groupes nominaux. Leur E/A se réalise à travers des types de tâches qui seront déclinés en tâches proposées aux élèves. Les types de tâches constituent le niveau de référentiel suivant et sont écrits selon la structure [verbe d'action + complément], où le complément est un objet d'E/A auquel est appliquée une action.

Les verbes utilisés dans la définition des types de tâches ont ensuite fait l'objet d'une réflexion approfondie sur la base des huit verbes d'opération mentale initialement proposés par Fabienne Vernet[3] (travail non publié, IUFM de Grenoble, 2008-2009). Fabienne Vernet avait analysé plusieurs moyens d'enseignement du fonctionnement de la langue et montré que l'ensemble des consignes utilisées revenait à mobiliser huit opérations mentales différentes. Nous avons alors considéré que le travail inverse (c'est-à-dire appliquer ces verbes à différentes entrées notionnelles) permettrait de décliner tous les types de tâches et de formaliser ce que l'élève a à faire lorsqu'il s'agit de français à l'école. Nous avons donc cherché à transposer ces verbes de référence à tous les domaines de l'E/A du français, donc à nos arborescences notionnelles (sans l'opération de comparaison qui est le plus souvent contenue dans les autres –Vernet n'en avait d'ailleurs relevé que peu d'occurrences dans son travail initial – mais en ajoutant l'opération de production).

On voit ici que nous avons changé d'approche par rapport aux premières recherches en lien avec la TAD dans le cadre du projet *Cartographie des Savoirs*. Plutôt que de partir d'une analyse des moyens d'enseignement disponibles dans une institution et chercher la liste exhaustive des types de tâches qui y sont proposés pour aboutir à une Organisation Praxéologique (Mathématique en l'occurrence ; Chaachoua, 2011), nous partons désormais des trois arborescences d'objets d'E/A et nous regardons s'il est possible d'appliquer les huit verbes d'opérations mentales (ou quels verbes peuvent être appliqués parmi les huit) pour produire un ensemble de types de tâches réalistes et pouvant être déclinés en tâches pour l'élève. Ce travail est donc plus large qu'à partir de la seule analyse des moyens d'enseignement (souvent limités et incomplets) et présente l'intérêt d'être directement lié à la théorie Θ de référence, avec des types de tâches décorrélés des niveaux scolaires, donc indépendants des programmes mais

[3] Fabienne Vernet a ensuite animé différentes conférences pédagogiques en Isère sur le thème de l'E/A de la grammaire en cycle 3. Ses propositions et outils peuvent toujours être consultés en ligne sur 2 sites de circonscriptions : http://www.ac-grenoble.fr/ien.bv/spip.php?article1497 ou http://www.ac-grenoble.fr/ien.bourgoin2/spip.php?article624.

facilement transposables à différentes institutions francophones[4]. Nous parvenons ainsi à un peu plus de 750 types de tâches qui représentent le domaine.

Nous posons que ces trois référentiels forment des Organisations Praxéologiques Linguistiques de référence (désormais OPL) qui concernent tous les degrés du primaire et du secondaire 1. Leur extension au secondaire 2 est un objectif, dans le cadre de la réforme des programmes français du lycée.

Pour vérifier le bien-fondé de ces OPL, nous avons cherché à y rattacher tous les types de tâches des praxéologies initiales de la *Cartographie des Savoirs* ; il apparait clairement que ces nouvelles organisations praxéologiques « contiennent » celle de la Cartographie mais vont aussi plus loin et avec une structure plus régulière.

Pour résumer, nous précisons ici ce que nous pensons être les premiers apports et forces de la construction des OPL :

- toutes les notions prises en charge actuellement, sont déclinées en types de tâches de manière homogène, en utilisant huit verbes d'opérations mentales ;
- tous les aspects de l'enseignement-apprentissage de la langue (à part la production orale, non modélisée) sont traités ;
- une OPL n'est pas limitée à une seule institution de référence mais engage une réflexion didactique plus large et transversale, liée à la théorie de référence.

3.2 La transposition des OPL à des programmes scolaires

La transposition des OPL à une institution, c'est-à-dire à des programmes spécifiques, signifie que chaque type de tâches va être associé au(x) degré(s) scolaire(s) où il est impliqué dans un objectif d'E/A. On voit alors apparaitre de nombreux types de tâches qui sont associés à plusieurs degrés scolaires parce qu'ils sont appliqués à des corpus différents au fil des degrés scolaires (par exemple, l'identification du verbe conjugué concerne progressivement la phrase simple puis complexe, des temps différents, …). Nous avons donc pensé des progressions linguistiques et didactiques qui permettent à leur tour de construire des corpus[5] indexés sur lesquels sont appliqués les types de tâches. Il y a donc une progression et une continuité des objets d'E/A en fonction de la transversalité d'un type de tâches à plusieurs degrés scolaires et de l'évolution des corpus d'application. Plus globalement, ce travail revient donc à une mise en cartographie des programmes d'une institution – notre objectif initial – modélisant la continuité et la progression des apprentissages, d'un degré à l'autre, d'un cycle à l'autre.

Notre travail de transposition de la TAD au domaine du français langue de scolarisation permet ainsi de développer une réflexion innovante concernant les fondements didactiques de l'EIAH. Cette réflexion se prolonge en trois points : 1) une nouvelle définition des compétences ; 2) une nouvelle approche de l'évaluation des compétences ; 3) l'élaboration de parcours d'enseignement-apprentissage inédits. Nous reprenons les deux premiers points dans la suite de cet article.

---

[4] On note ainsi un décalage d'une année scolaire entre programmes français et PER suisse pour l'introduction de l'attribut du sujet, de 2 années pour l'étude systématique du passé simple, mais les types de tâches restent les mêmes car liés à la langue et non à l'institution de référence.

[5] Ces corpus font l'objet d'un travail parallèle et sont progressivement publiés sur le site Scolagram.u-cergy.fr (Geoffre & Sautot, travail en cours de publication).

3.3 Le passage au référentiel des compétences

Nous définissons une compétence comme composite et constituée des types de tâches qu'elle mobilise pour être mise en œuvre.

Cette définition est un prolongement du travail de construction des OPL et fait écho à Michèle Artaud (2016) lorsqu'elle considère que « déclarer que ''X a atteint la compétence C'' signifie alors que X sait accomplir certains types de tâches selon les techniques en vigueur dans l'institution dont il est le sujet, ces techniques étant justifiées ou produites par un savoir déterminé ».

Inversement, on peut énoncer que la difficulté à atteindre une compétence visée relève de l'échec à un ou plusieurs types de tâches qui en sont constitutifs.

A partir des OPL et de cette définition, il est possible de construire un référentiel des compétences. Puisqu'issu des OPL, ce référentiel est initialement indépendant des programmes et est essentiellement constitué de compétences « intégrales » (assemblage de tous les types de tâches impliqués dans la compétence de plus haut niveau). L'application à une institution donnée permet ensuite de décliner une gradation de compétences enchâssées par degrés scolaires, ne recouvrant que les types de tâches travaillés jusqu'au degré concerné comme des *poupées russes*. Chaque compétence devient un objectif d'E/A dont on pourra définir la construction progressive à l'échelle de plusieurs degrés scolaires. Tous les types de tâches constitutifs d'une compétence ne sont en effet pas forcément liés à un même degré scolaire. Par ex., si l'on considère la compétence « Orthographier correctement une phrase », il est bien évident que sa portée sera très partielle en 2e primaire mais plus large en 5e année.

Dans une institution donnée, le référentiel des compétences sera donc transversal aux degrés scolaires et rendra compte de la progressivité spiralaire des apprentissages.

Les OPL et le référentiel des compétences « intégrales » permettent non seulement de penser les objets d'E/A avec la souplesse de s'adapter à des cadres institutionnels multiples, mais également de lier le développement actuel de la solution numérique éducative au travail initial de production de la *Cartographie des Savoirs*. En effet, les relations entre types de tâches sont également modélisées et permettent d'entrevoir deux modalités d'évaluation et de remédiation subséquentes :

- un type de tâche qui apparait initialement au degré scolaire N devient automatiquement un prérequis de ce même type de tâches au degré N+1 ou N+2 (par exemple, l'accord sujet-verbe dans des cas simples vu en 3e primaire est un prérequis des accords sujet-verbe complexes vus en 4e ou 5e année primaire). Une première modalité de remédiation peut donc être pensée au niveau d'un même type de tâches en modifiant (simplifiant ou adaptant) le corpus de travail ;
- les relations de prérequis entre types de tâches restent identifiables (par exemple, l'identification du sujet de la phrase est un type de tâches en soi et est aussi un prérequis à l'accord du verbe conjugué avec le sujet). Une deuxième modalité de remédiation repose donc sur le travail spécifique d'un type de tâches T_i parce qu'il est un prérequis du type de tâches T_j lui-même objet de l'apprentissage en cours (en d'autres termes, T_i entre dans la technique qui permet de réaliser T_j et l'erreur est alors due à une technique erronée parce qu'un prérequis est déficitaire).

Les conséquences sont importantes puisque, du point de vue de l'évaluation, nous pourrons, à partir de la *Cartographie des Savoirs*, des techniques répertoriées et des différents niveaux de référentiels, indexer les erreurs et déclencher des remédiations

individuelles selon la fréquence, la récurrence d'un type d'erreur. Le suivi du travail de l'élève en temps réel et de son évaluation formera alors le moteur de l'adaptive learning par l'exploitation des *praxéologies personnelles* (Chaachoua et Bittar, 2016 ; Bonnat, 2018) dont la détermination fera l'objet d'un article ultérieur.

4 Conclusion

Notre travail porte sur la transposition de la Théorie Anthropologique du Didactique au domaine de la langue (dans notre cas, le français langue 1 et langue de scolarisation) pour aboutir à des Organisations Praxéologiques Linguistiques de référence (OPL). Nous nous sommes détachés de l'approche praxéologique habituelle, qui revient à construire l'organisation praxéologique à partir de l'étude des moyens d'enseignement et programmes en vigueur au sein d'une institution, pour chercher des organisations praxéologiques plus générales et indépendantes des programmes.

Nous montrons comment cette réflexion permet ensuite de redéfinir la notion de compétence en lien avec l'approche praxéologique et de construire un référentiel des compétences ; OPL et référentiel des compétences pouvant alors être transposés à une institution donnée.

Ce cadre didactique ouvre la voie au suivi du travail de l'élève en temps réel et à son évaluation, donc au développement de l'adaptive learning, dans une perspective d'aide à l'enseignant et aux élèves, au sein de la classe ordinaire. La prochaine étape concerne le développement de plusieurs parcours de type « preuve de concept » (proof of concept) qui vont être testés dans des classes à partir de l'automne 2019.

Références

1. Artaud, M. : Fonctionnaliser la notion de compétence en la pensant avec le concept de praxéologie. Études de cas dans l'enseignement secondaire des mathématiques et la formation des professeurs. 3e congrès de l'ARCD, Marseille, 2013 (2016). https://www.researchgate.net/publication/303805776
2. Bonnat, C. : Modélisation de praxéologies personnelles a priori dans une situation de conception expérimentale en biologie. Préactes du 6ème Congrès International de la Théorie Anthropologique du Didactique, Autrans, France, Jan 2018 (2018). hal-01708033v2
3. Chaachoua, H. : La praxéologie comme modèle didactique pour la problématique EIAH. Étude de cas : la modélisation des connaissances des élèves. Habilitation à diriger des recherches. Université de Grenoble (2011). https://tel.archives-ouvertes.fr/tel-00922383
4. Chaachoua, H. H. & Bittar, M. : La théorie anthropologique du didactique : paradigme, avancées et perspectives. 1° simpósio latino-americano de didática de matemática. Bonito – Mato Grosso do Sul – Brasil. 1-5 novembre 2016 (2016). http://ladima.tuseon.com.br/uploads/file_manager/source/d7322ed717dedf1eb4e6e52a37ea7bcd/CONFERÊNCIA%202%20-%20FRANCÊS.pdf.
5. Chevallard, Y. : *La transposition didactique : du savoir savant au savoir enseigné* (2e ed.). Grenoble, France : Éd. La Pensée sauvage (1985/1991).

Modèle d'Exercices et Parcours d'Apprentissage Prenant en Compte le Raisonnement de l'Élève en Mathématiques au Collège

Elann Lesnes-Cuisiniez[1,2] et Brigitte Grugeon-Allys[1,3]

[1] Laboratoire de Didactique André Revuz – Université Paris Diderot, 8e étage, 8 place Aurélie Nemours, 75013 Paris, France
[2] elann.cuisiniez@etu.univ-paris-diderot.fr
[3] brigitte.grugeon-allys@u-pec.fr

Résumé En se plaçant d'un point de vue didactique, l'enjeu de cet article est de proposer, dans le cadre de la conception d'un E.I.A.H., un modèle d'exercices en géométrie à un niveau collège. Ce modèle d'exercices associé au profil de l'élève exprimé en termes de mode de raisonnement permettra, à un niveau local, la génération de feedbacks adaptés aux procédures de résolution des élèves, et à un niveau plus global, la création de parcours d'apprentissage adaptatifs qui prennent en compte leur raisonnement.

Mots clefs : didactique, modèle d'exercices, modèle de l'apprenant, parcours d'apprentissage, géométrie.

Abstract From a didactic point of view, the goal of this article is to offer, within the framework of the design of an I.T.S, a task model in geometry at a middle school level. This task model, combined with the student's profile expressed in terms of reasoning mode, will allow for, at a local level, the generation of feedbacks adapted to students' resolution procedures and, at a more global level, the creation of adaptive learning paths that take into account their reasoning.

Keywords : didactics, task model, learner model, learning paths, geometry.

1 Introduction

L'enjeu de cet article est de proposer un modèle d'exercices prenant en compte un modèle de l'apprenant afin d'implémenter un diagnostic et des parcours d'apprentissage adaptés aux modes de raisonnement des élèves en géométrie à la fin du collège.

Des travaux plus anciens affichant cet objectif de modélisation tels que les projets Pépite, Lingot et PepiMeP [6], en algèbre, visaient à développer un diagnostic global et multidimensionnel des connaissances des élèves. Ce diagnostic se fondait sur une étude cognitive, épistémologique et anthropologique de l'algèbre

élémentaire pour définir le rapport à l'algèbre des élèves à un niveau scolaire donné et leur proposer des exercices adaptés dans le cadre d'un parcours d'apprentissage. Depuis 2016, les diagnostics en 4e et 3e ont été développés sur un serveur WIMS [4] dans le cadre de l'ANR NéoPraéval [5].

Cet article se situe dans le contexte du projet MindMath qui vise l'élaboration d'une plateforme gamifiée et adaptative pour l'apprentissage de l'algèbre et de la géométrie au collège. Il s'agit donc de poursuivre le travail effectué dans le cadre des projets Pépite et PepiMep en implémentant et automatisant un diagnostic et les parcours d'apprentissage en algèbre élémentaire, mais également d'étendre le projet au domaine de la géométrie.

Nous nous placerons donc d'un point de vue didactique. Nous n'évoquerons ni l'aspect ludique de la plateforme, ni l'algorithme de recommandation d'exercices qui sera utilisé, ni la génération des feedbacks. En revanche, nous proposons un modèle d'exercices prenant en compte un modèle de l'apprenant, de ses réponses aux exercices et des parcours d'apprentissage, fondé sur une approche didactique. À ces modèles, s'adosse l'algorithme de machine learning qui calcule le parcours d'apprentissage le plus adapté en fonction des objectifs visés au niveau scolaire du collège, de l'apprenant, de ses réponses aux exercices et des parcours d'apprentissage des autres élèves.

Nous répondrons donc aux questions suivantes : en quoi l'appui sur les cadres théoriques de la didactique des mathématiques permet-il de fonder les modèles d'exercices et de l'apprenant en géométrie ? Quelles sont les caractéristiques des exercices sur lesquelles nous pouvons nous appuyer pour le diagnostic ? Comment utiliser le modèle de l'apprenant qui existe déjà en algèbre pour modéliser l'apprenant en géométrie et permettre la création de parcours d'apprentissage adaptatifs ?

2 Caractérisation des Modèles

2.1 Modèle d'Exercices

Une tâche s'exprime par un verbe, par exemple « développer l'expression littérale donnée » [3]. Plusieurs tâches peuvent relever du même type de tâches. Par exemple, les tâches « développer l'expression 2(x+3) » et « développer l'expression 7(9x+6) » relèvent toutes les deux du type de tâches « développer une expression de la forme a(bx+c) ».

Un exercice mathématique est une tâche qui relève d'un type de tâches [3], nous définissons donc des variables à deux niveaux. Le premier niveau est celui des variables de type de tâches qui spécifient les caractéristiques du type de tâches dont est issu l'exercice et l'apprentissage visé au regard du niveau scolaire en jeu et des programmes institutionnels, en particulier les propriétés mathématiques attendues dans la résolution des tâches d'un type donné. Pour l'exemple « développer une expression de la forme a(bx+c) », une variable de type de tâches serait la propriété de la distributivité simple. Le deuxième niveau

4. Serveur WIMS : `http://www.pmmp.jussieu.fr/wims/`
5. ANR NéoPraéval : `https://anr.fr/Projet-ANR-13-APPR-0002`

est celui des variables de tâche qui spécifient l'activité mathématique visée et sa complexité au sens de Robert [8]. Par exemple, la valeur du coefficient b.

Variables de Type de Tâches. En ce qui concerne la géométrie, nous nous appuyons notamment sur les travaux mathématiques d'Euclide [4] et les travaux didactiques de Celi [1] pour décomposer l'activité géométrique en cinq genres de tâches : construire, justifier, modéliser, calculer et représenter. Chacun de ces genres de tâches est exprimé sous la forme d'un générateur de types de tâches [2], c'est-à-dire GT = [un verbe d'action, un complément fixe ; un système de variables]. Les variables attachées à un générateur sont appelées variables de type de tâches, désignées par l'acronyme « Vtt ». Pour un générateur de type de tâches donné, on définit différents types de tâches en attribuant différentes valeurs aux variables de type de tâches.

Par exemple, pour le genre de tâches construire, nous définissons le générateur GT = [Construire, une figure de nature donnée, (Vtt1, Vtt2)] avec les variables de type de tâches suivantes :

— Vtt1 : nature de la figure à construire (cercle, triangle (isocèle, équilatéral, rectangle), quadrilatère (carré, rectangle, parallélogramme, trapèze, losange, cerf-volant), point, droite/segment) ;
— Vtt2 : registres de représentation en entrée (registre des figures géométriques, des schémas (à main levée), du langage naturel, des programmes de construction, des expressions algébriques).

Variables de Tâche Au delà des variables de type de tâches, nous caractérisons les tâches par des variables de tâche désignées par l'acronyme « Vt ». Ces variables sont liées à la complexité de la tâche, à son milieu (outils, environnement de géométrie dynamique, etc.) et à son énoncé. Ainsi, à partir d'un type de tâches, nous définissons différentes tâches de complexités différentes en attribuant des valeurs différentes aux variables de tâche :

— Vt1 : éléments donnés internes à la figure à construire (éléments géométriques donnés, positions de ces éléments, etc.) ;
— Vt2 : nombre de pas de construction ;
— Vt3 : éléments donnés externes à la figure à construire (position du centre de la transformation, présence d'un quadrilatère ou cercle dans lequel il faut construire la figure, etc.) ;
— Vt4 : forme de la réponse attendue de l'élève (figure robuste [6], programme de construction, etc.) ;
— Vt5 : outils géométriques du logiciel disponibles (cercle, médiatrice...).

2.2 Modèle de l'Apprenant

Les modes de raisonnement ont été développés dans le cadre du projet Pépite (et ses suites) [5]. Ils sont définis *a priori* grâce à une analyse épistémologique

6. figures géométriques robustes : « figures qui conservent la propriété visée au cours de déplacement de chacun de ses points » [9]

et didactique. Ils relèvent de types de tâches et caractérisent une activité mathématique en spécifiant des invariants dans les raisonnements des élèves.

La prise en compte des modes de raisonnement permet d'organiser une analyse non seulement au niveau microscopique (en étudiant les techniques mises en œuvre par les élèves) mais aussi au niveau macroscopique (au niveau des propriétés employées) sur un ensemble de tâches issues de types de tâches du genre de tâches considéré [6]. Les modes de raisonnement sont définis au regard de l'institution. L'appui sur les programmes scolaires permet d'inférer un mode de raisonnement idoine à un niveau scolaire donné.

En nous appuyant sur les paradigmes géométriques développés par Houdement et Kuzniak [7], nous définissons les modes de raisonnement en géométrie :

— G0 : associé au paradigme de la géométrie perceptive, usage de propriétés relevant de la perception, travail « au jugé », parfois approximatif ;
— G1 : associé au paradigme de la géométrie instrumentée, usage des instruments pour mesurer, construire, vérifier certaines propriétés, etc. ;
— G2 (attendu à la fin du collège) : associé au paradigme de la géométrie raisonnée, usage des propriétés géométriques (des figures et des outils), des théorèmes, des raisonnements hypothético-déductifs, etc.

Il existe aussi un mode de raisonnement G2-erroné lorsque les propriétés et raisonnements géométriques commencent à être utilisés mais des erreurs subsistent dans leur application (un pas de raisonnement déductif erroné par exemple).

3 Un Exemple de Diagnostic en Géométrie

3.1 Méthodologie

À partir de la définition des variables de type de tâches et de tâche, et l'explicitation *a priori* des différentes techniques en fonction du mode de raisonnement, nous repérons des éléments dans les productions de l'élève qui permettent d'inférer des aspects de son mode de raisonnement. Cette analyse menée de façon transversale sur tous les exercices du diagnostic nous permet de décrire le rapport à la géométrie de l'élève.

L'exemple présenté ci-après est celui d'un exercice de diagnostic proposé à dix élèves de troisième en février 2018. Le test a eu lieu dans une salle informatique équipée du logiciel GeoGebra (ce n'est pas le logiciel de géométrie dynamique qui est utilisé dans le projet MindMath). Les écrans ont été filmés et le son enregistré. Les chercheurs qui ont mené l'expérimentation n'ont pas aidé les élèves à l'exception des questions quant à l'utilisation de GeoGebra, les élèves n'étant pas tous familiers avec ce logiciel [7].

3.2 Analyse *a priori*

L'énoncé demande de « construire le carré ABCD et un demi-cercle de diamètre [AB] » à partir d'un segment [AB] donné. Cette tâche relève du générateur de

7. Même si l'expérimentation ne fait pas réellement partie du projet MindMath puisque la plateforme n'est pas encore développée, elle nous a permis de tester le modèle didactique présenté.

type de tâches GT_G = [Construire, une figure de nature donnée, (Vtt1, Vtt2)] avec les variables de type de tâches suivantes :

— Vtt1 : un carré et un demi-cercle ;
— Vtt2 : registre du langage naturel (énoncé qui donne la nature des figures).

Dans une phase de diagnostic, les variables de tâche sont choisies de manière à laisser apparaître les procédures des élèves liées à leur rapport à la géométrie :

— Vt1 : segment tracé dans une position non prototypique, déplaçable ;
— Vt2 : 9 pas de construction ;
— Vt3 : pas d'éléments externes à la figure à construire ;
— Vt4 : figure robuste ;
— Vt5 : tous les outils de GeoGebra sont disponibles sauf « tracer un demi-cercle à partir de deux points ».

Nous définissons *a priori* les techniques attendues pour cette tâche en fonction du mode de raisonnement :

— pour G0 : figure construite incorrecte, points placés au jugé, la figure n'est pas un carré, les instructions ne sont pas forcément respectées.
— pour G1 : figure correcte mais non robuste aux déplacements, utilisation des instruments de mesure pour la construction.
— pour G2 : figure robuste aux déplacements construite à partir des propriétés géométriques du carré et du demi-cercle, raisonnement par analyse-synthèse.

3.3 Analyse des Productions

En observant les vidéos des écrans des élèves ainsi qu'en les écoutant, nous analysons des techniques et éléments de discours relevant des trois modes de raisonnement.

L'élève E1 présente des aspects du mode de raisonnement G2. E1 utilise en effet à la fois les propriétés du carré et les propriétés des outils géométriques (cercle, perpendiculaire) pour sa construction qui aboutit à une figure robuste (cf. figure 1). C'est la procédure attendue à la fin du collège.

L'élève E2 présente des aspects du mode de raisonnement G1. En utilisant la souris, E2 déplace les points C et D jusqu'à avoir quatre longueurs égales et un angle droit en C, sa figure n'est pas robuste (cf. figure 1). E2 semble connaître au moins visuellement les propriétés du carré mais ne sait pas comment les mettre en oeuvre pour construire la figure.

E3 présente des aspects du mode de raisonnement G0. E3 fait apparaître, accidentellement, une grille en arrière plan et y déplace le segment [AB] donné (cf. figure 1). Le segment est alors dans une position prototypique pour le carré (à la verticale). E3 compte les carreaux pour placer les points C et D de manière à former visuellement un carré. Lorsqu'il s'agit de tracer le demi-cercle, E3 place un point « à peu près » au milieu du segment [AB] puisqu'il n'est pas possible de se référer à une intersection de la grille. E3 n'utilise ni les propriétés géométriques ni les instruments de mesure pour la construction mais se contente de placer les points au jugé.

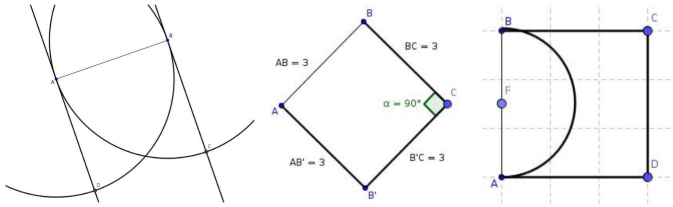


Fig. 1. De gauche à droite : construction de E1, de E2 et de E3.

4 Conclusion

Dans cet article, l'appui sur les théories de la didactique des mathématiques et sur une étude épistémologique et didactique du domaine mathématique considéré nous a permis de créer un modèle d'exercices et un modèle de l'apprenant opérationnels. En effet, lors de l'expérimentation, la suite d'exercices créés à partir d'un choix de valeurs de variables de type de tâches et de tâche pertinent nous a permis de décrire des éléments du rapport personnel des élèves à la géométrie.

À partir des variables de tâche, il est maintenant nécessaire de hiérarchiser les exercices en termes de complexité afin que la plateforme puisse calculer automatiquement des parcours d'apprentissage qui amèneront les élèves à utiliser les propriétés et raisonnements géométriques.

Références

1. Celi, V. : Une comparaison de l'enseignement de la géométrie dans le secondaire en france et en italie. Canadian Journal of Math, Science & Technology Education **5**(3), 377–399 (2005)
2. Chaachoua, H. : T4TEL un cadre de référence didactique pour la conception des EIAH. In : Séminaire national de didactique des mathématiques (2018)
3. Chevallard, Y. : Analyse des pratiques enseignantes et didactique des mathématiques : l'approche anthropologique. Actes de l'UE de la Rochelle pp. 91–118 (1998)
4. Euclide, Peyrard, F. : Les élémens de géométrie d'Euclide. chez F. Louis (1804)
5. Grugeon-Allys, B. : Modéliser le profil diagnostique des élèves dans un domaine mathématique et l'exploiter pour gérer l'hétérogénéité des apprentissages en classe : une approche didactique multidimensionnelle. e-JIREF **2**(2), 63–88 (2016)
6. Grugeon-Allys, B., Pilet, J., Chenevotot-Quentin, F., Delozanne, E. : Diagnostic et parcours différenciés d'enseignement en algèbre élémentaire. Recherches en Didactique des Mathématiques (Numéro spécial hors-série, Enseignement de l'algèbre élémentaire : bilan et perspectives), 137–162 (2012)
7. Houdement, C., Kuzniak, A. : Paradigmes géométriques et enseignement de la géométrie. Annales de didactique et de sciences cognitives **11**, 175–193 (2006)
8. Robert, A. : Outils d'analyse des contenus mathématiques à enseigner au lycée et à l'université. Recherches en didactique des mathématiques **18**(2), 139–190 (1998)
9. Soury-Lavergne, S. : De l'intérêt des constructions molles en géométrie dynamique. MathémaTICE Les nouvelles technologies pour l'enseignement des mathématiques (2011)

Session
Apprentissage de l'informatique

Reconnaissance de Motifs et Répétitions : Introduction à la Pensée Informatique

Yvan Peter[1], Marielle Léonard[2], Yann Secq[1]

[1] Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France
{prenom}.{nom}@univ-lille.fr
[2] Master IME, Poitiers - France-IOI
marielle.leonard59@gmail.com

Résumé Cet article présente une expérimentation d'initiation d'élèves de primaire à la pensée informatique. L'objectif est de développer la capacité de reconnaissance de motifs et leur expression sous la forme de répétitions via des activités déconnectées et en ligne. Une analyse des traces et des questionnaires pré et post activités ont été réalisés afin de mesurer l'évolution des compétences. Cet article porte sur la session 2018 impliquant 20 classes de CM1/CM2. Les résultats montrent un impact positif des activités et mettent en relief les compétences acquises.

Mots clés : Pensée Informatique · Programmation · Séquence pédagogique · École primaire

Abstract. This article presents an experimentation to introduce primary school pupils to Computational Thinking. The aim is to enhance their capability to spot repetitive patterns and to express them as loops. Unplugged and plugged-in activities are used to train the pupils. Trace analysis and pre and post questionnaires were used to measure the impact of the intervention. This article deals with the 2018 session involving 20 classes. The results show a positive impact of the activities and give information about the competencies acquired.

Keywords: Computational Thinking · Programming · Pedagogical Sequence · Primary school

1 Introduction

La pensée informatique[3] (PI) correspond à des compétences d'abstraction et de résolution de problèmes qui peuvent être exploitées dans de nombreux

3. Traduction imparfaite du terme *Computational Thinking* dont le périmètre et la nature sont encore sujet à discussion.

contextes et appliquées à des domaines divers, aussi bien à l'école qu'à l'extérieur. La PI en tant que compétence cognitive ne se réduit pas à la programmation. Toutefois, elle vise à exprimer les choses d'une manière qui se prête à un traitement informatique. En ce sens, elle permet d'aborder la dimension algorithmique et la programmation en général. Dans son article fondateur, J. Wing appelle à l'enseignement de la pensée informatique pour tous et assez tôt dans la scolarité [12]. Depuis, celle-ci a été appliquée à de nombreux domaines tels que les mathématiques et les sciences expérimentales [11], les arts [5] et même l'apprentissage des langues [8].

Cet article présente notre travail sur l'acquisition de compétences de base de la PI avec des élèves d'école primaire (8 - 10 ans). Nous nous sommes focalisés sur la capacité des élèves à reconnaître des motifs répétitifs et à les exprimer sous forme de boucles. Cette abstraction et cette expression ouvrent la voie à un traitement systématique des données. Nous présentons dans cet article les résultats de l'expérimentation de 2018. Celle-ci a impliqué 20 classes venant de 16 écoles pour un total de 447 élèves. Nous avons fait passer aux élèves des questionnaires avant et après l'activité pour mesurer leur capacité à identifier et coder les répétitions. Nous avons également collecté les traces d'activité de programmation sur la plate-forme.

Dans la suite, nous passons en revue l'état de l'art concernant la pensée informatique, comment elle est introduite à l'école et comment est réalisée l'évaluation des compétences. La section 3 présente la séquence pédagogique que nous avons définie et le contexte expérimental. Nous discuterons de l'analyse des résultats dans la section 4 avant de terminer sur la conclusion et les perspectives.

2 État de l'art

2.1 La pensée informatique

Les premiers concepts de PI datent des travaux de Seymour Papert avec le langage Logo [6]. Plus récemment, l'article de Jeannette Wing présentant la PI comme une compétence de base au même titre que la lecture, l'écriture et l'arithmétique a déclenché un vif intérêt au sein des communautés éducatives et de recherche [12]. Wing insiste sur le fait que la PI ne se réduit pas à la programmation mais correspond plutôt à la capacité de manipuler des abstractions et de résoudre les problèmes qui peut être appliquée à de nombreux domaines. Elle prône l'apprentissage de ces compétences dès l'école.

Depuis, un nombre croissant de travaux de recherche se sont développés pour explorer les moyens d'introduire la PI à l'école : quels sont les concepts fondamentaux à enseigner ? Quelles technologies se prêtent le mieux à ces apprentissages ? etc. Ces problématiques sont d'autant plus importantes que de nombreux pays ont commencé à introduire la PI et plus généralement «le code» dans les programmes à différents niveaux scolaires.

Rich *et al.* ont définit des *Trajectoires d'Apprentissage* (TA) en se basant sur une étude de la littérature. Une TA définit les concepts qui peuvent être abordés en fonction de la classe et avec quel niveau de détails [7]. Elle est formalisée

par un ensemble d'*objectifs d'apprentissage* associés à un *chemin d'apprentissage* permettant la réalisation de ces objectifs ainsi que des exemples d'activités qui peuvent être réalisées. Les auteurs relèvent que la plupart des travaux sont relatifs à un seul objectif d'apprentissage ou que, quand il y en a plusieurs, ils sont indépendants. Ils observent également que ces objectifs ont pu être appliqués à différents niveaux dans la mesure où ils s'adressent en général à des novices. Pour cette raison, ils se sont inspirés des approches pédagogiques et des programmes de mathématiques pour définir l'ordre d'introduction des concepts de la PI (chemin d'apprentissage). L'article illustre leur approche sur les concepts de Séquence, Répétition et Conditionnelle.

Ching *et al.* ont réalisé une étude des technologies disponibles pour l'enseignement de la PI [2]. Ils ont identifié *les robots pédagogiques*, *les kits robotiques*, *les jeux de plateau*, *les applications en réalité augmentée*, *les applications ou sites Web pour la programmation* (graphique ou textuelle), *les outils de développement orientés animation ou jeu.* Ces différentes catégories se distinguent selon qu'elles utilisent une manipulation physique ou une interaction sur l'écran et selon que la rétroaction est concrète (e.g. mouvement d'un robot) ou abstraite (i.e., retour sur l'écran). L'étude montre que les concepts abordés avec ces technologies vont des séquences et boucles à des concepts plus avancés et peuvent impliquer des activités créatives ou de résolution de problème.

Gouws *et al.* proposent un cadre de description des compétences liées à la PI [3]. En se basant sur une revue de la littérature, ils définissent un ensemble de compétences qui peuvent être abordées à travers la programmation : *Processus et Transformations*, *Modèles et Abstractions*, *Motifs et Algorithmes*, *Outils et Ressources*, *Inférence et Logique*, *Évaluations et Améliorations.* Ils combinent ces compétences avec un niveau de maîtrise inspiré de la taxonomie de Bloom : *Reconnaître*, *Comprendre*, *Appliquer et Assimiler.* Ce cadre peut être utilisé pour l'analyse et pour la conception d'activités.

Weintrop *et al.* s'intéressent à l'introduction des pratiques de la PI en mathématiques et en sciences de manière à permettre une définition des activités de PI indépendante de l'informatique [11]. Les auteurs définissent une taxonomie de 22 pratiques groupées selon les catégories suivantes : *Données et Informations*, *Modélisation et Simulation*, *Calcul*, *Résolution de problèmes* et *Approche Systémique.*

2.2 Évaluation des compétences de la pensée informatique

Brennan & Resnick articulent la PI autour de trois dimensions : Les *concepts algorithmiques* (boucles, parallélisme...), les *pratiques* (développement itératif, débugage...) et les *perspectives* (expression personnelle, connexion avec les autres...) [1]. Ils proposent d'évaluer ces dimensions à travers l'analyse des productions, des entretiens basés sur les productions et des projets de développement.

Le rapport du *Standford Research Institute* (SRI) produit par Snow *et al.* s'intéresse aux moyens d'évaluer les compétences de la PI (résolution de problème, abstraction...) dans le contexte de l'enseignement *Exploring Computer*

Science (ECS) qui se déroule sur un an au lycée aux États-Unis [10]. A cette fin, ils proposent des patrons de conception pour la création d'évaluations pertinentes des connaissances et des pratiques. Le rapport couvre l'évaluation des unités suivantes du cours : Interaction Homme-Machine, Résolution de problème, Conception Web et Introduction à la programmation. Ces évaluations incluent des quiz, des problèmes, de la lecture de code et le traçage de son exécution.

Grover *et al.* combinent évaluation formative et sommative dans le cadre d'un module d'introduction des concepts de la PI d'une durée de 6 semaines dans le secondaire [4]. Les évaluations reposent sur des questionnaires à choix multiples dont la plupart comprennent des extraits de code en Scratch. Certains exercices impliquent de remettre en ordre des blocs de code ou des activités de traçage et de debugage de code.

Seiter *et al.* proposent un cadre pour l'évaluation des compétences de la PI à l'école primaire baptisé *Progression of Early Computational Thinking* (PECT) [9]. Ce cadre s'appuie sur l'évaluation de programmes en Scratch (utilisation d'instructions spécifiques) dans le cadre de patrons de conceptions classiques liés à au développement d'animations ou de jeux (e.g., animation, gestion des collisions...). Ces patrons sont mis en relation avec des concepts de la PI. Le cadre proposé a été évalué sur la base des programmes disponibles sur le site Web de Scratch.

Les approches basées sur l'étude du code produit par les apprenants sont intéressantes à titre exploratoire mais chronophages. L'autre mode d'évaluation majoritaire utilise des QCM éventuellement basés sur de la lecture de code.

3 Cadre expérimental

3.1 Participants et organisation

L'expérimentation de 2018 a impliqué des élèves de CM1/CM2 de 16 écoles du secteur. Vingt classes ont participé pour un total de 447 élèves. L'âge des élèves était compris entre 8 et 10 ans avec un équilibre des genres (49% de filles). L'expérimentation s'est déroulée sur une semaine avec 5 classes par jour (hors mercredi). Les classes sont venues à l'université pour la journée (sous forme de sortie scolaire). Afin de prendre en charge un aussi grand nombre d'élèves, ceux-ci sont encadrés par des étudiants de deuxième année de DUT Informatique, eux-mêmes encadrés par les enseignants. La séquence pédagogique et les activités sont présentées au préalable aux étudiants de façon à ce qu'ils puissent encadrer et accompagner les élèves dans leurs apprentissages.

3.2 Progression pédagogique

La progression pédagogique est présentée figure 1. Celle-ci inclut des activités débranchées et en ligne liées à l'identification de motifs répétitifs ainsi que leur expression sous forme de séquences d'instructions et de boucles. Nous avons trois phases décrites dans la suite. Les deux premières durent 1h30 chacune, la dernière dure 2 heures. Les élèves suivent les deux premières phases le matin et la troisième l'après-midi.

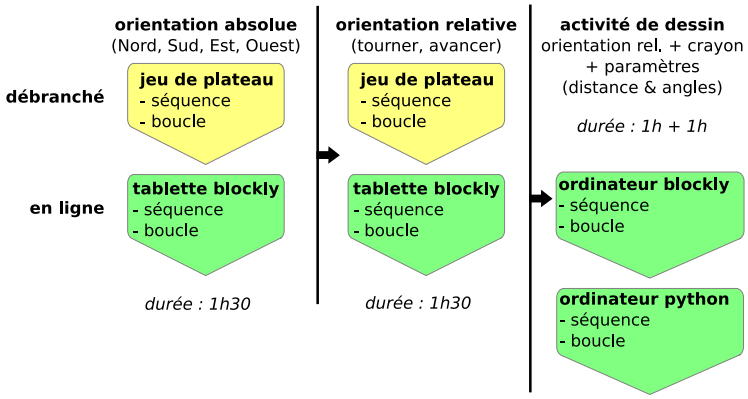


Figure 1. Progression pédagogique.

Orientation absolue Dans cette phase, les élèves doivent déplacer un personnage sur une grille jusqu'à une case définie en utilisant une orientation absolue (Nord, Sud, Est, Ouest). Ils démarrent l'activité avec un jeu de plateau (figure 2). Les élèves travaillent en groupe de 4 à 6 et prennent à tour de rôle différentes fonctions : élaborer une solution (i.e., algorithme), *pointeur d'instruction* (qui annonce l'ordre à exécuter), et *processeur* (qui exécute l'ordre en déplaçant le personnage). La complexité des chemins évolue avec l'adjonction d'obstacles et de bonus. Quand la séquence d'ordres gagne en longueur, la frustration des élèves augmente. C'est le bon moment pour introduire la notion de répétition (*répéter n fois*).

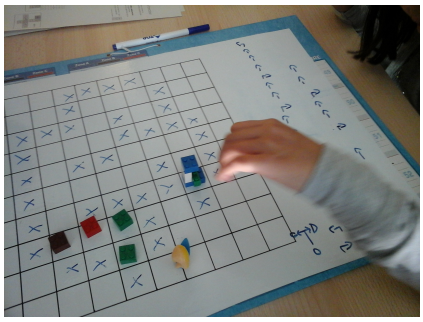


Figure 2. Jeu de plateau.

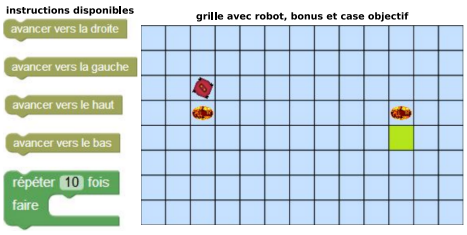


Figure 3. Puzzle sur tablette avec le jeu d'instruction

Quand les concepts principaux d'instruction, de séquence, de boucle, d'exécution (et de bug...) sont compris, les élèves passent sur tablette par binômes pour une activité similaire. Ils utilisent un langage de programmation visuelle basée sur des blocs (*Blockly*) (figure 3). Pour chaque puzzle, les élèves disposent

d'un jeu d'instructions spécifique et peuvent être limités en nombre d'instructions. Les activités évoluent à nouveau de la séquence à la boucle pour renforcer les concepts.

Orientation relative Cette phase suit une organisation similaire. Le changement principal tient à l'utilisation d'un personnage orienté qui implique un jeu d'ordres différent (*tourner à gauche*, *tourner à droite*, *avancer*). Cela implique également de se rappeler l'orientation du personnage dans la phase de planification. Le personnage tourne obligatoirement de 90°. Cette phase prépare la phase de dessin pour laquelle l'orientation est nécessaire.

Activité de dessin La dernière phase est réalisée en salle de TP avec un ordinateur par élève. Les élèves réalisent des activités de dessin avec une tortue (dans l'esprit de Logo). Le jeu d'instructions est similaire à la phase précédente avec en plus la gestion du crayon (lever et baisser pour dessiner) et la paramétrisation des instructions (e.g., `avancer(`*distance*`)` ou `tourner à droite(`*angle*`)`). Les élèves utilisent la même plate-forme que sur les tablettes. Dans la première partie, ils continuent d'utiliser *Blockly* pour dessiner. Dans la seconde partie, nous introduisons la programmation *Python*, les faisant ainsi passer d'une programmation par blocs à une programmation textuelle. Afin de faciliter les choses, les élèves utilisent des instructions en français (e.g., `avancer(10)`). Toutefois, ils utilisent la syntaxe *Python* pour les boucles.

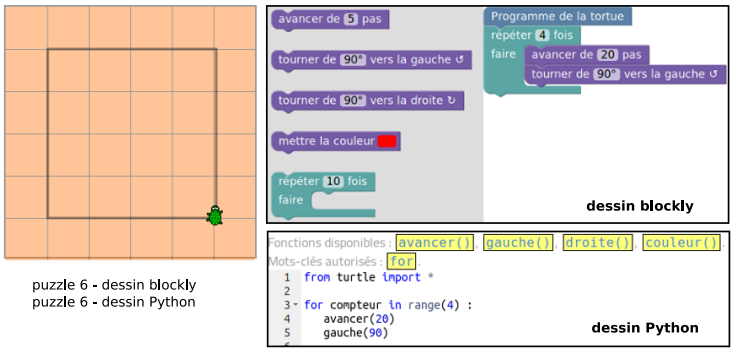
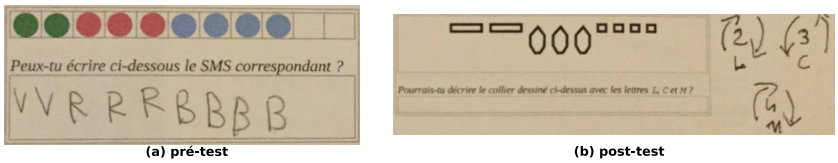


Figure 4. Passage de la programmation par blocs à textuelle.

Les langages textuels sont en général introduits beaucoup plus tard dans la scolarité. L'introduire dans la dernière séquence sur un contexte similaire et avec la même interface (figure 4) nous donne des indices sur la capacité des élèves à transférer leurs compétences de la programmation par blocs à textuelle pour cette classe d'âge.

3.3 Collecte des données

Pré et post tests Les élèvent passent un test au début et à la fin de la journée de manière à mesurer l'évolution de leur capacité à identifier des motifs répétitifs et à les exprimer sous la forme de répétitions qui peuvent être traitées par des boucles. Les tests portent sur le codage de motifs avec des lettres. Les élèves sont informés qu'ils peuvent utiliser la notation qui leur semble la plus appropriée, y compris raccourcie. Le pré test consiste à coder une représentation graphique d'une partition (figure 5(a)). Le post test porte sur le codage d'instructions pour la réalisation d'un collier de pâtes (figure 5(b)). Nous avons choisi deux contextes différents pour éviter que les élèves ne se contentent de se rappeler des motifs du pré test.



(a) pré-test (b) post-test

Figure 5. Motifs des pré et post tests.

Le tableau 1 présente les motifs utilisés, qui passent d'une séquence d'une instruction et jusqu'à trois instructions pour les plus complexes. La notation présentée ici correspond au pré test mais nous avons les mêmes motifs pour le post test. Le *motif* correspond à ce qui est présenté aux élèves et l'*encodage* à ce qui est attendu des élèves, la *correspondance* définit à quoi correspond le *type*.

Table 1. Motifs des tests

Type	Correspondance	Motif	Encodage
`1i`	motif 1 instruction	`RRRRRRRRRRR`	`11xR`
`Nx1i`	N x motifs 1 instruction	`VVRRRBBBB`	`2xV 3xR 4xB`
`2i`	motif 2 instructions	`BRBRBRBRBR`	`5x(BR)`
`3i+2i`	motifs 2 instructions + 1 instruction	`VRRVRRBBBBB`	`2x(VRR) 4B (2x(V 2R) 4B)`

Activités de programmation Les activités de programmation ont été réalisées sur la plate-forme de France-IOI[4]. Les quatre séquences comprennent un ensemble de puzzles de difficulté croissante. On retrouve successivement des séquences, des boucles à une instruction, un mélange des deux, des boucles à plusieurs instructions et jusqu'à des boucles imbriquées. La première séquence comporte 24 puzzles avec une difficulté très progressive afin que les élèves puissent

4. Association qui organise le concours Algorea.

construire leurs compétences. Les séquences suivantes comprennent de 15 à 18 puzzles. Elles démarrent avec des puzzles simples (séquence) de manière à ce que les élèves puissent transférer leurs compétences sur un nouveau jeu d'instructions puis la difficulté augmente. La figure 6 présente quelques puzzles qui sont analysés dans la suite.

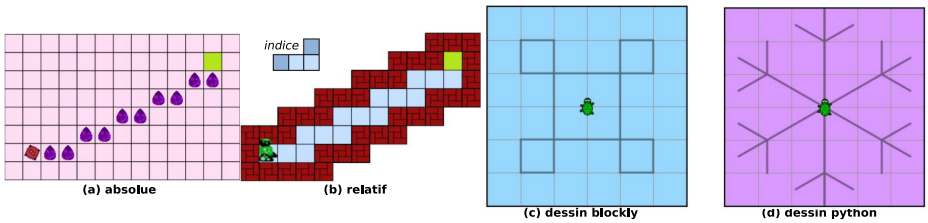


Figure 6. Puzzles difficiles dans chaque phase.

La plate-forme passe d'un puzzle au suivant après validation mais elle permet également de sélectionner un puzzle spécifique. Chaque séquence réalisée sur la plate-forme dure 30 à 45 minutes suivant les groupes. Cela explique que les élèves n'ont pas réalisé le même nombre de puzzles suivant leur rapidité. Quand de nouveaux concepts sont introduits (e.g., la boucle), le premier puzzle comprend une aide pour la réalisation.

4 Résultats

4.1 Analyses des questionnaires

Les questionnaires pré et post activité ont été codés en fonction de la reconnaissance ou non des motifs par l'élève pour arriver à une valeur entre 0 et 1. Pour les 447 élèves, nous obtenons une moyenne $m = 0,147$, $sd = 0,07$ pour le pré test et $m = 0,241$, $sd = 0,12$ pour le post test. Un t-test apparié nous donne une valeur $t(446) = -6,76$ ($p <, 0001$) qui confirme l'impact significatif de la séquence pédagogique sur la capacité des élèves à repérer et encoder des motifs répétitifs.

Table 2. Résultats détaillés.

	1i	Nx1i	2i	3i+2i
pré test	118 (27%)	102 (23%)	29 (8%)	14 (4%)
post test	166 (38%)	140 (32%)	70 (17%)	56 (14%)

Le tableau 2 présente un détail de résultats pour les différents types de motifs de complexité croissante. On note que la reconnaissance des motifs à une

instruction et des séquences de boucles à une instruction augmente de manière significative (+40% pour les deux). Le plus intéressant est probablement que la reconnaissance des motifs plus complexes (2 et 3 instructions) a progressé encore plus. Cela pourrait impliquer qu'après un apprentissage sur des motifs courts, la compétence est généralisée rapidement à des motifs plus complexes par les élèves. La figure 5 montre le cas idéal d'un élève qui n'a utilisé aucune notation de répétition sur le pré test mais l'a appliqué avec succès sur le post test.

4.2 Analyse des activités

Nous avons eu un export des traces d'activités des participants sur la plateforme. Celles-ci sont malheureusement limitées car nous n'avons que la dernière validation de l'élève pour chaque puzzle et le moment où cette validation est faite. Nous n'avons donc pas l'historique des essais des élèves. La validation du puzzle peut être positive (i.e., le puzzle est réussi) ou négative. Pour chaque séquence d'activité en ligne, nous présentons un graphe indiquant pour chaque puzzle le taux de succès (nombre validations positives / nombre de validations total) et le nombre total de validations (essais). Nous présentons également les transitions entre les niveaux de difficulté (e.g. de séquence à boucle) ce qui nous permet de visualiser les paliers de difficulté.

Pour les deux premières phases, les élèves partagent une tablette à deux et résolvent un puzzle chacun leur tour. En pratique, ils collaborent généralement pour résoudre le puzzle même s'ils n'en ont pas eu la consigne. Ceci explique le nombre d'essais maximal autour de 200. Pour la dernière phase, les élèves travaillent seuls sur ordinateur et nous avons potentiellement 447 essais (nombre d'élèves impliqués). Nous avons perdu des essais sur la première et la dernière séquence suite à des problèmes techniques, ce qui explique des chiffres inférieurs.

Orientation absolue La figure 7 montre les résultats de la première phase. Nous avons une progression mesurée de la difficulté des puzzles donnant un taux de succès supérieur à 90%. On observe une diminution du nombre d'essais quand on aborde les puzzles avec boucles qui indique que certains élèves commencent à rencontrer des difficultés. Toutefois, la chute brutale correspond aux boucles à plusieurs instructions (i.e., motifs plus longs) avec un taux de succès qui tombe à 66% pour le puzzle 19 (figure 6(a)) (le puzzle 18 est un tutoriel).

Orientation relative La figure 8 correspond à la seconde phase. Le taux de succès proche ou supérieur à 90% indique que les élèves ont réussi à transférer leur compétences sur le nouveau langage. On note également une amélioration de leur capacité à gérer les boucles à une instruction. En effet, on observe encore 137 essais avec un succès raisonnable pour le puzzle 7 (à comparer aux 51 essais du puzzle 17 de la phase précédente). A nouveau, passer aux boucles à plusieurs instructions présente une difficulté majeure. Le puzzle 8 (figure 6(b)) qui a un nombre d'essais significatif atteint un taux de succès de 56% seulement. On peut noter qu'étant le premier puzzle de ce type, il y avait une indication sur le motif à traiter.

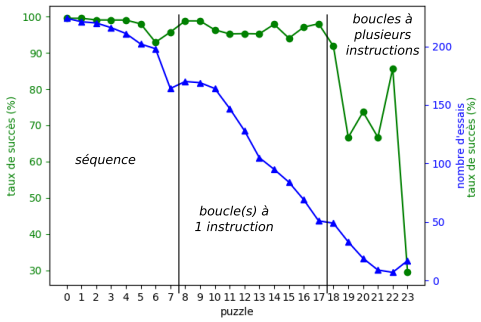


Figure 7. Orientation absolue.

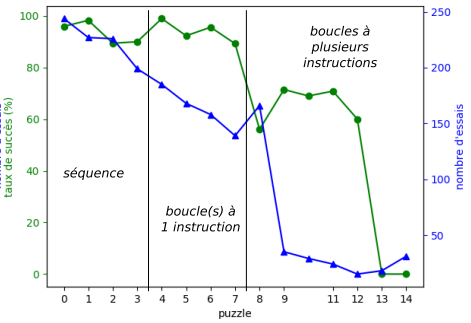


Figure 8. Orientation relative.

Dessin en Blockly La dernière phase introduit de nouvelles difficultés. D'abord les instructions sont paramétrées. Ensuite, on commence à avoir des angles différents de 90°que les élèves n'ont pas encore vu en classe.

A nouveau, le nombre d'essais et le taux de succès des premiers puzzles (y compris avec boucles) indiquent que le changement de langage n'est pas un problème pour les élèves une fois que les compétences liées aux séquences et boucles sont intégrées.

Les puzzles qui combinent séquences d'instructions et boucles à plusieurs instructions restent difficiles (puzzle 9 - 11, 9 fournissant des indices) dans la mesure où l'on voit le nombre d'essais chuter. La figure 6(c) montre le puzzle 11 pour lequel nous avons encore 129 essais mais avec un taux de succès de 58%. Peu d'élèves ont abordé les puzzles avec des boucles imbriquées mais avec un taux de succès légèrement supérieur à 60%.

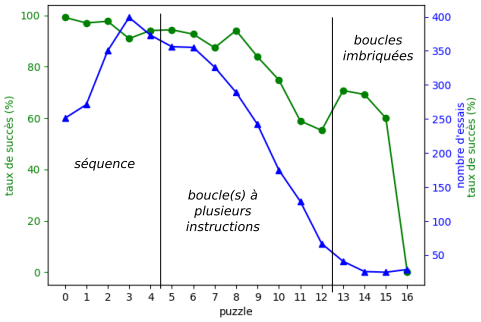


Figure 9. Dessin en *Blockly*.

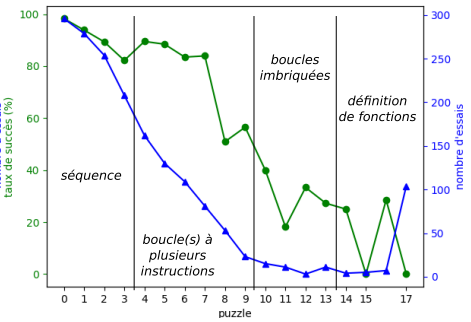


Figure 10. Dessin en *Python*.

Dessin en Python La figure 4 permet de voir que les élèves peuvent utiliser des boutons pour générer les instructions afin d'éviter d'avoir à tout taper au

clavier. Ils doivent malgré tout adapter les paramètres de ces instructions à leurs besoins. La première séquence de puzzles a toujours un bon taux de succès avec 82% et une bonne participation [5]. C'est un résultat intéressant qui montre que les élèves ont bien transféré leurs compétences d'un langage visuel à un langage textuel.

Les puzzles impliquant séquence et boucles sont toujours difficiles (puzzles 8 et 9) avec un taux de succès à peine supérieur à 50%. Les boucles imbriquées sont également un point difficile. Le puzzle 10 est un tutoriel. Le puzzle 11 présenté figure 6(d) atteint seulement 18% de taux de succès avec un faible nombre d'essais. Le dernier puzzle correspond à une activité où les élèves peuvent dessiner librement. Il n'y a donc pas de condition de validation. Le graphe montre qu'un bon nombre d'élèves en on profité.

5 Conclusion

Cet article porte sur l'apprentissage des concepts fondamentaux et des compétences de la pensée informatique par des élèves de 8 - 10 ans. Nous avons conçu une séquence pédagogique pour initier les élèves aux notions d'instruction, de séquence et de boucles et leur faire pratiquer ces compétences sur différents langages (libre (activité débranchée), blockly et textuel). Cette séquence a évolué depuis nos premières expérimentations en 2014 et nous avons un nombre croissant de classes impliquées. Nous cherchons également à améliorer nos instruments de mesure pour vérifier l'impact sur les élèves (questionnaires et analyse des activités).

Les résultats de l'expérimentation de 2018 sont encourageants. L'analyse statistique des questionnaires montre l'impact significatif de la séquence et nous avons une augmentation significative des élèves qui identifient les motifs répétitifs et sont capables de les synthétiser sous forme de boucles. Il semble également qu'une fois que les élèves ont acquis cette compétence pour des motifs de longueur 2 (2 instructions dans la boucle), ils sont capables de généraliser à des motifs plus longs.

Les résultats de l'analyse des activités sont à prendre avec précaution dans la mesure où, bien sûr, ils sont aidés par les étudiants voire les professeurs et parents accompagnant. Avec un étudiant pour 4 à 6 élèves, on peut toutefois espérer que cela n'a pas un impact trop fort sur les résultats.

Les élèves reçoivent un diplôme à la fin de la journée qui donne l'adresse de la plate-forme ainsi que leur code d'identification. Nous avons observé environ 300 accès dans les jours suivants et jusqu'à deux mois après (date de la récupération des données). Toutes les séquences ont été utilisées y compris le dessin en Python pour laquelle il y a eu 271 essais. Ces éléments sont encourageants et montrent la motivation des élèves.

Nous avons réalisé une nouvelle expérimentation en janvier 2019 qui doit encore être dépouillée. Nous avons bénéficié d'une nouvelle version de la plate-forme qui nous donnera des informations plus précises (début et fin des activités,

5. Nous avons perdu des traces suite à un problème technique.

essais réalisés). Cela nous permettra d'affiner notre analyse et d'aborder d'autres points comme les stratégies de résolution des élèves. Nous avons également fait évoluer nos questionnaires pour identifier un biais potentiel lié à la disposition des motifs à coder.

Remerciements

Nous remercions l'association France-IOI pour la mise à disposition de la plate-forme. Ce travail est partiellement soutenu par le projet Interreg Dig-e-lab.

Références

1. Brennan, K., Resnick, M. : New frameworks for studying and assessing the development of computational thinking. annual American Educational Research Association meeting, Vancouver, BC, Canada pp. 1–25 (2012)
2. Ching, Y.H., Hsu, Y.C., Baldwin, S. : Developing Computational Thinking with Educational Technologies for Young Learners. TechTrends (1980) (2018)
3. Gouws, L., Bradshaw, K., Wentworth, P. : Computational thinking in educational activities. In : Proceedings of the 18th ACM conference on Innovation and technology in computer science education - ITiCSE '13. p. 10 (2013)
4. Grover, S., Cooper, S., Pea, R. : Assessing computational learning in K-12. Proceedings of the 2014 conference on Innovation & technology in computer science education - ITiCSE '14 pp. 57–62 (2014)
5. Knochel, A.D., Patton, R.M. : If Art Education Then Critical Digital Making : Computational Thinking and Creative Code. Studies in Art Education : A Journal of Issues and Research **57**(1), 21–38 (2015)
6. Papert, S. : Mindstorms : Children, computers, and powerful ideas. Basic Books, Inc. (1980)
7. Rich, K.M., Strickland, C., Binkowski, T.A., Moran, C., Franklin, D. : K–8 Learning Trajectories Derived from Research Literature : Sequence , Repetition , Conditionals. acm Inroads **9**(1), 46–55 (2018)
8. Sabitzer, B., Jarnig, M. : Computational Thinking Through Modeling In Language Lessons. In : IEEE Global Engineering Education Conference (EDUCON). pp. 1913–1919 (2018)
9. Seiter, L., Foreman, B. : Modeling the learning progressions of computational thinking of primary grade students. Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research pp. 59–66 (2013)
10. Snow, E., Tate, C., Rutstein, D., Bienkowski, M. : Assessment Design Patterns for Computational Thinking Practices in Secondary Computer Science. Tech. Rep. December, SRI International (2017)
11. Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., Wilensky, U. : Defining Computational Thinking for Mathematics and Science Classrooms. Journal of Science Education and Technology **25**(1), 127–147 (2016)
12. Wing, J.M. : Computational Thinking. COMMUNICATIONS OF THE ACM **49**(3), 33–35 (2006)

Objet Tangible ou Simulation Numérique : Deux Situations Équivalentes pour l'Apprentissage de la Programmation ?

Grégoire Fessard[1], Patrick Wang[1] et Ilaria Renna[1]

Institut Supérieur d'Électronique de Paris, Paris, France
{gregoire.fessard, patrick.wang, ilaria.renna}@isep.fr

Résumé. La programmation physique est une approche visant à programmer et interagir avec un objet tangible programmable pour apprendre des concepts clés de la programmation. Cette approche présente de nombreux avantages en termes de motivation, de créativité, et d'apprentissage de la programmation. Cependant, ces apprentissages sont rarement comparés à ceux résultant de la programmation d'une simulation numérique de l'objet tangible utilisé. Cet article présente les résultats d'une telle étude comparative. Celle-ci cherche à déterminer si ces deux situations amènent à des différences significatives dans l'apprentissage de concepts clés de la programmation. Les participants à cette étude ont été séparés en deux groupes : un programmant une carte électronique (la BBC micro:bit), l'autre programmant une simulation de cette carte. Les résultats montrent qu'il y a un apprentissage des concepts fondamentaux de la programmation quel que soit le groupe. Cependant, aucune différence significative n'a été constatée lorsque les gains d'apprentissage ont été comparés entre les deux groupes.

Mots-clé: Apprentissage de l'informatique · Programmation physique · Programmation numérique · Programmation par blocs.

Abstract. Physical computing is about programming and interacting with a tangible object to learn key concepts of Computer Science. This approach presents several benefits, namely increased student motivation, creativity, and learning gains. Yet, these learning gains hardly are compared with those resulting from the programming of a simulation of a tangible object. This article presents the results of a comparative study we conducted to explore this issue. With this study, we aimed at determining whether the programming of a tangible object or its digital simulation yields significantly different learning gains. Participants were divided into two groups: one programmed an electronic board (the BBC micro:bit) while the other programmed a simulation of it. The results of this experiment suggest that, while both groups significantly improved their understandings of fundamental programming concepts, no significant difference was found when comparing the learning gains between the two groups.

Keywords: Computer Science Education · Physical computing · Digital computing · Block-based programming.

1 Introduction

La programmation physique (ou *physical computing* en Anglais) est un paradigme dont l'objectif est de « concevoir des objets tangibles et interactifs en utilisant des composants programmables » [14]. Cette approche se distingue d'un enseignement de la programmation utilisant historiquement la console comme seule interface lors de l'exécution d'un programme. Suite à l'émergence des robots et cartes programmables, la programmation physique est devenue de plus en plus exploitée et, en conséquence, a fait l'objet de nombreuses études portant sur l'initiation à la programmation, en particulier chez les plus jeunes [7,18]. Pour autant, et bien que focalisées sur les effets de la programmation physique chez les apprenants, ces études ne comparent pas leurs résultats avec des situations relevant de la programmation *numérique* (i.e., la programmation d'une simulation numérique d'un objet tangible).

Dans cet article, nous présentons une étude menée afin d'investiguer si une des deux situations (programmation physique ou numérique) conduit à des gains d'apprentissage en programmation plus importants que l'autre. Pour cett étude préliminaire, nous avons utilisé la carte électronique BBC micro:bit que nous avons proposée sous sa forme tangible ou simulée à des débutants en programmation afin de proposer deux expériences de programmation identiques. Un langage de programmation par blocs a été utilisé pour programmer cette carte. Pour cette étude comparative, nous nous sommes plus particulièrement intéressés à l'apprentissage de trois concepts fondamentaux de la programmation : les variables, les structures conditionnelles, et les structures itératives.

Dans la section qui suit, nous allons présenter un état de l'art portant sur l'enseignement de la programmation en nous intéressant principalement à la programmation physique et à l'utilisation de langages de programmation par blocs. En Sect. 3 et 4, nous décrivons l'environnement de programmation utilisé pour notre expérimentation ainsi que la conception du protocole expérimental. La Sect. 5 présente les résultats de notre expérimentation. Enfin, les deux dernières sections proposent une discussion de ces résultats et des pistes de travaux futurs.

2 État de l'art

L'apprentissage de la programmation présente des difficultés bien répertoriées dans la littérature. Par exemple, du Boulay identifie, entre autres, trois difficultés : la maîtrise de la syntaxe d'un langage de programmation, la compréhension de la structure d'un programme ou d'un algorithme, et la compréhension du rapport entre un programme et la machine qui peut ainsi être programmée [1].

La programmation visuelle, et en particulier la programmation par blocs, peut être utilisée pour pallier les deux premières difficultés. Ces langages de

programmation permettent de se focaliser sur la conception d'un algorithme sans se soucier de la syntaxe d'un langage de programmation. Une étude comparative a montré que la programmation par blocs entraînait une meilleure compréhension des concepts élémentaires de la programmation qu'un langage de programmation standard [21]. Cet aspect explique les nombreuses études mettant en œuvre des langages utilisant des blocs pour l'introduction à la programmation [12,21].

Le constructionnisme [13] est une théorie souvent mentionnée dans les travaux de recherche portant sur l'apprentissage de la programmation et qui peut pallier la troisième difficulté. Cette théorie stipule que l'acquisition de nouvelles connaissances peut se produire grâce à la *construction* d'un objet synthétisant des objectifs pédagogiques identifiés. Le constructionnisme a ouvert la voie à de nombreuses études mettant en œuvre des objets programmables tangibles [3,7,9] ou simulés [4,12] pour l'apprentissage de la programmation.

La programmation physique s'inspire donc du constructionnisme en proposant de programmer des robots ou des cartes électroniques. L'état de l'art sur la programmation physique met l'accent sur les bienfaits d'une telle approche : gain de motivation et de créativité [18], inclusion de populations sous-représentées [17], et apprentissage de la programmation [3]. Cependant, des résultats similaires concernant l'apprentissage de la programmation ont été obtenus lorsque des objets programmables numériques ont été utilisés [4,10,12].

Bien que ces deux situations mènent à un apprentissage de la programmation, peu de travaux ont cherché à déterminer si une situation est mieux adaptée que l'autre. Concernant l'apprentissage de la programmation, nous n'avons pu identifier que trois articles mettant en œuvre un objet tangible et sa simulation [2,5,6]. Cependant, ces études n'ont pas été conçues pour comparer les gains d'apprentissage entre ces deux situations mais plutôt pour proposer deux systèmes complémentaires. Notre étude cherche donc à combler ce vide dans l'état de l'art et à apporter une réponse à la question de recherche suivante : « y a-t-il des gains d'apprentissage différents lorsqu'un apprenant programme un objet tangible ou une simulation numérique de cet objet tangible ? » En particulier, nous nous focalisons sur l'apprentissage de trois concepts fondamentaux de la programmation : les variables, les structures conditionnelles, et les structures itératives. Nous avons donc reformulé notre hypothèse de recherche en trois sous-questions, une pour chaque concept de programmation mentionné précédemment.

3 Environnement de programmation

L'environnement de programmation mis en œuvre dans cette étude est composé de deux éléments : la BBC micro:bit en tant qu'objet programmable et une interface Web permettant de programmer une micro:bit grâce à un langage de programmation par blocs. Étant donné le contexte de notre étude, nous avons conçu deux versions distinctes de cet environnement : une première permettant de programmer une carte micro:bit tangible et une seconde permettant de programmer une simulation numérique de cette carte. Comme nous allons le décrire dans la suite de cette section, ces deux environnements ont été conçus pour pro-

poser une expérience de programmation pratiquement identique selon que l'on programme l'objet tangible ou la simulation.

3.1 Objet programmable

La micro:bit [1] est une carte programmable conçue en 2016 et ayant fait l'objet d'études dans le domaine de l'apprentissage de la programmation [16,18,19]. Cette carte est équipée, entre autres, de deux boutons-poussoirs et d'une grille de diodes électroluminescentes permettant la gestion d'entrées-sorties (voir Fig. 1). Ces LED peuvent être utilisées pour afficher aussi bien des informations alphanumériques que des formes pré-définies (voir la croche illustrée en Fig. 2), et un mécanisme de défilement permet d'afficher des messages arbitrairement longs.

La simulation numérique de la micro:bit a repris un programme *open source* écrit en Javascript [2] permettant de mimer à l'identique le comportement d'une carte micro:bit tangible. Ce programme nous a donc permis d'intégrer facilement la simulation de la micro:bit dans l'interface Web de programmation.

Sous sa version tangible, la micro:bit peut être reliée à un ordinateur par un câble USB. En transférant un fichier binaire dans son espace de stockage interne, il est possible de mettre à jour le programme exécuté par la carte [3]. La micro:bit simulée peut quant à elle être programmée en faisant appel à des fonctions Javascript implémentées dans l'interface Web de programmation.

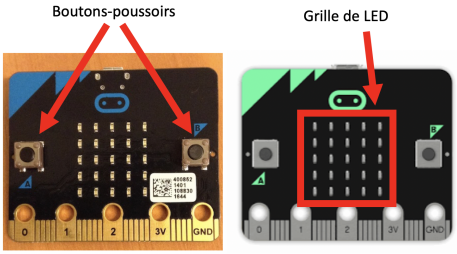


Fig. 1. Une carte micro:bit tangible (à gauche) et sa simulation numérique (à droite).

3.2 Interface de programmation

Pour programmer la micro:bit, nous avons développé une interface Web de programmation illustrée en Fig. 2. Cette interface implémente la bibliothèque Blockly [4] qui permet de construire des programmes à l'aide d'un langage de programmation par blocs.

1. https://microbit.org/
2. https://github.com/Microsoft/pxt-microbit
3. Cette étape a été automatisée pour rendre ce processus identique entre les deux objets, comme mentionné en Sect. 3.2.
4. https://developers.google.com/blockly/

La Fig. 2 illustre plus spécifiquement l'interface qui a été développée pour la programmation de la simulation de la micro:bit. Cette interface est divisée en trois zones. La Zone 1 affiche les énoncés des exercices sur lesquels devront travailler les participants à l'expérimentation. Un menu déroulant permet de passer d'un exercice à l'autre. La Zone 2 correspond à l'environnement de construction d'un programme. Celle-ci se compose d'une boîte à outils (sur fond gris) qui contient les blocs de programmation. Elle est composée de cinq catégories : Affichage, Math, Variables, Logique, et Boucle. Dans chacune de ces catégories, il est possible de trouver des blocs de programmation spécifiques. L'utilisateur peut ainsi construire des programmes en glissant et déplaçant ces blocs depuis la boîte à outils vers l'espace central (sur fond blanc) et en les emboîtant à la manière d'un puzzle. Enfin, la Zone 3 affiche la simulation de la carte micro:bit.

L'interface utilisée pour programmer l'objet tangible reprend à l'identique les Zones 1 et 2. Le contenu de la Zone 3 est simplement remplacé par un texte expliquant comment exécuter un programme sur la carte micro:bit tangible.

