



HAL
open science

Confinement et distancié forcés chez les enseignant.e.s des 1er et 2nd degrés : genèse d'une hybridation des pratiques ?

Gaëlle Lefer Sauvage, Sylvain Genevois, Jacques Kerneis, Nathalie Wallian

► To cite this version:

Gaëlle Lefer Sauvage, Sylvain Genevois, Jacques Kerneis, Nathalie Wallian. Confinement et distancié forcés chez les enseignant.e.s des 1er et 2nd degrés : genèse d'une hybridation des pratiques ?. EIAH, Jun 2021, FRIBOURG, Suisse. hal-03276466

HAL Id: hal-03276466

<https://hal.science/hal-03276466>

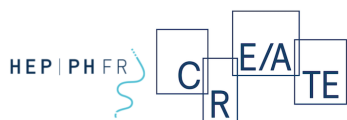
Submitted on 2 Jul 2021

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.

EIAH 2021

**Actes de la 10^e Conférence
sur les Environnements Informatiques pour l'Apprentissage Humain**



Lyon 1



Actes de la Conférence EIAH 2021

**10^e Conférence sur les
Environnements Informatiques
pour l'Apprentissage Humain**

**Du 7 au 10 juin 2021
Haute École Pédagogique Fribourg (HEP|PH FR), Suisse**

**Edités par
Marie Lefevre, Christine Michel,
Thierry Geoffre, Mireille Rodi, Lionel Alvarez, Aous Karoui**

**Site de la Conférence :
<https://blog.hepfr.ch/eiah2021/>**

**Association des Technologies de l'Information
pour l'Éducation et la Formation**

ISBN : 978-2-901384-03-8

Table des matières

Les comités	1
Introduction	5
Le mot du Comité d'Organisation	10
Conférencière invitée : Margarida Romero	11
Conférencière invitée : Hélène Sauzéron	20
Table ronde	28
Journée Ateliers	29
Formation Doctorale	31
Textes des communications retenues	32
Articles longs	33
Amélioration de l'interprétabilité du diagnostic cognitif de l'apprenant par catégorisation des composantes de connaissance, <i>Olivier Allègre, Amel Yessad, Vanda Luengo</i>	33
Vers la conception de <i>feedback</i> pour enseignants dans un contexte d'évaluation formative à grande échelle : une approche analytique, <i>Rialy Andriami- zeza, Franck Silvestre, Jean-François Parmentier, Julien Broisin</i>	46
Apprentissage de représentations pour l'enseignement de la programmation : une approche centrée enseignant, <i>Guillaume Cleuziou, Frédéric Flouvat, Matthieu Exbrayat, Julien Robert, Romuald Thion</i>	58

Suivi longitudinal de la fluence en lecture par évaluation automatique de la parole, <i>Erika Godde, Gérard Bailly, Anne Laure Piat-Marchand, Marie Line Bosse</i>	70
Rétroactions épistémiques dans une plateforme d'entraînement aux mathématiques : apports croisés de l'informatique et de la didactique, <i>Sébastien Jolivet, Amel Yessad, Mathieu Muratet, Vanda Luengo, Benjamin Reiter, Brigitte Grugeon-Allys, Elann Lesnes-Cuisiniez</i>	82
Le cursus académique personnalisé dans une approche par compétences avec érosion : étude d'un nouveau problème fondamental, <i>Alexis Lebis, Jérémie Humeau, Anthony Fleury, Mathieu Vermeulen</i>	94
Intention d'usage des tablettes en formation initiale d'enseignants à Mayotte, <i>Gaëlle Lefer Sauvage</i>	106
Expérimentation à grande échelle d'applications pour tablettes pour favoriser l'apprentissage de la lecture et de l'anglais oral, <i>Sonia Mandin, Ahmed Zaher, Svetlana Meyer, Mathieu Loiseau, Gérard Bailly, Coralie Payre-Ficout, Julien Diard, Fluence Group, Sylviane Valdois</i>	118
Comment guider les doctorants dans l'utilisation du Design Based Research ?, <i>Nadine Mandran, Mathieu Vermeulen, Estelle Prior</i>	130
Stratégies pour de nouvelles formes d'appropriation des ENT. Impact de l'école à la maison dans le développement des usages et pratiques numériques des enseignants, <i>Christine Michel, Laëtitia Pierrot</i>	142
Réalité virtuelle et formation à distance : impact de la représentation des avatars sur la coprésence des apprenants dans un environnement virtuel collaboratif, <i>Richard Ngu Leubou, Marc Trestini, Benoit Crespin</i>	154
Combiner les leviers de l'approche par compétences et de l'auto-régulation pour accompagner le travail en autonomie à l'université. Analyse du service COMPER, <i>Laëtitia Pierrot, Christine Michel, Julien Broisin, Nathalie Guin, Marie Lefevre, Rémi Venant</i>	166
Génération d'un profil dynamique du stress pour l'entraînement à la gestion de situations de crise, <i>Luca Pelissero-Witoslawski, Domitile Lourdeaux, Dominique Lenne</i>	178
Soutien aux pédagogies actives et transformations impulsées par une plateforme numérique dédiée aux sciences expérimentales, <i>Maelle Planche, Cédric d'Ham, Christian Hoffmann, Nadine Mandran, Isabelle Girault, Claire Wajeman, Patricia Marzin</i>	190
La prise en compte de l'incertitude dans le processus de diagnostic des compétences de l'apprenant, <i>Nesrine Rahmouni, Domitile Lourdeaux, Az-zeddine Benabbou, Taher Bensebaa</i>	202

Les modèles apprenant pour soutenir l'apprentissage tout au long de la vie : revue de littérature, <i>Sergio Iván Ramírez Luelmo, Nour El Mawas, Jean Heutte</i>	214
Conception et évaluation d'un processus de personnalisation fondé sur des référentiels de compétences, <i>Louis Sablayrolles, Nathalie Guin, Marie Lefevre, Julien Broisin</i>	226
Conception et expérimentation de tableaux de bord d'apprentissage pour les enseignants et les apprenants, <i>Yassine Safsouf, Franck Poirier, Khalifa Mansouri</i>	238
Détection de déficits d'auto-évaluation et d'auto-efficacité dans un logiciel enseignant la lecture et l'écriture, <i>Thomas Sergent, Morgane Daniel, François Bouchet, Thibault Carron</i>	250
Analyse de remédiations proposées par des enseignants ciblant des déficits d'auto-régulation de jeunes élèves dans une application d'apprentissage de la lecture, <i>Thomas Sergent, Morgane Daniel, François Bouchet, Thibault Carron</i>	262

Articles courts **274**

Fondements pédao-didactiques d'un environnement informatique d'apprentissage destiné aux jeunes apprenants de l'école primaire, <i>Lionel Alvarez, Aous Karoui, Mario Ramalho, Thierry Geoffre</i>	274
S'entraîner à l'asepsie avec la réalité virtuelle, <i>Leonardo Angelini, Alexandre Favre, François Magnin, Mariateresa De Vito Woods, Omar Abou Khaled, Elena Mugellini</i>	281
Outils numériques pour l'apprentissage de la grammaire en français langue d'enseignement : Quels analyseurs pour les appréhender ?, <i>Rosianne Arseneau, Thierry Geoffre</i>	288
Système de recommandation de ressources pédagogiques pour un apprentissage sur application mobile parascolaire, <i>Anaëlle Badier, Mathieu Lefort, Marie Lefevre</i>	294
Conception d'un compagnon digital pour assister le maître du jeu dans un environnement mixte, <i>Catherine Bonnat, Eric Sanchez, Matthieu Vallat, Omar Abou Khaled, Elena Mugellini</i>	300
Vers une méthodologie de co-conception de <i>learning games</i> par les étudiants en soins infirmiers, <i>Sebastian Gajewski, Nour El Mawas, Jean Heutte</i>	306
Aide au suivi de la progression de groupes d'apprenants pour la mise en place d'une pédagogie différenciée, <i>Fatima Harrak, François Bouchet</i>	312

Parcours adaptatifs au sein de la Plateforme Européenne d'Apprentissage Personnalisé des Langues PEAPL, <i>Aous Karoui, Lionel Alvarez, Mario Ramalho, Thierry Geoffre</i>	318
La chorégraphie des traces étudiantes : une nouvelle visualisation multidimensionnelle pour les acteurs pédagogiques, <i>Alexis Lebis, Nadine Mandran, Mathieu Vermeulen</i>	324
Les IPMC : une plongée au cœur de la mise en pratique des compétences, <i>Alexis Lebis, Estelle Prior, Nadine Mandran, Mathieu Vermeulen</i> .	330
Confinement et distanciel forcés chez les enseignant.e.s des 1er et 2nd degré : genèse d'une hybridation des pratiques ?, <i>Gaëlle Lefer Sauvage, Sylvain Genevois, Jacques Kerneis, Nathalie Wallian</i>	336
Vers une méthode de méta-design de jeux sérieux : Application pour l'apprentissage de la programmation à travers Blockly Maze, <i>Bertrand Marne, Mathieu Muratet, Karim Sehaba</i>	342
Vers l'utilisation de pictogrammes affichables sur des tablettes à picots rétractables pour l'illustration de livres tactiles : importance de la résolution de la tablette, <i>Caroline Masclé, Gwenaël Kaminski, Florence Bara, Christophe Jouffrais</i>	348
Représentations sociales de l'analytique des apprentissages avec le numérique, <i>Camila Morais Canellas, François Bouchet, Vanda Luengo</i>	354
Faciliter le recueil de traces par la numérisation d'un outil tangible de co-design : application à la conception de tableaux de bord d'apprentissage, <i>Katia Oliver-Quellenec, François Bouchet, Thibault Carron, Claire Pinçon</i>	360
Faciliter l'intégration des critères RGPD et éthique lors de la co-conception des jeux d'apprentissage numériques, <i>Estelle Prior, Mariem Jaouadi, Nadine Mandran, Eric Sanchez</i>	366

Posters et démonstrations **372**

Visualisation automatique de graphes pour l'apprentissage des algorithmes, <i>Pierre Baron, Laurent Roche, Thibault Raffailac</i>	372
Les <i>Teaching Analytics</i> pour l'auto-évaluation par analyse comportementale de l'enseignant sur un LMS, <i>Ibtissem Bennacer, Remi Venant, Sebastien Iksal</i>	377
Rapport au savoir en contexte muséal : le cas du jeu <i>Geome</i> , <i>Catherine Bonnat, Gil Oliveira, Simon Morard, Elsa Paukovics, Eric Sanchez</i>	381
Etude de cas : schèmes et invariants opératoires chez les professionnels de la formation par et pour le numérique - Projet Interreg - Teach Transition, <i>Elfi Casanave, Pierre-André Caron</i>	385

<i>DigiTrainer</i> , un outil pour former ceux qui forment, <i>Pierre-François Coen</i> , <i>Edmée Runtz-Christan</i>	389
Méthodologie d'analyse des contributions rédactionnelles lors d'une co-écriture en ligne, <i>Edith Durand</i> , <i>Paul Lotin</i> , <i>Damien Chabanal</i> , <i>Simon Ensor</i> , <i>Loïse Jeannin</i> , <i>Anne-Laure Foucher</i> , <i>Lidia Lebas</i> , <i>Véronique Quanquin</i> , <i>Ciara R. Wigham</i> , <i>Christine Blanchard</i>	393
Articuler présence et distance à la suite du confinement : genèse des pratiques hybrides?, <i>Gaelle Lefer Sauvage</i> , <i>Cendrine Mercier</i>	397
Modèle de méthode pour accompagner les équipes dans le processus de concep- tion des jeux épistémiques numériques, <i>Estelle Prior</i>	401
<i>Geome</i> et son Compagnon Digital : un dispositif permettant la mise en place de visites muséales ludiques, <i>Eric Sanchez</i> , <i>Catherine Bonnat</i> , <i>Gil Oli- veira</i> , <i>Matthieu Vallat</i> , <i>Kevin Manixab</i> , <i>Didier Crausaz</i> , <i>Nicolas Kra- mar</i> , <i>Sylvia Mueller</i> , <i>Omar Abou Khaled</i> , <i>Elena Mugellini</i>	405

Index des auteurs **409**

Membres des comités

Présidentes du comité de programme

Marie Lefevre	LIRIS, Université Lyon 1, France
Christine Michel	Techné, Université Poitiers, France

Comité de programme

Abel Marie-Hélène	HEUDIASYC, Université de Technologie de Compiègne, France
Amadiou Franck	CLLE, Université de Toulouse, France
Baron Georges-Louis	Université Paris Descartes, France
Baron Monique	LIP6, Sorbonne Université, France
Becerril-Ortega Raquel	CIREL, Université de Lille, France
Bonnat Catherine	Université Fribourg, Suisse
Bouchet François	LIP6, Sorbonne Université, France
Bouhineau Denis	LIG, Université Grenoble Alpes, France
Bourda Yolaine	LRI, CentraleSupélec, France
Bourdet Jean-François	Université du Maine, France
Bousbia Nabila	LMCS, Ecole nationale Supérieure d'Informatique, Algérie
Broisin Julien	IRIT, Université Paul Sabatier Toulouse 3, France
Caron Pierre-André	CIREL, Université de Lille, France
Carron Thibault	LIP6, Sorbonne Université, France
Céci Jean-François	Passages, Université de Pau et des Pays de l'Adour, France
Cerisier Jean-François	Techné, Université de Poitiers, France
Charlier Bernadette	Université de Fribourg, Suisse
Crétin-Pirolli Raphaëlle	IUT Laval, Université du Maine, France
De La Higuera Colin	LS2N, Université de Nantes, France
De Lievre Bruno	Université de Mons, Belgique
Delestre Nicolas	LITIS, Normandie Université, INSA Rouen Normandie, France
Depover Christian	Université de Mons, Belgique

Després Christophe	LIUM, Le Mans Université, France
Dessus Philippe	LaRAC, Université Grenoble Alpes, France
El Kechai Hassina	Techné, Université de poitiers, France
El Mawas Nour	CIREL, Université de Lille, France
Fleck Stephanie	PERSEUS, Université de Lorraine, France
Fluckiger Cédric	CIREL, Université de Lille, France
Garlatti Serge	Lab-STICC, IMT Atlantique, France
Geoffre Thierry	Haute École Pédagogique Fribourg, Suisse
George Sébastien	LIUM, Le Mans Université, France
Gilliot Jean-Marie	Lab-STICC, IMT Atlantique, France
Girault Isabelle	LIG, Université Grenoble Alpes, France
Grandbastien Monique	LORIA, Université de Lorraine, France
Groux-Leclot Dominique	MIS, Université de Picardie Jules Verne, France
Guin Nathalie	LIRIS, Université de Lyon, France
Huet Nathalie	CLLE-LTC, Université Toulouse 2, France
Iksal Sébastien	LIUM, Le Mans Université, France
Jaccard Dominique	University of Applied Sciences Western Switzerland, Suisse
Jean-Daubias Stéphanie	LIRIS, Université de Lyon, France
Jolivet Sébastien	LIP6, Sorbonne Université / LDAR, Université de Paris, France
Labat Jean-Marc	LIP6, Sorbonne Université, France
Lafifi Yacine	LabSTIC, University 8 May 1945 Guelma, Algérie
Laforcade Pierre	LIUM, Le Mans Université, France
Lavoué Elise	LIRIS, IAE Lyon, Université Jean Moulin Lyon 3, France
Lebis Alexis	IMT Lille Douai, France
Lefevre Marie	LIRIS, Université de Lyon, France
Lejeune Anne	LIG, Université Grenoble Alpes, France
Lenne Dominique	HEUDIASYC, Université de Technologie de Compiègne, France
Leroux Pascal	CREN, Le Mans Université, France
Libbrecht Paul	IUBH Fernstudium, Allemagne
Loisy Catherine	IFé, École Normale Supérieure de Lyon, France
Lourdeaux Domitile	HEUDIASYC, Université de technologie de Compiègne, France
Luengo Vanda	LIP6, Sorbonne Université, France
Mailles-Viard Metz Stéphanie	ADEF, Aix Marseille Université, France
Mandran Nadine	LIG, Université Grenoble Alpes, France
Marfisi-Schottman Iza	LIUM, Le Mans Université, France

Marne Bertrand	ICAR, Université Lyon 2, France
Marty Jean-Charles	LIRIS, Université de Savoie, France
Merceron Agathe	Beuth University of Applied Sciences Berlin, Allemagne
Meyer Florian	Université de Sherbrooke, Québec, Canada
Michel Christine	Techné, Université de Poitiers, France
Molinari Gaëlle	Unidistance & TECFA, Université de Genève, Suisse
Muratet Mathieu	LIP6, INS HEA, France
Nkambou Roger	Université du Québec, Canada
Nodenot Thierry	LIUPPA, IUT de Bayonne Pays Basque, France
Nogry Sandra	Paragraphe, Université Cergy-Pontoise, France
Oubahssi Lahcen	LIUM, Le Mans Université, France
Pélissier Chrysta	LHUMAIN, Université Montpellier 3, France
Peraya Daniel	TECFA, Université de Genève, Suisse
Perez-Sanagustin Mar	IRIT, Université Paul Sabatier Toulouse 3, France
Peter Yvan	Université de Lille, France
Pierrot Laëtitia	Techné, Université de Poitiers, France
Poyet Françoise	ELICO, Université Claude Bernard Lyon 1, INSPE de Lyon, France
Reffay Christophe	ELLIADD, Université de Franche-Comté, France
Reuter Robert	Université du Luxembourg, Luxembourg
Sanchez Eric	TECFA, Université de Genève, Suisse
Sehaba Karim	LIRIS, Université Lumière Lyon 2, France
Secq Yann	CRIStAL, Université de Lille, France
Seridi Hassina	LabGED, Université Badji Mokhtar-Annaba, Algérie
Silvestre Franck	IRIT, Université Paul Sabatier Toulouse 3, France
Sutter Widmer Denise	TECFA, Université de Genève / UER MI, Haute école pédagogique Vaud, Suisse
Temperman Gaëtan	Université de Mons, Belgique
Trestini Marc	LISEC, Université de Strasbourg, France
Vadcard Lucile	LARAC, Université Grenoble Alpes, France
Venant Rémi	LIUM, Le Mans Université, France
Vermeulen Mathieu	IMT Lille Douai, France
Yessad Amel	LIP6, Sorbonne Université, France

Président·e·s des ateliers

Laëtitia Pierrot Techné, Université Poitiers, France
Venant Rémi LIUM, Le Mans Université, France

Président·e·s de la formation doctorale

Catherine Bonnat TECFA, Université de Genève, Suisse
Jolivet Sébastien LIP6, Sorbonne Université / LDAR, Université de Paris, France
Mandran Nadine LIG, Université Grenoble Alpes, France
Venant Rémi LIUM, Le Mans Université, France

Président·e·s du comité d'organisation

Thierry Geoffre UR DidaLang/CRE/ATE, HEP|PH FR, Suisse
Mireille Rodi UR DidaLang, HEP|PH FR, Suisse
Lionel Alvarez CRE/ATE, HEP|PH FR, U-Fribourg, Suisse
Aous Karoui UR DidaLang/CRE/ATE, HEP|PH FR, Suisse

Bénévoles

Suzy Blondi
Quentin Brumeau
Nicolas Lefranc
Hugo Mora
Loïc Piernas
Mario Ramalho
Ania Tadlaoui

La conférence EIAH'21

Marie Lefevre¹, Christine Michel²

Présidentes du comité de programme EIAH'21

¹ Université de Lyon, CNRS, LIRIS UMR 5205, F-69222 Villeurbanne, France
marie.lefevre@liris.univ-lyon1.fr

² Université de Poitiers, TECHNE EA 6316, F-86073 Poitiers, France
christine.michel@univ-poitiers.fr

1 Introduction

La conférence Environnements Informatiques pour l'Apprentissage Humain (EIAH) est une conférence internationale, pluridisciplinaire et francophone sur la conception et l'analyse des environnements numériques pour l'éducation et la formation, placée sous l'égide de l'ATIEF (Association des Technologies de l'Information pour l'Éducation et la Formation).

La 10^{ème} édition de la conférence EIAH (EIAH'2021) a eu lieu en ligne du 7 au 10 juin 2021 et est organisée et portée par la Haute École Pédagogique Fribourg (HEP|PH FR). Elle succède à la conférence EIAH'2019 organisée à Paris en 2019.

La thématique scientifique cette année concerne la question des transformations dans le domaine des EIAH, en particulier décrire ou discuter les innovations technologiques et d'usage(s), leurs caractéristiques et leurs effets sur la construction des connaissances et l'expérience d'apprentissage, mais aussi plus globalement les transformations qu'elles induisent en fonction des contextes des usages.

La révolution des modalités d'apprentissage liée à la pandémie de COVID'19 nous a semblé porter beaucoup de potentiels en termes d'analyse des innovations technologiques et d'usage(s). Au-delà de cet aspect conjoncturel, nous souhaitons:

- favoriser les réflexions systémiques et situées sur les approches, outils et méthodes permettant de concevoir, déployer ou observer les dispositifs hybrides à des niveaux micro ou macro,
- mais également constituer un panorama de la variété des dispositifs proposés en considérant les supports technologiques utilisés, les activités proposées, les effets observés en termes d'apprentissage et d'expérience utilisateur, ou les dynamiques sociales soutenant leur émergence ou leur pérennité, en particulier en termes d'acceptation et d'appropriation.

2 Des innovations dans la forme

Au-delà des axes de recherche, que nous présentons dans la section suivante, deux formes d'innovations caractérisent la conférence EIAH cette année : le mode de rencontre et le mode de financement.

Le contexte sanitaire a imposé que la tenue de la conférence soit à distance. Plutôt que d'utiliser une plateforme de travail collaboratif classique, le comité d'organisation a proposé d'héberger la conférence dans le monde Laval Virtual World¹. Si les présentations restent relativement similaires à une visioconférence classique, les interactions sociales, en petits comités comme dans le cadre des posters, ou en plus grand comité comme dans les sessions de présentation, promettent d'être plus directes et conviviales. Nous avons hâte de les expérimenter.

Le financement de la conférence EIAH'21 est réalisé cette année exclusivement *via* un droit de publication des articles. Nous avons fait ce choix, avec l'équipe d'organisation et le CA de l'ATIEF, afin de favoriser la participation à cette édition virtuelle. La conférence ainsi que les événements associés ont été donc ouverts et gratuits pour tous les participants, la seule condition étant de s'inscrire sur la plateforme Laval Virtual. Nous espérons par ce biais faire découvrir les recherches de la communauté EIAH à un plus grand nombre de chercheurs ou d'industriels. Ce pari est gagné en termes d'inscription puisque plus de 300 personnes ont choisi de participer. Le mode d'audience sera sans doute différent à celui d'une conférence en présentiel, avec une sélection plus personnelle des sessions ou présentations par les participants. Nous sommes très intéressées de voir comment les participants vont s'approprier cette nouvelle forme, quelle sera leur expérience et comment cela pourra être valorisée à plus long terme pour la diffusion et l'animation de la recherche.

3 Une tradition dans les thématiques de recherche

En termes d'activité de recherche justement, 68 articles ont été proposés initialement et 61 ont été finalement soumis : 38 articles longs, 17 articles courts et 6 démos et posters.

L'évaluation des propositions s'est faite selon le principe du double anonymat (des auteurs et des relecteurs). Pour chaque article soumis, trois relecteurs ont proposé une évaluation : deux du domaine de la proposition et un hors domaine. Ensuite, les auteurs ont pu prendre connaissance des évaluations et éventuellement apporter une réponse. Pour finir, un méta-relecteur a animé une discussion afin de proposer une décision d'acceptation ou non. Ce travail a impliqué 67 relecteurs et 20 méta-relecteurs.

Tableau 1. Nombre et type d'articles publiés.

Type	Soumis	Acceptés			Rejetés	Retirés par		
		directement	sous réserve	en article court		en poster	auteurs	Publiés
Long	38	16	4	10	1	7	(-2)	20
Court	17	5	2	(+9)	5	3	(-3)	16
Poster	5	4			(+4)	1		8
Démo	1	1						1
Total	61	26	6	9	4	11	(-5)	45

¹ <https://association.laval-virtual.com/fr/laval-virtual-world/>

Au final, 45 articles ont été publiés : 20 en article long, 16 en articles courts et 9 en démo et posters (voir tableau 1). Le taux de sélection pour les articles longs est de 52%.

Les auteurs sont principalement Français avec une petite proportion d'auteurs venant des départements et territoires d'outre-mer (Mayotte, La Réunion et La Nouvelle Calédonie). Les auteurs internationaux sont Suisses, Marocains, Tunisiens, Algériens et Canadiens (voir Figure 1).

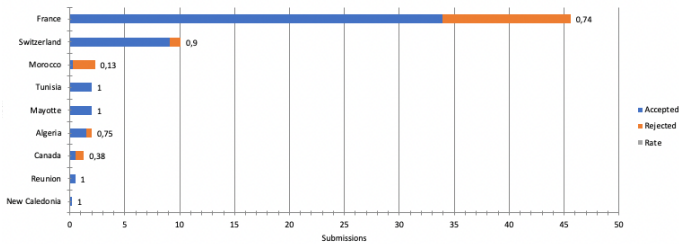


Fig. 1. Répartition des auteurs par pays.

La répartition entre STIC et SHS est équilibrée avec respectivement 22 et 23 articles (voir Figure 2). En ce sens, l'objectif d'avoir une conférence pluridisciplinaire est atteint.

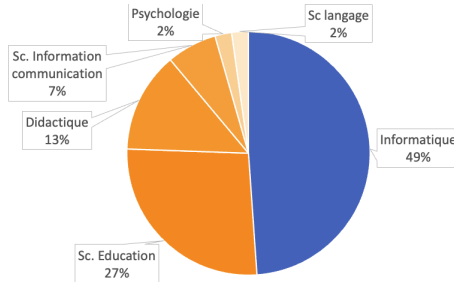


Fig. 2. Répartition des articles par discipline scientifique.

Les thématiques de travail qui ont structuré l'organisation des sessions portent sur les thématiques historiques de la conférence : la conception et plus globalement l'ingénierie des EIAH ainsi que l'analyse des usages et des pratiques ou l'observation des situations (voir Figure 3).

Les caractéristiques des EIAH qui sont plus spécifiquement travaillées sont, selon une dynamique qui se développe depuis plusieurs années de manière régulière : les

jeux, l'exploitation des traces d'apprentissage, les systèmes adaptatifs et la personnalisation. Ces constats restent assez cohérents avec les éditions 2017 et 2019.

Une thématique se développe : le travail sur les dimensions didactiques des EIAH.

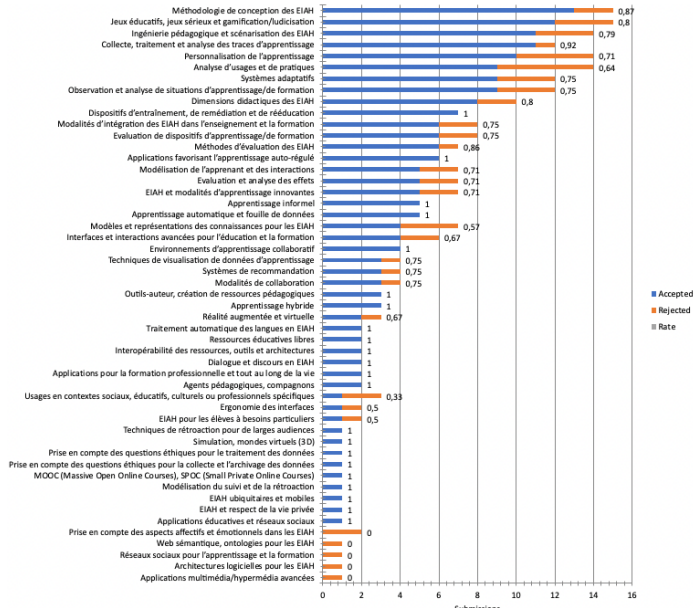


Fig. 3. Nombre de soumissions et taux d'acceptation par thématique

4 Remerciements

Nous tenions à remercier le comité d'organisation de la Haute École pédagogique de Fribourg qui a proposé de se jeter à l'eau et d'expérimenter la tenue de la conférence dans l'espace virtuel de la plateforme Laval Virtual. Ce choix est particulièrement challengeant pour eux, ce type d'organisation reste encore peu classique. Nous les remercions à la fois d'avoir osé le faire et de nous permettre à tous de l'expérimenter à grande échelle.

Nous remercions également tout particulièrement les membres du comité de programme, pour la qualité de leurs expertises et leur engagement. Les retours, conseils et avis donnés aux auteurs lors de la phase de rédaction ont été extrêmement constructifs pour affiner les contributions scientifiques et participer à la qualité de cette conférence.

Nous tenions enfin à remercier l'ATIEF qui a aidé à la tenue et la qualité de cette conférence par de multiples aspects. Sur le plan scientifique, la conférence EIAH'21 a

bénéficié d'une réflexion de fond, menée en concertation avec STICEF, sur les modalités d'évaluation des articles scientifiques. Les consignes et la grille d'évaluation utilisées pour la sélection des travaux ont permis aux relecteurs du comité de programme de produire des évaluations à la fois plus précises et informatives pour les auteurs, mais aussi pour les méta-relecteurs. Sur le plan pragmatique, une nouvelle bourse a financé l'inscription à la conférence de deux doctorants auteurs. De plus, les contraintes de gestions financières suisse et française étant incompatibles, une large partie des inscriptions a été prise en charge par le trésorier de l'ATIEF, Sébastien Iksal, que nous remercions chaleureusement.

Le mot du Comité d'Organisation

Thierry Geoffre

Haute École pédagogique Fribourg, 1700 Fribourg, Suisse
Thierry.Geoffre@eduf.fr.ch

L'idée de l'organisation de la conférence EIAH2021 en Suisse a été lancée lors de la 9^e Conférence, à Paris, en juin 2019. Accueillir un colloque sur les environnements informatiques pour l'apprentissage humain à la Haute École pédagogique de Fribourg, lieu dédié à la formation des futurs enseignant·e·s, à leurs apprentissages et aux apprentissages des élèves du primaire, nous a semblé particulièrement pertinent.

Comme les synergies entre l'UR DidaLang et le CRE/ATE (Centre de Recherche sur l'Enseignement/Apprentissage par les Technologies numériques) sont constantes, l'équipe s'est rapidement constituée entre Mireille Rodi, Lionel Alvarez et moi-même. La pandémie a ensuite fait table rase de l'organisation prévue : pas de présence à la HEP, pas de plaisir de se retrouver ou de se rencontrer, pas de chocolats suisses, pas de musée Tinguely ; mais l'environnement Laval Virtual (et une plage) en plus ! Ce choix a permis de renforcer l'investissement et les liens personnels des participants davantage que de simples sessions Teams. En accord avec le bureau de l'ATIEF, d'autres changements ont eu lieu ; les droits d'inscription sont ainsi devenus des droits de publication des articles et nous avons pu proposer des inscriptions gratuites et ouvertes à tou·te·s. L'environnement virtuel légitimait cette ouverture et ce sont finalement près de 350 communicant·e·s ou curieux·ses qui se sont inscrit·e·s.

Suite au choix de Laval Virtual, Aous Karoui, collaborateur scientifique au sein de l'UR, a rejoint le comité d'organisation et apporté sa connaissance du domaine et de l'environnement virtuel déjà utilisé lors de GALA2020.

Nous remercions ici l'équipe de Laval Virtual, en particulier Asma pour son écoute et son obstination à nous permettre de concrétiser la collaboration, et Pauline pour son soutien à l'élaboration du projet, sa disponibilité et son roadbook :-)

Nous remercions l'ATIEF, de nous avoir accordé sa confiance et son trésorier, Sébastien Iksal, qui a pris en charge les facturations et encaissements des droits de publications pour faciliter les démarches de paiement.

La HEP|PH FR a soutenu dès le début le projet d'organisation en son sein d'un colloque à la croisée de l'informatique et des sciences humaines. Nos vifs remerciements à Patrick Roy (Doyen de la recherche et prestations à des tiers) et Bertrand Forclaz (Adjoint du doyen et manager de la recherche) pour leur soutien et leur volonté de voir réussir cet événement au sein de la HEP|PH FR.

A titre personnel, je termine ces quelques lignes en remerciant Mireille, Lionel et Aous. Entre EIAH2021 et les projets, nous aurons passé l'année 2020-21 à distance mais sans se lâcher d'un pouce !

Du modèle de tâche au corpus de données pour l'analyse de la résolution créative de problème en contexte médiatisé

Margarida Romero¹, Louis Köhler¹

¹ Université Côte d'Azur, LINE, Nice, France
margarida.romero@univ-cotedazur.fr
louis.kohler@etu.univ-cotedazur.fr

Résumé. Dans le but d'analyser la résolution créative de problèmes dans le contexte d'une tâche de robotique pédagogique nous présentons un modèle de tâche sous forme de graphe orienté entre différents états entre la situation initiale et l'une des solutions potentielles. Pour analyser le processus de résolution créative de problèmes, des données sur le processus sont collectées suivant le même protocole. Un modèle de traces permet d'observer les différents états de la tâche. Dans cette étude nous centrons notre analyse sur les composantes de la pensée divergente liées à la fluidité, flexibilité et innovation de participants de différentes âges engagés dans la tâche de résolution créative de problèmes CreaCube. Les enfants présentent des scores plus importants sur l'ensemble des trois composantes créatives (fluidité, flexibilité et innovation) par rapport aux autres catégories d'âge, mais ils engagent également davantage de temps que les autres catégories d'âge.

Mots-clés : Traces, situation-problème, résolution créative de problèmes, robotique pédagogique, créativité.

Abstract. In order to analyze creative problem solving in the context of an educational robotics task we present a task model in the form of a directed graph between different states between the initial situation and one of the potential solutions. To analyze the creative problem-solving process, process data is collected using the same protocol. A trace model makes it possible to observe the different states of the task. In this study we focus our analysis on the components of divergent thinking related to the fluency, flexibility and innovation of participants of different ages engaged in the creative problem-solving task CreaCube. Children score higher on all three creative components (fluency, flexibility and innovation) compared to other age categories, but they also commit more time than other age categories..

Keywords: Learning analytics, problem-solving, creative problem-solving, educational robotics, creativity.

1 Résolution créative de problèmes

La résolution de problèmes est une compétence transversale considérée importante dans le contexte éducatif international [1]. Cependant, la résolution de problèmes englobe une grande diversité d'activités et de domaines disciplinaires, et la capacité de transférer des processus de résolution de problèmes d'une tâche à une autre doit être questionnée. Pour Kilpatrick [2] le niveau de compétence dans la résolution de problèmes spécifiques à un domaine dépend des connaissances spécifiques au domaine des apprenants, comprenant à la fois des connaissances factuelles, conceptuelles, procédurales et stratégiques. Tout en considérant l'importance des connaissances spécifiques tant du point de vue du domaine disciplinaire que d'un point de vue socio-culturel [3], nous considérons un ensemble de processus cognitifs et métacognitifs qui sont engagés dans la démarche générale de résolution de problèmes. Cette démarche est celle qui permet dépasser l'écart qui est posé entre une situation initiale et une situation désirée [4]. Dans les théories de résolution des problèmes il est habituel de retrouver tant des étapes liées à l'identification et catégorisation des problèmes, la construction d'une représentation mentale du problème, la recherche de stratégies ou procédures pour la résolution de problèmes, la mise à l'essai de ces procédures, leur évaluation et l'itération jusqu'à arriver à une solution satisfaisante. Ces étapes ne sont pas séquentielles, mais plutôt dynamiques et de nature itérative.

La résolution de problèmes comprend une diversité de caractéristiques [5], parmi lesquelles nous pouvons considérer la complexité du problème, ses modalités ou encore le degré de connaissance du modèle de résolution. Dans des problèmes avec un modèle de résolution connu, l'apprenant peut appliquer des connaissances existantes [6], tandis que sur des problèmes avec un modèle de résolution inconnu, l'apprenant doit se montrer créatif dans la manière de mobiliser ses connaissances, faire des mises à l'essai de solutions intermédiaires et tenter de trouver une solution convenable. Les problèmes sans modèle de résolution connu ont le potentiel d'engager l'apprenant dans une démarche créative. Sur ce type de problèmes « il existe de nombreuses façons de le résoudre, et les solutions sont influencées par le point de vue et l'expérience de chacun » [7, p. 28]. Sous une perspective similaire, Torrance et Goff [8] considèrent la résolution créative de problèmes comme la démarche consistant à trouver des réponses à des problèmes pour lesquels les individus ne peuvent pas simplement obtenir la réponse appropriée; ils doivent plutôt créer des solutions à ces problèmes. Dans cette démarche, nous retrouvons tant des processus de pensée divergente, que des prises de décision (pensée convergente) en lien aux différentes solutions produites. Parnes, Noller et Biondi [9] considèrent la résolution créative de problèmes comme une prise décision créative sur différentes options possibles : « nous spéculons d'abord sur "ce qui pourrait être"... nous sentons et anticipons toutes les conséquences ou les répercussions concevables... et nous choisissons et développons notre meilleure solution en pleine conscience » [p. 14].

2 Aspects développementaux dans la résolution créative de problèmes

La créativité est souvent perçue comme un trait propre de l'enfance [10], avec une propension davantage ouverte à l'expérience et à la nouveauté [11]. Cependant, elle présenterait des caractéristiques différentes selon les différents moments de la vie [12]. Chez les personnes âgées, malgré un déclin progressif des capacités physiques et cognitives, les connaissances acquises au long de la vie peuvent permettre de compenser certains processus. Dans le cadre de l'usage de technologies, leur approche est moins basée sur des essais et des erreurs, mais à une démarche davantage réflexive et utilitaire [13].

Dans le cadre de cette étude nous allons focaliser sur trois composantes de la pensée divergente à partir de Torrance [14] et Guilford [15]: la fluidité, la flexibilité et l'originalité ou innovation. « La fluidité correspond au nombre d'idées données, la flexibilité dans les réponses correspond au nombre de catégories dans lesquelles des idées peuvent être classées et l'originalité se réfère à la rareté statistique relative de chaque idée proposée en comparaison d'une population de référence » [16]. La flexibilité et l'originalité sont plus faibles chez les personnes plus âgées [17].

3 Résolution créative de problèmes avec des robots pédagogiques

Les robots pédagogiques ont été mobilisés en éducation dans une grande diversité d'activités, visant pour certaines des connaissances et des compétences disciplinaires ou encore transversales. Dans les activités de résolution de problèmes, les robots pédagogiques sont parfois mobilisés dans des problèmes bien définis, où l'apprenant dispose de toutes les règles pour résoudre l'activité, mais aussi peuvent donner lieu à des activités de résolution créative de problèmes qui permettent tant la considération d'une diversité de possibilités dans la démarche de résolution de problèmes que sur le type de solution pouvant donner réponse au problème [18]. Les tâches créatives de résolution de problèmes favorisent l'exploration des différentes stratégies pour aborder la ou les solutions à travers des approches de réflexion divergentes mais aussi l'évaluation de ces stratégies à travers un processus de réflexion convergent. Quand les activités sont instrumentées par des robots pédagogiques, il est possible d'observer par le biais de l'artefact les processus divergents par le biais des configurations réalisées par le participant. En soutenant la pensée divergente et convergente, les tâches de résolution créative de problèmes avec des robots pédagogiques créent les conditions pour l'émergence de processus créatifs.

La résolution créative de problèmes est un processus dynamique qui engage le participant dans une situation pour laquelle il faut créer une solution à partir des connaissances et des éventuels instruments mis à sa disposition. Dans des activités de résolution créative de problèmes avec des robots pédagogiques, les outils à disposition sont des artefacts technologiques tangibles et interactifs. La matérialité et interactivité des artefacts sont à considérer dans le processus de résolution créative de problèmes [19].

4 Méthodologie

Nous analysons la résolution créative de problèmes à partir de la tâche CreaCube, qui engage le participant dans une tâche de problèmes sans un modèle de résolution connu d'avance. La tâche a été développée dans le cadre du projet ANR CreaMaker. Dans cette tâche, les apprenants doivent combiner les cubes robotiques de manière à pouvoir passer d'un point initial à un point final. Les apprenants doivent découvrir les affordances des cubes nécessaires pour résoudre la tâche [20]. Les affordances sont des représentations internes des caractéristiques potentielles des outils numériques et non numériques dans l'environnement. Ils sont définis comme la relation entre les propriétés physiques d'un objet donné et les interactions potentielles que l'apprenant peut avoir avec l'objet. Pour résoudre la tâche, il faut considérer à la fois les contraintes matérielles et l'objectif de la tâche.

4.1 Procédure

Les participants réalisent la tâche CreaCube deux fois : une première activité au cours de laquelle ils sont face à des objets inconnus et une deuxième activité qui comprend la même consigne. La première activité de résolution créative de problèmes engage nécessairement le participant dans l'exploration du matériel et la création de connaissances en acte pour arriver à résoudre la tâche. Lors de la deuxième activité, le participant dispose déjà d'un modèle de solution qu'il peut appliquer, ou bien décider d'explorer une nouvelle solution. Les deux activités sont enregistrées afin d'analyser les configurations réalisées par les participants au cours de la tâche. La fluidité est analysée comme le nombre de figures réalisées par le participant au cours d'une activité. La flexibilité considère les figures différentes réalisées par un même participant. L'innovation considère uniquement les 5% des figures les plus rares développées par les différents participants.

4.2 Participants

Un ensemble de 342 participants ont participé au protocole. Des participants à partir de 7 ans ont été engagés dans la tâche CreaCube. Nous les classons selon les catégories d'âge suivantes : les enfants (*enf* : 7 à 12 ans), les adolescents (*ado* : 13 à 18 ans), les jeunes adultes (*jad* : 19 à 29 ans), les adultes (*adu* : 30 à 59 ans), les adultes âgés (*sen* : 60 à 79 ans).

5 Résultats

Nous présentons les résultats de la première activité en termes de pensée divergente (fluidité, flexibilité, innovation) mais également la durée de la tâche, et ensuite les résultats de la deuxième activité.

5.1 Composantes de la pensée divergente selon l'âge au cours de la première activité

Dans le tableau 1 nous présentons les résultats des composantes de la pensée divergente mais également la durée de la tâche selon l'âge durant la première activité.