Quel que soit l'objet programmable utilisé, l'exécution d'un programme se fait grâce au bouton « Lancer le programme » situé en bas de la Zone 2. Cela est rendu possible grâce à Blockly qui peut traduire chaque bloc en Javascript ou Python, deux langages utilisés pour programmer respectivement la simulation et la carte tangible. L'environnement de programmation propose ainsi la même expérience utilisateur quelle que soit la modalité utilisée.

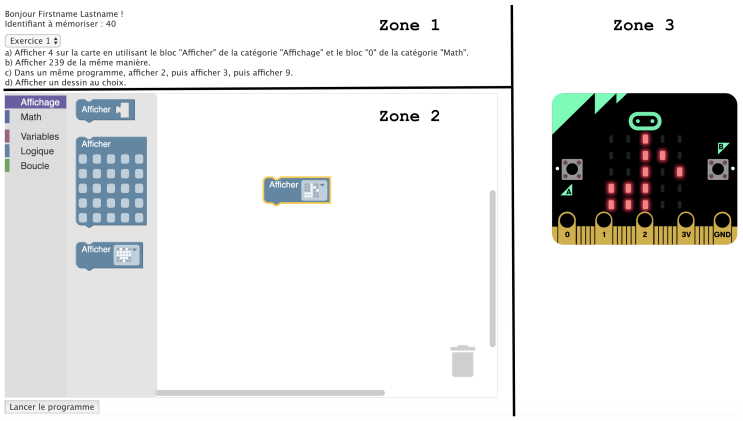


Fig. 2. Environnement de programmation par blocs présentant une simulation numérique de la micro:bit.

4 Protocole expérimental

Pour répondre aux questions de recherche formulées précédemment, nous avons conduit une expérimentation sur une demi-journée avec 36 collégiens et lycéens

âgés de 14 à 17 ans (dont 9 filles et 27 garçons, avec une moyenne d'âge de 15.2 ans). Sur ces 36 participants, 23 ont déclaré n'avoir aucune connaissance en programmation et 13 ont déclaré avoir des connaissances élémentaires de programmation. Ainsi, nous avions des participants capables de comprendre et d'apprendre des exercices proposés, ce qui nous permettait de quantifier l'utilité de l'environnement d'apprentissage conçu. Ces participants ont été répartis dans deux groupes : un groupe programmant l'objet tangible, l'autre programmant la simulation de la micro:bit. La composition des groupes a été décidée en amont de l'expérimentation et ce afin de garantir leur homogénéité en termes d'âge, de genre, et de connaissances initiales en programmation. Cette répartition a été validée par l'analyse *a posteriori* des résultats (voir Sect. 5.1). Le protocole expérimental conçu pour cette étude comporte un questionnaire pré-post test ainsi qu'une série d'exercices portant sur les concepts de variables, de structures conditionnelles, et de structures itératives et dont les énoncés sont affichés dans la Zone 1 de l'interface de programmation comme vu précédemment. L'hypothèse expérimentale que nous souhaitions tester était que la programmation d'un objet tangible ou d'une simulation équivalente de cet objet tangible conduirait à des gains d'apprentissage différents pour chacun de ces trois concepts, sans se risquer à émettre un jugement sur l'une ou l'autre de ces situations.

L'expérimentation a suivi le déroulement suivant. Dans un premier temps, l'environnement de programmation et le langage de programmation par blocs ont été présentés aux participants. Durant cette phase, aucune notion de programmation n'a été présentée en dehors de l'affichage d'informations en sortie (i.e., sur la grille de LED) puisque nécessaire pour la compréhension des questions du pré-post test. Puis, les participants ont répondu au pré-test, ont ensuite travaillé sur les exercices de programmation, et enfin ont répondu au post-test. Au cours de cette séquence, les participants devaient travailler de façon individuelle et pouvaient demander de l'aide uniquement en cas de problème technique.

4.1 Questionnaire pré-post test

Le questionnaire pré-post test est composé de 14 questions à choix multiple : 5 questions portent sur les variables, 5 sur les structures conditionnelles, et 4 sur les structures itératives. Nous nous sommes servis de la taxonomie de Bloom révisée [8] pour concevoir les énoncés des questions et pour définir un ordre de difficulté croissant pour les exercices de chaque notion.

Ces énoncés ont aussi été conçus pour que chaque question ne porte que sur un seul et unique concept de la programmation à la fois. Par exemple, cela implique que nous n'avons pas incorporé de question utilisant la notion de variable de boucle puisque celle-ci requiert des connaissances à la fois sur le concept de variable et de structure itérative. Cette précision est primordiale puisque, sans cette distinction, il aurait été beaucoup plus compliqué de quantifier les gains d'apprentissage pour chaque notion séparément.

La Fig. 3 présente trois exemples de questions utilisées dans le pré-post test. Pour chaque question, une seule réponse correcte est possible. Les réponses erronées reprennent quant à elles des erreurs classiques de programmation réper-

toriées dans la littérature [15]. Cela nous a permis de proposer des réponses plausibles qu'un débutant en programmation pourrait assimiler comme vraies.

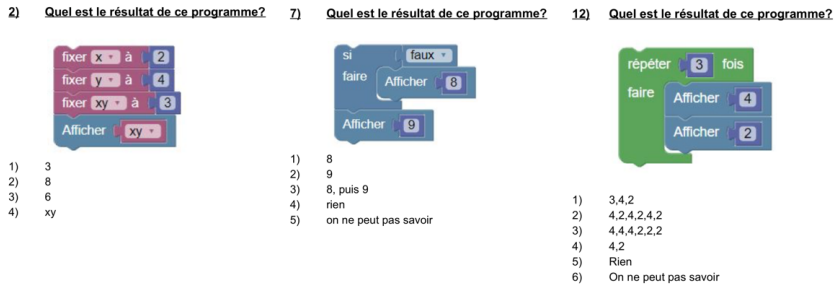


Fig. 3. Exemples de questions du pré-post test pour chaque concept étudié.

4.2 Exercices de programmation

Les participants ont travaillé sur 15 exercices de programmation portant sur les trois concepts mentionnés précédemment. Sur ces 15 exercices, 4 portent exclusivement sur le concept de variable, 3 sur le concept de structure conditionnelle, 4 sur le concept de structure itérative. Nous avons aussi proposé un exercice de prise en main de l'environnement de programmation (à réaliser avant tous les autres) et 3 exercices mêlant plusieurs concepts cités plus haut (à réaliser après tous les autres). Deux exemples d'exercices sont illustrés en Fig. 4, un portant sur les variables et l'autre sur les structures conditionnelles.

La conception de ces exercices a suivi une approche identique à celle utilisée pour la construction du questionnaire pré-post test. Nous nous sommes basés sur l'état de l'art de Qian et Lehman [15] portant sur les erreurs de programmation fréquemment commises par des apprenants pour concevoir des exercices pouvant amener nos participants à reproduire et à apprendre de leurs erreurs. Les exercices relatifs à un même concept ont aussi été proposés dans un ordre de difficulté croissant. Les premiers exercices de chaque concept avaient tous un énoncé détaillé avec une résolution guidée ; puis les exercices suivants demandaient des manipulations plus complexes tout en supprimant progressivement le guidage de l'énoncé. Par exemple, on peut voir en Fig. 4 que l'énoncé de l'exercice 6 (premier exercice sur les structures conditionnelles) est beaucoup plus détaillé que celui de l'exercice 5 (dernier exercice sur les variables).

Il était demandé aux participants de travailler les exercices dans l'ordre proposé. Cependant, aucun mécanisme de vérification n'a été implémenté et les participants étaient libres de passer à l'exercice suivant quand ils le souhaitaient. Ils pouvaient aussi revenir sur un exercice précédent et consulter leurs programmes tels qu'ils les avaient laissés dans la Zone 2 de l'interface de programmation.

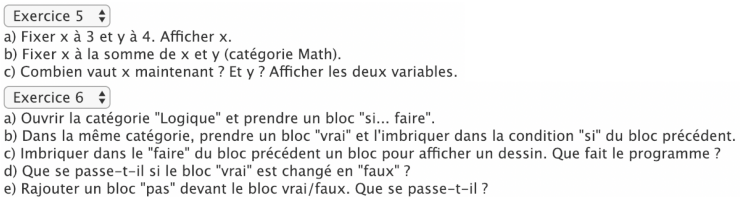


Fig. 4. Deux exemples d'exercices à réaliser.

5 Résultats

5.1 Validation du protocole expérimental

Comme précisé en Sect. 4, l'expérimentation a rassemblé 36 participants qui ont été séparés en deux groupes :

— 18 participants ont programmé la micro:bit tangible. Ce groupe était composé de 5 filles et 13 garçons avec une moyenne d'âge de 15.6 ans. Dans ce groupe, 11 participants ont déclaré n'avoir aucune connaissance en programmation et 7 ont affirmé avoir des connaissances élémentaires.

— 18 participants ont programmé la simulation de la micro:bit. Ce groupe était composé de 4 filles et 14 garçons avec une moyenne d'âge de 14.8 ans. Dans ce groupe, 12 participants ont déclaré n'avoir aucune connaissance en programmation et 6 ont affirmé avoir des connaissances élémentaires.

Nous avons d'abord cherché à confirmer l'homogénéité du niveau initial des participants entre les deux groupes. Pour cela, nous avons analysé les résultats obtenus au pré-test par les deux groupes avec un test de Student bilatéral pour deux échantillons non-appariés. Les résultats de ce test ($t = 1.499, p = 0.145 > 0.05$) indiquent que les moyennes au pré-test des deux groupes ne sont pas significativement différentes. Ce premier résultat confirme que la composition *a priori* des groupes n'a pas introduit de biais initial dans l'expérimentation.

5.2 Apprentissage de la programmation

Le Tableau 1 récapitule les résultats obtenus au pré-post test pour chaque groupe et en fonction de chaque concept évalué.

Tableau 1. Résultats au pré-post test pour chaque groupe de participants.

Groupe	Variables (5 pts)		Conditions (5 pts)		Boucles (4 pts)	
	Pré-test	Post-test	Pré-test	Post-test	Pré-test	Post-test
Tangible	$\mu = 4.167$	$\mu = 4.222$	$\mu = 1.667$	$\mu = 4.167$	$\mu = 2.389$	$\mu = 3.444$
	$\sigma^2 = 0.735$	$\sigma^2 = 0.889$	$\sigma^2 = 1.059$	$\sigma^2 = 1.441$	$\sigma^2 = 0.840$	$\sigma^2 = 0.497$
Simulation	$\mu = 3.611$	$\mu = 4.000$	$\mu = 1.389$	$\mu = 3.222$	$\mu = 2.222$	$\mu = 3.111$
	$\sigma^2 = 2.605$	$\sigma^2 = 0.941$	$\sigma^2 = 0.605$	$\sigma^2 = 1.595$	$\sigma^2 = 0.771$	$\sigma^2 = 0.928$

Apprentissage global. Nous souhaitions vérifier que le protocole expérimental conçu conduisait bien à un gain significatif d'apprentissage des concepts élémentaires de la programmation. Pour cela, nous avons calculé par deux fois la valeur de t dans le cadre d'un test de Student unilatéral pour des échantillons appariés : une fois pour le groupe programmant l'objet tangible, une seconde fois pour le groupe programmant la simulation. Concernant le groupe programmant l'objet tangible, les résultats ($t = 8.574, p \ll 0.05$) suggèrent que les notes obtenues au post-test sont significativement meilleures que celles obtenues au pré-test. Les résultats du groupe programmant la simulation ($t = 5.242, p \ll 0.05$) conduisent aux mêmes conclusions. Ce point valide la conception des exercices proposés aux participants dans le cadre d'une initiation à la programmation.

Nous avons ensuite comparé les gains d'apprentissage entre les deux groupes avec un test de Student bilatéral pour deux échantillons non-appariés. Le gain d'apprentissage d'un individu a été mesuré en soustrayant sa note du pré-test à sa note du post-test. Les résultats ($t = 0.687$ et $p = 0.497$) indiquent que les gains d'apprentissage sont équivalents entre les deux groupes.

Notion de variable. Nous avons commencé par analyser le gain d'apprentissage concernant la notion de variable pour chacun des deux groupes. Les résultats du test de Student unilatéral pour des échantillons appariés ($t = 0.325$ et $p = 0.375$ pour le groupe programmant la carte, $t = 1.046$ et $p = 0.155$ pour le groupe programmant la simulation) suggèrent qu'il n'y a pas eu d'apprentissage significatif pour aucun des deux groupes. Nous n'avons donc pas jugé utile de poursuivre l'analyse des données en comparant les gains d'apprentissage entre les deux groupes.

Notion de structure conditionnelle. Pour chaque groupe, nous avons cherché à déterminer si les notes au post-test étaient significativement meilleures que les notes au pré-test. Pour cela, nous avons conduit par deux fois un test de Student unilatéral pour des échantillons appariés. Les résultats ($t = 6.556$ et $p \ll 0.05$ pour le groupe programmant la carte, $t = 6.761$ et $p \ll 0.05$ pour le groupe programmant la simulation) suggèrent qu'il y a effectivement eu une amélioration significative des notes entre le pré-test et le post-test concernant la notion de structure conditionnelle.

Nous avons continué cette analyse en comparant les gains d'apprentissage entre les deux groupes en appliquant un test de Student pour deux échantillons non-appariés. Les résultats de ce test ($t = 1.425$ et $p = 0.163$) indiquent que les moyennes des gains d'apprentissage étaient similaires entre les deux groupes. Cela implique que, dans le cadre de notre expérimentation, la programmation d'un objet tangible ou d'une simulation n'a pas entraîné de différence significative dans l'apprentissage et la compréhension du concept de structure conditionnelle.

Notion de structure itérative. Enfin, nous avons procédé à une analyse similaire pour évaluer les apprentissages concernant la notion de structure itérative.

Les résultats des tests de Student unilatéraux pour des échantillons appariés ($t = 6.174$ et $p \ll 0.05$ pour le groupe programmant la carte, $t = 3.915$ et $p \ll 0.05$ pour le groupe programmant la simulation) suggèrent que les notes obtenues au post-test étaient significativement meilleures que celles du pré-test pour les deux groupes.

Cependant, les résultats du test de Student bilatéral pour deux échantillons non-appariés ($t = 0.586$ et $p = 0.562$) montrent encore une fois que les gains d'apprentissage ne sont pas significativement différents entre les deux groupes.

6 Discussion

Les résultats de notre étude suggèrent qu'il n'y a pas de différence significative d'apprentissage selon qu'un apprenant programme un objet tangible ou sa simulation. De plus, les résultats concernant la notion de variable n'ont pas montré de gain d'apprentissage significatif entre le pré-test et le post-test. Cela peut s'expliquer par les bons scores obtenus dès le pré-test pour les deux groupes de participants (voir Table 1). Nous y voyons deux explications possibles : les questions étaient trop simples ou alors la notion de variable est plus facile à comprendre (ce qui est aussi un résultat présenté dans d'autres études [21]).

L'expérimentation décrite dans cet article présente plusieurs points d'amélioration possibles. Premièrement, les participants n'ont pu interagir avec l'environnement de programmation qu'une demi-journée. Bien que nos résultats montrent un gain d'apprentissage significatif pour chacun des groupes, il serait intéressant d'analyser des données provenant d'une expérimentation plus longue et ancrée dans un contexte institutionnel. Un tel résultat permettrait de trancher quant à l'intérêt d'introduire des objets tangibles programmables dans les cours d'introduction à la programmation. Une telle expérimentation est d'ores et déjà prévue dans la suite de nos travaux.

Un second point d'amélioration porte sur la manipulation de l'objet programmable par les apprenants. En effet, les participants n'étaient pas incités à interagir physiquement avec la micro:bit tangible en raison de la conception des exercices de programmation : le maniement de la carte n'avait pas de répercussion sur le programme exécuté. Or, la micro:bit présente plusieurs dispositifs permettant de récupérer des informations en entrée tels que les deux boutons-poussoirs mentionnés en Sect. 3.1 ou encore un accéléromètre (qui sont d'ailleurs aussi disponibles sur la simulation). Puisque notre protocole expérimental n'incluait pas la notion de données en entrée/sortie (hormis l'affichage d'informations sur les LED), nous n'avons pas conçu d'exercice utilisant ces composants. Il serait donc pertinent de comparer nos résultats avec ceux d'une étude mettant plus l'accent sur un interaction conséquente avec l'objet programmable tangible.

nfin, nous avons fait le choix d'utiliser la BBC micro:bit pour cette expérimentation. Cette carte présente les avantages d'être peu encombrante, peu coûteuse, et disponible sous une forme numérique. Un point à explorer concerne donc l'utilisation d'un robot et de sa simulation afin de mener une étude comparative semblable à celle présentée dans cet article.

7 Conclusion et travaux futurs

Cet article présente une étude comparative des gains d'apprentissage lorsqu'un apprenant programme un objet tangible ou sa simulation. Les résultats indiquent que, bien qu'un apprentissage significatif soit observé dans les deux cas, ces situations semblent globalement équivalentes concernant l'apprentissage de la programmation, et plus spécifiquement des notions de structures conditionnelles et itératives. Ces résultats nous conduisent à investiguer plus en détail les différences entre programmation physique et numérique en concevant une nouvelle expérimentation insistant sur les manipulations de l'objet programmable.

En plus des pistes d'améliorations décrites plus haut, nous avons identifié trois axes de travaux futurs. Le premier concerne l'analyse des traces d'utilisation de l'environnement de programmation afin de déterminer les stratégies mises en place par les apprenants pour résoudre les exercices proposés. Cela nous permettrait aussi d'étudier l'évolution de la compréhension des concepts étudiés à travers les actions réalisées par chaque apprenant tout au long des exercices. Le second concerne l'identification des différences d'apprentissage lorsque la programmation physique est utilisée dans l'apprentissage des STEM (*Science, Technology, Engineering, Mathematics*). Des articles présentent déjà ce type de pratique [11,20] et il s'agit de conduire une étude comparative reprenant les caractéristiques de celle décrite dans cet article (i.e., homogénéité des connaissances initiales, moyennes d'âge similaires, répartition équilibrée entre filles et garçons). Enfin, le dernier axe de recherche porte sur l'étude comportementale des apprenants lorsqu'ils font face à un environnement de programmation impliquant un objet tangible ou une simulation. Ce dernier point permettrait de fournir des compléments de réponses quant à l'ampleur du gain de motivation ou de créativité résultant de la programmation physique et/ou numérique.

Références

1. du Boulay, B. : Some Difficulties of Learning to Program. Journal of Educational Computing Research **2**(1), 57–73 (1986)
2. Brauner, P., Leonhardt, T., Ziefle, M., Schroeder, U. : The Effect of Tangible Artifacts, Gender and Subjective Technical Competence on Teaching Programming to Seventh Graders. In : Teaching Fundamentals Concepts of Informatics. pp. 61–71 (2010)
3. Cliburn, D.C. : Experiences with the LEGO Mindstorms Throughout the Undergraduate Computer Science Curriculum. In : Proceedings. Frontiers in Education. 36th Annual Conference. pp. 1–6 (2006)
4. Cooper, S., Dann, W., Pausch, R., Pausch, R. : Teaching Objects-First in Introductory Computer Science. SIGCSE Bull. **35**(1), 191–195 (2003)
5. Fagin, B. : Ada/Mindstorms 3.0. IEEE Robotics & Automation Magazine **10**(2), 19–24 (2003)
6. Garrett, A., Thornton, D. : A Web-Based Programming Environment for LEGO Mindstorms Robots. In : Proceedings of the 43rd Annual Southeast Regional Conference - Volume 2. pp. 349–350 (2005)

7. Hodges, S., Scott, J., Sentance, S., Miller, C., Villar, N., Schwiderski-Grosche, S., Hammil, K., Johnston, S. : .NET Gadgeteer : A New Platform for K-12 Computer Science Education. In : Proceeding of the 44th ACM Technical Symposium on Computer Science Education. pp. 391–396 (2013)
8. Krathwohl, D.R. : A Revision of Bloom's Taxonomy : An Overview. Theory Into Practice **41**(4), 212–218 (2002)
9. Magnenat, S., Riedo, F., Bonani, M., Mondada, F. : A Programming Workshop Using the Robot "Thymio II" : The Effect on the Understanding by Children. In : 2012 IEEE Workshop on Advanced Robotics and its Social Impacts. pp. 24–29 (2012)
10. Major, L., Kyriacou, T., Brereton, P. : The Effectiveness of Simulated Robots for Supporting the Learning of Introductory Programming : a Multi-Case Case Study. Computer Science Education **24**(2-3), 193–228 (2014)
11. Maschietto, M., Soury-Lavergne, S. : Designing a Duo of Material and Digital Artifacts : the Pascaline and Cabri Elem e-Books in Primary School Mathematics. ZDM **45**(7), 959–971 (2013)
12. Meerbaum-Salant, O., Armoni, M., Ben-Ari, M. : Learning Computer Science Concepts with Scratch. Computer Science Education **23**(3), 239–264 (2013)
13. Papert, S., Harel, I. : Situating Constructionism. Constructionism **36**(2), 1–11 (1991)
14. Przybylla, M., Romeike, R. : Key Competences with Physical Computing. KEYCIT 2014–Key Competencies in Informatics and ICT (Preliminary Proceedings) p. 216 (2014)
15. Qian, Y., Lehman, J. : Students' Misconceptions and Other Difficulties in Introductory Programming : A Literature Review. ACM Trans. Comput. Educ. **18**(1), 1–24 (2017)
16. Schmidt, A. : Increasing Computer Literacy with the BBC micro:bit. IEEE Pervasive Computing **15**(2), 5–7 (2016)
17. Sentance, S., Schwiderski-Grosche, S. : Challenge and Creativity : Using .NET Gadgeteer in Schools. In : Proceedings of the 7th Workshop in Primary and Secondary Computing Education. pp. 90–100. ACM (2012)
18. Sentance, S., Waite, J., Hodges, S., MacLeod, E., Yeomans, L. : Creating Cool Stuff : Pupils' Experience of the BBC micro :bit. In : Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education. pp. 531–536 (2017)
19. Sentance, S., Waite, J., Yeomans, L., MacLeod, E. : Teaching with Physical Computing Devices : the BBC micro:bit Initiative. In : Proceedings of the 12th Workshop on Primary and Secondary Computing Education. pp. 87–96 (2017)
20. Wang, P., Renna, I., Amiel, F., Zhang, X. : Learning with Robots in CS and STEM Education : A Case Study with ISEP-R0B0. In : Proceedings of the 4th Workshop on Robots for Learning at ACM/IEEE HRI 2018. pp. 16–21 (2018)
21. Weintrop, D., Wilensky, U. : Comparing Block-Based and Text-Based Programming in High School Computer Science Classrooms. ACM Trans. Comput. Educ. **18**(1), 3 :1–3 :25 (2017)

Soutien à l'apprentissage de la programmation : conception et évaluation d'un indicateur sémantique

Julien Broisin[1] and Clément Hérouard[2]

[1] Université Paul Sabatier, IRIT UMR 5505, 118 route de Narbonne, F-31062 Toulouse, France julien.broisin@irit.fr
[2] ENS, Rennes, France
clement@herouard.fr

Résumé. L'accompagnement des débutants dans l'apprentissage de la programmation connaît un véritable engouement depuis plus d'une décennie. La majorité des travaux se focalise sur une évaluation syntaxique du code pour étudier le comportement des apprenants et fournir des aides appropriées. Nous proposons dans cet article la conception d'un indicateur reflétant la proximité sémantique de deux codes source distincts, dans l'objectif d'exprimer la capacité d'un apprenant à résoudre un problème donné. Cet indicateur s'inspire d'une approche dédiée à la comparaison de chaînes de caractères, et applique une approche par apprentissage automatique pour déterminer certaines caractéristiques à partir d'un jeu de données obtenu durant un cours d'introduction à la programmation Shell suivi par 166 étudiants. Les premiers résultats sont encourageants : l'indicateur permet de classer correctement un code source d'un point de vue sémantique dans 58% des cas, et il est corrélé avec les évaluations sommatives réalisées par les enseignants.

Mots-clé : Apprentissage de la programmation · Analyse sémantique · Fouille de données éducatives · Apprentissage automatique.

Abstract. Supporting beginners in learning programming is of great interest for more than a decade. Most of the work focuses on a syntactic evaluation of the code produced by learners to study their coding behaviour and provide them with appropriate support. We propose in this article the design of an indicator reflecting the semantic proximity of two distinct source codes, in order to express the learner's capacity to solve a given problem. This indicator is based on an approach dedicated to strings comparison, and applies a machine learning approach to explore certain characteristics from a dataset gathered during an introductory Shell programming course involving 166 students. The first results are encouraging: the indicator makes it possible to correctly classify source code from a semantic point of view in 58% of cases, and is correlated with the summative evaluations carried out by teachers.

Keywords: Learning programming · Semantic analysis · Educational data mining · Machine learning.

1 Introduction

L'apprentissage de l'informatique est actuellement en plein essor. En France, la réforme du baccalauréat qui sera opérationnelle dès la prochaine rentrée scolaire prévoit une option Numérique et Sciences Informatiques en classes de première et terminale pour dispenser plus de 350 heures d'informatique à un élève de lycée. En Europe, notamment en Allemagne et au Royaume-Uni, des transformations éducatives profondes sont également initiées depuis 2016 pour favoriser le développement et l'apprentissage du numérique dans les écoles et préparer les apprenants à l'acquisition des compétences du 21ème siècle. Cette intégration de l'informatique dans les cursus de formation dès le plus jeune âge requière non seulement des enseignants préparés à cette nouvelle discipline, mais également des solutions technologiques pour accompagner les instructeurs et les apprenants dans leurs pratiques quotidiennes.

Dans cet objectif, la communauté de recherche en Environnements Informatiques pour l'Apprentissage Humain (EIAH) s'intéresse à la conception de systèmes dédiés à l'accompagnement de l'apprentissage de la programmation, comme en témoignent les approches d'analyse du comportement des apprenants [1, 2]. Une des approches reconnues dans la littérature consiste à proposer des systèmes qui analysent le code produit par les apprenants d'un point de vue syntaxique. Ces systèmes visent à fournir un soutien approprié et constructif aux apprenants en réalisant une évaluation syntaxique du code pour détecter les erreurs grammaticales et/ou éviter les erreurs de misconceptions [3]. Pourtant, un des résultats d'apprentissage primordiaux de la programmation concerne bien le développement de la capacité d'un individu à résoudre un problème [4]. Une approche pour atteindre cet objectif repose sur la conception de systèmes capables d'analyser un code source du point de vue sémantique, c'est-à-dire capable de refléter comment un problème a été résolu. Bien que certaines initiatives tentent d'aborder l'analyse sémantique de codes informatiques [5], les travaux sur ce sujet sont insuffisants et peu de solutions ont été proposées. Ainsi, des études supplémentaires sont nécessaires pour mieux comprendre le potentiel de cette approche lorsqu'il s'agit d'accompagner l'apprentissage de l'informatique.

Nous proposons alors dans cet article la conception d'un indicateur reflétant la proximité sémantique de deux codes source distincts, dans l'objectif de refléter la pertinence d'une production d'un apprenant au regard d'un problème donné. Les questions de recherche traitées dans l'étude sont donc les suivantes :
Question de recherche 1 : Comment concevoir un indicateur sémantique reflétant la capacité d'un apprenant à résoudre un problème ?
Question de recherche 2 : La qualité sémantique d'une production d'un apprenant est-elle corrélée avec sa performance académique ?

Pour répondre à la première question, nous adoptons une approche par comparaison d'arbres de syntaxe abstraite. Nous nous appuyons sur la distance d'édition formulée par Levenshtein, un algorithme qui a démontré son efficacité pour comparer des chaînes de caractères, et proposons une approche par apprentissage automatique pour déterminer certains facteurs nécessaires au calcul de la proximité sémantique. En ce qui concerne la seconde question de recherche,

nous menons un ensemble d'études statistiques sur des données collectées dans un contexte écologique d'apprentissage afin d'évaluer la précision de l'indicateur au regard de la perception humaine. En particulier, à partir de codes source produits par 166 étudiants, nous analysons la corrélation entre la valeur de notre indicateur calculée automatiquement pour l'ensemble des scripts d'une part, et les scores assignés manuellement par des enseignants d'autre part.

2 Approches d'analyse automatique de code source

Dans le contexte de l'apprentissage de la programmation, l'analyse automatique de code source est utilisée pour atteindre différents objectifs tels que l'amélioration des feedbacks fournis aux apprenants [6, 7], la prédiction de leur performance [8, 9], ou leur suivi pendant la tâche d'apprentissage [10].

Ces analyses automatiques sont réalisées par compilation du code [11], par recherche d'erreurs classiques au sein d'un code source [7], ou par l'écriture de test unitaires par les enseignants [12]. Ces différents travaux guident les apprenants dans l'écriture de programmes syntaxiquement corrects, mais ne permettent pas d'évaluer le code source à un niveau sémantique : l'évaluation syntaxique n'est pas suffisante pour rendre compte de la pertinence d'une production d'un apprenant vis-à-vis d'un problème posé par l'enseignant.

Dans la méta-review proposée par Ihantola et al. [13] qui ont étudié pas moins de 118 articles de recherche dans le domaine de la fouille données éducatives pour la programmation, le mot *semantics* est absent du papier. De même, une recherche des termes *semantic analysis programming* dans Google Scholar et Web of Science retourne bien un nombre important de résultats, mais aucun article ne traite réellement de l'analyse sémantique de code.

Les travaux que nous avons recensés qui se rapprochent d'une analyse sémantique sont ceux proposés par Bey et al. [5]. Afin de mettre en œuvre une notation automatisée du code des apprenants dans le cas des MOOC, les auteurs proposent de comparer le graphe de flot de contrôle correspondant au code produit par l'apprenant, avec un ensemble de graphes issus de solutions expertes qui résolvent le problème donné [14]. Si le graphe de l'apprenant est reconnu parmi la base de graphes experts, alors un score et un feedback sont retournés à l'apprenant ; s'il n'est pas reconnu, alors la production de l'apprenant est évaluée manuellement par un expert humain afin d'enrichir la base de graphes.

L'absence de travaux concernant l'analyse automatique d'un point de vue sémantique peut s'expliquer par le fait que la sémantique d'un programme n'est pas calculable. Nos travaux tentent de dépasser cette limite en proposant la conception d'un indicateur reflétant la pertinence d'un code source par rapport à un problème donné.

3 Spécification de l'indicateur sémantique

Notre objectif est de concevoir un indicateur capable d'évaluer la distance d'édition entre deux codes source du point de vue sémantique, et de satisfaire les con-

traintes suivantes : la tâche des enseignants doit se limiter à la production d'une solution unique aux problèmes posés, au contraire des méthodes par tests unitaires qui requièrent de leur part un effort important ; la distance doit diminuer au fur et à mesure que l'étudiant se rapproche de la solution, puisqu'un script incomplet peut être évalué même s'il ne répond pas entièrement au problème posé. Ainsi, si une solution à un problème est fournie par l'enseignant, la distance entre le code source correspondant et les productions des apprenants devrait fournir un nouvel éclairage sur leur capacité à résoudre le problème.

Le domaine du traitement automatique des langues a depuis longtemps étudié des problématiques liées à la correction orthographique où la sémantique est naturellement prise en compte [15]. Nous nous sommes inspirés de ces approches pour proposer la conception de notre indicateur, et en particulier de la distance d'édition entre deux chaînes de caractères, ou distance de Levenshtein [16]. Cette distance est définie par un coût calculé à partir du nombre minimum d'opérations (i.e., insertion/suppression d'un caractère, et substitution d'un caractère par un autre) nécessaires pour passer d'une chaîne à l'autre. Dans notre contexte, nous souhaitons identifier les opérations à effectuer pour passer d'un code source à un autre, et ainsi calculer une distance correspondant à la somme des coûts de chacune des opérations à réaliser. Dans le cas de la comparaison de chaînes de caractères, les coûts associés à chacune de ces opérations sont tous fixés à 1, à l'exception du coût de substitution d'un caractère par lui-même qui est nul. Dans le cas de la comparaison de codes source, les coûts de ces opérations ne sont pas tous égaux : la substitution de deux instructions visant les mêmes objectifs ne doit pas avoir le même coût que la substitution d'une commande par une autre ayant une fonctionnalité très différente. Le challenge de notre approche consiste donc à attribuer un coût à chacune des opérations selon l'importance des modifications à réaliser pour passer d'un code source à un autre.

La méthodologie mise en œuvre pour concevoir notre indicateur est alors la suivante : formalisation du code source sous la forme d'un arbre de syntaxe abstraite ; transformation de cet arbre en une chaîne d'instructions élémentaires ; identification des coûts d'insertion, de suppression et de substitution des instructions élémentaires ; spécification d'une méthode d'apprentissage automatique pour déterminer les valeurs de ces différents coûts ; évaluation des résultats.

3.1 Formalisation de code source

Nous avons adopté les arbres de syntaxe abstraite (Abstract Syntaxic Tree, AST) afin de proposer une description formelle des codes source. En effet, au contraire des arbres d'analyse, les arbres de syntaxe abstraite ne représentent pas les nœuds et les branches qui n'affectent pas la sémantique d'un programme.

Nous nous appuyons sur un parseur écrit en JavaScript qui lit un script et retourne l'arbre de syntaxe abstraite correspondant au format JSON ; un exemple de conversion d'un code source Bash en AST est donné par la Figure 1.

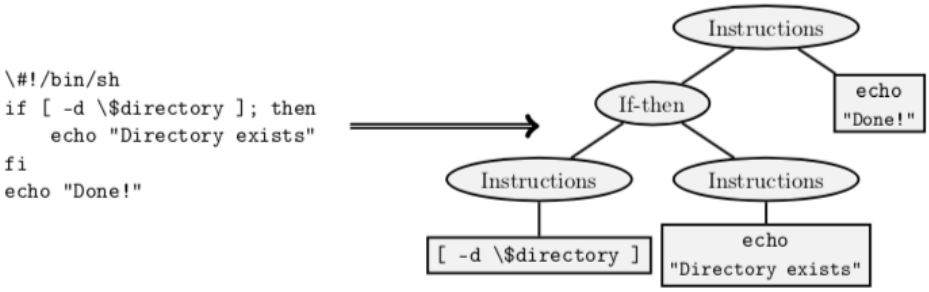


Fig. 1. Conversion d'un script en arbre de syntaxe abstraite.

3.2 Mise en œuvre de la distance d'édition

Ainsi, avec la représentation formelle proposée ci-dessus, la mesure de notre indicateur correspond à la distance d'édition entre deux arbres de syntaxe abstraite. Toutefois, afin d'adapter la distance de Levenshtein à notre contexte, nous souhaitons convertir l'AST généré à partir du code source en une chaîne de caractères ; du point de vue de la terminologie, nous nommons *tokens* les caractères résultant du processus de transformation de l'AST afin de ne pas les confondre avec ceux contenus dans le code source. Ainsi, nous allons effectuer un parcours en profondeur de l'AST pour créer la chaîne de tokens.

Production de la chaîne de tokens. Le langage Bash, objet de notre étude, propose 17 structures de contrôle correspondant chacune à un token différent. Un token peut par exemple être du type Command, du type Assignation pour les définitions de variables, ou du type While, Do ou Done pour signaler respectivement le début, le corps ou la fin de boucle. Notons que les tokens Command sont caractérisés par le nom de la commande et ses arguments. Un exemple de transformation d'un AST en chaîne de tokens est illustré par la Figure 2.

La distance entre deux chaînes de tokens est donc caractérisée par un certain nombre de coûts, ou paramètres, correspondant aux différentes opérations élémentaires d'insertion, de suppression et de substitution de chaque token. Les 17 tokens du langage Bash conduisent à la définition de 308 coûts : 17+17 coûts de création et de suppression d'un token, 16x17 coûts de substitution d'un token par un autre, 1 coût pour substituer un token Command en un autre dont le nom de la commande est différent, et enfin 1 coût pour substituer un token Command en un autre dont les noms de commande sont identiques et les arguments différents. Ce nombre de paramètres est trop élevé pour mettre en œuvre une méthode par apprentissage efficace. En effet, un nombre trop important de paramètres risque de diminuer la densité des tests et ainsi rendre la recherche de "bonnes" valeurs pour les paramètres plus difficile.

Réduction du nombre de paramètres. Une première opération pour réduire le nombre de paramètres à entraîner consiste à ignorer les tokens Until, Subshell

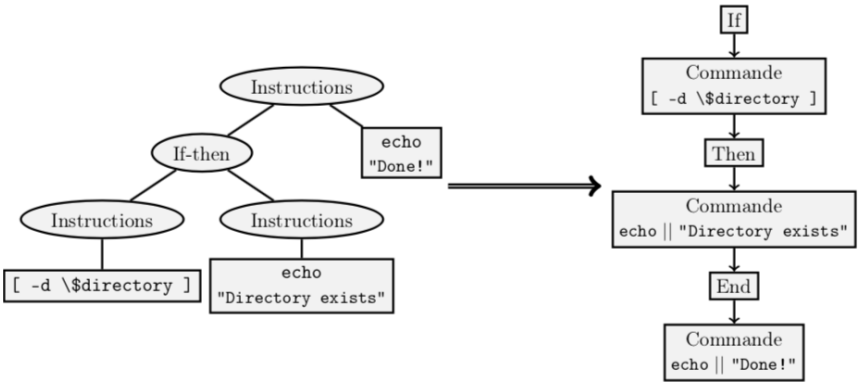


Fig. 2. Conversion d'un arbre de syntaxe abstraite en chaîne de tokens.

et Pipeline car les structures de contrôle correspondantes n'apparaissent pas dans le jeu de données de cette étude (cf. Section 4). Nous proposons également une symétrisation du problème en affectant des coûts égaux aux opérations d'insertion et de suppression d'un token ; de même, la substitution du token A par le token B a un coût égal à la substitution de B par A. Cette symétrisation rend la distance entre les scripts S_1 et S_2 égale à la distance entre S_2 et S_1, et permet de réduire le nombre de paramètres à 107. Enfin, nous posons une hypothèse plus lourde en considérant que seuls certains tokens peuvent se substituer à d'autres. Ainsi, seuls les groupes de tokens suivants peuvent se substituer entre eux : {Command, Assignation}, {If, Case}, {Then, Else, CaseItem}, et {For, While}. En effet, il est envisageable qu'un script contienne un Case là où un autre script emploie un If, par contre il est peu probable qu'une substitution d'un For par un If soit fréquente, ces deux instructions visant des objectifs très différents.

Ces simplifications diminuent considérablement le nombre de coûts puisque 23 paramètres doivent désormais être calculés. Notons θ le vecteur contenant ces 23 coûts. Notre indicateur, noté $d(S, C, \theta)$, est alors défini par la distance d'édition, sous les 23 paramètres θ, entre la chaîne de tokens d'un script S et celle d'un script C. Afin d'être en mesure de spécifier le critère d'apprentissage conduisant aux valeurs optimales de θ, nous présentons le jeu de données issu d'une situation d'apprentissage réelle qui a été utilisé pour réaliser cette étude.

4 Jeu de données

Le jeu de données fut obtenu dans un département Informatique d'un Institut Universitaire Technologique (IUT) durant un cours d'introduction à la programmation Shell suivi par 166 apprenants. Les étudiants découvrent des commandes courantes telles que echo, ls ou cat, la gestion des variables, ainsi que les structures du langage (e.g., for, while, if, case). Les étudiants mettent ensuite en

pratique ces notions lors de séances de travaux pratiques durant lesquelles ils produisent des scripts Bash pour tenter de solutionner une série de problèmes.

Le jeu de données comprend les productions des étudiants durant cinq séances de travaux pratiques d'une heure et demie chacune. A chaque enregistrement d'un script Bash par un étudiant, une copie de ce dernier accompagnée de son horodatage et de l'identifiant de l'étudiant est sauvegardée dans un répertoire afin d'enrichir le jeu de données. Ainsi, notre échantillon comprend 19232 scripts. Toutefois, des erreurs de nomenclature ou de numérotation des fichiers ont rendu 18% de ces scripts inexploitables : le jeu de données est constitué de 15794 scripts.

Cet échantillon comprend, pour un même exercice, plusieurs scripts produits par un même élève. Or les méthodes d'apprentissage sont efficaces lorsqu'elles s'appuient sur des données indépendantes les unes des autres. Alors nous ne considérons, par exercice et par étudiant, que le dernier script produit par l'apprenant car nous supposons qu'il est le plus abouti. Ce choix permet d'obtenir un jeu de données indépendantes, mais réduit l'échantillon à 2540 scripts.

Enfin, nous souhaitons classifier manuellement chaque script de l'échantillon dans deux catégories distinctes (i.e., *Correct* et *Incorrect*) exprimant leur justesse sémantique. Certains exercices dispensés lors de cette expérimentation ne permettent pas de concevoir une solution répondant totalement au problème posé, car certains problèmes concernent par exemple la compréhension de l'exécution d'un script fourni dans l'énoncé. Après cette opération de classification, l'échantillon final comprend 733 scripts répartis sur 11 exercices différents ; 397 scripts sont sémantiquement corrects, les autres étant sémantiquement incorrects.

Ce jeu de données a été divisé aléatoirement en un jeu d'apprentissage dont l'objectif est d'identifier les valeurs optimales du paramètre θ, et un jeu de test permettant d'évaluer la pertinence des résultats obtenus lors de la phase d'apprentissage. Ces deux échantillons contiennent le même nombre de scripts pour chaque exercice, et la proportion de scripts sémantiquement corrects et incorrects est préservée. De plus, notons qu'une solution a été fournie par un enseignant pour chacun des exercices.

5 Méthode d'apprentissage et résultats

Pour entraîner les bonnes valeurs des 23 coûts du paramètre θ, nous définissons le critère d'apprentissage $Score(\theta)$ à minimiser tel que :

$$Score(\theta) = \frac{\sum_{(S,C)\in Correct} \sqrt{d(S,C,\theta)}}{\sum_{(S,C)\in Incorrect} \sqrt{d(S,C,\theta)}} \tag{1}$$

où $Correct$ (resp. $Incorrect$) est l'ensemble des paires composées des scripts corrects (resp. incorrects) du jeu d'apprentissage et des corrections associées.

La fonction $Score$ est faible quand les scripts corrects sont associés à de faibles distances, et les scripts incorrects à de fortes distances. Nous sommons les racines carrées des distances pour réduire l'influence des fortes valeurs.

Nous pouvons remarquer une propriété de la fonction $d : \forall\lambda > 0, d(S, C, \lambda.\theta) = \lambda.d(S, C, \theta)$. Alors il est évident que $Score(\lambda\theta) = Score(\theta)$. Cela signifie que la fonction $Score$ est constante sur toutes les demies-droites partant de 0, et que l'exploration des différents coûts du paramètre θ, notés θ_i tels que $0.01 < \theta_i \leq 1$, est suffisante pour trouver les valeurs de θ_i minimisant $Score$.

Le jeu de données d'apprentissage a servi à initialiser l'algorithme de minimisation de descente de gradient. Le vecteur minimisant $Score$, noté $\overline{\theta}$, admet des valeurs de 0.01 (i.e., le minimum que nous avons fixé) presque partout sauf pour trois paramètres :

1. la création ou suppression d'un token Command, dont le coût est de 1 ;
2. la création ou suppression d'un token If, dont le coût est de 0.112 ;
3. la création ou suppression d'un token Function, dont le coût est de 0.022.

Nous disposons, à cette étape, du paramètre $\overline{\theta}$ qui renferme les valeurs des 23 coûts qui ont été obtenues par notre méthode d'apprentissage. Les tests menés dans la section suivante avec le paramètre $\overline{\theta}$ sur le jeu de données de test visent un double objectif : étudier la capacité de notre indicateur à différencier un script sémantiquement correct d'un script incorrect ; étudier la corrélation entre notre indicateur et des notes attribuées manuellement aux scripts par des enseignants.

6 Validation

6.1 Étude de la prédiction

Une première approche pour évaluer la pertinence de notre indicateur sous le paramètre $\overline{\theta}$ consiste à évaluer sa capacité à classer automatiquement un script comme étant correct ou incorrect d'un point de vue sémantique. Dans une première étape, nous calculons la distance $d(S, C, \overline{\theta})$ entre chaque script S du jeu de données de test et la correction C du problème correspondant, afin d'obtenir une valeur de notre indicateur pour chaque script correct et incorrect.

L'étape suivante consiste à évaluer le modèle de prédiction résultant des différentes valeurs de l'indicateur. Nous utilisons pour cela la mesure ROC qui est très répandue dans le domaine des statistiques pour observer la performance d'un classifieur binaire. Lorsque le classifieur calcule une mesure m qui est comparée à un seuil σ pour prédire la classe, l'idée de la courbe ROC est de faire varier σ de 1 à 0 et, pour chaque valeur de σ, reporter dans un graphique le taux de faux positifs en abscisse, et le taux de vrais positifs en ordonnée. L'aire sous la courbe (AUC) indique alors la probabilité pour que le classifieur place un positif devant un négatif. Un classifieur qui ne se trompe jamais présente une AUC égale à 1, alors qu'un classifieur aléatoire est caractérisé par une AUC égale à 0.5.

Dans notre étude, si $d(S, C, \overline{\theta}) < \sigma$, alors le script est classé dans la catégorie *Correct* ; sinon le script est classé dans la catégorie *Incorrect*. La courbe ROC de la Figure 3 illustre le ratio de vrais positifs (i.e., fraction des scripts corrects qui sont effectivement détectés par le classifieur) en fonction du ratio de faux positifs (i.e., fraction des scripts incorrects qui sont mal classés par le classifieur). La

courbe ROC est sensiblement au-dessus de la bissectrice, l'aire sous cette courbe est égale à 0.585. Notre classifieur est donc légèrement meilleur qu'un classifieur aléatoire, mais ses performances ne sont pas très élevées ; nous émettons des hypothèses pour expliquer ces résultats dans la Section 7.

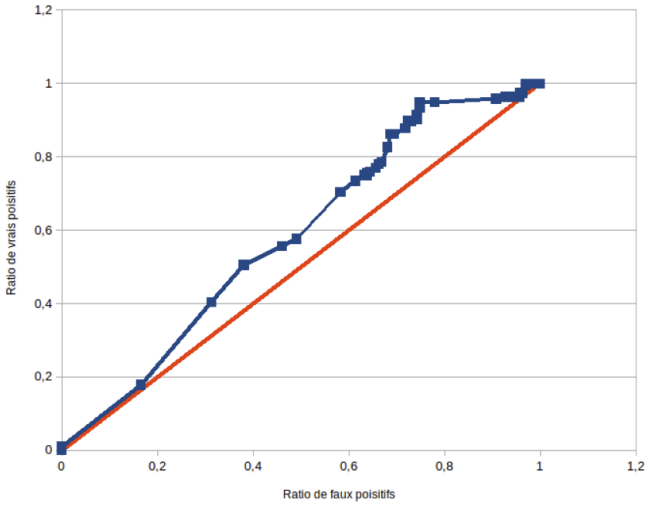


Fig. 3. Courbe ROC de notre classifieur (en bleu) et bissectrice du hasard (en rouge).

6.2 Étude de la corrélation avec une notation humaine

L'un des objectifs de notre indicateur est de refléter la progression de l'apprenant vers la solution au problème, c'est-à-dire que sa valeur est supposée diminuer au fur et à mesure qu'un script tend vers la solution. Nous étudions alors la corrélation entre la valeur de notre indicateur sous le paramètre $\overline{\theta}$ pour un script donné et la correction associée, et la note attribuée à ce script par un humain.

L'étude de cette corrélation est réalisée à l'aide d'un second jeu de données obtenu après l'examen terminal du cours décrit dans la Section 4. Parmi les 166 étudiants inscrits à ce cours, 163 ont participé à l'examen terminal qui consistait à produire un script répondant à un problème. Un enseignant a ensuite attribué un score sur 10 à chacun des 163 scripts. Après avoir éliminé les scripts grammaticalement incorrects, 105 scripts constituent ce second jeu de données.

La Figure 4 donne, pour chaque script, la valeur de l'indicateur en fonction de la note attribuée par l'enseignant. Ce graphe montre une corrélation négative : la corrélation de Kendall (non-linéaire) a une valeur égale à -0.319 (valeur p inférieure à $2.7 \cdot 10^{-6}$), la corrélation de Pearson (linéaire) ayant une valeur de -0.446 (valeur p égale à $1.9 \cdot 10^{-6}$). Nous avons donc une corrélation entre la note donnée par un humain et la valeur de notre indicateur. Cette corrélation n'est pas

extrêmement élevée, mais elle peut être suffisante pour estimer automatiquement la progression d'un étudiant vers la résolution d'un problème.

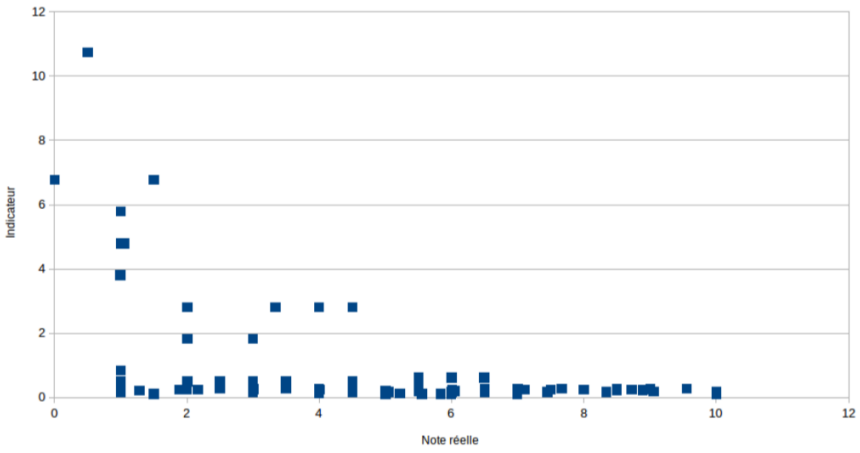


Fig. 4. Valeur de l'indicateur en fonction de la note attribuée par un enseignant.

7 Discussion

Notre indicateur permet de différencier un script sémantiquement correct d'un script incorrect, mais ses performances sont sensiblement supérieures à celles d'un classifieur aléatoire seulement. Une hypothèse pour expliquer ces performances émane du caractère hétérogène du jeux de données d'apprentissage, notamment en termes de taille des scripts qu'il contient. En effet, la distance calculée par notre indicateur tend à augmenter à chaque token du script. Donc plus un script est long, plus l'indicateur risque de retourner une valeur élevée (même si le script s'approche de la solution), alors qu'un script court et incorrect est évalué avec une faible valeur en raison du faible nombre de tokens.

L'étude de la corrélation réalisée sur le second jeu de données donne des indications sur une certaine qualité de généralisation de notre approche. En effet, contrairement au jeu de test issu du premier jeu de données, le jeu de scripts obtenu suite à l'examen terminal correspond à un exercice absent du jeu de données d'apprentissage. Notre méthode semble donc bien se généraliser à des exercices qui n'ont pas été rencontrés durant la phase d'apprentissage. Toutefois, cette corrélation n'est pas extrêmement forte et des scripts très mal évalués humainement obtiennent une valeur de notre indicateur quasiment nulle. Cela peut aussi s'expliquer par l'hypothèse formulée ci-dessus, puisque des scripts presque vides et donc incorrects sont mal notés par l'enseignant humain, alors que notre indicateur retourne une faible valeur. Des tests doivent être menés

pour étudier comment ajuster la méthode de calcul de l'indicateur, notamment en fonction du nombre de tokens du script représentant la solution au problème.

Des développements ont été initiés pour retourner cet indicateur aux apprenants pendant l'activité de programmation. Une première visualisation restitue, sous la forme d'une jauge, la valeur de l'indicateur à chaque tentative d'exécution du script. Une seconde visualisation expose la progression de l'apprenant dans la résolution du problème à partir des valeurs successives de l'indicateur pour un même script. Elle donne la variation de l'indicateur sous la forme d'un dégradé de couleurs pour fournir à l'étudiant une vue instantanée et facile à interpréter de l'évolution de la justesse de son script d'un point de vue sémantique. Ces outils destinés à supporter les apprenants seront expérimentés dès la prochaine année universitaire avec une nouvelle promotion d'étudiants d'IUT Informatique.

8 Conclusion

Nous avons cherché dans ce papier à concevoir un indicateur dont l'objectif est de servir de fondation à des systèmes de guidage automatique et à des outils d'awareness destinés à la fois aux apprenants et aux enseignants pendant la réalisation d'une activité pratique dédiée à l'apprentissage de la programmation. Si de nombreux travaux s'intéressent à la qualité syntaxique du code produit par les apprenants pour soutenir ces activités d'apprentissage, notre originalité consiste à étudier la qualité sémantique du code, c'est-à-dire à évaluer sa proximité avec la solution à un problème posé.

Ainsi, notre contribution principale concerne la conception d'un indicateur reflétant la capacité d'un apprenant à résoudre un problème. Pour répondre à la première question de recherche posée en introduction, nous avons adapté la distance de Levenshtein : à partir de deux chaînes de *tokens* représentant l'arbre de syntaxe abstrait des programmes correspondant, l'indicateur retourne une estimation de la distance d'édition séparant les deux scripts. Pour entraîner et tester cet indicateur, nous avons utilisé un jeu de données comprenant des scripts produits par des étudiants en situation d'apprentissage réelle. Les tests ont révélé que lorsqu'il s'agit de différencier des scripts sémantiquement corrects et incorrects, notre indicateur présente des performances légèrement supérieures à celles d'un classifieur aléatoire. D'autre part, nous avons observé une corrélation inverse entre la valeur de l'indicateur et le score attribué par une évaluation humaine : plus la notation humaine d'un programme est élevée, plus la distance entre ce programme et la solution du problème posé est faible.

Ces premiers résultats sont encourageants et laissent entrevoir l'opportunité de développer de nouveaux modèles et outils dédiés à l'analyse de l'apprentissage de la programmation. Toutefois, de nombreuses améliorations doivent être apportées et plusieurs pistes de recherche doivent être explorées pour améliorer la qualité de notre indicateur. En plus de données additionnelles nécessaires au raffinement des divers paramètres de notre indicateur, la généralisation de notre approche à différents langages de programmation doit être vérifiée, tout comme la prise en compte de programmes plus complexes.

References

1. Spacco, J., Denny, P., Richards, B., Babcock, D., Hovemeyer, D., Moscola, J., Duvall, R.: Analyzing student work patterns using programming exercise data. In: Proceedings of the 46th Technical Symposium on Computer Science Education, pp. 18–23. ACM (2015)
2. Sharma, K., Jermann, P., Dillenbourg, P.: Identifying Styles and Paths toward Success in MOOCs. In: Proceedings of the 8th International Conference on Educational Data Mining, pp. 408–411.
3. Taherkhani, A., Malmi, L.: Beacon-and schema-based method for recognizing algorithms from students' source code. Journal of Educational Data Mining, **5**(2), 69–101 (2013)
4. Saeli, M., Perrenet, J., Jochems, W. M., Zwaneveld, B.: Teaching programming in Secondary school: A pedagogical content knowledge perspective. Informatics in Education, **10**(1), 73–88 (2011)
5. Bey, A., Jermann, P., Dillenbourg, P.: A Comparison between Two Automatic Assessment Approaches for Programming: An Empirical Study on MOOCs. Educational Technology & Society, **21**(2), 259–272 (2018)
6. Denny, P., Luxton-Reilly, A., Carpenter, D.: Enhancing syntax error messages appears ineffectual. In: Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education, pp. 273–278. ACM (2014)
7. Karam, M., Awa, M., Carbone, A., Dargham, J.: Assisting Students with Typical Programming Errors During a Coding Session. In: Proceedings of the 7th International Conference on Information Technology, pp. 42–47. IEEE (2010)
8. Watson, C., Li, F. W., Godwin, J. L.: Predicting performance in an introductory programming course by logging and analyzing student programming behavior. In: Proceedings of the 13th International Conference on Advanced Learning Technologies, pp. 319–323. IEEE (2013)
9. Ahadi, A., Lister, R., Lal, S., Leinonen, J., Hellas, A.: Performance and Consistency in Learning to Program. In: Proceedings of the 19th Australasian Computing Education Conference, pp. 11–16. ACM (2017)
10. Alammary, A., Carbone, A., Sheard, J.: Implementation of a smart lab for teachers of novice programmers. In: Proceedings of the 14th Australasian Computing Education Conference, pp. 121–130. ACM (2012)
11. Carter, E., Blank, G. D.: A tutoring system for debugging: status report. Journal of Computing Sciences in Colleges **28**(3), 46–52 (2013)
12. Sharma, K., Mangaroska, K., Trætteberg, H., Lee-Cultura, S., Giannakos, M.: Evidence for Programming Strategies in University Coding Exercises. In: Proceedings of the 13th European Conference on Technology Enhanced Learning, pp. 326–339. Springer (2018)
13. Ihantola, P., Vihavainen, A., Ahadi, A., Butler, M., Börstler, J., et al.: Educational data mining and learning analytics in programming: Literature review and case studies. In: 2015 ITiCSE on Working Group Reports, pp. 41–63. ACM (2015).
14. Aiouni, R., Bey, A., Bensebaa, T.: An Automated Assessment Tool of Flowchart Programs in Introductory Programming Course using Graph Matching. Journal of e-Learning and Knowledge Society, **12**(2), 141–150 (2016)
15. Beijering, K., Gooskens, C., Heeringa, W.: Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm. Linguistics in the Netherlands, **15**, 13–24 (2008)
16. Levenshtein, V., I.: Binary codes capables of correcting deletions, insertions, and reversals. Soviet Physics Doklady, **10**(8), 707–710 (1966)

Session
Ingénierie des EIAH

Etudier des objets intermédiaires pour analyser les moments-clés de la conception d'un EIAH

Matthieu Cisel

Université Paris-Descartes, Laboratoire EDA, 45 rue des Saint-Pères,
75006 Paris, France
matthieucisel@gmail.com

Résumé. Cet article porte sur le processus de conception d'une application à destination du programme Savanturiers, au sein d'un consortium regroupant recherche académique, acteurs de l'éducation et une entreprise privée. Nous nous focalisons sur le processus d'arbitrage dans les orientations technologiques. Le cahier des charges n'est pas fixé au début du projet, la seule contrainte imposée étant que l'EIAH doit instrumenter une pédagogie par projet dans le domaine scientifique. Une tentative de consensus sur l'orientation technologique à impulser a été réalisée via un vote sur les différentes fonctionnalités possibles, vote fondé sur une « matrice de décision ». Nous mobilisons la notion d'objet intermédiaire et une analyse mathématique de la matrice sur laquelle ce vote repose pour montrer que cette approche ne permet pas de satisfaire l'objectif visé par la démarche : répondre au mieux aux besoins des utilisateurs finaux, les enseignants. Nous extrapolons plus largement sur l'intérêt des objets intermédiaires pour l'analyse d'un processus de conception.

Mots-clés. Conception, objet intermédiaire, analyse des besoins

Abstract. We analyze in this article the design of a learning environment intended for the Savanturiers program. It was produced by a consortium made of a laboratory, a group of teachers and a private company, with no clear specifications when the project was funded. The only constraint was that it was supposed to be used in inquiry learning. A consensus on the orientation to impulse to the project was searched through a vote at the beginning of the project. We analyze the tool that was used to perform that vote through the notion of intermediary object. We show, mathematically, that the process failed to reach its intended objective: be as close as possible to the needs of the users, aka, teacher. We extrapolate over the interest of intermediary objects to grasp the underlying dynamics of a design process.

Keywords. Design, intermediary object, needs analysis

1 Introduction

1.1 Un consortium visant à développer un EIAH pour les projets Savanturiers

Via des appels à projets comme le Plan d'Investissements d'Avenir (PIA) eFRAN, (pour Espaces de Formation, de Recherche et d'Animation Numérique), l'Etat

encourage la co-conception d'EIAH au sein de consortiums regroupant recherche académique, acteurs publics de l'éducation et entreprises privées. Avec la montée en puissance de ce type de consortiums hétéroclites, il devient de plus en plus pressant de mieux appréhender les questions que soulèvent ces collaborations, entre des acteurs appartenant à des mondes sociaux aussi distincts qu'une entreprise, des enseignants, et un laboratoire de recherche.

Si la question des dynamiques sous-jacentes au processus de conception n'est pas nouvelle dans la littérature consacrée aux EIAH, elle porte généralement sur des artefacts produits par des laboratoires. Aux exemples français comme Cabri-Géomètre [14] ou Pépite [11] s'ajoutent, dans la littérature anglophone, des artefacts comme le Knowledge Integration Environment [16], conçus pour instrumenter l'enseignement de la démarche d'investigation [26]. Pour nombre des EIAH issus de la recherche et dont la conception a été documentée, l'artefact a souvent vocation à mettre à l'épreuve des théories, de tester des hypothèses [23]. Ce n'est généralement pas le cas pour des consortiums où l'artefact devient la propriété d'une entreprise, qui tâchera de le vendre sur le marché éducatif. Dans cette dernière configuration, les éléments qui guident le processus de conception peuvent être plus pragmatiques. Il peut être intéressant pour la recherche, en particulier lorsqu'elle se situe dans une perspective de recherche d'accompagnement, d'identifier comment s'articulent les influences des différents acteurs, ou en d'autres termes, comment se manifeste la multivocalité du processus.

À cette fin, nous nous proposons, sur la base du cas d'étude qu'est le projet eFRAN « Les Savanturiers du Numérique », de nous pencher sur la question des dynamiques de co-conception d'un environnement dédié à l'instrumentation de la pédagogie par projet dans l'enseignement scientifique, le Cahier-Numérique de l'Elève-Chercheur (CNEC). Le consortium a été réuni de 2016 à 2019 pour fournir une application numérique aux projets Savanturiers, fondés notamment sur un partenariat [10] avec des chercheurs pour l'encadrement des projets. Ceux-ci mettent en œuvre une approche qu'ils nomment l'Education par la Recherche, relativement proche de la démarche d'investigation [18], au sein de l'école primaire et du collège. En termes de caractéristiques du CNEC, le consortium a reçu pour seule spécification la nécessité d'instrumenter les projets Savanturiers.

1.2 Les problématiques de recherche et leur instanciation à un cas d'étude

Notre analyse du processus de conception s'inscrit dans deux problématiques distinctes. L'une relève de la montée en puissance des consortiums pour la création d'EIAH : Lors de processus de conception, comment la multivocalité consubstantielle des consortiums se manifeste-t-elle dans les caractéristiques des artefacts produits ?

Nous analysons les difficultés du processus de conception au prisme de la notion d'objet intermédiaire [12,13,25]. Nous cherchons en particulier à montrer comment ces artefacts circulant entre les acteurs permettent de réifier divers aspects d'un processus de conception. Ce faisant, la visée de nos recherches est également, dans une large mesure, méthodologique. Elles visent en effet à illustrer une approche relativement marginale de l'analyse des objets intermédiaires, en fondant l'analyse sur leurs caractéristiques propres plus que sur les interactions qui les entourent. Cette considération nous a amené à développer dans notre recherche une seconde

problématique d'ordre méthodologique, ancrée dans les analyses de contradictions relevant de la théorie de l'activité [6] : Dans quelle mesure l'analyse des caractéristiques des objets intermédiaires peut-elle révéler les contradictions internes aux processus de conception ?

Si ces problématiques ont guidé notre travail de recherche pendant les trois années qu'a duré le consortium, nous nous focalisons uniquement dans le cadre de cette contribution sur un moment clé du projet eFRAN, le choix des orientations technologiques. Il s'est en partie basé sur un vote initialement considéré par le consortium comme déterminant dans les orientations technologiques du CNEC. La question de recherche qui a guidé cet article, dérivée des problématiques exposées dans les paragraphes précédents, peut être formulée en ces termes : Dans quelle mesure les caractéristiques d'un objet intermédiaire utilisé pour créer consensus lors d'un moment pivot, en l'occurrence un arbitrage-clé, peut-elle révéler les contradictions à l'œuvre du processus de conception ?

Nous nous proposons de porter l'attention sur un arbitrage, réalisé sous la forme d'un vote, en début de projet quant aux orientations technologiques à suivre, sur la base de ce qui fut nommé la matrice de décision. Ce vote visait à prioriser le développement des fonctionnalités du CNEC. Il a associé l'ensemble des acteurs, y compris des enseignants, afin « d'être au plus près des besoins des utilisateurs ». Nous cherchons à appréhender comment la matrice de décision révèle la manière dont le point de vue des enseignants a été pris en compte dans le choix des grandes orientations technologiques. Bien que centrée sur une étude de cas, cette contribution ne constitue pas une évaluation du processus de conception dans le projet eFRAN qui a servi de base à l'étude.

Si l'étude vise à illustrer des contradictions apparentes dans un consortium produisant un EIAH, le propos est avant tout méthodologique. Plus particulièrement, nous étayons l'idée selon laquelle, par la réification de la multivocalité qui caractérise une co-conception, les objets intermédiaires offrent au chercheur une perspective incontournable sur les dynamiques à l'œuvre. Nous cherchons ce faisant à diffuser auprès de la communauté EIAH une notion à la fécondité éprouvée, tout en la mobilisant selon une perspective originale. Nous visons également à étayer l'idée selon laquelle une réflexion sur un processus de conception ne constitue pas un simple « retour d'expérience », et que le dit processus de conception peut constituer un objet scientifique à part entière. Pour ce faire, la notion d'objet intermédiaire peut être utile à la théorisation des analyses effectuées.

1.3 Le cadre notionnel de l'objet intermédiaire

La notion d'objet intermédiaire est apparue dans les années 1990 [13] pour mieux appréhender les dynamiques sous-jacentes au processus de conception d'artefact. Elle fait écho, sans en être une simple prolongation, à des travaux anglo-saxons plus anciens sur les objets-frontière [22], plus centrés sur des objets qui partagent des codes communs aux différents mondes sociaux qu'ils traversent. Nous reprendrons ici la définition d'objet intermédiaire que donne Decortis [5], relativement proche de celle de Vinck [25] : « un artefact construit par les acteurs du processus de conception, fortement investi par ceux-ci et qui circule entre eux. L'objet intermédiaire représente, traduit des idées et sert de médiateur et d'outil d'échange ».

Il est fréquent, dans les travaux consacrés aux objets intermédiaires, de se concentrer non pas sur les interactions qui entourent ces objets, et en particulier sur les interactions orales entre les acteurs qui les mobilisent [8]. Leurs caractéristiques peuvent pourtant constituer la focale des recherches, tant elles révèlent des aspects saillants du processus de conception, sans interférer par trop avec le processus. Enregistrer et analyser des interactions orales constituait une procédure trop invasive par les différents acteurs du consortium concernés. Force est de constater que l'on retrouve assez peu de travaux s'inscrivant dans cette logique, et le fait que nous ayons pu identifier à travers l'analyse d'un objet intermédiaire unique certains obstacles complexes à mettre en évidence constitue selon nous un argument en la faveur de cette approche.

La matrice de décision présente un double avantage. Le premier est celui de réifier la manière dont se matérialise la multiivocalité qui définit le consortium eFRAN ; en effet, la plupart des acteurs ont pris part à ce vote. Pour reprendre le cadre notionnel mobilisé par Decortis, cette matrice constitue un objet intermédiaire de type médiateur ouvert [12]. Ces objets « laissent de nombreuses marges de manœuvre à ceux qui les utilisent, incitent à un travail d'interprétation et visent à fournir une variété de possibles en vue d'ouvrir des horizons de conception et des premières pistes de recommandation » [5]. Par ailleurs, un objet est médiateur lorsqu'il « peut contribuer à faciliter, rapprocher, dissuader, participer à l'élaboration de compromis entre les acteurs », ce qui constitue justement la fonction première de cette matrice (ibid, p.86).

Le second avantage de cette matrice est qu'elle se prête à une analyse mathématique, qui objective certaines difficultés à prendre en compte le point de vue des utilisateurs dans le choix des orientations technologiques. Nous présentons dans la section qui suit le contexte qui a entouré la genèse et l'utilisation de la matrice de décision, ainsi que l'approche statistique mobilisée.

2 Méthodologie

2.1 Retour sur la genèse de l'objet intermédiaire étudié

La matrice de décision est une initiative de l'industriel [2], Tralalère, qui utilise des outils analogues pour des prises de décision en interne de manière routinière. Précisons que la matrice de décision, et non les fonctionnalités qu'elle porte, représente l'objet intermédiaire d'intérêt. C'est la première fois que ce type de matrice est utilisé par l'entreprise pour réaliser des prises de décision collégiales avec des partenaires. La première mouture de la matrice a été réalisée par l'industriel au cours des premiers mois du projet. Elle a été enrichie par les divers chercheurs impliqués, sur la base d'une revue de la littérature portant sur les EIAH dans l'enseignement des sciences, et sur la base de réflexions faisant suite aux observations réalisées en classe [3]. A la suite de ce travail, trois réunions entre industriel et chercheurs ont eu vocation à harmoniser le vocabulaire utilisé, et à regrouper les fonctionnalités en grandes familles.

La chef de projet a proposé des scores au nom des Savanturiers, après avoir échangé avec des collègues. Un ingénieur de Tralalère a établi un score Stratégie, censé correspondre à la stratégie de l'entreprise, ainsi qu'un score de faisabilité estimée de développement. Il n'était pas inclus dans le score Stratégie de Tralalère, car la faisabilité des développements devait pour le partenaire industriel être tenue comme indépendante des orientations stratégiques qu'il défendait.

Les différents partenaires ont ensuite identifié ce qu'il convenait de soumettre aux enseignants pour le vote. Les fonctionnalités dont le développement était jugé indispensable au bon fonctionnement de l'application par l'industriel et les Savanturiers n'ont par exemple pas été soumises au vote, afin de ne pas disperser inutilement l'attention des enseignants. En définitive, trente-trois "fonctionnalités" ont été soumises au vote des enseignants. Les quatre enseignants investis ont voté lors d'une réunion organisée chez le partenaire industriel, nommée Comité de Conception, réunion à laquelle les chercheurs ont assisté.

En tant que chercheurs, nous nous sommes positionnés en observateurs, et n'avons pas voté. Au sein de ce consortium, nous travaillions dans une large mesure selon une perspective dite de Recherche Orientée par la Conception [19], ou Design-Based Research (DBR) dans la littérature anglo-saxonne [1]. Ce paradigme autorise une certaine réflexivité sur des actions dans lesquelles nous sommes impliqués. Néanmoins, nous étions en désaccord sur cette manière de procéder au choix des orientations technologiques. Ne pas voter nous permettait d'analyser cette étape en tant que simples observateurs, ce qui clarifiait notre positionnement épistémologique.

A l'issue de l'arbitrage, quatre scores ont été fournis pour chaque fonctionnalité : le score *Enseignants*, le score *Savanturiers*, le score *Stratégie* de l'entreprise Tralalère, et un score *Faisabilité*, lui aussi calculé par l'industriel, chacun étant compris entre 0 et 10. Notons que le score *Enseignants* a été calculé comme la moyenne des scores attribués individuellement par chaque enseignant. Les conséquences de ce mode de calcul seront traitées en détail dans les analyses qui seront présentées. Le score final, dont a dépendu dans une large mesure l'ordre de priorité de développement des fonctionnalités, correspond à la somme pondérée des scores attribués par les différents acteurs. Détail central, il a été décidé à l'initiative de l'industriel que la pondération des scores donnés par les enseignants serait doublée par rapport aux autres acteurs, afin d'accorder plus de poids au point de vue de l'enseignant et pour « être au plus près des besoins des utilisateurs ». Nous allons voir que malgré cette mesure, le poids des enseignants dans l'arbitrage a été plus limité que celui des autres acteurs.