Table 1. Composantes de la pensée divergente selon l'âge. (Activité 1)

Catégorie d'âge	Effectif	Fluidité	Flexibilité	Innovation	Temps (secondes)
Enfants (enf)	48	m : 8.3 sd : 8.1	m : 3.6 sd : 2.3	m : 1.1 sd : 1.3	m : 291.1 sd : 220.0
Adolescents (ado)	5	m : 3.2 sd : 1.6	m : 2.0 sd : 0.7	m : 0 sd : 0	m : 112.6 sd : 64.7
Jeunes adultes (jad)	39	m : 4.0 sd : 5.0	m : 1.9 sd : 1.1	m : 0.3 sd : 0.5	m : 139.0 sd : 88.4
Adultes (adu)	29	m : 3.0 sd : 3.3	m : 2.0 sd : 1.6	m : 0.5 sd : 0.9	m : 119.0 sd : 72.4
Seniors (sen)	5	m : 4.4 sd : 3.5	m : 2.2 sd : 0.8	m : 0 sd : 0	m : 220.8 sd : 154.1

Les enfants présentent des scores plus importants sur l'ensemble des trois composantes créatives (fluidité, flexibilité et innovation) par rapport aux autres catégories d'âge, mais ils engagent également davantage de temps que les autres catégories d'âge. Nous effectuons ensuite une représentation graphique des résultats :

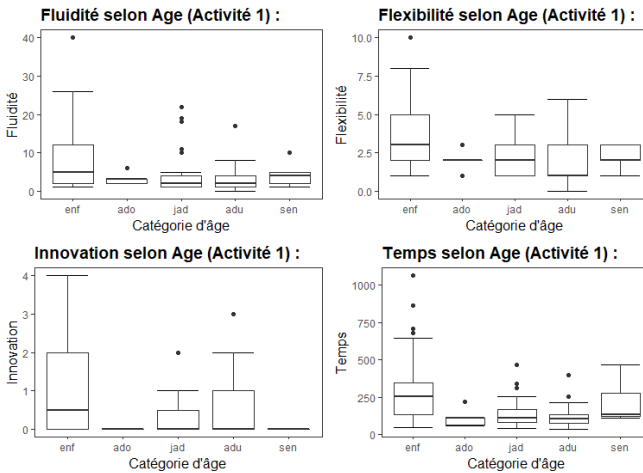


Fig. 1. Fluidité, flexibilité, innovation et temps durant l'activité 1.

L'asymétrie des boîtes étayant les tests visant à démontrer le caractère non-normal de la distribution statistique de nos populations étudiées. Par ailleurs, on remarque des différences significatives entre le groupe des enfants et les autres groupes.

Pour l'analyse des différences d'âge entre les différents groupes indépendants, nous réalisons le test de Kruskal-Wallis, en raison de la nature non paramétrique des données. Ce test a été conduit pour déterminer si l'âge avait un effet sur les composantes de la pensée divergente et le temps de réalisation d'une activité. Le test a révélé des différences significatives au cours de la première activité, $p=.0016$ (âge ~ fluidité), $p = .0014$ (âge ~ flexibilité), $p = .0074$ (âge ~ innovation) et $p = .0002$ (âge ~ temps de l'activité).

Des tests post-hoc de Dunn ont été réalisés pour situer les différences. Les différences étaient statistiquement significatives, selon le test de Dunn, entre les groupes d'âge "enfants/adultes" et "enfants/jeunes adultes", pour les composantes de fluidité et de flexibilité. Les différences étaient également significatives entre les enfants et toutes les autres catégories d'âge pour la composante d'innovation. Enfin, pour les temps d'activité, les différences étaient statistiquement significatives entre les groupes d'âge "enfants/adolescents", "enfants/jeunes adultes" et "enfants/adultes".

Ces résultats viennent confirmer les premières observations, émises lors de l'analyse des boîtes à moustaches, sur les différences entre les groupes.

5.2 Composantes de la pensée divergente selon l'âge au cours de la deuxième activité

Dans le tableau 2 nous présentons les résultats des composantes de la pensée divergente mais aussi la durée de la tâche selon l'âge durant la deuxième activité.

Table 2. Composantes de la pensée divergente selon l'âge. (Activité 2)

Catégorie d'âge	Effectif	Fluidité	Flexibilité	Innovation	Temps (secondes)
Enfants (enf)	48	m : 3.2 sd : 3.2	m : 2.0 sd : 1.7	m : 0.7 sd : 1.1	m : 117.0 sd : 69.8
Adolescents (ado)	5	m : 1.8 sd : 1.1	m : 1.2 sd : 0.4	m : 0 sd : 0	m : 77.2 sd : 42.9
Jeunes adultes (jad)	39	m : 2.1 sd : 1.9	m : 1.5 sd : 1.2	m : 0.2 sd : 0.6	m : 91.7 sd : 65.9
Adultes (adu)	29	m : 2.3 sd : 2.0	m : 1.7 sd : 1.0	m : 0.5 sd : 0.7	m : 93.5 sd : 51.5
Seniors (sen)	5	m : 1.6 sd : 0.9	m : 1.4 sd : 0.9	m : 0.2 sd : 0.4	m : 59.0 sd : 14.3

On remarque ici que les temps sont plus proches que précédemment. Par ailleurs, malgré ce rapprochement les composantes restent globalement plus élevées chez les enfants.

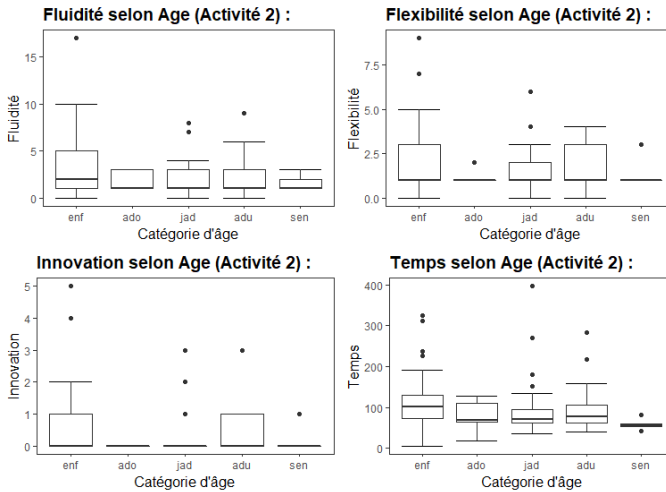


Fig. 2. Fluidité, flexibilité, innovation et temps durant l'activité 2.

La dispersion observée dans les boîtes à moustaches ci-dessus confirme un rapprochement statistique entre les groupes. De plus, l'asymétrie est, cette fois, moins marquée.

Le test de Kruskal-Wallis a, cette fois-ci, révélé des différences significatives pour l'influence de l'âge sur le temps de l'activité 2 seulement, $p = .0077$ ($\text{âge} \sim \text{temps de l'activité}$). Ceci suggère qu'en dehors du temps pris pour faire l'activité 2, les statistiques de composantes de pensée divergente entre les groupes sont plus proches comme pouvaient l'évoquer les boîtes à moustaches.

Le test post-hoc de Dunn réalisé a situé les différences statistiquement significatives entre les groupes d'âge "enfants/jeunes adultes", "enfants/adultes" et "enfants/seniors", suggérant une nouvelle fois une différence entre le groupe des enfants et les autres groupes.

6 Discussion

Dans le cadre de cette étude nous avons pu analyser les différentes composantes de la pensée divergente dans le cadre d'une tâche de résolution créative de problèmes. Nous observons que les enfants présentent à la fois des temps davantage longs mais aussi des scores plus importants dans les différentes composantes créatives. Nous pouvons considérer que les enfants sont moins orientés à performer en termes de temps, ce qui leur permet de s'engager dans une démarche davantage créative que les adolescents et adultes, qui ont développé la tâche de manière plus rapide et moins créative. Ces résultats peuvent être mis en perspective avec les études qui considèrent le lien entre le

temps de résolution de l'activité et la créativité. L'effet de la pression temporelle a des effets différentiels selon les participants [21]. Dans le cadre de la tâche CreaCube les enfants ne perçoivent pas la tâche sous une perspective de pression temporelle, mais comme une opportunité de créer des solutions créatives tant sur la première activité que sur la répétition de la tâche au cours de la deuxième activité. Ces résultats nous amènent à considérer l'importance d'engager les participants adolescents et adultes dans une démarche moins orientée à la performance temporelle dans le but de leur permettre de résoudre le problème de manière davantage créative.

Remerciements

Cette recherche est développée dans le cadre du projet Projet-ANR CreaMaker (18-CE38-0001). Nous remercions Eloïse Duhot-Prevot pour sa précieuse contribution à la plateforme d'analyse des vidéos.

References

1. C.-J. Chang *et al.*, « An analysis of student collaborative problem solving activities mediated by collaborative simulations », *Comput. Educ.*, vol. 114, p. 222-235, 2017.
2. J. Kilpatrick, J. Swafford, et B. Findell, *Adding it up: Helping children learn mathematics*. Citeseer, 2002.
3. L. DeBlois, S. Barma, et S. Lavallée, « L'enseignement ayant comme visée la compétence à résoudre des problèmes mathématiques: quels enjeux? », *Éducation Francoph.*, vol. 44, n° 2, p. 40-67, 2016.
4. R. E. Mayer et M. C. Wittrock, Éd., « Problem-solving transfer », in *Handbook of educational psychology*, New York : London: Macmillan Library Reference USA, Simon & Schuster Macmillan ; Prentice Hall International, 1996, p. 47-62.
5. M. Romero, L. DeBlois, et A. Pavel, « Créacube, comparaison de la résolution créative de problèmes, chez des enfants et des adultes, par le biais d'une tâche de robotique modulaire », *MathémaTICE*, n° 61, 2018, [En ligne]. Disponible sur: <http://revue.sesamath.net/spip.php?article1104>
6. G. Aldon et G. Olivier, « Un dispositif de recherche de problèmes de mathématiques au cycle 3 », *Repères IREM*, n° 108, p. 26-40, 2017.
7. N. L. Greenwald, « Learning from problems », *Sci. Teach.*, vol. 67, n° 4, p. 28-32, 2000.
8. E. P. Torrance et K. Goff, « A quiet revolution. », *J. Creat. Behav.*, 1989.
9. S. J. Parnes, R. B. Noller, et A. M. Biondi, *Guide to creative action*. MacMillan Publishing Company, 1977.
10. I. Capron Puozzo, *La créativité en éducation et formation: Perspectives théoriques et pratiques*, De Boeck. Louvain-la-Neuve, Belgium, 2017.
11. Y.-S. Lin, « Fostering creativity through education—a conceptual framework of creative pedagogy », *Creat. Educ.*, vol. 2, n° 03, p. 149, 2011.
12. P. Hyvönen, M. Romero, P. Hakkarainen, et N. Impiö, « Creative collaboration for enhancing older adult's ICT use », présenté à Biennial Conference of the European Association for Research in Learning and Instruction (EARLI), Munich, Germany,

août 2013.

13. M. D. Rama, H. de Ridder, et H. Bouma, « Technology generation and age in using layered user interfaces. », *Gerontechnology*, 2001.
14. E. P. Torrance, *Torrance tests of creative thinking*. Personnel Press, Incorporated, 1968.
15. J. P. Guilford, « Creativity: Its measurement and development », *Source Book Creat. Think.*, p. 151-167, 1962.
16. M. Besançon, B. Barbot, et T. Lubart, « Évolution de l'évaluation de la créativité chez l'enfant de Binet à nos jours », *Rech. Éducatives*, n° 5, p. 215-226, oct. 2011, doi: 10.4000/rechercheseducations.840.
17. R. Cera, C. Cristini, et A. Antonietti, « Conceptions of learning, well-being, and creativity in older adults », *J. Educ. Cult. Psychol. Stud. ECPS J.*, n° 18, p. 241-273, 2018.
18. V. Komis, M. Romero, et A. Misirli, « A Scenario-Based Approach for Designing Educational Robotics Activities for Co-creative Problem Solving », in *International Conference EduRobotics 2016*, 2016, p. 158-169.
19. A. Leroy, M. Romero, et L. Cassone, « Interactivity and materiality matter in creativity: educational robotics for the assessment of divergent thinking », *Interact. Learn. Environ.*, vol. 0, n° 0, p. 1-12, janv. 2021, doi: 10.1080/10494820.2021.1875005.
20. G. Kalmipourtzis et M. Romero, « Artifactual Affordances in Playful Robotics », in *Games and Learning Alliance*, Cham, 2020, p. 316-325. doi: 10.1007/978-3-030-63464-3_30.
21. A. L. Antes et M. D. Mumford, « Effects of Time Frame on Creative Thought: Process Versus Problem-Solving Effects », *Creat. Res. J.*, vol. 21, n° 2-3, p. 166-182, mai 2009, doi: 10.1080/10400410902855267.

Conception d'un Système Tutoriel Intelligent (STI) basé sur les progrès d'apprentissage : des étapes formelles aux étapes d'expérimentations chez les enfants

Hélène Sauzéon^{1,2}, Benjamin Clément¹, Cécile Mazon^{1,2}, Didier Roy¹ et Pierre-Yves Oudeyer¹

¹ Equipe FLOWERS, Inria, ENSTA Paris-Tech et Université de Bordeaux, F-33405, Talence, France

helene.sauzeon@inria.fr

² Equipe Handicap, Activité, Cognition et Santé, Université de Bordeaux - Inserm, Bordeaux Population Health Lab (U1219), F-33086 Bordeaux, France

Résumé. S'appuyant sur le modèle des Progrès d'Apprentissage (PA, Oudeyer & Gotlieb, 2016 [1]), une approche STI basée sur les PA a été conçue et implémentée dans un jeu sérieux de numératie (addition et soustraction de nombres dans des activités d'échanges monétaires) appelé KidLearn. Dans ce contexte, le STI personnalise les parcours d'apprentissage en optimisant les PA (algorithmes de type bandit multi-bras), en minimisant les présupposés sur les apprenants et le domaine de connaissance. Dans un premier temps, l'approche STI adoptée est présentée, suivie de sa validation par une étude à grande échelle auprès d'enfants de 6 à 8 ans. Puis, la capacité de généralisation de cette approche est illustrée à travers son transfert à un public cognitivement hétérogène (i.e. éducation spécialisée), et son transfert à un autre domaine d'apprentissage (i.e., éducation thérapeutique).

Mots-clés : STI, progrès d'apprentissage, motivation intrinsèque, études terrains.

Abstract. Based on the Learning Progression (LP) model of Oudeyer and Gotlieb, 2016 [1], an LP-based ITS approach was designed and implemented in a serious numeracy game (addition and subtraction of numbers in money exchange activities) called KidLearn. In this context, the ITS personalizes learning paths by optimizing LPs (multi-armed bandit-like algorithms), minimizing presuppositions about learners and knowledge domain. First, the adopted ITS approach is presented, followed by its validation through a large-scale study with 6-8 year old children. Then, the generalizability of this approach is illustrated through its transfer to a cognitively heterogeneous children (i.e., special education), and its transfer to another learning domain (i.e., health education).

Keywords: ITS, Learning progress, Intrinsic motivation, field studies.

1 Introduction

L'un des défis de l'école du 21^e siècle est de faire face à la grande diversité des apprenants en termes de capacités d'apprentissage et de motivations. Les nombreuses recherches sur l'éducation personnalisée et l'enseignement actif témoignent de cet enjeu sociétal critique dont l'objectif *in fine* est d'assurer l'égalité des chances à l'école.

Bien que les atouts de l'apprentissage personnalisé soient aujourd'hui bien documentés (e.g., [2], [3]), sa généralisation est défi face à la taille élevée des classes. C'est pourquoi de fortes attentes sont associées aux technologies éducatives *adaptatives* pour l'automatiser et soutenir les enseignants dans leurs missions. Parmi elles, les Systèmes tutoriels Intelligents (STI) occupent une place privilégiée puisqu'ils ont la particularité d'individualiser l'enseignement en tenant compte de multiples caractéristiques de l'élève et/ou d'offrir un ajustement du parcours pédagogique selon les réponses au fil de l'eau de l'apprenant.

Dans le présent travail, nous présentons le projet Kidlearn dont les objectifs sont les suivants : 1) Optimiser et personnaliser automatiquement les parcours d'apprentissage dans les STI, en minimisant les présupposés sur les apprenants et le domaine de connaissance ; 2) Proposer une nouvelle approche avec des algorithmes de type bandit multi-bras; et 3) Optimiser l'engagement des apprenants en s'appuyant sur le modèle des progrès d'apprentissage [1] en lien avec la théorie des motivations intrinsèques [4] [5].

2 Etat-de-l'art

D'une manière générale, deux grandes familles de conceptions des STI pour la personnalisation peuvent être distinguées [6]. La première, à savoir la personnalisation dite "linéaire", implique que tous les apprenants suivent un même chemin d'apprentissage, mais à des vitesses différentes ou avec un nombre différent d'exercices. La seconde, qualifiée "à chemins ramifiés", permet à chaque apprenant de suivre un chemin spécifique en fonction de ses réponses.

Plusieurs revues systématiques ont rapporté des résultats prometteurs, voire positifs, sur la valeur ajoutée des STI (e.g., [7] [8]), tout en soulignant les limites méthodologiques de ce champ de recherche (e.g., pas de groupe contrôle, pas d'équivalence initiale des groupes sur le critère de jugement d'efficacité de l'intervention STI, pas de mesures pré et post-intervention, voir par exemple, [9]) ainsi qu'une grande variété de contextes d'utilisation de ces STI (e.g., par exemple, les STI peuvent être utilisés seuls ou mixés à une intervention pédagogique menée par l'enseignant). Au final, les revues systématiques concluent qu'il est difficile aujourd'hui d'identifier les caractéristiques propres au STI qui sont essentielles à la réussite de l'apprentissage personnalisé [10] [6].

De manière plus ou moins explicite, les STI s'appuient sur le concept de la Zone Proximale de Développement (ZPD, [11]) où il s'agit de proposer à l'apprenant des activités pédagogiques ni trop difficiles, ni impossibles au regard de ses capacités afin qu'il puisse progresser dans ses acquisitions. Le concept ZPD est un concept bien connu

en psychologie du développement qui a inspiré de nombreux modèles d'apprentissage (par exemple, la théorie de la charge cognitive, [12]).

Récemment, avec l'intérêt croissant pour le phénomène de la curiosité (e.g. pour revue, [13]), ce concept a été revisité dans le modèle des Progrès d'Apprentissage (PA) [1]. Ce modèle propose que la ZPD conditionne PA mais également la motivation intrinsèque. Le modèle PA est un modèle contemporain d'apprentissage par récompense où la motivation intrinsèque et l'apprentissage sont liés par une boucle vertueuse : les tâches adaptées aux capacités de l'enfant sont plus motivantes et génèrent des PA qui, en retour, sont une récompense interne stimulante pour de nouvelles acquisitions.

Il est intéressant de noter que le modèle LP s'appuie également sur les théories des motivations intrinsèques, et en particulier, la théorie de l'autodétermination [4] [5], qui pose les motivations intrinsèques comme l'ingrédient essentiel à l'apprentissage. Les motivations intrinsèques se rapportent aux comportements ou aux activités réalisés pour le bien de l'individu ou pour satisfaire ses propres besoins de développement. Elles s'opposent aux motivations extrinsèques plus hétérogènes dans leur nature et, qui renvoient aux comportements motivés pour des raisons autres que les intérêts et désirs inhérents à l'individu (exigences environnementales, attentes ou récompenses sociales, etc.).

Sur la base du modèle des PA, le projet Kidlearn vise à développer un STI dont la particularité du mécanisme automatique de personnalisation est d'optimiser les progrès d'apprentissage de l'apprenant au fil de l'eau de ses réponses aux activités proposées. A ce titre, Kidlearn appartient à la famille des STI à personnalisation ramifiée.

3 Le Projet KidLearn

Dans un premier temps l'approche STI adoptée et les algorithmes développés sont présentés suivis de leur validation à travers une étude à grande échelle réalisée auprès d'enfants de 6 à 8 ans. Puis, la capacité de généralisation de cette approche est illustrée à travers son transfert à un public cognitivement hétérogène (i.e. étude en éducation spécialisée), et son transfert à un autre domaine d'apprentissage (i.e., éducation thérapeutique d'enfants avec asthme chronique).

3.1 Approche STI, algorithme ZPDES et Etude à grande échelle

Notre stratégie est de proposer à chaque apprenant les activités maximisant ses progrès. Ces activités ne sont pas nécessairement celles définies a priori dans les modèles cognitifs et apprenants, ce sont celles repérées en temps réel, grâce aux résultats de l'étudiant, comme étant les plus efficaces pour celui-ci. Cette approche présente trois principaux avantages :

1. *Une faible dépendance vis à vis des modèles cognitifs et apprenant.* Il est souvent très difficile pour l'enseignant d'identifier précisément les points forts et les points faibles de l'élève, et donc de déterminer quelles activités seront profitables. De là, un premier prérequis est que le STI explore et expérimente diverses activités afin d'évaluer leur potentiel didactique pour faire progresser l'apprenant concerné.

2. *La possibilité d'utiliser des méthodes efficaces d'optimisation*, qui ne font pas d'hypothèses spécifiques sur la façon dont les élèves apprennent et qui ont seulement besoin d'informations sur la progression de l'apprentissage estimée pour chaque activité. Nous faisons la simple hypothèse que les activités et leurs paramètres qui sont repérés comme fournissant un bon gain d'apprentissage, doivent être choisis prioritairement. Un formalisme efficace et bien étudié pour ce type de problèmes est celui du bandit multi-bras. Suivant une analogie avec des bandits manchots dans un casino, à chaque étape du processus d'optimisation nous choisissons une des machines à sous, nous mison puis nous observons le gain qui s'ensuit, l'objectif étant de découvrir le meilleur bras. Nous sommes dans un processus d'exploration/ expérimentation.

3. *Une motivation accrue*. Dans notre approche, les activités choisies doivent être en permanence celles qui possèdent le meilleur gain d'apprentissage pour être motivantes. Cette théorie des motivations intrinsèques renvoie à la zone proximale de développement (ZPD).

Dans le cadre de simulations d'étudiants virtuels, nous avons construit le logiciel Kidlearn sur l'apprentissage de l'utilisation de la monnaie (lecture, composition, addition, soustraction de nombres entiers ou décimaux) et implémenté trois approches :

- une séquence d'exercices prédéfinie construite par un expert en didactique,
- la même séquence optimisée par un algorithme, RiARiT, utilisant des informations préalables sur la difficulté et les caractéristiques des tâches proposées,
- la même séquence optimisée par un autre algorithme, ZPDES, sans information sur les tâches.

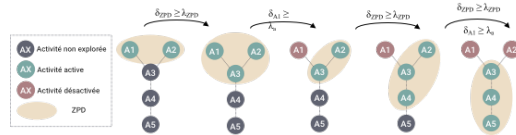


Fig. 1. Graphe de fonctionnement de ZPDES.