2.2 Analyse statistique de la matrice de décision

Nous avons réalisé dans le cadre de cette contribution une analyse statistique portant uniquement sur la matrice de décision. Nous nous sommes intéressés au poids des différents acteurs sur le score final attribué à chaque fonctionnalité. Nous avons procédé d'une part à une analyse de variance et d'autre part à une analyse en composantes principales. L'analyse de variance permet d'extraire la part de variance du score final qui est attribuable à un acteur donné. Ce chiffre correspond donc à la somme des carrés pour un critère donné, divisée par la somme des carrés totaux, il sera fourni dans une table d'ANOVA. Cette approche permet de quantifier

mathématiquement le poids relatif des différents acteurs dans le score final, et, ce faisant, sur cette étape du processus de conception.

Ce poids peut dépendre de plusieurs facteurs : la pondération attribuée à chaque acteur, mais aussi la variabilité dans les scores attribués et la corrélation avec les scores des autres acteurs. Dans la mesure où le score final est une combinaison linéaire des scores attribués par les différents acteurs, plus la variance des scores attribués par un acteur est élevée, plus la probabilité d'influer sur le score final s'accroît. En d'autres termes, un acteur plus enclin à mettre des notes élevées et des notes basses pèsera davantage qu'un acteur qui donne des notes relativement homogènes. Ensuite, le poids d'un acteur est fonction de la plus ou moins forte corrélation de sa manière de noter avec celle des autres acteurs. A cet égard, l'analyse en composantes principales permet de visualiser les corrélations entre les différentes variables, et, à travers une telle analyse, les convergences entre les points de vue des acteurs au moment du vote.

Notons que cette analyse statistique ne constitue pas à elle seule une démonstration des difficultés rencontrées eu égard à la prise en compte des besoins des utilisateurs, dans la mesure où elle ne porte que sur un moment délimité du processus du conception. Elle vise en revanche à s'ajouter à un faisceau de preuves, issu essentiellement du travail d'observation de trois ans mené par les chercheurs, et dont l'exposé, qui nécessiterait de longs développements, dépasse le cadre de cette contribution.

3 Résultats

3.1 Des utilisateurs qui pèsent peu dans le choix des orientations prises

Nous commencerons par montrer que les enseignants ont eu un poids plus faible que les autres acteurs dans l'attribution du score final, à travers une analyse multivariée de la matrice de décision qui permet de visualiser les convergences et divergences entre acteurs dans la manière de noter. Pour expliquer cet état de fait, nous reviendrons plus en détail sur les scores attribués par les enseignants, ce qui nous permettra d'identifier, d'un point de vue mathématique, les déterminants de leur faible influence. Nous allons montrer que les modalités d'articulation des scores qui ont été retenues, doublées par les divergences dans les notes attribuées par les enseignants, ont induit un lissage sur le calcul de leur score, ce qui a réduit mathématiquement leur influence au cours de l'arbitrage analysé.

La première étape de notre analyse consiste en une ANOVA visant à expliquer le poids de différents facteurs explicatifs sur la variation du score final, score qui détermine la priorité des fonctionnalités, et donc celles qui seront retenues. Calculer l'impact sur la variance du score final constitue une manière d'objectiver le poids d'un acteur sur la détermination du score, et donc, dans un certain sens, de la forme que prend l'artefact final. L'ANOVA, dont les statistiques sont présentées dans le Tableau 1, montre que le score *Enseignant*, malgré la pondération plus forte qui a été établie, n'explique que 12% de la variance expliquée du Score final. En revanche le score *Stratégie* de l'industriel explique 41% de cette variance, et la faisabilité,

également établie par l'industriel, 23%. Enfin, les Savanturiers pèsent pour 25% de la variance de ce score.

Tableau 1. Table d'ANOVA du modèle permettant d'expliquer le score total attribué à une fonctionnalité en fonction des scores attribués par les différents acteurs du consortium ayant voté

	DDL	Somme des carrés	Part de la variance expliquée	F-value	Pr(>F)
Faisabilité estimée	1	299	23%	1,70E+31	< 2.2e-16
Stratégie industriel	1	532	41%	3,03E+31	< 2.2e-16
Score Savanturiers	1	325	25%	1,85E+31	< 2.2e-16
Score enseignants	1	144	12%	8,16E+30	< 2.2e-16
Résidus	26	0			

3.2 Des enseignants différents, dont les points de vue s'annulent mutuellement

Plusieurs éléments sont susceptibles d'expliquer le faible poids du score enseignant sur le score final. Nous avons émis l'hypothèse d'un « effet lissage », lié notamment au fait que les différents enseignants impliqués seraient porteurs de visions assez divergentes des fonctionnalités à prioriser, de sorte que le calcul d'un score moyen lisse considérablement la variabilité du score enseignant. Nous allons voir que cette hypothèse est soutenue par l'analyse multivariée des scores attribués par les enseignants. Pour la mettre à l'épreuve, nous avons calculé de nouveaux scores finaux, mais en identifiant le score enseignant au score attribué par un enseignant en particulier, pour mimer la situation au cours de laquelle un seul enseignant choisirait le score enseignant, au lieu de calculer la moyenne de quatre scores. En d'autres termes, nous nous penchons sur quatre situations, au cours desquelles chacun des quatre enseignants aurait le monopole de l'établissement du score.

On constate que le score enseignant, lorsqu'il est calculé de cette manière, influe considérablement sur le score final, puisqu'il explique entre 22% et 40% de la variance du score final, selon l'enseignant considéré, alors qu'il ne correspond qu'à 12% lorsqu'il est calculé comme la moyenne des différents scores enseignants. Si chaque enseignant décidait seul du « score enseignant », il aurait une influence très importante sur le score final, et donc la forme finale de l'artefact, ce qui s'explique également par le fait qu'ils ont une pondération plus importante de leur score.

Nous en avons conclu que le fait de calculer le score enseignant sur la base d'une moyenne conduit à un lissage du score, qui diminue mécaniquement la variabilité de ce score enseignant, et, par conséquent, la part de variance du score final qu'il explique. En somme, le choix de la procédure adoptée - établir un point de vue qui correspond à la moyenne des points de vue des enseignants, est en grande partie responsable de la faiblesse de l'influence des enseignants dans le processus de vote. Néanmoins, l'intensité de ce lissage est fonction de la corrélation entre les scores

attribués par les enseignants. Plus ces corrélations sont faibles, ou en d'autres termes plus les notes attribuées par les enseignants divergent, plus le lissage est important et plus faible est leur influence. Cette considération nous a amené à réaliser une analyse en composantes principales (Figure 1), qui permet de visualiser les corrélations entre les scores des quatre enseignants impliqués en les projetant sur un plan bidimensionnel composé de deux dimensions représentant respectivement 32% et 27% de la variance du jeu de données.

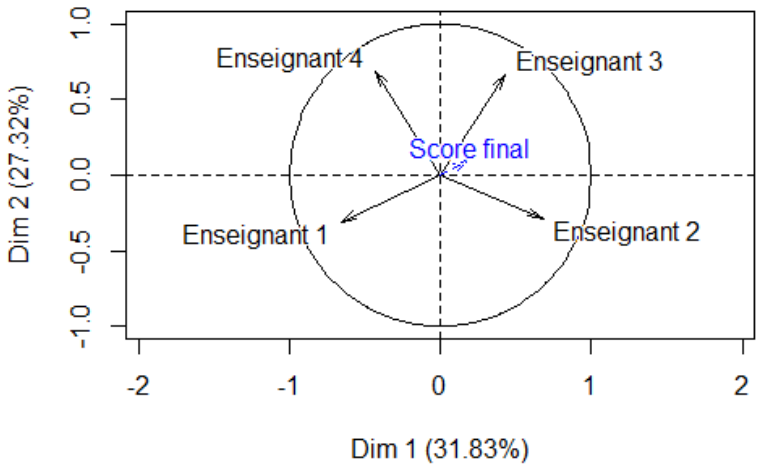


Fig. 1. Cercle des corrélations de l'analyse en composante principale. Il permet de visualiser les divergences entre les scores attribués par chacun des enseignants pour les différentes fonctionnalités soumises au vote. Le « Score final », en bleu, est présenté comme variable illustrative.

Le cercle des corrélations permet de visualiser la faible corrélation entre ces scores. Le tableau des corrélations, qui n'est pas fourni ici, montre que la corrélation entre les scores de deux enseignants atteint au mieux 0,17, et qu'elle est négative dans la majorité des cas. Ce phénomène se traduit dans le cercle des corrélations entre les scores de chaque enseignant, qui permet de visualiser les divergences dans les manières de noter d'un enseignant à l'autre (Figure 1). Cette analyse confirme dans une large mesure que la divergence des points de vue des enseignants quant à la priorisation à accorder aux fonctionnalités a considérablement atténué leur poids sur le score final, et partant de là, sur l'arbitrage effectué. Les différences d'approches pédagogiques, de matières et de niveaux complexifient l'émergence d'un consensus, ce qui questionne la pertinence qu'il y a à vouloir prioriser des besoins selon une logique de vote d'un échantillon limité d'enseignants, dont on peut discuter la représentativité. Cette considération va nous amener à discuter, dans les paragraphes qui suivent, des limites d'une approche inductiviste de la spécification des besoins.

4 Discussion

L'arbitrage que nous avons analysé ici a conduit le consortium à écarter un certain nombre de propositions. Le CNEC, dont un prototype a été finalisé en 2019, partage en définitive beaucoup de caractéristiques avec des environnements comme le WISE (Web-Based Inquiry Science Environment) [17,21]. Il inclut au terme du processus – les dernières versions du prototype ont été codées – une palette d'artefacts articulés entre eux, allant d'outils de brainstorming à des outils de structuration d'écrits analogues à ceux du Knowledge Forum [15,20]. Sa forme finale reflète dans une large mesure l'esprit dans lequel la matrice de décision a été construite : une agrégation de fonctionnalités qui viennent remplir des besoins relativement distincts. Les tests utilisateurs réalisés jusqu'à présent questionnent la cohérence de l'artefact pris dans son ensemble, et révèlent un certain nombre de difficultés que nous attribuons en partie à l'inductivisme qui a caractérisé le choix des orientations technologiques suivies.

4.1 Une approche inductiviste de la conception qui nuit à la cohérence de l'artefact

Les perspectives des différents acteurs ont largement divergé quant à la manière d'identifier les besoins des utilisateurs. Ces résultats suggèrent que la négociation d'une position commune a fini par nuire à la prise en compte du point de vue des utilisateurs, alors même que celui-ci avait vocation à être central. Cette constatation souligne selon nous les difficultés qui émergent d'une vision inductiviste du processus de conception [3]. Par approche inductiviste, nous entendons une vision qui consiste à refuser la moindre idée préconçue sur ce que devrait instrumenter l'artefact à concevoir, et vouloir identifier les fonctionnalités d'intérêt principalement sur la base de discours d'utilisateurs et d'observations de terrain, avec une interprétation aussi minimaliste que possible. Cette approche fait écho à la multiplication des injonctions des politiques publiques visant à intégrer davantage les enseignants dans le processus de conception d'un EIAH, selon une logique de conception participative [7].

On constate à la lueur de nos analyses l'écart qui existe entre une volonté proclamée de s'inscrire dans une logique de conception participative, avec une pondération plus importante du vote des utilisateurs, et l'influence effective des utilisateurs finaux à un moment supposé clé du processus de décision. Cet inductivisme se manifeste par un discours, et comme nous le voyons ici, des pratiques, qui consistent à affirmer que les besoins doivent « émerger du terrain », mais sans qu'il n'y ait de réel consensus sur la méthodologie pour les faire émerger.

En effet, l'approche suivie au cours du vote rend le processus de décision particulièrement sensible aux modalités d'articulation des points de vue d'une poignée d'enseignants, sans régler la question de leur représentativité. L'analyse de cet arbitrage constitue selon nous une nouvelle illustration des limites d'une démarche inductiviste, qui mettrait au second plan des modèles théoriques préexistants. Cette considération nous amène à souligner l'importance qu'il y a, pour appréhender les dynamiques traversant un processus de conception, à identifier les quelques objets intermédiaires qui jouent un rôle central dans la direction que prend le processus, et

que Decortis qualifie d'objet intermédiaires pivot [5, p.85], c'est-à-dire qui acquièrent un rôle central dans la conception, vont impulser de manière irréversible des orientations au projet dans son ensemble.

4.2 Identifier les objets intermédiaires pivots d'un processus de conception

L'idée d'utiliser le CNEC en premier lieu comme un terrain pour tester des hypothèses scientifiques [23] n'a pas eu d'écho favorable au sein du consortium, dans la mesure où une telle initiative paraissait incompatible avec l'intention dans laquelle était conçu l'artefact : produire un outil industrialisable. En revanche, les chercheurs ont tenté de proposer des modèles théoriques en guise d'objets intermédiaires pivot.

Dans une recherche qu'elle mène sur POGO, un artefact dédié à l'écriture créative, Decortis [5] avait utilisé un modèle théorique, le modèle NAM, pour orienter la conception. Ce modèle avait connu de multiples simplifications successives, et permis une meilleure intercompréhension entre industriel et chercheurs. L'auteur souligne le rôle cadrant du modèle pour la production d'artefacts [5, p.106]. Cette démarche d'adaptation réalisée par des chercheurs est nécessaire si l'on veut que de tels modèles aient une influence sur le processus de conception.

Le principal obstacle que nous avons rencontré a été la difficulté à trouver un consensus quant au modèle théorique le plus approprié, et si le modèle de la co-élaboration de connaissances [20] emporta un temps l'adhésion d'un certain nombre des membres du consortium, ce ne fut jamais au point de devenir un objet pivot du processus de conception.

L'intérêt qu'il y a à identifier et à analyser la genèse et les caractéristiques d'un objet pivot est que ceux-ci permettent de simplifier la compréhension du processus dans son ensemble en réifiant un certain nombre de dynamiques. La conception d'un artefact comme le CNEC est faite de centaines, voire de milliers de décisions prises dans des lieux variés. Il est impossible de documenter, dans une optique de recherche, la totalité des réunions et des décisions prises, et à plus forte raison d'analyser l'ensemble des interactions afférentes. L'analyse que nous avons présentée vise à étayer l'idée selon laquelle, en sélectionnant un nombre réduit d'OIP, et en se centrant sur leurs caractéristiques plus que sur les interactions qui les entourent [8], le chercheur peut appréhender les tenants des moments clés du processus de conception.

5 Conclusion

L'une des limites de notre recherche réside dans le fait que l'analyse de la matrice ne présage en rien de l'influence effective des enseignants impliqués dans la conception du CNEC à l'échelle de l'ensemble du projet - ils sont intervenus à bien d'autres moments de la conception pendant près de deux années après cet arbitrage. Il serait nécessaire de prolonger le travail que nous avons mené ici pour mieux comprendre comment les « besoins des utilisateurs » ont été pris en compte par les différents acteurs du consortium, tout au long du processus de conception.

Cette étude suffit néanmoins à pointer le fait que, pour mesurer l'influence des enseignants sur un projet de conception, on ne peut se satisfaire de bilans comptables

du nombre d'heures que ceux-ci auraient investi dans les réunions. L'injonction à impliquer davantage les enseignants dans la conception des technologies éducatives subventionnées par l'Etat résulte possiblement d'une volonté d'anticiper des critiques putatives contre le financement public de solutions déconnectées des problèmes du terrain.

Néanmoins, puisqu'il n'y a pas de consensus sur ce que sont ces besoins, ni sur la manière dont les acteurs du système éducatif peuvent contribuer à les identifier, on peut se demander dans quelle mesure le fait de collaborer étroitement avec des enseignants améliore effectivement l'adéquation des EIAH conçus avec les besoins des praticiens, quelle que soit la manière dont on les définit.

Nous conclurons sur le fait que l'ancrage empirique d'une stratégie gagnerait à être accompagné d'une analyse approfondie de l'offre technologique existante, et de ses évolutions futures. Ce travail inclurait non seulement un état de l'art sur les technologies éducatives publiées dans la littérature, réalisé par ailleurs dans le cadre de ce projet [4], mais aussi sur celles qui sont utilisées au quotidien, dans les salles de classe, sans nécessairement avoir reçu l'attention de la part de la communauté scientifique. Reste à déterminer dans quelle mesure l'analyse de cette offre logicielle constitue un objet légitime pour les chercheurs en sciences de l'éducation.

6 Références

1. Amiel, T. et Reeves, T. C. (2008). Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda. *Journal of Educational Technology et Society*, *11*(4), 29–40
2. Cisel, M., Beauné, A, Bernard, F., Voulgre et E. Baron, G. (2017, juin). Analyse d'un outil de décision mobilisé dans la conception d'un EIAH. Dans *Actes de la 8ème conférence sur les Environnements Informatiques pour l'Apprentissage Humain,* Strasbourg, France. (p.382-385).
3. Cisel, M. et Baron, G.-L. (2018, juin). Conception d'un EIAH à destination du programme Savanturiers : difficultés engendrées par une approche inductiviste de la spécification des besoins. Dans Boulc'h, L, Voulgre, E (Eds.), *Actes de la 3ème conférence Ecole et Technologies de l'Information et de la Communication (ETIC3)*, Paris, France.
4. Cisel, M. et Baron, G. (2019, avril). Vers des intelligences artificielles pour l'enseignement du raisonnement scientifique. Dans *Actes du Séminaire Inter-Laboratoires sur l'Education Scientifique et Technologique.* Patras, Grèce.
5. Decortis, F. (2015) *L'ergonomie orientée enfants. Concevoir pour le développement*. Paris, France : PUF.
6. Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki, Finland: Orienta-Kosultit Oy.
7. Frascara, J. (2003). *Design and the Social Sciences: Making Connections*. CRC Press.
8. Grégori N., Blanco E., Brassac C. et Garro O. (1998). Analyse de la distribution en conception par la dynamique des objets intermédiaires, p. 135-154, dans Trousse B., Zreik K. (éds). *Les objets dans la conception*. Paris, France : Europia.
9. Hasni A., Bousadra F. et Lebeaume J. (dirs) (2018). *Les démarches d'investigation scientifique et de conception technologique. Regards croisés sur les curriculums et les pratiques en France et au Québec.* Saint Lambert, Québec : Cursus universitaire.
10. Houseal, A. K., Abd-El-Khalick, F., et Destefano, L. (2014). Impact of a student–teacher–scientist partnership on students' and teachers' content knowledge, attitudes toward

science, and pedagogical practices. *Journal of Research in Science Teaching, 51(1), 84–115.*

11. Jean-Daubias, S. (2000). *Pépite, un système d'assistance au diagnostic de compétence.* (Thèse de doctorat non publiée). Université du Mans. France
12. Jeantet, A., Tiger, H., Vinck, D. et Tichkiewitch, S. (1996). La coordination par les objets dans les équipes intégrées de conception de produit. *Coopération et conception*, Toulouse, France : Octares.
13. Jeantet, A. (1998). Les objets intermédiaires dans la conception. Éléments pour une sociologie des processus de conception. *Sociologie du travail*, *40*(3), 291–316.
14. Laborde, C. (2002). Integration of technology in the design of geometry tasks with Cabri-Geometry. *International Journal of Computers for Mathematical Learning*, *6*(3), 283-317.
15. Laferriere, T. et Lamon, M. (2010). Knowledge Building/Knowledge Forum®: The Transformation of Classroom Discourse. Dans M. S. Khine et I. M. Saleh (Eds.), *New Science of Learning (pp. 485–501).* New York, NY : Springer.
16. Linn, M. C. (2000). Designing the Knowledge Integration Environment. *International Journal of Science Education,* 22(8), 781–796.
17. Linn, M. C. (2013). *Internet Environments for Science Education*. Abingdon-on-Thames, UK : Routledge.
18. Quintana, C., Zhang, M. et Krajcik, J. (2005). A Framework for Supporting Metacognitive Aspects of Online Inquiry Through Software-Based Scaffolding. *Educational Psychologist, 40(4),* 235–244.
19. Sanchez, É. et Monod-Ansaldi, R. (2015). Recherche collaborative orientée par la conception. *Education et didactique*, *9*(2), 73–94.
20. Scardamalia, M. et Bereiter, C. (2003). Knowledge building environments: Extending the limits of the possible in education and knowledge work. *Encyclopedia of Distributed Learning*, 269–272.
21. Slotta, J. D. et Linn, M. C. (2009). *WISE Science: Web-Based Inquiry in the Classroom*. New York, NY : Teachers College Press.
22. Star S.L. et Griesemer J. (1989). Institutionnal ecology, 'Translations' and Boundary objects: amateurs and professionals on Berkeley's museum of vertrebate zoology. *Social Studies of Science, 19*, 387-420.
23. Tchounikine, P. (2011). *Computer Science and Educational Software design - A Resource for Multidisciplinary Work in Technology Enhanced Learning*. Berlin, Allemagne : Springer.
24. Tricot, A., Plégat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G. et Morcillo, A. (2003, avril). Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH. Dans Cyrille Desmoulins, Pascal Marquet, Denis Bouhineau (Eds), *Actes de la conférence the Environnements Informatiques pour l'Apprentissage Humain* (EIAH 2003) (p. 391–402). Strasbourg, France
25. Vinck, D. (2009). De l'objet intermédiaire à l'objet-frontière. *Revue d'anthropologie des connaissances, 1*(1), 51–72.
26. Wu, H.-K. et Hsieh, C.-E. (2006). Developing Sixth Graders' Inquiry Skills to Construct Explanations in Inquiry-based Learning Environments. *International Journal of Science Education*, *28(11)*, 1289–1313.

Muse-Logique : Un Système Tutoriel Intelligent pour l'apprentissage du raisonnement logique

Roger Nkambou[1], Serge Robert[1], Janie Brisson[1], Ange Tato[1], Clauvice Kenfack[2], Pamela Kissok[1]

[1] Université du Québec à Montréal, 201 Av. du Président Kennedy, Montréal, QC, Canada
nkambou.roger@uqam.ca, robert.serge@uqam.ca
[2] Université de Yaoundé 1, Cameroun
clauvicek74@gmail.com

Résumé. Nous présentons une démarche de conception du Système Tutoriel Intelligent (STI) Muse-Logique conçu pour favoriser le développement des compétences en raisonnement logique dans différents situations et contextes. L'ensemble des composantes développées ont été validées et argumentés par des experts tout au long du processus (informaticiens, logiciens et psychologues du raisonnement). Un catalogue des erreurs (syntaxique et sémantique) a été établi, en plus d'une représentation explicite des connaissances sémantiques et des structures et méta-structures sous-jacentes au raisonnement conditionnel. Une expérimentation a permis de collecter des données de près de 300 étudiants ayant traité 48 activités de raisonnement. Ces données ont contribué à l'élaboration d'un modèle psychométrique utilisé pour initialiser le modèle de l'apprenant et pour valider la structure du réseau bayésien. Muse-Logique innove d'une part par l'approche qui a mené à son développement une co-construction donnant lieu à des structures fortes sur lesquelles le raisonnement est basé, et d'autre part par son approche hybride à la construction du réseau bayésien utilisé dans le modèle de l'apprenant.

Mots-clés. Systèmes Tutoriels intelligents, Modèle de l'apprenant, Diagnostic cognitif, Apprentissage du raisonnement logique

Abstract. In this paper, we present the design and implementation Logic-Muse, an Intelligent Tutoring System. Logic-Muse aims at helping learners to develop their logical reasoning skills in different problem-solving situations. All the components developed have been validated and argued by experts (ITS researchers, logicians and reasoning psychologists) throughout the process. A catalog of reasoning errors has been established, in addition to an explicit representation of semantic knowledge and structures as well as meta-structures underlying conditional reasoning. We carried out a study to collect data from nearly 300 students who had processed 48 reasoning activities. These data contributed to the development of a psychometric model used as part of the learner model. Muse-Logique innovates by its participatory development approach (leading in strong and valid structures on which reasoning is based), but also by its hybrid approach to the building of the Bayesian learner model.

Keywords. Intelligent Tutoring System (ITS), Logical Reasoning Learning, Cognitive Diagnosis, Bayesian Student Modeling

1 Introduction

De nombreuses expériences en sciences cognitives montrent que des erreurs systématiques sont courantes chez l'humain dans l'usage du raisonnement logique [7,17]. Un certain nombre de questions se posent lors de la recherche de solutions pour améliorer les compétences humaines dans ce domaine: quels sont les phénomènes impliqués dans l'acquisition de compétences de raisonnement logique? La modélisation permet-elle de les obtenir? Quelles sont les stratégies d'adaptation pour favoriser le développement de la compétence logique? Quelles sont les caractéristiques d'un système de tutorat intelligent (STI) pour soutenir cet apprentissage? Les réponses à ces questions ne peuvent être apportées sans une explication et une compréhension appropriée des connaissances sous-tendant le raisonnement logique et les erreurs commises par les humains, couplées avec la participation active d'experts pertinents, notamment des logiciens, des psychologues du raisonnement et des professionnels de l'éducation en logique.

Dans le projet Muse-Logique, ces conditions sont remplies avec une équipe multidisciplinaire mise en place il y a quelques années. L'objectif est d'étudier les bases de l'acquisition des compétences de raisonnement logique, de comprendre les difficultés liées à cet apprentissage et de créer un STI capable de détecter, de diagnostiquer et de corriger les erreurs de raisonnement dans diverses situations. Pour ce faire, notre approche de conception participative vise à: 1) fournir un catalogue des erreurs de raisonnement logique; 2) développer une théorie originale des causes de ces erreurs, y compris une théorie de la compétence en raisonnement logique; 3) analyser les structures formelles du traitement de l'information que l'on trouve dans les systèmes logiques; 4) développer un système de tutorat intelligent pour aider les apprenants à améliorer leurs capacités de raisonnement et ultimement, 5) aider les experts à redéfinir la nature de la rationalité humaine en utilisant les STI comme un banc d'essai pour tester d'autres hypothèses de la science cognitive du raisonnement.

L'approche participative adoptée a mené à l'élaboration de composants essentiels pour la construction du STI Muse-Logique. Ceux-ci ont été préalablement élicités, validés et argumentés par les experts. Ces composants incluent le catalogue des erreurs, les structures et méta-structures du raisonnement logique et les stratégies remédiatives favorisant l'émergence de bons modèles mentaux. Muse-Logique dans sa version actuelle implémente ces composants pour la logique classique des propositions et offre une panoplie d'activités permettant à un apprenant de développer sa compétence du raisonnement logique dans plusieurs classes de situations que nous avons établies avec les experts : concrètes, contrefactuelles et abstraites.

Cet article décrit la démarche de conception participative de Muse-Logique, notamment les fonctionnalités des trois modules issus de celle-ci. La section 2 décrit les fondements théoriques sous-jacents à Muse-Logique suivi par les sections 3, 4 et 5 qui décrivent les résultats de la conception participative des trois principaux composants du système : le modèle du domaine, le modèle de l'apprenant et le tuteur. La section 6 illustre l'implémentation de Muse-Logique suivie d'une conclusion qui résume les contributions et quelques travaux futurs.

2 Cadre théorique

Pourquoi construire un système pour le raisonnement logique? Cette idée part des constats suivants : 1) Une partie importante de la connaissance humaine relève du traitement de l'information par des raisonnements, 2) Le raisonnement logique occupe une place importante dans la cognition, 3) Le raisonnement logique humain est plus souvent non classique que classique, 4) En tant que machine cognitive en situation de lutte pour sa survie, l'humain tend à faire des erreurs systématiques dans ses raisonnements logiques, 5) Ces erreurs sont dues à l'emploi de stratégies de découverte (inductives, analogiques, abductives) dans un contexte de justification. 6) Apprendre à raisonner logiquement, c'est apprendre comment fonctionne le traitement cognitif de l'information, tout en apprenant les lois et procédures valides du raisonnement logique. 7) Les mécanismes cognitifs sont le produit de modules cérébraux relativement autonomes mais également en interaction étroite entre eux.

A partir de ces constats, l'objectif de nos travaux a consisté à rassembler les développements récents dans les domaines des Systèmes Tutoriels Intelligents, des Sciences cognitives, et d'Intelligence Artificielle pour proposer un environnement d'apprentissage capable d'aider les étudiants à améliorer leurs compétences en logique. Notre perspective pour étudier le raisonnement logique est naturaliste (les mécanismes du raisonnement sont le résultat d'une histoire évolutionnaire), cognitiviste (la cognition est du traitement d'information) et méliorativiste (nous pouvons améliorer nos compétences logiques). L'information y est définie comme de la régularité, par opposition à la probabilité. Le raisonnement consiste à tirer de l'information (conclusion), à partir d'information donnée (prémisses), en appliquant une règle (implicite ou explicite). Le type de règle définit la nature du raisonnement. Quand ces règles de raisonnement sont logiques, elles sont non ampliatives, c'est-à-dire qu'elles n'augmentent pas l'information entre les prémisses et la conclusion, par opposition aux raisonnements créatifs, qui sont ampliatifs, qui augmentent l'information entre les prémisses et la conclusion. Ainsi, notre hypothèse sur l'erreur logique est que lorsqu'elle est syntaxique, elle vient du traitement du raisonnement logique comme s'il s'agissait d'un raisonnement créatif. Les sophismes sont en effet des inférences qui créent de l'information sans justification pour le faire [15]. Elles surviennent aussi lorsque les contenus alternatifs ne sont pas considérés. Des études ont prouvé que la sensibilité aux contre-exemples était en effet un moyen d'inhiber les sophismes de l'implication [12].

Les Systèmes Tutoriels Intelligents ont fait avancer la compréhension de l'apprentissage en général, mais ont aussi fait leurs preuves quant à leur capacité à soutenir l'apprentissage du raisonnement en sciences, mathématiques, et même en logique [2,10,20]. Nous pouvons formuler deux limites des STI proposés en logique : 1) Ces systèmes ne se fondent pas sur une élicitation des connaissances et structures logiques essentielles au raisonnement, limitant ainsi leur portée explicative du raisonnement de l'apprenant; 2) Ils ne se fondent pas sur des théories fondamentales de l'erreur de raisonnement au centre desquelles se trouve la théorie des processus duaux [7,19]. Celle-ci identifie deux types de processus mentaux qui peuvent être activés dans le raisonnement logique : des processus de type 1, spontanés, ancestraux et probablement innés, et des processus de type 2, réflexifs, récents et appris. Les

processus de type 1 nous permettent de faire certains raisonnements logiques valides, mais ils sont aussi responsables de plusieurs erreurs. Ces processus de type 1 ne peuvent pas être éliminés, ils sont trop ancrés en nous, mais nous pouvons apprendre à les inhiber et à activer des processus de type 2. Cette activation accroît considérablement notre compétence logique en nous permettant d'éviter les erreurs causées par le type 1. Nous appuyant sur ces thèses, nous explorons les différences entre ces deux types de processus et les stratégies pour inhiber les premiers et activer les seconds.

L'activation des processus de type 2 et la correction de l'erreur logique ne peuvent se faire sans l'inhibition des processus de type 1 et cette inhibition exige de prendre conscience de l'erreur. En voulant améliorer la théorie du développement cognitif de Piaget, Moshman a démontré qu'une prise de distance métacognitive est essentielle à cette prise de conscience de l'erreur et à son inhibition [13]. Notre projet explore cette thèse et tente de montrer comment des détours métacognitifs sont essentiels à l'acquisition de la compétence logique.

Par exemple lorsque plongée dans différents contextes, mais avec un même problème soumis, une personne éprouverait de la difficulté à bien conclure, même dans une logique apparemment simple comme la logique des propositions. Elle réussirait dans un contexte plus commun qu'un contexte plus abstrait par exemple. Le contexte ici renvoie à la manière dont est formulé le problème. Il peut être : familier (causal), abstrait ou encore formel. Les théories évolutionnistes ont pu faire ressortir l'influence des biais et des heuristiques liés aux effets de contexte. Ces théories soutiennent que même ayant la technique logique appropriée pour raisonner, certains facteurs tels que les biais de croyance et les heuristiques avec lesquels nous évoluons toute notre vie influenceront notre raisonnement [5,21]. Grâce à ces théories, il devient possible d'expliquer les erreurs de raisonnement invalides.

Selon Robert, le processus d'opérations mentales met en relation le raisonnement et la catégorisation de l'information [15]. Pour raisonner, il faut organiser et structurer l'information en catégories, établir des liens entre elles, sélectionner et appliquer des règles logiques nécessaires sur ces informations. La conception de notre Muse-Expert (le modèle du domaine de Muse-Logique) a fait appel à la mise en œuvre de ces théories. La catégorisation dans ce cas réfère à l'organisation de la connaissance, la distinction des règles d'inférences (valides et invalides), la décomposition des informations de chaque problème sur lesquels le système doit raisonner et la représentation des structures/modèles dont il est doté.

3 De la conception participative du module expert

Sur le plan architectural, Muse-Logique est constitué des trois modules classiques d'un système tutorial intelligent [14]: un module expert, un module tuteur et un modèle de l'apprenant. L'environnement d'apprentissage donne accès à une banque de problèmes de raisonnement logique générés dans des situations variées. Chaque module offre une interface qui permet sa manipulation individuelle par un utilisateur.

Comme présenté dans la figure 1, le module Muse-Logique Expert met en œuvre aussi bien les habiletés et connaissances en raisonnement logique que les mécanismes

de raisonnement connexes (syntaxique et sémantique des règles du système logique donnée). En fait, puisque nous prévoyons couvrir plusieurs systèmes logiques, Muse-logique comporte un contrôleur général dont le but est de sélectionner la logique la plus pertinente dans laquelle la situation (de l'énoncé du problème actuel) s'adapte. Pour se faire le système s'appuie sur des méta-connaissances pour rechercher des motifs (marqueurs de chaque système logique fournis par les experts) dans la situation qui offre une meilleure prédiction, conduisant à la sélection du système logique approprié. Dans cet article, nous nous concentrons uniquement sur un système logique (la logique classique des propositions).

Fig. 1.: Composants de Muse-Expert

Spécification de la mémoire sémantique et procédurale de la logique propositionnelle. La conception participative de chaque composant du module expert a été soigneusement réalisée dans l'équipe. Tout d'abord, nous avons étudié le domaine de la logique propositionnelle ce qui nous a amené à la spécification complète de tous les concepts qui lui sont liées, puis conduit à une ontologie de domaine que nous avons implémenté comme un modèle OWL. L'ontologie formelle a ensuite été validée lors d'une séance de travail avec les experts logiciens.

L'équipe a ensuite entrepris d'éliciter le processus de déclenchement des connaissances procédurales. Ce processus a abouti à la spécification de toutes les règles d'inférence logique dans la logique propositionnelle y compris les règles valides et non-valides. 8 règles d'inférence valides ont été mises en œuvre, deux pour chacun des quatre opérateurs binaires de base de la logique propositionnelle (conjonction, disjonction, implication et incompatibilité). Durant ce processus, les experts logiciens et psychologues du raisonnement ont travaillé ensemble sur la validation empirique de certaines preuves liées à des concepts logiques "mal définis" tels que les constructions d'incompatibilité. L'étude des huit autres règles invalides a conduit à la mise en œuvre du catalogue des erreurs de raisonnement. Les erreurs de raisonnement ont été divisées en deux catégories: les sophismes (inférences invalides) et les suppressions d'inférences valides.

Les *Sophismes* sont des erreurs de raisonnement logique dans lesquelles le raisonneur ne parvient pas à trouver le bon lien entre les prémisses et la conclusion des règles en cause, quant aux suppressions d'inférences valides, elles se produisent quand il ou elle ajoute de nouvelles informations dans les prémisses rendant impossible la conclusion.

Élaboration des méta-structures du raisonnement en logique propositionnelle. L'équipe a examiné en détail l'encrage sémantique du raisonnement à travers les méta-structures. En fait, le raisonnement logique a été considéré à la fois dans ses

composantes syntaxiques et sémantiques. Quand la logique associe ainsi le syntaxique et le sémantique, la notion de structure devient centrale. Un système logique est fait de structures. Une structure est un ensemble d'objets affectés par des relations (comme >, =…) et/ou des opérations (comme +, x...). Une structure qui permet de composer de nouveaux objets par des opérations est une structure algébrique, tandis qu'une structure de relation entre des objets est une structure d'ordre. Chaque structure formelle se distingue par ses connecteurs ou ses opérateurs et par les propriétés que possèdent ces relateurs ou opérateurs. La logique classique est régie par des structures logiques telles que l'algèbre de Boole et une structure d'ordre appelée « treillis de Boole ». Les logiques non classiques peuvent être obtenues par des affaiblissements ou des élargissements de l'algèbre de Boole ou du treillis de Boole. Dans cette perspective, l'apprentissage de différents systèmes logiques peut être vu comme étant l'acquisition de différentes structures algébriques pour catégoriser le monde et les différentes structures d'ordre pour établir des hypothèses de dépendance causale entre ces catégories. Nous traitons alors les stratégies de type 1 causant les erreurs logiques comme des simplifications des structures à l'œuvre dans les raisonnements. D'un point de vue métacognitif, nous utilisons ces métastructures pour illustrer le raisonnement erroné, interprété comme un mauvais parcours de celle-ci ou comme un écrasement de structure (groupe de Klein) [16].

Des situations de raisonnement multiples. Le raisonnement logique n'est pas un processus absolu dans le sens que le même type de raisonnement peut être difficile à appliquer d'une classe de situations à une autre. Être compétent dans l'usage du modus ponendo ponens (MPP) dans une situation donnée (par exemple en situation descriptive causale avec peu d'alternatives) n'implique pas qu'on le sera dans une autre (par exemple en situation descriptive causale avec beaucoup d'alternatives). Cette thèse est soutenue par une étude menée par les membres l'équipe Muse-Logique dans laquelle certains sophismes étaient attribuables à la difficulté de trouver des contre-exemples dans un raisonnement mettant en œuvre l'affirmation du conséquent [4,11]. Nous avons examiné avec les experts logiciens et les psychologues du raisonnement différentes familles de situation pouvant engendrer une variation du niveau de difficulté d'un type de raisonnement. Nous avons ainsi établi sept familles de situations (table 1).

<table>
<tr><td colspan="4">DESCRIPTIVE</td><td>NORMATIVE</td></tr>
<tr><td colspan="2">Causal</td><td colspan="2" rowspan="2">Abstract</td><td rowspan="4"></td></tr>
<tr><td>Concret</td><td>Contrary-to-fact</td></tr>
<tr><td>Few Alternatives (DCCFA)</td><td>Few Alternatives (DCCFFA)</td><td rowspan="2">Formal (AF)</td><td rowspan="2">Informal (AI)</td></tr>
<tr><td>Many Alternatives (DCCMA)</td><td>Many Alternatives (DCCFMA)</td></tr>
</table>

Table 1: Familles de situation de raisonnement (classe de problèmes)

La catégorisation fournit non seulement une base permettant de classer les compétences en raisonnement, mais aussi permet d'organiser les activités d'apprentissage du raisonnement. Le raisonnement normatif a été exclu parce que les expérimentations du raisonnement sur du contenu normatif ont montré une excellente

performance tant des adultes que des enfants [6,8]. Une explication de cet effet de contenu est que l'esprit humain comporte un module spécialisé pour raisonner sur les normes sociales [8].
La catégorisation nous a permis d'établir une mémoire procédurale de la logique des propositions comportant 96 éléments de connaissances, soit 48 éléments de connaissances valides et 48 non valides. Par ailleurs cette contextualisation aide à classifier les exercices à soumettre à l'apprenant et défini un cadre fin pour l'évaluation des compétences de l'apprenant. Par exemple catégorie **Causale avec peu d'alternatives** comporte des exercices avec un contenu familier, facile à comprendre, et dont l'antécédent présente très peu d'alternatives *: Si un chien a des puces, alors il se gratte constamment. Alternatives : avoir la peau sèche ou une maladie de la peau*; alors que dans la catégorie **Causale avec beaucoup d'alternatives**, en plus d'un contenu familier, facile à comprendre, il y a beaucoup d'alternatives causales quand on y réfléchit bien. Exemple : *Si on lance une roche dans la fenêtre, alors la fenêtre brisera.* ***Alternatives*** *: lancer une chaise dans la fenêtre, lancer une brique dans la fenêtre, collision avec une voiture, tempête tropicale, frapper la fenêtre avec un bâton, se projeter dans la fenêtre*, etc.
Tous les exercices sont présentés sous forme de syllogisme et offre des choix de réponse. Exemple de syllogisme :

Règle : Si je lance une roche dans la fenêtre, alors la fenêtre brisera.
Information : La fenêtre est brisée.
Que peut-on conclure ?

- *On a lancé une roche dans la fenêtre*
- *On n'a pas lancé une roche dans la fenêtre*
- *On ne peut conclure*

4 Le modèle de l'apprenant

Le modèle de l'apprenant dans Muse-Logique comporte plusieurs dimensions (Figure 3). La mémoire épisodique garde la trace des exercices effectués par l'apprenant ainsi que tous les éléments de performance reliés. Le modèle cognitif est principalement un modèle bayésien dont les nœuds sont les 96 unités de connaissances liées au raisonnement. Les probabilités a priori ainsi que les relations d'influence entre les différentes unités de connaissance ont été initialement établies par les experts. Certains nœuds sont directement connectés aux exercices de raisonnement (items). Certains nœuds du réseau sont bien entendu liés aux inférences invalides (sophismes ou des suppressions d'inférence). On entretient ainsi et de manière automatique l'état de l'apprenant sur chaque unité de connaissances (valide ou erronée). Les statistiques sur la performance sont calculées à partir des données stockées dans la mémoire épisodique.

Allié au réseau bayésien construit avec les experts, un modèle de diagnostic cognitif de type CDM (*Cognitive Diagnosis Model* [18]) a été construit à partir des données collectées lors d'une expérimentation avec 300 participants sur une banque de 48 exercices.

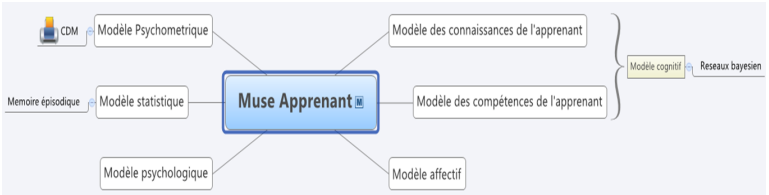


Fig. 3: Composants de modèle de l'apprenant dans Muse-Logique

Pour ce faire, une Q-Matrix (voir extrait à la figure 4(a)) mettant en évidence le croisement entre les items (48 regroupés en 16 classes) et les unités de connaissances ou habiletés cognitives (9 classes) a été soigneusement élaborée avec l'aide des experts logiciens. Nous avons ensuite entraîné un modèle *DINA* sur les 300 patrons de réponses, ce qui a permis d'estimer des paramètres utiles comme le « guess », la probabilité qu'un apprenant puisse répondre correctement à un exercice sans avoir les habiletés nécessaires; « slip » pour la probabilité d'une mauvaise solution alors qu'on disposait des habiletés nécessaires.

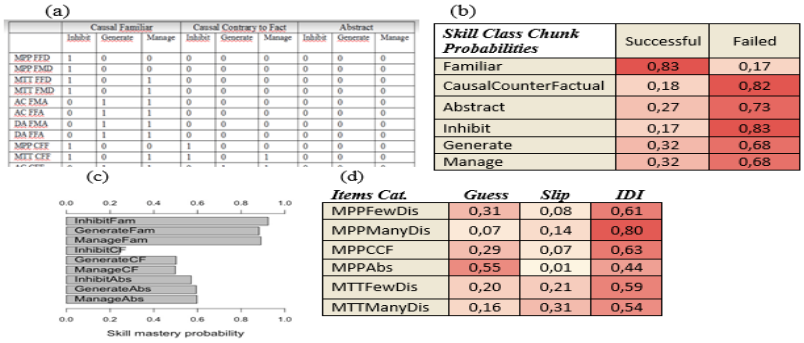


Fig. 4. Quelques résultats du modèle (probabilités des habiletés et des items)

La figure 4 illustre quelques résultats du modèle construit : en (c), la probabilité marginale des habiletés, en (d) les probabilités des paramètres des items et en (b) les probabilités d'échec/succès liées aux habiletés. Les détails de ce modèle ne sont pas présentés dans cet article. Toutefois, il est important de noter qu'il fait partie intégrante du modèle de l'apprenant et est utilisé notamment pour l'initialisation du modèle (soit la prédiction du niveau initial d'un apprenant). Cette prédiction permet d'estimer les probabilités a priori du réseau bayésien (plutôt que de considérer la chance). Par ailleurs, ce modèle a été utilisé pour valider et ajuster la structure du réseau initialement construit avec les experts, aboutissant ainsi à une structure plus efficace informée à la fois par les données (bottom-up) et les experts (top-down). Enfin, bien que prévu, le modèle affectif n'a pas été implémenté.

5 Fonder la rétroaction tutorielle à la fois sur le catalogue des erreurs et les structures métacognitives de la logique

L'objectif principal dans Muse-Logique est d'aider l'apprenant à devenir un bon raisonneur logique quelles que soient les situations de raisonnement. Muse-Logique Tuteur doit être en mesure de diagnostiquer la procédure mentale de raisonnement de l'apprenant pour détecter les erreurs de raisonnement et l'aider à les corriger. Pour ce faire, en collaboration avec les experts, nous avons précisé les interventions du tuteur lorsque des erreurs sont détectées ainsi que la façon dont le tuteur apporte son aide pour les corriger. Par exemple, dans un problème simple de syllogisme en situation concrète causale avec beaucoup d'alternatives (situation DCCMA), si l'apprenant décide de conclure à un problème invalide, le système doit vérifier si la réponse de l'apprenant vient du fait qu'il n'a pas examiné d'autres alternatives possibles de la situation mentionnée dans la prémisse. Si tel est le cas, le tuteur lui recommandera alors de penser à l'existence d'alternatives. Si l'apprenant refuse de conclure à un problème valide, le tuteur lui demandera la raison de cette abstinence. Si sa réponse montre qu'il doute de la véracité de la prémisse majeure, le système lui rappellera l'importance de toujours considérer celle-ci comme vraie.

Nous avons implémenté une base de telles règles d'intervention fournies par les experts logiciens pour chaque erreur possible suivant la situation (type de contenu) dans laquelle le raisonnement est opéré. En outre, l'un de nos objectifs était de mettre en place un soutien métacognitif à l'apprenant. Nous avons examiné attentivement les méta-structures logiques (pour la logique propositionnelle classique) comme décrit dans les sections précédentes et avons analysé la façon dont elles seront utilisées pour fournir une rétroaction visuelle à l'apprenant afin qu'il ou elle soit consciente de ses propres erreurs de raisonnement.

6 Implémentation et intégration des composants validés

Muse-Logique offre quatre niveaux d'activités d'apprentissage pour l'apprenant, organisés en trois groupes de services d'apprentissage : 1) le ***service d'exploration du domaine*** qui utilise l'ontologie du domaine; 2) ***le service des exercices*** permettant à l'apprenant de résoudre une variété de problèmes de raisonnement incluant les syllogismes et les polysyllogismes, mais aussi des exercices généraux sur les concepts du raisonnement y compris la conversion de formules logiques dans des situations de raisonnement, la vérification de formules bien formées, la construction de table de vérité, etc. Ce service intègre un générateur automatique de problèmes (voir plus loin); 3) ***l'espace de visualisation métacognitif du raisonnement*** qui offre entre autres la possibilité d'explorer et d'illustrer les erreurs par les méta-structures de la logique (treillis de Boole et groupes de Klein pour la logique propositionnelle), complété par un sous-espace d'autodiagnostic offert à l'apprenant et lui permettant de simuler son raisonnement à travers des structures sémantiques comme une table de vérité par exemple. Dans cette dernière perspective, les figures 5 et 6 font suite à une erreur d'inférence sur un exercice de raisonnement, obligeant le Tuteur à *prompter*

l'apprenant pour qu'il le réalise par lui-même par une simulation interactive de son raisonnement à travers une table de vérité.

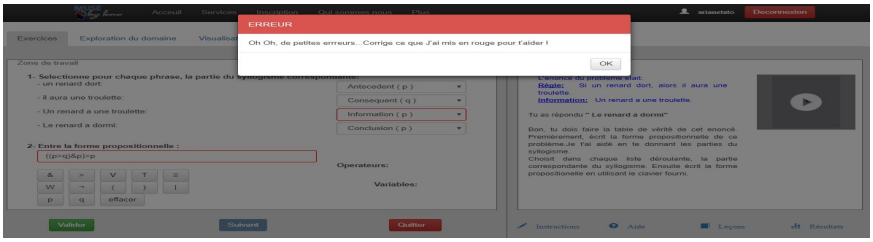


Fig. 5. Espace de simulation du raisonnement suite à une erreur de raisonnement de l'apprenant

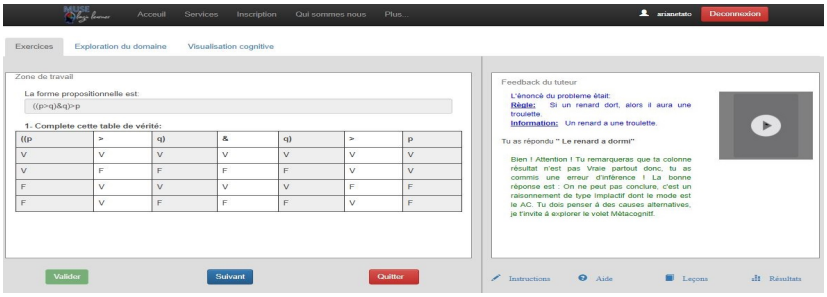


Fig. 6. Table de vérité remplie par l'étudiant et qui explique l'erreur d'inférence

Gérer les erreurs de raisonnement dans Muse-Logique. Les erreurs systématiques dans la logique classique et dans la logique probabiliste sont identifiées dans la littérature; les erreurs dans d'autres systèmes seront recueillies par des expériences. Dans Muse-Logique, les expériences sur les erreurs de raisonnement sont introduites à travers une variété d'activités de raisonnement, y compris syllogismes et polysyllogismes. Dans certains d'entre eux, l'apprenant suggère sa réponse en choisissant dans une liste (celle étant valide ou celle se rapportant aux erreurs de type 1), le système vérifie (grâce à Muse-Expert) si la réponse suggérée est logiquement valide ou non et fournit des commentaires pertinents après un diagnostic cognitif. Certaines questions offrent un temps de réponse limité de quelques secondes, de sorte que le caractère spontané des réponses de type 1 sera plus facilement déterminé. D'autres invitent les participants à décrire brièvement la procédure utilisée pour répondre, de sorte que cela peut aider pour l'interprétation des résultats et peut révéler des différences de procédure entre les réponses de type 1 et de type 2.

Génération automatique de problèmes de raisonnement. Les problèmes de raisonnement de type syllogismes et polysyllogismes sont des problèmes typiques de résolution de problème en raisonnement logique car ils nécessitent la capacité de raisonner en conformité avec les règles d'inférence pertinentes pour trouver la solution

(parvenir à une conclusion logiquement valide). Un syllogisme catégorique est composé de trois parties, la prémisse majeure, la mineure et la conclusion. Exemple: ***Prémisse majeure****: Je vais choisir la soupe ou je vais choisir la salade;* ***Prémisse mineure****: Je ne vais pas choisir la soupe;* ***Conclusion****: Je vais choisir la salade*.

Selon le type de l'opérateur utilisé dans la prémisse majeure, un syllogisme pourrait être disjonctif, implicatif ou d'incompatibilité, soit trois types de raisonnements. L'exemple précédent est un syllogisme disjonctif. Dans ces syllogismes, la prémisse majeure peut toujours être considérée comme un antécédent (A) liée à un conséquent (C) par un opérateur logique. En outre, la prémisse mineure est une déclaration spécifique, qui est soit une affirmation de A ou C, ou la négation de A ou C. Chaque cas correspond à une seule règle d'inférence : Par exemple, la règle d'inférence valide associée à un syllogisme est un modus ponens ponendo (MPP) sur le conjonctif conditionnel si la prémisse mineure est l'affirmation de A. Par ailleurs, la conclusion est soit l'affirmation ou la négation de la partie de la prémisse qui n'est pas utilisée dans la mineure. Sur cette base, la prémisse majeure est suffisante pour générer un problème de syllogisme. Donc, notre banque de problèmes contient un stock de prémisses majeures auquel on ajoute des alternatives éventuelles pour A ou C. Pour chaque problème de syllogisme, le système fournit à l'apprenant trois choix possibles. Les deux premiers choix sont l'affirmation et la négation de la partie de la prémisse majeure qui n'est pas présente dans la mineure. Le troisième choix est "*Vous ne pouvez pas conclure*". Un de ces trois choix de réponses est toujours l'erreur spontanée typique de Type 1. Notre algorithme est capable de détecter automatiquement ce choix ainsi que la bonne réponse. Le système tient compte de l'évolution de l'étudiant dans ses apprentissages (état du réseau bayesien) et de la stratégie pédagogique en jeu (règles du tuteur sur l'enchaînement des questions) pour la génération du prochain problème. Les problèmes générés sont situés dans une des 6 classes de situations identifiées dans la table 1.

Conclusion

Le travail qui est présenté dans cet article est issu d'une initiative multidisciplinaire qui nous a permis de nous inscrire dans une perspective de conception participative ayant mené à la co-construction d'un ensemble de composants validés par les experts du raisonnement logique et de la psychologie du raisonnement avant leur implémentation informatique. Cet article avait pour but de partager cette expérience unique avec le lecteur. Nous avons présenté les éléments qui nous ont motivés à proposer un tel système, puis nous avons présenté la démarche entreprise pour définir et expliciter les éléments de référence que nous avons utilisés. Par ailleurs, un aperçu des différents modules et leurs composants a été proposé. Les fonctionnalités du système implémenté ont été décrites et illustrées. Une première expérimentation a permis de collecter des données pour la construction de la composante psychométrique du modèle de l'apprenant, utilisé entre autres pour initialiser celui-ci et pour valider et renforcer le réseau bayésien de l'apprenant.

Muse-Logique sera déployé cet automne dans le cadre d'un cours de logique offert aux étudiants de première année à l'Université de Québec à Montréal. Ce déploiement

nous permettra de collecter des données empiriques afin de mieux mesurer l'apport du système en termes de gain d'apprentissage, et l'ajuster en conséquence. Une fois ce rodage complété, nous itérerons alors sur les autres logiques et implémenterons les méta-connaissances de choix de logiques afin de conférer à Muse-Logique, l'intelligence de passer d'un système logique à une autre selon la situation.

References

1. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. Int. J. Digit. Libr. 1 (1997) 108–121
2. Barnes, T., et Stamper, J. (2010). Automatic Hint Generation for Logic Proof Tutoring Using Historical Data. Educational Technology & Society, 13 (1), 3–12.
3. Brisson, J., de Chantal, P-L., Lortie-Forgues, H. & Markovits, H (2014). Belief bias is stronger when reasoning is more difficult, Thinking & Reasoning.
4. Cloutier, A. (2016). Métacognition, Raisonnement Logique et Philosophie pour enfants. (Mémoire de maîtrise). Université du Québec à Montréal.
5. Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. Cognition, 31(3), 187-276.
6. Evans, J.S.B.T., Newstead, S.E. & Byrne, R.M.J. (1993), Human Reasoning, Hove U. K., Lawrence Erlbaum Publishers.
7. Girotto, V., Gilly, M., Blaye, A. & Light, P. (1989).Children's performance in the sélection task: Plausibility and familiarity. British Journal of Psychology, 80(1), 79-95
8. Lesta, L. et K. Yacef (2002). "An Intelligent Teaching-Assistant System for Logic". Proceedings of International Conference on Intelligent Tutoring Systems (ITS'02) S. Cerri et F. Paraguo (eds) Biarritz, France, Springer-Verlag.
9. Markovits H. (2014). On the road toward formal reasoning: Reasoning with factual causal and contrary-to-fact causal premises during early adolescence. Journal of Experimental Child Psychology, 128, 37-51.
10. Markovits, H., Brunet, M.-L., Thompson, V., & Brisson, J. (2013). Direct evidence for a dual process model of deductive inference. Journal of Experimental Psychology: Learning, Memory, and Cognition, 39(4), 1213.
11. Moshman, D. (2004). From inference to reasoning: The construction of rationality David Thinking & Reasoning 10(2), 221–239
12. Nkambou, R., Bourdeau, J. et Mizoguchi, R. (Eds.), (2010). Advances in Intelligent Tutoring Systems. Springer, Studies in Computational Intelligence, Volume 308, 510p.
13. Robert, S. (2005). Categorization, reasoning and memory. In Cohen, H. & Lefebvre, C. (Eds). Handbook of Categorization in Cognitive Science, pp. 699-717 Elsevier.
14. Robert, S., Brisson, J. (2016). The Klein Group, Squares of Opposition and the Explanation of Fallacies in Reasoning, Logica Universalis 10(2-3), 377-389. Springer International Publishing.
15. Rossi, S. et van der Henst (dir.) (2007). Psychologies du raisonnement, Bruxelles: de Boeck.
16. Rupp, A., Templin, J. & Henson, R. (2010). Diagnostic measurement: Theory, methods and applications. New York: Guilford Press
17. Stanovich, K. E. (2011). Rationality and the reflexive mind. Oxford: Oxford Univ. Press.
18. Tchetagni, J., Nkambou, R. et Bourdeau, J. (2007). Explicit Reflection in Prolog-Tutor, International Journal on Artificial Intelligence in Education, 17 (2), pp. 169-217.
19. Varin, C. (2007). L'Enseignement du raisonnement conditionnel : De la logique aux neurosciences. (Mémoire de maîtrise), UQAM (p. 22-71).

Modèle et outil pour assister la scénarisation des activités pédagogiques orientées RV

Oussema Mahdi, Lahcen Oubahssi, Claudine Piau-Toffolon et Sébastien Iksal

Le Mans Université, LIUM, EA 4023, Laboratoire d'Informatique de l'Université du Mans, Avenue Messiaen, 72085 LE MANS CEDEX 9, France.
{oussema.mahdi, lahcen.oubahssi, claudine.piau-toffolon, sebastien.iksal}@univ-lemans.fr

Résumé. L'apprentissage humain est devenu une discipline émergente pour la réalité virtuelle. Dans ce cadre, nous nous intéressons aux EVAH qui visent à mettre l'apprenant en situation pédagogique dans un environnement de réalité virtuelle. Nous avons constaté, d'après la littérature, que très souvent les EVAH existants dépendent d'un domaine ou d'un contexte particulier et qu'ils ne permettent pas aux enseignants de définir ou d'adapter leurs modèles de scénarios en fonction des nouvelles situations pédagogiques qu'ils pourraient imaginer. Pour aider les enseignants à concevoir et générer des EVAH adaptés à leurs besoins, notre approche vise à définir un processus de conception et de production d'EVAH qui puisse être instancié dans différentes situations pédagogiques. Nos contributions portent sur la définition d'un modèle de scénario et le développement d'un éditeur permettant la spécification de scénarios et des activités pédagogiques basés sur des objets pédagogiques orientés RV.

Mots-clés. EVAH, EIAH, activité pédagogique, scénario pédagogique

Abstract. Human learning has become an emerging discipline for virtual reality. In this context, we are interested in VRLE, which aims to put the learner in a pedagogical situation in a virtual reality environment. We have found from the literature that existing VRLEs are dependent on a particular field or context and do not allow teachers to define or adapt their scenario models to new pedagogical situations they might imagine. To help teachers in designing and generating VRLE adapted to their needs, our approach aims at defining a process for the design and production of VRLE that can be instantiated in different pedagogical situations. Our contributions focus on the definition of a scenario model and the development of an editor allowing the specification of scenarios and pedagogical activities based on VR-oriented pedagogical objects.

Keywords. VRLE, TEL, pedagogical activity, learning scenario

1 Introduction

La Réalité Virtuelle (RV) est devenue un champ disciplinaire à part entière, offrant à l'utilisateur une expérience immersive exceptionnelle. Un Environnement Virtuel (EV) est représenté par un modèle 3D de données réelles ou imaginaires qui offrent des possibilités d'interaction et d'immersion toujours plus performantes [1]. Ces

possibilités trouvent un grand intérêt dans le domaine de l'apprentissage en permettant de créer des situations d'apprentissage originales et dynamiques, détachées des contraintes qui peuvent exister lors de formations réelles (danger, coût, incertitude) et apportent des avantages spécifiques (enrichissement des situations, rejeu, etc.) [2]. L'ensemble de ces dispositifs d'apprentissage utilisant les techniques de la réalité virtuelle peuvent être regroupés sous l'acronyme EVAH (Environnements de réalité Virtuelle pour l'Apprentissage Humain). Mikropoulos, et Natsis [3] ont défini un EVAH comme un «*environnement virtuel basé sur un modèle pédagogique particulier, intégrant ou impliquant un ou plusieurs objectifs didactiques, offrant aux utilisateurs des expériences qu'ils ne pourraient vivre autrement dans le monde réel et limitant des résultats d'apprentissage spécifiques*». Cependant, la conception et l'intégration des situations d'apprentissage dans un EVAH sont des tâches à la fois complexes et coûteuses. Les difficultés peuvent être d'ordre technique, induites par l'interdisciplinarité intrinsèque à la RV ou encore d'ordre cognitif, héritées des EIAH [4] [5]. Nous avons constaté, d'après la littérature, que les EVAH existants dépendent très souvent d'un domaine ou d'un contexte particulier, ils ne permettent pas aux enseignants de définir ou d'adapter facilement leurs modèles de scénarios en fonction des nouvelles situations pédagogiques qu'ils pourraient imaginer. Cet article présente un travail de recherche visant à proposer des solutions pour aider les enseignants à concevoir, réutiliser et déployer leurs scénarios pédagogiques dans les EVAH. Notre objectif est de proposer des solutions techniques et méthodologiques, suffisamment génériques pour être applicables à plusieurs environnements, quel que soit le domaine ou le type de tâche à réaliser. Nos questions de recherches sont relatives à la conception des situations pédagogiques orientées RV, et à la production des EVAH qui répondent mieux aux besoins des enseignants. Il s'agit ainsi de répondre aux questions suivantes : Comment aider les enseignants à formaliser leurs besoins pédagogiques à l'aide de solutions de conception éprouvées, réutilisables et adaptables et les opérationnaliser/déployer dans un environnement cible tout en respectant les intentions pédagogiques et en limitant les pertes sémantiques ? Quelle architecture utiliser pour définir des services permettant de réutiliser/adapter des environnements 3D existants ? Comment assurer l'interopérabilité des différents environnements 3D ? Comment faire face aux limites de compatibilité des composants techniques. Les pratiques de conception des enseignants sont itératives et participatives, une réflexion peut avoir lieu avant, pendant et après la mise en œuvre d'une situation pédagogique [6] ? Nous proposons ainsi une approche de conception itérative centrée sur l'enseignant. Cet article est structuré en quatre parties. Nous commençons par dresser un état de l'art des travaux et des outils orientés RV et pédagogie. Ensuite, nous décrivons notre processus de conception et de production d'EVAH. Les parties qui suivent présentent nos contributions qui portent sur la définition d'un modèle de scénario et le développement d'un éditeur permettant la spécification de scénarios et des activités pédagogiques basés sur des objets pédagogiques orientés RV.

2 EVAH : une revue de littérature

2.1 Modèles de conception et de production des EVAH

La conception et l'intégration des EVAH dans une formation sont des processus complexes et coûteux [2]. Le processus de conception et de production d'un EVAH doit prendre en compte les exigences pédagogiques des enseignants afin de répondre à leurs besoins. Pour produire un EVAH, l'approche la plus utilisée, consiste à résoudre les contraintes techniques avant d'apporter des solutions aux exigences pédagogiques des enseignants. Par exemple, Trinh et al. [7] fournissent des modèles d'exploitation des connaissances pour les agents virtuels peuplant des environnements virtuels. Ces connaissances portent sur la structure et la dynamique de l'environnement, ainsi que sur les procédures que les équipes peuvent exécuter dans cet environnement. Mellet d'Huart [8] propose un modèle basé sur les neurosciences, il vise une problématique de conception des EVAH dans des contextes quasi-industriels avec une démarche heuristique. Cette approche s'appuie sur des savoirs issus de recherches fondamentales dans différents domaines (en particulier en science cognitive, en biologie de la cognition et en neurophysiologie de l'action). Son modèle est théorique (sous forme d'une méthodologie), il n'a pas été validé/expérimenté et il n'est pas opérationnalisable. Chen et al. [9] proposent un cadre théorique pour orienter la conception d'un EVAH. Ce cadre est structuré en deux parties : une partie "macro-stratégie" qui fait référence à la conception globale des EVAH (Identification d'objectifs d'apprentissage, identification des scénarios pédagogiques, etc.) et une partie "micro-stratégie" qui fait référence à l'adaptation des scénarios pédagogiques en fonction du type d'EVAH que l'enseignant souhaite concevoir. Chen et Teh [10] proposent quelques améliorations dans ce cadre, à l'aide de la recherche formative, du modèle de conception pédagogique de l'environnement virtuel proposé par Chen et al. [9] Leur modèle permet de développer et d'évaluer de manière formative les simulations sur un système virtuel non immersif. Ritz [11] fournit des lignes directrices pour les meilleures pratiques d'intégration de la réalité virtuelle immersive, en particulier de CAVE (Cave Automatic Virtual Environment), dans l'enseignement. Ces directives répondent à un besoin concret en informant et en aidant les enseignants à adapter la conception pédagogique aux technologies émergentes. Sur la base de l'étude de ces travaux de recherche, nous avons constaté que les limites de ces propositions sont liées à la difficulté de mise en œuvre de leurs modèles de conception et à l'absence ou l'insuffisance de définition de modèles adaptables et réutilisables par les enseignants non-informaticiens dans des contextes différents afin de faciliter la conception et la production d'un EVAH [12].

2.2 Modèles de scénarios dans les EVAH

Nombreux sont les travaux qui ont abordé la question de la conception des situations pédagogiques dans des environnements virtuels. Marion et al [4] proposent un modèle de scénario d'apprentissage, POSEIDON dont l'objectif est de concevoir des activités pédagogiques orientées RV. Il permet de décrire les composantes d'un scénario pédagogique comportant des activités en environnement virtuel. Pour décrire ces

activités, chaque scénario POSEIDON s'appuie sur un modèle métier explicite et indépendant de l'environnement d'apprentissage. Ce modèle décrit à la fois les caractéristiques de l'environnement virtuel (entités, activités, etc.) et les concepts entrant en jeu dans l'apprentissage. L'objectif de ce travail est d'assurer la généricité de la modélisation, quels que soient la nature ou le domaine de l'EVAH en se reposant sur MASCARET [13]. Il s'agit d'un méta-modèle d'environnement virtuel qui fournit une représentation abstraite de la structure et du domaine des environnements qui permet la description des activités pédagogiques. Il a recours à différents niveaux de modélisation afin de décrire respectivement les concepts d'un environnement, sa dynamique et les activités qu'il est possible d'y réaliser. MASCARET a été utilisé aussi pour assurer le côté générique du modèle de scénario pédagogique POSVET [14], ce dernier permet de réutiliser des scénarios pédagogiques sur différentes plateformes. Ce travail offre une base intéressante pour adapter et générer les scénarios pédagogiques orientés réalité virtuelle, mais n'offre pas de solutions pour aider les enseignants dans leur processus de conception. Chen et al. [9] proposent un cadre théorique qui identifie quatre principes pour concevoir les scénarios pédagogiques: 1) le principe conceptuel qui guide l'apprenant vers l'information qu'il doit prendre en compte; 2) le principe de métacognition qui guide le processus de réflexion de l'apprenant pendant l'apprentissage; 3) le principe de procédure qui indique comment utiliser les informations disponibles dans l'environnement virtuel; 4) le principe "stratégique" qui permet à l'apprenant d'analyser la tâche d'apprentissage ou le problème à résoudre. Le Corre et al. [15] ont proposé un ITS appelé CHRYSAOR basé sur POSEIDON en réponse à certaines faiblesses de l'ITS (Intelligent Tutorial System) PEGASE [16] telles que le manque de lien avec le scénario pédagogique ou le manque de modularité et de personnalisation. Carpentier et al. [5] et Trinh et al. [7] proposent des modèles basés sur un contrôle centralisé et indirect d'une simulation émergente à partir d'un modèle de contenu de scénario d'apprentissage. Dans l'approche de [5], l'environnement est peuplé de personnages virtuels autonomes et l'utilisateur est libre de ses actions. La conception du scénario d'apprentissage est réalisée en deux étapes : les objectifs dynamiques sont déterminés à partir de l'activité de l'utilisateur, puis un scénario d'apprentissage est généré par ces objectifs et mis en œuvre par le biais d'ajustements de simulation. Nous constatons que ces modèles de scénarios ne peuvent pas être facilement adaptés en fonction des nouvelles situations pédagogiques que l'enseignant pourra imaginer contrairement à leurs pratiques [17].

2.3 Architecture fonctionnelle et technique pour produire un EVAH

Parmi les travaux qui ont abordé la question de l'architecture fonctionnelle et technique des EVAH, nous avons retenu ceux de Lanquepin et al. [18] qui proposent une plate-forme nommée HUMANS (Human Models based Artificial eNvironments Software). HUMANS est générique et produit des environnements avec des AVATAR, il est destiné aux experts et non aux enseignants qui cherchent à produire des situations pédagogiques virtuelles (avec ou sans AVATAR). HUMANS offre via son module SELDON une approche de scénario adaptatif. Le modèle SELDON est extrinsèque, ce qui signifie que le scénario est considéré comme une étape supplémentaire dans le cadrage d'un environnement virtuel existant, et non comme une partie intégrante du

processus de conception pour cet environnement. L'objectif de l'équipe qui a travaillé sur le système HUMANS est de modéliser l'activité humaine et sociale et de proposer une scénarisation adaptative, qui guide dynamiquement le scénario et les comportements de personnages virtuels autonomes dans l'environnement virtuel vers des contenus personnalisés. En revanche, l'objectif de notre travail est d'aider et accompagner les enseignants qui cherchent à modéliser et opérationnaliser des situations pédagogiques virtuelles. Le projet GVT (Generic Virtual Training) [19] a pour objectif de développer une plateforme qui permet de produire des activités pédagogiques de type procédures de maintenance. Cette plateforme repose sur des métaphores visuelles. Ce concept est important, car il met l'accent sur les interactions avec les objets à l'aide d'un menu d'icônes représentant les interactions possibles entre l'objet et l'utilisateur. La génération des scénarios décrivant le déroulement de ces activités dans un ordre bien défini s'avère une tâche compliquée dans GVT. Pour cela, Mollet et Arnaldi [20] ont eu recours au langage LORA. Ce langage permet la création et l'édition des scénarios. Chaque scénario est composé d'états (ou étapes) représentant les actions et des liens entre elles (les transitions). GVT distingue le scénario d'activité qui décrit les procédures à réaliser dans l'environnement, et le scénario pédagogique qui favorise la réutilisabilité des scénarios existants. Une limite de GVT selon Marion [12] est que, du fait du contexte industriel dans lequel s'inscrit le projet, il ne peut être utilisé que pour l'apprentissage de procédures [12] difficilement adaptable pour d'autres contextes.