Nos simulations concernaient une population de 1000 étudiants, chaque étudiant étant amené à résoudre 100 exercices afin d'évaluer comment et à quelle vitesse chaque algorithme estimait et proposait les activités les plus pertinentes aux étudiants. RiARiT et ZPDES se sont avérées plus efficaces que la séquence prédéfinie ; les algorithmes pouvaient effectuer une réorganisation dynamique et modifier les paramètres des exercices, en exploitant pour cela leur exploration des résultats de l'élève [14].

Après une première expérimentation en situation réelle avec 100 élèves de CE1 ayant donné des résultats similaires aux simulations, nous avons réalisé une étude à grande échelle auprès d'élèves de CE1, basée, sur le logiciel Kidlearn en y incluant des activités plus variées et élaborées (Figure 2). Nous avons comparé deux conditions de personnalisation : une condition de personnalisation linéaire appelée séquence pré-définie (établie par un dictateur pour constituer la condition contrôle) et une condition de personnalisation basée sur les PA, à l'aide de l'algorithme ZPDES (selon deux versions avec ou sans optimisation des PA). Chacune de ces conditions de personnalisation était subdivisée en deux autres conditions, l'une donnant à l'élève la possibilité de choisir l'objet de la transaction monétaire, et l'autre lui imposant l'objet de la transaction. Cette dernière manipulation visait à évaluer l'impact d'une prise de décision motivée par un choix personnel, sans pour autant impacter la chaîne des opérations calculatoires à opérer par l'enfant.



Fig. 2. Visuels de l'application KidLearn

Pour évaluer les progrès des élèves au sein de l'application Kidlearn, nous avons comparé au cours des essais successifs le niveau de difficulté atteint dans l'espace tâche ; ainsi que le niveau de réussite des élèves (Figure ci-dessous) [15]. A cela, des évaluations pré-/post-interventions ont été ajoutées pour renseigner les progrès réels des élèves et leur motivation selon 5 stratégies.

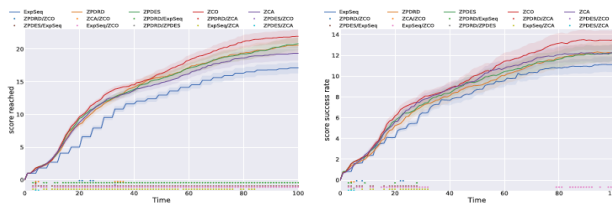


Fig. 3. Evolution dans le temps des Scores atteints et Scores de succès selon les 5 stratégies proposées (Exseq- Séquence pré-définie ; ZPDRD- ZPDES sans optimisation ; ZPDES - ZPDES avec choix d'objets ; ZCO - ZPDES avec choix d'activité Achat/Vente).

3.2 Etudes de Généralisation de notre approche STI

Fort de ces résultats, pour en vérifier la généralisation possible, nous avons opté pour deux cas d'étude : 1) éprouver l'approche auprès d'apprenants atypiques (i.e., résistance à la diversité cognitive des apprenants) ; et 2) la tester dans un autre domaine de connaissance (i.e., transfert inter-domaine).

Nous avons ainsi proposé la version ZPDES de KidLearn à des enfants avec autisme scolarisés en Unité Locale d'inclusion Scolaire (Mazon et al., *Article en préparation*). Du fait de la forte variabilité cognitive dans l'autisme, de nombreux enfants et adolescents éprouvent des difficultés dans l'apprentissage des mathématiques, et sont même en retard sur leurs camarades à l'école. Une étude de faisabilité a alors été menée auprès de 24 élèves (13-14 ans) atteints d'autisme avec déficience intellectuelle (QI autour de 60, équivalent à un âge mental de 7-8 ans) : 14 élèves ont reçu l'application KidLearn-ZPDES, et les dix autres une application « contrôle » (jeu sans calcul). Les mesures de calcul et surtout de motivation intrinsèque ont été améliorées dans la condition KidLearn comparé à la condition contrôle. Aussi, l'amélioration en calcul était de même amplitude que celle observée pour les enfants de 6-9 ans. Pris ensemble, cela indique que ZPDES est une intervention prometteuse pour l'apprentissage du calcul dans le cadre de l'éducation spécialisée.

Dans le cadre d'un programme d'éducation thérapeutique d'enfants asthmatiques de 8 à 11 ans, le logiciel KidBreath a été conçu et déployé pendant un mois auprès de 29 enfants en condition réelle d'usage et dans un environnement non contrôlé, soit selon une séquence prédéfinie de progression d'apprentissage, soit selon une séquence personnalisée avec ZPDES [16]. Pour chaque condition, l'impact de la pratique de KidBreath a été évalué en termes d'efficacité pédagogique (connaissance de la maladie), et d'efficience pédagogique (e.g., type de motivation suscitée). En condition personnalisée avec ZPDES, l'efficacité pédagogique a été légèrement moins concluante que la condition prédéfinie : bien que les enfants améliorent leur connaissance générale sur l'asthme, cette amélioration n'est significative. De manière intéressante, le niveau atteint de connaissance avec ou sans personnalisation était similaire alors que le nombre de contenu et le temps consacré dans la condition « ZPDES » étaient bien inférieurs à ceux observés dans la condition prédéfinie. Aussi, un plus fort impact de la condition ZPDES est observé sur la motivation intrinsèque. Ainsi, l'apprentissage personnalisé via ZPDES a produit une efficience pédagogique accrue en termes de motivation intrinsèque, de durée d'apprentissage et d'étendue des contenus explorés, mais sans augmentation significative de l'efficacité pédagogique comparé à la condition prédéfinie. Une durée d'apprentissage plus longue qu'un mois aurait sans doute permis de mieux démontrer la meilleure efficacité pédagogique de ZPDES, du fait de l'impact de la motivation dans le temps.

4 Conclusion et Perspectives

En conclusion, l'ensemble des étapes expérimentales de validation de notre approche STI de personnalisation dirigée par les progrès d'apprentissage (ZPD) produit des résultats prometteurs en termes à la fois, d'efficacité (progrès observés pré- et post-intervention) et d'efficience (motivation suscitée post-intervention, notamment) pédagogique. Et, le succès de son transfert multi-domaines et de généralisation à la diversité cognitive sont des résultats encourageants. Dans une visée éducative, l'algorithme ZPDES fait aujourd'hui l'objet d'une intégration dans une suite d'outils d'apprentissage des mathématiques (programme « adaptative math », <https://www.adaptivemath.fr/team>). Enfin, pour aller plus loin dans sa généralisation, il fait aussi l'objet d'un projet applicatif au domaine de la rééducation cognitive visant un public varié en âge et en profil cognitifs.

References

1. Oudeyer, P. Y., Gottlieb, J., & Lopes, M. Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. *Progress in brain research*, 229, 257-284 (2016).
2. Deunk, M., Doolaard, S., Smale-Jacobse, A., & Bosker, R. J. Differentiation within and across classrooms: A systematic review of studies into the cognitive effects of differentiation practices. Groningen: Gion Onderwijs/Onderzoek (2015).
3. Iterbeke, K., De Witte, K., & Schelfhout, W. The effects of computer-assisted adaptive instruction and elaborated feedback on learning outcomes. A randomized control trial. *Computers in Human Behavior*, 106666 (2020).
4. Ryan, R. M., & Deci, E. L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68 (2000).
5. Ryan, R. M., & Deci, E. L. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860 (2020).
6. Bartolomé, A., Castañeda, L., & Adell, J. Personalisation in educational technology: the absence of underlying pedagogies. *International Journal of Educational Technology in Higher Education*, 15(1), 1-17 (2018).
7. Alevén, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. Instruction based on adaptive learning technologies. In R. E. Mayer, & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 522-560), Routledge (2017).
8. Faber, J. M., Luyten, H., & Visscher, A. J. The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & Education*, 106, 83-96 (2017).
9. Cheung, A. C. K., & Slavin, R. E. The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review*, 9, 88-113 (2013).
10. Hew, K. F., & Cheung, W. S. (2011, July). Using Web 2.0 Technologies in K-12 School Settings: Evidence-Based Practice?. In International Conference on ICT in Teaching and Learning (pp. 319-328). Springer, Berlin, Heidelberg (2011).
11. Vygotsky, L. S. Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press (1978).

12. Sweller, J. Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37-76). Academic Press (2011).
13. Murayama, K., FitzGibbon, L., & Sakaki, M. Process account of curiosity and interest: A reward-learning perspective. *Educational Psychology Review*, 31(4), 875-895 (2019).
14. Clement B., Roy D., Oudeyer P.-Y., Lopes M., Multi- Armed Bandits for Intelligent Tutoring Systems, *Journal of Educational Data Mining* (JEDM) (2015).
15. Clément, B. Adaptive Personalization of Pedagogical Sequences using Machine Learning. Thèse - <https://hal.inria.fr> (2018).
16. Delmas, A., Clement, B., Oudeyer, P. Y., & Sauzéon, H. (2018). Fostering health education with a serious game in children with asthma: pilot studies for assessing learning efficacy and automatized learning personalization. *Frontiers in Education*, 3, p. 99 (2018).

Table Ronde : “Conduite de la recherche sur les EIAH - Ancrages épistémologiques et méthodes de recherche”

Nadine Mandran ¹, Éric Sanchez ²

¹ Grenoble Alpes Université, CNRS, France

² Université de Genève, TECFA, 1211 Genève 4, Suisse

Description

Cette table ronde, organisée par le Groupe de travail¹ « Epistémologie et Méthodes de conduite de la recherche » de l'ATIEF, a pour objectif d'informer sur les concepts liés à l'épistémologie du chercheur et la cohérence entre l'épistémologie et les méthodes, et d'échanger sur les différents courants méthodologiques présents dans la communauté des EIAH et de l'importance de ces éléments pour la constitution d'un projet.

En préambule et pour montrer l'importance d'être clairs sur nos méthodes de recherche, nous ferons un retour d'expérience sur la participation de certain-e-s d'entre nous à des comités d'évaluation des projets scientifiques déposés à l'ANR.

Participants

La table ronde sera animée par Eric Sanchez, professeur en technologies éducatives au TECFA / Université de Genève, et fera intervenir :

- Catherine Thomas, Professeur en sciences de gestion, Université Nice Côte d'Azur
- Nathalie Guin, Maitre de conférences HDR en Informatique, LIRIS, Université Claude Bernard - Lyon 1
- Elise Lavoué, Maitre de conférences HDR en Informatique, LIRIS, Université Jean Moulin Lyon 3
- Gaëlle Molinari, Professeur assistante en Psychologie de l'éducation et en Technologies éducatives, Université de Genève & UniDistance
- Catherine Bonnat, Collaboratrice scientifique, Didactique, TECFA Université de Genève
- Laëtitia Pierrot, Chercheure Postdoc, Sciences de l'information et de la communication, Université de Poitiers et Montpellier
- Nadine Mandran, Ingénieure de Recherche en méthodologie, LIG, Université de Grenoble
- Amel Yessad, Maitre de conférences en Informatique, LIP6, Sorbonne Université

¹ <https://lig-gt-epistemology.imag.fr/>

Les ateliers thématiques participatifs

Laëtitia Pierrot¹ et Rémi Venant²

¹ Université de Poitiers, TECHNE EA 6316, F-86073 Poitiers, France
laetitia.pierrot@univ-poitiers.fr

² Le Mans Université, LIUM EA 4023, F-72085 Le Mans, France
remi.venant@univ-lemans.fr

1 Introduction

La conférence EIAH 2021 a été un moment de rencontres et d'échanges scientifiques, en particulier dans le cadre d'ateliers et de symposiums thématiques. Organisés autour de thématiques spécifiques, ces derniers visent à identifier le périmètre actuel des recherches et à en dessiner les axes d'exploration futurs.

Cette année, trois grandes thématiques se sont distinguées par un atelier et cinq symposiums, sur les huit soumissions initiales (deux ateliers ont été annulés par manque de participants). Les interactions homme-machine ont été abordées à travers le prisme de l'hybridation des enseignements, prise dans sa globalité, mais aussi dans l'étude des relations entre environnements de simulation (réalité virtuelle, augmentée ou mixte) et *embodiment*. La conception des EIAH eux-même a aussi été à l'étude, focalisée d'une part sur les tableaux de bord dans une approche collaborative, et sur la prise en compte dans ces environnements de l'adaptation et la génération automatique d'autre part. Enfin, l'enseignement du numérique lui-même a attiré l'attention, d'une part analysé dans son périmètre le plus large par les humanités numériques, d'autre part traité pour l'apprentissage de l'informatique, le rapprochement des problématiques se manifestant par l'organisation d'une table ronde commune.

Avec une participation importante (le nombre d'inscriptions moyen étant de 90), ces ateliers ont suscité un fort intérêt et s'assurent une pérennité, également de part leur ancrage dans des groupes de travaux préexistant comme les GT ATIEF, GIS2if, EduIHM ou encore APIMU.

1 Les interactions médiées pour l'apprentissage

Symposium "Enseignement et apprentissage en situation(s) hybride(s) : vers une réinvention de la forme scolaire ?"¹. Organisé dans le cadre du groupe de travail GTNUM#InteractionHybridation et impliquant le GT EduIHM, soutenu par l'ATIEF et l'AFIHM, par Stéphanie Fleck (Université de Lorraine), Audrey Serna (INSA de Lyon), Franck Silvestre, (Université Toulouse Capitole), Sébastien Kubicki (Ecole Nationale d'Ingénieurs de Brest) et Luc Massou (Université de Lorraine). Ce symposium

¹ <https://projet.liris.cnrs.fr/InteractionHybridation/#about>

a exploré les formes et scénarios d'hybridation en jeu en contexte d'enseignement-apprentissage, travaillé sur les différentes définitions possibles de l'hybridation scolaire avant d'en interroger les enjeux.

Symposium “Simulation numérique et apprentissage humain : le rôle de la cognition incarnée”². Organisé par Rawad Chaker (Université Lyon 2) et Nady Hoyek (Université Lyon 1), ce symposium a traité des liens entre les environnements de simulation numérique et la cognition incarnée (ou *embodiment*) comme instance, ou dimension mobilisée pour l'apprentissage humain, et perçue comme paradigme essentiel pour la communauté EIAH.

2 La conception des EIAH

Symposium “Outil de Conception participative de tableaux de bord d'apprentissage”³. Organisé par Jean-Marie Gilliot (IMT Atlantique), Laurent Neyssensas (Ecole de Design de Nantes Atlantique) et Madjid Sadallah (IMT Atlantique), ce symposium a eu pour objectif d'aborder le développement d'outils permettant l'appropriation des Learning Analytics par les utilisateurs potentiels et la capture des besoins utilisateurs.

Journée de travail “Adaptation et génération dans les EIAH”⁴. Organisée dans le contexte d'un GT ATIEF par Pierre Laforcade (Le Mans Université), Marie Lefevre (Université Lyon 1) et Sébastien Jolivet (Université Grenoble Alpes), la journée de travail a traité de la complexité et la variété des activités d'adaptations dans le cadre de la génération d'exercices adaptés à l'apprenant.

3 L'apprentissage du numérique

Symposium HUMANE “Humanités numériques, éducation et formation”⁵. Organisé dans le contexte du GTNUM GIS2IF #HN_éducation par Béatrice Drot-Delange (Université Clermont-Auvergne) et Thibaud Hulin (Université Franche-Comté), ce symposium s'est intéressé aux humanités numériques dans le champ de l'éducation et de la formation, pour décrire et partager les pratiques enseignantes et pour analyser les conceptions du numérique mises en œuvre.

Atelier APIMU “Apprentissage de la pensée informatique de la maternelle à l'Université : retours d'expériences et passage à l'échelle”⁶. Organisé par Julien Broisin (Université de Toulouse), Christophe Declercq (INSPÉ de Nantes), Cédric Fluckiger (Université de Lille), Yannick Parmentier (Université de Lorraine), Yvan Peter (Université de Lille) et Yann Secq (Université de Lille), cette 5ème édition de l'atelier a abordé la problématique de l'enseignement de la pensée informatique de la maternelle à l'université et de la formation et de l'accompagnement des enseignants.

² <https://cieseiah.hypotheses.org/>

³ <https://padlad.github.io/EIAH2021/>

⁴ <https://projets-lium.univ-lemans.fr/gt-adaptation/eiah21/>

⁵ <https://humane-eiah-21.sciencesconf.org/>

⁶ <https://hal-udl.archives-ouvertes.fr/APIMU-2021>

Formation Doctorale : Prendre du recul sur son travail de thèse, se poser quelques questions et échanger avec des membres de la communauté

Catherine Bonnat¹, Sébastien Jolivet², Nadine Mandran³, Rémi Venant⁴

¹ Université de Genève, TECFA, Genève, Suisse
Catherine.bonnat@unige.ch

² Sorbonne Université, LIP6, Paris, France
sebastien.jolivet@lip6.fr

³ Université Grenoble-alpes, CNRS, Grenoble, France
nadine.mandran@univ-grenoble-alpes.fr

⁴ La Mans Université, LIUM, Le Mans, France
reми.venant@univ-lemans.fr

1 Pour qui ?

Cette session s'adresse à tous les doctorant-e-s, en début comme en fin de thèse. Elle se veut un temps de formation scientifique et méthodologique. Elle est destinée à des chercheurs de toutes les disciplines (STIC ou SHS) dont le travail est en lien avec les EIAH. Pour les doctorants, en plus de pouvoir bénéficier d'un regard expert sur leurs travaux, c'est aussi un moment d'intégration dans la communauté EIAH.

2 Principe de la session

Le principe est de permettre aux doctorants de présenter l'avancement de leurs travaux, de (se) poser des questions sur les points importants et de bénéficier du retour de leurs pairs et de chercheurs confirmés, à la fois dans leur domaine mais aussi de manière plus transversale sur des aspects méthodologiques. L'objectif est de favoriser des échanges assez approfondis, pour cela des modalités spécifiques ont été élaborées, articulant une soumission écrite, la relecture d'une autre soumission et une présentation orale. Les doctorants ont ainsi l'occasion de se former à différents aspects du métier de chercheur : présenter leur travail à l'écrit et à l'oral, interagir et réagir suite à la présentation et relire des travaux d'un autre chercheur. Diverses ressources leur ont été proposées pour les accompagner et les guider dans ces différents aspects, ressources qui restent pertinentes au-delà de la session pour un large public et sont accessibles depuis le site de la conférence.

**Textes des communications
retenues**

Articles longs

Amélioration de l'interprétabilité du diagnostic cognitif de l'apprenant par catégorisation des composantes de connaissance

Olivier Allègre^{1,2}, Amel Yessad¹, and Vanda Luengo¹

¹ MOCAH, LIP6, Sorbonne Université, Paris, France

² Kartable, www.kartable.fr

Abstract. Le diagnostic cognitif permet de modéliser les connaissances d'un apprenant à partir de ses interactions avec un EIAH pour la recommandation adaptative de ressources ou la prédiction de l'état futur de l'apprenant. En étudiant les performances des algorithmes des modèles *knowledge tracing* et *factor analysis* sur des jeux de données restreints aux composantes de connaissance (CC) déclaratives ou procédurales, nous montrons que les modèles *factor analysis* sont plus performants sur des CC déclaratives, alors que les modèles *knowledge tracing* réagissent mieux sur les CC procédurales, même s'il est compliqué de comparer les deux modèles l'un à l'autre. Ces résultats suggèrent l'intérêt d'une catégorisation des CC par leur nature dans le choix des algorithmes de diagnostic cognitif, et que cette catégorisation permet l'amélioration de l'interprétabilité du modèle de l'apprenant.

Keywords: Apprentissage adaptatif · Diagnostic cognitif · Fouille de données éducationnelles · Modélisation de l'apprenant.

Abstract. Cognitive diagnosis is used to model the learner knowledge from the learners' traces on an ITS for adapted feedback or prediction of learner future state. By studying the performance of algorithms of *knowledge tracing* and *factor analysis* models on datasets restricted to either declarative or procedural knowledge components (KC), we show that *factor analysis* models outperform on declarative KCs while *knowledge tracing* ones outperform on procedural KCs, even if it is complicated to compare them. These results suggest the interest of categorizing the KCs by nature in cognitive diagnosis algorithms and that this categorization allows the enhancement of learner model interpretability.

Keywords: Adaptive learning · Cognitive diagnosis · Educative data mining · Learner modeling.

1 Introduction

Personnaliser l'enseignement est une des promesses de l'intelligence artificielle en éducation. Par l'intermédiaire de la modélisation de l'apprenant, il est possible de proposer aux élèves des ressources adaptées à leur niveau, dans un

but d'apprentissage ou d'évaluation. La visée du modèle de l'apprenant est de déterminer l'état dans lequel ce dernier se situe. Via la modélisation des capacités cognitives, de la motivation ou encore de la méta-cognition, l'état latent prédit de l'apprenant peut permettre d'adapter la rétroaction et la recommandation du EIAH.

Nous ne nous intéresserons dans ce papier qu'à la modélisation du savoir et des capacités de l'apprenant, très utilisée dans les systèmes éducatifs adaptatifs [11]. Afin de définir de tels modèles, l'usage le plus courant est de définir et de modéliser le contenu d'un domaine d'étude, par l'intermédiaire de composantes de connaissance (CC) qui correspondent aux constituants du savoir (e.g. "Connaître la table de multiplication de 3" ou "Accorder un adjectif régulier dans un groupe nominal minimal"). Les CC sont associées aux ressources pédagogiques consultables par l'apprenant ainsi qu'aux exercices permettant de l'évaluer. À l'aide d'algorithmes de diagnostic cognitif, il est enfin possible de déterminer des métriques de maîtrise de ces CC à partir des traces des apprenants sur les exercices. Beaucoup de caractéristiques de l'apprenant sont modélisables [2] et, même au sein de la modélisation de son savoir, il est possible d'adopter des stratégies de diagnostic cognitif extrêmement diverses. Le choix du modèle n'est pas simple : il est nécessaire d'avoir à la fois de bonnes performances sur la modélisation de l'apprenant et, spécifiquement dans le domaine de l'éducation, une interprétabilité du modèle par les enseignants [12] [10] et parfois même par les apprenants [4].

Cohen et Squire [5] introduisent en 1980 le terme de mémoire déclarative, mémoire du conscient incluant principalement des faits et des événements, qu'ils opposent à la mémoire procédurale, inconsciente cette fois, permettant d'acquérir capacités cognitives et motrices. En différenciant la nature déclarative ou procédurale du savoir de l'apprenant, nous pouvons donc émettre l'hypothèse que les CC, elles-mêmes, correspondent à l'une ou à l'autre. Nous nous interrogerons dans ce papier sur l'impact du choix des algorithmes de diagnostic cognitif sur le modèle de l'apprenant et ses constituantes procédurales et déclaratives. La catégorisation en nature des CC permet-elle une meilleure interprétabilité du diagnostic cognitif ?

Dans un premier temps, nous étudions la modélisation de l'apprenant et nous focalisons sur les algorithmes permettant la prédiction du savoir de l'apprenant. Nous détaillons ensuite les paramètres de ces algorithmes. Enfin, nous présentons l'étude expérimentale que nous avons menés sur des jeux de données de mathématiques et de français filtrés par la nature des CC et nous discutons les résultats obtenus.

2 Travaux associés

Le choix des éléments modélisés dans la modélisation de l'apprenant est large, mais nous ne nous intéressons qu'aux modèles basés sur les connaissances de l'apprenant, par le biais d'un modèle du domaine sous forme de décomposition

en CC. Cette structure de modèle de l'apprenant est représentée dans la Figure 1.

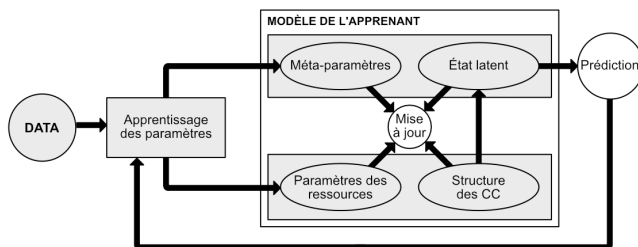


Fig. 1. Structure du modèle de l'apprenant, illustration inspirée de [11] : le modèle de l'apprenant comporte des équations de mise à jour de prédiction afin de déterminer l'état latent de l'apprenant à partir des paramètres de l'apprenant (méta-paramètres) et des ressources, eux même appris à partir des données issues des traces des apprenants.

La majorité des systèmes éducatifs adaptatifs présentés dans la littérature sont basés sur ce type de modèle [11] [1]. Chaque système s'appuie sur des équations de mise à jour et de prédictions spécifiques et des paramètres liés à l'apprenant (méta-paramètres) ou au domaine (paramètres des ressources), qui rendent possible l'interprétation du modèle de l'apprenant.

Nous nous focalisons sur les modèles *knowledge tracing* et *factor analysis* dans ce papier. On indexe les apprenants par $i \in \llbracket 1, I \rrbracket$, les exercices par $j \in \llbracket 1, J \rrbracket$ et les CC par $k \in \llbracket 1, K \rrbracket$. On note $Y_{i,j} \in \{0, 1\}$ la justesse de la réponse, binaire, de l'élève i à l'exercice j , p_{ij} la probabilité que l'élève i réponde correctement à la question j et σ la fonction logistique.

2.1 Knowledge tracing

Le modèle *knowledge tracing* repose sur la modélisation des processus d'apprentissage et d'oubli. Le *bayesian knowledge tracing* (BKT), un modèle caché de Markov, est le modèle *knowledge tracing* le plus répandu, où l'apprentissage est modélisé par une transition d'un état de non maîtrise d'une composante de connaissance à un état de maîtrise [6] à l'aide de quatre paramètres : P_i la probabilité a priori, P_l la probabilité de transition entre l'état de non-maîtrise à l'état de maîtrise, P_g la probabilité de deviner la bonne réponse à un exercice sans maîtriser la CC associée, P_s la probabilité de ne pas répondre correctement à un exercice malgré la maîtrise de la CC associée. L'apprentissage de tous

ces paramètres est généralement exécuté par un algorithme de maximisation de l'espérance. Nous considérons dans ce papier la version classique de BKT (que nous notons BKT) et trois de ses variantes :

- IDE [9], dans lequel les probabilités P_g et P_s sont dépendantes des exercices
- ILE [13], qui considère la probabilité P_i dépendante de l'exercice
- PPS [8], dans lequel la probabilité P_i dépend de la justesse de la première tentative de l'apprenant

2.2 Factor analysis

Le modèle *factor analysis* est fondé sur l'analyse de certaines observations par le biais d'un modèle logistique. Il permet de déterminer les facteurs communs parmi les interactions de l'ensemble des apprenants sur les ressources pédagogiques. Le choix des observations diffère entre les algorithmes du modèle *factor analysis* et mène à un ensemble de paramètres sur les apprenants et les ressources. Nous étudions ici quatre algorithmes de *factor analysis*, dont on précise les paramètres et l'équation de mise à jour.

- L'algorithme *item response theory* (IRT) modélise la capacité θ_i de l'apprenant i ainsi que la difficulté d_j de la question j .

$$p_{ij} = \sigma(\theta_i \cdot d_j - \delta_j) \quad (1)$$

- L'algorithme *additive factor model* (AFM) modélise l'apprenant à l'aide de la difficulté β_k de la CC k , du biais γ_k de maîtriser la CC k pour chaque essai et N_{ik} le nombre d'essais de l'apprenant i sur un exercice associé à la CC k

$$p_{ij} = \sigma \left(\sum_{k \in KC(j)} \beta_k + \gamma_k N_{ik} \right) \quad (2)$$

- L'algorithme *performance factor analysis* (PFA) modélise l'apprenant à l'aide du biais β_k de la CC k , des biais γ_k et δ_k de maîtriser la CC k pour chaque essai réussi ou non, et du nombre de succès $S_{i,k}$ et d'échecs $F_{i,k}$ sur des exercices associés à la CC k .

$$p_{ij} = \sigma \left(\sum_{k \in KC(j)} \beta_k + \gamma_k S_{i,k} + \delta_k F_{i,k} \right) \quad (3)$$

- L'algorithme *learning factor analysis* (LFA) modélise l'apprenant à l'aide du biais α_i de chaque apprenant, et comme pour l'additive factor model, la difficulté β_k de la CC k , des biais γ_k et δ_k de maîtriser la CC k pour chaque essai réussi ou non, et du nombre de succès $S_{i,k}$ et d'échecs $F_{i,k}$ sur des exercices associés à la CC k .

$$p_{ij} = \sigma \left(\alpha_i + \sum_{k \in KC(j)} (\beta_k + \gamma_k S_{i,k} + \delta_k F_{i,k}) \right) \quad (4)$$

3 Méthodologie

3.1 Données exploitées

Nous utilisons les données issues de traces d'apprenants de CE1 et CE2 sur des exercices interactifs sur l'ensemble du programme scolaire de français et de mathématiques de CE1 et CE2, issus de la plateforme web Kartable (www.kartable.fr). La majorité des exercices sont des questions à choix multiples, où le nombre de réponses attendues est donné (environ 90% ; le reste correspond à des questions libres, des couples à relier ou des glisser-déposer). Les observations présentes dans ces jeux de données sont les suivantes : utilisateur, exercice réalisé, CC associées, nombre d'essais, nombre de succès, nombre d'échecs et date de l'interaction.

La Table 1 décrit le jeu de données en fonction de la matière des exercices desquels sont issues les traces des apprenants. Nous avons choisi de prendre deux matières (mathématiques et français) dont la structure pédagogique est très différente afin de vérifier si nos résultats dépendent de la matière. Par la suite, les mathématiques seront désignées par MAT et le français par FRA.

Table 1. Récapitulatif des ressources et des utilisateurs du jeu de données utilisé dans ce papier. Le terme “exercice” correspond aux questions présentées aux élèves tandis que le terme “interactions” correspond à la réponse d'un apprenant à cette question.

Matière	Nombre d'apprenants	Nombre d'exercices	Nombre d'interactions
MAT	6734	2314	1 393 525
FRA	7309	2880	2 093 946

3.2 Catégorisation des composantes de connaissance

Les composantes de connaissance sont catégorisées en fonction de leur nature déclarative ou procédurale. Cette catégorisation des CC est effectuée à partir de la définition des mémoires déclarative et procédurale [3]. Les CC déclaratives portent sur un ensemble de connaissances factuelles ou explicites. Les CC procédurales, elles, portent sur des schémas de raisonnements implicites ou tacites.

La typologie des CC renseigne en partie la nature de ces dernières : l'usage d'un verbe d'action est d'usage dans la dénomination des CC et il donne une indication sur leur caractère déclaratif ou procédural. Une CC nommée “Savoir...” ou “Connaître...” aura tendance à être une CC déclarative, tandis qu'une CC nommée “Calculer...” ou “Déterminer...” sera plus facilement une CC procédurale. La catégorisation est automatiquement déterminée de cette façon, puis elle est comparée à une classification effectuée par deux enseignants sur un échantillon de 30 CC issues de l'ensemble des CC étudiées dans notre papier. L'accord entre la classification automatique et celle des enseignants a été mesuré par le calcul du kappa de Cohen $\kappa = (p_a - p_e) / (1 - p_e)$, où p_a est l'accord relatif

entre des enseignants avec la classification automatique et p_e la probabilité d'un accord aléatoire. Nous obtenons $\kappa = 0.7$ en mathématiques et $\kappa = 0.1$ en français. Malgré une volonté d'uniformisation des définitions, il semble complexe pour certains domaines (tel que le français) de définir la nature de toutes les CC. Nous considérons donc les mathématiques comme une matière où la catégorisation est plutôt correcte et le français comme une matière où la catégorisation est très bruitée, ce qui doit être pris en compte dans l'analyse des résultats. Dans la suite, les CC déclaratives seront désignés par DEC, les CC procédurales par PRO.

3.3 Mesure de la performance

Afin de mesurer la capacité de nos modèles à prédire la réponse d'un apprenant i à un exercice j , nous utilisons les métriques suivantes :

- AUC (pour *area under curve*), une métrique allant de 0 (lorsque les prédictions sont toutes erronées) à 1 (traduisant une classification parfaite). Ce score est calculé par l'intermédiaire de la courbe ROC.
- RMSE (pour *root-mean-square error*), mesurant la précision de la prédiction. Plus la corrélation des prédictions avec les données est forte, plus RMSE tend vers 0.

La performance des algorithmes étudiés est principalement mesurée à l'aide des valeurs d'AUC. La métrique RMSE n'est utilisée ici que pour départager deux algorithmes dont la métrique AUC serait égale : à performance de classification équivalente, on compare l'écart à la moyenne pour déterminer quel algorithme pourrait donner des résultats trop éloignés de la réalité.

On procède à une validation croisée à $k = 5$ blocs. Pour les modèles de *knowledge tracing*, la validation croisée est basée sur les CC : le jeu de données de la phase de test porte sur des CC différentes du jeu de données de la phase d'apprentissage. La validation croisée pour les modèles logistiques est, elle, basée sur les apprenants : le jeu de données de la phase de test porte sur des apprenants différents du jeu de données de la phase d'apprentissage. La différence entre les deux types de validation croisée est représentée dans la Figure 2. Les deux types de modèles de prédiction ne seront donc pas comparés, car les métriques utilisées ne concordent pas, malgré le fait qu'elles portent sur le même jeu de données. La validation croisée basée sur les apprenants mène à des valeurs d'AUC plus élevées que celle basée sur les CC [7].

4 Résultats

4.1 Comparaison des algorithmes de diagnostic cognitif

Les résultats détaillés des algorithmes que l'on a utilisé pour ce papier sont présentés dans la Fig. 3. On remarque que la variante KT-IDE (pour le modèle BKT) et IRT (pour le modèle *factor analysis*) ont de meilleures performances sur

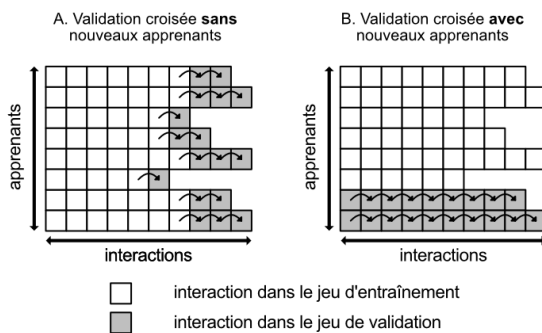


Fig. 2. Comparaison de la méthode de validation entre (A) le modèle de *knowledge tracing* et (B) le modèle de *factor analysis*, illustration inspirée de [11].

notre jeu de données que les autres algorithmes et variantes étudiés. Les résultats présentés dans les parties 4.2 et 4.3 sont restreints à ces deux algorithmes. On note tout de même que des résultats de même tendance ont été observés pour les autres algorithmes de leur modèle respectif.

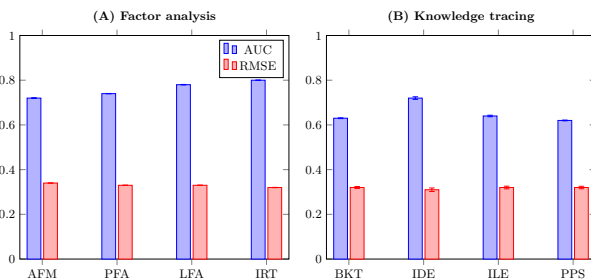


Fig. 3. Performance des différents algorithmes testés pour les modèles (A) *factor analysis* et (B) *knowledge tracing* : AUC et RMSE sont calculés pour chaque algorithme selon la méthodologie décrite dans la partie 3.3.

4.2 Filtrage par nature des composantes de connaissance et matière associée

Chaque interaction entre un apprenant et un exercice peut être filtrée en fonction de la nature des CC associées et de la matière concernée. La Figure 4 présente les métriques de performance de l'algorithme IRT sur les différents jeux de données filtrés. Les résultats sont groupés par matière concernée afin de pouvoir étudier la différenciation des données par la nature des CC sur les deux matières étudiées. Nous remarquons une meilleure performance d'IRT sur le jeu de données ne portant que sur les CC déclaratives (AUC +0,05 / RMSE -0,06 par rapport au jeu de données restreint aux CC procédurales). En comparant ces mêmes résultats en fonction de la matière associée aux CC par l'algorithme IRT, nous remarquons une meilleure performance globale de l'algorithme sur le jeu de données de mathématiques : l'écart de performance entre les mathématiques et le français est assez faible sur le jeu de données restreint aux CC procédurales (AUC +0,01 / RMSE -0,02), mais devient bien plus important sur le jeu de données restreint aux CC déclaratives (AUC +0,05 / RMSE -0,09).

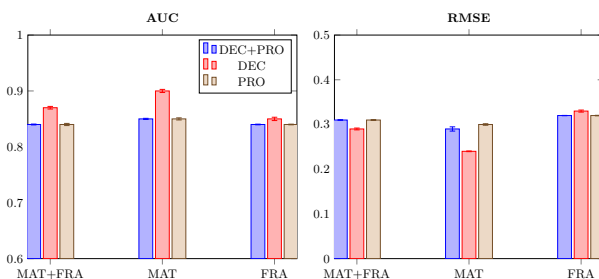


Fig. 4. Performances de l'algorithme IRT sur les données filtrées par nature de CC et par matière associée. Les matières sont en abscisse pour comparer les performances en fonction du filtrage par nature de CC. À gauche, AUC. À droite, RMSE.

De la même manière, la Figure 5 présente les performances de l'algorithme KT-IDE en fonction de la nature des CC associées pour étudier l'impact de la différenciation des données par la nature des CC sur les deux matières étudiées. Cette fois, nous observons que ce sont sur les CC procédurales que l'algorithme de *knowledge tracing* réalise de meilleures prédictions (AUC +0,07 / RMSE -0,05 sur le jeu de données sur les deux matières). La matière concernée n'influe que très peu sur les résultats, nous ne remarquons qu'une légère amélioration sur les mathématiques (AUC = / RMSE -0,02 pour les mathématiques par rapport au français sur le jeu de données restreint aux CC procédurales).

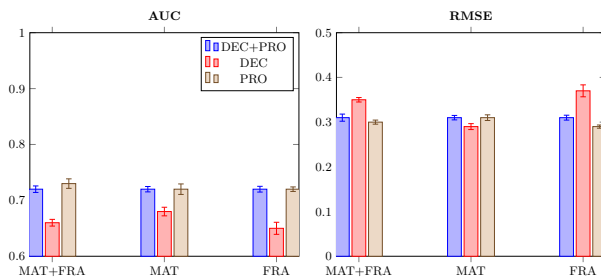


Fig. 5. Performances de l'algorithme IDE sur les données filtrées par nature de CC et par matière associée. Les matières sont en abscisse pour comparer les performances en fonction du filtrage par nature de CC. À gauche, AUC. À droite, RMSE.

5 Discussion

5.1 Comparaison des algorithmes de diagnostic cognitif

Comme nous l'avons indiqué dans la partie 3.3, la façon dont sont calculées les métriques de performance des différents algorithmes rend impossible la comparaison inter-modèles. Néanmoins, il est possible de comparer les algorithmes appartenant à la même catégorie de modèle de l'apprenant.

Une nette domination d'IRT parmi les algorithmes de *factor analysis* sur toutes les métriques présentées est observée dans la partie 4.1. Du côté des algorithmes de *knowledge tracing*, les meilleurs résultats sont ceux de la variante KT-IDE. La nature des exercices sur lesquels sont récoltées les traces des apprenants peut être une des raisons de tels résultats. D'une part, les paramètres de *slip* et *guess* individualisés pour chaque exercice permettent une étude plus précise de tous les exercices sur lesquels sont récoltées les traces des apprenants. D'autre part, la théorie de réponse à l'item est plus performante sur des questions à choix multiples.

5.2 Influence de la nature des composantes de connaissance

Nos résultats sur les performances des deux modèles de l'apprenant deviennent bien plus éloquentes lorsque l'on s'intéresse à la comparaison des performances des algorithmes sur les jeux de données filtrés par la nature des CC. D'une part, pour le modèle *factor analysis*, le jeu de données limité aux CC déclaratives donne lieu à de meilleures performances que sur les jeux de données limités aux CC procédurales. Les algorithmes de *factor analysis* sont plus adaptés à la prédiction de réussite d'un apprenant sur des exercices portant sur des CC déclaratives. Les modèles logistiques semblent à même de modéliser la mémoire

(à savoir la mémoire déclarative de Cohen [5]). D'autre part, le modèle *knowledge tracing* est plus performant sur le jeu de données limité aux CC procédurales. Là encore, la structure de ce modèle semble plus adaptée pour étudier le phénomène de compréhension, propre aux tâches techniques. Il existe un gain de performance à limiter les algorithmes de *knowledge tracing* aux seules CC procédurales.

Nous comprenons ici l'importance de l'interprétabilité des paramètres utilisés dans le modèle de l'apprenant. L'augmentation des performances par la catégorisation des CC pour l'un et l'autre des modèles utilisés révèle un choix pertinent des observations par rapport à l'objet à modéliser. Les paramètres, résultant du choix des observations, sont ainsi plus significatif dès lors que leur détermination dépend de la nature des CC sur lesquels ils sont calculés. Le modèle *knowledge tracing* repose sur des méta-paramètres particulièrement adaptés aux CC procédurales, tandis que le modèle *factor analysis*, lui, considère des paramètres associés aux CC déclaratives. En restreignant l'usage de BKT aux CC procédurales, le paramètre lié à l'apprentissage (noté P_i dans la partie 2), au delà de permettre une bonne prédiction des résultats de l'apprenant, récupère tout le sens qu'il est censé avoir : il mesure la capacité de l'apprenant de passer d'un état de non-maîtrise d'une CC à un état de maîtrise, et modélise le moment du déclin, propre aux tâches techniques associées aux CC procédurales. Nous comprenons qu'il n'y aurait que très peu de sens à un tel paramètre pour une CC déclarative. De la même manière, un paramètre tel que le niveau de l'apprenant peut avoir un champ d'interprétation extrêmement large. Cependant, une fois limité aux seules CC déclaratives, il ne s'agit plus d'un niveau global censé mesurer la capacité d'un élève à comprendre, appliquer, etc, mais bien de la seule capacité d'apprendre et de retenir des faits. Comme le périmètre d'observation est plus restreint (car seule une nature de CC est considérée lors de l'apprentissage), les paramètres sont plus spécialisés et sont destinés à mesurer une métrique propre à la fonction cognitive sollicitée. Ils deviennent donc plus facilement interprétables.

Dans l'optique d'un système éducatif adaptatif où l'on chercherait l'interprétabilité du modèle de l'apprenant, le fait d'avoir des paramètres compréhensibles (voir partie 2) et adaptés en fonction des ressources étudiées devient très intéressant. De telles métriques permettent d'accéder à une analyse sur les différentes fonctions cognitives de l'apprenant. De plus, la catégorisation des CC permet en effet une amélioration des prédictions du modèle de l'apprenant.

5.3 Influence de la matière

Nous avons pu également observer des différences entre les deux matières observées. L'écart de performance entre *factor analysis* et *knowledge tracing* en fonction de la nature des CC est également visible lorsque l'on restreint les jeux de données aux mathématiques ou au français. Néanmoins, le contraste est plus marqué pour les mathématiques que pour le français. Nous rappelons que les deux matières ont été sujettes à une catégorisation de qualité différentes ($\kappa = 0.7$ pour les mathématiques, $\kappa = 0.1$ pour le français). Nous expliquons cette

différence entre les deux matières par les divergences remarquées dans la partie 3.3 sur la catégorisation des CC. D'une part, il apparaît que la catégorisation des CC procédurales est plus fiable que celle des CC déclaratives dans les deux matières, ce qui explique les similitudes des résultats entre mathématiques et français sur le jeu de données restreint aux CC procédurales. D'autre part, il existe d'après notre étude auprès des experts plus de cas de CC considérées comme procédurales mais déclarées déclaratives en français qu'en mathématiques. La performance obtenue sur le jeu de données restreint aux CC déclaratives est donc plus contrastée, avec de meilleurs résultats sur les mathématiques que sur le français.