L'analyse de ces architectures fonctionnelles et techniques pour produire un EVAH nous a permis de constater qu'elles ont été développées pour des domaines particuliers, elles ne traitent pas les problèmes de la conception (adaptation ou réutilisation) et de l'opérationnalisation des modèles de scénario directement par les enseignants en fonction de leurs situations pédagogiques.

3 Proposition d'un processus de conception orientée RV

Dans le but d'aider les enseignants à produire ou à générer un EVAH, nous proposons une solution méthodologique basée sur un processus de conception qui comporte plusieurs étapes allant de la définition de la situation d'apprentissage jusqu'à son opérationnalisation. Ce processus a été défini à partir de processus existants en EIAH itératif, centré-utilisateur et adapté à la spécificité des EVAH [21] [22]. Dans une première étape, les enseignants expriment leurs besoins pédagogiques. Puis ils sont amenés à les formaliser en fonction de leurs situations d'apprentissage. Notre proposition de modèle de scénario orienté RV (section 4) permet d'offrir un cadre théorique et de structurer les activités RV. Il permet ainsi de répondre à plusieurs questions concernant la scénarisation pédagogique orientée RV (citées début de cet article). La deuxième étape du processus consiste à identifier et à adapter l'environnement 3D dans lequel sera instancié le scénario formalisé ainsi que les outils de RV nécessaires. Il s'agit ici d'un service, qui permet la réutilisation et l'adaptation des environnements 3D existants pour les rendre compatibles avec les situations voulues par l'enseignant. La troisième étape du processus consiste à opérationnaliser des scénarios sur un ou plusieurs environnements 3D. Le service d'adaptation et l'API

de l'opérationnalisation ne seront pas détaillés dans cet article, ils feront l'objet d'une future publication. La quatrième étape permet de simuler et tester l'EVAH généré. Précédé d'une étape d'apprentissage, le processus recourt à une phase d'adaptation des situations modélisées selon les résultats obtenus. Notons que la revue de la littérature nous a permis de faire ressortir deux constats : d'une part le besoin d'un modèle de scénario spécifique aux scénarios orientés réalité virtuelle et d'un outil capable de fournir les éléments nécessaires à la modélisation de telles situations pédagogiques [6]. Ainsi nous proposons deux contributions pour la phase de conception du processus, un modèle et un éditeur de scénarios pédagogiques orientés réalité virtuelle (sections 4 et 5). Nous présentons dans la partie suivante les premières étapes du processus de conception qui visent à produire des scénarios pédagogiques orientés RV.

4 Modèle de scénario pédagogique orienté RV

Pour [23], la scénarisation pédagogique consiste à spécifier et modéliser les situations pédagogiques. C'est « *avant tout un travail de conception de contenu, d'organisation des ressources, de planification de l'activité et des médiations pour induire et accompagner l'apprentissage, et d'orchestration* » tenant compte de l'approche pédagogique suivie. Un scénario pédagogique permet de structurer le contexte de l'apprentissage et de l'organiser dans l'environnement virtuel et dans le temps. L'activité pédagogique et l'objet pédagogique sont deux composants indispensables dans le scénario. L'activité définit les modalités précises d'acquisition, de validation, de communication d'une ou plusieurs connaissances et l'objet permet la réalisation de ces activités [24]. Avant de présenter notre modèle de scénario, nous allons détailler ces deux concepts, car ils représentent la base de notre modèle de scénario.

4.1 Activité pédagogique orientée RV

Dans le domaine des EVAH, la situation pédagogique est considérée comme une composition d'activités réalisées par un ensemble d'acteurs dans un environnement donné. Marion [12] définit une activité pédagogique orientée réalité virtuelle comme étant la tâche à réaliser par l'apprenant dans l'environnement virtuel (le quoi), décrite par l'enseignant. À cette description, ce dernier peut ajouter des informations pédagogiques ayant pour objectif d'instrumenter cette tâche dans un but pédagogique précis (le comment). L'EVAH est caractérisé (notamment par Roussou [25]) par deux éléments indissociables : l'immersion dans un monde virtuel et l'interaction avec des objets 3D modélisés (objets pédagogiques, avatars). Le déroulement des activités est décrit sous une forme hiérarchique. Dans notre modèle, les activités peuvent être organisées de manière séquentielle ou parallèle.
Comme illustré dans la figure 1, une activité est caractérisée par un ensemble de prérequis et d'objectifs pédagogiques que l'enseignant souhaite atteindre. Chaque activité peut être découpée en une séquence d'actions. Nous avons regroupé ces actions en quatre catégories que nous appelons actions de type "Primitives Comportementales virtuelles" (PCV ou VBP) [26] [27] : (a) ***Observer*** le monde virtuel : permet d'obtenir une liste des objets qui apparaissent dans le champ de vision de l'utilisateur ainsi que

leurs distances et leurs orientations par rapport à l'utilisateur ; (b) ***Se déplacer*** dans le monde virtuel : permet d'obtenir l'orientation moyenne, de savoir quel est l'objet vers lequel l'utilisateur se dirige globalement, de connaître la vitesse instantanée de l'utilisateur, de définir des zones et de savoir à quelles distances de ces zones l'utilisateur se trouve. Les déplacements peuvent être des déplacements en transition ou en rotation ; (c) ***Agir*** sur le monde virtuel : permet d'intercepter les actions sur les objets de l'environnement. Les actions sur les objets peuvent être sous forme de sélection ou bien de manipulation ; (d) ***Communiquer*** avec autrui ou avec l'application : permet de détecter l'intention et de présenter l'information à l'apprenant soit en communiquant avec d'autres utilisateurs, avec des personnages virtuels ou bien en communiquant avec l'application pour modifier sa configuration, pour donner des ordres au système informatique, etc.

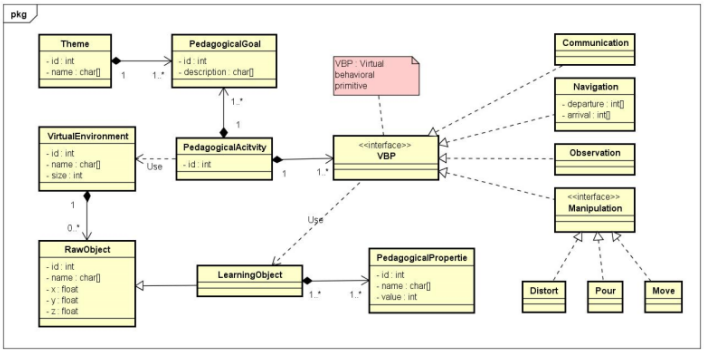


Fig 1. Modèle d'une activité pédagogique orientée réalité virtuelle

4.2 Objet pédagogique orienté RV

Dans certains travaux de recherche, un objet pédagogique est présenté avec le terme "*ressource pédagogique*" ou encore "*objet d'apprentissage*" (LO) [24] [28]. L'objet pédagogique est considéré comme un ensemble d'informations assemblées pour atteindre un objectif d'apprentissage. Il est conçu par l'enseignant et peut adopter des formes variées [24]. Pour notre approche, nous nous intéressons à la notion d'objet pédagogique orientée réalité virtuelle. Un objet pédagogique orienté réalité virtuelle est présenté sous la forme d'un objet brut possédant des propriétés pédagogiques et techniques. Un objet brut, appelé aussi objet 3D ou objet graphique, est une entité d'acquisition de connaissances. Quant aux propriétés, elles sont utilisées pour stocker des valeurs associées à ces objets. Certaines propriétés techniques sont communes à tous les objets (notamment ceux qui régissent la position, la forme ou la couleur des objets), tandis que d'autres sont des propriétés spécifiques à l'objet ou au domaine de l'apprentissage. Prenons l'exemple d'une solution H2O comme objet pédagogique. Elle possède des propriétés telles que la concentration, le volume et la température de vaporisation. Nous nous proposons dans le cadre de ces travaux de recherche de définir une plateforme d'objets pédagogiques virtuels. L'objectif visé est d'avoir un environnement avec des règles qui permettent de décrire le comportement dynamique

des objets bruts ainsi que leurs propriétés pédagogiques. Ces règles définissent la valeur des propriétés de l'objet en fonction des actions subies sur l'objet ou sur l'environnement où il sera utilisé. Par exemple, un cube (objet brut) doit avoir les propriétés techniques poids et position et si on le lâche, il tombe et se déforme. On pourra lui associer des propriétés pédagogiques liées à la gravitation pour une utilisation dans un contexte pédagogique comme un cours de physique. L'objectif d'avoir une plateforme d'objets pédagogiques orientés RV est d'assurer leur réutilisation dans différentes situations indépendamment du contexte d'apprentissage.

4.3 Modèle de scénario pédagogique orienté RV

Notre proposition d'un modèle de scénario pédagogique a été élaborée à partir de l'analyse théorique des différents modèles de scénarios existants dans les domaines des EIAH et EVAH, et de la conception de trois exemples de situations pédagogiques différentes. Nous illustrons notre proposition à l'aide d'un exemple dans la section suivante. Un de nos objectifs est de développer un éditeur de scénarios pédagogiques RV qui embarque notre modèle permettant aux enseignants de modéliser et d'adapter facilement leurs situations (sous forme de scénarios) et de générer leur propre EVAH.
D'après Marion, un scénario pédagogique orienté réalité virtuelle décrit l'organisation des activités pédagogiques dans un environnement virtuel, les objectifs pédagogiques qui leur sont associés — en termes de connaissances ou compétences —, les prérequis, les rôles des différents intervenants dans la situation pédagogique — qu'ils soient enseignant ou apprenant — et les outils et ressources nécessaires à la réalisation de ces activités virtuelles (objets virtuels (bruts ou pédagogiques)) [12].
Marion adopte une approche qui fait le lien entre la description de l'activité d'apprentissage, telle que décrite avec IMS-LD sous forme d'unités d'apprentissage (unit of learning), et l'activité de l'apprenant dans l'environnement virtuel. Il faut noter qu'IMS-LD a ses propres limites, car il est complexe et manque d'expressivité concernant la description des interactions entre les utilisateurs dans les tâches virtuelles ou non virtuelles, les obligeant à recourir à un ingénieur pédagogique. Quant à notre approche, une formalisation des scénarios par patrons est vue comme une solution aux problèmes liés aux difficultés d'expressivité que pose ce type de standard. L'application de l'approche patron reste un des objectifs à long terme de ce travail de recherche, nous souhaiterions dans un premier temps valider la faisabilité de notre processus de modélisation et l'opérationnalisation des situations pédagogiques virtuelles. La figure 2 illustre notre proposition de modèle de scénario pédagogique orienté RV. Ce modèle a été établi à partir des différents concepts nécessaires à la description d'une activité pédagogique en environnement virtuel. Parmi les spécificités de notre proposition, les moyens d'interaction (outils de RV), les types d'interactions (Primitives Comportementales virtuelles : PCV) et l'utilisation de la notion d'objet pédagogique virtuel. Notre proposition est un enrichissement des modèles existants par une description fine d'activités (utilisation du concept PVC) où les apprenants interagissent avec un dispositif afin de tenir compte des spécificités induites par le déroulement d'activités en environnement virtuel. Notre approche vise à créer un modèle qui permet de faire le lien entre la description de l'activité pédagogique, et l'activité de l'apprenant dans l'environnement virtuel. Le modèle de scénario doit garder le caractère générique

et flexible pour pouvoir s'adapter à des contextes de réalisation de scénario différents sans avoir à modifier sa description.

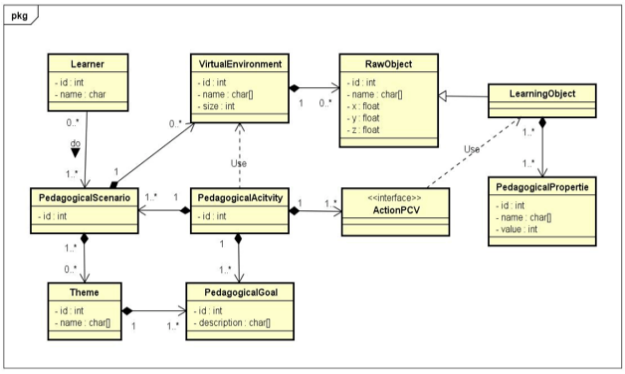


Fig. 2. Modèle d'un scénario pédagogique orienté RV

4.4 Illustration du modèle via un exemple

Pour élaborer notre modèle de scénario, nous avons travaillé sur la conception de trois situations pédagogiques différentes. Ce travail est réalisé selon le processus de conception présenté dans la partie 3, les différents enseignants participent à la conception des différents modèles de situations pédagogiques selon une approche itérative, centrée-utilisateur. La situation pédagogique retenue ici pour illustration concerne la "Réalisation d'un dosage volumétrique" pour un enseignement de chimie (Table 1). L'objectif pédagogique de cette situation est "*la maîtrise des étapes nécessaires à la conception du protocole permettant de déterminer la concentration d'une solution inconnue par dosage volumétrique*".

Table 1. Exemple d'activité pédagogique virtuelle (choix des réactifs)

Situation pédagogique RV	Réalisation d'un dosage volumétrique
Environnement virtuel	Laboratoire de chimie virtuel
Activité pédagogique RV	Choix des réactifs
Actions RV (PCV ou VBP)	Observer les réactifs
	Placer l'objet pédagogique "réactif NaOH" sur le chariot
	Placer l'objet pédagogique "réactif H2O" sur le chariot

L'enseignant a d'abord identifié des activités pédagogiques qui répondent à ses objectifs pédagogiques. L'apprenant doit commencer par choisir les réactifs nécessaires à la réalisation du dosage et ajuster la concentration des solutions et les volumes utilisés. Ensuite, il choisit le matériel nécessaire à la réalisation du dosage et les équipements individuels de protection. Finalement, la dernière activité pédagogique est la réalisation du dosage. La figure 3 illustre comment l'enseignant scénarise cette activité via l'éditeur proposé.

5 Editeur proposé

Dès lors que l'on se place dans un contexte interactif, la scénarisation pédagogique ne consiste plus uniquement en « *la modélisation d'un scénario* », mais également en « *la mise en place des mécanismes* » nécessaires à la réalisation de ce scénario [29]. Par ce fait, nous avons développé un éditeur de scénarios qui facilite la création des situations pédagogiques modélisées. L'objectif est de permettre la création et l'adaptation des différentes composantes d'un scénario pédagogique notamment les environnements d'apprentissage, les activités pédagogiques et leur organisation.

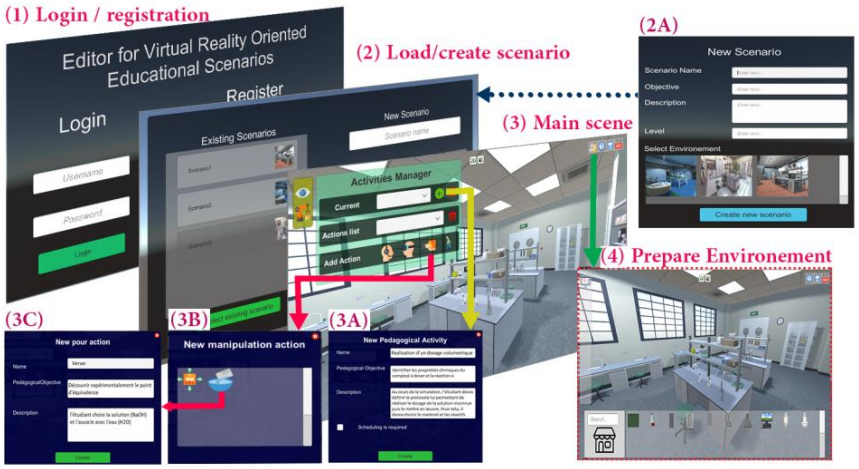


Fig. 3. Définition des actions de RV à réaliser au sein de l'activité pédagogique

La figure 3 illustre les interfaces principales de notre éditeur. Dans un premier temps, l'enseignant commence par la création d'un nouveau projet de scénario (fig.3 (2)). Ensuite, il choisit un environnement de RV adapté à sa situation pédagogique (fig.3 (2A)). Il faut noter que l'adaptation des environnements RV et des objets pédagogiques virtuels est réalisée au niveau de la plateforme d'objets pédagogiques virtuels. Nous rappelons que l'objectif de cette plateforme est d'avoir un environnement avec des règles qui permettent de décrire le comportement dynamique des objets bruts ainsi que leurs propriétés pédagogiques. Après avoir choisi son environnement, l'enseignant a la possibilité de choisir entre deux types de vue (2D ou 3D) (fig.3 (3) et (4)). Ensuite il peut sélectionner depuis l'inventaire les objets et les place dans son environnement (fig.3 (4)). Comme expliqué dans les parties précédentes chaque objet pédagogique possède une liste de propriétés, l'enseignant défini alors les valeurs attendues. À partir de là, commence l'étape de la définition du scénario dans laquelle l'enseignant décrit ses activités pédagogiques (fig.3 (3)). Ensuite, l'enseignant décrit pour chaque activité les actions souhaitées en se servant de la barre contenant la liste des actions possibles (fig.3 (3B) et (3C)). Les actions représentent l'ensemble des opérations accomplies par l'apprenant (par exemple : déplacer, verser, déformer, couper, etc.). Finalement, il gère l'ordre des actions pédagogiques dans une telle activité pédagogique et l'ordre du déroulement des activités pédagogiques au sein de l'environnement virtuel.

6 Conclusion

Notre travail porte sur la problématique de la conception et la production de situations pédagogiques dans des EVAH. À partir des questions de recherche formulées précédemment, nous avons constaté que dans les différents travaux existants, les EVAH ou les modèles d'architecture proposés soit, dépendaient d'un domaine et d'un contexte particulier, soit n'étaient pas assez complets pour pouvoir assister l'enseignant dans la description, l'adaptation ou la réutilisation des scénarios pédagogiques. Notre approche vise à proposer des solutions à la fois techniques et méthodologiques pour aider les enseignants. Dans un premier temps, nous avons élaboré un processus de conception et de production d'un EVAH. Dans ce papier, nous nous sommes intéressés à la première partie de ce processus dédiée à la conception de situation pédagogique orientée RV. Nous avons cherché dans un premier temps à apporter des solutions pour structurer les situations pédagogiques sous forme de modèles de scénarios réutilisables. Lors de cette étape, nous avons travaillé en partenariat avec des équipes pédagogiques. Nous avons modélisé les activités et les objets pédagogiques décrits dans la situation pédagogique et proposé un modèle de scénario et une première version d'un prototype auteur. Dans un second temps, nous souhaiterions proposer des outils et des services expérimentaux offrant des fonctionnalités complémentaires pour notre éditeur qui permettent d'intégrer, de réutiliser et d'adapter les modèles de scénarios pédagogiques orientés RV dans les nouveaux EVAH afin de faciliter la tâche de conception des enseignants. Des tests et des expérimentations sont à venir afin de mettre à l'épreuve le modèle et l'éditeur développés et montrer le caractère générique de notre approche.

Références

1. Cazeaux, E., Devillers, F., Saint-Romas, C., Arnaldi, B., Maffre, E., Mollet, N., Tisseau, J. (s.d.) (2005). Giat Virtual Training : formation à la maintenance. Dans Laval Virtual 2005. Laval.
2. Oubahssi, L., Piau-Toffolon, C. (2018). Virtual Learning Environment Design in the Context of Orientation Skills Acquisition for LUSI Class. 10th International Conference on Computer Supported Education, (pp. 47-58).
3. Mikropoulos, T. A., Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). Computers & Education, 56(3), 769-780.
4. Marion, N., Querrec, R., Chevaillier, P. (2009). Integrating Knowledge from Virtual Reality Environments to Learning Scenario Models-A Meta-modeling Approach, (pp. 254-259).
5. Carpentier, K., Lourdeaux, D. (2014). Generation of Learning Situations According to the Learner's Profile Within a Virtual Environment. Communications in Computer and Information Science, 245 - 260.
6. Bennett, S., Agostinho, S., Lockyer, L. (2017). The process of designing for learning: understanding university teachers design work. Educational Technology Research and Development, 65(1), 125–145. https://doi.org/10.1007/s11423-016-9469-y
7. Trinh, T.-H., Querrec, R., De Loor, P., Chevaillier, P. (2010). Ensuring semantic spatial constraints in virtual environments using UML/OCL VRST '10, 2010. 219-226.
8. Mellet-d'Huart, D. (2004). De l'intention à l'attention : contributions à une démarche de conception d'environnements virtuels pour apprendre à partir d'un modèle de l'(én)action. Thèse de doctorat en Informatique, Le Mans Université.

9. Chen, C. J., Toh, S. C., Fauzy, W. M. (2004). The Theoretical Framework for Designing Desktop Virtual Reality-Based Learning Environments. Journal of Interactive Learning Research, 15(Norfolk, VA: Association for the Advancement of Computing in Education (AACE)), 147-167.
10. Chen, C. J., Teh, C. S. (2013). Enhancing an Instructional Design Model for Virtual Reality-Based Learning. Australasian Journal of Educational Technology. 2013, 29(5)
11. Ritz, L. T. (2015). Teaching with CAVE virtual reality systems: Instructional design strategies that promote adequate cognitive load for learners. SMTC Plan B Papers. 5.
12. Marion, N. (2010). Modélisation de scénarios pédagogiques pour les environnements de réalité virtuelle d'apprentissage humain. Thèse de doctorat. Brest.
13. Buche, C., Querrec, R., De Loor, P., Chevaillier, P. (2004). Mascaret: Pedagogical multi-agents system for virtual environment for training. Journal of Distance Education Technologies, 4(2), 41-61.
14. Fahim, M., Jakimi, A., El Bermi, L. (2016). Pedagogical Scenarization for Virtual Environments for Training: Towards Genericity, Coherence and Adaptivity. International Journal of Advanced Engineering Research and Science, 3(12), 96-103.
15. Le Corre, F., Hoareau, C., Ganier, F., Buche, C., Querrec, R. (2014). A Pedagogical Scenario Language for Virtual Environment for Learning based on UML Meta-model. Application to Blood Analysis Instrument. Proceedings of the 6th International Conference on Computer Supported Education (CSEDU), (pp. 301-308). Spain.
16. Buche, C., Bossard, C., Querrec, R., Chevaillier, P. (2010). PEGASE: A generic and adaptable intelligent system for virtual reality learning environments. International Journal of Virtual Reality, 9(2), 73-85.
17. Goodyear, P. (2015). Teaching as design. HERDSA Review of Higher Education, 27-50.
18. Lanquepin, V., Carpentier, K., Lourdeaux, D., Lhommet, M., Barot, C., Amokrane, K. (2013). HUMANS: a HUman Models based Artificial eNvironments software platform. VRIC, (pp. 59-68). Laval, France.
19. Gerbaud, S., Mollet, N., Ganier, F., Arnaldi, B., Tisseau, J. (2008). GVT: a platform to create virtual environments for procedural training. Virtual Reality Conference, (pp. 476-481). Reno, NE, USA.
20. Mollet, N., Arnaldi, B. (2006). Storytelling in Virtual Reality for Training. Dans Z. Pan, R. Aylett, H. Diener, X. Jin, S. Gobel, L. Li, S. B. Heidelberg (Éd.), Technologies for E-Learning and Digital Entertainment (pp. 334-347). Berlin, Heidelberg.
21. Abedmouleh, A., Oubahssi, L., Laforcade, P., Choquet, C. (2012). An Analysis Process for Identifying and Formalizing LMS Instructional Language. ICSOFT (pp. 218-223). Rome, Italie.
22. Choquet, C. (2007). Ingénierie et réingénierie des EIAH : l'approche REDiM. In U. d. Mans (Ed.), HDR. Université du Mans.
23. Henri, F., Compte, C., Charlier, B. (2007). La scénarisation pédagogique dans tous ses débats. Revue Internationale des Technologies en Pédagogie Universitaire, 4, 14-24.
24. Pernin, J.-P. (2003). Objets pédagogiques : unités d'apprentissage, activités ou ressources ? Sciences et Techniques Educatives, Hors série, 179-210.
25. Roussou, M. (2004). Learning by doing and learning through play: An exploration of interactivity in virtual environments for children. Computers in Entertainment, 2, 10.
26. Fuchs, P., Moreau, G., Burkhardt, J.-M. (2006). Le traité de la réalité virtuelle Volume 2 - L'interfaçage : l'immersion et l'interaction en environnement virtuel. Presse des Mines.
27. Schlemminger, G., Roy, M., Veit, M., Capobianco, A., Noeppel, G. (2013). Réalité virtuelle et jeux : de nouveaux outils pour des apprentissages plurilingues ? Education et sociétés plurilingues (n° 35, décembre 2013), p. 29-41.
28. Learning Objects Metadata Working Group (2001).
29. Barot, C. (2014). Scénarisation d'environnements virtuels. Vers un équilibre entre contrôle, cohérence et adaptabilité. https://tel.archives-ouvertes.fr/tel-00980537

Session
Analyse d'usage

Ingénierie d'un EIAH doté d'étayages adaptatifs

Anne Lejeune[1], Catherine Bonnat[1], Viviane Guéraud[1], Patricia Marzin-Janvier[2,1]

[1]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, F-38000 Grenoble France
anne.lejeune, catherine.bonnat, viviane.gueraud @univ-grenoble-alpes.fr
[2]Univ. Brest, Univ. Rennes, CREAD, F-29200 Brest, France
patricia.marzin-janvier@espe-bretagne.fr

Résumé. Le projet DIABOLO[1] a confronté des didacticiens et des informaticiens au défi de concevoir et expérimenter, en un an, un EIAH doté de rétroactions dépendant du diagnostic des difficultés rencontrées par les élèves. La tâche d'apprentissage, au programme de SVT au lycée en France, concerne l'étude d'un métabolisme fermentaire. Elle avait fait l'objet d'une modélisation didactique qu'il s'agissait de transposer informatiquement afin d'analyser l'impact des étayages proposés. Nous explorons les solutions d'ingénierie susceptibles de satisfaire le cahier des charges puis présentons l'EIAH conçu.

Mots-clés. Ingénierie des EIAH, diagnostic, étayage, traces

Abstract. DIABOLO project has confronted didactic and computer scientists with the challenge of designing catand testing, within a year, an EIAH providing dynamic feedback when student's difficulties are diagnosed. The learning task, part of the SVT program in French high school, is centred on the study of a fermentative metabolism. It had been modelled by didactic scientists and was to be computationally transposed in order to evaluate the feedbacks' impact. We browse engineering solutions that may satisfy the specifications and then present the resulting EIAH.

Keywords. EIAH design, diagnostic, scaffolding, traces

1 Introduction

DIABOLO (projet EMERGENCE 2018), a été le cadre du croisement de nos travaux en didactique des sciences et en ingénierie des EIAH pour concevoir et expérimenter - *sur un an* - un EIAH fondé sur une modélisation didactique préalable.

L'activité d'apprentissage concerne la conception d'un protocole sur le métabolisme fermentaire, étudié en SVT au lycée (B.O. spécial n° 8 du 13 octobre 2011). Dans des travaux de thèse préalables au projet Diabolo, Bonnat [1] a effectué la modélisation didactique de la tâche et des connaissances en jeu, en prenant pour cadre le modèle T4TEL [2] - extension de l'approche praxéologique [3]. S'intéressant aux difficultés méthodologiques et épistémologiques rencontrées par les élèves dans l'écriture d'un protocole [4] [5], Bonnat a étudié les étayages pouvant être mis en

1 DIAgnostic des connaissances en Biologie fondé sur une mOdéLisation didactique et infOrmatique - Projet EMERGENCE 2018

place dans un EIAH pour y remédier. Elle a ainsi catégorisé ces difficultés en fonction des concepts auxquels elles se rapportaient, qu'ils soient objet d'apprentissage au lycée (concepts OA), ou qu'ils ne le soient pas (concepts NOA). Conformément au modèle proposé par Croset et Chaachoua [6], elle a ensuite modélisé les techniques erronées, potentiellement sources des erreurs tenaces relatives aux concepts OA. Puis elle a défini les rétroactions à fournir suite au diagnostic de ces erreurs et a classé ces rétroactions par niveaux, caractérisés par le type (question ou information) et la nature (technique ou technologie) du message délivré à l'apprenant.

Donner suite à ces travaux de thèse, nécessitait de disposer d'un EIAH transposant informatiquement tous les volets de la modélisation didactique. Cet EIAH devait générer des données permettant de répondre à la question de recherche : *Quel est l'impact sur les protocoles conçus par les élèves, de l'étayage dynamique de leurs erreurs relatives à des concepts OA ?* Cet EIAH devait également disposer de mécanismes de diagnostic automatique et être suffisamment flexible pour que les étayages mis en place puissent être adaptés suite à l'analyse des données recueillies.

Dans la section suivante, nous confrontons les propriétés requises par le cahier des charges, aux solutions informatiques permettant de les satisfaire. Puis nous présentons l'architecture et les composants de l'EIAH développé.

2 Ingénierie de l'EIAH produit

2.1 Propriétés requises et solutions envisageables

Transposition informatique de la modélisation didactique préalable. Selon la modélisation didactique préalable, la tâche d'apprentissage visée est constituée de 4 phases consécutives : (1) prise de connaissance du métabolisme étudié, (2) collecte du matériel expérimental, (3) conception du protocole et (4) étude de la réaction résultant de l'exécution (simulée) du protocole. Chaque phase fait référence à la conception expérimentale en laboratoire et pour chacune, les connaissances en jeu et le protocole « expert » ont été spécifiés, de même que les erreurs les plus fréquentes et les étayages susceptibles d'y remédier.

La modélisation didactique de l'activité cible convient tout à fait à sa *transposition informatique* dans une simulation pédagogique [7], [8].

Cependant, intégrer dans une simulation pédagogique les mécanismes nécessaires à la génération et à l'enregistrement des données permettant d'analyser l'efficience des étayages, susciterait des développements lourds. De plus, l'adaptation ultérieure de ces étayages nécessiterait sa réingénierie.

Disposer de données pertinentes sur l'activité de l'apprenant. Bien que l'utilisation du numérique génère automatiquement de nombreuses données, il est pour nous fondamental de décider, dès la conception d'un EIAH, ce qu'il convient de tracer et quand il faut le faire, ainsi que de spécifier les informations qui doivent être enregistrées pour que les traces générées soient porteuses de sens. Ne pas consentir à cet effort revient à laisser libre cours à l'interprétation abusive de ces données.

Plusieurs, dont Iksal [9], partagent ce point de vue. Dans le cadre du projet DIABOLO, pour que les chercheurs en didactique des sciences puissent analyser l'impact des étayages des erreurs relatives aux concepts OA, ils devaient disposer de données caractérisant le contexte de chaque étayage : le matériel rassemblé et le protocole conçu par l'élève ; ils devaient aussi disposer de données caractérisant l'activité de l'élève après étayage : documentation consultée, évolution du protocole, modification du matériel rassemblé. Un inventaire des informations à recueillir au fil de l'activité de l'élève devait ainsi être mené en amont du développement de l'EIAH.

Flexibilité. La *flexibilité* est une des premières qualités d'un EIAH construit selon une approche Learning Design [10]. Cette approche conduit à définir un objet informatique - que nous appelons « scénario » [11] - ayant la vocation d'être partageable et surtout *adaptable* sans nécessiter la réingénierie de l'EIAH. Le scénario modélise et structure l'unité d'apprentissage : il détermine et organise l'accès aux ressources pédagogiques mises à la disposition des utilisateurs, il contrôle les productions des apprenants tout au long de l'activité proposée et modélise les notifications (conseils ou feedbacks) pouvant leur être adressées en retour de ces productions. Notifier l'apprenant suite à ses productions relève de mécanismes de *diagnostic en temps réel*.

En conséquence, suivre une approche Learning Design présente l'avantage de pouvoir distinguer clairement à partir de la modélisation pédagogique préalable :

- ce qui relève des qualités attendues d'une simulation pédagogique dédiée à la tâche cible conformément à l'analyse didactique qui en a été faite ;
- ce qui doit pouvoir être facilement adaptable, soit l'étayage des erreurs relatives aux concepts OA, afin que l'EIAH conçu remplisse à terme sa mission de remédiation de ces erreurs.

Faire cette distinction revient (1) à considérer la simulation comme une ressource pédagogique mise à disposition de l'élève par le scénario de l'EIAH et, (2) à déléguer au scénario la modélisation des rétroactions délivrées à l'élève quand il fait une erreur relative à un concept OA lors de ses interactions avec la simulation. Cela suppose de fait, que le scénario connaisse les variables de la simulation permettant de caractériser son état au fil des interactions de l'élève.

2.2 Architecture et composants de l'EIAH construit

Nous avons appliqué les solutions étudiées dans la section précédente en utilisant les outils logiciels dont nous disposions et en concentrant l'effort de développement sur la production d'une simulation pédagogique dédiée à la tâche cible. La figure 1 décrit l'architecture et les composants de l'EIAH proposé.

SimDiabolo : simulation implémentant la modélisation didactique à l'exception des étayages relatifs aux concepts OA. SimDiabolo propose l'étude de la fermentation alcoolique des levures (*consommation de glucose, production d'alcool et de dioxyde de carbone - réaction dépendante de la teneur du milieu en dioxygène et de sa température*). Elle est composée de 5 « pages » : une page d'accueil, les 4 autres

représentant les phases de l'activité (cf. 2.1). L'élève peut naviguer d'une page à l'autre. Les étayages directement proposés par SimDiabolo sont des étayages fixes tels que la mise à disposition de documents ressources (quelle que soit la page active) et la liste des actions disponibles pour construire le protocole, assortie du contrôle de saisie de leurs paramètres, ainsi que des rétroactions dynamiques délivrées suite au choix inapproprié d'un microorganisme, dispositif ou réactif (concepts NOA).

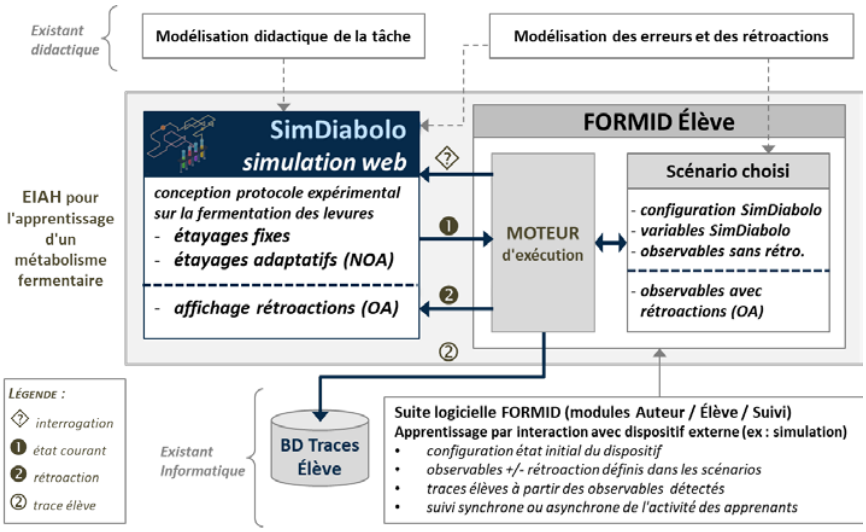


Fig. 1. Architecture de l'EIAH (cadre central) et apports de l'existant (haut et bas)

FORMID-Élève - diagnostic, étayages d'erreurs sur concepts OA, génération et enregistrement des traces élève. La suite logicielle FORMID [12], [13], est dédiée à la *scénarisation*, à l'*exécution* et au *suivi* de situations d'apprentissages en ligne ; elle relève des principes du Learning Design [9] par le fait que la mise en œuvre de ces situations est orchestrée par un scénario exécutable. Chaque scénario choisit un dispositif externe avec lequel l'élève interagira (ici : SimDiabolo), importe la liste de ses variables et configure son état initial (ici : solutions, verrerie, autres matériels disponibles). Il définit également des observables [14] sur des états particuliers de la simulation que le moteur d'exécution de FORMID-Élève vérifiera (mécanisme *observable/observé*). Un observable peut être associé à une rétroaction délivrée à l'élève en cas de détection et est caractérisé par son intention [15].

Au fil de l'activité de l'élève, le moteur d'exécution de FORMID-Élève interroge l'état de SimDiabolo, lequel est représenté par les valeurs des variables importées dans le scénario (*page active, dernier document ressource ouvert, matériel rassemblé, actions du protocole, solutions créées*). Si cet état satisfait la condition de détection d'un observable, l'éventuelle rétroaction qui y est associée est fournie à SimDiabolo pour affichage. Chaque fois qu'un observable est détecté ou n'est plus détecté, une trace élève est générée et enregistrée en base de données. Les traces générées comportent les informations suivantes : *instant du recueil (h/mn/s/ms) / identifiants de*

la session, du groupe d'élèves, de l'élève à l'origine de la trace / identifiant de l'observable / début ou fin de détection / état du dispositif externe.

Scénarios FORMID. Nous avons transposé dans des scénarios FORMID la modélisation didactique des rétroactions relatives aux erreurs portant sur des concepts OA, en définissant des observables sur les états de SimDiabolo représentatifs de ces erreurs. Selon la modélisation didactique, des rétroactions de différents niveaux peuvent étayer une même erreur sur un concept OA (cf. section 1). Pour faciliter l'analyse de l'impact de ces rétroactions, nous avons décidé de définir 4 scénarios, chacun implémentant un niveau différent. Hormis cette distinction, chaque scénario détermine la même configuration de SimDiabolo ; chacun définit également les mêmes observables sur les interactions de l'élève avec SimDiabolo afin que les traces générées puissent rendre compte aussi bien du comportement de l'élève que de ses erreurs, quels que soient les concepts sur lesquels elles portent (cf. Tableau 1).

L'EIAH a ainsi été décliné en 4 « versions » ne différant l'une de l'autre que par le type et la nature du message délivré en cas d'erreur relative à un concept OA.

Tableau 1. Répartition des observables selon le moment de l'activité de l'élève

Page Active	Catégorie	Nombre d'observables
Page Introduction	*Comportement*	**4**
Page Matériel	*Comportement*	**15**
	Erreur sur concept NOA	**1**
Page Protocole	*Comportement*	**11**
	Erreur sur concept NOA	**13**
	Erreur sur concept OA	**17**
Page Expérience	*Comportement*	**11**

3 Bilan et perspectives

Courant novembre 2018, nous avons conduit des expérimentations dans des classes de SVT pour un effectif global de 90 élèves de seconde générale et 21 élèves de terminale scientifique. Chaque version de l'EIAH a été testée dans au moins deux lycées différents et nous avons recueilli un volume de traces conséquent (plus de 26 000 en seconde et environ 2 800 en terminale).

Les premiers résultats d'analyse des traces recueillies confirment que leur sémantique permet d'apprécier l'effet, sur les productions des élèves, des rétroactions dynamiques relatives aux concepts OA. D'autre part, l'activité des élèves pourra être analysée ultérieurement à d'autres fins en exploitant l'ensemble des observables que nous avons définis.

L'approche d'ingénierie que nous avons suivie rend possible l'évolution de l'EIAH proposé sans effort de réingénierie. Ainsi, prendre en compte la diversité des savoirs des élèves en leur délivrant des rétroactions progressives ou dégressives selon le diagnostic en temps réel de leurs difficultés, ne demandera que de définir de nouveaux scénarios.

Pour conclure, le projet DIABOLO a été le terrain d'échanges d'une grande richesse autant sur le plan scientifique que sur le plan humain. Il a aussi permis aux didacticiens de pouvoir mieux projeter la transposition informatique de l'analyse didactique d'une situation d'apprentissage et offert aux informaticiens l'occasion trop rare de penser l'ingénierie d'un EIAH sans s'éloigner de sa raison d'être.

Références

1. Bonnat, C. (2017). Étayage de l'activité expérimentale par un EIAH pour apprendre la notion de métabolisme cellulaire en terminale scientifique. Thèse de doctorat, Université Grenoble-Alpes. Grenoble
2. Chaachoua, H. (2018). T4TEL un cadre de référence didactique pour la conception des EIAH. [éd.] J., Vendeira, C. Pilet. Séminaire ARDM, pp. 5-22.
3. Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique. La Pensée Sauvage, Recherches en Didactique des Mathématiques, Vol. 19, pp. 221-266.
4. Marzin-Janvier, P. (2013). Comment donner du sens aux activités expérimentales ? Mémoire d'habilitation à diriger des recherches, Université Joseph-Fourier, Grenoble I, France.
5. Quintana, C., Reiser, B., J., Davis, E., A., Krajcik, J., Fretz, E., Duncan, R., G., Kyza, E., Edelson, D., Soloway, E. (2004). A Scaffolding Design Framework for Software to Support Science Inquiry. Journal of the Learning Sciences, Vol. 13, pp. 337-386.
6. Croset, M-C., Chaachoua, H. (2016). Une réponse à la prise en compte de l'apprenant dans la TAD : la praxéologie personnelle. La Pensée Sauvage, Recherche en Didactique des Mathématiques, Vol. 36, pp. 1-34.
7. De Jong, T., van Joolingen, W. R. (1998). Scientific Discovery Learning with Computer Simulations of Conceptual Domains. Review of Educational Research, Vol 68(2), pp 179-202. https://doi.org/10.3102/00346543068002179
8. Thomas, R.C., Milligan, C.D. (2004), Putting Teachers in the Loop: Tools for Creating and Customising Simulations. Journal of Interactive Media in Education (Designing and Developing for the Disciplines Special Issue), Vol. 15, pp 1-22. ISSN:1365-893X.
9. Iksal, S. (2011). Tracks Analysis in Learning Systems: A Prescriptive Approach. International Journal for e-Learning Security, Vol. 1(1), pp. 3-9.
10. Koper, R. (2006). Current Research in Learning Design. Journal of Educational Technology & Society, Vol. 9(1), pp. 13-22.
11. Lejeune, A., Pernin, J.P. (2004), A taxonomy for scenario-based engineering, Proceedings of the IADIS International Conference, Lisbon, Portugal, pp. 249-256
12. Guéraud, V., Cagnat, J.M. (2006). Automatic semantic activity monitoring of distance learners guided by pedagogical scenarios. [éd.] Springer. 1st EC-TEL Conference, pp. 476-481
13. Guéraud, V., Lejeune, A., Adam, J.M. Dubois, M., Mandran, N. (2009). Flexible environment for supervising simulation-based learning situations. 14th AIED Conference, pp. 698-700.
14. Lejeune, A., Guéraud, V. (2012). Embedding observation means into the learning scenario: Authoring approach and environment for simulations-based learning. 12th IEEE ICALT Conference, pp. 273-275.
15. Lejeune, A., Guéraud, V., Mandran, N. (2016). Improving Usage of Learning Designs by Teachers: A Set of Concepts for Well-Defined Problem Resolution. 13th EC-TEL Conference, Vol. 9891/2016, pp. 404-409. LNCS.

« Marketplaces » de l'apprentissage en ligne : appréhender les stratégies des instructeurs

David Pontalier, Matthieu Cisel

Université Paris-Descartes, Laboratoire EDA, 45 rue des Saint-Pères,
75006 Paris, France
matthieucisel@gmail.com

Résumé. A côté des plateformes de MOOC dominées par les établissements d'enseignement supérieur, se développent des plateformes ouvertes au grand public, comme *Udemy*, où tout un chacun peut concevoir un cours, le plus souvent payant. On parle de « marketplaces » de l'apprentissage. Nous proposons à travers cet article une étude exploratoire de l'offre d'une de ces plateformes, *Skillshare,* afin de mieux comprendre les stratégies des concepteurs de cours. Nous utilisons un logiciel de type « web scraper » pour extraire automatiquement les données partagées sur les pages d'accueil de plus de douze mille cours. Cette première analyse corrobore notamment l'hypothèse selon laquelle les concepteurs de cours, sur ces plateformes, privilégient des sujets très opérationnels, qui n'intersectent que partiellement les sujets couverts par les établissements d'enseignement supérieur. Nous essayons d'appréhender l'influence du modèle par abonnement, avec rémunération fondée sur le niveau de visionnage des vidéos, sur les stratégies des concepteurs.

Mots-clés. Open education, apprentissage en ligne, crowdsourcing, analyse de contenu, marketplace

Abstract. Besides MOOC platforms, that are dominated by institutions of higher education, some platforms like *Udemy* allow anyone to host their own courses. These marketplaces allow course designers to earn money, but they share revenues with the hosting platform. We propose here an exploratory study of the offering of such a platform, Skillshare. We used a web scraper to automatically extract the data from more than twelve thousand course landing pages: number of registered learners, course duration, topic, and cost. The first analyses notably corroborate the hypothesis stating that course designers using these platforms favor concrete topics that only partially overlap with topics covered by institutions of higher education, and manage to attract hundreds, sometimes thousands of registrations, even when they are not free.

Keywords. Open education, online learning, crowdsourcing, content analysis, place de marché

1 Introduction

Il existe à côté des fournisseurs de MOOC [2] comme Coursera un type de plateforme, moins connu tant des médias que de la communauté scientifique : les « marketplaces » de l'apprentissage en ligne. Des plateformes comme *Skillshare* ou *Udemy*

comptabilisent des millions d'inscriptions et des milliers, voire parfois des dizaines de milliers d'instructeurs. Les cours sont produits par des individus non nécessairement affiliés à une institution particulière. Nous ne rangeons pas ces « marketplaces » parmi les plateformes de MOOC, comme peuvent le faire certains auteurs [5]. En effet, il n'y a généralement pas de délimitation des cours dans le temps, et le système a été pensé pour que le règlement des inscriptions constitue le principal modèle économique. Nous sommes bien loin du « Open » de « Open registration » [2] qui fit la notoriété des MOOC.

Tout un chacun peut utiliser ces plateformes pour y développer un cours, à condition de partager les bénéfices éventuels avec l'entreprise qui héberge la formation. Les modèles varient selon les acteurs. Là où les instructeurs sont libres de fixer le prix de la formation chez Udemy, Skillshare repose sur un modèle par abonnement. Pour une somme de quelques dizaines d'euros par mois, les utilisateurs ont accès à l'ensemble des cours dits « Premium » de la plateforme. Les utilisateurs n'ayant pas souscrit à un abonnement n'ont accès qu'aux seuls cours gratuits, largement minoritaires, comme nous allons le voir. Les concepteurs de cours, ou instructeurs, sont rémunérés ensuite au prorata du nombre de visionnages de vidéos que leur contenu engendre, c'est le modèle économique dit, du « streaming ». S'il existe des ouvrages [3] visant à aider de tels concepteurs dans leur initiative, la littérature scientifique sur le sujet reste embryonnaire ; les rares publications mentionnant Udemy sont issues de l'équipe de l'entreprise [6], et portent avant tout sur des considérations techniques.

Une meilleure compréhension du fonctionnement de ces plateformes nécessite de mieux appréhender les thématiques qu'elles portent, mais également de saisir les stratégies mises en place par les instructeurs pour accroître leurs profits. Nous proposons ici une première étude exploratoire visant à décrire les principales caractéristiques d'une plateforme représentative des marketplaces de l'apprentissage, Skillshare, et à travers ce travail de répondre à plusieurs questions relatives aux stratégies des acteurs de ces plateformes. Quelles sont les stratégies des instructeurs de ces plateformes en matière de création de cours ? Dans quelle mesure le modèle économique choisi par une plateforme influe-t-il sur ces stratégies ?

Nous avons opté pour un cadre théorique régulièrement mobilisé dans les travaux sur les communautés d'enseignants producteurs de ressources en ligne : la dynamique des communs d'Elinor Ostrom [4]. Comme le soulignent Baron et Zablot [1], « La théorie met l'accent sur l'étude des règles régissant la communauté en offrant un cadre stable d'attentes ». Le caractère marchand de l'entreprise Skillshare ne constitue pas à nos yeux un obstacle à l'utilisation de ces cadres. En effet, si les entreprises marchandes de production de ressources sont évoquées dans les recherches mobilisant ces cadres (ibid), elles restent à la marge des travaux sur les communautés créatrices de ressources. Les marketplaces comme Skillshare constituent une opportunité d'explorer l'intérêt de tels cadres.

Dans le cadre théorique mobilisé, trois niveaux de règles sont distingués : « les règles opérationnelles, qui affectent les décisions à prendre de manière routinière (déterminer l'information qui peut être diffusée ou non, définir les sanctions et les récompenses, …) » (ibid). Le choix des modèles économiques relève par exemple de ce premier niveau. Puis viennent « les règles concernant des choix collectifs qui affectent les règles précédentes ». Enfin viennent « les règles associées à des « choix constitutionnels » qui

déterminent l'ensemble de règles de choix collectifs ». Ces deux derniers niveaux de règles ne sont pas traités dans cet article.

A travers ces réflexions sur les stratégies des instructeurs et sur les règles opérationnelles qui régissent leur action, nous souhaitons attirer l'attention sur un objet encore trop peu étudié, et la recherche présentée ici est avant tout abductive. C'est-à-dire qu'après une première analyse exploratoire quantitative du jeu de données, nous émettons un certain nombre d'hypothèses qu'il faudra étayer via des travaux plus qualitatifs. Notre principale hypothèse que les instructeurs adaptent les caractéristiques de leurs cours (nombre, durée, thématique) de sorte à maximiser leur profit, en prenant en compte l'impact du modèle dit de « streaming » dans leur stratégie de conception. Pour maximiser le temps de vidéo vue, ils produisent par exemple des formations courtes contenant peu de vidéo, quitte à multiplier le nombre de cours produits. Afin de mettre à l'épreuve cette hypothèse, nous avons réalisé une analyse des caractéristiques de plus de douze mille cours de la plateforme Skillshare via un web scraper, logiciel qui extrait automatiquement des informations d'intérêt sur une page Internet.

2 Méthodologie

Le projet initial consistait à analyser l'offre de Udemy, l'une des marketplaces les plus anciennes et les plus reconnues en termes d'apprentissage en ligne. Néanmoins, le site a techniquement empêché l'extraction de données par des web scrapers, et explicitement interdit, dans ses conditions générales d'utilisation, d'utiliser ce type de logiciels sur son site. Nous nous sommes dès lors concentrés sur des plateformes équivalentes, mais ne présentant pas ce type de contraintes.

Nous avons utilisé le web scraper *Data Miner* pour extraire les données d'intérêt. Nous l'avons installé comme plugin sur le navigateur Google Chrome. Les cours de la plateforme sont divisés en quatre catégories – Business, Création, Style de vie, et Technologie – ces catégories étant divisées en trente-quatre sous-catégories, allant de la photographie à la comptabilité. Il n'est pas possible d'afficher tous les cours de la plateforme sur une même page. En revanche, il est possible d'afficher en même temps les trois cents cours les plus populaires d'une sous-catégorie – si tant est qu'il y en ait au moins trois cents, ce qui n'est pas le cas de toutes les sous-catégories. Le terme « populaire » utilisé par la plateforme fait référence au nombre d'inscrits sur la plateforme. Au-delà de cette limite de trois cents, les quelques cours qui ne sont pas pris en compte n'ont souvent que quelques inscrits, voire aucun, et peuvent dès lors être considérés comme quantité négligeable en termes d'activité de la plateforme.

Pour extraire les données, nous nous sommes connecté sur une page correspondant à une sous-catégorie de cours. Le scraper parcourt l'ensemble de la page après l'avoir fait défiler automatiquement, jusqu'à afficher tous les cours les plus populaires disponibles (option « scroll »). Nous avons ensuite extrait pour chaque cours les informations suivantes : nom du cours, nom de l'instructeur, catégorie et sous-catégorie du cours, nombre d'étudiants inscrits, et durée du cours, en heures et minutes. Nous avons répété manuellement l'opération, sous-catégorie de cours par sous-catégorie, jusqu'à obtenir une base de données de 12314 cours. Cette base a été nettoyée et analysée avec le logiciel R 3.2.

3 Résultats

Le premier élément qui distingue les marketplaces des principales plateformes de MOOC comme Coursera ou edX est le nombre d'inscriptions récoltées, sachant que la plupart sont le fait d'utilisateurs qui ont payé un abonnement. L'abonnement est nécessaire pour accéder aux cours dit « Premium » ; ils représentent 80,4% des formations de la plateforme. Environ 18% des cours concernés ont moins de dix étudiants, environ un tiers entre dix et cent, un autre tiers entre cent et mille, et un peu plus de 10% ont plus de mille inscriptions (Table 1), le maximum étant de 60007. Skillshare apparaît dès lors comme une plateforme à deux vitesses, avec pour une moitié des cours ne rassemblant que quelques dizaines d'inscriptions tout au plus, et pour l'autre moitié des cours en rassemblant plusieurs centaines voire plusieurs milliers ou dizaines de milliers. Les revenus des concepteurs de ces cours étant proportionnels au nombre d'inscriptions engrangées, on peut voir se dessiner à cet égard deux catégories d'instructeurs. Le propos est néanmoins à nuancer par le fait que les instructeurs sont susceptibles de diffuser plusieurs cours sur la même plateforme.

Table 1. Répartition des cours selon le nombre d'inscriptions qu'ils engrangent, et selon leur importance dans l'activité de la plateforme (N= 12314)

Nombre d'inscrits dans le cours	Proportion des cours	Durée moyenne du cours (minutes)	Proportion des inscriptions sur la plateforme
Aucun	2,9	76,2	0,0
Entre 0 et 10	15,1	74,1	0,1
Entre 10 et 100	34,8	70,4	2,3
Entre 100 et 1000	35,5	70,7	17,9
Entre 1000 et 10000	10,4	68,3	47,7
Plus de 10000	1,2	63,5	32,0

Il est probable que les instructeurs visent un optimum en termes de durée pour maximiser leurs profits. Plus les cours sont longs, plus ils peuvent facturer à Skillshare des minutes visionnées. Néanmoins, ils doivent prendre en compte le fait qu'au-delà d'une certaine durée de vidéo, les utilisateurs de cours en ligne ne persistent que rarement jusqu'au terme de la formation. Avec des cours trop longs, les instructeurs risqueraient non seulement de dépenser inutilement des ressources pour la création de vidéos, mais aussi de décourager certains utilisateurs de débuter leurs cours. Cette considération nous a amené à nous pencher sur les durées moyennes des cours en fonction du nombre d'inscrits, pour constater que les formations qui attirent le plus d'apprenants sont aussi les plus courtes (Table 1), et sur la manière dont cette durée varie selon les thématiques (Table 2).

On constate que les cours relevant de la technologie durent 113 (±141) minutes en moyenne, là où les cours des autres catégories durent autour d'une heure environ. Une ANOVA montre que la différence est statistiquement significative. Ceci tient probablement au fait que les cours à dominante technologique impliquent davantage de vidéos relevant de savoir-faire (tutoriels, etc.). On peut aussi émettre l'hypothèse selon

laquelle ces cours attirent moins d'étudiants en moyenne – 181(±755) contre 1452 (±3644) pour les formations relevant de la catégorie « Création » (Photographie, etc.) – et que la production de vidéos, notamment sous forme de tutoriels, constitue une stratégie pour accroître ces revenus malgré une plus faible audience. En effet, la technologie n'est pas la thématique qui attire le plus d'apprenants 4,8% des inscriptions pour 17,8% des cours (Table 2), les cours relevant de la « Créativité » (Photographie, etc.) engrangeant l'essentiel des inscriptions.

Table 2. Une plateforme structurée en différentes catégories (N=12314). Les chiffres entre parenthèses correspondent aux écarts-type, particulièrement élevés.

Catégorie de cours	% des cours de la plateforme	Nombre moyen d'étudiants	% des inscriptions de la plateforme	Durée moyenne (minutes)
Business	31,0%	488 (±2208)	22,7%	66 (±88)
Créativité	28,7%	1452 (±3644)	62,5%	64 (±78)
Style de vie	22,4%	296 (±1022)	9,9%	53 (±77)
Technologie	17,8%	181 (±755)	4,8%	113 (±141)

Si le fait de dépasser une certaine longueur en termes de vidéos peut être considéré comme un risque pour la collecte d'inscriptions, une stratégie employée par certains instructeurs peut probablement consister à multiplier les cours payants. Cette considération nous a amené à nous pencher sur la question de l'équilibre entre cours gratuits et cours payants (Table 3). Nous avons constaté qu'une minorité des concepteurs suivaient une démarche de productions de cours gratuits. Près de la moitié des instructeurs (42.5%) ne produisent qu'un seul cours sur la plateforme, payants, et 28,7% deux cours ou davantage, le tout sans produire de cours gratuits. Une partie non négligeable des instructeurs (21,3%) produisent un ou plusieurs cours gratuits, sans proposer de cours payants. Seule une minorité (7,4%) produisent simultanément des cours gratuits et des cours payants.

Table 3. Répartition des instructeurs selon le nombre de cours payants et de cours gratuits qu'ils produisent, en pourcentage des concepteurs de cours (N= 4541). Entre parenthèses sont affichés la proportion des cours de la plateforme représentés par une catégorie d'instructeurs.

N Cours payants N Cours gratuits	Aucun	Un	Deux et plus
Aucun		42,5% **(17,5%)**	28,7% **(46,7%)**
Un	14,0% **(5,0%)**	1,3% **(1,0%)**	2,2% **(4,7%)**
Deux plus	7,3% **(14,0%)**	1,2% **(0,7%)**	2,7% **(10,2%)**

Les instructeurs proposant plusieurs cours payants sont minoritaires, ils représentent néanmoins 46,7% des cours (Table 3). Par ailleurs, un instructeur qui ne conçoit qu'un cours recueille 724 inscriptions en moyenne, pour 2043 inscriptions lorsqu'il produit

entre deux et quatre cours, et 13069 pour les instructeurs qui produisent plus de vingt cours. Par ailleurs, même si plus de la moitié (58%) produisent des cours dans une seule catégorie, les instructeurs qui conçoivent plusieurs cours varient fréquemment les thématiques qu'ils choisissent : 33% produisent des formations relevant de deux catégories différentes. Ce résultat nous amène à émettre l'hypothèse qu'une partie de ces instructeurs ne cherche pas uniquement à partager des savoir-faire qu'ils possèdent déjà, mais aussi à identifier plusieurs thématiques susceptibles de rapporter de revenus, quitte à se former sur un sujet pour pouvoir le monétiser.

4 Discussion et conclusion

Nous constatons que les instructeurs de Skillshare favorisent la création de cours payants, dont la durée est bien inférieure à celle des MOOC [2]. Ceci tient vraisemblablement au fait que les concepteurs sont rémunérés en fonction du temps de visionnage de leurs vidéos, conséquence des règles opérationnelles [4] qui régissent le fonctionnement de la plateforme.

L'une des limites de ce travail est de se cantonner à une approche quantitative. Une enquête qualitative en cours, menée auprès de concepteurs de la plateforme et des dirigeants de la plateforme, nous permettra de mieux appréhender les mécanismes qui régissent l'action des instructeurs, et ainsi de corroborer et affiner nos hypothèses relatives à leur stratégie de production de contenus. Une étude sur les apprenants est également envisagée, pour comprendre l'interaction entre leur utilisation des cours et les stratégies mises en œuvre par les instructeurs.

Enfin, suivant les suggestions de Baron et Zablot [1], nous proposons dans de futurs travaux de mobiliser plus avant le cadre de la dynamique des communs pour appréhender comment d'éventuelles arènes de choix collectifs [4] se mettent éventuellement en place pour fixer les règles constitutionnelles de la plateforme, et dans quelle mesure les instructeurs peuvent peser sur les orientations prises par la plateforme.

Références

1. Baron, G.-L. et Zablot, S. (2017). De la constitution de ressources personnelles à la création de communautés formelles : étude de cas en France. *Review of Science, Mathematics and ICT Education, 11*(2), p. 27-45.
2. Daniel, J. (2012). Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. *Journal of Interactive Media in Education,* 3(0).
3. Mardan, A. (2018). *Using Your Web Skills To Make Money: Secrets of a Successful Online Course Creator and Other Income Strategies that Really Work*. New York ; NY : Apress.
4. Ostrom, E. (1990). *Governing the commons: the evolution of institutions for collective action*. Cambridge ; NY : Cambridge University Press.
5. Tovar, E., Dimovska, A., Piedra, N. et Chicaiza, J. (2013, Mars). OCW-S: Enablers for building sustainable open education evolving OCW and MOOC. Dans *Proceedings of IEEE Global Engineering Education Conference (EDUCON 2013)* (p. 1262–1271).
6. Wai, L. (2016, Décembre). Data science at Udemy: Agile experimentation with algorithms. Dans *Proceedings of 2016 Future Technologies Conference (FTC 2016)* (pp. 355-360).

Écriture numérique à distance : interventions de co-écriture dans un environnement dynamique

Christine Blanchard Rodrigues[1], Marion Andriamasinoro[2]

[1]Université Clermont Auvergne, LRL, F-63000 Clermont-Ferrand, France
christine.blanchard@uca.fr
[2]Université Clermont Auvergne, LRL, F-63000 Clermont-Ferrand, France
marion.andriamasinoro@etu.uca.fr

Résumé. Il existe aujourd'hui de nombreux dispositifs permettant l'écriture synchrone et asynchrone, seul ou en groupes, utilisés dans des contextes variés d'apprentissage. Nous nous intéressons plus particulièrement au système Etherpad, qui permet une activité scripturale d'un ou plusieurs participants sur une page blanche (ou pad), simultanément ou non. Ce système est de plus intéressant pour la recherche, puisqu'il conserve un historique dynamique d'écriture, favorisant ainsi le suivi pas à pas du développement de la production écrite par un ou plusieurs participants. Nous décrivons ici le dispositif mis en place pour des étudiants de Licence, qui avaient pour objectif l'écriture en groupes (de 3 à 5 étudiants) d'un article traitant des Sciences du Langage. Nous présentons également les résultats préliminaires de l'analyse des interventions des pairs, en vue de comprendre les processus de co-écriture à l'œuvre.

Mots-clés. pad, co-écriture, intervention, processus rédactionnel

Abstract. Many courses are nowadays in place for synchronous and asynchronous writing, either alone or in groups, in various learning contexts. We are particularly interested in the Etherpad system, which allows one or more participants to engage in scriptural activity on a blank page ("pad"), synchronously or asynchronously. This system is indeed interesting for research, since it retains a dynamic history of writing, thus allowing to follow step by step the development of a written production by one or more participants. In this presentation, we describe the scheme set up for students that were asked to write in groups (from 3 to 5 students) an article on Language Sciences. Preliminary results of peer intervention analysis are also presented to understand co-writing processes.

Keywords. pad, co- writing, intervention, writing process

1 Introduction

De nombreuses études montrent les apports de l'écriture, d'un ou de plusieurs participants, dans des systèmes numériques en ligne ([1], [2], [3], [4], [5]). Dans de tels dispositifs, les scripteurs sont encouragés à suivre des phases telles que la planification, la mise en texte, la révision et la publication ([6], [7], [8]), chacune étant dépendante des autres, avec des retours en arrière possibles. Cette activité contribue au

développement d'une littératie - compétences rédactionnelles - mais aussi d'une littératie numérique ([9]), conformément aux normes ISTE ([10]). Elle pose également la question de l'accompagnement, du rôle de l'enseignant ([11], [12]), et de la collaboration ([13]) dans de tels dispositifs.

Notre contexte de recherche vise la compréhension du développement des littératies et des difficultés inhérentes. Le processus rédactionnel est en effet complexe, et nécessite un accompagnement dans toutes ses phases ([14]). Il est d'autant plus complexe lorsqu'il s'agit d'une écriture à plusieurs, lors de la co-création d'un texte.

Dans le cadre de la présente étude, nous nous intéressons plus particulièrement au développement et/ou à l'amélioration de compétences scripturales à travers l'intervention des pairs dans l'écriture-même. Nous traitons la question de recherche suivante : lors du processus rédactionnel en contexte numérique d'écriture synchrone, quelles formes prennent les interventions des participants lors d'une co-écriture ? Pour y répondre, nous prenons appui sur un dispositif de formation intégrant la co-écriture d'un article, à l'aide du système Etherpad. Nous émettons l'hypothèse qu'un outil permettant une réelle utilisation synchrone du même support, à plusieurs claviers, favorise non seulement l'écriture coopérative (acteurs agissant ensemble), mais aussi l'écriture collaborative, incluant des corrections par un scripteur de la production d'un autre scripteur (par exemple, correction de langue, correction de la structure syntaxique), dans une démarche d'intertextualité, dans laquelle "les interventions langagières se construisent les unes à partir des autres" [16]. Nous verrons ainsi si l'observation d'une démarche plus coopérative (contribution indépendante, relevée, par exemple, dans l'étude [17]) que collaborative avec ce système de co-écriture numérique, se vérifie dans notre contexte.

2 Contexte de l'étude : dispositif centré sur la co-écriture en ligne

Notre projet d'étude du processus co-rédactionnel prend appui sur les données issues d'un enseignement intitulé "Plateformes de communication à distance", dispensé en 3ème année de Licence Sciences du Langage (SDL), à l'UFR Lettres, Culture et Sciences Humaines (LCSH), Université Clermont Auvergne (UCA), France. L'objectif consiste en la production d'articles, qui forment un numéro de la revue étudiante "Le Gergo Parleur"[1]. Les sujets traités sont choisis par les participants.

Dans ce cours (8 séances de 3 heures), des groupes de 3 à 5 étudiants élaborent un écrit en utilisant une variété d'outils de communication et de création, notamment des pads (générés depuis Framapad©[2]), pour produire du texte et/ou intervenir sur les productions des membres de leur groupe. La production écrite de chaque participant est identifiée dans le pad par une couleur et par le prénom du scripteur (qui apparait lorsque l'on survole le texte avec la souris). Par ailleurs, le pad enregistre la saisie et l'effacement de tout caractère. Le dispositif pédagogique prévoit une écriture simultanée dans le pad, le cours ayant lieu toujours en synchrone, en présentiel (une séance) et à distance (une séance avec clavardage, deux séances en visio-conférence).

[1] Projet coordonné par Christine Blanchard et Lidia Lebas, enseignantes au département de Linguistique, soutenu par les Aides à Projets Pédagogiques de l'UFR LCSH, UCA.

[2] https://framapad.org/fr/.

2.1 Scénario pédagogique

Le cours est structuré de manière à s'accorder sur le processus d'écriture, tout en accompagnant les étudiants dans la démarche de projet en mode hybride. La formation concernée a ainsi été décomposée en 6 grandes étapes (tableau 1), et comprend des séances synchrones, soit en présentiel, soit à distance.

Le cours amène ainsi les étudiants à utiliser des outils de communication à distance mais aussi des outils de co-création en ligne. Les participants disposent pour cela d'espaces de travail sur la plateforme numérique de l'université (un espace commun à l'ensemble des participants, et un espace collaboratif pour chaque groupe). Les outils mis à leur disposition sont variés : Google document©, Mindomo© et Framapad© pour l'écriture, Adobe Connect©, forums de discussion, clavardage via la plateforme Moodle© de l'université pour la communication.

Tableau 1. Etapes, procédures et modalités du scénario pédagogique du cours.

Etapes	Procédure et déroulement des séances	Phases
1 - Analyse de l'existant (1 séance)	- Présentation du cours et constitution des groupes - Analyse de l'existant (***revues***) et réflexion	Planification
2 – Choix des thématiques (1séance)	- Observation du schéma de l'activité d'écriture - Choix des sujets traités (***Google© doc***)	
3- Choix des sujets (1 séance)	- Elaboration d'une carte conceptuelle (***Creately©*** ou ***Mindmeister©***)	
4 - Élaboration du contenu (4 séances)	- Planification (Diagramme de Gantt prévisionnel) - Création du contenu textuel (***Framapad©***) - Recueil et/création de ressources	Mise en texte
	- Vérification de l'intention d'écriture, amélioration, corrections (***outils au choix)*** - Diagramme de Gantt final	Révision
5 - Mise en place de la revue (1 séance)	- Finalisation de l'ensemble de la production finale - Intégration à la maquette de la revue - Réflexion sur le projet (message sur le ***forum)***	Publication
6 - Diffusion de la revue (fin du cours)	- Impression de la revue - Retour sur le produit : réflexion ***(forum)***	Auto-évaluation

2.2 Compétences développées

Le cours et son organisation visent le développement de différentes compétences, telles que :

- Être capable de produire un texte conforme, en respectant les phases du processus rédactionnel.
- Savoir participer à un projet d'écriture en vue d'une production diffusée à un public non spécialiste.
- Comprendre et appliquer la démarche de gestion de projet, à l'aide d'outils tels que le diagramme de Gantt (que les étudiants retrouveront tant dans la poursuite d'études que dans la vie professionnelle).

- Savoir utiliser les outils de communication et de création (carte scripturale, outils de co-écriture en ligne, clavardage et visioconférence) permettant un travail coopératif et collaboratif en présentiel et à distance.
- Savoir adopter une démarche réflexive sur le processus d'écriture, sur la communication, sur la gestion de projet et sur la formation professionnalisante apportée par cet enseignement.

2.3 Aspects coopératifs et collaboratifs du dispositif

Le dispositif mis en place peut se définir comme une situation problème, qui "vise l'émergence d'un conflit sociocognitif, pour que les représentations initiales de chacun soient éprouvées, puis majorées" ([18]). Il comporte en outre une dimension coopérative et collaborative.

Le projet est coopératif au niveau de la conception du numéro de revue : en effet, chaque groupe d'étudiants écrit un texte, qui constitue l'un des articles de la revue. Le dispositif comporte aussi une dimension coopérative, en termes de projet d'écriture d'un article en groupe. Les étudiants doivent en effet se répartir les tâches. Ils sont ainsi initiés à la démarche de projet, notamment par la réalisation d'un diagramme de Gantt prévisionnel puis final, pour comprendre comment cette répartition amène à une production collective. La collaboration, quant à elle, est envisagée et souhaitée dans l'écriture-même de l'article. Elle accompagne une écriture coopérative (chaque participant pouvant développer la production écrite).

3 Contexte de l'étude et analyse préliminaire

Nous considérons présentement les données issues du cours dispensé lors du semestre 2 de l'année universitaire 2017/2018. Pour l'année 2018, le cours comprenait 47 participants, qui ont formé 11 groupes de travail (cf. Tableau 2). Il y avait 40 étudiantes et 7 étudiants, en formation initiale.

Tableau 2. Groupes et participants.

G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
G1Ea G1Eb G1Ec G1Ed G1Ee	G2Ea G2Eb G2Ec G2Ed G2Ee	G3Ea G3Eb G3Ec G3Ed	G4Ea G4Eb G4Ec G4Ed	G5Ea G5Eb G5Ec G5Ed G5Ee	G6Ea G6Eb G6Ec G6Ed G6Ee	G7Ea G7Eb G7Ec G7Ed	G7Ea G7Eb G7Ec G7Ed	G7Ea G7Eb G7Ec G7Ed	G10Ea G10Eb G10Ec G10Ed	G11Ea G11Eb G11Ec

3.1 Données considérées et méthodologie d'analyse

Nous avons observé l'écriture sur les 11 pads utilisés lors du cours. Nous procédons par une observation par balayage ([19]), en vue de relever des éléments les plus

pertinents[3]. Pour cette étude préliminaire, nous avons observé les historiques de constitution des pads, à travers la fonctionnalité "historique dynamique" de Framapad©. Les caractéristiques de ce logiciel nous ont permis d'identifier les différentes contributions, grâce à la fonction que nous avons décrite précédemment (identification en couleur et prénom du scripteur lors du survol du texte).