Il est nécessaire ici de prendre en compte le niveau correspondant aux exercices étudiés. En classes de CE1 et CE2, la différence entre exercices portant sur des composantes de connaissance déclarative et procédurale semble bien plus marqué en mathématiques qu'en français. En mathématiques, les CC déclaratives portent sur de la restitution pure et les CC procédurales sur des tâches mécaniques demandant une rigueur méthodologique. En français, par contre, la nature des composantes de connaissance est plus floue : comment déterminer le type de mémoire sollicité par un exercice demandant le pluriel d'un nom ? Les kappa témoignent de la difficulté pour des experts de discriminer entre des CC déclaratives et procédurales. Les différences de performance entre les matières, elles, mettent en lumière l'importance de la classification par leur nature des CC pour un gain réel en interprétabilité.

6 Conclusion et perspectives

Nous avons mis en évidence dans ce papier une différence de performance des algorithmes modélisant l'apprenant en fonction de la nature des composantes de connaissance sur lesquelles sont récoltées les données issues des traces d'élèves. Cette différence a été étudiée dans deux matières différentes (les mathématiques et le français), ce qui permet d'améliorer la vraisemblance de notre hypothèse.

Nous comprenons alors l'importance de l'interprétabilité des paramètres utilisés dans ces algorithmes qui, au delà d'obtenir une précision suffisante, peuvent et doivent permettre d'accéder à des informations sur l'apprenant. Il devient possible de déterminer l'aisance d'un apprenant vis-à-vis de sa mémoire déclarative ou procédurale, tout en ne modélisant que son savoir, ce qui ouvre le champ à de nouveaux diagnostics cognitifs.

Le corollaire de ces résultats est l'intérêt pour un ELAH adaptatif de combiner des modèles permettant de modéliser les deux types de mémoire, déclarative et procédurale, d'un apprenant. Sous la réserve que les modèles de *knowledge tracing* puisse avoir des performances équivalentes aux modèles d'additive factors, la combinaison des deux laisse entrevoir un gain significatif pour le diagnostic. La détermination de la performance d'un tel modèle de diagnostic hybride devrait faire l'objet de futures recherches.

References

1. Vincent Alevan and Kenneth R Koedinger. Knowledge component (kc) approaches to learner modeling. *Design Recommendations for Intelligent Tutoring Systems*, 1:165–182, 2013.
2. Vincent Alevan, Elizabeth A McLaughlin, R Amos Glenn, and Kenneth R Koedinger. Instruction based on adaptive learning technologies. *Handbook of research on learning and instruction*, pages 522–560, 2016.
3. John R Anderson. Act: A simple theory of complex cognition. *American psychologist*, 51(4):355, 1996.
4. Susan Bull and Judy Kay. Open learner models. In *Advances in intelligent tutoring systems*, pages 301–322. Springer, 2010.
5. Neal J Cohen and Larry R Squire. Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210(4466):207–210, 1980.
6. Albert T Corbett and John R Anderson. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, 4(4):253–278, 1994.
7. Mohammad Khajah, Robert V. Lindsey, and Michael Mozer. How deep is knowledge tracing? In *Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, June 29 - July 2, 2016*. International Educational Data Mining Society (IEDMS), 2016.
8. Zachary A Pardos and Neil T Heffernan. Using hmms and bagged decision trees to leverage rich features of user and skill from an intelligent tutoring system dataset. *Journal of Machine Learning Research W & CP*, 40, 2010.
9. Zachary A Pardos and Neil T Heffernan. Kt-idem: Introducing item difficulty to the knowledge tracing model. In *International conference on user modeling, adaptation, and personalization*, pages 243–254. Springer, 2011.
10. Zachary A Pardos, Qing Yang Wang, and Shubhendu Trivedi. The real world significance of performance prediction. *International Educational Data Mining Society*, 2012.
11. Radek Pelánek. Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. *User Modeling and User-Adapted Interaction*, 27(3-5):313–350, 2017.
12. Radek Pelánek, Jirí Rihák, and Jan Papoušek. Impact of data collection on interpretation and evaluation of student models. In *Proceedings of the sixth international conference on learning analytics & knowledge*, pages 40–47, 2016.
13. Sarah Schultz and Trenton Tabor. Revisiting and extending the item difficulty effect model. In *In Proceedings of the 1st workshop on massive open online courses at the 16th annual conference on artificial intelligence in education*, pages 33–40. Citeseer, 2013.

Vers la conception de *feedback* pour enseignants dans un contexte d'évaluation formative à grande échelle : une approche analytique

Rialy Andriamizeza¹, Franck Silvestre¹, Jean-François Parmentier², and Julien Broisin¹

¹ Institut de recherche en Informatique de Toulouse, Université de Toulouse, 118 Route de Narbonne, 31400 Toulouse, France

² Toulouse INP-ENSEEIH, 2 Rue Charles Camichel, 31000 Toulouse

Abstract. Pour résoudre le problème de l'augmentation du nombre d'étudiants dans l'enseignement supérieur, l'évaluation formative assistée par la technologie et ses processus variés ont émergé. Parmi eux, nous nous intéressons aux processus à deux votes et à ses interactions pour identifier des *feedback* à fournir aux enseignants afin de les aider dans leur prise de décision. À partir de la littérature et d'un jeu de données collectées en contexte écologique, nous avons mobilisé différentes techniques d'analyse pour identifier des informations pouvant aider les enseignants à mener des séquences d'évaluation formative dans le contexte des processus à 2 votes : (1) Plus le pourcentage de réponses correctes au premier vote est proche de 50, plus la séquence a des chances d'être bénéfique ; (2) Plus l'évaluation par les pairs est cohérente, plus la séquence a des chances d'être bénéfique ; (3) La cohérence du degré de confiance des apprenants est corrélée avec la cohérence de l'évaluation par les pairs. Les prochains travaux viseront à implémenter ces *feedback* dans notre système.

Keywords: évaluation formative; *learning analytics*; instruction par les pairs; *feedback* pour enseignants

Abstract. To face the increasing number of students in higher education, technology-enhanced formative assessment and its processes emerged. We focus on two-votes-based processes, and its interactions that could be used to identify appropriate feedback to support teachers. Based on literature and using a dataset gathered from an authentic learning contexts, we use learning analytics to identify different information that will help teachers conduct formative assessment sequences in two-votes-based processes: (1) The percentage of correct answers at the first vote should be close to 50; (2) The consistency of peer assessment is correlated with sequence benefits; (3) The consistency of learners' confidence degree is correlated to the consistency of peer assessment. The next step of these works will consist in implementing meaningful guidance tools based on these indicators into our system.

Keywords: formative assessment; learning analytics; peer instruction; teacher feedback

1 Introduction

L'évaluation formative a pour objectif d'améliorer l'enseignement et l'apprentissage en fournissant des *feedback* aux enseignants et apprenants afin de les aider à adapter leurs comportements. Sadler définit le *feedback* comme "une information concernant le succès d'une tâche" [28]. Les *feedback* auxquels nous nous intéressons dans ce papier sont les informations utiles pour les enseignants concernant le succès des tâches accomplies par les apprenants, et ce, afin de faciliter la prise de décision pour la suite de la séquence. Cependant, selon Andersson, l'évaluation formative est souvent utilisée de manière informelle ou approximative [1]. Ellis met l'accent sur la tâche complexe consistant à collecter toutes les interactions d'apprentissage en face à face [14]. Ainsi, fournir des *feedback* de qualité est une tâche complexe, surtout dans les configurations à grand échelle où les interactions d'apprentissage augmentent avec le nombre d'apprenants.

Pour résoudre ce problème et pour faire face à l'augmentation du nombre d'étudiants dans l'enseignement supérieur, l'évaluation formative assistée par la technologie a émergé. Différents systèmes implantent des processus variés permettant aux enseignants de mettre en œuvre des séquences d'évaluation formative dans un contexte d'enseignement de masse. Une famille de processus, à savoir "les processus à deux votes", demande aux apprenants de voter deux fois durant la séquence comprenant 5 phases distinctes : (1) l'enseignant pose une question à la classe ; (2) les apprenants donnent leurs premières réponses ; (3) les apprenants réfléchissent aux réponses de leurs pairs et remettent éventuellement en cause leurs connaissances ; (4) les apprenants donnent leurs secondes réponses à la même question ; (5) l'enseignant et les apprenants échangent sur les résultats restitués. Le Peer Instruction (PI), tel que décrit par Mazur [9], est l'une des premières implantations de ce type de processus d'évaluation formative. Ces travaux sur le Peer Instruction [22] ainsi que des études sur d'autres processus à deux votes [7] ont montré un impact positif sur l'engagement des apprenants et leurs apprentissages. L'objectif des processus à deux votes est d'obtenir une augmentation du nombre de réponses correctes entre le premier et le second vote grâce à la confrontation de points de vue entre apprenants. Lorsque c'est le cas, cela signifie qu'en moyenne, le niveau de compréhension des apprenants a augmenté [32]. Ainsi, dans la suite de ce papier, nous qualifierons de "bénéfique" toute séquence d'évaluation formative à deux votes pour laquelle la proportion de bonnes réponses a augmenté lors du second vote.

Ces 5 phases comportent une grande variété d'interactions pouvant générer des traces exploitables. Cependant, étant donné le manque de données relatives aux processus à deux votes [3], peu de travaux ont exploré l'usage de ces interactions pour les transformer en *feedback* pertinents. C'est pourquoi, dans ce papier, nous considérons la question de recherche suivante : Quels *feedback* pourraient être fournis aux enseignants tout au long des différentes phases d'un processus d'évaluation formative à deux votes dispensé à grande échelle, afin d'en améliorer les bénéfices en terme d'apprentissage ? Pour répondre à cette question, nous avons adopté la démarche traditionnelle illustrée par Lebis [20], qui consiste, à

partir des traces de l'EIAH (Environnement Informatique pour l'Apprentissage Humain), à appliquer différentes techniques d'analyse aboutissant à de nouvelles connaissances. Le jeu de données dont nous disposons dans cette étude provient de l'utilisation en contexte écologique, dans l'enseignement supérieur, d'un outil implantant un processus à deux votes.

Le papier est structuré comme suit. La section 2 introduit l'évaluation formative et montre que les *feedback* existants dans les outils d'évaluation formative peuvent être améliorés pour aider les enseignants à prendre des décisions à partir des données collectées au fil de l'activité. La section 3 décrit l'outil d'évaluation formative que nous avons utilisé pour notre étude. La section 4 fournit des détails sur les analyses que nous avons effectuées ainsi que sur les résultats obtenus. Enfin, la section 5 conclut et évoque des pistes pour nos travaux futurs.

2 L'évaluation formative et le *feedback*

2.1 Définitions

Bien que l'évaluation soit souvent utilisée en tant qu'évaluation *de* l'apprentissage, elle peut aussi être utilisée *pour* l'apprentissage [21]. D'un côté, l'évaluation sommative est utilisée pour mesurer le niveau des apprenants à la fin d'une unité d'instruction. De l'autre, l'évaluation formative est cruciale pour permettre aux enseignants d'évaluer la compréhension des apprenants et pour adapter le cours [11]. Hattie évoque l'évaluation formative comme l'une des méthodes à privilégier pour améliorer les résultats des apprenants [17]. En 1998, Black et William définissent l'évaluation formative comme "l'ensemble des activités entreprises par les enseignants, et/ou par leurs élèves, qui fournissent des informations qui seront utilisées comme *feedback* pour modifier les activités d'enseignement et d'apprentissage dans lesquelles ils sont engagés" [4]. Cette définition met en lumière l'importance des données et de leur traitement afin de proposer des *feedback* conçus pour améliorer l'apprentissage et l'enseignement.

Par exemple, lors de cours en face à face, Meltzer et Mannivan ont considéré l'utilisation d'artefacts visuels (tels qu'une feuille de papier ou des cartons) pour permettre aux apprenants de répondre aux questions posées par les enseignants [23]. La présentation des artefacts par les apprenants constituent un *feedback* adressé à l'enseignant lui permettant de prendre connaissance des réponses des apprenants en un coup d'oeil et d'adapter son enseignement. Malheureusement, cette méthode s'applique difficilement à l'enseignement à grande échelle car les enseignants peuvent difficilement visualiser plusieurs dizaines de réponses et les assimiler en temps réel.

2.2 Les *feedback* dans les TEFA

Parmi les solutions qui ont émergé pour appliquer l'évaluation formative à grande échelle avec *feedback* immédiat, les TEFA (*Technology-Enhanced Formative Assessment*) se sont développés [30]. Parmi les systèmes au coeur des TEFA, les CRS (*Classroom Response System*) sont les plus utilisés.

La forme la plus simple d'évaluation formative peut être mise en oeuvre avec des CRS tels que Clickers [6] ou avec des plateformes disponibles sur le web telles que Poll Everywhere [8]. Elle permet aux enseignants de poser une question à choix et aux apprenants de sélectionner une réponse. Des *feedback* représentant la répartition des votes sont ensuite présentés aux enseignants. Ils ont pour but de faciliter l'échange entre les enseignants et apprenants à la fin de la séquence. Kahoot [18], Socrative [2] et Plickers [19] supportent le même processus mais proposent en plus d'une vue d'ensemble du vote des apprenants, un *feedback* représentant une vue d'ensemble de la séquence ainsi que les réponses par apprenant sur différentes séquences.

Tirant partie d'une des cinq stratégies caractéristiques de l'évaluation formative identifiées dans la théorie de Black et William [5], ComPAIR [27] utilise les apprenants comme ressources d'apprentissage au sein de son processus d'évaluation formative. Plus précisément, il permet aux apprenants de fournir une réponse textuelle à une question ouverte posée par les enseignants. Par la suite, une phase de revue par les pairs s'engage. Les apprenants doivent fournir un *feedback* textuel sur les réponses de deux de leurs pairs ainsi qu'une justification concernant la raison pour laquelle une des réponses leur paraît plus pertinente qu'une autre. Pendant et après cette phase, les enseignants accèdent à un *feedback* présentant, pour chaque apprenant, la réponse choisie, le *feedback* textuel qu'il a fourni et les comparaisons soumises pour chaque paire de réponses présentée.

Elaastic [13] et myDalite [7] supportent tous deux un processus à deux votes illustré par la Figure 1. Il consiste à demander aux apprenants de voter pour une première réponse et de fournir une explication textuelle pour justifier leur choix. Par la suite, les apprenants peuvent voter une seconde fois. Concernant cette seconde tentative, les deux plateformes se distinguent. myDalite permet aux apprenants de choisir une justification écrite par un pair comme second vote. Ils peuvent également ne pas changer d'avis et garder leurs justifications du premier vote. Les enseignants disposent d'un *feedback* qui détaille le nombre d'apprenants qui sont passés d'une réponse incorrecte à correcte, correcte à incorrecte, incorrecte à incorrecte et correcte à correcte. De son côté, Elaastic implique les apprenants dans une phase d'évaluation avant qu'ils puissent soumettre leur second vote. Il leur est demandé d'évaluer les justifications écrites de leurs pairs. À tout moment, les enseignants peuvent consulter les premiers et seconds votes des apprenants et accéder à l'ensemble de leurs justifications écrites ainsi qu'à la note moyenne attribuée par les pairs à chacune d'entre elles. Plus de détails sont donnés dans la section 3.1.

Les *feedback* présentés dans cette section ont pour objectif de restituer aux enseignants une synthèse du déroulement de la séquence lorsque cette dernière est terminée. Ils ne sont pas conçus pour aider les enseignants à orchestrer des séquences d'évaluation formative en leur fournissant, par exemple, tout au long des différentes phases du processus à deux votes, des informations sur le comportement des apprenants. Pour combler cette lacune, nous souhaitons étudier comment la variété de données générées par les différentes interactions

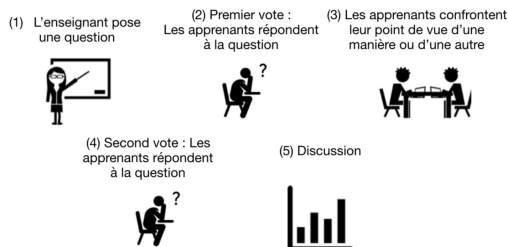


Fig. 1: Les 5 phases du processus à deux votes.

des processus à deux votes constitue une opportunité de construire des *feedback* utiles à la prise de décision.

3 Conception du jeu de données

Nous présentons ici la plateforme d'évaluation formative utilisée ainsi que le jeu de données collectées par son utilisation dans l'enseignement supérieur.

3.1 Elaastic, une technologie au service de l'évaluation formative

Elaastic est une plateforme web [26] utilisée depuis 2015 dans différents contextes d'enseignement dans le supérieur (face à face, distance ou hybride) pour des disciplines telles que la physique, l'informatique et le management de projet.

Durant la phase 1, les enseignants posent une question. Si la question est fermée, elle peut être à choix exclusif ou à choix multiples. Durant la phase 2, les apprenants répondent à la question et fournissent une explication écrite pour justifier leur(s) choix. Ils doivent également fournir leur degré de confiance sur une échelle de Likert à 4 items (voir Figure 2). Durant la phase 3, les apprenants s'engagent dans une activité d'évaluation par les pairs [31]. Plus précisément, il leur est demandé d'exprimer leur degré d'accord avec chaque justification. Ce degré d'accord est fourni sur une échelle de Likert à 5 items (voir Figure 3). Afin de réduire les biais sociaux, les justifications présentées aux apprenants sont anonymes. Les enseignants peuvent configurer le nombre de justifications (5 maximum) présentées à chaque apprenant. Ensuite, la phase 4 commence et les apprenants ont l'opportunité de voter une seconde fois pour la(les) réponse(s) qu'ils pensent être la(les) bonne(s). Enfin, les enseignants démarrent la phase 5. Les votes des apprenants, l'ensemble des justifications écrites ainsi que leurs notes moyennes attribuées par les pairs sont communiquées aux étudiants pour une discussion.

▼ Question de mathématiques [QUESTION À CHOIX EXCLUSIF]

$g(x) = ax^2 + 24$

Pour la fonction f définie ci-dessus, a est une constante et $g(4) = 8$. Quelle est la valeur de $g(-4)$? 1.8 : 2.0 : 3.-1: 4.-8

Réponse

Veuillez soumettre votre réponse

Votre réponse : 1 2 3 4

Réponse textuelle

B I U | \int $\frac{1}{x}$ $\frac{1}{x^2}$ $\frac{1}{x^3}$ $\frac{1}{x^4}$ $\frac{1}{x^5}$ $\frac{1}{x^6}$ $\frac{1}{x^7}$ $\frac{1}{x^8}$ $\frac{1}{x^9}$ $\frac{1}{x^{10}}$ $\frac{1}{x^{11}}$ $\frac{1}{x^{12}}$ $\frac{1}{x^{13}}$ $\frac{1}{x^{14}}$ $\frac{1}{x^{15}}$ $\frac{1}{x^{16}}$ $\frac{1}{x^{17}}$ $\frac{1}{x^{18}}$ $\frac{1}{x^{19}}$ $\frac{1}{x^{20}}$ $\frac{1}{x^{21}}$ $\frac{1}{x^{22}}$ $\frac{1}{x^{23}}$ $\frac{1}{x^{24}}$ $\frac{1}{x^{25}}$ $\frac{1}{x^{26}}$ $\frac{1}{x^{27}}$ $\frac{1}{x^{28}}$ $\frac{1}{x^{29}}$ $\frac{1}{x^{30}}$ $\frac{1}{x^{31}}$ $\frac{1}{x^{32}}$ $\frac{1}{x^{33}}$ $\frac{1}{x^{34}}$ $\frac{1}{x^{35}}$ $\frac{1}{x^{36}}$ $\frac{1}{x^{37}}$ $\frac{1}{x^{38}}$ $\frac{1}{x^{39}}$ $\frac{1}{x^{40}}$ $\frac{1}{x^{41}}$ $\frac{1}{x^{42}}$ $\frac{1}{x^{43}}$ $\frac{1}{x^{44}}$ $\frac{1}{x^{45}}$ $\frac{1}{x^{46}}$ $\frac{1}{x^{47}}$ $\frac{1}{x^{48}}$ $\frac{1}{x^{49}}$ $\frac{1}{x^{50}}$ $\frac{1}{x^{51}}$ $\frac{1}{x^{52}}$ $\frac{1}{x^{53}}$ $\frac{1}{x^{54}}$ $\frac{1}{x^{55}}$ $\frac{1}{x^{56}}$ $\frac{1}{x^{57}}$ $\frac{1}{x^{58}}$ $\frac{1}{x^{59}}$ $\frac{1}{x^{60}}$ $\frac{1}{x^{61}}$ $\frac{1}{x^{62}}$ $\frac{1}{x^{63}}$ $\frac{1}{x^{64}}$ $\frac{1}{x^{65}}$ $\frac{1}{x^{66}}$ $\frac{1}{x^{67}}$ $\frac{1}{x^{68}}$ $\frac{1}{x^{69}}$ $\frac{1}{x^{70}}$ $\frac{1}{x^{71}}$ $\frac{1}{x^{72}}$ $\frac{1}{x^{73}}$ $\frac{1}{x^{74}}$ $\frac{1}{x^{75}}$ $\frac{1}{x^{76}}$ $\frac{1}{x^{77}}$ $\frac{1}{x^{78}}$ $\frac{1}{x^{79}}$ $\frac{1}{x^{80}}$ $\frac{1}{x^{81}}$ $\frac{1}{x^{82}}$ $\frac{1}{x^{83}}$ $\frac{1}{x^{84}}$ $\frac{1}{x^{85}}$ $\frac{1}{x^{86}}$ $\frac{1}{x^{87}}$ $\frac{1}{x^{88}}$ $\frac{1}{x^{89}}$ $\frac{1}{x^{90}}$ $\frac{1}{x^{91}}$ $\frac{1}{x^{92}}$ $\frac{1}{x^{93}}$ $\frac{1}{x^{94}}$ $\frac{1}{x^{95}}$ $\frac{1}{x^{96}}$ $\frac{1}{x^{97}}$ $\frac{1}{x^{98}}$ $\frac{1}{x^{99}}$ $\frac{1}{x^{100}}$

Si $g(4)$ a retourné un résultat positif alors $g(-4)$ devait retourner un résultat négatif

Votre degré de confiance

Pas du tout confiant(e) Pas vraiment confiant(e) Confiant(e) Tout à fait confiant(e)

Soumettre

Fig. 2: Elaastic : formulaire de soumission du premier vote.

Confrontation de points de vue

Ci-dessous sont présentées une ou plusieurs réponses alternatives. Merci de bien vouloir indiquer dans quelle mesure vous êtes d'accord avec ces réponses.

Choix [1]
 x^2 retourne le même résultat pour une valeur et son opposé. Donc $g(x)$ et $g(-x)$ retournent le même résultat. Cette règle s'applique également lorsque $x = 4$. Donc si $g(4) = 8$ alors $g(-4) = 8$

Votre évaluation :

Je ne me prononce pas 1 2 3 4 5

Choix [2]
 Si $g(4) = 8$ alors $g(-4) = 0$ car $4+4 = 8$ et $(-4) + 4 = 0$

Votre évaluation :

Je ne me prononce pas 1 2 3 4 5

Vous disposez d'une deuxième chance pour changer votre réponse et votre degré de confiance.

Votre réponse : 1 2 3 4

Soumettre

Fig. 3: Elaastic : formulaire de soumission du second vote.

3.2 Description du jeu de données

Jusqu'à présent, la plateforme a été utilisée lors de 623 séquences menées par 53 enseignants au sein desquelles 1769 apprenants ont fourni 8757 réponses ainsi que 9256 évaluations de leurs pairs.

Une séquence est caractérisée par un contexte d'apprentissage (face-à-face, à distance ou hybride), les réponses soumises avant et après la confrontation de point de vue entre pairs, ainsi que le nombre de participants. Pour chaque réponse, les données suivantes ont été collectées : l'identifiant de l'apprenant,

la justification textuelle, le score de la réponse, ainsi que le choix sélectionné dans le cas d'une question à choix. Concernant les réponses correspondant à un premier vote, des informations additionnelles sont collectées telles que la note moyenne attribuée par les pairs et le degré de confiance de l'apprenant qui a fourni la réponse. Les questions sont composées de leur contenu textuel, du type (ouverte, fermée, à choix exclusif, à choix multiples, etc.) et, dans le cas des questions à choix, du nombre de différents choix proposés aux apprenants. Enfin, pour chaque évaluation par les pairs, la justification écrite, la note attribuée, l'identifiant de l'évaluateur et l'identifiant de l'évalué sont fournis.

4 Analyse de données

Afin de réduire l'influence de facteurs externes et des valeurs aberrantes, nous avons filtré notre jeu de données. Dans un premier temps, nous n'avons gardé que les questions à choix afin de pouvoir distinguer de manière automatique une réponse fautive d'une réponse juste. Dans notre analyse, nous considérons une réponse comme correcte si le score obtenu correspond au score maximum. De plus, nous n'avons gardé que les séquences en face-à-face. Puis nous avons écarté les séquences comportant moins de 10 apprenants car nous souhaitons nous concentrer sur les contextes correspondant à un plus grand nombre d'apprenants. Finalement, nous considérons deux variables p_1 et p_2 qui sont respectivement la proportion d'apprenants ayant répondu correctement au premier et second vote. Les séquences où $p_1 = 0$ ont été enlevées car la confrontation de point de vue ne peut pas avoir lieu (il n'y a pas de justification associée à une bonne réponse pour convaincre ceux qui se sont trompés). Les séquences où $p_2 = 1$ ou $p_1 = 1$ ont également été enlevées. En effet, elles représentent des questions considérées comme trop faciles pour mesurer un effet. Après avoir nettoyé nos données, nous obtenons 104 séquences menées par 21 enseignants au sein desquelles 616 apprenants ont fourni 1981 réponses ainsi que 4072 évaluations de leurs pairs.

Pour mener nos analyses, nous avons utilisé l'estimation de la taille d'effet de Cohen d proposée par Parmentier [25] qui s'appuie sur p_1 et p_2 et est calculée comme suit : $d = 0.6 \ln \left(\frac{p_2}{1-p_2} \frac{1-p_1}{p_1} \right)$. Il est important de mentionner qu'une séquence est bénéfique lorsque $d > 0$ car cela implique que $p_1 < p_2$. La taille d'effet [29] permet de mesurer l'impact de l'évaluation par les pairs sur les votes et, ainsi, de quantifier les bénéfices apportés par une séquence. De plus, notre volume de données est suffisamment grand (> 40) pour utiliser des tests paramétriques sans satisfaire la contrainte de normalité [16]. À partir de ces éléments, nous avons pu exprimer et vérifier nos hypothèses.

4.1 Hypothèse 1 : Les bénéfices d'une séquence sont liés à la distance de p_1 à 50%

Watkins et Mazur [22] ont constaté que leur application du Peer Instruction était la plus bénéfique pour les apprenants lorsqu'entre 30% et 70% d'entre eux ont répondu correctement à la question lors du premier vote. Selon eux, un trop

petit pourcentage de réponse correcte indique des connaissances insuffisantes sur la question pour entamer des discussions productives. À l'inverse, un trop haut pourcentage de bonnes réponses signifie que la question est trop facile pour provoquer une confrontation. À partir de ces travaux, nous émettons l'hypothèse que les bénéfices apportés par une séquence sont liés à la distance de $p1$ à 50%.

Les résultats de la Figure 4 indiquent que la taille d'effet diminue lorsque la distance de $p1$ à 50% augmente. La corrélation de Pearson de $|p1 - 0.5|$ avec d

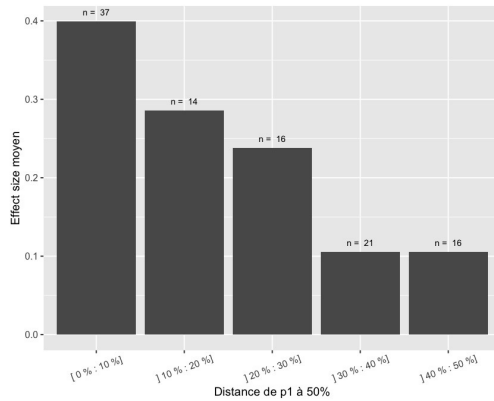


Fig. 4: Taille d'effet moyenne d selon la distance de $p1$ à 50%.

est égale à -0.31 avec une p -value égale à 0.001 et un intervalle de confiance à 95% égal à $[-0.48; -0.13]$. Ainsi, plus la proportion de bonne réponse s'éloigne de 50%, plus les bénéfices de la séquence ont des chances de diminuer.

En synthèse, la distance de $p1$ à 50% est un indicateur utile pour prédire si une séquence a des chances d'être bénéfique. Un *feedback* restituant $p1$ pourrait être fourni aux enseignants après le premier vote pour les aider à prendre une décision. Plus précisément, si la distance de $p1$ à 50% est supérieure à 30%, la Figure 4 montre que la taille d'effet moyenne est significativement plus basse. Dans ce cas, l'enseignant peut sauter la phase de confrontation entre les pairs.

4.2 Hypothèse 2 : La cohérence de l'évaluation par les pairs est liée aux bénéfices d'une séquence

Double et al. montrent que l'utilisation de l'évaluation par les pairs comme pratique formative permet d'améliorer les performances académiques [12]. Ils affirment que réfléchir sur les réponses des pairs favorise l'augmentation du

pourcentage de réponses correctes. Etant donné qu'il est attendu des apprenants ayant fourni une bonne réponse qu'ils convainquent les autres apprenants, nous émettons l'hypothèse que la cohérence de la phase d'évaluation par les pairs est liée aux bénéfices d'une séquence.

Afin de mesurer la cohérence de l'évaluation par les pairs lors d'une séquence, nous définissons ρ_{peer} qui est la corrélation d'un degré d'accord d'un pair à une justification avec l'exactitude de la réponse correspondante. Ces deux variables étant des variables latentes [15], la corrélation polychorique est l'outil statistique approprié [24]. Par exemple, si les justifications des apprenants qui se sont trompés sont bien notées et que celles des apprenants ayant répondu correctement sont mal notées, alors ρ_{peer} sera plutôt proche de -1. À l'inverse, si lors d'une séquence les justifications liées à des réponses correctes sont bien notées et que les justifications liées à des réponses incorrectes sont mal notées, alors ρ_{peer} sera proche de 1. La Figure 5a montre la taille d'effet d en fonction de ρ_{peer} .

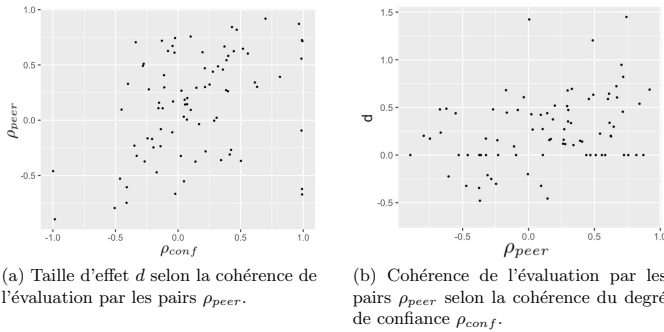


Fig. 5: Taille d'effet d selon ρ_{peer} , et ρ_{peer} selon ρ_{conf} .

La corrélation de Pearson de ρ_{peer} avec d est égale à 0.34 avec une p-value inférieure à 0.002 et un intervalle de confiance à 95% égal à [0.14;0.52]. Ainsi, plus l'évaluation par les pairs est cohérente, plus les bénéfices de la séquence ont des chances d'être élevés. De plus, ρ_{peer} est indépendante de la distance de $p1$ à 50% (p-value = 0.25), ce qui signifie que les Sections 4.1 et 4.2 ont permis d'identifier deux prédicteurs distincts des bénéfices d'une séquence.

À partir de la valeur de ρ_{peer} , nous proposons la fourniture d'un *feedback* aidant les enseignants à sélectionner les réponses à commenter lors de la phase 5. Par exemple, si $\rho_{peer} < 0$, alors les réponses incorrectes sont plus populaires que les réponses correctes. Ces réponses incorrectes doivent être adressées par les enseignants car elles sont potentiellement des conceptions erronées.

4.3 Hypothèse 3 : la cohérence du degré de confiance est liée avec la cohérence de l'évaluation par les pairs

Curtis a utilisé le degré de confiance des apprenants concernant leur réponse comme un moyen d'identifier les apprenants mal informés [10] qu'il définit comme étant des apprenants très confiants mais ayant fourni une mauvaise réponse. Etant donné qu'il est attendu des apprenants ayant répondu incorrectement qu'ils soit convaincus par les autres, nous pensons que des apprenants mal informés ne peuvent pas correctement noter les justifications de leurs pairs. Plus généralement, nous émettons l'hypothèse que la cohérence du degré de confiance des apprenants est liée à la cohérence de l'évaluation par les pairs.

Pour les mêmes raisons que dans la section 4.2, la cohérence du degré de confiance peut-être calculée en utilisant ρ_{conf} qui est la corrélation polychorique du degré de confiance des apprenants avec l'exactitude de leurs réponses au premier vote. La figure 5b est un nuage de point de la cohérence de l'évaluation par les pairs ρ_{peer} en fonction de la cohérence du degré de confiance ρ_{conf} .

La corrélation de Pearson de ρ_{conf} avec ρ_{peer} est égale à 0.38 avec une p-valeur inférieure à $4e - 4$ et un intervalle de confiance à 95% égal à [0.18:0.55]. Cela suggère que la cohérence du degré de confiance des apprenants est liée à la cohérence de l'évaluation par les pairs. Ainsi, plus le degré de confiance des apprenants est cohérent, plus l'évaluation par les pairs a des chances de l'être aussi.

À partir de la valeur de ρ_{conf} , il est envisageable de concevoir un *feedback* pour les enseignants leur indiquant si leurs apprenants sont mal informés. Si $\rho_{conf} < 0$ alors les apprenants ayant répondu incorrectement sont plus confiants que ceux ayant répondu correctement. Dans ce cas, dès la fin du premier vote, un *feedback* pourrait avertir les enseignants que la phase d'évaluation par les pairs est susceptible d'être de mauvaise qualité et ainsi les aider à décider s'ils veulent immédiatement traiter la conception erronée ou raccourcir la séquence.

5 Conclusion

Ce papier s'intéresse à l'évaluation formative, plus précisément, aux bénéfices et défis que constitue la fourniture de *feedback* aux enseignants dans des contextes d'enseignement massif. Nous avons exploré la contribution de la technologie pour adresser les défis de l'évaluation formative et nous avons constaté que les *feedback* conçus pour assister les enseignants pour une prise de décision peuvent être améliorés. Nous avons introduit les processus à deux votes comme une famille de processus au sein de laquelle un vote a lieu avant et après une confrontation de points de vue. Nous nous sommes intéressés à l'identification de nouveaux *feedback* à proposer aux enseignants lors des différentes phases du processus à deux votes afin d'en améliorer les bénéfices. À cette fin, nous avons émis 3 hypothèses en nous appuyant sur une revue de la littérature et avons exploré les données issues de Elaastic pour les vérifier.

Nos résultats suggèrent que les bénéfices d'une séquence dépendent de la cohérence de la phase d'évaluation par les pairs ainsi que de la proportion de

réponses correctes lors du premier vote. De plus, la cohérence de l'évaluation par les pairs dépend de la cohérence du degré de confiance. Ainsi, nous pouvons utiliser ces indicateurs pour accompagner les enseignants dans leur prise de décision au fil de la séquence. Après le premier vote, nous pouvons avertir les enseignants qu'il y a une proportion trop basse ou trop haute de réponses correctes pour que la séquence soit bénéfique. Nous pouvons également utiliser la cohérence du degré de confiance afin d'avertir les enseignants que la phase d'évaluation par les pairs peut ne pas provoquer l'effet attendu. Puis, après l'évaluation par les pairs, nous pouvons dire à l'enseignant si cette dernière était cohérente ou non afin de détecter et traiter les conceptions erronées.

Dans cette étude, nous nous sommes concentrés sur l'identification de *feedback* utiles aux enseignants. Les prochains travaux s'intéresseront à l'implantation de ces *feedback* au sein d'Elaastic. Nous mesurerons ainsi leurs l'impact sur la prise de décisions des enseignants pour mener les séquences. De plus, l'impact de ces décisions sur les bénéfices des séquences sera également étudié.

References

1. Andersson, C., Palm, T.: The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction* **49**, 92–102 (Jun 2017)
2. Awedh, M., Mueen, A., Zafar, B., Manzoor, U.: Using socrative and smartphones for the support of collaborative learning. arXiv preprint arXiv:1501.01276 (2015)
3. Bhatnagar, S., Zouaq, A., Desmarais, M.C., Charles, E.: A dataset of learnersourced explanations from an online peer instruction environment. *International Educational Data Mining Society* **13**, 350–355 (2020)
4. Black, P., Wiliam, D.: Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice* **5**(1), 7–74 (Mar 1998)
5. Black, P., Wiliam, D.: Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)* **21**(1), 5 (2009)
6. Bruff, D.: Classroom response systems (“clickers”) (2015)
7. Charles, E.S., Lasry, N., Bhatnagar, S., Adams, R., Lenton, K., Brouillette, Y., Dugdale, M., Whittaker, C., Jackson, P.: Harnessing peer instruction in-and out-of class with mydalite. In: *Education and Training in Optics and Photonics*. p. 11143_89. Optical Society of America, SPIE, Quebec City, Quebec, Canada (2019)
8. Clark, S.: Enhancing active learning: Assessment of poll everywhere in the classroom. Tech. rep., University of Manitoba (2017)
9. Crouch, C.H., Mazur, E.: Peer instruction: Ten years of experience and results. *American journal of physics* **69**(9), 970–977 (2001)
10. Curtis, D.A., Lind, S.L., Boscardin, C.K., Dellings, M.: Does student confidence on multiple-choice question assessments provide useful information? *Medical education* **47**(6), 578–584 (2013)
11. Davis, M.: Technology fed growth in formative assessment. *Education Week* p. 11 (2015)
12. Double, K.S., McGrane, J.A., Hopfenbeck, T.N.: The impact of peer assessment on academic performance: A meta-analysis of control group studies (2020)

13. Elastic: <https://elaastic.irit.fr>, dernière consultation: 2021-04-23.
14. Ellis, C.: Broadening the scope and increasing the usefulness of learning analytics: The case for assessment analytics. *British Journal of Educational Technology* **44**(4), 662–664 (2013)
15. Everett, B.: An introduction to latent variable models. Springer Science & Business Media (2013)
16. Ghasemi, A., Zahediasl, S.: Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism* **10**(2), 486 (2012)
17. Hattie, J.: Visible learning for teachers: Maximizing impact on learning. Routledge, 711 Third Avenue, New York, NY 10017 (2012)
18. Ismail, M.A.A., Mohammad, J.A.M.: Kahoot: A promising tool for formative assessment in medical education. *Education in Medicine Journal* **9**(2) (2017)
19. Krause, J.M., O'Neil, K.: Assessment Tool for K–12 and PETE Professionals. *Strategies, A Journal for Physical and Sport Educators* **30**, 7 (2017)
20. Lebis, A.: Capitaliser les processus d'analyse de traces d'apprentissage: modélisation ontologique & assistance à la réutilisation. Ph.D. thesis, Sorbonne Université (2019)
21. Martinez, M.E., Lipson, J.I.: Assessment for learning. *Educational Leadership* **46**(7), 73–75 (1989)
22. Mazur, E., Watkins, J.: Just-in-time teaching and peer instruction. In: *Just-in-time Teaching: Across the Disciplines, Across the Academy*, pp. 39–62. Stylus Publishing, LLC, 22883 Quicksilver Drive, Sterling, Virginia 20166-2102 (2010)
23. Meltzer, D.E., Manivannan, K.: Transforming the lecture-hall environment: The fully interactive physics lecture. *American Journal of Physics* **70**(6), 639–654 (2002)
24. Olsson, U.: Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* **44**(4), 443–460 (1979)
25. Parmentier, J.F.: How to quantify the efficiency of a pedagogical intervention with a single question. *Physical Review Physics Education Research* **14**(2), 020116 (2018)
26. Parmentier, J.F., Silvestre, F.: La (dé-)synchronisation des transitions dans un processus d'évaluation formative exécuté à distance : impact sur l'engagement des étudiants. In: 9ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2019). pp. 97–108. ATIEF, Sorbonne Université, LIP6Paris, France (2019)
27. Potter, T., Englund, L., Charbonneau, J., MacLean, M.T., Newell, J., Roll, I., et al.: Compar: A new online tool using adaptive comparative judgement to support learning with peer feedback. *Teaching & Learning Inquiry* **5**(2), 89–113 (2017)
28. Sadler, D.R.: Formative assessment and the design of instructional systems. *Instructional Science* **18**(2), 119–144 (Jun 1989)
29. Sawilowsky, S.S.: New effect size rules of thumb. *Journal of Modern Applied Statistical Methods* **8**(2), 26 (2009)
30. Spector, J.M., Ifenthaler, D., Sampson, D., Yang, J.L., Mukama, E., Warusavitarana, A., Dona, K.L., Eichhorn, K., Fluck, A., Huang, R., et al.: Technology enhanced formative assessment for 21st century learning. *International Forum of Educational Technology and Society* **19**(3), 58–71 (2016)
31. Topping, K.J.: Peer assessment. *Theory into practice* **48**(1), 20–27 (2009)
32. Tullis, J.G., Goldstone, R.L.: Why does peer instruction benefit student learning? *Cognitive Research: Principles and Implications* **5**(1), 15 (Dec 2020)

Apprentissage de représentations pour l'enseignement de la programmation : une approche centrée enseignant

Guillaume Cleuziou^{1,2,3}, Frédéric Flouvat¹, Matthieu Exbrayat², Julien Robert³ et Romuald Thion^{1,4}

¹ Université de la Nouvelle-Calédonie, ISEA EA 7484, Nouméa
`prenom.nom@unc.nc`

² Université d'Orléans, INSA Centre Val de Loire, LIFO EA 4022, Orléans

³ Université d'Orléans, IUT Département Informatique, Orléans
`prenom.nom@univ-orleans.fr`

⁴ Université Lyon 1, LIRIS CNRS, Villeurbanne
`prenom.nom@univ-lyon1.fr`

Résumé. L'utilisation de plates-formes de validation de programmes est de plus en plus fréquente pour l'apprentissage des bases de l'informatique. Si l'automatisation des tests de validation libère du temps pour l'enseignant et ludifie le parcours de l'apprenant, on observe également certains biais induits par ces dispositifs pédagogiques. Des recherches en Intelligence Artificielle (IA) visent à pallier certains de ces biais. Pour cela, elles se heurtent au problème de l'analyse informatique de ces programmes par les algorithmes d'IA. Cet article traite d'une expérience pratique d'usages pédagogiques d'une plate-forme d'entraînement à la programmation. Nous présentons une nouvelle méthode pour construire des représentations vectorielles des programmes. Dans une approche centrée 'enseignant', nous proposons enfin d'exploiter ces nouveaux espaces de représentation pour détecter automatiquement des programmes validés par la plate-forme mais devant être signalés à l'enseignant.