3.2 Premiers résultats

A travers notre observation, nous avons relevé quatre formes d'intervention dans l'utilisation en groupe du document d'écriture numérique :

- Compléter les phrases d'un co-scripteur. Il s'agit d'ajouter de la production écrite, souvent au sein du même paragraphe, précédemment ou à la suite d'un énoncé. C'est la stratégie la plus récurrente dans les pads étudiés.
- Corriger la production du co-scripteur. Cette stratégie, en lien avec la phase de révision de l'activité d'écriture, est toutefois peu répandue dans les pads observés.
- Commenter la production écrite. Cela est observé sur quelques pads. Il faut noter que le système de Framapad© ne permet pas d'ajouter des commentaires (contrairement à un document Google ou Office).
- Communiquer dans le document. Cette stratégie s'observe toutefois assez peu, Framapad© contenant un système de clavardage qui permet de remplir cette fonction.

Cette observation nous permet d'apporter des éléments de réponse à notre question de recherche : les participants utilisent le support en réalisant majoritairement des interventions de complétion d'énoncés et/ou de paragraphes. Ces interventions s'apparentent à une démarche collaborative, comme définie par [17] : "Un énoncé comportant au moins deux couleurs peut ainsi être caractérisé comme 'collaboratif' ".

4 Conclusion et perspectives

Cette étude préliminaire nous permet de relever quelques formes d'utilisation du document en ligne dans la co-écriture des textes. Ces résultats seront à préciser et à renforcer, en relevant notamment les aspects linguistiques ayant fait l'objet d'une intervention des pairs, via commentaires ou écriture, et à compléter par le croisement avec diverses données, telles que les productions écrites finales, ainsi que les clavardages et séances en visio-conférence. La place de l'enseignant ([20]) pourra également être étudiée par un croisement des données en lien avec les interventions tutorales. Nous espérons ainsi mettre en évidence un certain nombre de paramètres qui entrent en jeu dans cette co-construction de connaissances, en vue d'être en mesure de proposer des recommandations adaptées pour les scripteurs et les enseignants.

[3] Un projet de recueil des données pour une mise en corpus est en cours.

Références

1. Hsu, H.-C., Lo, Y.-F., Using wiki-mediated collaboration to foster L2 writing performance. In: Language Learning & Technology, Vol. 22 n°3. (2018) 103–123.
2. Strobl, C., Satar, H. M. (eds), Multimodality in Electronic Feedback in Writing. Writing and Pedagogy, Vol. 9 n°3. (1998).
3. Hulin, T., Enseigner l'activité "écriture collaborative". In: Tic&société, Vol. 7 n°1. (2013) http://journals.openedition.org/ticetsociete/1314.
4. Rodrigues, C., Outils collaboratifs pour la production écrite en anglais. In: Écrire avec le numérique, Les Langues Modernes n°3. (2014).
5. Blanchard Rodrigues, C., Lebas, L., Collaborative writing as means of enhancing interlocutive attitudes and skills in written communication. Table ronde Collaborative writing and commentary in digital environments : an e-learning perspective. (2017) Universität Duisburg-Essen, Essen: Allemagne.
6. Flower, L.S., Hayes, J.R., A cognitive process theory of writing. In: College Composition and Communication, Vol. 32 n°4. (1981) 365–387.
7. Van Galen, G.P., Handwriting: Issues for a psychomotor theory. In: Human Movement Science, Vol. 10. (1991) 165–191.
8. Levelt, W. J. M., Roelofs, A., Meyer, A. S., A Theory of Lexical Access in Speech Production. In: Behavioral and Brain Sciences, Vol. 22. (1999) 1–75.
9. Centre Canadien d'Éducation aux médias et de littératie numérique, Les fondements de la littératie numérique, Éducation aux médias et de littératie numérique. http://habilomedias.ca/
10. International Society for Technology in Education, Les normes ISTE pour les élèves. (2016) https://www.iste.org/docs/pdfs/iste-standards-s-2016-french-2017.pdf
11. Rodrigues, C., Stratégies lexicales d'intercompréhension pour l'apprentissage de l'anglais en IUT. In Degache C., Garbarino S. (eds.), Itinéraires pédagogiques de l'alternance des langues : l'intercompréhension. (2017) Grenoble : UGA Editions, collection Didaskein.
12. Rodrigues, C., Analyse des potentialités d'un dispositif médiatisé pour les régulations pédagogiques dans la production écrite en FLE. In: Alsic, Vol.15 n°3. (2012).
13 Thibert, R., Quelles pratiques collaboratives à l'heure des TIC ? In: Dossier d'actualité de l'INRP, n° 43. (2009).
14. Bishop, M.-F., "Racontez vos vacances…" : Histoire des écritures de soi à l'école primaire (1882-2002). Grenoble : Presses universitaires de Grenoble (2010).
15. Dumouchel, G., L'écriture collaborative à l'aide des cartes scripturales : Proposition d'une approche innovante. In : Québec français, n°173. (2014) 70–73.
16. Baker, M.J., Formes et processus de la résolution coopérative de problèmes : des savoirs aux pratiques éducatives. In: Rouiller Y., Lehraus K. (eds.), Vers des apprentissages en coopération : rencontres et perspectives. Berne : Peter Lang (2008) 107–130.
17. Temperman, G., Walgraeve, S., De Lievre, B., Boumazguida, K., Développer des compétences de conceptualisation et d'analyse avec un forum de discussion et un etherpad. In: Sticef, Vol. 24, n°1. (2017) 153-181.
18. Connac, S., Coopérer ? Quel bazar ! In: Les cahiers pédagogiques, n° 505 (2013) 12–13
19. Atifi, H., Gauducheau, N., Marcoccia, M., Les manifestations des émotions dans les forums de discussion. Journées d'étude Emotions et interactions en ligne, ICAR ENS LSH/Lyon 2. (2005).
20. Peraya D., Dumont P., Interagir dans une classe virtuelle : analyse des interactions verbales médiatisées dans un environnement synchrone. In: Revue française de pédagogie n°145. (2003) 51–61.

Outils en réseau et interactions entre pairs dans la réalisation d'activités collectives à distance

Sonia Androwkha

Université de Lille, Laboratoire CIREL, équipe Trigone (EA 4354)
Cité Scientifique, 59655 Villeneuve d'Ascq, France
sonia.proust@univ-lille.fr

Résumé. Cette contribution a pour cadre la réalisation d'activités collectives en formation à distance. Elle interroge le lien entre les outils en réseau et les interactions socio-cognitives et socio-affectives entre pairs. S'appuyant sur un ensemble de données qualitatives recueillies auprès d'apprenants adultes constitués depuis peu en trinômes, cette étude présente une analyse de données textuelles. Les résultats montrent que les apprenants privilégient des outils déjà usités avant l'entrée en formation et les utilisent en continu. Ces outils, prioritairement sollicités pour effectuer les tâches collectives prescrites, participent significativement au développement des interactions socio-affectives.

Mots-clés. outils en réseau ; interactions sociales ; travail collectif ; e-learning

Abstract. The following contribution is about the achievement of distance collective activities. It will examine the link between the networked tools and the socio cognitive or socio affective interactions between peers. This paper is based on a set of qualitative data obtained among adult learners who have recently been working as a trio, and it presents a lexicometric analysis. The results suggest that learners prefer tools already in use before starting learning and use them continuously. These tools, which are primarily used to carry out the prescribed collective tasks, contribute significantly to the development of socio-affective interactions.

Keywords. networked tools ; social interactions ; collective work ; e-learning

1 Introduction

L'émergence d'internet et du web 2.0 a renouvelé la formation à distance en lui offrant l'opportunité de mettre au cœur de ses dispositifs les interactions entre pairs. Dans le cadre de la réalisation d'activités collectives, ces interactions, appuis cognitifs et soutiens affectifs, sont rendues possibles par les outils en réseau (outils de communication et/ou de collaboration).

Bien que l'usage de ces outils soit de plus en plus intégré au dispositif de formation, il reste une innovation pour laquelle peu de connaissances empiriques stabilisées sont disponibles à ce jour. Dans ce contexte, la présente contribution expose une recherche quantitative à partir de données qualitatives recueillies auprès d'apprenants adultes en formation à distance réalisant des activités collectives en trinôme. L'enjeu de cette

étude est d'identifier les types d'outils en réseau utilisés pour le travail collectif et leurs fonctions dans le développement des interactions socio-cognitives et socio-affectives entre pairs. Elle constitue le premier volet d'une recherche longitudinale qui ambitionne de questionner l'évolution dans le temps de ces liens entre outils et interactions sociales.

Après que la notion d'interaction sociale d'un point de vue cognitif puis d'un point de vue affectif a été définie, les outils en réseau, supports des interactions, sont introduits. Sont présentés ensuite les principaux aspects méthodologiques de recueil des données collectées sur le terrain. Enfin, une synthèse des résultats est proposée et des perspectives sont envisagées.

2 Cadrage conceptuel

2.1 La dimension socio-cognitive des interactions entre pairs

Dans son modèle de la présence en *e-learning*[1], Jézégou [3] décrit les interactions socio-cognitives comme « des interactions sociales de confrontation de points de vue, d'ajustements mutuels, de négociations et de délibérations pour résoudre de façon commune et conjointe une situation problématique ». Dans ce contexte de co-résolution, les recherches soulignant l'efficacité des stratégies collectives sur l'apprentissage attestent de l'importance de l'interdépendance sociale, c'est-à-dire des liens que tissent des individus dans un groupe [1]. Cette interdépendance, si elle est positive, appelle les membres du groupe à exprimer leurs intérêts et objectifs au regard d'un but partagé et à fournir un effort de coordination constant [2] ; l'interdépendance sociale s'inscrit dans une dynamique d'actions où la responsabilité est collective.

2.2 La dimension socio-affective des interactions entre pairs

Outre-Atlantique, les nombreuses recherches qui se sont intéressées aux interactions entre apprenants au sein d'un environnement distant se réfèrent principalement au modèle de la *Community of Inquiry*[2] [4] et notamment à l'une de ses dimensions, la présence sociale, décrite comme étant la projection des caractéristiques personnelles des apprenants exprimant leurs émotions, communiquant ouvertement et s'efforçant par divers moyens d'établir la cohésion au sein de leur communauté d'apprentissage [5]. Développer des relations interpersonnelles affectives susceptibles de bénéficier au groupe devient alors un enjeu majeur pour ses membres. Dans son modèle de la présence en *e-learning*, Jézégou [3] insiste également sur la nécessité d'instaurer des formes d'interactions nécessitant « symétrie de la relation et aménité ». Un rapport

[1] La présence en *e-learning* est définie comme issue de « certaines formes d'interactions sociales entre les apprenants, entre le formateur et les apprenants lorsque ces derniers sont engagés dans une démarche de collaboration à distance au sein d'un espace numérique de communication. ».

[2] Le modèle de *Community of Inquiry* est un cadre de référence au soutien de l'apprentissage en ligne par l'utilisation de la communication assistée par ordinateur.

interactionnel authentique requiert de la part des membres du groupe une attitude d'acceptation les obligeant à considérer l'autre tel qu'il se présente et à reconnaître la valeur de ses contributions.

2.3 Les interactions entre pairs par les outils en réseau

Lors de la réalisation d'une activité en groupe et à distance, les apprenants doivent être en capacité de percevoir l'activité de leurs pairs et, en retour, de rendre visible et intelligible leur propre activité. Dans ce contexte, la recherche convoque le concept de *group awareness*, soit « prise de conscience de groupe » [6], définie comme une compréhension partagée qui offre à chacun des membres du groupe un contexte à sa propre activité et assure la cohérence des contributions individuelles pour le bénéfice du groupe. Ce sont les outils en réseau, supports des interactions entre pairs, qui permettent au groupe de disposer de signaux émanant de chacun de ses membres, de sorte qu'il puisse se représenter les forces en présence en son sein [7] et mesurer les contributions de chacune d'entre elles à la dynamique collective. En outre, ces informations offrent aux apprenants la possibilité d'anticiper les attentes de leurs pairs et d'apporter à ces derniers une aide en cas de difficulté, qu'elle soit ou non exprimée.

3 Cadre méthodologique

3.1 Contexte de l'étude empirique et recueil de données

Le terrain de recherche est le Master deuxième année *Digital Learning Manager* (promotion 2018-2020) proposé en formation à distance à l'Université de Lille par le département Sciences de l'Éducation et de la Formation des Adultes. La formation prescrit des activités individuelles, ainsi que des activités collectives que les apprenants sont invités à réaliser en trinôme en optant, pour interagir et collaborer, pour les outils de leur choix. Cette étude intervient alors que les trinômes ont réalisé deux à trois activités collectives depuis leur constitution. Un recueil de données par entretiens individuels a été effectué auprès de 18 apprenants volontaires (13 femmes et 5 hommes ayant entre 28 et 56 ans) dans l'objectif d'appréhender la façon dont ceux-ci réalisent les activités avec leur trinôme et le sens qu'ils donnent aux interactions entre pairs.

3.2 Analyse des données

Les propos recueillis ont fait l'objet d'analyses statistiques (au moyen du logiciel libre IRaMuteQ) s'appuyant sur la méthode de classification décrite par Reinert [8], qui appréhende de manière quantitative et synthétique les informations contenues dans un ensemble de données textuelles. Une classification hiérarchique descendante a permis de grouper les segments de texte les plus proches et de répartir les formes issues de ces segments en classes lexicales. À partir de celle-ci, une analyse factorielle des correspondances (AFC) a été générée afin d'étudier la proximité des classes entre elles.

4 Synthèse des résultats

Le résultat de l'AFC (Fig.1) résume 73,66% de l'information globale du corpus et projette sur un plan, selon deux facteurs d'inertie, quatre classes de formes lexicales colorées.

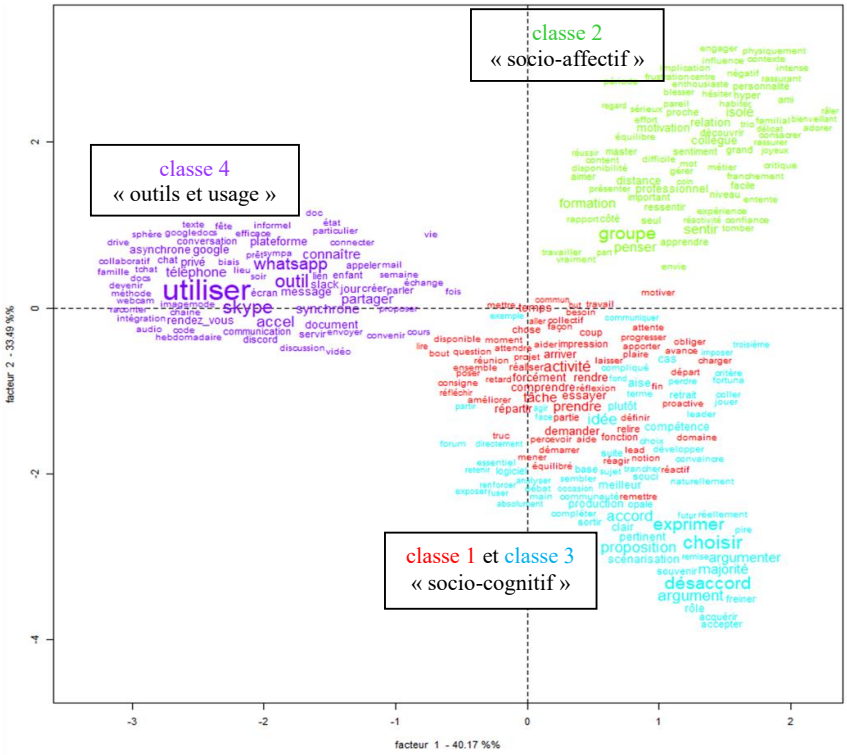


Fig. 1. Analyse Factorielle des Correspondances (Méthode Reinert)

L'axe horizontal (facteur 1) montre une opposition assez marquée entre la classe 4 et les classes 1, 2 et 3. Il renvoie aux composantes du travail collectif à distance (de la composante technique en abscisse négative à la composante humaine en abscisse positive). La classe 4 se rapporte aux outils et à leurs usages. Les classes 1 et 3 se regroupent et appréhendent la réalisation de l'activité collective d'un point de vue socio-cognitif : apparaissent notamment au sein de ces deux classes les mots « désaccord », « exprimer », « argumenter », « tâche », « répartir ». La classe 2, quant à elle, appréhende le groupe d'un point de vue socio-affectif : sont contenus dans cette classe les mots « ressentir », « collègue », « relation », « découvrir ».

L'axe vertical (facteur 2) renvoie à la dimension sociale du travail collectif à distance : il oppose assez nettement les classes 1 et 3 (dimension socio-cognitive en ordonnée négative) à la classe 2 (dimension socio-affective en ordonnée positive). La classe 4 se projette de part et d'autre du niveau 0 du facteur 2 ; elle supporte les interactions sociales d'un point de vue technologique.

Afin d'investiguer plus finement les formes projetées par l'AFC, l'usage du concordancier a permis de revenir sur les segments de texte du corpus et de saisir la fonction des outils dans le cadre des interactions entre pairs.

Types d'outils en réseau utilisés pour le travail collectif

L'usage du concordancier à partir de la forme *outil*, présente significativement dans la classe 4 (extrait ci-dessous), montre que les apprenants privilégient des outils en réseau déjà usités avant l'entrée en formation et de préférence accessibles sur smartphone : visioconférence, éditeur de texte partagé, messagerie instantanée, plateforme de communication collaborative. Ces outils sont perçus comme propices au travail de groupe et employés aussi bien en mode synchrone qu'asynchrone.

****** *etu_15 *gpe_5**

les **outils** qu on utilise aujourd hui sont ceux qu on utilisait les uns les autres dans notre vie personnelle on a été au plus simple en fait

Outils en réseau et interactions socio-cognitives

La perception des interactions entre pairs est majoritairement liée à la réalisation de la tâche collective prescrite : sur l'axe horizontal de l'AFC, la classe 4 (*outils et usage*) apparaît plus proche des classes 1 et 3, liées au *socio-cognitif*, que de la classe 2, relative au *socio-affectif*. En début d'activité, les apprenants ressentent le besoin de s'assurer d'une compréhension partagée de la consigne afin d'orienter les actions collectives et, dans ce but, privilégient souvent les points synchrones par visioconférence, permettant des échanges spontanés, fluides, ainsi que le débat. Dans ce cadre, la visioconférence est couplée à un éditeur de texte partagé servant à ébaucher la structure de l'activité (voir extrait du concordancier ci-dessous, *etu_6*gpe_2) ; par la suite, ce second outil est plutôt utilisé en asynchrone. Les messageries instantanées répondent au besoin de réactivité des apprenants : elles sont adoptées dans le but d'effectuer des points organisationnels réguliers ou d'interagir autour de notions de cours simples. Le partage d'écran est perçu comme un moyen de s'entraider pour surmonter des obstacles techniques liés à l'activité. Ces outils, par leur complémentarité, permettent aux apprenants d'entretenir un lien permanent à la tâche et d'avoir une vision structurée de la progression de l'activité (voir extrait du concordancier ci-dessous, *etu_10*gpe_4).

****** *etu_6 *gpe_2**

on commence par vérifier que l on comprend de la même manière la consigne ça ça se fait plutôt sur **whatsapp** et sur un document partagé

****** *etu_10 *gpe_4**

on utilise whatsapp sur notre **téléphone** cet outil nous sert au quotidien pour se dire où on en est dans la progression de l activité

Outils en réseau et interactions socio-affectives

Lors de la réalisation d'activités collectives se développent progressivement de nombreuses interactions socio-affectives, même si elles ne sont pas prioritairement recherchées : sur l'axe vertical de l'AFC, la classe 4 (*outils et usage*) tend à se rapprocher de la classe 2 (*socio-affectif*). Les outils en réseau utilisés offrent un espace de dialogue informel, ils permettent d'interagir sur d'autres sujets que ceux strictement liés aux activités et d'ainsi mieux se connaître (voir extrait du concordancier ci-dessous,

*etu_8*gpe_3). La messagerie instantanée instaure un lien de proximité, elle est un espace de soutien moral, de réassurance. La visioconférence humanise les relations en ouvrant une fenêtre sur l'environnement privé. En communication asynchrone, les messages sont agrémentés d'émojis dont les apprenants usent pour apporter chaleur humaine, humour et légèreté ; les émojis sont également des didascalies qui rendent les messages plus intelligibles et en minorent le risque de mauvaise interprétation. Les interactions régulières sont perçues par la plupart des trinômes comme un rempart à l'isolement (voir extrait ci-dessous, *etu_17*gpe_6).

****** *etu_8 *gpe_3**

en ce qui concerne les relations humaines au_delà du master on est en contact permanent sur **whatsapp** ça permet de discuter d autre chose aussi

****** *etu_17 *gpe_6**

on s envoit des petits messages on partage nos états d âmes des messages un peu plus **informels** ça rassure je ne me sens pas seule

Perspectives

Une seconde étude, menée six mois après la présente et la complétant, permettra de révéler l'évolution dans le temps des liens entre outils et interactions sociales.

Remerciements : L'auteure remercie le programme de coopération transfrontalière territoriale européenne Interreg Dig-e-Lab qui soutient cette recherche.

Références

1. Johnson, D. W. et Johnson, R. T. (1999). The three Cs of school and classroom management. Dans H. Freiberg (dir.), Beyond behaviorism : Changing the classroom management paradigm (p.119-144). Boston : Allyn & Bacon.
2. Dillenbourg, P., Baker, M., Blaye, A. et O'Malley, C. (1995). The Evolution of Research on Collaborative Learning. Dans E. Spada et P. Reiman (dir.), *Learning in Humans and Machine:Towards an Interdisciplinary Learning Science* (p.189-211). Oxford : Elsevier.
3. Jézégou, A. (2014). Le modèle de la présence en e-learning. Une modélisation théorique au service de la pratique, notamment en contexte universitaire. Dans G. Lameul, C. Loisy (dir.), *La pédagogie universitaire à l'heure du numérique* (p. 111-120). Bruxelles : De Boeck.
4. Garrison, D. R., Anderson, T. et Archer, W. (2000). Critical Inquiry in a Text-based Environment : Computer Conferencing in Higher Education. *Internet and Higher Education*, *2*(2), 87-105.
5. Rourke, L., Anderson, T., Garrison, D. R. et Archer, W. (2001). Assessing teaching presence in a computer conferencing environment. *Journal of Asynchronous Learning Networks*, *5*(2), 1-17.
6. Dourish, P. et Bellotti, V. (1992). Awareness and Coordination in Shared Workspaces. *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*, 107-114). Toronto, Ontario, Canada.
7. Carroll, J.M., Rosson, M.B., Convertino, G. et Ganoe, C. (2006). Awareness and teamwork in computer-supported collaborations. *Interacting with Computers*, *18*, 21-46.
8. Ratinaud, P. et Déjean, S. (2009). IRaMuTeQ : implémentation de la méthode ALCESTE d'analyse de texte dans un logiciel libre, *Modélisation Appliquée aux Sciences Humaines et Sociales*, 8-9.

Session
Adoption des EIAH

TGRIS : Outil de simulation pour la gestion des émotions dans la formation à l'entretien des conseillers pédagogiques

Iza Marfisi-Schottman[1], Isabelle Vinatier[2] et Elisabetta Bevacqua[3]

[1] LIUM, EA 4023, Le Mans Université 72085, le Mans, France
iza.marfisi@univ-lemans.fr
[2] CREN, EA 2661, Université de Nantes 44300, Nantes, France
isabelle.vinatier@univ-nantes.fr
[3] Lab-STICC, UMR 6285, CNRS, ENIB 29280, Plouzané, France
bevacqua@enib.fr

Résumé. Se former à la gestion des émotions dans les métiers de l'humain passe généralement par des jeux de rôle. Il est difficile de systématiser ce procédé car il est coûteux (comédiens) où difficile à mettre en place si les étudiants jouent un rôle en devenir pour eux. Pour la formation de conseillers pédagogiques de l'enseignement (CP), nous proposons TGRIS, un outil de simulation d'entretien, en Réalité Virtuelle, qui permet aux CP de discuter naturellement avec un enseignant débutant virtuel (*i.e.* simulé par un agent conversationnel animé). Cet outil présente la particularité d'utiliser des échanges prélevés en situation réelle d'entretien de conseil problématique. De plus, les réactions verbales, non-verbales et l'état émotionnel de l'enseignant débutant virtuel sont pilotées par le formateur. La recherche explore comment ce système permet aux CP d'apprendre à dialoguer avec un enseignant débutant qui résiste au conseil. L'expérimentation avec 12 CP montre le potentiel d'un dispositif de formation instrumenté avec TGRIS.

Mots-clés. Simulateur, formation, émotion, réalité virtuelle, outil auteur, agent conversationnel animé.

Abstract. The ability to deal with one's emotions in professional situations that involve human relations is generally taught with role playing. However, role playing is expensive (comedians) or difficult to set up if the students are playing a role that is not theirs yet. In order to improve the training offered to Education Counselors (ECs), we propose the TGRIS virtual reality simulator, that allows ECs to naturally discuss with a virtual novice teacher (i.e. simulated by an embodied conversational agent). This tool has the particularity of using exchanges extracted from real problematic counselling situations. In addition, the verbal and non-verbal responses as well as the emotional state of the virtual novice teachers are piloted by the teacher. This research explores how this system allows ECs to learn how to dialogue with novice teachers who resist to their help. The experimentation lead with 12 ECs shows the potential of instrumented training with TGRIS.

Keywords. Simulator, teaching, emotions, virtual reality, authoring tool, Embodied conversational agents.

1 Introduction

Apprendre à gérer les charges émotionnelles inhérentes aux situations professionnelles est très important dans de nombreux domaines alors même que le travail sur et avec les émotions reste généralement invisible [1]. Une caractéristique des émotions est leur communicabilité entre les personnes en raison des neurones miroirs [2]. En effet, dans l'interaction en face à face, se trouve à l'œuvre un processus d'échoïsation corporelle définit comme : « *le mécanisme par lequel un observateur vit en miroir l'état affectif de celui qu'il observe et à qui il s'adresse. Il s'agit d'un mécanisme instinctif de captation des indices non verbaux d'expression émotionnelle* » [3].

L'apprentissage du contrôle des émotions est particulièrement développé dans le secteur de la santé [4][5] mais également dans le secteur du commerce et de la communication pour former au management, à la négociation, etc. Dans le domaine de l'enseignement, la prise en compte des émotions fait l'objet d'un engouement récent. Il est constaté une augmentation des références aux émotions dans les programmes scolaires et les initiatives pédagogiques sur ce thème sont diverses [6]. Un autre aspect se développe, il concerne la formation des enseignants par la prises en compte des émotions dans l'analyse de leurs pratiques professionnelles.

C'est le cas notamment avec des Conseillers Pédagogiques (CP) qui doivent mener des entretiens avec des enseignants novices résistants. Le travail des CP consiste à former les enseignants débutants pendant leurs deux premières années. Trois à cinq fois par an, ces derniers sont observés par les CP pendant qu'ils enseignent. Ces observations se prolongent par un entretien de conseil avec, en vue, l'amélioration de leurs pratiques d'enseignement. Or, il arrive que les CP se trouvent profondément déstabilisés et démunis face à des enseignants en difficulté et agressifs. La formation institutionnelle qu'ils suivent ne les aide pas à faire face à ces situations d'entretiens difficiles. Il y a six ans, 15 CP de l'académie de Nantes ont pris l'initiative de partager leur expérience dans le cadre de l'Association Nationale des Conseillers Pédagogiques (ANCP, 44). L'accompagnement de l'analyse de leur activité par un chercheur a débouché sur la mise en place d'une formation officielle, en collaboration avec le laboratoire de recherche du CREN [7].

Comme il est impossible de filmer les entretiens problématiques, étant donné le contexte déjà tendu, les CP enregistrent et retranscrivent leurs entretiens pour les analyser. Or, ces transcriptions ne contiennent ni des données sur la communication non verbale (c.-à-d. gestes, attitudes et postures) ni la prosodie et modulations de la voix qui accompagnent les prises de parole. Si les enjeux en tension dans les entretiens ont clairement été identifiés (enjeux épistémiques, enjeux relationnels et pragmatiques [8]), les émotions auxquelles se trouvaient confrontés les CP avaient besoin d'être explicitées et débattues dans un cadre sécurisant, celui de la formation. Le système de simulation d'entretien que nous proposons est centré sur des situations réelles de travail problématiques. La conception de formation visée est basée sur l'idée qu'elle doit être conçue à partir de l'analyse de l'activité réelle des formés et notamment à partir de leurs situations professionnelles problématiques. Cette conception s'inscrit dans le champ théorique de « *la didactique professionnelle* » [9].

Dans la section suivante nous faisons un point sur des travaux portant sur la prise en compte des émotions dans la formation, pour ensuite positionner notre propre conception de formation.

2 Les émotions dans la formation des enseignants

La formation à la gestion des émotions dans l'enseignement est basée sur l'analyse de vidéos de classe et/ou sur l'invitation à exprimer le vécu subjectif lié à telle ou telle situation difficile [10]. Dans ce domaine l'intérêt porté à la gestion des émotions concerne spécifiquement les enseignants entrant dans le métier et qu'il s'agit d'aider à gérer leur stress [11]. Pour Rouve et Ria [12][13], les enseignants débutants doivent comprendre que les difficultés qu'ils rencontrent ne sont pas imputables à des traits personnels mais plutôt à des traits génériques d'une communauté professionnelle en cours de développement (celle des enseignants débutants). Les raisons les plus souvent évoquées par un enseignant pour justifier un niveau de stress élevé sont liées à l'interaction avec un élève agressif et à la charge de travail [11]. Tel que relevé par Théorêt *et al* [14], le facteur de protection le plus souvent mis en évidence par les enseignants résilients sont leurs habiletés et compétences professionnelles.

De manière générale, les formations qui intègrent la question des émotions dans l'enseignement ne portent pas sur une ré-évocation vécue de situations réelles porteuses de tension dans le cadre de la formation. En ce qui nous concerne, analyser collectivement des cas réels et problématiques rapportés par des CP permet de cerner le sens de ce qu'ils font effectivement [15]. Même si ces séances d'analyse sont très utiles, la formation à la gestion de leurs émotions est encore une dimension de leur activité professionnelle largement à explorer. C'est la raison pour laquelle il nous a semblé opportun d'associer à ces analyses un travail sur les émotions via un outil de simulation d'entretien, avec une question fondamentale : **quel est l'outil numérique en mesure d'immerger les CP dans des situations réalistes pour leur faire vivre des émotions, et de les analyser dans le cadre sécurisant de la formation ?**

Plusieurs outils de simulation d'entretien ont été développés au cours des dernières années pour offrir une solution à ce problème. Dans la section 3, nous analysons les fonctionnalités de ces outils. Dans la 4ème section, nous proposons TGRIS, un simulateur d'entretien qui permet au formateur de créer des situations émotionnelles authentiques. Cet outil a été mis au point avec le groupe de CP cité ci-dessus, dans une démarche de conception centrée sur l'utilisateur [16]. Pour finir, nous présentons les résultats de la première expérimentation avec un groupe de CP débutants.

3 Outils de simulation d'entretien

Ces dernières années, plusieurs entreprises ont tenté de répondre au besoin de simulateurs d'entretien. *SIMmersion*[1] (fig. 1), par exemple, offre un système de simulation avec des vidéos d'acteurs pour se former sur plusieurs scénarios d'entretien d'embauche et de vente. *VTS Editor*[2] (fig. 1) propose la même chose mais avec un Agent Conversationnel Animé (ACA). Il permet également de créer de nouveaux scénarios en configurant l'arrière-plan, l'apparence des ACA et l'arborescence de dialogue qui contient toutes les réponses possibles pour l'ACA et l'apprenant. Toutefois, les inte-

1 http://simmersion.com/
2 https://seriousfactory.com/logiciel-auteur-vts-editor/

ractions que proposent ces simulations ne permettent pas de créer des situations authentiques, porteuses de charges émotionnelles. En effet, ces interactions présentent trois inconvénients majeurs. Premièrement, l'apprenant clique sur des phrases et ne travaille donc pas ses compétences en communication orale. Deuxièmement, les réponses sont prédéfinies et l'apprenant ne travaille donc pas la construction de ses propres phrases. Pour finir, la posture et les expressions faciales de l'apprenant ne sont pas du tout prises en compte par le système.

Fig. 1. Choix des réponses dans SIMmersion (gauche) et VTS Editor (droite)

Certains systèmes comme INOTS [17], utilisés par l'armée américaine, proposent une liste de réponses (ex : accepter) et demandent aux apprenants de formuler leurs réponses à voix haute (ex : « oui, commandant, j'accepte cette mission ») avant de cliquer sur la réponse. Même si les apprenants s'entraînent donc à créer et dire leurs phrases, leurs performances orales ne sont pas prises en compte par le système. *ViTA* [18] *et MACH* [19] offrent des interactions beaucoup plus naturelles. Ces systèmes permettent à l'apprenant de discuter naturellement avec l'ACA, jouant le rôle du recruteur. Ce système est capable de détecter certains mots clés pour adapter les interactions de l'ACA. Cependant, ce système ne fonctionne que pour des entretiens où l'ACA est le moteur de la conversation et où il suffit d'adapter légèrement ses réactions à ce que dit l'apprenant. *Tardis* [20] offre un système similaire. Cependant, il utilise aussi certains signaux non verbaux exécutés par l'apprenant, tels que l'expression faciale, la posture, la direction de la tête, et l'intonation de la voix pour adapter le comportement émotionnel et social de l'ACA.

La Réalité Virtuelle (RV) est une des dernières avancées dans le domaine de la simulation. Plusieurs chercheurs ont prouvé l'efficacité de ces environnements pour des besoins de formations spécifiques : Ponder *et al.* [21] ont utilisé des simulations de RV pour s'entraîner à la prise de décision rapide dans des situations stressantes, Bell et Weinstein [22] pour aider les personnes atteintes de troubles mentaux à se former aux entretiens d'embauche et Brinkman *et al.* [23] pour aider les personnes à surmonter leur phobie sociale. Afin de créer des simulations authentiques, qui déclenchent des émotions, les ACA et les éléments de l'environnement virtuel sont pilotés, en temps réel, par un enseignant ou un thérapeute à travers une interface de contrôle.

Le groupe de CP, pourrait grandement bénéficier de ce type de simulation immersive, propice à faire naître des émotions : dans la section suivante, nous proposons donc un outil de RV, permettant aux formateurs de créer leurs propres situations d'entretien et de les analyser en collaboration avec celui qui les expérimente et les formés qui observent la scène (double débriefing).

4 TGRIS : Simulateur d'entretien émotionnel

Nous proposons le dispositif TGRIS (Teacher-Guided Realistic Interview Simulator) dont l'objectif est de permettre aux CP de créer leurs propres situations d'entretien, porteuses de charges émotionnelles, pour former leurs collègues débutants. Ce dispositif n'a pas vocation à remplacer l'analyse des transcriptions ou des séances de jeux de rôle. Il s'agit d'un outil de prise de conscience à des situations d'entretien difficiles, auxquelles ils seront confrontés. Nous posons les trois questions suivantes :

Q1. Quel dispositif de formation permet à des formés de faire l'expérience d'émotions *in situ* ? Nous pensons qu'une formation par les pairs, instrumenter par un dispositif de simulation d'entretien peut répondre à cette question. Au vu de l'état de l'art, nous pensons que ce dispositif doit être piloté par le formateur et proposer des interactions naturelles, c'est-à-dire dans lequel le CP formé parle simplement à un ACA (enseignant débutant résistant).

Q2. Quelles fonctionnalités sont nécessaires pour que les formateurs créaient des situations de simulation possédant un fort pouvoir d'évocation ? Nous pensons qu'un dispositif de simulation d'entretien qui permet au formateur de choisir différents visages pour l'ACA ainsi que de piloter ses émotions, ses gestes et ses interventions peut répondre à cette question.

Q3. Est-ce que la RV aide les formés à prendre conscience plus facilement de leurs émotions ? On peut raisonnablement estimer qu'un dispositif en RV, plus immersif que la version écran, aidera les formés à prendre conscience de leurs émotions.

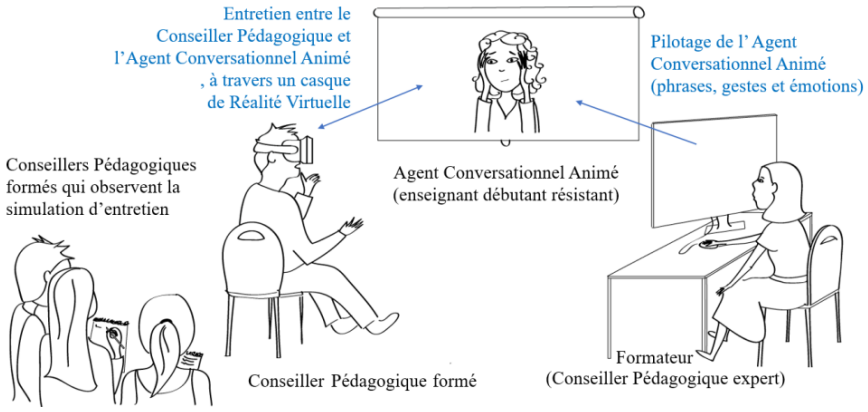


Fig. 2. Dispositif TGRIS

Le dispositif TGRIS est présenté dans la fig. 2. Le CP formé est assis et parle directement à l'ACA via un casque de RV. Grace à un casque audio haute performance, il est coupé du monde réel. Avant de lancer la simulation, le formateur peut choisir l'apparence de l'ACA et choisir un contexte d'entretien (courte introduction qui aidera le formé à situer l'entretien « L'enseignant se montre peu investi et les élèves montent sur les tables »). Les autres formés observent la simulation qui sera analysée lors d'une séance de débriefing juste après.

Pendant la simulation, les réactions de l'ACA sont déclenchées par le formateur en temps réel, à travers une interface de contrôle (fig. 3). Le formateur peut ainsi déclencher ses réactions verbales (interventions), non-verbales (*e.g.* croiser les bras, reculer le torse) et changer ses émotions (*e.g.* colère, anxiété, tristesse). Par défaut, l'interface propose des interventions génériques (*e.g.* « oui », « je ne sais pas ») qui peuvent fonctionner dans n'importe quel entretien. TGRIS permet aussi aux formateurs de créer leurs propres réactions, porteuses de charges émotionnelles spécifiques à la situation d'entretien visée et en fonction de leurs objectifs pédagogiques. Dans notre cas, ce sont les CP qui ont listé les interventions problématiques auxquelles ils avaient été confrontés lors de l'entrée dans l'entretien avec des enseignants débutants.

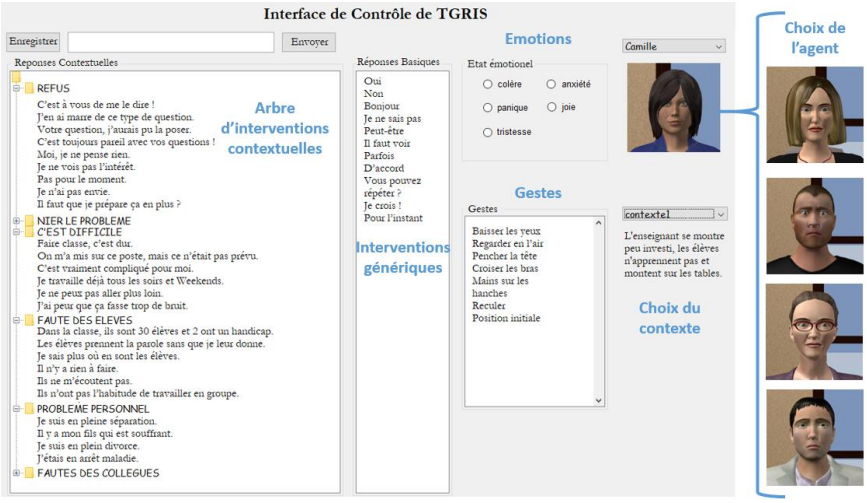


Fig. 3. Interface de contrôle de TGRIS

Ces interventions ont été regroupées en six catégories, représentées sous la forme d'un arbre. Par exemple, la catégorie *Refus* contient les interventions telles que « *J'en ai marre de ce type de question ; c'est à vous de me le dire.* » qui montrent un refus d'écoute de la part de l'enseignant débutant. Certaines de ses catégories sont associées à une émotion (colère pour la catégorie *Refus*). Le formateur peut soit ouvrir la catégorie et choisir une intervention spécifique, soit cliquer sur la catégorie, pour lancer une des réactions au hasard. Les formateurs peuvent ajouter leurs interventions (catégories, interventions et émotions associées) à partir d'un document texte mais également ajouter de nouvelles interventions directement via l'interface de contrôle

TGRIS est construit sur la plateforme d'ACA GRETA. Cette dernière a la spécificité de fournir une très grande palette d'émotions multimodales, combinant le ton de la voix, l'expression faciale et le langage corporel [24]. De plus, les agents de GRETA (fig. 3 droite) réagissent au son de la voix de l'utilisateur en hochant la tête et en changeant la direction de leur regard, donnant l'impression qu'ils écoutent. Même si l'aspect visuel des ACA ne semble pas très réaliste, les auteurs montrent que l'aspect immersif dépend surtout de son comportement et de la cohérence de l'environnement.

TGRIS a été conçu de façon itérative, en collaboration avec les CP. Lors d'une formation en 2016, nous leur avons fait tester la RV et présenté un bref état de l'art sur les outils de simulation. Ils ont immédiatement manifesté leurs intérêts pour ces simulateurs qui leur semblaient propices à faire naître des émotions. Ce sont les CP qui ont souhaité centrer la simulation sur l'entrée de l'entretien avec les enseignants débutants puisqu'il s'agit d'un moment clé, dont dépend la poursuite de l'entretien. Ils ont également travaillé sur les maquettes de l'interface et fourni l'intégralité des interventions issues de leurs analyses d'entretiens problématiques [16].

5 Première expérimentation et résultats

Le but de cette première expérimentation était de trouver des réponses à nos trois questions de recherche, mais aussi de mesurer l'impact de l'intégralité du dispositif de formation instrumentée avec TGRIS.

TGRIS a été utilisé dans le cadre d'une journée de formation par les pairs en juin 2018 avec 12 CP, dont quatre sont expérimentés. Ces derniers ont souhaité jouer le rôle du formateur et piloter l'ACA par binôme. Afin de réduire la taille du groupe, nous l'avons divisé en deux : le matin, un des groupes faisait des jeux de rôle pendant que l'autre testait TGRIS. Les groupes ont été intervertis l'après-midi. Les CP ont testé deux versions de TGRIS (fig. 4). En effet, afin de repérer si la RV facilite la prise de conscience des émotions (Q3), nous avons développé une version de TGRIS dans laquelle l'ACA est affiché sur un écran. Cette version nous permettait également de filmer l'intégralité du visage des CP, ce qui était primordial pour analyser les émotions ressenties. Nous avons également filmé les séances de débriefing qui avaient lieu après chaque simulation d'entretien (jeu de rôle, TGRIS RV ou TGRIS écran).

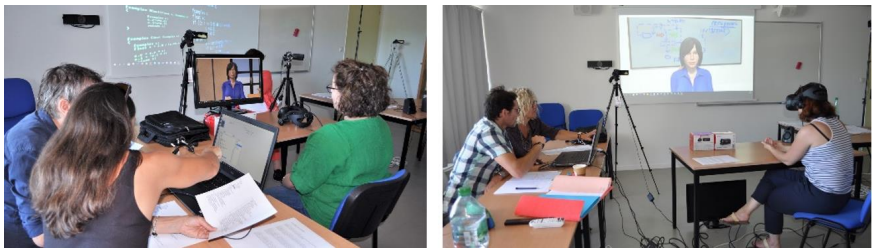


Fig. 4. Expérimentation de TGRIS (deux versions ont été testé : Ecran et Réalité Virtuelle)

Nous avons demandé aux CP de répondre à plusieurs questionnaires ouverts. Le premier questionnaire était donné aux CP juste après chacune de leurs simulations d'entretien vise à savoir s'ils ont vécu des émotions (Q1) : Qu'avez-vous vécu pendant l'entretien ? Cette situation était-elle évocatrice d'une situation vécue ?... Un deuxième questionnaire était donné à la fin de la journée pour mesurer l'utilité des fonctionnalités de TGRIS (Q2) et les différences entre les deux versions (Q3) : Avez-vous apprécié le fait que l'ACA exprime des émotions ? Voyez-vous l'intérêt de piloter l'ACA par le formateur ?...

Sur la journée d'expérimentation, nous avons pu faire passer sept simulations sur TGRIS RV et sept sur TGRIS Ecran. Chaque simulation a duré environ quatre minutes, suivi de 13 minutes de débriefing. Pour chaque simulation, les formateurs ont déclenché en moyenne 10 intervention de l'ACA, cinq gestes et trois émotions. Voici un exemple d'entretien :

Contexte annoncé : les élèves n'apprennent pas et montent sur les tables.

1. CP : Ce matin, j'ai observé ta classe. Je te propose qu'on fasse une petite analyse de ce que j'ai pu percevoir. J'ai vu que c'était un peu compliqué pour toi ce matin et j'aurais voulu savoir comment toi tu te situes dans cette classe.
2. ACA : On n'a pas perdu trop de temps.
3. CP : Alors oui, on n'a pas perdu de temps, pourtant il m'a semblé que c'était un peu compliqué au niveau de la gestion des élèves. L'entrée en classe a été un peu problématique. Est-ce que toi tu as ressenti cette gêne au moment de l'accueil ?
4. ACA : [Émotion : tristesse] C'est toujours la même chose.
5. CP : Alors qu'est-ce que tu entends par là ? Est-ce que tu peux définir un peu mieux ?
6. ACA : Les élèves prennent la parole sans que je leur donne.
7. CP : Effectivement, les élèves prennent la parole sans que tu leur donnes. Est-ce que tu as réfléchi à ce que tu pourrais mettre en place pour essayer de canaliser leurs paroles ?
8. ACA : [Émotion : colère] Ce n'est pas moi qui ai décidé. C'est la titulaire qui m'a dit de faire comme ça. [Geste : bras croisés]
9. CP : D'accord, donc là tu suis exactement ce que vous avez défini avec la titulaire. Est-ce que cela te convient ?
10. ACA : [Émotion : tristesse] Avec la collègue titulaire, c'est difficile….
11. CP : Effectivement, c'est un problème qu'on peut rencontrer quand on travaille à temps partiel. Est-ce que tu as essayé de demander à des collègues de servir de médiateurs pour que vous puissiez avoir une conversation apaisée ?
12. ACA : Non. [Geste : regard baissé]
13. CP : alors, je comprends que ce soit difficile pour toi. Après ce sont des situations qui sont assez courantes quand on débute. Est-ce que tu as essayé d'en parler à tes collègues ?
14. ACA : L'accompagnant des élèves en situation de handicap n'est pas toujours là.

Cet entretien se termine avec une phrase inadaptée de l'ACA. Cela se produit quand les formateurs eux-mêmes sont déstabilisés : ils sont mis en difficulté face à l'habileté d'un conseiller pédagogique qui relance les échanges alors même que l'ACA représente un débutant qui se montre rétif aux échanges.

Tout d'abord, on peut noter que les simulations avec TGRIS ont duré beaucoup plus longtemps que prévu. Alors que nous avions prévu que le CP et l'ACA fassent trois interventions, ils en ont fait plus de 10 en moyenne (soit 20 interventions). C'est très long pour un entretien qui n'avance pas, selon les dire des CP : « c'est très déstabilisant quand l'entretien tourne en rond », « on n'arrive pas à s'accrocher à quelque chose ». Les CP se sont investis dans l'entretien et ne se sont pas sentis bloqués par des réponses inadaptées de l'ACA. De plus, sans qu'ils en aient reçu la consigne, nous pouvons constater que les formateurs cherchent spontanément à déstabiliser les CP novices en simulant des enseignants butés et renfermés. Lors du débriefing, un des formateurs explique son choix : « Ce n'est pas quand quelque chose marche comme on veut que l'on se questionne. ». **Les CP formateurs ont en effet cherché à déstabiliser leurs collègues pour favoriser l'auto-analyse de la situation vécue**. Même si l'expérimention n'a impliqué que 12 participants, les réactions des CP sont frap-

pantes : on repère que les CP bredouillent, ralentissent leur débit de parole, baissent la voix, restent bouche ouverte, se placent en position basse dans les échanges, disent avoir chaud et manifestent leur déstabilisation. Les analyses qualitatives des simulations tendent donc à montrer que le **dispositif TGRIS permet de créer des situations expérientielles engageantes.**

L'objectif de l'expérimentation était également de vérifier si TGRIS permet de créer des situations porteuses de charges émotionnelles (Q1). L'analyse des visages des CP pendant les simulations avec TGRIS Ecran montre des mimiques de stress, des regards fuyants, des sourires nerveux et des pincements de lèvres : autant de signes qui attestent des émotions vécues chez le CP effectuant la simulation, mais également chez les CP qui observent. D'après les questionnaires, plusieurs fonctionnalités de TGRIS ont contribué au fait que les formés vivent des émotions *in situ.* Tout d'abord, la RV, qui était une expérience nouvelle pour tous les CP, en a déstabilisé quelques-uns : *« Bluffant, il faut s'habituer au monde virtuel ! »*. Cependant, **les CP ont surtout été déstabilisés par les réactions des ACA (enseignants débutants), et notamment leur posture de résistance** : *« Une situation de communication virtuelle qui déstabilise, surtout les propos plus les expressions du visage de l'ACA »*, *« Une situation de stress », « Réaction très forte aux propos de l'ACA. » « Très déstabilisant ! », « J'ai été déstabilisée par les propos tenus par l'ACA et désemparée, car contrainte à partir sur un terrain qui n'était pas prévu. »*. Il s'agit d'une différence majeure avec les simulations sous la forme de jeu de rôle, qui n'ont pas du tout bousculé ou déstabilisé les CP : *« Intéressant, en tant qu'acteur, j'aime bien les situations d'improvisation, car cela peut faire apparaître nos habitudes non conscientes »*.

Cette déstabilisation est due au fait qu'un ACA est moins conciliant qu'une vraie personne, qui a tendance à ménager son interlocuteur, surtout quand il s'agit d'un collègue. De plus, le fait d'être face à un ACA, et non pas un humain, permet aux CP de prendre plus de recul et d'être plus honnêtes vis-à-vis des émotions qu'ils ressentent et ainsi de mieux les analyser lors de la séance de débriefing : *« Le fait d'être face à un écran donne une autre distance. On est plus en recul dans notre façon de répondre. », « L'outil permet de se confronter brutalement à la résistance des enseignants. Cette résistance est aussi là avec un vrai enseignant, mais elle est masquée par la politesse. », « L'enchaînement des émotions de l'ACA si elles sont inattendues, suffisent à créer la déstabilisation. Cette déstabilisation et l'analyse à chaud avec les collègues ouvrent un axe de travail personnel. ».* Pour finir, même si la façade virtuelle est importante pour déstabiliser les CP, le fait que les réponses de l'ACA soient pilotées par un humain rend l'entretien crédible et engageant. En effet, 10/12 CP ont trouvé qu'il était bien que l'ACA soit piloté par un humain *« Le formateur peut orienter les réponses pour doser la résistance de l'agent aux conseils du CP ».* En réponse à Q1, on peut donc dire que TGRIS permet de créer des situations expérientielles provoquant des déstabilisations conscientes et donc discutables lors des debriefing.

Le deuxième objectif de l'expérimentation était de mesurer le pouvoir d'évocation du dispositif TGRIS (Q2). D'après les questionnaires, **les CP ont jugé que 12/14 simulations avec TGRIS avaient un fort pouvoir d'évocation** : *« La situation est crédible (même si on reste conscient du factice) », « Plusieurs réponses de l'agent sont en lien avec des entretiens réels vécus dans des situations antérieures », « Des similitudes nombreuses par rapport à des entretiens vécus en termes de résistance des enseignants débutants ».* Les réponses à la question « Qu'avez-vous vécu ? » mon-

trent aussi que les CP, ayant déjà fait face à des enseignants débutants résistants, ont ressenti les mêmes émotions : *« Les émotions ressenties sont très proches de celles vécues face à un enseignant, en situation d'observation. Sentiment de ne plus avoir de réponse à apporter à un moment ! ».* Ces éléments tendent donc à penser que l'ACA proposé par TGRIS et ces interactions, pilotées par le formateur, permettent de recréer des situations proches du réel et donc pertinentes pour la formation des CP débutants.

Pour finir, l'analyse entre les simulations avec TGRIS Ecran et TGRIS RV montre que **la RV aide les formés à prendre conscience plus facilement de leurs émotions** (Q3). En effet, les CP ayant utilisé TGRIS RV ont été autant déstabilisés qu'avec la version écran mais il est apparu que l'effet immersif de la RV a créé une expérience plus intense et les a naturellement amenés à plus se concentrer sur les gestes, mimiques et émotions de l'agent alors qu'ils ne les avaient presque pas perçus avec la version sur ordinateur : *« Une immersion plus totale. Un meilleur ressenti et perception des émotions que la version ordi. Une meilleure implication dans l'échange. », « Cette entrée virtuelle permet de se focaliser sur les mots et l'attitude de l'ACA. ».*

Cette expérimentation a aussi mis en relief le potentiel de l'utilisation de TGRIS comme outil pour la formation. Tout d'abord, les **CP ont trouvé TGRIS très utile pour la formation** (score de 4.5/6 pour TGRIS Ecran et de 4.75/6 pour la version RV). Les séances de débriefing après les simulations ont beaucoup été appréciées par les CP. Il s'agit de moments privilégiés de partage pendant lesquels tous les participants prennent la parole. Cette contribution collective a réellement aidé les CP à questionner leur posture. Voici quelques phrases dites à la fin de la formation qui montrent que les CP ont su porter un regard nouveau sur leurs pratiques : *« Je me suis détaché du fond pour me concentrer sur ce que l'on ne voit pas. », « Quand on va les voir, on a l'impression que les enseignants débutants répondent et qu'ils suivent notre scénario alors que pas du tout ! On prend un raccourci avec l'outil, parce que l'ACA est plus brut que des vrais enseignants ».*

Malgré la conception itérative centrée utilisateur, l'utilisation de TGRIS dans un contexte de formation réel a révélé certaines limites. Tout d'abord, certains des CP les moins expérimentés étaient réticents et stressés à l'idée d'utiliser TGRIS. Au contraire, les CP expérimentés étaient visiblement bien plus à l'aise avec TGRIS et se portaient volontaires. Plus les CP ont travaillé antérieurement à l'analyse de leurs propres activités (ayant donc acquis des clés d'intelligibilité de leur pratique) et plus ils se montrent ouverts à poursuivre ce travail d'analyse sur leurs émotions. Les CP formateurs ont également trouvé qu'il manquait des interventions basiques qu'ils ont pu ajouter directement via l'interface de pilotage. Ils ont également noté qu'il était difficile pour eux de choisir les interventions, les émotions et les gestes rapidement. Pour finir, la prononciation de certains mots dit par l'ACA n'était pas claire.

Pour la prochaine version, nous avons donc convenu d'améliorer la synthèse vocale et de déclencher des gestes automatiquement avec certaines répliques (comme c'est déjà le cas avec des émotions). Nous utiliserons également une interface tactile pour que les deux CP formateurs puissent collaborer plus facilement qu'avec une seule souris. Nous comptons également améliorer le réalisme de l'environnement virtuel en ajoutant des éléments de décoration d'une salle de classe et le bruit de fond d'une récrée (moment pendant lequel les CP font l'entretien). *In fine*, la pertinence d'un tel dispositif devrait se repérer à l'usage qu'ont fait les CP de cette expérimentation dans l'exercice de leur activité réelle.

5 Conclusion et travaux à venir

Dans cet article, nous proposons TGRIS, un outil de simulation d'entretien en Réalité Virtuelle (RV). TGRIS propose des interactions naturelles puisque l'apprenant parle directement à l'Agent Conversationnel Animé (ACA), comme il le ferait avec un humain. C'est le formateur qui déclenche les phrases, les gestes et les émotions de l'ACA, de façon à créer des situations conflictuelles, porteuses de charges émotionnelles. De plus, le formateur peut paramétrer TGRIS avec de nouvelles phrases et émotions afin de l'adapter à ses objectifs pédagogiques.

Notre première expérimentation avec 12 Conseillers Pédagogiques (CP) a montré l'utilité de TGRIS pour leur formation. En effet, le dispositif a permis de créer des situations expérientielles engageantes dans lesquelles les CP ont mené un entretien de conseil avec un enseignant débutant virtuel. Les CP expérimentés, qui jouaient le rôle du formateur, ont naturellement simulé des débutants fermés au dialogue pour mettre les CP formés en difficulté et favoriser leur auto-analyse de la situation vécue. Les simulations ont semblé pertinentes, car les CP ont réellement été déstabilisés par les réactions de l'ACA, et notamment sa posture de résistance. Ces simulations ont été porteuses d'une charge émotionnelle importante que les CP n'avaient pas vécue pendant les jeux de rôle. Ces émotions sont d'autant plus fortes avec la version de TGRIS en RV. Les phrases de débriefing conduites juste après chaque simulation ont été particulièrement riches du fait que les CP avaient été déstabilisés. Ces séances ont permis aux CP d'analyser les émotions ressenties et les situations d'entretien réelles que la simulation leur avait évoquées mais surtout d'apporter un regard critique sur leurs pratiques.

Au vu du potentiel du dispositif TGRIS et l'enthousiasme des CP, nous avons prévu une autre expérimentation en mars 2019. Nous cherchons également d'autres terrains d'expérimentation pour valider le fait que TGRIS puisse s'adapter à d'autres types de formation. Notre défi pour la prochaine version de TGRIS est d'améliorer l'interface de pilotage de l'agent. Celle-ci doit offrir une grande variété de réponses tout en étant assez simple pour que le formateur puisse déclencher la bonne réponse rapidement. Nous envisageons d'ajouter des fonctions pour automatiser les gestes de l'agent, pour que le formateur puisse se concentrer uniquement sur l'anticipation des interventions. Nous travaillons aussi sur la possibilité d'utiliser de la retranscription automatique pour analyser les interventions dites par l'apprenant afin de présélectionner les réponses possibles de l'agent. Pour finir, nous aimerions ajouter une fonction d'enregistrement afin que le formateur et les apprenants puissent rejouer la simulation lors des séances de débriefing.

Remerciements. Ces travaux de recherche ont été financés par Atlanstic 2020.

Références

1. Molinier P., « Éditorial », *Travailler*, Vol. 9, No. 1 (2003) 5–8
2. Rizzolatti G. et Sinigaglia C., *Les neurones miroirs*. Odile Jacob (2008)

3. Cosnier J., *Emotions et sentiments* (version actualisée de *Psychologie des émotions et sentiments*. Paris : Retz, 1994) (2006)
4. Saint-Dizier de Almeida V., « Comment améliorer la compréhension de l'entretien d'annonce de diagnostics médicaux sérieux », *Interv. sur Trav.*, Vol. 10, No. 2 (2013)
5. Béguin P. et Weill-Fassina A., « De la simulation des situations de travail à la situation de simulation », Dans La simulation en ergonomie, Octarès, Toulouse (1997) 5–28
6. Cuisinier F., « Émotions et apprentissages scolaires : quelles pistes pour la formation des enseignants ? », *Rech. Form.* (2016) 9–21
7. Vinatier I., « Des entretiens de conseil problématiques entre conseillers pédagogiques et enseignants débutants : dissémination ou congruence. », Dans La part du langage : pratiques professionnelles en formation, Paris, Harmattan (2015) 61–92
8. Vinatier I., *Le travail de l'enseignant : Une approche par la didactique professionnelle*, Bruxelles, De Boeck Education (2013)
9. Pastré P., « La didactique professionnelle. Approche anthropologique du développement chez les adultes. », *Formation et Pratiques Prof.,* Vol. 8, No. 2, (2011)
10. Cuisinier F., Tornare E., et Pons, « Les émotions dans les apprentissages scolaires : un domaine de recherche en émergence. », *Appr. Neuro. Apprent. Enfant*, Vol. 137 (2015) 1–9
11. Royer *et al.*, « Le stress des enseignants québécois à diverses étapes de leur carrière », *Vie pédagogique* (2001) 5–8
12. Ria L., « Les émotions au cœur de l'action des enseignants novices. », In : Les émotions, Revue E.P.S., Paris (2005) 99–113
13. Rouve M.-E. et Ria L., « Analyse de l'activité professionnelle d'enseignants néo-titulaires en réseau « ambition réussite » : études de cas. », *Trav. Form. En Éducation*, Vol. 1 (2019)
14. Théoret M. *et al., Analyse de la résilience chez les personnels scolaires œuvrant en milieux défavorisés.* Univ. Montréal, CSIM (2004)
15. Vinatier I., « Pour une didactique professionnelle de l'enseignement », *Rev. Fr. Pédagogie*, Vol. 175 (2011) 155–156
16. Vinatier I. et Marfisi-Schottman I., « L'irruption des émotions entre conseillers pédagogiques et enseignants débutants », Dans L'analyse des interactions dans le travail, Raison et Passions, Dijon (2018)
17. Hays M. J. *et al.,* « Can Role-Play with Virtual Humans Teach Interpersonal Skills? », *In proceedings of Interservice Industry Training, Simulation, and Education Conf.* (2012)
18. Burke S. L. *et al.*, « Using Virtual Interactive Training Agents with Adults with Autism and Other Developmental Disabilities », *J. Autism Dev. Disord.*, Vol. 48, No. 3 (2018) 905–912
19. Hoque M., *et al.*, « MACH: My Automated Conversation Coach », *Proceedings of the Int. Joint Conf. on Pervasive and Ubiquitous Computing*, NY, USA (2013) 697–706
20. Anderson K. *et al.,* « The TARDIS Framework: Intelligent Virtual Agents for Social Coaching in Job Interviews », *Proceedings of the Int. Conf. on Advances in Computer Entertainment*, Berlin, Heidelberg, Vol. 8253, (2013) 476–491
21. Ponder M. *et al.,* « Immersive VR decision training: telling interactive stories featuring advanced virtual human simulation technologies », *Proceedings of the workshop on Virtual environments*, NY, USA (2003) 97–106
22. Bell M. D. et Weinstein A., « Simulated Job Interview Skill Training for People with Psychiatric Disability: Feasibility and Tolerability of Virtual Reality Training », *Schizophr. Bull.*, Vol. 37, No. suppl 2 (2011) 91–97
23. Brinkman W.-P. *et al.,* « A Virtual Reality Dialogue System for the Treatment of Social Phobia », *Proceedings of Computer Human Interact.*, NY, USA (2012) 1099–1102
24. Bevacqua, E., Prepin, K., Niewiadomski, R., de Sevin, E., Pelachaud, C. GRETA: Towards an Interactive Conversational Virtual Companion, In Y. Wilks, J. Benjamins, Artificial Companions in Society: Perspectives on the Present and Future (2010) 143–156

Impact de la personnalisation d'une interface graphique sur l'intention d'adoption d'un EIAH

Arnaud Zeller[1], Najoua Mohib[2]

[1] Université de Strasbourg, LISEC EA 2310, 7 rue de l'université, F-67000 Strasbourg, France
zellera@unistra.fr
[2] Université de Strasbourg, LISEC EA 2310, 7 rue de l'université, F-67000 Strasbourg, France
najoua.mohib@unistra.fr

Résumé. Dans cette étude, nous étudions l'impact d'une activité de personnalisation d'une interface graphique par les apprenants sur leur intention d'adoption de l'EIAH. L'analyse conduite s'appuie sur une exploitation des traces d'interactions de l'apprenant avec l'interface d'un logiciel prototype de traitement de texte doté d'un module de collecte des données et de fonctions de personnalisation avancées. Les résultats suggèrent que plusieurs variables relatives à l'intention d'usage d'un EIAH, du point de vue de la flexibilité de son interface, expliquent, en partie, son acceptation dans le processus d'apprentissage. L'activité de personnalisation d'un environnement se présente comme un élément déterminant de l'effet de ces variables.

Mots-clés. Modèle UTAUT, *learning analytics*, personnalisation de l'interface

Abstract. This research investigates the impact of an activity of personalization of a graphical user interface by the learners, on their behavioral intention to use the VLE. The analysis conducted is based on the exploitation of interaction traces between the learner and the interface of a word processor with a management module of learning analytics and advanced personalization functions. Several variables related to easiness of use of the VLE partly explain its acceptance. The activity of personalizing an environment appears to be a decisive element of the effect of these variables.

Keywords UTAUT model, learning analytics, interface personalization

1 Introduction

La question de la personnalisation de l'apprentissage est un sujet central dans le domaine de la recherche sur les EIAH (Environnements Informatiques pour l'Apprentissage Humain). Si les environnements numériques supports sont aujourd'hui, du point de vue du système [1] de plus en plus fonctionnels, la question de leur adoption par les apprenants reste posée malgré les efforts et les moyens d'ingénierie pédagogique consentis. A défaut de facilité et de rapidité de leur prise en main, l'accroissement des fonctionnalités généralement implémentées permet d'adapter des trajectoires d'usages qui se trouvent souvent contrariées par l'émergence de conflits instrumentaux [2]. Le développement de systèmes adaptatifs et de tuteurs

intelligents [3], [4], [5], [6] ne dispensent pas les apprenants des efforts requis par la genèse instrumentale [7], processus dans lequel nous pensons que l'interface de l'EIAH, du point de vue de son niveau de flexibilité, occupe une place décisive.

Dans cet article, nous nous intéressons aux possibilités offertes à l'utilisateur, au sein d'un EIAH, d'adapter lui-même l'interface. D'après Tchounikine [8], un EIAH « est un logiciel spécifiquement conçu dans le but d'amener un apprenant à développer une activité favorable à l'atteinte des objectifs de la SPI [situation pédagogique informatisée] ou de la SAI [situation d'apprentissage informatisée] considérée (…) Le logiciel utilisé dans une SPI peut être un logiciel conçu spécifiquement pour des situations pédagogiques ou un logiciel quelconque utilisé, en la circonstance, à des fins pédagogiques ». C'est le cas dans notre étude puisque le logiciel de traitement de texte, dont il est question, a été spécialement conçu pour permettre à des élèves d'apprendre à personnaliser une interface afin de faciliter l'appropriation du langage de commande. Notre travail repose sur l'idée que, dans ce type de logiciel, les caractéristiques des objets métaphoriques, déclencheurs d'actions et dépendants de l'interface graphique, peuvent être des facteurs d'acceptation de l'environnement numérique. Dit autrement, nous nous attendons à ce que la personnalisation d'une interface graphique favorise l'intention d'adoption d'un EIAH.

2 État de l'art

2.1 UTAUT et intention d'adoption

De nombreux travaux ont été menés sur l'adoption d'un produit ou d'une technologie notamment dans la communauté IHM (Interfaces Homme Machine). Cependant, la plupart des études restent focalisées sur l'évaluation de spécifications fonctionnelles d'un produit, laissant de côté l'évaluation des aspects cognitifs et psychologiques. A cet égard, les notions d'acceptabilité, d'acceptation, d'appropriation et d'intention d'usages viennent compléter l'évaluation purement fonctionnelle.

Rappelons-en brièvement leurs définitions. Barcenilla et Bastien [9] définissent le terme d'acceptabilité comme le degré d'intégration et d'appropriation d'un objet dans un contexte d'usage. L'acceptation se définit quant à elle par l'évaluation de l'usage pour les tâches pour lesquelles le système a été conçu [10]. L'appropriation renvoie à la façon dont l'individu investit personnellement l'objet ou le système selon l'adéquation qu'il y trouve avec ses valeurs personnelles et culturelles, cette adéquation lui donnant envie d'agir au-delà de subir l'usage [9]. L'intention d'adoption qui influence et rend prédictible le comportement individuel [11], [12] est un préalable au comportement d'usage en référence à la théorie unifiée de l'acceptation des technologies (UTAUT) de Venkatesh *et al*. [13].

D'après ce modèle, quatre facteurs peuvent expliquer l'intention d'adoption et d'utilisation des technologies (*behavioral intention and use behavior*) : 1) les attentes en termes de performances (*performance expectancy*) associées au degré selon lequel un individu croit que l'utilisation du système l'aidera à atteindre un but ; 2) les attentes en termes d'effort (*effort expectancy*) associées au degré de facilité associé à l'usage du système ; 3) les influences sociales (*social influence*) associées au degré selon lequel

un individu perçoit que les personnes qui lui sont importantes croient qu'il ou elle devrait utiliser le nouveau système et 4) les conditions facilitatrices (*facilitation condition*) associées au degré selon lequel un individu perçoit qu'une organisation ou une aide existent pour l'aider dans l'utilisation du système. A ces facteurs, s'ajoutent quatre autres variables modératrices (l'âge, le sexe, l'expérience et le volontarisme « d'utilisation ») qui influenceraient également l'intention d'usage.

Nous proposons ici de mobiliser le modèle UTAUT pour étudier l'impact de la personnalisation d'une interface graphique sur l'intention d'adoption d'un EIAH par des apprenants.

2.2 Interface graphique

Une interface vise à permettre de communiquer avec la machine afin de réaliser une ou des tâches sur un ordinateur qui met en œuvre des procédures afin d'y parvenir. Historiquement, le modèle mental retenu dans la conception de l'interface homme-machine est celui de l'utilisateur considéré comme système interactif [14], [15]. D'abord représentés sous forme de commande en ligne, les interfaces sont devenues graphiques, à partir des années 70-80 [16], en s'appuyant, encore aujourd'hui, sur la métaphore du bureau. Elles se caractérisent désormais par leur plasticité [17], [18] et leur flexibilité [19]. La plasticité des interfaces rend compte des variations de l'utilisateur de la plate-forme et de l'environnement, comme forme d'adaptation aux conflits dans un contexte d'usage, sans altérer l'utilité et l'utilisabilité. La flexibilité de l'interface renvoie aux moyens mis à disposition des utilisateurs pour personnaliser l'interface afin de rendre compte de leurs stratégies ou habitudes de travail et des exigences de la tâche. En nous référant aux travaux de Venkatesh *et al*. [11], nous envisageons l'interface graphique comme une « infrastructure technique » reliant les dimensions pédagogiques et la partie système de l'EIAH.

Nous considérons sous l'angle de la flexibilité et au-delà de la plasticité, le processus d'appropriation d'une interface graphique d'un EIAH comme contribuant aux intentions d'adoption. A cet égard, les logiciels de traitement de texte ne présentent pas *a priori* d'intention pédagogique car d'abord destinés à la production de document. Ils deviennent un EIAH quand ils sont associés à une situation pédagogique informatisée (SPI), telle que l'écriture inventée ou l'apprentissage de compétences orthographiques par l'usage du correcteur ou une activité de collecte de traces [8]. Nous pensons que, dans le cas du traitement de texte, l'appropriation d'une interface graphique peut être mise au jour à travers l'analyse des traces des activités de navigation, de saisie et d'édition de texte qui concourent à la personnalisation de l'interface, participant ainsi à l'intention d'adoption.

Les recherches menées dans le domaine des *Learning Analytics*, portant sur l'étude de traces à partir des interfaces, s'appuient habituellement sur des *plug-in* tels que TraceMe, des environnements dotés de modules de collecte tels que Moodle, Tactiléo Map et sur des outils d'analyse (CAM, NSDL Paradata, UTL, LCDM) qui tendent à structurer les données dans le format xAPI [20]. Si ces analyses peuvent être confrontées à des problèmes de reproductibilité [21] pour des raisons de contraintes techniques ou de dépendances contextuelles, elles offrent un potentiel heuristique

notamment par les spatialités [22] qu'elles révèlent. Nous mobilisons ici ce potentiel pour étudier le fonctionnement d'un EIAH du point de vue de son interface. Un des objectifs de notre étude est de voir dans quelle mesure la flexibilité d'une interface facilite l'utilisation d'un logiciel de traitement de texte et l'apprentissage des utilisateurs.

2.3 Problématique

Le but de notre recherche est donc de vérifier, à travers l'analyse des traces, si l'activité de personnalisation relevant de l'appropriation de l'interface favorise l'intention d'adoption d'un EIAH. Notre hypothèse est que la personnalisation de l'interface graphique permet de réduire les attentes en termes de performances (*performance expectancy – PE*) et d'efforts (*effort expectancy – EE*) mais aussi d'atténuer les influences sociales (*social influence – SI*) et d'améliorer les conditions facilitatrices (*facilitation condition – FC*) de l'environnement pour utiliser le logiciel de traitement de texte.

3 Méthodologie

3.1 Présentation du logiciel DOCYRUS

Nous avons développé un logiciel de traitement de texte, DOCYRUS, que nous avons doté d'un module de collecte de traces et de fonctions avancées de personnalisations (*cf.* Fig. 1) en nous inspirant du projet AMBRE-Add [23] qui vise l'automatisation de la personnalisation d'un l'EIAH par les enseignants à partir des profils d'élèves. Dans le logiciel DOCYRUS, la personnalisation est manuelle et effectuée par l'apprenant. Le choix du traitement de texte s'explique par la place centrale qu'occupe ce type de logiciels dans les suites bureautiques et dont les finalités et les interfaces sur lesquelles ils reposent sont pluriels et pluri-formes (ex. écriture d'invention à des finalités de création, de révision voire de mémorisation).

Nous avons ajouté à DOCYRUS, un module de collecte de traces, présent dans d'autres environnements, tels que Input-log [24]. Nous avons ainsi renforcé les capacités de personnalisation de l'interface du logiciel par l'apprenant en lui permettant, et ce, à tout moment du déroulement de l'activité, de changer de thème graphique, de libellés des boutons ou des icônes voire de réinitialiser tout ou partie de l'infrastructure technique personnalisée (*cf.* Fig. 2).

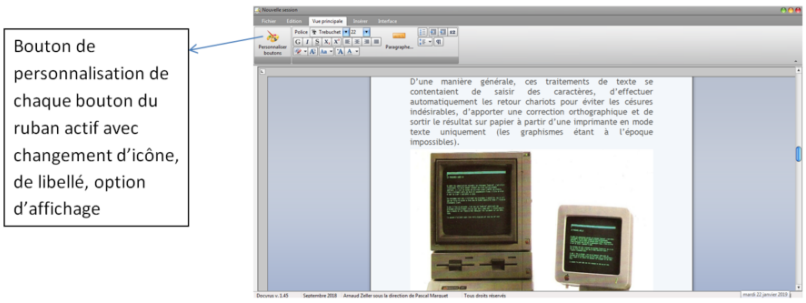


Fig. 1. Interface de Docyrus (thème graphique MacSkin)

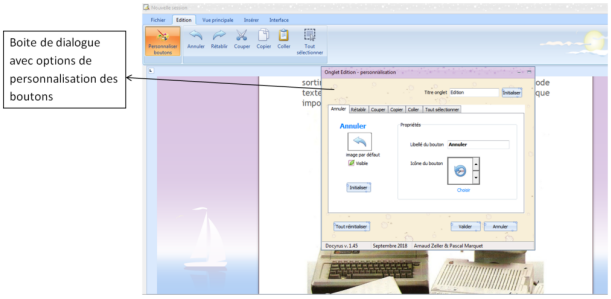


Fig. 2. Boite de dialogue de personnalisation (thème graphique Summer 2008)

3.2 Contexte de la recherche

A l'occasion de la Fête de la science 2018, nous avons organisé, un atelier pédagogique de découverte et d'appropriation du logiciel DOCYRUS dans un centre de culture scientifique et technique destiné à des enfants âgés de 3 à 15 ans. Nous avons choisi de ne pas fournir de ressources documentaires pouvant aider les participants à utiliser le logiciel mis à leur disposition. Plus de 120 personnes, principalement des enfants scolarisés à l'école maternelle, élémentaire et au collège (88 %), ont participé à l'expérience. Placés en situation d'apprentissage, les participants devaient répondre, seuls, par essai-erreur, à la consigne suivante : « *vous allez créer et personnaliser votre traitement de texte. Des connaissances en informatique ne sont pas nécessaires* ».

L'activité proposée s'est déroulée en trois temps et la durée des ateliers était de 30 minutes. Les participants ont d'abord saisi des informations personnelles les concernant (âge, sexe, niveau scolaire, raison de leur présence : contrainte ou libre). Ils ont, ensuite, réalisé l'activité d'apprentissage, proprement dite, organisée en deux phases. Premièrement, chaque utilisateur devait créer l'interface du traitement de texte afin de s'approprier les fonctions du logiciel DOCYRUS. Cette phase, au cours de laquelle aucune activité de saisie, d'édition ou de mise en forme du texte n'était possible,

correspond au temps 1 d'implémentation de l'interface et des blocs fonctionnels du traitement de texte. Deuxièmement, une fois l'interface créée, les apprenants pouvaient utiliser les différentes fonctionnalités du logiciel en vue d'éditer et de remettre en forme un document illustré. Lors de ce temps 2 d'utilisation, il était tout à fait possible d'enclencher, à nouveau, des actions de personnalisation de l'interface. A l'issue de l'activité, les utilisateurs ont été invités à compléter un questionnaire en ligne avant de visualiser les traces de bas et de haut-niveau liées aux opérations qu'ils ont effectuées en utilisant le logiciel de traitement de texte.

Nous avons choisi d'utiliser le questionnaire de Venkatesh *et al*. [11], constitué d'une échelle de Likert à 7 points et reprenant 17 items (sur les 27 qu'il comporte) en lien avec les variables de notre étude (*cf*. Tableau 1)[1]. Le but de cette mesure était de connaître le degré d'acceptation du logiciel de traitement de texte DOCYRUS. Il s'agissait notamment de vérifier si les capacités de personnalisation du logiciel étaient de nature à influencer l'intention d'usage. Nous n'avons pas organisé de pré-test car nous n'avons pas souhaité familiariser les participants à la tâche de personnalisation de l'interface pour ne pas influencer l'acceptabilité de l'EIAH par une expérience antérieure. Nous avons ainsi pu récupérer et exploiter 50 questionnaires complets (N= 50) et comptabiliser 56461 instants que nous avons agrégés en traces enregistrées séquentiellement par horodatage.

3.3 Variables et mesures

Tableau 1. Présentation des variables liées au modèle UTAUT [11]

Variable	Code	Indices
Performance Expectancy – PE (Attentes en termes de performance)	V1_PE	U6 - Cette interface me serait utile RA1 - Je gagne du temps en utilisant cette interface RA5 - Je suis plus efficace en utilisant cette interface
Effort Expectancy – EE (Attentes en termes d'efforts)	V2_EE	EOU3 - Ce serait facile pour moi d'utiliser correctement cette interface EOU5 - Je trouverais cette interface facile à utiliser EU4 - Cela serait facile pour moi d'apprendre à utiliser cette interface
Social Influence – SI (Influences sociales)	V3_SI	SN1 - Les personnes qui influencent mon comportement pensent que je devrais utiliser cette interface SN2 - Les personnes qui sont importantes pour moi pensent que je devrais utiliser cette interface SF4 - En règle générale, l'école (proposant cette interface) soutiendrait l'utilisation de cette interface

[1] Nous avons supprimé les items qui faisaient référence à une comparaison étant donné l'absence de pré-test. Nous avons vérifié la cohérence interne des items restants par un alpha de Cronbach ($\alpha > 0.7$). Nous avons également observé que d'autres chercheurs [25], [26], [27] ont modifié ce questionnaire en supprimant et/ou ajoutant de nouvelles variables au modèle UTAUT, parfois empruntés à d'autres modèles d'acceptation et d'utilisation des technologies.

Facilitation Condition – FC (Conditions facilitatrices)	V4_FC	PBC2 - J'ai le matériel nécessaire pour utiliser cette interface PBC3 - J'ai les connaissances nécessaires pour utiliser cette interface PBC5 - Cette interface n'est pas compatible avec les autres logiciels (proposés par l'école) que j'utilise FC3 - Une personne serait disponible pour m'aider en cas de problème dans l'utilisation de cette interface
Behavioral Intention to use the system – BI (Intention d'utilisation du système)	V5_BI	BI1a - J'ai l'intention d'utiliser cette interface dans les 3 prochains mois BI1b - J'ai l'intention d'utiliser cette interface dans les 6 prochains mois BI2a - Je prédis que je vais utiliser cette interface dans les 3 prochains mois BI2b - Je prédis que je vais utiliser cette interface dans les 6 prochains mois

Afin de tester notre hypothèse principale, nous avons croisé les variables liées au processus d'appropriation de l'interface qui renvoient aux activités de personnalisation de l'interface (VPI), de navigation (VNA), de saisie de texte (VST) et d'édition et mise en forme du texte (VFT) avec les variables déterminantes du modèle UTAUT (*cf.* Tableau 1), à savoir les attentes en termes de performance (V1_PE), les attentes en termes d'effort (V2_EE), les influences sociales (V3_IS), les conditions facilitatrices (V4_CF), considérées comme ayant une influence sur l'intention d'utilisation (V5_BI). Il convient d'ajouter que l'activité de personnalisation (VPI) de l'interface graphique repose sur deux types d'action. La première renvoie à la personnalisation des boutons (VPI_PB) (ex. renommage des libellés des boutons, changement des icônes des boutons, affichage ou non de ces boutons à l'écran) tandis que la deuxième tient au choix du thème graphique du logiciel (VPI_TG) (ex. couleur globale des fenêtres, icônes « ouverture » et « fermeture » de la fenêtre, canevas).

Notre méthodologie repose sur l'analyse des traces d'usage des apprenants. Nous avons donc exploité les traces « informatives », issues de l'activité d'apprentissage réalisée, qui provenaient des évènements contrôleurs déclenchés par les apprenants tels que les survols, les sélections et clics depuis la souris et le clavier, des objets de l'interface (ruban, zones, boutons, boîtes de dialogue) auquel s'ajoutait l'activité de saisie de texte.

4 Résultats

4.1 La personnalisation de l'interface au cœur du processus d'appropriation

Les résultats obtenus montrent que l'activité de saisie du texte (VST) possède la moyenne la plus élevée (M = 184,08) d'évènements contrôleurs (nombre de caractères dactylographiés émis sur le traitement de texte). Pour autant, elle ne constitue pas l'activité principale des apprenants (*cf.* Tableau 2). En effet, le nombre de mots pouvant être formés reste relativement faible (les mots tiennent sur deux lignes) et la valeur de l'écart-type est très élevée (σ = 262,705). En réalité, les trois autres activités liées au

processus d'appropriation de l'interface (personnalisation, navigation, édition et mise en forme du texte) nous semblent mieux décrire l'activité principale des apprenants. Il ressort ainsi que la personnalisation de l'interface (VPI ; M = 27,14) présente la moyenne la plus élevée d'entre elles, suivie de la navigation (VNA ; M = 26,98) puis de l'édition et de la mise en forme du texte (VFT ; M=17,76). Le nombre maximal d'évènements contrôleurs (n = 96) reste proche de l'édition et la mise en forme de texte (n = 99). Il convient de noter le caractère volontaire de la participation des étudiants qui s'élève à 90 %.

Tableau 2. Statistiques descriptives relatives au processus d'appropriation

Variables	Code	N	Min.	Max.	Moy.	Écart-type
Personnalisation de l'interface	VPI	50	2	96	27,14	21,248
Saisie du texte	VST	50	0	1197	184,08	262,705
Navigation	VNA	50	12	60	26,98	13,84976
Édition et mise en forme du texte	VFT	50	1	99	17,76	18,35874
Caractère volontaire (Variable modératrice)		20	0	1	0,90	0,303

Nous observons également, en ce qui concerne l'activité de personnalisation, une évolution différente des actions de personnalisation des boutons (VPI_PB) et des choix du thème graphique (VPI_TG) entre le temps 1 d'implémentation et le temps 2 d'utilisation (*cf*. Tableau 3). Les résultats mettent en évidence, entre T1 et T2, une augmentation de VPI_PB (M = 18,80 > 8,34) et une diminution de VPI_TG (M = 4,24 < 10,74). Ils indiquent, par ailleurs, que la moyenne, en T2, de VPI_TG est très inférieure à celle de VPI_PB (4,24 < 18,80). De tels constats révèlent que les apprenants n'associent pas de propriétés fonctionnelles à la personnalisation du thème graphique du logiciel, si ce n'est sa capacité à modifier l'aspect esthétique du logiciel. En revanche, ils attribuent des propriétés appropriatives à la personnalisation des boutons. L'amplification des opérations liées à VPI_PB, en T2, s'explique par un effet d'aller et retour avec les activités de navigation, d'édition et de mise en forme de texte.

Tableau 3. Statistiques descriptives relatives à la personnalisation de l'interface

Variables liées à l'activité de personnalisation de l'interface	Code	N	Min.	Max.	Moy.	Écart-type
Personnalisation des boutons	VPI_PB – T1 Implémentation	50	1	38	8,34	6,877
Personnalisation des boutons	VPI_PB – T2 Utilisation	50	1	**74**	**18,80**	17,358
Choix du thème graphique	VPI_TG – T1 Implémentation	50	1	50	10,74	10,368
Choix du thème graphique	VPI_TG – T2 Utilisation	50	1	31	**4,24**	5,263

4.2 Effets de la personnalisation de l'interface sur l'intention d'utilisation

L'analyse des données concernant l'intention d'utiliser le logiciel (V5_BI) montre que les apprenants émettent une intention d'utilisation (V5_BI_IU3 et V5_BI_IU6) supérieure à 50 % (*cf*. Tableau 4), et ce, quelle que soit la temporalité considérée (M = 3,91 à 3 mois ; M = 3,88 à 6 mois). La prédiction d'utilisation (V5_BI_PU3 et V5_BI_PU6) apparaît également plus forte à moyen terme qu'à court terme (M = 4,15 ; M = 3,38). Un tel résultat laisse penser que les utilisateurs ont anticipé et/ou compris le délai nécessaire à la mise en production permettant au logiciel d'être accessible en téléchargement après des derniers tests de vérification d'anomalies.

Tableau 4. Statistiques descriptives relatives à l'intention d'utilisation du système (*Behavioral Intention to use the system – V5_BI)*

Variable	Code	N	Min.	Max.	Moyenne	Écart-type
Intention d'utilisation à 3 mois	V5_BI_IU3	50	1	7	3,91	2,183
Intention d'utilisation à 6 mois	V5_BI_IU6	50	1	7	3,88	2,237
Prédiction d'utilisation à 3 mois	V5_BI_PU3	50	1	7	3,38	2,122
Prédiction d'utilisation à 6 mois	V5_BI_PU6	50	1	7	4,15	2,340

Afin d'examiner les éventuels liens entre la personnalisation de l'interface et l'intention d'adoption du logiciel de traitement de texte, nous avons calculé une corrélation de Spearman entre les variables déterminant l'intention et le comportement d'usage et les variables liées au processus d'appropriation de l'interface graphique (*cf*. Tableau 5).

Tableau 5. Corrélations non paramétriques

Rho de Spearman		VPI	VFT	VNA	VST
V1_PE	Coefficient de corrélation	0,142	0,232	0,064	0,233
	Sig. (bilatéral)	0,326	0,105	0,656	0,104
V2_EE	Coefficient de corrélation	**0,492****	0,185	0,263	**0,323***
	Sig. (bilatéral)	0,000	0,197	0,065	0,022
V3_SI	Coefficient de corrélation	0,237	0,173	0,231	0,245
	Sig. (bilatéral)	0,097	0,229	0,107	0,086
V4_FC	Coefficient de corrélation	**0,335***	0,173	**0,338***	**0,372****
	Sig. (bilatéral)	0,017	0,230	0,016	0,008

* La corrélation est significative à 0.05 (bilatéral).
** La corrélation est significative à 0.01 (bilatéral).

Les résultats révèlent un lien entre les actions de personnalisation de l'interface (VPI) déclenchées par les apprenants et leurs attentes en termes d'effort (V2_EE). Le rho donne un résultat significatif ($\varrho = 0{,}492$; $P < .01$). Il en va de même pour le croisement

de cette variable avec les conditions facilitatrices de l'environnement support (V4_FC) avec un rho de Spearman qui donne un résultat significatif (ϱ = 0,335 ; P < .05). Toutefois, il n'est pas possible d'établir de corrélations significatives entre les actions de personnalisation de l'interface (VPI) et les variables relatives aux attentes en termes de performance (V1_PE ; ϱ = 0,142 ; ns) et aux influences sociales (V3_SI ; ϱ = 0,237 ; ns).

Nous observons qu'il existe également une corrélation significative (ϱ = 0,323 ; P< .05) entre les opérations d'édition et de mise en forme du texte (VFT) et les attentes en termes d'effort (V2_EE). Cette corrélation est légèrement supérieure avec la variable V4_FC avec un rho qui donne aussi un résultat significatif (ϱ = 0,372 ; P < .01). La variable relative aux actions de navigation (VNA) peut elle aussi être corrélée avec V4_FC (ϱ = 0,338 ; P < .05). Cependant, l'analyse de corrélation ne révèle pas de relations significatives entre la saisie de texte (VST) et les 4 variables explicatives du modèle UTAUT.

5 Conclusion

La communauté EIAH s'est beaucoup intéressée, jusqu'à présent, à l'évaluation des apprentissages consécutifs à la personnalisation automatique d'interfaces ou d'environnement dits adaptatifs. Ces travaux ont abouti à la capacité des environnements, à partir de l'analyse en mode synchrone des actions de l'utilisateur, à s'adapter aux apprenants, dans le respect de la scénarisation pédagogique et en veillant à ne pas les détourner de la tâche prescrite devant être réalisée sur l'EIAH. Dans cette recherche, nous avons considéré, au contraire, que l'évaluation et la personnalisation interviennent dans un processus dynamique multidimensionnel facilitant l'intention d'adoption de l'environnement par les conditions facilitatrices mise en œuvre (UTAUT) conduisant à une meilleure utilisabilité.

Les résultats confirment, en partie, l'hypothèse de départ et permettent de conclure que la personnalisation d'une interface graphique intervient dans le processus d'adoption, par les apprenants, d'un EIAH de type logiciel de traitement de texte. Il ressort principalement de notre étude que la personnalisation de l'interface est corrélée avec les attentes en termes d'effort et les conditions considérées comme facilitatrices de l'environnement support. Moyennant quoi, les opérations d'édition et de mise en forme du texte sont elles aussi corrélées aux mêmes attentes et aux mêmes conditions facilitatrices. A titre d'illustration, nous avons mesuré le nombre de fois où les apprenants ont, par exemple, modifié le libellé d'un bouton permettant de déclencher une opération de traitement sur un texte (ex. mettre en gras, en italique, …), changé l'icône d'un bouton, caché ou au contraire rendu visible des boutons du ruban selon qu'ils les considéraient utiles ou non, *etc*.

Dans une perspective instrumentale, l'apprenant commence son activité par l'appropriation de l'interface, ici, l'artefact technique. Il poursuit son apprentissage par la mise en relation de cet artefact avec l'artefact pédagogique, c'est-à-dire, la navigation appréhendée comme une auto-scénarisation, puis enfin l'artefact didactique, à savoir la saisie, l'édition et la mise en forme de texte à travers la réalisation des activités habituellement réalisées sur un logiciel de traitement de texte. Mais dans ce processus

itératif, c'est par l'utilisation et donc l'acquisition de connaissances portant à la fois sur le fonctionnement de l'EIAH et la réalisation de l'activité elle-même que l'apprenant réalise l'appropriation individuelle de ces trois familles d'artefacts dans une zone dite d'adaptation [28]. Nous pensons que ces processus s'inscrivent dans la genèse instrumentale des artefacts complexes présents dans l'EIAH, artefacts qui entrent en conflit en générant des difficultés d'instrumentalisation et d'instrumentation, qui doivent se réaliser simultanément [2].

Ces résultats pourront, par la suite, être complétés par l'analyse des temporalités de l'activité dans DOCYRUS dans lesquelles s'inscrivent ces processus. L'étude du rythme et de la fréquence d'apparition des évènements permettront de caractériser les interactions entre les activités de personnalisation et les activités de navigation lors de la réalisation d'une activité sur un EIAH. Les variations observées pourraient avoir un impact sur le déroulement de l'activité d'apprentissage.

Références

1. Dillenbourg, P., Tchounikine, P.: Flexibility in macro-scripts for computer-supported collaborative learning. Journal of Computer Assisted Learning. 23(1) (2007) 1–13
2. Marquet, P.: e-Learning et conflit instrumental. Entre didactique, pédagogie et technique. Recherche et formation. 3(68) (2011) 31–46
3. Balacheff, N.: Didactique et intelligence artificielle. Recherche en didactique des mathématiques. 14 (1994) 9–42
4. Tchounikine P.: Pour une ingénierie des Environnements Informatiques pour l'Apprentissage Humain. Revue I3 Information-Interaction-Intelligence. 2(1) (2002) 59–95
5. Desmarais, M.C., Baker, R.J.d.: A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments. User Modeling and User Adaptive Interaction. 22(1-2) (2012) 9–38
6. Aleven, V., McLaren, B. M., Sewall, J., Van Velsen, M., Popescu, O., Demi, S., Koedinger, K. R.: Example-Tracing tutors: Intelligent tutor development for non-programmers. International Journal of Artificial Intelligence in Education. 26(1) (2016) 224–269
7. Rabardel, P.: Les hommes et les technologies : approche cognitive des instruments contemporains. Armand Colin, Paris (1995)
8. Tchounikine, P.: Précis de recherche en ingénierie des EIAH. (2009). Repéré à http://lig-membres.imag.fr/tchounikine/Articles/PrecisV1.pdf
9. Barcenilla, J., Bastien, J.-M.-C.: L'acceptabilité des nouvelles technologies : quelles relations avec l'ergonomie, l'utilisabilité et l'expérience utilisateur ? Le travail humain. 72(4) (2009) 311–331
10. Dillon, A., Morris, M. G.: User acceptance of new information technology: theories and models. Annual review of information science and technology. 31 (1996) 3–32
11. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly. 27(3) (2003) 425–478
12. Venkatesh, V., Zhang, X.: Unified theory of acceptance and use of technology: US vs. China. Journal of global information technology management. 13(1) (2010) 5–27

13. Venkatesh, V., Thong, J. Y., Xu, X.: Unified theory of acceptance and use of technology: A synthesis and the road ahead. Journal of the Association for Information Systems. 17(5) (2016) 328–376
14. Bush, V.: As We May Think. In: Greif, I. (ed.), Computer-Supported Cooperative Work: A Book of Readings. Morgan Kaufmann Publishers, San Mateo, CA (1945) 17–34
15. Sutherland, I.E.: Sketchpad: a man-machine graphical communication system (Thesis). Massachusetts Institute of Technology (1963)
16. Delozanne, E.: Interfaces en EIAH. In: Grandbastien, M., Labat, J.-M. (eds.): Environnements informatiques pour l'apprentissage humain. Hermès-Lavoisier, Paris (2006) 223–244
17. Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., Vanderdonckt, J. A.: Unifying Reference Framework for Multi-Target User Interfaces. Interacting with Computers. 15(3) (2003) 289–308
18. Thevenin, D., Haue, J. B, Calvary, G.: Cadre de conception pour la plasticité des interfaces : démarche ingénierie, démarche centrée utilisateur. Revue d'Interaction Homme-Machine. 4(1) (2003) 119–148
19. Dewan, P., Choudhary, R.: A high-level and flexible framework for implementing multiuser user interfaces. ACM Transactions on Information Systems. 10(4) (1992) 345–380
20. Venant, R.: Évaluation et visualisation de la performance de l'apprenant lors d'une activité pratique. In : Proceedings of the 6th RJC EIAH (2016)
21. Lebis, A., Lefevre, M., Luengo, V., Guin, N.: Capitalisation of analysis processes: Enabling reproducibility, openness and adaptability thanks to narration. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge (2018)
22. Noucher, M., Mericskay, B., Roche, S.: Usages des traces numériques en géographie : Potentiels heuristiques et enjeux de recherche. L'information géographique. 2(2) (2018) 39–61
23. Guin, N., Lefevre, M., Jean-Daubias, S.: Personnalisation de l'apprentissage dans l'EIAH Ambre-add. In Proceedings of EIAH 2011 (2011)
24. Leijten, M., Van Horenbeeck, E., Van Waes, L.: Analyzing writing process data: A linguistic perspective. In: Cislaru, G. (ed.), Writing(s) at the crossroads: the process-product interface. John Benjamins Publishing Company, Amsterdam/Philadelphia (2015) 277–302
25. Ibanescu, G.: Facteurs d'acceptation et d'utilisation des technologies d'information : une étude empirique sur l'usage du logiciel « Rational Suite » par les employés d'une grande compagnie de services informatiques (Dissertation). UQAM (2011)
26. Viet, D. T.: Antécédents individuels de l'utilisation du gouvernement électronique au niveau municipal (Thesis). Université de Montréal (2012)
27. Kouakou, K. S.: Adoption des réseaux sociaux numériques par les bibliothécaires des universités ivoiriennes : une approche par l'UTAUT. Les Cahiers du numérique. 11(2) (2015) 167–202
28. Denami, M., Luft, B.: IAM, PNA: from Instructional Design to a Usability Test for Learning. Understanding, Designing and Analysing the Technology Appropriation Process. Journal of e-Learning and Knowledge Society. 13(1) (2017) 65–76

L'acceptabilité des tablettes tactiles en classe : étude de l'évolution des perceptions d'élèves du secondaire

Julie Mulet[1], Cécile Van de Leemput[2], Sandrine Rome[1], Jan Van der Linden[2], Franck Amadieu[1]

[1] Université Toulouse 2 Jean-Jaurès, Laboratoire CLLE-LTC (UMR 52-63), 5 Allée Antonio Machado, 31100 Toulouse, France
julie.mulet@univ-tlse2.fr

[2] Université libre de Bruxelles
Centre de Recherche en Psychologie du Travail et de la Consommation

Résumé. De nombreuses études montrent que les étudiants expriment des perceptions positives envers l'usage de tablettes tactiles en classe. Cependant, seule une minorité de ces études s'appuie sur un cadre théorique précis. De plus, la plupart n'évaluent pas l'évolution de ces perceptions dans une perspective longitudinale. Cet article présente deux études conduites pendant plusieurs mois auprès d'élèves d'écoles secondaires issus de deux établissements différents, avec pour cadre théorique les modèles d'acceptabilité. Les résultats montrent une évolution des perceptions, mais le sens de ces évolutions diffère. Dans l'étude 1, les niveaux de la plupart des perceptions diminuent tandis qu'ils augmentent principalement dans l'étude 2. Des études précédentes montrent que les perceptions d'acceptabilité varient en fonction des expériences d'utilisation : les résultats obtenus ici seront expliqués à la lumière des différences de conditions d'implantation et d'utilisation des tablettes au sein des deux établissements.

Mots-clés. Acceptabilité ; perceptions ; tablettes tactiles ; élèves du secondaire ; sentiment d'auto-efficacité

Abstract. Many studies show that students express positive perception about the use of tablet for learning, but only a minority of them relies on clearly identified theoretical frameworks. Moreover, these studies often do not take account the temporal component of these perceptions. This paper presents two studies conducted in secondary education from two separate schools, based on the technology acceptance models. Results show a shift in students' perceptions, but in opposite directions. In study 1, most of perceptions decrease whereas in study 2, most of them increase. Previous studies show that students perceptions toward a new technology depend on experiences of use. Thus, these results will be explained through differences between conditions of use found in the two schools.

Keywords. Acceptance ; perceptions ; tablets PC ; secondary students ; self-efficacy

1 Introduction

Cet article s'intéresse à l'étude longitudinale des perceptions des apprenants vis-à-vis de l'usage de tablettes tactiles en classe. Cette question est abordée au regard des modèles de l'acceptabilité des nouvelles technologies qui ont été utilisés comme cadre théorique pour la conduite des deux études présentées.

1.1 Usage des tablettes tactiles en classe : quels résultats ?

Parmi les outils numériques utilisés en contexte scolaire, les tablettes tactiles ont une place privilégiée et sont fortement encouragées par les politiques éducatives. Leurs caractéristiques physiques (taille et poids, portabilité, absence de périphériques de contrôle tels que les claviers ou les souris…) en font des outils facilement utilisables en classe. Par ailleurs, leurs caractéristiques techniques (autonomie, connectivité, interactivité, possibilités de personnalisation, fonctionnalités multiples) permettent de répondre à une grande variété d'objectifs pédagogiques. Une récente revue de la littérature de Haßler, Major et Hennessy [1] conclut que dans la majorité des études, l'usage de tablettes dans l'enseignement primaire et secondaire soutient le processus d'apprentissage. Enfin, les tablettes, et plus largement les outils permettant l'apprentissage mobile (« *m-learning* »), rendraient l'apprentissage plus motivant pour les élèves [2]. D'après de nombreuses études, les élèves trouvent les tablettes faciles à utiliser [3]. Ils estiment aussi que leur usage en classe rend l'apprentissage plus facile, plus intéressant et plus agréable [4, 5]. Néanmoins, une récente revue de la littérature démontre que ces perceptions positives ne sont pas constatées dans toutes les études [6]. Dans certains cas, les apprenants expriment même des critiques à l'égard de l'outil : les tablettes induiraient une fatigue oculaire [4], seraient sujettes à des bugs et dysfonctionnements qui perturberaient l'apprentissage [7] et ne seraient pas adaptées à certaines tâches académiques telles que les tâches d'écriture [8]. Ainsi, les perceptions positives supposées des apprenants à l'égard de l'usage des tablettes tactiles en classe sont à relativiser. De plus, des critiques méthodologiques et théoriques peuvent être formulées. D'un point de vue méthodologique, une majorité des études s'intéresse aux perceptions des élèves à un instant donné, consécutivement à une période plus ou moins longue d'usage de l'outil. Pourtant, ces perceptions sont susceptibles d'évoluer en fonction de multiples facteurs. Au plan théorique, peu d'études évaluent les perceptions des élèves dans un cadre théorique bien défini. Les modèles de l'acceptabilité des nouvelles technologies offrent un cadre utile pour comprendre l'intégration par les utilisateurs d'une nouvelle technologie.

1.2 Les modèles d'acceptabilité pour comprendre les perceptions des élèves

Les modèles d'acceptabilité visent à définir les facteurs qui contribuent à la détermination des intentions d'usage d'un système ou d'une technologie. L'acceptabilité correspond à la perception qu'a un individu de la valeur de cette technologie. D'après les modèles, une bonne acceptabilité d'un outil est une condition

préalable à son utilisation. Un des modèles les plus utilisés dans le domaine est le TAM (*Technology Acceptance Model*) [9]. D'après le TAM, les intentions d'utiliser un outil sont déterminées par les attitudes générales, définies comme les réactions affectives à l'égard de la technologie évaluée [10]. Les attitudes sont elles-mêmes déterminées par deux perceptions instrumentales de l'outil : son utilité perçue et sa facilité d'usage perçue. L'utilité perçue correspond au degré auquel une personne croit que l'utilisation de l'outil améliorera ses performances. La facilité d'usage perçue correspond au degré auquel une personne croit que l'utilisation de l'outil lui demandera peu ou pas d'effort [11]. Plus récemment, des antécédents aux perceptions instrumentales ont été ajoutés dans les modèles. Le sentiment d'auto-efficacité (SAE) lié à l'usage de l'outil (*i.e.,* la perception qu'une personne a de ses capacités à utiliser la technologie) a notamment été intégré en tant que déterminant de la facilité d'usage perçue dans la troisième version du TAM [12].

Les modèles d'acceptabilité ont parfois été utilisés pour évaluer l'adoption des tablettes tactiles par les élèves [13], les étudiants [14] ou les enseignants [15]. Néanmoins et comme évoqué plus haut, la plupart des études ne s'intéressent pas à l'évolution de ces perceptions, qui sont pourtant susceptibles de fluctuer. Par exemple, dans deux études menées auprès d'étudiants, Bhattacherjee et Premkumar [16] ont constaté une diminution de l'attitude générale et de l'utilité perçue après 6 mois d'usage d'une technologie informatisée. Ce résultat peut être rapproché de « l'effet de nouveauté » constaté dans certaines études lors de l'introduction d'une nouvelle technologie en classe [17] : l'intérêt suscité par la nouveauté s'estompe et fur et à mesure des usages académiques [18], en particulier lorsque les étudiants ont été confrontés à des expériences négatives (bugs, difficultés d'utilisation…) [19, 20]. A l'inverse, les usages répétés d'un nouvel outil peuvent avoir une influence positive sur d'autres perceptions. L'expérience d'utilisation des outils numériques pour l'apprentissage est en effet associée à une augmentation du SAE lié à l'utilisation du numérique [21] et à une diminution de l'anxiété [22] : au fur et à mesure des usages, les apprenants développeraient des compétences qui les rendraient plus confiants en leur capacités à utiliser l'outil.

1.3 Objectif des études et hypothèses

Deux études sont présentées dans cet article. Les recueils de données mis en place visent à évaluer de façon longitudinale les perceptions des élèves du secondaire vis-à-vis des tablettes tactiles à trois moments du processus d'adoption de l'outil : avant l'utilisation des tablettes en classe, après quelques semaines d'utilisation puis après plusieurs mois d'utilisation (à l'issue de l'année scolaire). La première étude s'est déroulée dans une classe de lycée général et la seconde étude a eu lieu auprès d'élèves de collège. Les deux études ont été conduites au cours de l'année scolaire 2017-2018.

Les perceptions évaluées sont principalement issues des modèles de l'acceptabilité. Sur la base du TAM [11], nous avons mesuré les perceptions générales et instrumentales suivantes : les attitudes générales, l'utilité perçue et la facilité d'usage perçue. Sur la base du TAM 3 nous avons aussi mesuré le SAE des élèves relativement à deux aspects d'utilisation des tablettes tactiles : l'usage général (SAE général) et l'usage spécifique pour des tâches académiques (SAE usage académique). Enfin, nous

avons ajouté une nouvelle variable : l'utilité motivationnelle, définie comme la perception que l'usage de l'outil améliorera l'attractivité des tâches scolaires [23].

En référence aux études mettant en avant un « effet de nouveauté » lors de l'intégration d'une nouvelle technologie [16, 18], on s'attend à ce que les attitudes générales des élèves (H1), l'utilité perçue de la tablette (H2) et l'utilité motivationnelle (H3) diminuent au fur et à mesure des utilisations. En référence aux travaux montrant une amélioration des compétences perçues et fur et à mesure des expériences avec un outil numérique [21], nous nous attendons à ce que les deux dimensions du SAE (usage général (H4) et usage académique (H5)) augmentent en cours d'année. Enfin, le SAE étant un déterminant direct de la facilité perçue d'après le TAM 3, nous nous attendons à ce que la facilité perçue augmente en cours d'année (H6).

2 Etude 1 : utilisation collective de tablettes en lycée

L'étude s'est déroulée dans un lycée général français. Au début de l'année scolaire 2017-2018, l'établissement scolaire a été doté de 12 tablettes Android. Du fait de ce faible nombre, il a été décidé que les tablettes ne seraient employées que pour des tâches collaboratives. L'équipe enseignante a introduit cet outil au sein d'une classe pilote de 2nde. Deux enseignants ont choisi d'utiliser occasionnellement les tablettes en classe : l'enseignant de sciences et vie de la terre et celui de sciences physiques.

2.1 Méthode

Participants

Sur les 31 élèves de la classe, 21 ont participé à l'ensemble des trois recueils de données. Parmi ces 21 élèves, 62% avaient déjà utilisé la tablette à titre individuel pour des tâches de divertissement ou de communication.

Mesures

Les perceptions des élèves ont été mesurées par des questionnaires. Chaque perception était mesurée par différents items. Les attitudes générales ont été mesurées par des items construits par Courtois et al. [24]. La facilité d'utilisation, l'utilité et l'utilité motivationnelle perçues et le SAE relatif à l'usage général ont été mesurées par des items construits par van der Linden et al. (2017). Les items du SAE relatif à l'usage académique ont été construits pour l'expérience. Pour chaque item, les élèves devaient évaluer leur degré d'accord sur une échelle de Likert en 7 points allant de 1 ("absolument pas d'accord") à 7 ("tout à fait d'accord"). Le tableau 1 présente les différentes dimensions mesurées ainsi qu'un exemple d'item pour chacune.

Tableau 1. Exemple d'items pour chaque dimension

Variable étudiée	Exemple d'item	Nombre d'items
Attitudes	« C'est intéressant que l'on travaille avec les tablettes en classe »	3
Facilité d'utilisation perçue	« La tablette sera facile à utiliser en classe »	4
Utilité perçue	« Avec la tablette en classe, je travaillerai plus rapidement »	4
Utilité motivationnelle	« Utiliser la tablette en classe va rendre les exercices plus amusants »	4
SAE – Usage général	« Je peux utiliser une tablette sans problèmes sans avoir besoin d'aide. »	4
SAE – Usage académique	« Je pense pouvoir réaliser efficacement les activités en cours avec la tablette. »	3

Protocole expérimental

Les questionnaires de perception ont été remplis en ligne pendant les heures de cours en respectant l'anonymat des élèves. Chaque recueil a duré environ 30 minutes. Les trois recueils ont eu lieu dans la salle informatique du lycée : le premier s'est déroulé en novembre 2017, avant usage des tablettes, le second en mars 2018, le dernier en juin 2018. Les tablettes ont été utilisées à partir de décembre 2017.

2.2 Résultats

Après avoir contrôlé la cohérence interne des items relatifs à chaque dimension évaluée (calcul des coefficients alpha de Crombach avec cohérence acceptée lorsque le coefficient est supérieur à .70, en référence à Nunnally [25], nous avons conduit des t-test à mesure répétées afin d'évaluer l'évolution entre chaque moment de mesure (mesure 1 et 2 ; mesure 2 et 3). Les données descriptives relatives à l'étude 1 sont présentées dans le tableau 2.

Evolution des perceptions générales et instrumentales

Les attitudes générales des élèves avant l'utilisation des tablettes sont très positives (5.8 sur une échelle à 7 niveaux). Toutefois, après les premières utilisations des tablettes, ces perceptions positives diminuent significativement, passant à une moyenne de 4.57 en novembre ($t(20)=6.20$, $p<.001$). Cette chute se poursuit entre les deuxième et troisième vagues de mesures puisque le score moyen passe à 4.01 en mai ($t(20)=2.43$, $p=.03$).

Les perceptions relatives aux qualités instrumentales de la tablette en classe sont elles aussi positives, en particulier en début d'année. La facilité d'utilisation reste stable

dans le temps. Toutefois, l'utilité perçue en classe diminue entre novembre et mars ($t(20)=5.57$, $p<.001$) ce qui indique qu'au fur et à mesure des premières utilisations, les élèves ont trouvé la tablette de moins en moins utile pour apprendre. L'utilité perçue est restée stable entre mars et juin.

Evolution des perceptions motivationnelles

L'utilité motivationnelle perçue diminue significativement après les premières utilisations de la tablette ($t(20)=4.05$, $p=.001$), mais reste statistiquement stable entre les mois de mars et de juin. Le SAE lié à l'usage académique suit un pattern similaire, : il diminue entre les deux premières vagues de mesure ($t(20)=2.03$, $p=.05$) et ne varie pas significativement entre les deux dernières vagues de mesure. En revanche, le SAE relatif à l'utilisation générale de la tablette reste statistiquement stable dans le temps.

Tableau 2. Données descriptives de l'étude 1

Variable étudiée	Score moyen (et écart type) en novembre	Score moyen en mars	Score moyen en juin
Attitudes	5.8 (0.95)	**4.57** (1.29)**	**4.01* (1.31)**
Facilité perçue	5.42 (1.17)	5.16 (1.09)	4.81 (0.92)
Utilité perçue	5.42 (1.17)	**4.16** (1.25)**	3.74 (1.40)
Utilité motivationnelle	5.46 (0.98)	**4.41** (1.31)**	3.96 (1.34)
SAE – Usage général	5.68 (0.81)	5.50 (1.22)	4.90 (1.46)
SAE – Usage académique	5.76 (1.13)	**5.30* (1.15)**	4.91 (1.60)

Synthèse des résultats de l'étude 1 et éléments de discussion

Comme prédit par nos hypothèses, les attitudes générales, l'utilité perçue et l'utilité motivationnelle ont diminué entre les deux premières mesures. Cette diminution est cohérente avec l'hypothèse d'un effet de nouveauté [16].

En revanche, le SAE relatif à l'usage académique a diminué entre les deux premières vagues de mesure. Ce résultat n'était pas attendu, mais il peut s'expliquer les difficultés rencontrées (problèmes de fonctionnement et de mises à jour des outils numériques, coupures durables du réseau internet…) qui ont les enseignants à ne plus utiliser les tablettes à partir du printemps, ce qui explique aussi que la majorité des perceptions n'a significativement pas diminué entre les deux dernières mesures. Ainsi, nos résultats montrent que la difficile expérience utilisateur dans cet établissement scolaire a négativement impacté les perceptions des élèves. L'impact de l'expérience utilisateur sur les perceptions d'acceptabilité a été montré dans des études antérieures [20, 26]. En revanche, lorsque les tablettes ne sont plus utilisées (à partir du mois de mars), les perceptions ne diminuent plus significativement, à l'exception des attitudes générales. Les attitudes correspondent à des perceptions générales qui, d'après les modèles de l'acceptabilité, peuvent être impactées par d'autres dimensions qui n'ont pas été mesurées ici, telles que la norme subjective (définie comme la croyance d'un l'individu vis-à-vis des perceptions qu'ont les autres personnes par rapport à l'outil évalué [27]) ou la présence d'une assistance technique [28]. Les difficultés rencontrées, l'absence

d'assistance technique et l'abandon des tablettes en cours d'année ont certainement impacté les perceptions des enseignants (norme subjective). Ces éléments ont donc pu contribuer à la chute des attitudes générales, et ce même lorsque les tablettes n'étaient plus utilisées.

Deux dimensions ne semblent pas avoir été affectées par ces problèmes techniques puisqu'elles sont restées statistiquement stables à travers le temps : il s'agit de la facilité perçue et du SAE relatif à l'utilisation générale de la tablette. Autrement dit, seules les perceptions liées à l'usage de la tablette en contexte académique ont été négativement impactées par l'expérience utilisateur. Ce résultat laisse à penser que les expériences utilisateur n'ont un impact que sur les dimensions liées au contexte spécifique d'utilisation et non sur les dimensions plus générales.

3 Etude 2 : utilisation individuelle de tablettes en collège

L'étude 2 s'est déroulée dans un collège général français. L'établissement scolaire a fait le choix de proposer une tablette personnelle (iPad) à chaque élève. Ces derniers avaient la possibilité d'amener la tablette à leur domicile hors du temps scolaire (hormis pendant la période estivale).

A la rentrée 2017, année de démarrage de l'étude, l'établissement était engagé depuis 2 ans dans la distribution progressive des tablettes. Ainsi, la plupart des enseignants utilisaient déjà l'outil en classe. Néanmoins, les élèves interrogés dans le cadre de l'expérience n'y avaient pas encore eu accès.

3.1 Méthode

Participants

L'étude s'est centrée sur les élèves de 6ème et de 5ème : au moment du premier recueil de données, ces derniers n'avaient pas encore été équipés de tablettes tactiles, contrairement aux élèves des classes supérieures. Au total, environ 170 élèves ont participé à chaque recueil de données. Cependant, nous n'avons pu à terme conserver les données que des 51 élèves qui ont participé à l'ensemble des trois recueils. Plus de 80% d'entre eux avaient déjà utilisé une tablette tactile à titre personnel pour des activités de divertissement ou de communication.

Instrument de mesure et protocole expérimental

Les questionnaires utilisés étaient similaires à ceux de l'étude 1. Ils ont été remplis en ligne pendant les heures de cours en respectant l'anonymat des élèves. Les recueils se sont déroulés dans la salle informatique du collège. Trois recueils ont eu lieu : le premier s'est déroulé en décembre 2017, avant usage des tablettes ; le second en avril 2018 ; le dernier en juin 2018. Les tablettes ont été utilisées à partir de janvier 2018.

3.2 Résultats

La procédure d'analyse des données était similaire à celle employée pour l'étude précédente. Les données descriptives obtenues à l'issue de l'étude 2 sont indiquées dans le tableau 3.

Evolution des perceptions générales et instrumentales

Les attitudes générales, bien qu'élevées tout au long de l'année, baissent légèrement en avril ($t(50)=3.32$, $p=.002$) puis augmentent en fin d'année ($t(50)=-2.47$, $p=.02$) pour retrouver un niveau similaire à celui constaté avant l'adoption des tablettes. La facilité d'utilisation perçue reste stable entre les mois de novembre et d'avril puis augmente entre avril et juin ($t(50)=-2.52$, $p=.02$). La perception d'utilité de la tablette suit un pattern similaire **:** elle reste stable entre les deux premiers moments du recueil puis augmente au mois de juin ($t(50)=-2.93$, $p=.005$).

Evolution des perceptions motivationnelles

L'utilité motivationnelle diminue légèrement après les premières expériences d'utilisation, mais cette diminution est marginalement significative ($t(50)=1.86$, $p=.07$) puis augmente significativement en juin (les scores obtenus en décembre et en juin ne diffèrent pas significativement**).** Le SAE relatif à l'usage académique augmente lui aussi significativement entre la mesure 2 et la mesure 3. En revanche, le SAE relatif à l'usage général reste stable tout au long de l'année.

Tableau 3. Données descriptives de l'étude 3

Variable étudiée	Score moyen (et ET) en décembre	Score moyen (et ET) en avril	Score moyen (et ET) en juin
Attitudes	6.10 (0.88)	**5.64* (1.02)**	**6.04* (1.14)**
Facilité perçue	5.72 (1.03)	5.96 (0.95)	**6.25* (0.93)**
Utilité perçue	4.95 (1.40)	4.80 (1.29)	**5.23* (1.45)**
Utilité motivationnelle	5.27 (1.33)	4.97 (1.27)	**5.47* (1.46)**
SAE – Usage général	5.59 (1.05)	5.83 (1.05)	5.96 (1.07)
SAE – Usage académique	5.92 (1.20)	5.88 (1.17)	**6.26* (1.09)**

Synthèse des résultats de l'étude 2 et éléments de discussion

La plupart des perceptions vis à vis de l'usage des tablettes tactiles pour l'apprentissage sont restées stables entre le premier et le deuxième moment de mesure, puis ont augmenté significativement entre la seconde et la dernière mesure. Ainsi, contrairement à nos hypothèses, les attitudes générales (H1), l'utilité perçue (H2) et l'utilité motivationnelle (H3) n'ont pas diminué. On peut supposer que les expériences utilisateur positives vécues par les élèves ont permis de limiter « l'effet de nouveauté » [16] qui avait justifié nos hypothèses 1, 2 et 3.

Comme nous l'attendions, le SAE relatif à l'usage académique et la facilité d'utilisation perçue ont augmenté ; en revanche, le SAE relatif à l'usage général est resté stable tout au long de l'année. Comme évoqué dans l'étude précédente, il semble que seules les perceptions spécifiques au contexte d'utilisation étudié sont affectées par les expériences utilisateur.

4 Synthèse des résultats et discussion générale

Dans l'ensemble, les élèves interrogés ont des perceptions d'acceptabilité de l'usage des tablettes tactiles en classe très positives, et ce tout au long de l'année. On constate aussi des variations en cours d'année des différentes perceptions étudiées. Néanmoins, ces évolutions ne vont pas dans le même sens. Dans l'ensemble, les élèves du lycée (étude 1) ont des perceptions de moins en moins positives, en particulier lorsque l'on compare les mesures recueillies avant utilisation des tablettes (mesure 1) à celles recueillies en mars (mesure 2). A l'inverse, les élèves du collège (étude 2) ont des perceptions qui demeurent stables ou diminuent légèrement après la phase d'adoption mais qui évoluent positivement en fin d'année (mesure 3). Pourquoi de telles divergences dans l'évolution des perceptions des élèves ? Des éléments contextuels relatifs à l'implantation des tablettes dans les établissements scolaires peuvent apporter des éléments d'explication. En effet, les contextes d'implantation et les modalités d'utilisation ont beaucoup différé entre les deux établissements.

Comme présenté dans les synthèses de résultats, les expériences d'utilisation des tablettes en classe ont été dans l'ensemble positives pour les élèves de collège (étude 2) et plutôt négatives pour les élèves de lycée (étude 1). Les expériences négatives liées à la rencontre de problèmes techniques amènent à une chute des perceptions positives dans diverses études antérieures [20, 26]. Nos résultats semblent aller dans le même sens : aussi, il semble indispensable de prendre en compte l'impact des expériences utilisateurs spécifiques dans l'étude de l'évolution des perceptions vis-à-vis des nouvelles technologies pour l'apprentissage. A ce titre, les modèles récents de l'acceptabilité proposent d'intégrer le vécu subjectif de l'utilisateur lors de l'usage d'un nouvel outil : c'est notamment le cas du *Component of User Experience model* (CUE-model) [29].

Par ailleurs, les enseignants de lycée étaient confrontés à l'utilisation des tablettes en classe pour la première fois, tandis que les enseignants du collège expérimentaient cet usage pour la troisième année consécutive. Or, le niveau d'expérience des enseignants avec l'usage des tablettes est un facteur des perceptions des élèves dans certaines études [19, 30] : il impacte la capacité à résoudre les problèmes techniques rapidement, soit parce que les enseignants sont aguerris à leur résolution soit parce que l'établissement dispose d'un support technique rapide et efficace. La rapidité de la résolution des problèmes techniques est un facteur important de l'acceptabilité d'un nouvel outil [28] et de la réussite d'un projet d'apprentissage mobile [31]. Enfin, les élèves du collège (étude 1) se voyaient attribuer une tablette personnelle pour toute la durée de l'année scolaire tandis que les élèves de lycée (étude 2) ne disposaient pas de tablettes individuelles. Or, d'après Naismith et Corlett [31], l'accès régulier à la technologie (notamment hors contexte scolaire) et le fait de pouvoir s'approprier

personnellement l'outil sont aussi des facteurs qui impactent l'intégration des outils numérique mobiles pour l'apprentissage.

Il est à noter que dans les deux études, le SAE relatif à l'utilisation générale de la tablette reste stable. Ce résultat suggère que les conditions d'accès à une nouvelle technologie et les expériences d'utilisation impactent les perceptions spécifiques au contexte d'utilisation, mais pas les perceptions plus générales. Rappelons que dans les deux cas, la majorité des élèves avaient déjà utilisé une tablette tactile à titre personnel. Il est possible que le SAE général se soit aussi construit au travers des expériences vécues hors contexte académique, neutralisant ainsi l'effet de celles vécues en classe.

5 Conclusion et préconisations

Si beaucoup d'études qualitatives ou quantitatives mettent en avant les perceptions positives qu'entretiennent les élèves vis-à-vis de l'usage des tablettes tactiles en classe, peu se sont intéressées à l'évolution de ces perceptions en cours d'année scolaire. Aussi, l'objectif des deux études présentées ici était d'évaluer de façon longitudinale les perceptions des élèves du secondaire vis-à-vis des tablettes tactiles. Si nos résultats confirment la forte adhésion des élèves, ils montrent aussi que ces perceptions évoluent différemment en fonction des conditions d'accès à la technologie d'une part et des expériences d'utilisation d'autre part. Ces résultats vont dans le sens de modèles récents de l'acceptabilité qui proposent de prendre en compte l'impact de l'expérience utilisateur pour analyser le processus d'acceptabilité d'un nouvel outil. Néanmoins, nos résultats suggèrent aussi que seules les perceptions spécifiques au contexte d'usage sont impactées par les expériences vécues.

Au plan pratique, ces études soulèvent l'importance d'accompagner l'implantation des outils numériques dans les classe, de proposer une assistance efficiente et de former les enseignants à la gestion technique de ces outils. Ces éléments apparaissent être des facteurs clés pour l'adoption des outils numériques en classe, sans quoi ces outils risquent fort de rester inutilisés.

References

1. Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. Journal of Computer Assisted Learning, 32(2), 139–156. https://doi.org/10.1111/jcal.12123
2. Jones, A., Issroff, K., Scanlon, E., Clough, G., McAndrew, P., & Blake, C. (2006). Using mobile devices for learning in informal settings: is it motivating? In: IADIS International Conference on Mobile Learning. IADIS Press, (January), 251–255.
3. Ciampa, K. (2014). Learning in a mobile age: An investigation of student motivation. Journal of Computer Assisted Learning, 30(1), 82–96. https://doi.org/10.1111/jcal.12036
4. Dündar, H., & Akçayir, M. (2014). Implementing tablet PCs in schools: Students' attitudes and opinions. Computers in Human Behavior, 32, 40–46. https://doi.org/10.1016/j.chb.2013.11.020
5. Soykan, E. (2015). Views of students',teachers' and parents' on the tablet computer usage in education. Cypriot Journal of Educational Sciences, 10(3), 228. https://doi.org/10.18844/cjes.v1i1.68.

6. Mulet, J., van de Leemput, C.,& Amadieu, F. (2019). A critical literature review of perceptions of tablets for learning in primary and secondary schools. Educational Psychology Review, 1-32. https://doi.org/10.1007/s10648-019-09478-0
7. Kontkanen, S., Dillon, P., Valtonen, T., Eronen, L., Koskela, H., & Väisänen, P. (2017). Students' experiences of learning with iPads in upper secondary school—a base for proto-TPACK. Education and Information Technologies, 22(4), 1299–1326. https://doi.org/10.1007/s10639-016-9496-7.
8. Ling, G. (2016). Does It Matter Whether One Takes a Test on an iPad or a Desktop Computer? International Journal of Testing, 16(4), 352–377. https://doi.org/10.1080/15305058.2016.1160097
9. Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science, 35(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982
10. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. Journal of Chemical Information and Modeling, 53(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004
11. Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319. https://doi.org/10.2307/249008
12. Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x
13. van Deursen, A. J. A. M., ben Allouch, S., & Ruijter, L. P. (2014). Tablet use in primary education: Adoption hurdles and attitude determinants. Education and Information Technologies, 21(5), 971–990. https://doi.org/10.1007/s10639-014-9363-3
14. Moran, M., Hawkes, M., & Gayar, O. (2010). Tablet Personal Computer Integration in Higher Education: Applying the Unified Theory of Acceptance and use Technology Model to Understand Supporting Factors. Journal of Educational Computing Research, 42(1), 79–101. https://doi.org/10.2190/EC.42.1.d
15. Ifenthaler, D., & Schweinbenz, V. (2013). The acceptance of Tablet-PCs in classroom instruction: The teachers' perspectives. Computers in Human Behavior, 29(3), 525–534. https://doi.org/10.1016/j.chb.2012.11.004
16. Bhattacherjee, A., & Premkumar, G. (2004). Understanding Changes in Belief and Attitude toward Information Technology Usage: A Theoretical Model and Longitudinal Test. MIS Quarterly, 28(2), 229–254. https://doi.org/ 10.2307/25148634
17. Couse, L. J., & Chen, D. W. (2010). A Tablet Computer for Young Children? Exploring its Viability for Early Childhood Education. Journal of Research on Technology in Education, 43(1), 75–96. https://doi.org/10.1080/15391523.2010.10782562
18. Song, S., & Keller, J. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. Educational Technology Research and Development. Retrieved from http://link.springer.com/article/10.1007/BF02504925
19. Montrieux, H., Raes, A., & Schellens, T. (2017). 'The best app is the teacher' Introducing classroom scripts in technology-enhanced education. Journal of Computer Assisted Learning, 33(3), 267–281. https://doi.org/10.1111/jcal.12177
20. Pamuk, S., Ergun, M., Çakir, R., Yilmaz, H. B., & Ayas, C. (2013). The use of tablet pc and interactive board from the perspectives of teachers and students: Evaluation of the fatih project. Kuram ve Uygulamada Egitim Bilimleri, 13(3), 1815–1822. https://doi.org/10.12738/estp.2013.3.1734
21. Jan, S. K. (2015). The relationships between academic self-efficacy, computer self-efficacy, prior experience, and satisfaction with online learning. American Journal of Distance Education, 29(1), 30–40. https://doi.org/10.1080/08923647.2015.994366
22. Pruet, P., Ang, C. S., & Farzin, D. (2015). Understanding tablet computer usage among primary school students in underdeveloped areas: Students' technology experience, learning

styles and attitudes. Computers in Human Behavior, 55, 1131–1144. https://doi.org/10.1016/j.chb.2014.09.063
23. Van der Linden, J., Amadieu, F., & Van de Leemput, C. (2017). L'importance de l'expérience vécue pour l'usage des tablettes à l'université. In C. Hellemans, A. Casini, & A. Van Daele (Eds.), Bien-Être et diversité des situations de travail - Mobilités, conciliations et violences au travail (L'Harmatta, pp. 63–77). Paris.
24. Courtois, C., Montrieux, H., De Grove, F., Raes, A., De Marez, L., & Schellens, T. (2014). Student acceptance of tablet devices in secondary education: A three-wave longitudinal cross-lagged case study. Computers in Human Behavior, 35, 278–286. https://doi.org/10.1016/j.chb.2014.03.017
25. Nunnally, J. (1978). Psychometric theory. 2nd Edition, McGraw-Hill, New York.
26. Montrieux, H., Courtois, C., De Grove, F., Raes, A., Schellens, T., & De Marez, L. (2014). Mobile Learning in Secondary Education: Teachers' and Students' Perceptions And Acceptance Of Tablet Computers. International Journal of Mobile and Blended Learning, 6(2), 26–40. https://doi.org/10.4018/ijmbl.2014040103
27. Fishbein, M., & Ajzen, I. (1975). Belief , attitude , intention and behaviour : An introduction to theory and research. Reading. MA: Addison-Wesley.
28. Ifenthaler, D., & Schweinbenz, V. (2016). Students' Acceptance of Tablet PCs in the Classroom. Journal of Research on Technology in Education, 48(4), 306–321. https://doi.org/10.1080/15391523.2016.1215172
29. Thuring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. International Journal of Psychology, 42(4), 253–264. https://doi.org/10.1080/00207590701396674
30. Tay, H. Y. (2016). Longitudinal study on impact of iPad use on teaching and learning. Cogent Education, 3(1). https://doi.org/10.1080/2331186X.2015.1127308
31. Naismith, L., & Corlett, D. (2006). Reflections on success: A retrospective of the mLearn conference series 2002-2005. Proceedings of MLearn 2006 – the 5th World Conference on M-Learning, 1–29.

Session
Métadonnées

Un schéma OWL pour la description d'éléments de formation

Daniel Antelme, Nicolas Delestre, Nicolas Malandain

Laboratoire d'Informatique, de Traitement de l'Information et des Systèmes
INSA Rouen Normandie prenom.nom@insa-rouen.fr
Normandie Université

Résumé. De nos jours les établissements d'enseignement publient leurs offres de formation uniquement sur le Web historique, celui des documents à destination des humains. Pour que des programmes informatiques puissent tirer parti de ces informations, il faut qu'elles soient aussi publiées sur le Web des données. Malheureusement il n'existe pas d'ontologie permettant de bien décrire la pluralité de ces offres et les évolutions des modalités d'enseignement. Cet article propose un schéma OWL pour la description des éléments de formation qui a été testé pour décrire plusieurs types d'offres de formation : initiale, continue, en présentiel, à distance, proposée par un ou plusieurs établissements.

Mots-clé. Éléments de formation, Web des données, modèle conceptuel, schéma OWL2

Abstract. Nowadays, educational institutions publish their programme offers only on the historical Web, therefore for humans. Softwares would be able to take advantages of those informations but only if they were published on the Web of Data. Unfortunately, there is no ontology to properly describe the plurarity of these offers and the recent changes in teaching methods. This article proposes an OWL schema of educational item descriptions that has been tested for the description of several types of training offers : initial, continuing, face-to-face, distance learning, offered by one or more institutions.

Keywords: Educational items, linked data, conceptual model, OWL2 schema

1 Introduction

Lorsqu'un organisme ou établissement de formation veut publier son offre de formation, il utilise aujourd'hui principalement le Web. Cette publication peut prendre plusieurs formes : du texte rédigé, des listes hiérarchiques, des tableaux ou encore des figures. De plus, très souvent, le vocabulaire utilisé est propre à l'institution.

Les programmes informatiques, comme les moteurs de recherche, ne savent pas inférer de sens pertinent ou élaboré à partir de ces informations, mais juste

cibler des mots-clé. Il apparaît donc judicieux de transcrire – en plus – ces informations de façon structurée à destination des machines, dans le Web des données tel que le préconisait Tim Berners Lee dès 1996. Encore faut-il que ces données soient issues d'un modèle assez riche pour représenter toutes les diversités de l'enseignement. Notre proposition tente de répondre à ce problème en permettant de décrire tout type d'élément de formation (quels que soient le cadre de formation et sa modalité) à l'aide des outils du Web des données.

Nous allons tout d'abord présenter quelques exemples d'éléments de formation qui nous semblent caractéristiques. Nous ferons alors un état de l'art des descriptions des offres de formation, ce qui nous amènera à constater qu'aucun modèle n'est totalement satisfaisant. Nous proposerons ainsi un modèle conceptuel de données, puis une implantation à l'aide d'OWL2. Enfin, avant la conclusion, nous donnerons quelques liens vers des exemples d'offres de formation décrites avec notre modèle et des exemples de requêtes SPARQL.

2 Des exemples d'éléments de formation à décrire

Afin de bien vérifier si l'existant permet de bien décrire les propriétés d'un élément de formation (EF) (que l'on qualifierait de « cours » dans le vocabulaire quotidien mais cela serait trop vague), nous allons présenter trois exemples d'EF réels.

Le premier, « Introduction à la compilation », du département Architecture des Systèmes d'Information (ASI) de l'INSA Rouen Normandie, a vu au cours du temps une évolution de sa modalité de transmission et de ses objectifs pédagogiques : initialement en présentiel, il a ensuite été enseigné à distance et est maintenant enseigné en *blended learning*. Enfin en 2017, le contenu du cours a en partie changé puisque le langage d'application n'est plus le C mais le Java, et lex et yacc ont été remplacés par Javacc.

Le deuxième, « Document et Web Sémantique », toujours enseigné en ASI, revêt deux formes : l'une plutôt « classique » où les étudiants sont évalués à l'aide d'un examen sur table et de la réalisation d'un projet, et l'autre dite « de découverte », dont l'évaluation ne porte que sur le suivi des CM, l'assiduité à des TP et la réalisation d'un projet. Bien entendu, la charge de travail n'étant pas la même, le nombre de crédits ECTS est différent.

Le troisième, le MOOC « Soyez acteurs du Web ! » est dispensé sur la plateforme FUN (`https://www.fun-mooc.fr/`). La première session d'enseignement a eu lieu en 2014, et jusqu'en 2017, il y a eu une session par an. Durant ces 4 années, bien que l'objectif pédagogique soit le même il y a eu une évolution importante quant aux connaissances et compétences enseignées.

Pour décrire ces EF nous avons donc besoin d'exprimer le fait :

EF-1 qu'un EF puissent être décrit en terme de prérequis et de compétences acquises à l'issue de son enseignement ;

EF-2 qu'un EF puisse être dispensé à plusieurs reprises que nous nommerons « session » ;

EF-3 qu'une session est composée d'un ensemble d'activités pédagogiques ;

ACTI-1 qu'une activité pédagogique, tout comme un EF, a des prérequis et des acquis ;

ACTI-2 qu'une activité pédagogique se décrit aussi en modalité d'enseignement (présentiel, à distance ou mixte) ;

ACTI-3 qu'une activité pédagogique utilise des ressources pédagogiques ;

ACTI-4 qu'une activité pédagogique peut donner lieu à un ou plusieurs actes pédagogiques.

3 État de l'art

Voici les modèles, schémas et outils existant actuellement qui permettent de décrire un EF voire une offre de formation :

CDM *Course Description Metadata* a été créé début des années 2000 par la Norvège pour décrire les établissements d'enseignements supérieurs , leurs enseignants et leurs parcours de formations. Une adaptation française a été développée en 2004 (CDM-fr) et une version pour les Web des données (SemCDMfr) a été proposée par [2] en 2014 ;

Lhéo *Langage Harmonisé d'Échange d'informations sur l'Offre de formation* est un langage de description des offres de formation de l'enseignement professionnel. Tout comme le CDM, Lhéo n'a pas vraiment de modèle conceptuel, c'est un langage XML d'échange d'information ;

MLO *Metadata for learning opportunities* est un modèle conceptuel pour faciliter l'intéropérabilité de schémas de description d'offres de formations. Il ne préconise pas d'implantation informatique. Segedinac & al. [5] ont proposé une version OWL de ce modèle mais qui ne semble pas être publique ;

AIISO est un schema RDF (RDFS) qui permet de décrire principalement les organismes de formation académiques et de manière plus succincte les formations proposées [6] ;

TEACH est un schéma OWL1 qui permet de décrire principalement les cours et leurs réalisations [4] ;

Schema.org permet d'insérer au sein des pages HTML des métadonnées, donc des données structurées que pourront interpréter les moteurs de recherche pour mieux comprendre les pages HTML [3].

Le tableau 1 synthétise cet état de l'art au regard des critères listés précédemment. Nous avons constaté qu'aucun de ces schéma ou modèle ne permet – seul – de représenter toutes les caractéristiques listées précédemment. Cependant, certains des aspects qu'ils expriment nous semblent pertinents, voire indispensables. C'est pourquoi nous les avons alignés afin d'identifier les classes et propriétés communes et celles qui sont propres à un schéma ou un modèle.

	CDM	Lhéo	MLO	AIISO	TEACH	Schema.org
EF-1	0 *formalPrerequisites* *recommandedPrerequisites*	0 *conditions-specifiques*	0 *prerequisite*	--	--	0 *coursePrerequisites*
EF-2	--	+ *session*	+ *specifies*	--	--	+ *hasCourseInstance*
EF-3	--	--	+ *hasPart*	--	--	--
ACTI-1	--	--	0 *prerequisite*	--	--	--
ACTI-2	+ *formOfTeaching*	+ *modalites-pedagogiques*	--	--	--	+
ACTI-3	--	--	--	--	+ *hasCourseMaterial*	--
ACTI-4	--	0 *sous-module*	--	--	--	0 *subEvent*

Tableau 1. Capacité de représentation des caractéristiques listées dans la section 2

4 Un modèle conceptuel décrivant des éléments de formation

Pour définir ce modèle conceptuel (MC) nous allons suivre une démarche *bottom-up*. Au niveau le plus bas, il y a la réalisation d'un acte pédagogique (classe `ActePédagogique`). Cet acte pédagogique réunit dans un endroit donné (réel ou virtuel) des apprenants et des enseignants ou tuteurs.

Plusieurs actes pédagogiques avec des apprenants différents et éventuellement des enseignants différents peuvent représenter la même activité pédagogique (classe `ActivitePédagogique`). Une activité pédagogique permet à l'aide de ressources pédagogiques (classe `RessourcePédagogique`) d'acquérir des compétences (classe `Compétence`). Une activité pédagogique est aussi caractérisée par des prérequis et par une approche, une méthode pédagogique, appelée aussi type (classe `TypePédagogique`), qui dépend de la nature des acquis. Par exemple si la compétence à acquérir est épistémique, le type de l'activité pourrait être un cours magistral, le visionnage d'un film ou la lecture d'un document. Il est à noter qu'une évaluation est aussi une activité pédagogique.

La réalisation d'un EF (classe `RéalisationDÉlementDeFormation`) permet de représenter une session de formation, le fait d'enseigner l'EF sur une période donnée. Elle organise les activités pédagogiques. Cette organisation peut prendre plusieurs formes mais elle commence obligatoirement par une activité et se termine par une activité. Entre ces deux activités, les activités peuvent se suivre (organisation séquentielle) ou être en parallèle (formant ainsi un treillis). Ce sont les prérequis et acquis qui organisent les activités entre elles.

Enfin un EF (classe `ÉlémentDeFormation`) nécessite de la part des apprenants des prérequis et a pour objectif d'enseigner des acquis sur un sujet donné. Au cours du temps un élément de formation peut évoluer en ces termes.

La figure 1 présente l'ensemble de ces classes reliées les unes aux autres par des relations structurelles. Bien entendu, bien que non présentées ici, ces classes possèdent aussi des attributs qui permettent de les caractériser.

5 Un schéma pour le Web des données

Le W3C propose deux méta-modèles pour représenter notre modèle conceptuel : les RDFS [1] et l'OWL2 [7]. Nous avons retenu OWL2 car outre le fait qu'il

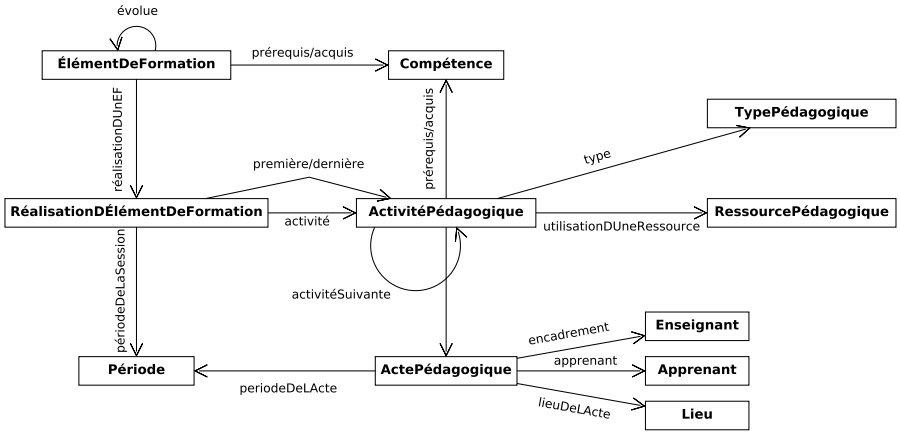


Fig. 1. Modèle conceptuel (diagramme de classes UML)

permette de vérifier la cohérence du modèle et des données, il a un pouvoir d'expression supérieur et il permet d'aligner des modèles.

Les classes de notre MC vont naturellement être représentées à l'aide des classes OWL2. Les relations entre les classes du MC sont représentées à l'aide des `owl:objectProperty`. Enfin, les attributs de classes du MC, non présentés dans la figure, sont représentés à l'aide d'`owl:objectProperty` quand leurs valeurs ne sont pas des types simples et d'`owl:dataProperty` lorsqu'elles le sont. En plus du *domain* et du *range*, OWL2 nous permet d'ajouter des propriétés aux `owl:objectProperty`. Par exemple la propriété `prérequis` est transitive, ou encore des propriétés (non représentées sur la figure 1 pour ne donner qu'une vue d'ensemble) peuvent être symétriques, d'autres fonctionnelles, etc. Enfin des cardinalités minimales et maximales peuvent être précisées. Toutes ces informations pourront être exploitées par le raisonneur OWL2.

Le Web des données permet d'interconnecter les individus entre eux mais aussi d'aligner les classes et propriétés des schémas pré-existants. Il est important que les individus instances des classes de notre schéma OWL2 soient aussi instances des classes de ces schémas lorsque les concepts sous-jacents sont les mêmes. Il en est de même pour les relations. C'est pourquoi nous avons aligné notre schéma avec les schémas Teach, AIISO et Schema.org.

6 Validation

Valider un MC et son implantation est toujours délicat. Nous ne pouvons pas appliquer des méthodes numériques fondées sur les statistiques ou sur des algorithmes (comme par exemple la validation croisée). Seules l'expérience et leur utilisation dans le temps valideront ou invalideront ces schémas. C'est pour cela que nous avons mis en ligne sur `http://iflot.insa-rouen.fr` notre schéma

OWL2, trois exemples bien différents de formation et des exemples de requêtes utilisant l'entrée SPARQL `http://linkeddata.insa-rouen.fr`.

7 Conclusion

Dans cet article nous avons proposé un nouveau modèle conceptuel permettant de décrire des éléments de formation. Ce modèle a été implanté en OWL2 et des exemples de descriptions de formation ont été proposés.

Toutefois, il reste à mieux faire connaître ce modèle et son schéma OWL. D'une part, en le rendant pérenne à l'aide d'URI déréférançable et d'autre part en le référençant dans des annuaires de schémas du Web des données tel que Link Open Vocabularies (`https://lov.linkeddata.es`).

De plus, cette contribution devrait être intégrée dans un travail de plus grande envergure qu'est NoDE-fr. Le NoDE-fr est un projet de normalisation AFNOR décomposé en trois parties, la première pour décrire les ressources pédagogiques (partie qui sera normalisée en 2019), la seconde pour décrire les offres de formations et enfin la troisième pour décrire les compétences. Notre travail s'inscrit donc dans cette deuxième partie.

Une fois que ces trois parties seront normalisées et que les établissements de formation publieront la description de leurs offres de formation en « open data » à l'aide de mapping pour le Web des données de ces trois normes, tout un chacun pourra développer des applications « innovantes » proposant de nouveaux services.

Nous remercions la fondation UNIT pour avoir co financé ce travail.

Références

1. Brickley, D., Guha, R. : Rdf schema 1.1 (2014), `https://www.w3.org/TR/rdf-schema/`
2. Delestre, N., Malandain, N., Ouchenne, B. : Une ontologie OWL pour le CDM-fr. In : Conférence des Technologies de l'Information et de la Communication pour l'Enseignement. p. 12. Béziers, France (2014)
3. Guha, R.V., Brickley, D., Macbeth, S. : Schema.org : Evolution of structured data on the web. Commun. ACM **59**(2), 44–51 (2016)
4. Kauppinen, T., Trame, J., Westermann, A. : Teaching core vocabulary specification (2012), `http://linkedscience.org/teach/ns/\#`
5. Segedinac, M., Konjovic, Z., Surla, D., Savic, G. : An OWL representation of the MLO model. In : SISY. pp. 465–470. IEEE (2012)
6. Styles, R., Shabir, N. : Academic institution internal structure ontology (aiiso) (2008), `http://vocab.org/aiiso/schema`
7. W3C OWL Working Group : Owl 2 web ontology language document overview (second edition) (2012), `https://www.w3.org/TR/owl2-overview/`

NoDEfr-1 : Métadonnées pédagogiques et Web des données

Nicolas Delestre[1], Yolaine Bourda[2],Erwan le Gall[3], Rosa-Maria Gomez de Regil[4], Yohan Colmant[5], Marie Peterlongo[6], Romuald Verrier[3]

[1] LITIS Insa Rouen Normandie, nicolas.delestre@insa-rouen.fr
[2] LRI, CentraleSupélec, Université Paris Saclay, Gif-Sur-Yvette, France. Yolaine.Bourda@lri.fr
[3] Ministère de l'Éducation nationale et de la Jeunesse, prenom.nom@education.gouv.fr
[4] CNED, Rosa-Maria.Gomez@cned.fr
[5] Université Polytechnique Hauts-de-France, yohan.colmant@univ-valenciennes.fr
[6] Unisciel, marie.peterlongo@unisciel.fr

Résumé. Les descriptions normalisées des nombreuses ressources d'apprentissage existantes reposent actuellement majoritairement sur un des profils d'application du LOM. Ces descriptions (i.e. métadonnées) sont accessibles soit via des portails soit sous forme de données ouvertes mais dans des formats difficilement traitables par des ordinateurs. D'un autre côté, la montée en puissance des données ouvertes et liées (LOD) permet de favoriser la réutilisation et le partage des données entre applications et d'imaginer de nouveaux services. La structure du LOM (arbre) rend difficile la création d'une expression RDF/OWL (graphe) faisant consensus. Enfin, l'ISO a normalisé le MLR (Metadata for Learning Resources) qui bénéficie d'une expression en RDF/OWL. Cet article commence par présenter brièvement le MLR et son profil d'application français, le NoDEfr-1. Ensuite, nous présenterons le *binding* RDF/OWL conçu ainsi qu'un outil de transformation automatique des métadonnées existantes.

Mots-clé. Description de ressources pédagogiques · Métadonnées · Profil d'application · Web des données · Données ouvertes et liées

Abstract. The descriptions of many existing educational resources are currently based mainly on LOM application profiles. These descriptive records (i.e. metadata) are accessible either via portals or as open data sets but in formats that are difficult to process by humans and machines alike. On the other hand, the rise of linked open data (LOD) makes it possible to promote the understanding and consumption of this metadata, reuse and sharing of data between applications and to imagine new services. The LOM structure (a tree) makes it difficult to reach a consensus on an RDF/OWL binding (a graph). Hence ISO has standardized the MLR (Metadata for Learning Resources) which has an RDF/OWL binding. In this paper, we begin with a brief presentation of the MLR and its French application profile : NoDEfr-1. Then we present the RDF/OWL binding and a tool for automatic transformation of existing metadata.

Keywords: Description of educational resources · Metadata · Application profile · Web of Data · Linked Open Data

1 Introduction

De très nombreuses ressources pédagogiques sont actuellement disponibles librement sur le Web. Néanmoins ces ressources ne peuvent pas toujours être indexées de façon satisfaisante sur la seule base de leur contenu, l'indexation « plein texte » étant notamment source d'ambigüités sémantiques. Par conséquent des métadonnées leur sont généralement associées. Le schéma le plus courant pour la description de celles-ci est le LOM[7]. Cette solution a néanmoins de nombreux inconvénients : les différents profils d'application peuvent être incompatibles entre eux, la sémantique de certaines métadonnées est équivoque, le multilinguisme n'est pas natif, certaines informations sont dupliquées, il n'existe aucun consensus sur une expression RDF/OWL, etc.

En France, depuis 2003, le MESRI soutient les UNT (Universités Numériques Thématiques) qui ont produit de nombreuses ressources de qualité, pour la plupart librement accessibles. Celles-ci sont décrites par des métadonnées conformes au schéma SupLOMFR - profil d'application du LOMFR, lui-même profil d'application du LOM - dédié à l'enseignement supérieur français. Dans ce contexte, les métadonnées produites par différents établissements sont publiées grâce à un réseau de portails OAI-PMH. Bien qu'utile, ce type de service est dédié uniquement à des personnes et ne permet pas l'indexation de matériel pédagogique par les moteurs de recherche généralistes. Cela limite donc la réutilisation de ces métadonnées par des logiciels et empêche leurs liaisons avec d'autres types de données.

Par ailleurs, l'explosion du Web sémantique et des données liées [2] a permis la production d'un grand nombre de données avec une sémantique univoque, augmentant la pertinence des résultats de recherche qui les exploitent et donc la visibilité de ces données. Dans notre contexte, le respect des principes des données liées peut être bénéfique à bien des égards. En effet, si des ressources pédagogiques, issues de différents entrepôts, sont liées entre elles ou avec d'autres types de ressources : (1) des services inter-entrepôts peuvent être créés ; (2) les ressources d'apprentissage peuvent référencer des ressources proches de leurs sujets pour offrir des informations utiles (i.e. DBpedia [3]) ; (3) les ressources d'apprentissage peuvent être liées à des taxonomies spécifiques ; (4) les ontologies sous-jacentes peuvent être mises en correspondance et servir de base à des inférences.

UNIT (l'UNT dédiée aux sciences et technologies) a ainsi financé deux projets : (1) le projet SemUNIT ([4]), preuve de concept de l'insertion des descriptions des ressources des UNT dans le LOD ; (2) le projet LOFAR qui propose un moteur de recherche de ressources pédagogiques à partir de sélection de cours [5]. Ces deux projets ont mis en évidence l'intérêt d'une telle approche et ont permis de sensibiliser une partie de la communauté française, peu au fait de ces technologies.

Cet article commence par présenter brièvement la norme ISO MLR. Ensuite nous présentons notre contribution : tout d'abord le profil d'application français NoDEfr-1, puis son expression en OWL pour le Web des données et enfin la transformation automatique des métadonnées respectant un des dérivés du LOMFR vers cette expression.

2 MLR

L'objectif de la norme ISO/IEC 19788 Metadata for Learning Resources (Métadonnées pour ressources d'apprentissage) dite « MLR » est de décrire des ressources pédagogiques ainsi que des ressources qui leur sont liées dans un contexte international multilingue et multiculturel. Le MLR, reposant sur des fondations explicites et précises, permet, non seulement, de pallier les insuffisances du LOM, largement connues, mais surtout il offre de nouvelles facilités dans la description des ressources pédagogiques ainsi que dans la diffusion de ces descriptions. La visibilité des ressources en est ainsi élargie. Dans le contexte du MLR, un élément de métadonnées est défini comme un élément de données utilisé pour décrire une ressource. Par la suite, nous parlerons donc d'éléments de données (et non plus d'éléments de métadonnées).

Le MLR est une norme en plusieurs parties facilement extensible, au gré des besoins. Chaque partie du MLR a un objectif très précis :

Partie 1 – *Framework* : elle sert de fondation à l'ensemble de la norme. Elle décrit comment spécifier : (1) les éléments de données et leurs caractéristiques, (2) les classes et leurs caractéristiques, (3) les règles de contenu applicables aux valeurs que peuvent prendre les éléments de données, (4) les profils d'application, (5) les vocabulaires.

Partie 2 – Éléments du Dublin Core : elle spécifie chacun des éléments de données du Dublin Core en utilisant la partie 1 du MLR.

Partie 3 – Profil d'application de base : elle définit un profil d'application fondé sur la partie 2 du MLR. Elle exprime un exemple de spécification d'un profil d'application.

Partie 4 – Éléments techniques : elle contient les spécifications des éléments de données permettant de décrire des informations de type technique.

Partie 5 – Éléments pédagogiques : elle contient les spécifications des éléments de données permettant de décrire des informations de type pédagogique.

Partie 7 – Expressions : elle propose une expression en RDF des différentes entités constituant le MLR. Elle explicite la représentation des identifiants MLR en IRI et le *mapping* des différentes entités introduites dans le MLR en RDF. Elle fournit de plus une ontologie OWL2 des classes et propriétés introduites dans les parties du MLR connues à ce jour.

Partie 8 – Enregistrements MLR : elle contient les spécifications des éléments de données permettant de décrire des enregistrements MLR.

Partie 9 – Éléments pour les personnes : elle contient les spécifications des éléments de données permettant de décrire des personnes physiques ou morales (individus ou organisations).

Partie 11 – Migration du LOM vers le MLR : elle est un guide pour la conversion d'enregistrements au format LOM en ensemble d'éléments au format MLR.

3 NoDEfr-1

La norme « Normalisation des descriptions pour l'éducation en France – Partie 1 : description des ressources » dite « NoDEfr-1 », profil d'application du MLR est en cours de finalisation. Cette norme a vocation, à terme, à remplacer le LOMFR et ses deux profils d'application que sont le ScoLOMFR (à destination des acteurs de l'enseignement primaire et secondaire) et le SupLOMFR (à destination des acteurs du supérieur).

Les travaux autour du NoDEfr-1 ont été menés dans des cadres différents (AFNOR, InterUNT...) avec des objectifs différents. D'une part, dans le cadre de l'AFNOR, il s'agit de produire un document normatif (i.e. une norme française) qui a vocation à faire foi mais qui, comme tout document normatif, peut être difficile à lire. D'autre part, dans un groupe plus informel mêlant acteurs du MENJ, du MESRI, de l'inter-UNT, l'objectif est de produire : (1) une proposition d'expression en RDF/OWL, non normative, et mise à la disposition de la communauté française voire francophone afin que celle-ci puisse l'utiliser pour ses besoins, (2) un document d'accompagnement explicitant entre autres les différences avec le LOMFR (3) une possibilité de transformation automatique de descriptions de ressources respectant le LOMFR (ou un de ses dérivés) exprimées en XML en descriptions de ressources respectant le NoDEfr-1 exprimées sous forme de triplets RDF.

3.1 La norme

La définition de la norme NoDEfr-1 a suivi les recommandations de la partie 3 du MLR. À cela s'ajoutent trois règles que s'étaient fixés les experts de l'AFNOR. Tout d'abord il fallait que l'on retrouve dans le NoDEfr-1 tous les éléments de données du LOMFR qui étaient réellement utilisées. Certains étaient présents dans le MLR, d'autres non, ces derniers ont donc été ajoutés. Ensuite lorsque des éléments de données ont été ajoutés, le « type IRI » a été privilégié, favorisant ainsi l'insertion des descriptions des ressources dans le Web des données. Enfin certains éléments de données du MLR n'ont pas d'utilité pour la communauté francophone. Ils ont donc été écartés de la norme. La figure 1 synthétise ce travail. Les classes reprises du MLR ou créées pour le NoDEfr sont reliées par des éléments de données (DES ou SED). Ces éléments de données issues du MLR ou créés *ex nihilo* sont utilisées pour décrire les classes.

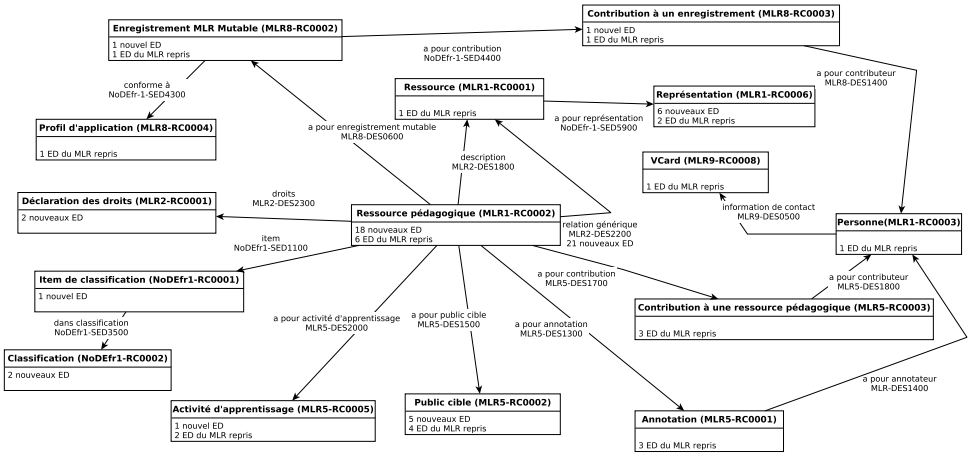


Fig. 1. Classes, nombre d'éléments de données (ED) et relations du NoDEfr-1

3.2 NoDEfr-1 et le Web des données

Représentation en OWL2 du NoDEfr-1 Le MLR, dans sa partie 7, propose une représentation du MLR en OWL2. Nous nous sommes donc fondés sur cette implantation pour définir le schéma OWL2 du NoDE-fr. Nous avons ajouté les classes et propriétés spécifiques au NoDEfr-1. Nous avons aussi redéfini certaines classes du schéma initial de façon à prendre en compte les cardinalités définies par le NoDEfr-1. Nous avons de plus privilégié l'utilisation des `owl:ObjectProperty` plutôt que les `owl:DatatypeProperty` ayant un intervalle de type `xsd:anyURI`. Enfin nous avons utilisé le schéma ≪ Collection ≫ pour représenter des ensembles ou des listes ordonnées [8]. Ce schéma est disponible à l'adresse suivante : `https://sourcesup.renater.fr/nodefr`

Transformation automatique Afin de favoriser l'acceptation du NoDEfr-1, il est indispensable que toutes les descriptions réalisées, et elles sont nombreuses, utilisant le LOMFR ou l'un de ses profils d'application (ScoLOMFR ou SupLOMFR) puissent être automatiquement transformées en descriptions conformes au NoDEfr-1 et à son expression définie ci-dessus. Les descriptions existantes étant conformes à un schéma XML, il s'agit donc ici de peupler une ontologie OWL, à partir de documents XML. Ce problème n'est pas nouveau et il a donné lieu à différentes approches le plus souvent basées sur XSLT dont certaines sont référencées par le W3C (cf. `https://www.w3.org/wiki/ConverterToRdf`).

La solution que nous avons développée a été réalisée en SWI-Prolog. Cette transformation a été testée sur plus de 50 000 fichiers contenant des descriptions reposant soit sur le ScoLOMFR soit sur le SupLOMFR.

4 Conclusion et travaux futurs

Après avoir présenté le MLR, nous avons décrit nos trois contributions principales : le profil d'application français du MLR : NoDEfr-1, sa représentation pour le Web des données et un outil de traduction des données XML du LOMFR vers des triplets RDF. Il faut maintenant que tout ceci soit adopté par la communauté. Outre un travail de communication, de formation, d'accompagnement et de maintenance , deux axes de travail sont envisagés. Le premier concerne la liaison des données. En effet il faut que les données publiées soient les plus formelles possible et surtout liées. Par exemple les personnes ou les établissements ne doivent plus être désignés et décrits à l'aide de chaînes de caractères mais bien identifiés par des IRI. Ce travail de désambigüisation est un problème difficile mais indispensable. Le second concerne les deux autres parties de la norme NoDEfr. Tout d'abord un groupe d'experts de l'AFNOR travaille actuellement sur le représentation des offres de formation. L'objectif est de proposer une norme en 2019. Ensuite, il faudra créer un profil d'application de la future norme ISO-IEC 22602 *Competency models expressed in MLR* qui permet de représenter les compétences. Une fois ces trois parties normalisées, les acteurs de la communauté éducative pourront publier des données ouvertes structurées qui permettront à quiconque de les réutiliser et de proposer des applications innovantes.

Références

1. Mohan, P., Brooks, C. : Learning objects on the semantic web. In IEEE International Conference on Advanced Learning Technologies (ICALT), 195–199 (2003)
2. Heath, T., Bizer, C. : Linked Data : Evolving the Web into a Global Data Space. 1st edition. Synthesis Lectures on the Semantic Web : Theory and Technology. Morgan & Claypool (2011)
3. Waitelonis, J., Sack, H. : Towards exploratory video search using linked data. Multimedia Tools and Applications, v59, N.2, 645–672 (2012)
4. Isaac, Y., Bourda, Y., Grandbastien, M. : SemUNIT - French UNT and Linked Data. In Proceedings of the 2nd International Workshop on Learning and Education with the Web of Data, LiLe@WWW 2012, Lyon, France (2012)
5. Delestre, N., Malandain, N., Ouchenne, B. : Une ontologie OWL pour le CDM-fr.In : Conference des Technologies de l'Information et de la Communication pour l'Enseignement. Beziers, France (2014)
6. Bourda, Y., Gauthier, G., Gomez de Regil, R-M., Catteau, O. : Métadonnées pour ressources d'apprentissage (MLR) - Nouvelle norme ISO de description de ressources pédagogiques. Rubrique de la revue STICEF, Volume 17, ISSN : 1764-7223, mis en ligne le 10/05/2011, `http://sticef.org` (2010)
7. IEEE, 1484.12.1-2002 - IEEE Standard for Learning Object Metadata, `https://ieeexplore.ieee.org/document/1032843`, (2002)
8. Ciccarese, P., Peroni, S. : Collections Ontology : Creating and handling collections in OWL 2 DL frameworks, Journal : Semantic Web, vol. 5, no. 6, 515–529, (2014)
9. Wielemaker, J., Schrijvers, T., Triska, M. and Lager, T. : SWI-Prolog. Theory and Practice of Logic Programming **12**(1-2), 67–96 (2012)

Posters

Des jeunes passeurs de pratiques : étude de l'appropriation à partir d'une analyse des données numériques

Laëtitia Pierrot

Université de Poitiers, TECHNÉ - EA 6316 86000 Poitiers, France
laetitia.pierrot@univ-poitiers.fr

Résumé. Qu'elle soit individuelle ou sociale [1], l'appropriation du numérique fait l'objet de nombreux travaux et revêt des définitions différentes selon les disciplines qui la mobilisent. Dans nos travaux, l'appropriation est abordée selon l'approche structurationniste [2], la genèse instrumentale [3], [4] et la sociologie des usages [5]. Dans ce contexte, l'article a pour but de présenter la manière dont, selon un objectif d'étude de l'appropriation, un cadre d'analyse de données numériques a été défini. Mobilisées dans le cadre d'une thèse portant sur les pratiques numériques de collégiens et de lycéens scolarisés de l'académie de Caen, les techniques d'analyses exploitées ont permis de mettre en avant un rôle de passeur de pratiques chez certains jeunes. L'objet de cet article est de présenter ce rôle qui facilite l'appropriation entre pairs.

Mots-clés. Appropriation, Activité instrumentée, Pratiques numériques juvéniles, Analyse de traces, Théorie des réseaux

Abstract. Many studies focus on the appropriation of digital technologies, whether individual or social [1]. According to the disciplinary approach, this process related to the use of a technical object is observed differently. In our work, the appropriation is studied according to the structurationist approach [2], the instrumental genesis [3], [4] and the sociology of uses [5]. In this context, this paper aims to present the way in which, with the aim of studying the appropriation, a framework of analysis of numerical data has been defined. Mobilized as part of a doctoral thesis on the digital practices of middle and high school students from the French academy of Caen, the analysis techniques used have made it possible to highlight a role as a practice facilitator amongst the students. The purpose of this paper is to present the very role that facilitates peer appropriation.

Keywords. Appropriation, Instrumented Activity, Juvenile Digital Practice, Trace Analysis, Network Theory

1 Cadre théorique et contexte de la recherche

Cet article s'appuie sur un travail de recherche doctorale qui a pour objet les pratiques numériques des jeunes. En particulier, la recherche a porté sur la circulation sociale des pratiques numériques, comme moyen d'observer l'appropriation dans une approche sociotechnique. L'appropriation recouvre différentes définitions selon

l'appartenance disciplinaire du chercheur qui l'étudie. Le cadre dans lequel nous nous inscrivons envisage l'appropriation comme un processus (1) situé puisque la formation des usages et des pratiques numériques qui traduisent l'appropriation s'inscrit dans un contexte et des circonstances précises [2], [5], (2) instrumenté, dans la mesure où l'appropriation du numérique suppose ce temps long d'instrumentation [3], [4] et (3) construit par les usagers, individuellement et socialement [1].
La difficulté dans l'analyse des usages ou des pratiques numériques, surtout pour des jeunes usagers, est qu'ils ne sont pas toujours conscients de leur activité. Et l'appropriation s'inscrivant dans un temps long, il peut résulter que l'accès au terrain est difficile ou contraint. Dans ces conditions, le travail de recherche a notamment consisté à rechercher un cadre d'analyse différent, en recourant aux données numériques. De fait, les données numériques présentent le double avantage d'être déjà disponibles, puisque générées automatiquement, et marquées temporellement [6], ce qui permet alors de suivre les pratiques numériques dans le temps, de suivre leur caractère instable car en construction [7], [8].

La problématique traitée est la suivante : comment rendre compte de l'appropriation du numérique par des jeunes, en se basant sur des données numériques ? Pour y répondre, notre travail s'appuie sur une piste méthodologique pour rendre compte de la circulation des pratiques numériques entre jeunes.

2. Méthodologie et présentation des données numériques collectées

Les données ont été collectées sur l'année scolaire 2015-2016, dans deux établissements de l'académie de Caen. Pour cela, les logs du proxy (service de cache Squid), correspondant à la navigation sur Internet des élèves dans leur établissement scolaire, ont été utilisés comme source principale. Un collège, l'établissement Jacques Monod (JM) et un lycée professionnel, le lycée Camille Claudel (CC) ont été sélectionnés pour cette étude en raison du large volume de données disponibles, parmi les traces des 17 établissements auxquels nous avons eu accès[1].

Puisque les données issues des logs du proxy comportent uniquement des traces de la consultation en ligne, nous avons élaboré une méthode qui consiste à établir une arborescence des ressources consultées dans le temps pour un groupe donné, la classe scolaire. Dans une classe où nous avions accès aux traces de plusieurs élèves, nous avons cherché à définir comment une même ressource (un site consulté) passe d'un élève à un autre dans le temps.

A partir d'une matrice des actions extraites des traces, nous distinguons l'élève qui a consulté une ressource pour la première fois dans les traces avec un rôle de « déclencheur » des autres élèves qui ont consulté la même ressource plus tard dans le temps, avec un rôle de « déclenché ». En supposant que la consultation effective de ces ressources est au départ restreinte à un élève puis partagée par plusieurs élèves, cette méthode permet de rendre visible une forme de circulation. Et, en conservant des indications sur le contexte, cela nous est utile pour identifier des situations favorables

[1] Les données numériques collectées ont été transmises par le service informatique du rectorat de l'académie de Caen, par l'intermédiaire de la Direction du numérique pour l'éducation du MENJ.

à la circulation, ou à l'inverse d'autres où les actions restent personnelles. En ce sens, cette méthode se rapproche des méthodes portant sur l'analyse de popularité des retweets [9].

Pour la visualiser, nous avons eu recours à des graphes, en adoptant les principes définis par l'analyse des réseaux sociaux. En particulier, nous prenons en compte le diamètre (la plus grande distance possible entre deux nœuds), la densité (le rapport du nombre de liens existants sur le nombre de liens possibles dans le graphe) et la modularité (score exprimé de -1 à 1 pour identifier une partition en sous-groupes) calculés sur le logiciel Gephi. La modularité a été calculée en utilisant l'algorithme de Louvain pour détecter des sous-groupes. Avec un score de modularité supérieur à 0,3, on considère que les classes comportent plusieurs groupes sociaux [10].

3. Résultats

Au collège, parmi les 14 classes tracées, nous avons identifié 2 classes où le score de modularité est supérieur à 0,3, attestant de sous-groupes parmi les élèves. Les graphes obtenus permettent de mettre en évidence que les ressources relèvent principalement des catégories thématiques d'information-documentation. Les autres pratiques, des réseaux sociaux ou de consultation audio ou vidéo sont restreintes à un petit groupe d'élèves. Les pratiques numériques apparaissent comme plus diffuses parmi les collégiens de 3ème que chez les plus jeunes. Ainsi, au collège, l'analyse des traces des élèves de 4e et de 3e marque une diversification des pratiques numériques qui semblent aller de pair avec la constitution de groupes sociaux distincts.

Au lycée, dans la classe de 1ère étudiée, les interactions sont moins concentrées sur les catégories thématiques liées à l'information-documentation. Si les pratiques d'information-documentation sont plus présentes, cette classe est celle où la plus grande variété de pratiques numériques est observée. En effet, dans les interactions se retrouvent aussi des pratiques liées à la communication, aux réseaux sociaux, aux jeux, à la consultation audio ou vidéo et aux transactions.

En étudiant la circulation des pratiques numériques, nous avons remarqué que des élèves (au collège et au lycée) concentrent les interactions au sein des groupes sociaux. D'après les observations faites sur le cas de Caen, si nous arrivons bien à trouver des groupes sociaux à partir des traces, les interactions entre les lycéens, qui témoignent d'une circulation, se font au sein des groupes et entre les groupes. Ainsi, la circulation entre groupes sociaux identifiés dépend, de fait, d'élèves qui jouent un rôle d'intermédiaire, de passeurs de pratiques numériques qui appuient alors cette circulation. Pour l'expliquer, nous avons noté que la plupart de ces élèves sont ceux qui ont une activité numérique la plus importante.

Pour prolonger cette analyse rendue possible grâce aux données numériques, il nous semble intéressant de réinvestir les concepts de capitaux culturel et social (au sens de Bourdieu [7]), à la manière d'Ellison [12], qui a montré que selon le capital social des usagers, leurs rapports aux autres sur des plateformes diffèrent, ou de Bourdeloie [13], qui considère l'appropriation comme résultant d'un capital culturel. Pour elle, plus les jeunes disposent de ressources culturelles, plus ils sont à même de développer de nouvelles pratiques. Ces premières conclusions ont pu être mises en

évidence grâce aux données numériques et surtout grâce aux instruments d'analyse utilisés. Parce que les données utilisées sont issues d'historiques de navigation, l'analyse de données similaires pour une nouvelle population permettra de confirmer ces observations et de valider la méthode employée.

Références

[1] S. Proulx, « Penser les usages des TIC aujourd'hui : enjeux, modèles, tendances », in *Enjeux et usages des TIC : aspects sociaux et culturels, t. 1*, L. Vieira et N. Pinède, Éd. Bordeaux: Presses universitaires de Bordeaux, 2005, p. 7‑20.

[2] G. DeSanctis et M. S. Poole, « Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory », *Organization Science*, vol. 5, nº 2, p. 121‑147, 1994.

[3] P. Rabardel, *Les hommes et les technologies : approche cognitive des instruments contemporains.* Paris: Armand Colin, 1995.

[4] P. Rabardel et P. Pastré, *Modeles du sujet pour la conception : dialectique activites développement*. Toulouse: Octares Editions, 2005.

[5] Venkatesh, Morris, Davis, et Davis, « User Acceptance of Information Technology: Toward a Unified View », *MIS Quarterly*, vol. 27, nº 3, p. 425, 2003.

[6] A. Cordier et B. Fuchs, « Interprétation Interactive de connaissances à partir de traces », in *IC2016:27es Journées francophones d'Ingénierie des Connaissances*, Montpellier, France, 2016.

[7] J.-C. Domenget, « La fragilité des usages numériques », *Les Cahiers du numérique*, vol. 9, nº 2, p. 47‑75, oct. 2013.

[8] A. Coutant et J.-C. Domenget, « Un cadre épistémologique pour enquêter sur les dispositifs sociotechniques d'information et de communication », p. 13, 2013.

[9] Q. Grossetti, C. Du Mouza, et N. Travers, « Tweet, Retweet et Follower : que recommander et à qui ? », in *AISR2017*, Paris, France, 2017.

[10] M. E. J. Newman, « Modularity and community structure in networks », *PNAS*, vol. 103, nº 23, p. 8577‑8582, juin 2006.

[11] P. Bourdieu, « Le capital social », *Actes de la Recherche en Sciences Sociales*, vol. 31, nº 1, p. 2‑3, 1980.

[12] N. Ellison et A. Thierry, « Réseaux sociaux, numérique et capital social », *Hermès, La Revue*, nº 59, p. 21‑23, 2011.

[13] H. Bourdeloie, « L'appropriation des dispositifs d'écriture numérique : translittératie et capitaux culturel et social », *Études de communication. langages, information, médiations*, nº 38, p. 23‑36, juin 2012.

Expérimentation du passage d'une activité débranchée au programme

Michel Adam[1], Moncef Daoud[1] et Patrice Frison[1,2]

[1] Université Bretagne Sud, Campus de Tohannic, 56017 Vannes Cedex, France
prenom.nom@univ-ubs.fr
[2] Laboratoire IRISA

Résumé L'apprentissage de la programmation peut se faire de différentes manières. On distingue habituellement les activités débranchées de l'activité branchée de programmation. Cependant le passage à l'activité de programmation est difficile. Dans un précédent article [4], nous avons proposé une démarche pour passer progressivement de l'activité débranchée à la programmation en utilisant AlgoTouch, un outil de programmation par manipulation directe. Une expérience basée sur cette démarche a été réalisée auprès d'un groupe de 39 collégiens âgés de 14 ans (en classe de troisième). Dans ce papier nous décrivons les différentes étapes de l'expérimentation et analysons les résultats obtenus.

Mots-clés : Algorithmes, Programmation pour débutants, Programmation par manipulation directe, Activités débranchées, Visualisation d'algorithmes.

1 Introduction

En France, depuis 2016, l'enseignement de la programmation doit désormais être dispensé depuis l'école primaire. Depuis cette réforme, la communauté scientifique s'intéresse à l'enseignement de l'informatique pour tout cycle d'apprentissage [1]. Il ne s'agit pas de se focaliser sur la programmation et l'apprentissage d'un langage. L'approche la plus communément admise consiste à développer d'abord des activités débranchées [2] et à réaliser des algorithmes sans utiliser de machines. L'idée étant de distinguer le concept d'algorithme, permettant de résoudre un problème, de celui d'un programme, associé à un langage de programmation, exécuté par une machine. Ces concepts (algorithme, machine, langage, information) sont les bases pour l'enseignement de la discipline informatique d'après la Société Informatique de France (SIF).

Cependant, le passage à l'activité de programmation est difficile [5]. Dans [4], nous avons proposé comment passer progressivement d'une activité débranchée à la réalisation d'un programme, en utilisant un outil de programmation par manipulation directe appelé AlgoTouch [3]. À partir d'une activité débranchée, nous avons montré comment transformer les objets et actions de cette activité en objets numériques et opérations de la machine. Dans un premier temps, l'apprenant va réaliser l'algorithme sur lequel il travaillait avec les éléments de la programmation (variables, opérateurs, comparateurs) disponibles avec AlgoTouch. Dans

un second temps, ayant constaté que certaines opérations se répètent, il enregistre ces opérations pour pouvoir les répéter automatiquement. AlgoTouch se charge de créer le programme correspondant. Ainsi, un algorithme complet est transformé progressivement en programme.

Dans le but de valider cette démarche, nous avons réalisé une première expérimentation. Les objectifs étaient triples : présenter et illustrer les notions d'algorithme et de programme ; faire réaliser un programme avec AlgoTouch ; évaluer l'intérêt des élèves sur la démarche et l'outil. Cet article détaille cette expérimentation réalisée sous forme d'un atelier et son évaluation.

2 Expérimentation

Un atelier a été réalisé avec des élèves de collège en classe de 3ème. Nous avons accueilli 39 élèves répartis en 3 groupes encadrés chacun par l'un des auteurs et deux étudiants de l'université (moniteurs). Les 6 moniteurs ont eu une courte formation au logiciel AlgoTouch ainsi qu'une préparation à l'atelier. Leur rôle principal était d'assister les élèves pendant la séance de travaux pratiques avec AlgoTouch. L'atelier, intitulé "Jouer avec des algorithmes et programmer sous AlgoTouch", s'est déroulé l'après-midi pendant deux heures. Son objectif était d'introduire d'une façon originale l'activité de programmation en illustrant les concepts d'algorithme, de données, de programme associé à un langage et exécuté par une machine. L'expérimentation s'est déroulée en trois phases : évaluation préliminaire des connaissances des élèves, réalisation pratique, évaluation de l'atelier.

Évaluation préliminaire Afin d'évaluer les connaissances acquises à la suite de l'atelier, un questionnaire préliminaire a été rempli par les élèves. Ils devaient estimer leurs connaissances sur certains concepts de l'informatique ainsi que sur quelques outils. La table 1 donne les résultats.

Table 1. Évaluation des connaissances des élèves selon 4 critères (par ordre croissant) : *Sans* ; *Débutant* ; *Base* ; *Avancé*.

Notions	Sans	Débutant	Base	Avancé
Activités débranchées	29	8	2	0
Algorithme	19	15	5	0
Langages - Programmation	10	25	4	0
Outil Scratch	0	16	23	0
Outils pour la robotique	8	22	9	0

Étonnamment, les élèves affirment n'avoir pas ou peu de connaissances sur la plupart des notions de base (algorithme, langage, programmation) alors qu'ils ont programmé avec Scratch et, pour la plupart, avec des robots.

Atelier Dans un premier temps (30 minutes), la notion d'algorithme a été expliquée aux élèves. Puis, ils ont pratiqué quelques activités débranchées. Elles consistaient à réaliser des algorithmes simples sans utilisation d'un ordinateur, comme par exemple la sélection du seau le plus lourd parmi cinq, à l'aide d'une balance à plateaux. Dans un second temps (30 minutes), on a montré comment transformer avec AlgoTouch un algorithme en un programme uniquement par manipulation des données. Les notions de variables, d'affectation, de séquentialité, d'alternative et de boucle ont été ainsi illustrées. La notion de langage a été introduite par la lecture du code généré. Les programmes ainsi obtenus ont été exécutés sur plusieurs jeux de données. Dans un troisième temps (une heure), les élèves ont joué avec les données d'un problème puis ont réalisé leur premier programme : recherche de la plus grande valeur d'un tableau simulant les poids des seaux de l'activité débranchée.

Pendant la séance, ils ont vite compris comment réaliser des affectations, des opérations et des comparaisons. À la fin de l'expérience, la plupart ont réussi à réaliser et exécuter le programme correspondant à l'activité débranchée.

Table 2. Évaluation de l'atelier. Pour chaque question, la réponse était soit : (1) Pas du tout d'accord ; (2) Pas d'accord ; (3) Assez d'accord ; (4) Tout à fait d'accord. La partie droite de la table indique le nombre de réponses pour chaque cas.

Questions	(1)	(2)	(3)	(4)
Q1. J'ai apprécié les activités débranchées	0	0	17	22
Q2. L'activité débranchée m'a aidé à comprendre la résolution de problème	2	7	12	18
Q3. La transition vers le numérique avec AlgoTouch m'a permis de mieux comprendre la notion de programme	2	4	22	11
Q4. J'ai compris la représentation (le stockage) de l'information	5	5	24	5
Q5. J'ai compris le fonctionnement de la machine	0	9	21	9
Q6. AlgoTouch permet l'apprentissage progressif du concept de langage de programmation	0	4	21	14
Q7. J'ai compris le principe et le fonctionnement des algorithmes	2	6	15	16
Q8. AlgoTouch m'a aidé à la construction progressive d'un programme	2	5	19	13
Q9. Je trouve qu'AlgoTouch est facile à utiliser	5	7	15	12
Q10. J'ai apprécié les activités branchées avec AlgoTouch	2	2	14	21
Total par choix de réponse	20	49	180	141
En pourcentage	5%	13%	46%	36%

Évaluation de l'atelier En fin de séance, une deuxième fiche proposait dix questions. Les résultats présentés dans la table 2 montrent :

1. une dominance pour les réponses de type "Assez d'accord" et "Tout à fait d'accord", les apprenants sont majoritairement d'accord à 82% ;

2. en général, une satisfaction des élèves pour les activités débranchées (Q1, Q2) et branchées (Q8, Q10) réalisées avec des taux respectifs de 100% et 90% ;
3. une tendance forte pour affirmer avoir compris les concepts de base (Q3, Q4, Q5, Q6, Q7) : algorithme, machine, programmation, information, ce qui marque un progrès par rapport au début de l'atelier ;
4. une satisfaction à l'utilisation d'AlgoTouch (Q9).

Ces réponses sont à prendre avec précaution. Comme il s'agit d'une auto-évaluation par les élèves, les résultats sont en général positifs.

3 Conclusion et perspectives

Dans ce papier, nous avons relaté une expérience avec des élèves de collège de classe de 3ème. À partir d'une activité débranchée, en utilisant AlgoTouch et par manipulation des données, ils ont réalisé le programme correspondant sans écrire une ligne de code. Ensuite, les élèves ont répondu à un questionnaire purement qualitatif. Ils indiquent avoir pris conscience de la notion d'algorithme et de sa traduction en programme. Ils déclarent avoir compris la notion de stockage de l'information et le fonctionnement de la machine. Par la pratique, avec AlgoTouch, ils ont pu découvrir la notion de langage de programmation.

L'approche proposée paraît adaptée pour l'introduction de la programmation à des élèves. En effet, il semble que l'objectif initial d'illustrer les concepts d'algorithme, de machine, de langage, d'information, ait été atteint. De nouvelles expérimentations doivent être réalisées pour confirmer notre approche et mesurer le savoir-faire des participants. Il s'agira, sur une durée plus longue, de réaliser plusieurs programmes de difficultés croissantes.

Références

1. Baron, G.L., Drot-Delange, B. : L'éducation à l'informatique à l'école primaire. 1014, Bulletin de la Société informatique de France **8**, 73–79 (Nov 2016), https ://hal.archives-ouvertes.fr/hal-01403598
2. Bell, T., Alexander, J., Freeman, I., Grimley, M. : Computer science unplugged : School students doing real computing without computers. The New Zealand Journal of Applied Computing and Information Technology **13**(1), 20–29 (2009)
3. Frison, P. : A teaching assistant for algorithm construction. In : Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education. pp. 9–14. ITiCSE '15, ACM, New York, NY, USA (2015). https://doi.org/10.1145/2729094.2742588
4. Frison, P., Daoud, M., Adam, M. : Transition didactique de l'activité débranchée à la programmation avec AlgoTouch. In : Didapro 7 – DidaSTIC. De 0 à 1 ou l'heure de l'informatique à l'école. pp. 1–17. Lausanne, Switzerland (Feb 2018), https ://hal.archives-ouvertes.fr/hal-01753119
5. Guzdial, M. : Programming environments for novices. Computer science education research **2004**, 127–154 (2004)

Un Karaoké pour Entraîner Prosodie et Compréhension en Lecture

Erika Godde[1,2], Gérard Bailly[1], Marie-Line Bosse[2]

[1] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-Lab, 38000 Grenoble, France
(Erika.godde; Gerard.bailly)@gipsa-lab.fr
[2] Univ. Grenoble Alpes, CNRS, LPNC, 38000 Grenoble, France
Marie-Line.bosse@univ-grenoble-alpes.fr

Résumé. La prosodie en lecture est une compétence dont le développement débute avec l'acquisition du décodage et se poursuit tout au long de la scolarité. Elle joue notamment un rôle dans la compréhension des textes lus. Nous utilisons ici RAKE, un karoké de lecture permettant un entrainement audio-visuel de la prosodie. Avec des enfants de CE2, CM1 et CM2, décodeurs mais faibles lecteurs, nous testons ici l'effet de RAKE, et notamment du surlignage en groupe de souffle, sur le phrasé, l'expressivité et la compréhension, aprés 10 séances d'utilisation.

Mots-clés. Lecture, prosodie, compréhension écrite, entrainement, application éducative

Abstract. Reading prosody is a skill developing from the early reading acquisition to the end of education. Prosody development has an impact on reading comprehension. The RAKE app is a reading karaoke used to perform an audio-visual enhanced reading-while-listening task. In this study we used RAKE in a 10 weeks training program with grade 3 to grade 5 pupils, poor readers but with good decoding skills. It aims at improving phrasing and expressivity, and so comprehension, by highlightening breath group.

Keywords. Reading, prosody, written comprehension, intervention, educational app

1 Introduction

Dans l'apprentissage de la lecture, l'accent a longtemps été mis sur le décodage, l'automatisation et la vitesse. Pourtant un élève qui lit vite, même sans erreur n'est pas forcément un bon lecteur. Depuis une dizaine d'années plusieurs auteurs soulignent l'importance de faire évoluer la définition de la fluence en y ajoutant la prosodie en lecture [1].La prosodie est liée à la compréhension de l'auditeur, mais aussi du lecteur. Cette relation est une piste qu'il peut être intéressant d'explorer dans le cas des enfants « bons décodeurs, mauvais compreneurs ».

La prosodie recouvre deux types de compétences : le phrasé, caractérisé par les pauses et les variations d'intonation, étroitement lié à la syntaxe et à la ponctuation, et

l'expressivité, traduite par les variations d'intonations et d'intensité de la voix, donnant vie au texte [2].

Le lien entre prosodie et compréhension varie au cours de l'acquisition de la lecture. Un lien direct entre prosodie, en particulier pause et respect de la syntaxe, et compréhension apparaît en CE2 [3]. Ce lien se renforce avec l'âge et la maitrise des compétences en lecture [4], la prosodie expliquant une part de plus en plus importante de la compréhension [5]. Une intervention sur la prosodie pourrait donc permettre d'améliorer également la compréhension si elle est effectuée durant la période adéquate.

L'application pour tablette que nous présentons ici est un karaoké de lecture permettant d'entrainer le phrasé et l'expressivité. Lors d'un pré-test du karaoké avec des CE2, nous avons pu montrer, sur une seule lecture, que le type de surlignage avait un effet sur les pauses respiratoires, et donc sur le phrasé [6]. Nous poursuivons ces travaux dans cette étude en utilisant le karaoké durant 10 séances, afin de tester les effets d'une utilisation sur un plus long terme. Nous nous intéressons ici à des enfants décodeurs qui ne parviennent pas à entrer dans l'expressivité et la compréhension des textes. La question principale est l'effet de l'utilisation régulière et répétée de l'application sur les compétences prosodiques des textes entrainés et l'effet du type de surlignage sur l'acquisition de ces compétences. Nous nous intéresserons aussi au transfert de ces compétences à un texte non entrainé et à leur effet sur la compréhension écrite.

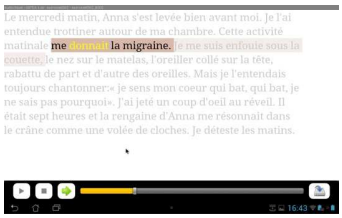


Fig. 1. Interface du karaoke de lecture RAKE (ici un surlignage groupe de souffle).

2 RAKE : un karaoké de lecture

Présentation. RAKE (Reading Assistance by KaraokE) est une application de karaoké de lecture développé par le laboratoire GIPSA. L'interface est présentée en figure 1. L'application propose un entrainement audio-visuel à la lecture, de type « lire en écoutant ». L'utilisateur muni d'un casque entend un lecteur expert lire le texte affiché à l'écran. Simultanément à la lecture experte, le texte lu est mis en surbrillance. Quatre niveaux de mise en surbrillance sont possibles : mot seul, mot et syllabe, mot et groupe de sens, mot et groupe de souffle (fig.1). L'utilisateur doit essayer de lire à voix haute en se synchronisant avec la voix du lecteur expert. Nous nous intéressons ici particulièrement au niveau groupe de souffle, le plus complexe, visant à entrainer spécifiquement le phrasé et la respiration.

Principe. L'apprenti lecteur est entrainé explicitement, par un indiçage du texte, et par le modèle de prosodie donné par le lecteur expert. L'indiçage de texte par le surlignage des groupes de sens et de souffle vise à améliorer le phrasé des élèves [7], permettant de repérer et placer ses pauses sur des frontières syntaxiques [6]. Le lecteur expert entendu donne le modèle d'un phrasé et d'une intonation experts, permettant les progrès de l'apprenti en termes de vitesse, phrasé et expressivité [8]. Ce type d'entrainement permet également un apprentissage plus implicite. En effet, l'apprenti lecteur suit le lecteur expert dans un délai très court (100 ms). Ce close-shadowing permet le transfert des habilités motrices, ici la respiration [9].

3 Entrainement

Participants. L'entrainement concerne 97 élèves de CE2, CM1 et CM2 scolarisés dans 8 classes de 2 écoles élémentaires de l'agglomération de Grenoble. Suite à des pré-tests, nous avons sélectionné 53 enfants correspondants au profil recherché : ayant acquis le décodage, mais avec un score de prosodie, évalué par l'échelle de fluence multidimensionnelle adaptée pour le français [1], inférieur à 2 sur 4 en phrasé et expressivité. Ces élèves sont répartis en deux groupes s'entrainant avec des surlignages différents : un groupe bénéficiant uniquement d'un surlignage des mots (groupe M) et un groupe bénéficiant d'un surlignage des groupes de souffle (groupe S). Un groupe contrôle est constitué avec 44 élèves des mêmes classes, non sélectionnés (groupe C).

Protocole. Les séances d'entrainement se font en petits groupes de 5 à 8 participants, hors de la classe, sous la supervision d'un chercheur, par séance de 15 min. Après une première de séance découverte de l'application, les élèves entrainés participent à 3 séances par semaine, pendant 3 semaines. Durant ces séances, les élèves lisent à 3 reprises un texte avec le karaoké, le texte change chaque semaine.

Lors des pré-tests et post-tests nous avons mesuré la compréhension écrite globale (texte avec question), locale (Maze-test), la conscience syntaxique, et la vitesse de lecture. Les enfants ont également été enregistrés pendant la lecture des 3 textes entrainés et d'un texte de référence, non entrainé.

4 Résultats préliminaires

Prosodie. Tout d'abord, une analyse acoustique permet de caractériser précisément les pauses respiratoires et non respiratoires (nombre, durée, placement) et les variations d'intonation. Ces analyses sont en cours. Nous évaluons également subjectivement les enregistrements des textes lus en utilisant une échelle d'évaluation subjective de la prosodie [1]. Cette échelle donne un score sur 4 pour l'expressivité, le phrasé, le décodage et la vitesse, soit un score total sur 16. L'évaluation est effectuée par 3 évaluateurs indépendants. Nous disposons pour l'instant des évaluations de 2 évaluateurs. Comme attendu, les groupes S et M ont progressé sur tous les paramètres

des textes entrainés. Ces premiers résultats semblent également montrer un effet du surlignage, avec des progrès plus importants en phrasé et expressivité pour le groupe S. Ces premiers résultats restent à confirmer avec les données du 3ème évaluateur.

Conscience syntaxique. Les premiers résultats montrent un effet du surlignage groupe de souffle sur la conscience syntaxique (fig. 2). Les enfants du groupe S font des progrès significatifs, ce qui n'est pas le cas des groupes C et M.

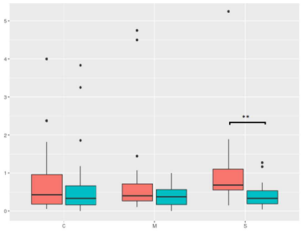


Fig. 2. Evolution du score de conscience syntaxique (0 étant le score maximal) entre pré-test (rose) et post-test (bleu) pour les différents groupes.

De nombreuses données restent à traiter, mais ces premiers résultats sont encourageants et semblent aller dans le sens de nos hypothèses, avec une amélioration de l'expressivité et du phrasé, et un impact sur la conscience syntaxique, particulièrement quand les groupes de souffle sont explicitement mis en valeur.

Références

1. Rasinski, T. V.: Assessing reading fluency. Pacific Resources for Education and Learning (PREL) (2004)
2. Bolinger, D., & Bolinger, D. L. M.: Intonation and its uses: Melody in grammar and discourse. Stanford university press. (1989)
3. Veenendaal, N. J., Groen, M. A., & Verhoeven, L.: The role of speech prosody and text reading prosody in children's reading comprehension. British Journal of Educational Psychology, 84(4) (2014) 521-536
4. Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F.: The role of reading fluency in children's text comprehension. Frontiers in psychology, 6, 1810 (2015)
5. Yildirim, K., Rasinski, T., & Kaya, D.: Fluency and comprehension of narrative texts in Turkish students in grades 4 through 8. Education 3-13 (2018) 1-10
6. Godde, E., Bailly, G., Escudero, D., Bosse, M. L., Bianco, M., & Vilain, C.: Improving fluency of young readers: introducing a Karaoke to learn how to breath during a Reading-while-Listening task. In 7th ISCA Workshop on Speech and Language Technology in Education (SLaTE 2017) (2017)
7. LeVasseur, V. M., Macaruso, P., & Shankweiler, D.: Promoting gains in reading fluency: A comparison of three approaches. Reading and Writing, 21(3) (2008) 205-230
8. Chard, D. J., Vaughn, S., and Tyler, B.-J.: A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. Journal of learning disabilities, 35(5) (2002) 386–406
9. G. Bailly, A. Rochet-Capellan, and C. Vilain,: Adaptation of respiratory patterns in collaborative reading in Interspeech, (2013) 1653-1657

Valorisation des données pour l'amélioration des plateformes d'hébergement de MOOC

Frédérick Bigrat[1], Mélody Laurent[2], Cécile Méadel[3]

[1] Université Paris 2 Panthéon Assas / IFP - CARISM, 5/7 avenue Vavin, 75006 Paris, France
frederick.bigrat@ac-paris.fr
[2] Chancellerie des universités de Paris / SIRIS – 46, rue Saint Jacques, 75005 Paris, France
melody.laurent@siris.sorbonne.fr
[3] Université Paris 2 Panthéon Assas / IFP - CARISM, 5/7 avenue Vavin, 75006 Paris, France
cecile.meadel@u-paris2.fr

Résumé. Durant toute la durée d'un MOOC (Massive Open Online Courses), les professeurs, les ingénieurs pédagogiques et les apprenants génèrent soit de manière directe soit de manière indirecte des données en quantité très importante. Certaines sont utilisées par les hébergeurs de plateformes pour alimenter leurs statistiques ou pour faciliter l'expérience de l'apprenant et d'autres sont encore peu valorisées. Comment peut-on construire des indicateurs efficients sur la base d'un système de gestion de traces modélisées s'appuyant sur les forums d'une plateforme d'hébergement de MOOC ?

Mots-clés. MOOC, Données, Plateformes, Technologies, Agent conversationnel

Abstract. Throughout the duration of a Massive Open Online Courses (MOOC), Professors, instructional-designers and learners either directly or indirectly generate very large amounts of data. Some are used by platform hosts to feed their statistics or to facilitate the learner experience and others are still undervalued. How can we build efficient indicators based on a modeled track management system relying on the forums of a MOOC Hosting platform ?

Keywords. MOOC, Data, Platforms, Technology, Chatbot

1 Introduction

Dans le secteur spécifique de la formation à distance [1] nous pouvons différencier deux types de données, les statiques pour lesquelles nous distinguerons celles qui concernent les caractéristiques du cours proposé (cours, documents en ligne, liens, etc.) ainsi que les données personnelles relatives plus spécifiquement à l'apprenant (nom, date de naissance, etc.), et les données dynamiques qui sont celles générées par les apprenants tout au long de la formation (forum, quizz, interactions vidéo, etc.) mais également par les plateformes elles-mêmes (connexions, etc.).

Dans le cadre de cet article, nous définirons les plateformes d'hébergement de MOOC comme étant un ensemble d'outils logiciels dont l'objectif est de permettre le pilotage d'enseignements proposés à distance, et conçues pour la gestion des parcours le suivi des étudiants et la diffusion des contenus d'apprentissages, suivant la définition proposée par [2].

Le problème principal rencontré par les utilisateurs dans les forums est de trouver des réponses aux questions qu'ils posent avec un faible temps d'attente et surtout de savoir qui sera leur interlocuteur [3]. L'hypothèse que nous formulons est que

des indicateurs basés sur un système de modélisation des traces des forums devraient permettre de dégager des indicateurs fonctionnels qui pourraient contribuer à améliorer l'usage d'un environnement informatique pour l'apprentissage humain (EIAH).

Nous commencerons par apporter une définition des données qui nous semblent pertinentes dans le cadre de cet article, puis nous aborderons l'évolution des dispositifs à destination des apprenants, nous étudierons ensuite les éléments relatifs à notre proposition de construction d'indicateurs à partir d'un système de gestion de traces modélisées, avant d'aborder notre conclusion.

2 Les données générées

Le développement des MOOC en France offre la possibilité de recueillir de grandes quantités de données [4]. Ces nombreuses données sont produites par les acteurs suivants, les enseignants-chercheurs, les ingénieurs pédagogiques et les apprenants. Toutes ces données qu'elles soient statiques ou dynamiques peuvent fournir des informations non déclaratives [5] sur les formations proposées et l'utilisation des plateformes.

Nous définissons les données que nous retenons dans cet article, comme étant celles qui sont générées spécifiquement dans les forums, à savoir : les commentaires, les échanges de liens web et de vidéos, et les partages d'expériences. Nous prenons également en compte les données supplémentaires produites par les enseignants-chercheurs lors de leurs interactions avec les apprenants.

Au travers des différents types d'hébergeurs de MOOC [6] qui existent aujourd'hui en France dans le secteur de l'enseignement, il est déjà possible d'obtenir des niveaux d'agrégation de certaines données en vue de statistiques grâce aux dispositifs de « Learning Analytics ». Il nous semble également possible d'utiliser ces dispositifs pour valoriser les données que nous venons de définir au profit cette fois non pas des hébergeurs mais des apprenants afin de faciliter leur expérience utilisateur.

3 Evolution des dispositifs à destination des apprenants

De la mise en ligne d'un MOOC jusqu'à son évaluation, les apprenants sont immergés dans un environnement relativement figé et représenté par des dispositifs de transmission directe du savoir que sont, de manière non exhaustive : l'enregistrement vidéo, les documents pédagogiques et des dispositifs d'interactions que sont les quizz, les forums et les visioconférences.

Nous pensons que les interactions entre les intervenants, comme d'ailleurs cela peut-être le cas en présentiel, sont vecteurs d'une grande richesse et formalisent des endroits où les principaux utilisateurs et acteurs des plateformes d'hébergement de MOOC se retrouvent pour d'une part échanger sur leurs propres expériences. Dans le cadre de l'enseignement à distance, les plateformes doivent en permanence évoluer afin d'être efficientes et répondre aux attentes des apprenants. Selon [7] le problème que les utilisateurs rencontrent dans les forums est de trouver des réponses aux questions qu'ils posent avec un faible temps d'attente et surtout de savoir qui sera son interlocuteur.

4 Vers l'amélioration des plateformes grâce aux données

Notre proposition au vu de ce que nous venons d'exposer précédemment, est que les données produites, plus spécifiquement dans les forums des lieux d'échanges entre étudiants mais aussi entre ces derniers et l'équipe enseignante devraient selon nous pouvoir bénéficier d'un traitement particulier.

Ainsi, au vu des interactions qui se déroulent quotidiennement dans les forums [8], il nous semble intéressant de prendre en compte les nombreuses questions et réponses associées qui y sont identifiées afin de modéliser la nature et les flux de données pour proposer la mise en ligne d'un « agent conversationnel » (chatbot)

Pouvoir valoriser les données des utilisateurs des forums nous semble être une approche très intéressante, car les forums sont souvent décrits comme un ingrédient essentiel d'un cours en ligne efficace, et cette communication asynchrone représente selon [9] la principale interaction éducative sur ce type de cours, et un nombre croissant d'études posent ainsi la question de l'utilisation du forum, de son rôle (forum de discussion ou bien site de questions/réponses) ou du rôle des modérateurs et des divers participants.

5 Conclusion et perspectives

Cet article de recherche en cours, a présenté dans le cadre de la formation à distance comment les données qui sont générées sur les plateformes d'hébergements de MOOC en France dans les forums pourraient-être utilisées à des fins d'amélioration de dispositifs technologiques et pédagogiques existants.

L'hypothèse que nous avons formulée est que l'utilisation d'une interface homme-machine pourrait permettre d'améliorer les échanges avec les apprenants. Notre objectif est d'améliorer les recommandations d'utilisation des contenus d'une plateforme d'hébergement de MOOC pour les apprenants par la mise en œuvre d'un « agent conversationnel ».

Ce dispositif pourrait être de nature à faciliter l'expérience utilisateur et permettrait de prendre en compte les demandes, souvent légitimes, des apprenants qui soient découvrent ce nouvel environnement d'apprentissage, soient cherchent à enrichir les informations qu'ils possèdent déjà, surtout pour les questions techniques ou de navigation qui sont très fréquentes et énergivores pour les ingénieurs pédagogiques lors de la première semaine d'un MOOC.

Une perspective à ces travaux en cours est que l'utilisation de cette technologie étant de plus en plus courante sur internet, les interactions avec les apprenants pourraient à terme se faire de manière automatisée.

Références

1. Jézégou, A. : La formation à distance : enjeux, perspectives et limites de l'individualisation. L'Harmattan, (1998).
2. Prat, M. : Réussir votre projet e-learning : pédagogie, méthodes et outils de conception, déploiement, évaluation. Herblain : ENI éd, (2012)
3. Ezen-Can, A., Boyer KE., Kellogg, S., Booth, S. : Unsupervised modeling for understanding Mooc discussion forums : a learning analytics approach. LAK'15 : Proceedings of the Fifth International Conference on Learning Analytics and Knowledge, ACM (2015).
4. Coulombe, C. : Vers une pédagogie pilotée par les données. TELUQ – Université du Québec, (2017).
5. Boelaert, J., Khaneboubi, M. : Une méthode de visualisation de traces d'activités sur la plateforme open edX : le cas du Mooc « Enseigner et former avec le numérique ». Adjectif.net, (2015).
6. Cisel, M., Bruillard, E. : Chronique des MOOC. Sciences et technologies de l'Information et de la Communication pour l'Éducation et la Formation, ATIEF, (2013).
7. Pomerol JC., Epelboin, Y., Thoury, C. : Les MOOC : conception, usages et modèles économiques. DUNOD, (2014).
8. Bar-Hen, A., Javaux, H., Villa-Vialaneix, N. : Analyse statistique des profils et de l'activité des participants d'un MOOC. Revue internationale des technologies en pédagogie universitaire, (2015).
9. Anderson, Huttenlocher, Kleinberg et Leskovec : Steering user behavior with badges, Proceedings of the 22nd international conference on World Wide Web (2014).

Apports d'un outil d'intelligence artificielle à l'enseignement-apprentissage des langues

Grégory Miras[1], Marie Lefevre[2], Najib Arbach[3], Louis Rapilly[3], Théo Dumarski[3]

[1] Université de Rouen Normandie, EA 7474, 7 rue Thomas Becket
76830 Mont Saint Aignan Cedex, France
gregory.miras@univ-rouen.fr
[2] Université de Lyon, CNRS
Université Lyon 1, LIRIS, UMR 5205, F-69622, France
marie.lefevre@liris.cnrs.fr
[3] FRELLO, 34 rue Raymond Aron, 76130 Mont-Saint-Aignan, France
team@frello.fr

Résumé. Le projet Apprentissage Adaptatif des Langues par le NUmérique (2ALN) prend la forme d'une plateforme en ligne FRELLO qui vise à proposer des parcours individualisés d'apprentissage des langues. Ces parcours adaptatifs sont assistés par des techniques d'intelligence artificielle. Les algorithmes adaptatifs créés ont pour but de déterminer des liens, de manière prédictive, entre les compétences langagières et culturelles visées. L'hypothèse est qu'une bonne personnalisation facilitera le processus d'apprentissage, augmentera l'investissement des apprenants et pourra être une aide de prise à la décision pour les enseignants qui utilisent la plateforme en hybride.

Mots-clés. Apprentissage adaptatif, intelligence artificielle, langues étrangères, numérique, personnalisation

Abstract. The objective of project FRELLO is to build an online platform to teach and learn foreign languages. The learning programs will be based on adaptive learning tools, allowing a more personalized learning experience. This individualization relies on developing artificial intelligence for education. The created adaptive algorithms should help to foresee links between language and cultural competences. The hypothesis is that an accurate personalization will ease the learning processes, increase learners' investment and help teachers adapt their teaching programs.

Keywords. Adaptive learning, artificial intelligence, foreign languages, digital, individualization

1 Introduction

Comme toutes les activités sociales, les pratiques d'enseignement-apprentissage des langues évoluent en fonction des technologies disponibles dans la société. Depuis plusieurs années, l'intelligence artificielle a permis de résoudre un certain nombre de défis dans des domaines variés comme la linguistique appliquée [1], notamment dans

le traitement automatique des langues, mais aussi la didactique des langues-cultures. Cependant, même si l'introduction des techniques d'IA dans les EIAH pour proposer des systèmes adaptatifs a fait l'objet de nombreux travaux [2], l'étude de son impact dans l'apprentissage des langues spécifiquement reste plus rare, notamment face à des préoccupations qui émergent des enseignants (économies, fracture numérique, etc.) et plusieurs rapports constatent les difficultés que rencontre l'apprentissage adaptatif malgré les investissements consacrés. Le projet Apprentissage Adaptatif des Langues par le NUmérique (2ALN) prend la forme d'une plateforme en ligne FRELLO qui vise à proposer des parcours individualisés d'apprentissage des langues. Nous chercherons à déterminer si l'intelligence artificielle, au service de l'enseignement/apprentissage des langues étrangères, peut être un levier utile à l'apprenant comme à l'enseignant. Nous nous appuierons sur la création de parcours en français langue étrangère (FLE).

2 Intelligence artificielle et enseignement-apprentissage des langues étrangères

Les premières applications des technologies de l'information et de la communication pour l'enseignement (TICE) étaient, comme le mentionne Desmet [3] : "complètement contrôlées par le programme : le logiciel proposait un parcours d'apprentissage fixe que l'apprenant était censé parcourir, sans guère d'intervention dans ce processus". Depuis plusieurs décennies, en éducation, le développement de l'apprentissage adaptatif est corrélé à celui des technologies en général, et les investissements consacrés (financiers et humains) sont considérables [4], sans qu'il n'y ait encore certitude sur son efficience réelle [5]. Les conditions d'un apprentissage adaptatif efficace nécessitent donc une phase de recherche et développement. Dans le même temps, les travaux du domaine [6] montrent que pour qu'il y ait apprentissage, il est nécessaire de stimuler des besoins sociaux réalistes (le socio-constructivisme) pour favoriser l'investissement tout mettant en place des exercices structuraux pour stabiliser des processus neurocognitifs (le béhaviorisme). L'une de nos hypothèses est que cette dimension sociale et cette personnalisation réduiront le fort décrochage des utilisateurs-apprenants sur les solutions d'apprentissage des langues en ligne. Les outils d'intelligence artificielle peuvent rendre possible cette individualisation en l'automatisant partiellement afin de libérer du temps à l'enseignant qui pourra mieux accompagner l'apprenant tout en réduisant les coûts [7]. Lors des tâches d'entraînement, permettant une forme d'automatisation d'éléments notamment linguistiques, l'apprenant travaille en relative autonomie tout en disposant d'un parcours personnalisé à t+1 ; et lors des tâches sociales, qui réintroduisent une intentionnalité, l'enseignant peut se focaliser sur les difficultés ayant émergé et non traitées. Les systèmes de ce type disponibles actuellement, en France, sont rares et l'accès pour les équipes pédagogiques aux paramètres des algorithmes complexe.

3 Apprentissage adaptatif des langues par le numérique

3.1 La plateforme FRELLO

Ce projet prend la forme d'une plateforme numérique en ligne (https://frello.eu/) dont la première version a été lancée en 2018. Il s'appuie sur une équipe pluridisciplinaire composée des acteurs de l'entreprise, de chercheurs en didactique des langues-cultures, en informatique mais aussi de développeurs web. Le processus général de la plateforme est modélisé par les étapes suivantes. L'enseignant assigne un module à l'apprenant qui a alors une tâche d'entraînement à effectuer. Son apprentissage et ses interactions avec la plateforme génèrent des données d'apprentissage qui permettent à l'enseignant de suivre l'évolution de ses compétences. Les données d'apprentissage sont envoyées au Learning Record Store (LRS) qui les stocke et les rend exportables, au format xAPI, pour l'algorithme de recommandation d'activités. Cette étape permet la définition de recommandations, en fonction des données d'apprentissage et de stratégies pédagogiques, qui sont ainsi renvoyées à la plateforme. L'apprenant reçoit des recommandations de modules pour atteindre ses objectifs en continuant son parcours individualisé d'apprentissage.

3.2 Structure pédagogique et algorithme adaptatif

Dès le début, la structure pédagogique de la plateforme a été pensée pour, à la fois, correspondre aux attentes d'un curriculum pédagogique en langues mais aussi pour rendre possible la mise en place d'un algorithme adaptatif par une organisation en blocs de modules. L'architecture du projet repose sur une progression par niveaux de langue, chaque niveau étant composé de plusieurs unités. Une unité se décompose à son tour en plusieurs leçons qui visent des savoir-faire communicatifs. Ces derniers intègrent différents modules de langue (lexique, grammaire, etc.) indispensables pour la maîtrise des savoir-faire. L'algorithme d'apprentissage adaptatif repose sur le modèle PERSUA2 [8] qui permet à une équipe pédagogique de définir ses propres besoins de personnalisation. Dans le cadre de la plateforme FRELLO, l'algorithme est paramétré sur trois principales règles : 1) la cohérence entre les modules et les savoir-faire ; 2) une progression en difficulté (entre les niveaux, les unités, les leçons et les modules) ; 3) les précédentes interactions entre les modules pédagogiques et les apprenants-utilisateurs. A ce niveau, le paramétrage de l'algorithme devient spécifique à la structuration de l'apprentissage des langues-cultures en introduisant la notion de savoir-faire communicatif. Ces règles permettent de proposer trois recommandations de modules dès qu'un apprenant a terminé trois modules. A ce stade, la recherche visera à déterminer si les recommandations personnalisées permettent un développement langagier effectif des apprenants et si les recommandations sont une aide à la prise de décision des enseignants.

3.3 Données préliminaires

Cette étude est au stade expérimental puisque la mise en place de l'algorithme d'apprentissage adaptatif est en phase de déploiement. Les données produites sont multiples puisque le système permet de récupérer des traces à chaque moment de l'interaction apprenant-système (temps de visionnage, clics, parcours, etc.), enseignant-système (choix, retours formatifs, etc.) et système-algorithme (choix du moteur de recommandation en fonction des données produites par les utilisateurs). Les premières

données sont issues des parcours d'apprentissage d'environ 150 apprenants de niveaux A1 et A2 du Cadre Européen Commun de Référence pour les Langues, tous adultes, aux origines linguistiques différentes. Les parcours socio-culturels sont aussi très épars (exilés non scolarisés, étudiants, entrepreneurs étrangers). Différents modules sont automatiquement assignés dans la progression d'une "leçon". Au fur et à mesure de la réalisation de ces modules, l'algorithme adaptatif fait des recommandations à l'apprenant en fonction de sa réussite ou non-réussite. L'apprenant peut alors continuer son parcours assigné ou faire d'autres modules pour renforcer certains points avant de revenir plus tard à la progression de la leçon. A terme, les recommandations permettront de créer automatiquement des parcours non linéaires et individualisés.

4 Conclusion et perspectives

Le projet 2ALN tire sa particularité du fait de l'adaptation d'un algorithme de recommandation proposé pour n'importe quelle discipline, dans le contexte des EIAH, à un contexte particulier nécessitant une expertise linguistique. L'observation de nos premières données montre que l'outil d'apprentissage adaptatif permet une pédagogie différenciée qui optimise l'intervention en présentiel. En termes de perspectives, si nous parvenons à vérifier nos hypothèses empiriquement, nous envisageons de combiner les potentialités de l'apprentissage hybride et adaptatif avec celles de la ludification pour nous rapprocher du modèle ALGAE (Adaptive Learning GAme dEsign [9]).

References

1. Miras, G., Boulton, A., Kübler, N., Narcy-Combes, J.-P.: Association Française de Linguistique Appliquée (AFLA). European Journal of Applied Linguistics. 6, 2 (2018) 315-326.
2. Markowska-Kaczmar U., Kwasnicka H., Paradowski M.: Intelligent Techniques in Personalization of Learning in e-Learning Systems. Studies in Computational Intelligence, Computational Intelligence for Technology Enhanced Learning, Vol. 273 (2010) 1-23.
3. Desmet, P.: L'enseignement/apprentissage des langues à l'ère du numérique : tendances récentes et défis. RFLA. Vol. xi, 1, (2006) 119-138.
4. Ball, S. J.: *Global Education Inc.*, Oxon: Routledge, Abingdon (2012).
5. SRI Education: *Lessons Learned from Early Implementations of Adaptive Courseware*. Menlo Park, CA: SRI Education (2016).
6. Narcy-Combes, J-P.: *Didactique des langues et TIC : vers une recherche-action responsable*, Paris: Éditions OPHRYS (2005).
7. Brudermann, C.: Tutorat en ligne et rétroactions correctives à distance – Vers un modèle de médiation pour la production en langue étrangère. ALSIC [En ligne]. Vol. 16 (2013).
8. Lefevre L., Guin N., Jean-Daubias S.: Personnaliser des activités pédagogiques de manière unifiée : une solution à la diversité des dispositifs. In STICEF, 19 (2012) 309-351.
9. Lavieri Jr, E. D.: A study of adaptive learning for educational game design (Thèse de doctorat, Colorado Technical University) (2014).

Évaluation de la satisfaction en formation pour adulte

Laurie Acensio[1,2], Frédéric Hoogstoel[2,], Luigi Lancieri[2]

[1] Lexiane Formation
[2] Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille
Université de Lille
F-59000 Lille, France
laurie.acensio@lexiane.fr, frederic.hoogstoel@univ-lille.fr, luigi.lancieri@univ-lille.fr

Résumé. Le contexte sociétal de la Formation Tout au Long de la Vie encourage les apprenants à s'engager dans une action de formation durant leur carrière professionnelle. L'action de formation pour adulte se concrétise principalement à travers un groupe de formation en présentiel évalué selon l'appréciation individuelle et immédiate de l'apprenant en fin de session. Dans cet article, l'analyse statistique consiste à identifier les critères qui influencent favorablement l'évaluation de la formation. La finalité d'application est de déterminer des critères de décision pour la modélisation d'un système d'aide à la composition des groupes de formation.

Mots-clés. Evaluation de la formation ; satisfaction ; FTLV

Abstract. The societal context of Lifelong Learning encourages learners to engage in training action during their professional career. The training action for adult is concretized mainly through a group of face-to-face training evaluated according to the immediate appreciation of the learner at the end of the session. In this article, statistical analysis consists of identifying the criteria that favorably influence the evaluation of training. The purpose of application is to determine decision criteria for the modeling of a support system for the composition of training groups.

Keywords. Evaluation of the training ; satisfaction ; Lifelong Learning.

1 Introduction

Initié dans le domaine de la psychologie du travail, le concept de la satisfaction a été particulièrement théorisé pour la prédiction du comportement de l'utilisateur dans une perspective marketing. Dans un contexte de formation continue, la notion de « satisfaction » n'est pas clairement définie car peu étudiée mais reste étroitement liée à l'évaluation de la formation. L'évaluation consiste à apprécier la valeur d'une formation sur la base de standards ou de critères spécifiques (…) dans une démarche de qualité [1]. Selon [2], l'apprenant est un acteur incontournable de l'évaluation : la qualité d'une formation résulte d'une adéquation entre ce dont l'apprenant est légitimement en droit d'attendre et la réalité évaluée. En pratique, l'évaluation se concrétise par l'administration d'un questionnaire en fin de formation. L'analyse des

réponses vise à vérifier l'atteinte d'objectifs définis préalablement. L'appréciation de l'action de formation se limite donc à formaliser la perception de l'apprenant sur une expérience de formation unique. Largement utilisé par les professionnels, les différentes versions du questionnaire d'évaluation de [3], [4] se focalisent sur quatre niveaux (réaction, apprentissage, comportement, résultat). Le caractère hiérarchique et causal du modèle émet l'hypothèse que le premier niveau (réaction) lié à la satisfaction générale conditionne la suite du processus de formation. Dans le cadre de cette étude, nous nous limitons à analyser les scores émis par les apprenants liées à la satisfaction globale et la satisfaction liée à la collaboration. Par la suite, les résultats visent à déterminer à priori les critères optimaux c'est-à-dire qui anticipent les interactions sociales et la collaboration au sein du groupe de formation.

2 Méthodologie

L'ensemble de données concerne 53 groupes de formation organisés de 2015 à 2018 incluant 387 observations[1]. La collecte des données s'effectue à travers un questionnaire dont la conception est basée sur un modèle simplifié de l'évaluation de Kirkpatrick [3], [4]. Le premier niveau lié à l'évaluation des réactions des participants à l'issue de la formation regroupe au total 20 questions composées principalement de questions fermées selon une échelle de Likert [5] sur cinq points[2]. Le design du questionnaire est classé selon les sections suivantes : identité professionnelle, satisfaction globale liée au contenu pédagogique, à l'organisation de la formation, le transfert en situation de travail et la satisfaction de collaboration lié l'animation du formateur et les interactions au sein du groupe. Les questionnaires individuels sont administrés en fin de sessions de formation et sur place. Par conséquent, l'aspect temporel de nature immédiate de l'administration du questionnaire peut induire une perception subjective et approximative. De plus, la non-anonymisation peut également constituer un biais d'évaluation qui peut freiner les apprenants à exposer ouvertement leurs déceptions et induire ainsi des réponses positives. Les méthodes d'analyses sont choisis en fonction de la nature de la variable à analyser : les variables quantitatives sont traités par le test de Student (appelé également test t) et la régression linéaire simple tandis que les variables catégorielles (appelés également variables qualitatives) sont traités par l'analyse de la variance[3] et le test du Chi2. Les conclusions des tests utilisés sont fondées sur une valeur significative (valeur de p[4] appelé également degré de signification).

3 Résultats

3.1 Statistique descriptive

La répartition des profils est composé selon les descriptions suivantes : une forte majorité de profils ayant le statut de «Cadre» (81,7%) composé d'environ deux tiers

[1] Significatif pour des estimations avec un taux de confiance suffisant
[2] 0-Pas du tout satisfait, 1-Insatisfait, 2-Neutre, 3-Plutôt satisfait et 4-Très satisfait.
[3] ANOVA (ANalysis Of Variance) et ANCOVA (ANalysis of COVariance)
[4] Valeur limite inférieure à 5%, est communément admise comme étant significative

de femmes (62%) et un tiers d'hommes (37%). La moyenne d'âge est d'environ 40 ans pour une ancienneté en moyenne entre 6 et 7 ans (38,7%). Nous notons également que le profil de l'apprenant diffère selon le mode organisationnel. Les groupes de formation organisés selon un mode intra-organisationnel[5] sont essentiellement composés de profils d'apprenants ayant le statut de « Non-Cadre » (78%). Cette tendance est significativement différente pour les groupes de formation organisés selon un mode inter-organisationnel[6], qui sont essentiellement composés de profil ayant le statut « Cadre » (62%). La moitié des apprenants (50,4%) s'engagent dans une action de formation selon leur propre initiative tandis que la présence pour 25,6% des participants[7], relève d'une exigence hiérarchique. L'objectif principal de la formation s'inscrit dans une démarche volontariste dont l'objectif pédagogique est : (1) de se perfectionner et ou d'acquérir des nouvelles compétences (62,5%) et (2) de s'adapter à une évolution réglementaire (53,6%). Les groupes sont principalement restreints selon la définition de [6] soit moins de 7 participants (60,4%), mixte (environ 83%)[8] et organisé selon un mode inter-organisationnel (62,3%). Le cycle de vie du groupe est dépendant de la durée des sessions de formations qui se déroule principalement sur une journée (52%)[9]. Compte tenu que la majorité des apprenants sont des femmes, les groupes dont le nombre de femmes est supérieurs aux hommes sont majoritaires (35,8%).

3.2 Analyse statistique

Les premiers résultats font apparaître une satisfaction globale nettement positive pour l'ensemble des formations suivies. En effet, la majorité est plutôt satisfaite (51,9%) voire très satisfaite (29,2%). Ce score axé sur la satisfaction globale est à nuancer car nous notons que le score de la satisfaction axée sur la collaboration relève des résultats négatifs (environ 10% se déclare peu satisfait).
Les facteurs pédagogiques liés à la qualité d'animation du formateur et le contenu du support pédagogique influencent très significativement la satisfaction globale et la satisfaction de la collaboration ($p<0,0001$). Néanmoins, il apparaît que les facteurs intrinsèques du groupe de formation soit les critères d'ordre organisationnels (taille, mixité du groupe, durée de la formation, nature de la composition du groupe et le mode organisationnel) sont significatifs au niveau de la satisfaction globale mais diminuent fortement sur la satisfaction de la collaboration, synthétisé dans le tableau suivant (Table 1) :

Table 1. Tableau comparatif des analyses bi-variées et la significativité des résultats de la satisfaction globale et la satisfaction de la collaboration selon la p-value associée

[5] Groupe d'apprenant issu d'une même organisation
[6] Groupe d'apprenant d'organisation distincte
[7] 24% ne l'ont pas précisé
[8] 17% des groupes sont non-mixte (13,2% sont composé exclusivement de femmes et 4% sont composés exclusivement d'hommes)
[9] 34% des sessions de formations ont une durée de deux jours et 17%ont une durée de trois jours

		Variable explicative	
		Satisfaction globale	Satisfaction de la collaboration
		Signification (p-value)	
Variable discriminante	Taille	<0,0001	0,23
	Mixité	<0,0001	0,15
	Nature de composition (Homogène/Hétérogène)	<0,0001	0,123
	Durée	0,02	0,38
	Mode organisationnel (inter/intra)	0,001	0,549

4 Conclusion et perspectives

L'évaluation de la satisfaction des apprenants démontre que le rôle du formateur, caractérisé par la qualité de l'animation, est un critère essentiel mais reste corrélé selon des multiples critères liés aux caractéristiques intrinsèques du groupe de formation. Parmi les observations étudiées, la nature de composition du groupe de nature homogène influence favorablement le score de la satisfaction globale[11] et augmente en fonction de la durée de la formation. Concernant les variables liés à la taille du groupe et le mode organisationnel, les moyennes estimées sont équilibrées et par conséquent, aucunes modalités ne se démarquent nettement permettant d'affirmer un critère prépondérant. Cette analyse vise à être approfondie pour identifier les facteurs qui résultent un score d'insatisfaction parmi l'ensemble des groupes d'observations. En complément, une analyse qualitative vise à traiter les réponses des questions ouvertes sur les points forts et faibles de la formation afin de confirmer ces premiers constats.

L'évaluation du score de la satisfaction globale des apprenants nous permet d'éliciter les critères de décision pour la modélisation d'un système d'aide à la composition des groupes. L'analyse statistique par sa nature longitudinale détermine ces critères selon des résultats justifiés en évitant ainsi des généralisations abusives.

Références

1. McCain, D.V : Evaluation Basics. ASTD Press, Alexandria. (2005).
2. Hadji C. : Faut-il avoir peur de l'évaluation ? Bruxelles : De Boeck. (2012).
3. Kirkpatrick, D.L. : Techniques for evaluating training programs. Journal of American Society for Training and Development 13, 9.
4. Kirkpatrick, D.L. : The four level of evaluation. Info Line 0701, 1–16, (2007).
5. Likert, R. : A technique for the measurement of attitudes. Archives of psychology, (1932).
6. Anzieu, D., & Martin, J. Y. : La dynamique des groupes restreints. Presses universitaires de France, (1971).

[11] Moyenne de 3,2

Conception d'un outil d'aide à la compréhension écrite pour les apprenants de français langue étrangère

Yuming Zhai, Gabriel Illouz et Anne Vilnat

LIMSI-CNRS, Univ. Paris-Sud, Univ. Paris-Saclay, France
{prénom.nom}@limsi.fr

Résumé. En vue d'aider les apprenants de français langue étrangère en compréhension écrite, nous concevons un outil qui propose des réécritures en contexte pour un mot ou une suite de mots. Nous nous appuierons sur les équivalences de traduction pour proposer ces réécritures, et un système de classification automatique des procédés de traduction au niveau sous-phrastique permettra de contrôler et d'interpréter le processus.

Mots-clés. TAL en EIAH, big data pour l'éducation, dispositif d'entraînement

Abstract. In order to help learners of French as a foreign language in reading comprehension, we design a tool which propose rewritings in context for a word or a sequence of words. We will rely on translation equivalences to propose these rewritings, and an automatic classification of sub-sentential translation processes will allow us to control and to interpret the searching process.

Keywords. NLP in Intelligent Learning Environment, big data for education, training device

1 Introduction

L'acquisition des langues étrangères est importante pour les étudiants, surtout quand ils veulent poursuivre des études dans un pays étranger. Au cours de l'apprentissage d'une langue étrangère, avoir recours à sa langue maternelle ou à une autre langue plus proche et déjà apprise est une pratique courante [9]. La traduction est ainsi la méthode la plus utilisée pour mieux assimiler et comparer les connaissances sur la langue. Une autre stratégie consiste en la reformulation paraphrastique [8,10,5]. Une telle compétence aide les apprenants à élargir leur vocabulaire et le répertoire d'expressions, à prendre l'habitude de réfléchir dans la langue étrangère au lieu de la langue maternelle. Elle est aussi importante pour la production et la compréhension écrite : pour simplifier ou rendre plus complexe leurs énoncés [2,3] ; pour transformer le langage du texte en ses propres termes, ce qui est crucial pour vérifier leur compréhension et la lier avec leurs connaissances préalables en vue de faire des inférences [6].

Compte tenu des recherches précédentes, nous proposons la conception d'un nouvel outil pour aider les apprenants (*a priori* chinois) de Français Langue Étrangère (FLE) en compréhension écrite, via la réécriture des mots ou des suites de mots en contexte. Nous ciblons les apprenants chinois, pour tester notre hypothèse que l'apprentissage d'une troisième langue étrangère (français) peut bénéficier de l'apprentissage d'une deuxième langue étrangère similaire (anglais). [1] Étant donné ce but précis, les réécritures proposées doivent être contrôlées et de bonne qualité.

1. Une majorité des étudiants chinois apprend l'anglais depuis l'école primaire.

L'extraction de réécritures suivra l'approche d'extraction des paraphrases par des traductions communes dans des corpus parallèles bilingues [1]. Le développement de cet outil sera aussi un cadre d'application possible de nos travaux de recherche en détection automatique de procédés de traduction au niveau sous-phrastique [12,13]. Cela distingue les traductions littérales de celles non littérales [11,4], en les étiquetant comme par exemple *Généralisation, Particularisation, Modulation Sémantique, etc*. Nous intégrons la classification automatique de procédés de traduction pour avoir un meilleur contrôle sémantique pendant la recherche de réécritures, en vue de faciliter la compréhension des justifications du système et de comparer les connaissances dans des langues différentes.

2 Difficultés en compréhension écrite

Les difficultés en compréhension écrite évoluent selon le niveau de langue des apprenants. Par exemple, les débutants peuvent confondre des mots homographes, sans pouvoir distinguer leur catégorie grammaticale et leur usage en contexte. Pour les apprenants dans un niveau intermédiaire même avancé, les phénomènes qui leur posent des difficultés sont : 1) des mots inconnus des domaines généraux ; des mots déjà vus mais avec un sens inconnu ou figuré dans un contexte donné ; 2) des termes spécifiques d'un certain domaine ; 3) des expressions figées, des idiomes ; 4) des structures syntaxiques complexes, des phrases longues avec une logique complexe. Les difficultés sur la structure syntaxique dépassent la portée de notre travail. Cet outil vise à aider les apprenants à mieux comprendre le texte au niveau sous-phrastique, en proposant des réécritures en contexte.

3 Conception du prototype

Outils existants Face aux définitions parfois obscures des dictionnaires, les apprenants peuvent être découragés en tentant d'identifier le sens exact selon le contexte par eux-mêmes. *DeepL*, le concurrent de *Google Translate*, permet de varier les expressions sous-phrastiques dans la langue cible. *Linguee*, un concordancier bilingue qui souligne la traduction correspondante (mais elle pourrait être erronée), montre l'usage des segments en contexte. *Writefull* exploite les informations de fréquence à partir de larges bases de données textuelles, permettant aux utilisateurs de vérifier l'usage de segments en production écrite grâce à ses multiples fonctions. *Rewordify* simplifie automatiquement un texte anglais en enlevant les ambiguïtés, pour faciliter la lecture aux apprenants débutants ou intermédiaires.

Notre conception De notre côté, nous visons à proposer des réécritures françaises en contexte en nous basant sur la désambiguïsation sémantique, la traduction et l'alignement de mot automatique [1,7]. Nous nous appuierons sur des traductions pivots anglaises et chinoises pour les apprenants chinois. Nous utiliserons aussi l'information sur les procédés de traduction au niveau sous-phrastique pour contrôler le processus.

Notre outil sera développé sous forme d'une application web. Chaque apprenant peut s'inscrire sur le site pour sauvegarder son historique de requêtes. Nous avons besoin des informations telles que sa langue maternelle, les autres langues étrangères maî-

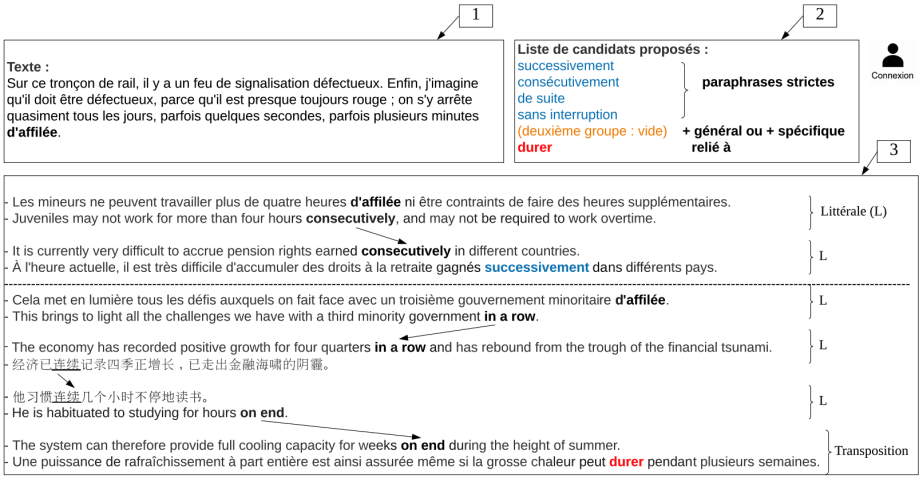


Fig. 1: Interface du prototype : réécrire en contexte pour mieux comprendre.

trisées et le niveau atteint, *etc.* pour créer son profil. L'interface la plus importante est montrée dans la figure 1. L'apprenant charge un texte dans le bloc 1 (au moins une phrase complète), et à chaque fois il sélectionne un mot ou une suite de mots qui lui pose des difficultés de compréhension. Le système récupère la phrase entière où cette entrée cherchée apparaît, afin de proposer une liste de candidats selon le contexte, affichés dans le bloc 2.

Les réécritures seront classées et contenues dans trois couleurs différentes : 1) réécritures avec une équivalence sémantique stricte, à savoir des paraphrases, et la phrase après la substitution reste grammaticale ; 2) réécritures en relation d'implication (plus générale ou plus spécifique), si elles existent ; 3) réécritures qui sont reliées avec le mot ou la suite de mots d'une certaine manière (*e.g.* ils appartiennent au même champ sémantique ; il faut adapter la structure de la phrase pour la substitution).

L'apprenant peut cliquer sur chaque candidat proposé, et le bloc 3 s'affichera pour justifier la réécriture. Ici l'entrée cherchée est *d'affilée*, et nous montrons le processus pour trouver le candidat *successivement* et *durer*, respectivement. Pour chaque paire de traduction trouvée, nous appliquons une classification automatique pour informer si elle est littérale ou non littérale. C'est le chemin avec chaque étape typée jusqu'au candidat qui permet d'interpréter et de classer les résultats. Par exemple, la locution adverbiale *on end* est traduite par un verbe français *durer* dans la dernière paire de phrases. La classe grammaticale change avec le procédé de traduction *Transposition*, ainsi ce candidat n'est pas une paraphrase stricte en contexte. L'apprenant peut décider de n'afficher que les paires de traduction non littérale.

Nous garderons les traces d'utilisation des apprenants pour étudier la façon dont ils apprennent, voir quel procédé de traduction est plus utile pour aider la compréhension, *etc.* À long terme, nous souhaitons que cet outil puisse aussi aider les apprenants à maîtriser la traduction selon les différents procédés de traduction.

4 Conclusion

Nous proposons la conception d'un nouvel outil pour aider les apprenants de FLE en compréhension écrite, via la proposition de réécritures sous-phrastiques en contexte. L'extraction de réécritures s'appuiera sur des traductions communes dans des corpus parallèles bilingues, et nous intégrerons la classification automatique de procédés de traduction pour classer les candidats et rendre la recherche plus interprétable. Nous planifions une évaluation de notre outil à la fois quantitative avec des tests en compréhension écrite disponibles, pour observer si les apprenants obtiennent de meilleures notes en utilisant l'outil, et qualitative en classe avec la supervision d'un enseignant qui peut proposer ses propres tâches pédagogiques en utilisant ce site.

Références

1. Bannard, C., Callison-Burch, C. : Paraphrasing with bilingual parallel corpora. In : Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. pp. 597–604. Association for Computational Linguistics (2005)
2. Chachu, S. : The intermediary role of reformulation in learning French as a foreign language : The case of students at the University of Ghana. SHS Web of Conferences **38**, 00007 (01 2017). https://doi.org/10.1051/shsconf/20173800007
3. Chen, M.H., Huang, S.T., Chang, J., Liou, H.C. : Developing a corpus-based paraphrase tool to improve EFL learners' writing skills. Computer Assisted Language Learning (2013). https://doi.org/10.1080/09588221.2013.783873
4. Chuquet, H., Paillard, M. : Approche linguistique des problèmes de traduction anglais-français. Ophrys (1989)
5. Eshkol-Taravella, I., Grabar, N. : Repérage et analyse de la reformulation paraphrastique dans les corpus oraux. In : TALN2014. pp. 304–315 (2014)
6. Kletzien, S.B. : Paraphrasing : An effective comprehension strategy. The Reading Teacher **63**(1), 73–77 (2009)
7. Kok, S., Brockett, C. : Hitting the Right Paraphrases in Good Time. In : Human Language Technologies : Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 2-4, 2010, Los Angeles, California, USA. pp. 145–153 (2010)
8. Martinot, C. : La reformulation : de la construction du sens à la construction des apprentissages en langue et sur la langue. Corela. Cognition, représentation, langage (HS-18) (2015)
9. Nation, P. : The role of the first language in foreign language learning. Asian EFL **5**(2), 1–8 (2003)
10. Rossari, C. : Les opérations de reformulation : analyse du processus et des marques dans une perspective contrastive français-italien. Sciences pour la communication, Peter Lang (1994)
11. Vinay, J.P., Darbelnet, J. : Stylistique comparée du français et de l'anglais : méthode de traduction. Bibliothèque de stylistique comparée, Didier (1958)
12. Zhai, Y., Max, A., Vilnat, A. : Construction of a Multilingual Corpus Annotated with Translation Relations. In : First Workshop on Linguistic Resources for Natural Language Processing. pp. 102–111 (2018)
13. Zhai, Y., Safari, P., Illouz, G., Allauzen, A., Vilnat, A. : Towards Recognizing Phrase Translation Processes : Experiments on English-French. In : (Preprint version) 20th International Conference on Computational Linguistics and Intelligent Text Processing (2019)

Soutien à la régulation cognitive et émotionnelle de l'apprentissage à distance : Intégration d'outils au design pédagogique d'un cours hybride

Maud Bernies[1], Gaëlle Molinari[2]

[1] Université de Genève, TECFA, Boulevard du Pont d'Arve 40, CH-1211 Genève 4, Suisse
maud.bernies@etu.unige.ch
[2] UniDistance, Überlandstrasse 12, Postfach 265, CH-3900 Brig, Suisse
gaelle.Molinari@unige.ch

Résumé. Comparée à la formation présentielle, la formation à distance requiert une prise de contrôle plus importante de l'apprentissage. La réussite dans ce type de formation nécessite la mobilisation de stratégies d'autorégulation efficaces sur toutes les dimensions de l'apprentissage (cognitive, affective, comportementale et contextuelle). Cet article présente les outils de soutien à la régulation développés dans le cadre d'un cours hybride. Nos choix de conception ont été guidés par trois objectifs : favoriser l'engagement en soutenant l'auto-observation et la co-régulation (l'entraide entre étudiants), ce par l'utilisation d'outils de reporting et de visualisation partagée ; intégrer ces outils à chaque étape du scénario d'apprentissage dans la plateforme utilisée, Moodle ; et encourager les étudiants à rendre compte à la fois de leurs états cognitifs et émotionnels. Une analyse d'usages et de perceptions par les étudiants de ces outils sera menée et servira à l'amélioration du dispositif.

Mots-clés. Formation à distance/hybride ; auto- et co-régulation ; auto-observation ; visualisation partagée ; états cognitifs et émotionnels.

Abstract. Compared to face-to-face learning, distance learning is recognized as requiring a higher degree of control over the learning process. Success in distance learning is dependent upon the use by students of effective self- regulated learning strategies on all dimensions of learning (cognitive, affective, behavioral and contextual). This paper describes the support system we developed for learner's self-regulated learning in a blended learning environment. Our design choices were guided by three goals: to foster engagement by supporting self-observation and co-regulation (through mutual help between students) using self-reporting and shared visualization tools; to embed these tools to the learning design into the platform used, Moodle; and to encourage learners to self-report both their cognitive and emotional states. An analysis of students' use and perceptions of the self- reporting and visualization tools will be conducted with the aim of refining the support system.

Keywords. Distance/blended learning; self- and co-regulation; self-observation; shared visualizations; cognitive and emotional states.

La formation à distance se caractérise par une demande d'autonomie élevée [1]. Elle peut susciter un sentiment d'isolement qui, couplé à un faible sentiment d'appartenance, peut augmenter la probabilité d'abandonner [2]. Un tel contexte contraint les apprenants à déployer des stratégies d'autorégulation et volitionnelles pour s'engager, prendre contrôle de leur apprentissage, et se maintenir au travail [3].

L'autorégulation s'opère sur toutes les dimensions de l'apprentissage, cognitive, motivationnelle, émotionnelle, comportementale et environnementale [2], [4]. Selon [3], l'auto-observation est l'une des conditions de l'autorégulation. Il s'agit d'un processus de monitorage de l'activité au cours duquel l'apprenant prend conscience et évalue ses états cognitifs, émotionnels et motivationnels. L'apprenant peut réaliser des jugements cognitifs sur son niveau de compréhension ou d'apprentissage. Il peut identifier ses émotions en lien avec l'activité, et les utiliser pour faire un diagnostic de son niveau d'engagement et de progression dans la tâche [5].

L'observation de soi est un processus complexe qu'il est nécessaire d'instrumenter par le biais de dispositifs techniques permettant aux apprenants de créer et mettre à jour une représentation la plus fidèle possible de leur activité [3]. Dans le contexte des EIAH, ces dispositifs renvoient aux outils d'*awareness* qui proposent aux apprenants une visualisation de leurs données auto-rapportées et des traces de leur activité (et/ou de celles de leurs pairs). Ces outils visent à soutenir la régulation "en vol" et à encourager l'activité réflexive ainsi que l'entraide [6]. DDART [7] est un exemple d'EIAH conçu pour encourager la régulation dans les situations d'apprentissage en mode projet. Sa particularité est d'offrir aux apprenants la possibilité de construire leurs propres indicateurs à partir de traces automatiques et déclaratives. DDART se réfère également à la taxonomie des indicateurs proposée par [8] pour permettre une régulation des différentes dimensions de l'apprentissage. Il intègre ainsi des indicateurs cognitifs (niveaux de connaissances et résultats d'apprentissage), affectifs (états émotionnels) et sociaux (fréquence des interactions entre apprenants).

Notre étude se situe dans le prolongement du projet EmoViz dont l'un des buts était d'offrir aux étudiants à distance des outils de support à la prise de conscience et la régulation de leurs émotions [9]. Les outils de soutien à la régulation présentés ici, ont été conçus sur la base de 2 résultats relatifs à l'analyse d'usage d'EMORE-L, un outil pour le report quotidien des émotions vécues lors des activités d'apprentissage à distance : le besoin des étudiants de rattacher leurs émotions à un contexte, et celui de partager leurs émotions avec leurs pairs. Le besoin de contexte nous ont amenées à proposer des outils qui incitent les apprenants à reporter non seulement leurs émotions mais également leurs sentiments de compréhension au regard de la tâche, le but étant de les aider à faire des liens entre leurs états émotionnels et cognitifs. Par ailleurs, les étudiants sont sollicités à faire des auto-évaluations cognitives et émotionnelles à chaque étape du scénario d'apprentissage. Enfin, en réponse au besoin de partage, les informations émotionnelles et cognitives collectées font l'objet de visualisations partagées afin d'encourager la co-régulation et l'entraide entre pairs [10].

Le Tableau 1 montre comment notre dispositif de soutien à la régulation a été intégré au design pédagogique d'une des 5 périodes (de 3 semaines) à distance d'un cours hybride sur la psychologie de l'éducation (37 étudiants, niveau Bachelor). Il illustre également la façon dont nos choix techno-pédagogiques ont été guidés par notre cadre théorique : favoriser l'observation de soi et des autres sur différentes dimensions de l'apprentissage (cognitive, émotionnelle et volitionnelle); encourager la co-régulation et l'entraide. Les outils de reporting et de visualisation partagée sont des outils de la plateforme Moodle (sondages, questionnaires, liens URL) qui héberge le cours.

Tableau 1. Période à distance 1 : Scénario (phases, activités), reporting et visualisation.

Scénario (phases, activités)		Reporting	Visualisation partagée
Ce qui est attendu	• Document des tâches • Guide de lecture	Ma compréhension des attentes **(auto-observation cognitive)**	Les sentiments de compréhension de la promo **(observation cognitive de soi et des autres)**
		Mon baromètre émotionnel par rapport aux attentes **(auto-observation émotionnelle)**	Le baromètre émotionnel de la promo **(observation émotionnelle de soi et des autres)**
Etude des ressources	• Vienneau • Hubermann • Zampa	Ma position par rapport à la compréhension des lectures **(auto-observation cognitive)**	Les positions de la promo **(observation cognitive de soi et des autres)**
		Mon baromètre émotionnel par rapport aux-ressources **(auto-observation émotionnelle)**	Le baromètre émotionnel de la promo **(observation émotionnelle de soi et des autres)**
Productions	• Résumés des lectures (forum) • Devoir 1 évalué par le pair (atelier) • Débat collectif	Share my feelings **(auto-observation cognitive, volitionnelle, émotionnelle)**	Les "share my feelings" dans la promo **(observation cognitive, volitionnelle, émotionnelle de soi et des autres)**
Echanges et entraide	• Echanges sur la période 1 (forum)	Ma disponibilité à aider ou mon besoin d'aide **(co-régulation, entraide)**	Les disponibilités à aider et besoins d'aide de la promo **(co-régulation, entraide)**

Dans la phase "Ce qui est attendu", les étudiants évaluent leur compréhension des objectifs de la période (cocher 1 item parmi 4 ; e.g., "je ne comprends pas du tout ce qui est attendu de moi") puis choisissent leur émotion dominante au regard des objectifs parmi 6 émoticônes (baromètre émotionnel) : confiant(e), enthousiate, confus(e),

anxieux(se), énervé(e), découragé(e). Dans la phase "Etude des ressources", ils évaluent leurs difficultés à comprendre les lectures proposées, informent l'état d'avancement de cette activité (cocher 1 item parmi 5 ; e.g., "j'ai terminé la lecture et je n'ai pas rencontré de difficultés particulières") puis complètent le baromètre émotionnel. Après le rendu des "productions", les étudiants répondent au questionnaire "Share my feelings" : ils évaluent l'effort mental et la difficulté perçus pour chaque production (items 1 et 2) ; expriment les émotions qu'ils ressentent à la fin de la période (items 3 à 8) ; évaluent leurs difficultés à se mettre au travail et persévérer (item 9) ; indiquent dans quelle mesure ils ont le sentiment d'avoir appris (item 10). Enfin, le sondage de la section "Echanges et entraide" invite les étudiants à indiquer leur disponibilité pour offrir de l'aide et/ou leur besoin d'être aidé(e) à tout moment au cours de la période. Après avoir reporté leurs réponses aux sondages et questionnaires, les étudiants ont la possibilité de consulter celles de leurs pairs via les visualisations partagées (liens URL). Les réponses sont anonymes, à l'exception de celles du sondage d'entraide.

References

1. Cosnefroy, L.: Autonomie et formation à distance. *Recherche et formation*, (69), (2012) 111-118
2. Jézégou, A. Créer de la présence à distance en e-learning. *Distances et savoirs*, *8*(2), (2010) 257-274.
3. Cosnefroy, L.: L'apprentissage autorégulé: perspectives en formation d'adultes. *Savoirs*, (2), (2010) 9-50.
4. Pintrich, P. R.: A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational psychology review*, *16*(4), (2004) 385-407.
5. Lavoué, É., Molinari, G., Prié, Y., & Khezami, S.: Reflection-in-action markers for reflection-on-action in Computer-Supported Collaborative Learning settings. *Computers & Education*, *88*, (2015) 129-142.
6. George, S., Michel, C., & Ollagnier-Beldame, M.: Usages réflexifs des traces dans les environnements informatiques pour l'apprentissage humain. *Intellectica*, *59*(1), (2013) 205-241.
7. Michel, C., Lavoué, E., George, S., & Ji, M.: Supporting Awareness and Self-Regula-tion In Project-Based Learning through Personalized Dashboards. *International Journal of Technology Enhanced Learning*, *9*(2/3), 204-226.
8. Dimitracopoulou, A., & Bruillard, É.: Enrichir les interfaces de forums par la visualisation d'analyses automatiques des interactions et du contenu. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation, 13*, (2006)
9. Molinari, G., Trannois, M., Tabard, A., & Lavoué, E.: October). EMORE-L: un outil de reporting des émotions pour l'apprentissage à distance. In *Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine* (2016) 167-176.
10. Temperman, G., De Lièvre, B., & Depover, C.: D'un outil d'awareness à un outil d'encadrement de l'apprentissage: Que nous disent les apprenants du tableau de bord? *Revue Internationale des Technologies en Pédagogie Universitaire*, *4*(1), (2007) 38-46.

La collaboration dans une recherche orientée par la conception : l'exemple de PLAY

Elsa Paukovics

Université de Fribourg - LIP, Rue P.-A. de Faucigny 2, 1700 Fribourg, Suisse
elsa.paukovics@unifr.com

Résumé. La recherche orientée par la conception est une méthodologie de recherche qui permet la collaboration entre différentes communautés sur la conception et l'expérimentation de dispositifs techno-pédagogiques en articulant des objectifs pratiques et scientifiques. Le projet PLAY s'insère dans cette démarche à travers la conception d'un jeu dédié aux visites de classe au musée de la nature à Sion (Suisse). Cette contribution vise à souligner et questionner les éléments liés à la co-construction des savoirs, caractéristique d'une recherche orientée par la conception, et à discuter de la collaboration entre les différents acteurs avec le concept d'objet frontière.

Mots-clés. Recherche orientée par la conception, collaboration, objet frontière, learning game, dispositif techno-pédagogique

Abstract. Design-based research is a methodological approach that allows collaboration between diverse communities of expertise by articulating both practical and scientific goals. Based on this approach, PLAY aims to design and develop a learning game dedicated to classroom visits in the Museum of Nature in Sion (Switzerland). This contribution highlights and questions the co-construction of knowledge in PLAY. The context is analyzed on the basis of the characteristics of design-based research, then collaboration will be discussed with the concept of boundary object.

Keywords. Design-based research, collaboration, boundary object, learning game, technology-enhanced learning environments

1 Introduction

La recherche en éducation en Suisse romande tend à s'orienter dans des démarches collaboratives entre praticiens et chercheurs, permettant ainsi le rapprochement entre recherche et besoins du terrain [1]. PLAY est une *recherche orientée par la conception* (RoC)[2] mise en place pour répondre à la demande concrète du musée de la Nature de Sion : accompagner les visites de classe à travers un dispositif techno-pédagogique innovant. Le jeu pédagogique, conceptualisé pour répondre à cette demande, est en cours de développement et repose sur des technologies numériques notamment dans l'utilisation d'une plateforme informatique, de la réalité augmentée et dans le traitement des traces numériques. Le développement, la réalisation et l'analyse de cet

environnement informatique pour l'apprentissage humain (EIAH) implique la collaboration du musée de la Nature avec le laboratoire d'innovation pédagogique (LIP), la Haute école pédagogique du Valais ainsi que de la Haute école spécialisée de Suisse occidentale. Par ailleurs, didacticiens, enseignant, spécialiste de l'apprentissage par le jeu, informaticien et personnel du musée sont amenés à travailler ensemble. Toutefois, la co-construction des savoirs qui découle de cette collaboration entre différentes communautés ne va pas de soi [3]. Il est nécessaire de se questionner sur cette collaboration et les éléments qui la caractérisent afin d'optimiser les dispositifs de conception et de recherche.

2 Recherche orientée par la conception et objet frontière

Un nouveau paradigme méthodologique se développe sous la forme de recherches collaboratives dans le champ des sciences de l'éducation en encourageant les recherches non pas *sur*, mais *avec* les praticiens [4]. La RoC [2] ou *design based research* [5] vise le développement et l'implantation d'EIAH en mettant la conception au service de la recherche. Les chercheurs initient ou rejoignent des projets innovants afin d'éprouver des modèles théoriques en conditions écologiques. Les objectifs de recherche s'articulent avec les objectifs du terrain à partir de la collaboration entre différentes communautés (chercheurs, designers, informaticiens, enseignants). La pratique est ainsi considérée comme un moyen au service de la recherche [2] ; les situations professionnelles sont problématisées et les expérimentations permettent de faire évoluer les modèles théoriques. En plus de son caractère contributif, la RoC repose sur un processus itératif [5]. La conception d'EIAH se trouve souvent au cœur des recherches orientées par la conception, reposant sur des savoirs scientifiques [6], didactiques et informatiques. La rencontre entre ces communautés soulève des questions d'ordres épistémologique et méthodologique.

Le concept d'objet frontière permet de problématiser la co-construction du savoir dans une RoC [7]. Selon Leigh Star [8] « les objets frontières sont un arrangement qui permet à différents groupes de travailler ensemble sans consensus au préalable » (p.19). Ancrés dans le temps, les actions et les réflexions, ces objets doivent être à la fois flexibles, afin de présenter les caractéristiques spécifiques à chaque communauté en conservant leur autonomie, mais aussi être fixes en favorisant les échanges et en permettant la construction d'éléments communs entre les communautés [9]. Les contours d'un objet frontière évoluent en fonction des actions et des réflexions de chacun. Il permet à des communautés distinctes de collaborer en partageant des similitudes et en soulignant des différences [9]. Ainsi, des objectifs communs s'articulent afin que chacun puisse travailler sur cet objet dans son champ d'activités. La forme, l'usage et les représentations soumises à l'objet différent selon la personne qui le considère, lui inférant une *flexibilité interprétative* [8]. La compréhension de cet outil repose sur la mise en place d'outils et de procédures communes [9]. Les EIAH, en tant qu'artefacts co-construits par différentes communautés, constituent des candidats au statut d'objets frontière.

A la lumière de ces apports théoriques, nous formulons les questions suivantes : (1) quels sont les éléments du projet qui relèvent d'une recherche orientée par la

conception ? (2) Dans quelle mesure l'EIAH développé représente-t-il un objet frontière ?

3 Méthode et analyse préalable

PLAY fait actuellement l'objet d'une étude de cas menée dans le cadre d'une thèse de doctorat qui regroupe des données sur les questions méthodologiques du projet. La méthode de l'étude de cas permet de combiner plusieurs techniques de collecte de données afin de saisir un phénomène qui est partie prenante du contexte lui-même [10]. L'objectif de la thèse est d'analyser la co-construction des savoirs dans le cadre d'une recherche collaborative au regard de la posture épistémique des acteurs impliqués. Ces données, collectées par observation participante [10], sont actuellement en cours d'analyse. Sur la base de ces données (enregistrements audio et vidéo d'une phase test du jeu et de trois séances de travail de conception/adaptation du jeu menées entre novembre 2017-2018), nous proposons, dans cette contribution, une brève analyse *a priori* du dispositif de recherche afin d'esquisser quelques pistes de réponse à nos questions.

(1) Les objectifs du projet sont doubles ; à la fois scientifiques et pragmatiques [1 ; 2]. Engagé à la demande des acteurs du terrain pour répondre à un besoin d'ordre pratique (accompagner les visites de classe au musée et sensibiliser les élèves au concept d'anthropocène), le jeu Pearl'Arbor est co-conçu par des praticiens et des chercheurs. Mettant la conception au service de la recherche, les objectifs scientifiques ont été définis conjointement aux objectifs de terrain afin de comprendre les effets du jeu sur les interactions entre les élèves et la muséographie. L'analyse scientifique des effets permet de comprendre les mécanismes du jeu et de proposer des pistes de réingénierie [2] à travers deux phases itératives [5] et trois différentes versions du jeu (la version 3 est en cours de développement). Ces réingénieries du prototype de Pearl'Arbor reposent sur la collaboration entre : le directeur du musée, une médiatrice, un informaticien, une didacticienne/enseignante, un chercheur sur les jeux dédiés à l'apprentissage et une doctorante. Bien que les statuts des différents acteurs puissent évoluer selon les circonstances, nous distinguons quatre communautés : celle des praticiens du musée, celle des didacticiens, celle des spécialistes en jeu de l'apprentissage et celle des informaticiens. Ces communautés différentes sont ainsi amenées à collaborer sur et dans l'EIAH co-construit.

(2) Le jeu Pearl'Arbor représente un objet commun élaboré en prenant en compte les savoirs spécifiques à chaque communauté engagée dans la recherche. Les contenus scientifiques (liés à l'anthropocène) sont apportés par le personnel du musée, puis, en collaboration avec les didacticiens et les spécialistes du jeu de l'apprentissage, ces *savoirs savants* sont transposés en *savoirs à enseigner* [11] à travers un dispositif ludique. La communauté informatique intègre ces savoirs à enseigner dans le jeu à partir des possibilités informatiques. Le jeu présente donc une *flexibilité interprétative*, c'est à dire qu'il est soumis, dans le groupe de travail, à différentes représentations, usages et valeurs [8]. Cette flexibilité est caractéristique de la co-construction d'un objet commun au sein de plusieurs communautés, et serait aussi issue des expériences personnelles des acteurs. En effet, les parcours et expériences individuelles semblent

jouer un rôle dans la construction de cet artefact que nous considérons donc être un objet frontière. Nous observons un processus de conception partagé par tous les membres du projet. Les frontières de Pearl'Arbor évoluent selon les apports de chaque communauté à partir des éléments apportés par chacun des acteurs. Par exemple, lorsque l'informaticien ajoute une fonctionnalité dans le jeu, tous les acteurs du projet ne sont pas capables de la comprendre pourtant cette action se répercute sur l'ensemble du jeu, notamment sur ses composantes didactiques dans l'ajout de feedbacks.

Cette contribution présente PLAY comme une RoC puisque le projet vise la conception d'un dispositif techno-pédagogique en conditions écologiques à travers des cycles itératifs et qu'il articule des objectifs à la fois pragmatiques et scientifiques [2]. Il parait raisonnable de considérer Pearl Arbor comme un objet frontière qui permet d'articuler et de valoriser les apports propres à chacune des communautés impliquées dans le développement du jeu. Pour poursuivre ces réflexions, nos perspectives de recherche dans le cadre de la thèse de doctorat qui est menée reposent sur le rapport personnel des acteurs aux savoirs co-construits dans une RoC.

Références

1. Sanchez, E., Müller, S., Kramar, N., Widmer, A., & Paukovics, E.: Students' Conducts during a Digital Game-Based Museum School Visit. In R. Bottino & D. Passey (Eds.), Empowering Learners for Life in the Digital Year, proceedings of the OCCE 2017: Springer (forthcoming).
2. Sanchez, E., Monod-Ansaldi, R.: Recherche collaborative orientée par la conception Un paradigme méthodologique pour prendre en compte la complexité des situations d'enseignement-apprentissage. Education et didactique. 9, 73–94 (2015).
3. Audoux, C., Gillet, A.: Recherche partenariale et co-construction de savoirs entre chercheurs et acteurs : l'épreuve de la traduction. Revue Interventions économiques. 43, 1–8 (2011).
4. Desgagné, S., Bednarz, N., Lebuis, P., Poirier, L., Couture, C.: L'approche collaborative de recherche en éducation: un rapport nouveau à établir entre recherche et formation. Revue des sciences de l'éducation. 27, 33 (2001).
5. Kennedy-Clark, S.: Research by design: Design-based research and the higher degree research student. Journal of Learning Design. 8, 108–122 (2015).
6. Wang, F., Hannafin, M.J.: Design-based research and technology-enhanced learning environments. Educational Technology Research and Development. 53, 5–23 (2005).
7. Arzarello, F., Robutti, O., Sabena, C., Cusi, A., Garuti, R., Malara, N., Martignone, F.: Meta-Didactical Transposition: A theorical Model for Teacher Education Programmes. In: Clark-Wilson, A., Robutti, O., and Sinclair (eds.), The Mathematics Teacher in the Digital Era. pp. 347–372. Springer, Dordrecht (2014).
8. Leigh Star, S.: Ceci n'est pas un objet-frontière !: Réflexions sur l'origine d'un concept. Revue d'anthropologie des connaissances. Vol 4, 1, 18 (2010).
9. Star, S.L., Griesemer, J.R.: Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social Studies of Science. 19, 387–420 (1989).
10. Albarello, L.: Choisir l'étude de cas comme méthode de recherche. De Boeck, Bruxelles, Belgique (2011).
11. Perrenoud, P.: La transposition didactique à partir de pratiques : des savoirs aux compétences. Revue des sciences de l'éducation. 24, 487–514 (1998).

Interactions en jeu dans les situations d'apprentissage instrumenté en Arts Visuels

Ana-Vanêssa Lucena[1]

[1] Haute Ecole Pédagogique BEJUNE, rue du 1er-Aout 44, 2300 La Chaux-de-Fonds, Suisse
AnaVanessa.Lucena@hep-bejune.ch

Résumé. La contribution que je propose ici est en lien avec mes cours de didactique en Education Visuelle à la HEP-BEJUNE en Suisse et à l'introduction actuelle du nouveau Plan d'Etude Romand. L'objectif de cette communication descriptive d'une démarche de formation est amplement exploratoire. Ma présentation s'appuiera sur des données écrites, vidéos et photos ayant servi à l'analyse de projets d'activités d'interventions scolaires menés par des étudiants stagiaires et assumera un triple intérêt : exposer le développement et les premiers résultats d'une expérience en formation en médiation culturelle instrumentée ; rendre compte du processus d'appropriation des tablettes tactiles par les étudiants et de voir la façon dont elles s'intégreraient à l'élaboration d'une activité au musée ; et enfin, soumettre cette façon de faire au débat.

Mots-clés. Dispositif d'enseignement, didactisation, outils numériques, instrumentalisation, instrumentation

Abstract. The contribution I offer to make here is linked to my course in Visual Education didactics at the HEP-BEJUNE in Switzerland and to the current introduction of the new Romand Study Plan. The objective of this descriptive talk about a teaching approach is very much exploratory. My presentation will be based on written data, videos and pictures instrumental to the analysis of teaching projects carried out by students during their professional internships. It will carry three points of interest: present the development and initial results of an instrumented cultural mediation training; report on the tablet devices adoption process by the students and on the way these integrate into the conception of a teaching activity at the museum; and finally, submit this approach to a debate.

Keywords. Teaching approach, didactisation, digital tools, instrumentalisation, instrumentation

1 Introduction

La contribution que je propose d'apporter à la 9ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain est en lien avec mes cours de didactique en Education Visuelle à la Haute Ecole Pédagogique de Berne, Jura et Neuchâtel (HEP-BEJUNE) en Suisse et à l'introduction actuelle du nouveau Plan d'Etude Romand (PER).

L'objectif de cette communication descriptive d'une démarche de formation est amplement exploratoire. Il s'appuie sur les échanges théoriques pendant les cours de didactique, sur l'analyse des activités proposées par les étudiants stagiaires à intérieur d'un Musée de Beaux-Arts ainsi que sur leurs façons de réaliser ces activités.

J'examinerai successivement l'approche de Rabardel concernant les situations d'activités instrumentées [1][2], puis j'exposerai des exemples de structuration de séquences et des réflexions par les étudiants concernant l'utilisation des tablettes numériques, étayant leurs choix des contenus et des activités en lien avec les activités médiatisées proposées aux élèves par le PER.

2 Descriptif et Analyse du Dispositif d'Enseignement

En partant du postulat que les activités artistiques à l'école ont toujours visé le développement de connaissances autant que de compétences nous débutons par saisir l'acuité de la question de l'utilisation des nouvelles technologies dans la composition de séquences d'enseignement artistique à l'école primaire. Une réflexion de fond est engagée par et avec les étudiants à propos de la didactisation de ces technologies, spécifiquement les usages didactiques potentiellement pertinents des tablettes tactiles pour l'enseignement des AV et à leurs possibilités formatrices pour les étudiants. Comment « didactiser » l'appropriation culturelle des œuvres d'arts dans le choix des contenus à enseigner et par ce biais-là produire des situations spécifiques de travail en lieu culturel par et avec les outils numériques ?

A la place de superposer les pratiques, les différentes activités artistiques nécessitent d'entretenir des rapports logiques de continuité, de complémentarité et de finalité pour être cohérentes avec la progression des apprentissages. En formation, il est demandé aux étudiants de construire une cohérence didactique entre les activités proposées aux élèves pendant leurs stages. L'idée de départ serait de se placer du point de vue de l'élève afin que celui-ci ne se situe pas, pendant les activités artistiques, dans des réalisations éclatées développant des compétences non transférables.

Trouver des situations didactiques dans lesquelles les connaissances en Éducation Visuelle sont fonctionnelles peut devenir complexe pour des futurs enseignants généralistes. Il faut arriver à concevoir une activité cohérente, créer une tâche exigeant un raisonnement dont les compétences à évaluer seraient les composantes et dont la présence serait certifiée principalement par l'agir conscient de l'élève plutôt que par une restitution explicite dans une production.

Une compétence acquise n'est pas uniquement la reproduction d'un mécanisme, d'un *savoir-refaire* [3]. Elle présume un mouvement d'adaptation et de dépassement, c'est-à-dire, d'apprentissage continu. Concevoir des dispositifs de formation en tant que situations d'apprentissage serait alors mettre en place les conditions dans lesquelles l'engagement demandé à l'étudiant l'impose à construire ou à réinvestir les compétences précises à la réussite de l'activité. C'est à partir de la fonction essentielle d'apprentissage que l'étudiant/stagiaire peut essayer de différencier les activités applicables dans une classe. Et c'est seulement parce qu'il va se placer du point de

vue de la construction des compétences et d'une construction transférable à d'autres situations, que ce type de dispositif aura un sens.

Ce type de dispositif de formation s'inscrit dans le cadre de l'interactionnisme social et de la médiation sociocognitive s'appuyant sur deux présupposés théoriques : celui de *l'origine culturelle de la pensée et de l'apprentissage* et celui de *la nécessaire médiation* [4]. L'idée étant que les tablettes, devenant l'instrument de médiation de l'activité au musée, seraient les déclencheurs des échanges entre les élèves et l'étudiant et influenceraient les transactions d'ordre sociocognitives et réaliseraient chez ces deux acteurs des changements au plan conceptuel. Mais pour arriver à cela, je suis consciente que le contenu de l'activité élaborée par l'étudiant et proposée dans la tablette doit constituer le lieu de réalisation de la pensée en dépassant la vision de cet outil comme simple médium.

La non connaissance du fonctionnement standard des outils numériques ainsi que de leurs possibilités – logiciels, applications, appareil photo, etc. – pour élaborer l'activité en médiation culturelle, empêcherait les étudiants de se concentrer sur le savoir travaillé. Si le dispositif de formation a comme souhait que les échanges entre étudiants soient primordialement sur les contenus et que les tablettes aient une place d'outil intégré, constitutif de l'activité, les objectifs sont éloignés.

Si on observe ce dispositif en rapport aux situations d'activités instrumentées [2] il est possible de comprendre le besoin des étudiants de développer une appropriation des tablettes par une utilisation plus soutenue. Rabardel met en évidence une double dimension du processus d'appropriation des outils pour pouvoir aboutir à son utilité réelle – de l'artefact à l'instrument – au service des apprentissages : une dimension d'ordre technique mais aussi psychologique [2]. Par ce concept, un artefact devient vraiment un instrument, un outil qui permet d'agir en contexte seulement si l'utilisateur, dans ce cas l'étudiant, l'intériorise en tant que tel. Ce processus d'intériorisation est appelé genèse instrumentale par Rabardel, et se compose par deux mouvements : le mouvement du sujet vers l'artefact (l'instrumentalisation) et le mouvement de l'artefact vers le sujet (l'instrumentation) [2]. Le premier consiste à la compréhension par l'étudiant de l'existence de l'artefact tablette numérique et de ses fonctions, pendant que le deuxième mouvement consiste dans la modification de ses schèmes d'actions et de pensée vis-à-vis de l'utilisation de la tablette. Orientés par les besoins de l'activité, ces deux mouvements sont réalisés par l'étudiant. Si l'instrumentalisation s'oriente vers la composante artefactuelle de l'instrument, ici toutes les fonctions d'une tablette numérique, l'instrumentation s'oriente vers les besoins spécifiques de l'étudiant lui-même dans la composition de son activité de médiation culturelle au MBA.

3 Conclusion

L'approche de Rabardel concernant les situations d'activités instrumentées permet de regarder comment l'artefact tablette numérique peut, dans des activités de formation, être incorporé aux activités de médiations culturelles dans la formation des étudiants pour acquérir le statut d'instrument (outil didactique). L'idée étant d'amener

les étudiants à cette appropriation, de faire avancer les étudiants dans cette double situation d'apprentissage (TIC-AV).

Pour le moment le retour d'expérience est essentiellement descriptif et je peux présenter les modalités d'utilisation, les situations favorables et les contraintes techniques éprouvées par les étudiants. Je n'ai, actuellement, aucunes données sur des changements induits par l'utilisation du numérique dans la médiation entre les élèves et les objets culturels de référence au musée ; uniquement sur le rapport au contenu du savoir envisagé par les étudiants dans les guides réalisés pour arriver à cette médiation.

La communication présentée s'appuiera sur des données écrites, vidéos et photos ayant servi à l'analyse de projets d'activités d'interventions scolaires menés par des étudiants stagiaires. Le fait qu'il existe une production artistique entre les deux sujets du dialogue (stagiaire et élève) a permis de revenir aux différents points de vue de la production et, ainsi, sur les différences de compréhension de l'action dont elle faisait l'objet.

La présentation de ma pratique assume ici un triple intérêt : exposer le développement et les premiers résultats d'une expérience en formation en médiation culturelle instrumentée ; rendre compte du processus d'appropriation des tablettes tactiles par les étudiants et de voir la façon dont elles s'intégreraient à l'élaboration d'une activité au musée ; et enfin, soumettre cette façon de faire au débat.

References

1. Rabardel, P.: Les hommes et les technologies : Approches cognitives des instruments contemporains. Armand Colin, Paris (1995).
2. Rabardel, P., Bourmaud, G.: Instruments et systèmes d'instruments. In P. Rabardel et P. Pastré (dir) : Modèles du sujet pour la conception – Dialectiques activités développement. p. 211-229. Octarès, Toulouse (2005).
3. Roegiers, X.: Une pédagogie de l'intégration. Compétences et intégration des acquis dans l'enseignement. de Boeck, (2000).
4. Vygotski, L. : Pensée et Langage. La Dispute, Paris (1934).

Prototype de feedback visuel des productions écrites d'apprenants francophones de l'anglais sous Moodle

Nicolas Ballier[1], Thomas Gaillat[1,2] Erin Pacquetet[1]

[1] Université de Paris, CLILLAC-ARP, 7 rue Thomas Mann, F-75013 Paris, France
nballier@free.fr, erin.pacquetet@yahoo.fr
[2] Université de Rennes 1, SCELVA - 263, av. du Gl Leclerc F-35000 Rennes, France
thomas.gaillat@univ-rennes1.fr

Résumé. Cette contribution présente une expérimentation en cours à l'université de Rennes 1 avec des étudiants d'anglais non spécialistes. L'EIAH cherche à présenter un retour quasi-immédiat aux étudiants sur leur niveau de performance dans leurs productions écrites. Nous présentons la chaîne de traitement à partir des textes extraits sous Moodle ainsi que les visualisations possibles. Nous centrons notre analyse sur les métriques produites par des analyseurs automatiques de sophistication lexicale, de complexité lexicale et syntaxique dont nous extrayons les caractéristiques les plus pertinentes.

Mots-clés. anglais L2, métriques de lisibilité, métriques de complexité lexicale, acquisition du lexique, sophistication lexicale, ICALL

Abstract: This case study illustrates the benefits of the visualisation of text metrics for French learners of EFL. Several metrics are used to provide immediate visual feedback on written productions. We present the data processing pipeline as well as the vizualisations obtained for the learners. These visualisations inform the learner on linguistics aspects of his/her writing.

Keywords. English L2, Learner corpus research, ITS, distance to natives, Learning management systems (LMS), ICALL

1 Introduction

Cet article décrit un prototype de dispositif automatisé de visualisation des productions écrites d'apprenants à partir de plusieurs critères linguistiques. Il s'inscrit dans une pédagogie d'évaluation formative visant à proposer des feedback spécifiques sans jugement de valeur [1]. Nous présentons ici une première expérimentation d'un prototype d'EIAH résultant d'un travail de plusieurs années sur les données d'apprenants [2] [3]. Il s'agit d'un dispositif d'analyse automatique de productions écrites mises en regard avec des productions de natifs. Le but est de visualiser les écarts existants entre ces deux types de production sur la base de critères linguistiques objectifs tels que la complexité lexicale, la lisibilité et la complexité syntaxique. Ce dispositif repose sur l'ingénierie d'une chaîne de traitement associant recueil de textes, enrichissement par annotation automatique, calcul de mesures textuelles et

visualisations sous formes de graphiques. La section 2 présente le dispositif de collecte et la visualisation des données. La section 3 présente nos perspectives.

2 Caractérisation du dispositif et visualisation

Après saisie dans une base de données Moodle [4, 5], les textes sont traités automatiquement avec les outils Quandeta [6], L2SCA [7] et TAALES [8].

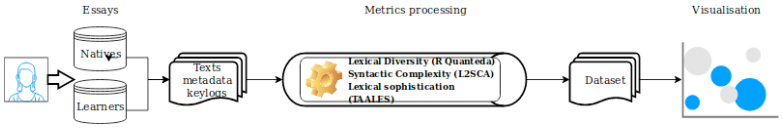


Fig 1 : Chaîne de traitement sous R des données collectées sous Moodle

Pour le positionnement relatif aux performances des natifs, nous avons utilisé un 20 écrits comparables extraits du corpus ICE-GB [9]. Cela représente 43 000 tokens, soit une moyenne d'un peu plus de 2 100 tokens par texte. Dans l'échantillon analysé du Corpus d'Étude des Langues Vivantes Appliquées à une Spécialité (CELVA.Sp), nous disposons de 286 productions d'étudiants LANSAD collectées à l'université de Rennes 1, soit un total de 62 724 tokens. Les textes sont classés par niveau de langue CECRL [10] selon les résultats obtenus par les étudiants au module écrit du test Dialang [11].

L'outil de visualisation est organisé comme un « tableau de bord » portant sur les différentes dimensions linguistiques à diagnostiquer. Nous renvoyons à [12] pour des présentations plus fouillées des métriques de lisibilité, à [13] pour la complexité syntaxique et à [14] pour la sophistication. Notre sélection des métriques est guidée par les premiers résultats de l'analyse de leur efficacité dans la discrimination des niveaux d'apprenants francophones [2] ou de leur corrélation avec la qualité des écrits en L2 [15, 16]. La Figure 2 présente la visualisation des métriques d'un texte d'apprenant (points étiquetés « learner ») par rapport aux métriques des textes de natifs. La première ligne (LexDiv.x) visualise la diversité lexicale à partir des trois mesures suivantes : D, K et TTR. D et K sont des mesures indiquant le taux de répétition des mots [14]. Il s'agit à terme de contribuer à la prise de conscience chez les étudiants de l'existence d'une grammaire textuelle, et à défaut d'une analyse en règle de la progression thème/rhème (*information packaging*). La deuxième ligne (Read.x) correspond aux indicateurs de lisibilité. Nous avons suivi [3] et sélectionné le FOG, le LIW et le RIX qui sont des indicateurs prenant en compte à la fois le nombre de mots mais aussi le caractère polysyllabique des mots. L'hypothèse étant que la longueur des mots accroît la difficulté de lecture. La troisième ligne (Synt.x) vise à représenter la grammaire de phrase et comporte les métriques issues du L2SCA [7]. Nous avons retenu des indicateurs intuitivement clairs pour les apprenants qui se fondent sur la notion de proposition. Le rapport entre le nombre de propositions et le nombre de phrases (C/S), le rapport entre le nombre de subordonnées (DC pour Dependent Clauses) et le nombre de propositions (C) renseignent sur le degré

d'hypotaxe d'un texte. La longueur moyenne des propositions est aussi prise en compte au sein du rapport entre le nombre de mots et le nombre de propositions. La visualisation de ces différents rapports peut fonctionner comme un tableau de bord des automatismes à acquérir et des objectifs à viser. Par exemple, si on prend la visualisation LexDiv.D, on voit que l'étudiant produit un degré de répétition de niveau relativement bas mais comparable à celui d'un groupe d'une dizaine de natifs.

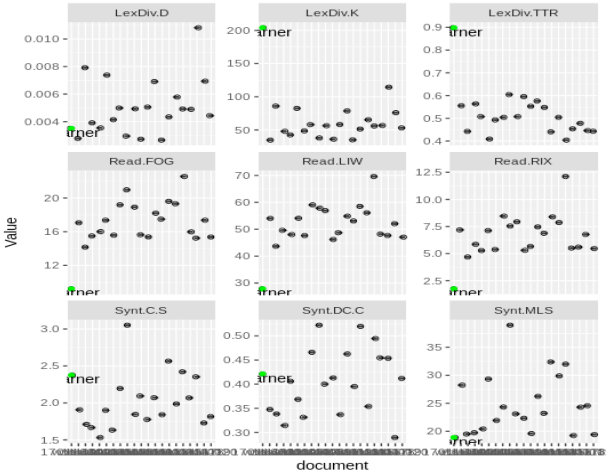


Fig. 2 : Visualisation des scores d'un apprenant (premier point à gauche étiqueté 'learner') en fonction des scores du groupe d'étudiants natifs anglophones (points noirs) du corpus ICE-GB.

3 Discussion et perspectives

L'enjeu est de contextualiser cette expérimentation de la visualisation de données au sein d'une problématique plus générale de la fouille de données éducatives. Au plan de l'enseignement, il s'agit de bénéficier de l'automatisation des procédures pour guider les étudiants à analyser leurs productions. Le pari est aussi celui d'une initiation de nos étudiants à ces problématiques plus linguistiques, qui permettrait le développement d'une certaine réflexivité sur les apprentissages. Au plan de la recherche, cette collecte et cette interprétation des traces numériques peuvent participer de l'analytique de l'apprentissage.

Sans doute sommes-nous encore au début de l'exploitation expérimentale de ces métriques. Les études à venir contribueront à faire le tri des métriques les plus pertinentes pour les différents stades d'apprentissage ([2], [13], [14]). En particulier, le "tableau de bord" pourrait prioriser / afficher des métriques différentes selon les niveaux et surtout selon le retour étudiant sur ce type de feedback formatif [1].

A moyen terme, nous voudrions affiner l'analyse des productions écrites par l'analyse des traces numériques claviers (*keylogs*) : repérer les zones d'hésitation des

apprenants, essayer de corréler la longueur des jets textuels [17], travailler sur les structures préfabriquées du lexiques (unités phraséologiques). Pour finir, l'enregistrement des traces numériques permet d'obtenir une vue dynamique sur la production. Ainsi, les corrections et différentes éditions deviennent visibles et peuvent, elles aussi, permettre de juger du niveau de maîtrise de la langue de l'apprenant.

Références

[1] V. J. Shute, "Focus on formative feedback," *Review of educational research*, vol. 78, no. 1, (2008), 153–189

[2] T. Arnold, N. Ballier, T. Gaillat, and P. Lissón, "Predicting CEFRL levels in learner English on the basis of metrics and full texts," arXiv preprint arXiv:1806.11099, (2018)

[3] P. Lissón et N. Ballier 'Investigating lexical progression through lexical diversity metrics in a corpus of French L3', *Discours,* 23, 2, December 2018, à paraître début juin 2019

[4] M. Dougiamas and P. Taylor, "Moodle: Using learning communities to create an open source course management system," in EdMedia: World Conference on Educational Media and Technology, Association for the Advancement of Computing in Education (AACE), (2003) 171–178

[5] C. Romero, S. Ventura, and E. Garcıa, "Data mining in course management systems: Moodle case study and tutorial," Computers & Education, vol. 51, no. 1, pp. 368–384, (2008).

[6] K. Benoit, K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, and A. Matsuo, "Quanteda: An R Package for the Quantitative Analysis of Textual Data," *Journal of Open Source Software*, vol. 3, no.30, (2018), 774.

[7] X. Lu, "Automatic analysis of syntactic complexity in second language writing," *International journal of corpus linguistics*, vol. 15, no. 4, (2010), 474–496.

[8] K. Kyle, S. Crossley, and C. Berger, "The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0," *Behavior research methods*, vol. 50, no. 3, pp. 1030–1046, (2018).

[9] G. Nelson, S. Wallis, and B. Aarts, *Exploring natural language: working with the British component of the International Corpus of English*, vol. 29. John Benjamins Publishing, (2002)

[10] Conseil de l'Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). Common European Framework of Reference for Languages: learning, teaching, assessment. Cambridge University Press.

[11] J. C. Alderson, and A. Huhta. "The Development of a Suite of Computer-Based Diagnostic Tests Based on the Common European Framework." *Language Testing* vol. 22, no. 3 (2005) 301–20

[12] T. François, *Les apports du traitement automatique du langage à la lisibilité du français langue étrangère*. PhD thesis, Ph. D. thesis, Université Catholique de Louvain, (2011)

[13] X. Lu, *Computational methods for corpus annotation and analysis*. Springer (2014)

[14] F. J. Tweedie and R. H. Baayen, "How variable may a constant be? measures of lexical richness in perspective," *Computers and the Humanities*, vol. 32, no. 5, (1998) 323–352

[15] P. Lissón, "Investigating the use of readability metrics to detect differences in written productions of learners: a corpus-based study," *Bellaterra Journal of Teaching & Learning Language & Literature,* vol. 10, no. 4, pp. 68–86, (2017)

[16] M. Kim and S. A. Crossley, "Modeling second language writing quality: A structural equation investigation of lexical, syntactic, and cohesive features in source-based and independent writing," *Assessing Writing*, vol. 37, (2018) 39–56

[17] G. Cislaru, et T. Olive, *Le processus de textualisation: Analyse des unités linguistiques de performance écrite*. De Boeck Supérieur, (2018)

Index des auteurs

A
Acensio, Laurie 375
Adam, Michel 359
Amadieu, Franck 331
Ana-Vanêssa, Lucena 391
Andriamasinoro, Marion 295
Androwkha, Sonia 301
Antelme, Daniel 343
Arbach, Najib 371
Arnaud, Zeller 319

B
Bailly, Gérard 363
Ballier, Nicolas 395
Barré, Vincent 109
Bernies, Maud 383
Bettenfeld, Vincent 151
Bevacqua, Elisabetta 307
Beyne-Rauzy, Odile 157
Bigrat, Frédérick 367
Blanchard Rodrigues, Christine 295
Bonnat, Catherine 193, 283
Bonvin, Guillaume 61
Bosse, Marie-Line 363
Bouchet, François 121
Bourda, Yolaine 349
Broisin, Julien 235

C
Caron, Pierre-André 25
Catteau, Olivier 157
Celso, Gonçalves 61
Choquet, Christophe 151
Cisel, Matthieu 13, 247, 289
Colmant, Yohan 349
Colombier, Nathalie 199
Cuko, Kostanca 37

D
d'Ham, Cédric 49, 139
Da Costa, Julien 181
Delestre, Nicolas 343, 349
Dumarski, Théo 371

E
Etienne, Mael 91

F
Fessard, Grégoire 223
Frison, Patrice 359

G
Gaillat, Thomas 395
Galaup, Michel 145
Galez, Christine 91
Gandon, Cécile 109
Gay, Dominique 85
Geoffre, Thierry 37, 199
Girault, Isabelle 49, 133, 139
Godde, Erika 363
Gomez, Rosa Maria 349
Grugeon-Allys, Brigitte 205
Gueraud, Viviane 283
Guinebert, Mathieu 79

H
Harrak, Fatima 121
Hérouard, Clément 235

I
Iksal, Sébastien 271
Illouz, Gabriel 379

K
Karoui, Aous 133

L
Laforcade, Pierre 169
Lagarrigue, Pierre 145
Laghouaouta, Youness 169
Laurent, Mélody 367
Le Gall, Erwan 349
Lefebvre, Nathalie 109
Lefevre, Marie 371
Lejeune, Anne 193, 283
Lesnes-Cuisiniez, Elann 205
Luengo, Vanda 79, 121
Léonard, Marielle 211

M

Mahdi, Oussema 271
Malandain, Nicolas 343
Mandran, Nadine 133, 139
Marfisi-Schottman, Iza 307
Marne, Bertrand 1
Marzin-Janvier, Patricia 49, 133, 139, 283
Mayere, Anne 157
Mercadier, Catherine 145
Meyer, Florian 109
Miras, Grégory 371
Molinari, Gaëlle 383
Moncef, Daoud 359
Mouaici, Mohamed 91
Mueller, Andreas 181
Mulet, Julie 331
Muratet, Mathieu 79
Méadel, Cécile 367

N
Najoua, Mohib 319
Nkambou, Roger 259

O
Oubahssi, Lahcen 271

P
Pacquetet, Erin 395
Parmentier, Jean-Francois 97
Paukovics, Elsa 387
Peter, Yvan 211
Peterlongo, Marie 349
Petit, Matthieu 109
Piau-Toffolon, Claudine 151, 271
Pierrot, Laëtitia 355
Pingaud, Hervé 145
Planche, Maëlle 133, 139
Pons Lelardeux, Catherine 145
Pontalier, David 289

R
Ramolet, Arthur 145
Raphaëlle, Crétin-Pirolli 151
Rapilly, Louis 371
Renna, Ilaria 223
Rodi, Mireille 37
Rome, Sandrine 331

S
Sacksick, Marie 73
Sanchez, Eric 61
Savy, Nicolas 157
Sebastien, Didier 85
Sebastien, Véronique 85
Secq, Yann 211
Serge, Robert 259
Silvestre, Franck 97
Szilas, Nicolas 181

V
Van de Leemput, Cécile 331
Van der Linden, Jan 331
Verrier, Romuald 349
Vignollet, Laurence 91
Vilnat, Anne 379
Vinatier, Isabelle 307

W
Wajeman, Claire 49, 139
Wang, Patrick 223

Y
Yessad, Amel 79

Z
Zhai, Yuming 379