Mots-clés : Enseignement de la programmation, Apprentissage de représentations, Réseaux de Neurones, SVM.

Abstract. The use of program validation platforms is becoming more and more common for learning the basics of computer science. While the automation of validation tests frees up time for the teacher and facilitates the learner's path, there are also certain biases induced by these pedagogical tools. Numerous works in Artificial Intelligence (AI) aim to overcome some of these biases. For that, they come up against the problem of the analysis of these programs by AI algorithms. In this paper, we rely on a practical experience of pedagogical uses of a programming training platform. We present a new method to construct vectorial representations of programs. In a teacher-centered approach, we finally propose to use these new representation spaces to automatically detect programs validated by the platform but which must be signaled to the teacher.

Keywords: Coding teaching, Representation Learning, Neural Networks, Support Vector Machine (SVM).

1 Introduction

L'utilisation de plates-formes d'apprentissage s'est considérablement développée ces dernières années au sein des formations en informatique, en particulier pour l'enseignement des fondamentaux de l'algorithmique et de la programmation [5, 2, 3]. Celles-ci prennent le plus souvent la forme d'une interface Web sur laquelle un apprenant soumet un programme comme solution à un exercice conçu par son enseignant ; il reçoit, en réponse, le résultat d'une évaluation/validation automatique de son programme. Cet article s'intéresse à cette pratique pédagogique dans sa dimension formative (et non évaluative), autrement dit dans ce qu'elle offre comme leviers pour améliorer l'acquisition des compétences pour l'apprenant.

Dans ce contexte, les intérêts de recourir à ces plates-formes sont multiples et principalement d'ordres quantitatif et qualitatif : l'automatisation d'une partie de l'accompagnement pédagogique peut permettre le pilotage d'un groupe plus large d'apprenants ; déléguer à la plate-forme la validation des programmes permet d'améliorer qualitativement l'accompagnement en consacrant plus de temps à l'analyse/remédiation sur des aspects plus fondamentaux. Du point de vue de l'apprenant, ce dispositif pédagogique apporte une dimension ludique et une autonomie susceptibles de favoriser son implication et sa progression.

Les plates-formes d'entraînement à la programmation présentent cependant des biais importants qui constituent de véritables verrous pédagogiques, opposables à cette modalité d'apprentissage [10]. Le risque majeur pour l'enseignant est de manquer de visibilité et d'exhaustivité dans son appréhension et sa maîtrise du groupe. On observe également un risque pour l'apprenant de s'enfermer progressivement dans une pratique de type essai-erreur improductive et susceptible d'induire une forme de découragement voire de désapprentissage.

Pour y remédier, les recherches actuelles s'orientent vers le dépassement de la simple tâche de validation et tentent d'introduire un réel accompagnement pédagogique, notamment au moyen de *feedbacks* plus nombreux, plus précis et générés automatiquement [11]. Cette ambition s'avère cependant assez limitée en pratique. Ces recherches s'appuient essentiellement sur des techniques d'Intelligences Artificielles (IA) qui se heurtent rapidement au problème de l'analyse des programmes informatiques par les algorithmes d'IA, ces derniers requérant généralement des données (représentations) vectorielles. Au-delà de ce verrou technique, ces recherches reposent sur l'hypothèse forte et risquée que l'expertise pédagogique d'un enseignant spécialisé en informatique peut être (au moins en partie) automatisée et ainsi mimée par une IA. Nous jugeons cette hypothèse 'risquée' pédagogiquement puisque la moindre erreur, par exemple dans un *feedback* généré, impactera directement l'apprenant.

Dans ce contexte, les contributions présentées dans cet article sont les suivantes : à partir de l'observation des usages d'apprenants et d'enseignants sur une plate-forme d'apprentissage, nous proposons d'exploiter des techniques récentes d'IA avec un point de vue centré 'enseignant'. L'objectif devient ainsi d'offrir à l'enseignant une assistance au pilotage et au développement de sa pratique pédagogique à travers la plate-forme. Ce positionnement présente trois avantages majeurs :

1. répondre conjointement aux deux verrous pédagogiques mentionnés : redonner de la visibilité à l'enseignant et lui permettre ainsi d'intervenir efficacement auprès des étudiants en situation de blocage,
2. éliminer les risques pédagogiques : l'enseignant conservant la maîtrise des interventions auprès des apprenants,
3. s'affranchir de l'hypothèse selon laquelle il existerait *une unique* pratique d'accompagnement pédagogique qui serait de surcroît automatisable.

Cette étude s'appuie sur une plate-forme d'entraînement à la programmation dont nous présentons le principe général de fonctionnement en section 2. Nous discutons des principaux biais pédagogiques observés (Section 3). Nous réalisons ensuite une preuve de concept sur la possibilité d'introduire de l'IA dans une approche centrée enseignant. Cette contribution repose sur une nouvelle méthode d'apprentissage machine (*machine learning*) [4] adaptée à l'analyse de programmes informatiques (section 4). Celle-ci est ensuite utilisée pour détecter automatiquement des programmes validés par la plate-forme mais sur lesquels un enseignant pourrait juger intéressant d'intervenir. Cette expérimentation encourageante, menée sur plusieurs milliers de programmes Python d'étudiants débutants en programmation, est présentée en Section 5.

2 Plate-forme d'entraînement à la programmation

Cette étude repose sur l'usage pédagogique d'une plate-forme d'entraînement à la programmation (ou exerciceur) développée à l'IUT d'Orléans. Elle prend la forme d'une interface Web sur laquelle les étudiants saisissent et évaluent leurs solutions aux exercices proposés. Bien que la plate-forme offre un large spectre de langages (Python, Java, SQL, etc.), nous limitons notre étude à l'apprentissage de la programmation et illustrons ce travail avec le langage Python.

Cet exerciceur est utilisé pour des enseignements d'initiation à l'algorithmique et à la programmation, depuis 2016 par des étudiants en 1ère année de DUT informatique à Orléans (environ 120 étudiants chaque année) et depuis 2020 par des étudiants en 1ère année de Licence Informatique à Nouméa (environ 60 étudiants chaque année). Signalons que même si le contenu de ces enseignements considère les bases de l'algorithmique/programmation, le niveau de maîtrise et de pratique des étudiants intégrant ces formations est hétérogène. Certains étudiants ont déjà acquis des connaissances en programmation au Lycée ou en autodidactes, alors que d'autres n'ont jamais pratiqué la programmation.

L'exerciceur Python implémenté sur la plate-forme permet une évaluation automatique de *fonctions* Python à partir d'*entrées* pré-définies (cas de tests). Ainsi, lors du dépôt d'un exercice, l'enseignant doit fournir :

- un *énoncé* : consigne expliquant la tâche que doit réaliser la fonction,
- une *solution* : proposition d'un code solution,
- une série d'*entrées visibles* : cas de tests illustrant les résultats attendus de la fonction (affichés à la suite de l'énoncé) et servant à construire les *feedbacks*,

- une série d'*entrées invisibles* : cas de tests utilisés lors de l'évaluation du programme de l'étudiant mais qui ne sont pas portés à sa connaissance⁵.

Exercice 2.1: Minimum

Ecrire une fonction `minimum` qui retourne le minimum d'une liste de nombres passée en paramètre.

- Sur l'entrée `[12, 1, 25, 7]` votre solution doit renvoyer `1`.
- Sur l'entrée `[-3, 7, 18, -12]` votre solution doit renvoyer `-12`.
- Sur l'entrée `[-12, -1, -25, -2]` votre solution doit renvoyer `-25`.
- Sur l'entrée `[]` votre solution doit renvoyer `None`.

```

1 def minimum(liste):
2     res = 0
3     for i in range(1, len(liste)):
4         if liste[i] < res:
5             res = liste[i]
6     return res

```

Votre fonction ne fait pas ce qui est attendu

- Sur l'entrée `[12, 1, 25, 7]` votre programme a renvoyé `0` alors qu'on attendait `1`
- Sur l'entrée `[]`, votre programme a renvoyé `0` alors qu'on attendait `None`
- Votre fonction ne renvoie pas ce qui est attendu pour au moins une entrée invisible

Fig. 1. Interface Web de la plate-forme d'entraînement à la programmation.

Dans l'interface présentée à l'étudiant (Figure 1), le menu à gauche de l'écran permet de naviguer entre les exercices. Pour chaque exercice, l'étudiant dispose de l'énoncé complété d'exemples basés sur les entrées visibles. Il dispose ensuite d'une zone de texte pour saisir son programme puis d'un bouton "Envoyer" pour déclencher l'évaluation automatique de sa proposition. Celle-ci consiste à comparer les résultats d'exécution du programme de l'étudiant d'une part et la solution de l'enseignant d'autre part sur chacune des entrées visibles et invisibles. Un feedback est alors affiché, explicitant notamment les erreurs (différences) observées sur les entrées visibles. Dans l'exemple illustré en Figure 1, l'étudiant a défini une fonction *minimum* syntaxiquement correcte ; cependant son algorithme présente une erreur dans l'initialisation de la variable `res` et ne gère pas le cas où la `liste` passée en paramètre est vide. Le feedback affiché au terme de l'évaluation oriente l'étudiant vers la compréhension de ses deux erreurs.

La plate-forme a été librement mise à disposition des enseignants de sorte que l'on observe une grande diversité dans les modalités pratiques de son intégration dans les enseignements. Certains enseignants l'ont adoptée comme environnement de programmation lors des travaux pratiques, d'autres ont également systématisé son usage mais seulement pour la validation, enfin une autre pra-

⁵ L'ajout d'entrées invisibles permet d'éviter la validation d'un programme centré exclusivement sur les exemples portés à la connaissance de l'étudiant (entrées visibles).

tique a consisté à circonscrire l'usage de l'exerciseur comme un outil d'aide au travail personnel (en dehors des heures de formation). Depuis 2020, les traces d'utilisation de la plate-forme sont conservées à des fins d'exploitation scientifique et pédagogique. Nous conservons les programmes soumis pour évaluation⁶ ainsi que les méta-données potentiellement utiles à l'analyse et la modélisation des processus d'apprentissage (identifiant étudiant, date/heure). La diversité d'usages a un impact significatif sur la nature des données récoltées.

3 Discussion sur la modalité pédagogique

Le dispositif pédagogique présenté offre des avantages indéniables parmi lesquels : d'une part, pour l'étudiant, une plus grande autonomie et une ludicisation de son apprentissage et d'autre part, pour l'enseignant, une délégation à la machine de la tâche répétitive de validation des programmes⁷, profitable à sa pédagogie. Il nous paraît néanmoins essentiel de mener une discussion concernant les biais inhérents à ce dispositif, observés en situation réelle d'apprentissage. Les verrous pédagogiques qu'ils révèlent ne remettent pas en cause le recours à ces plateformes mais constituent en revanche autant de pistes d'amélioration.

Tout d'abord, bien que l'outil numérique soit présenté à l'apprenant comme une solution d'apprentissage et non d'évaluation, le principe même de ludicisation *via* la validation automatique des solutions peut, chez certains apprenants, favoriser un comportement de type '*gamer*'. Il se met en quête de *trouver* la solution, y compris auprès d'autres apprenants, sans chercher à la *construire*.

Ensuite, il convient d'observer que le 'dialogue pédagogique' entre l'apprenant et la machine ne porte que sur des programmes complets. Contrairement à l'enseignant qui peut accompagner l'étudiant dans l'écriture de son programme, l'exerciseur n'intervient quant à lui qu'une fois un (premier) programme totalement écrit. Ceci précise les conditions dans lesquelles cet exerciseur doit être utilisé : en parallèle d'un accompagnement (synchrone) par l'enseignant et/ou limité à la réalisation de programmes courts ; cette dernière condition étant généralement vérifiée dans des enseignements d'initiation à la programmation.

L'observation des traces d'activités des étudiants sur la plate-forme met en évidence des situations de blocage, des étudiants réalisant parfois plusieurs dizaines de tentatives pour un même exercice⁸ et parfois sans parvenir à une solution valide. Par exemple, pour un exercice on observe un étudiant qui a réalisé au total 98 tentatives - en deux périodes de temps (43 tentatives en moins d'une heure puis 55 tentatives sur 1h30 deux jours plus tard) - sans réussite finale. Au-delà du risque de découragement induit par ce comportement de type 'essai-erreur' improductif [8], il est plus inquiétant d'observer des situations de 'désapprentissage' pour lesquelles l'étudiant dégrade une première solution initialement bien conçue et proche d'une solution valide. L'analyse de sa séquence de

⁶ Les traces sont conservées de façon anonyme et avec l'accord des utilisateurs assurant au passage une couverture plus large et systématique.

⁸ jusqu'à 174 tentatives réalisées en 3 jours.

tentatives révèle sa propension à remettre en cause toutes notions préalablement acquises, qu'elles soient d'ordre conceptuelle ou purement syntaxique.

Enfin, lorsque l'enseignant analyse les programmes validés par la plate-forme, il constate une part significative de solutions non souhaitées (environ 1/3 pour notre expérience). Il peut s'agir de solutions fausses (~15% des cas) révélant la nécessité de corriger les spécifications de l'exercice (solution enseignant, entrées visibles et/ou invisibles) mais aussi parfois de bonnes solutions que l'enseignant souhaiterait ne pas accepter (~85% des cas) parce qu'il s'agit de mauvaises pratiques ou d'un contournement de la compétence ciblée par l'exercice (p.ex. l'utilisation d'une fonction de l'API au lieu du codage de l'algorithme associé).

Ces différents biais observés - *non construction des propositions*, *accompagnement limité*, *blocage*, *désapprentissage* et *validations non-souhaitées* - constituent, de notre point de vue, de véritables verrous pédagogiques qui, bien qu'ils existent dans des situations classiques d'apprentissage, sont susceptibles d'être accentués par l'usage d'outils numériques de type exerciceur. Une manière de surmonter ces difficultés consiste à replacer l'enseignant au centre du dispositif. En le considérant comme expert métier, il s'agit alors d'exploiter les données générées par ses apprenants pour l'aider à décider efficacement des interventions pédagogiques nécessaires. Dans tout processus d'aide à la décision, il est utile de proposer à l'expert une représentation/visualisation précise des données. Dans notre contexte, cela pose le problème de la représentation des programmes.

4 Apprentissage de représentations de programmes

L'utilisation des plates-formes d'entraînement à la programmation soulève donc un certain nombre de défis pédagogiques, mais elle ouvre aussi de nouvelles perspectives grâce à la multitude de données collectées. Par exemple, [2] exploitent des méthodes de fouille de texte pour prédire la justesse d'une réponse donnée par un étudiant (un programme Python). [11] utilisent quant à eux des réseaux de neurones pour propager automatiquement les *feedbacks* des enseignants. [12] utilisent les données collectées pour prédire les erreurs faites par les étudiants. De manière générale, les approches basées sur des représentations vectorielles, appelées encore *embeddings*, suscitent beaucoup d'intérêts [9].

Le principe de ces approches est de représenter chaque programme par un vecteur de réels (*embedding*). Cette représentation est ensuite utilisée en entrée d'un algorithme d'apprentissage pour construire un modèle prédictif. L'intérêt de ces *embeddings* est de projeter (ou 'plonger') tout un vocabulaire dans un espace vectoriel de faible dimension, tout en capturant des relations sémantiques complexes à partir de simples opérations sur ces vecteurs (p.ex. $\text{vecteur}(\text{"king"}) - \text{vecteur}(\text{"man"}) + \text{vecteur}(\text{"woman"}) \approx \text{vecteur}(\text{"queen"})$). Différentes approches ont été proposées pour construire des *embeddings* de programmes [1]. Toutefois, elles ont encore du mal à capturer totalement la sémantique sous-jacente, notamment les aspects fonctionnels et de style.

Face aux limites des approches existantes, nous avons proposé l'approche *code2aes2vec* exploitant instructions, structure du code et traces d'exécution,

afin de construire des *embeddings* plus fins. Nous présentons brièvement cette approche dans la suite de cette section, une description détaillée est présente dans [4]. Elle procède en deux étapes : 1) représentation des programmes sous forme de séquences d'exécution abstraites (étape *code2aes*); 2) utilisation de l'algorithme *doc2vec* [6] pour apprendre des *embeddings* de programmes à partir de ces séquences (étape *aes2vec*). Contrairement aux autres propositions, cette approche générique et non-supervisée intègre des éléments de fonctionnalités, de style et d'ordinaire. Ce dernier point est primordial dans notre contexte car il faut pouvoir différencier des programmes répondant à un même exercice (i.e. implémentant les mêmes fonctionnalités) mais de façons différentes (en terme de stratégie ou d'efficacité).

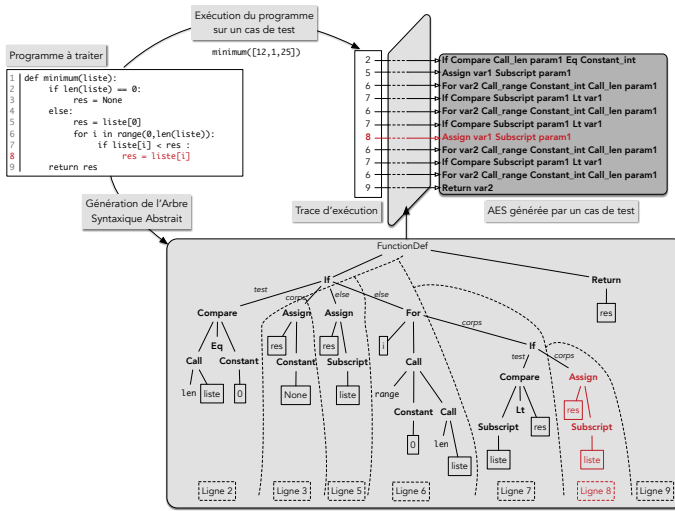


Fig. 2. Processus de construction des AES à partir d'un cas de test et d'un programme Python soumis en réponse à l'exercice 'rechercher le minimum dans une liste'.

L'étape *code2aes* consiste à traduire un programme sous forme de séquences d'exécution abstraites, ou AES (*Abstract Execution Sequence*). Chaque AES traduit la trace d'exécution du programme sur un cas de test (Figure 2), où chaque instruction est représentée par les "mots" du sous-arbre associé dans l'AST (parcours en profondeur). L'AES finale d'un programme est obtenue par concaténation des AES résultantes de chaque cas de test.

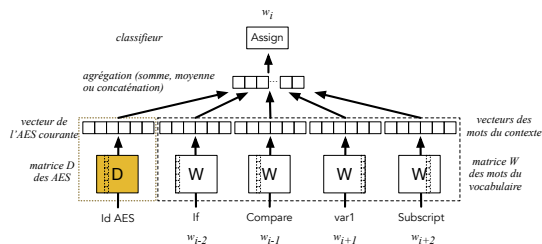


Fig. 3. Réseau de neurones utilisé pour prédire un mot w_i à partir de l'identifiant de son AES d'origine et de son contexte (deux mots précédents et deux mots suivants).

Ensuite, l'étape *aes2vec* exploite l'approche *doc2vec* [6] pour construire les *embeddings* des programmes à partir de leur AES. Pour cela, elle utilise un réseau de neurones (Figure 3). Ce réseau est entraîné à prédire le mot courant w_i d'une AES à partir de l'identifiant de l'AES, des mots précédents et des mots suivants. Chaque AES est associée à une colonne d'une matrice D , et chaque mot à une colonne d'une matrice W . Ces vecteurs sont ensuite agrégés par concaténation pour prédire le mot courant en utilisant un classifieur multi-classes du type *softmax*. Les paramètres sont mis à jour par descente de gradient stochastique. Une fois le modèle entraîné, la matrice D renferme les *embeddings* des programmes utilisés pour cet entraînement.

5 Exploitation des *embeddings* de programmes

Nous avons montré dans [4] que les *embeddings* appris avec *code2aes2vec* étaient suffisamment précis pour identifier la fonctionnalité d'un programme mais également pour distinguer des variantes de style. Ainsi, sur un corpus de 5690 programmes extrait des traces d'activités de nos étudiants sur 66 exercices, un simple classifieur supervisé (plus proches voisins ou SVM) permet d'identifier correctement l'exercice associé à un programme dans plus de 80% des cas.

Ces *embeddings* permettent aussi d'offrir à l'enseignant une visualisation globale des programmes de ses apprenants. L'utilisation d'un algorithme de réduction de dimensions adapté aux espaces d'*embeddings* (p.ex. une projection non-linéaire telle que proposée par la méthode t-SNE [7]), permet en effet de générer une carte en deux dimensions (ou plus si besoin) dans laquelle les programmes s'organisent selon leur typologie (syntaxe et fonctionnement). La Figure 4 illustre ce résultat sur l'un des exercices du corpus des 5690 programmes.

Pour aller plus loin, nous présentons dans cet article une preuve de concept montrant de façon concrète qu'il est possible d'exploiter les représentations issues

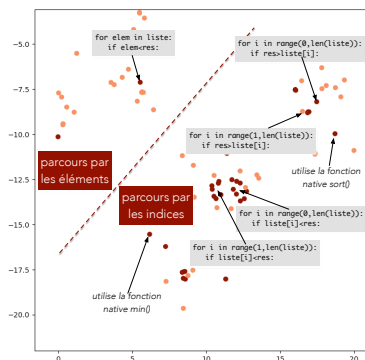


Fig. 4. Exemple de visualisation des programmes soumis par les étudiants pour l'exercice "rechercher le minimum dans une liste". Chaque point représente un programme soumis valide (en rouge) ou invalide (en rose).

de *code2aes2vec* pour aider l'enseignant à cibler efficacement ses interventions. Nous considérons plus particulièrement le problème - identifié en Section 3 - des *validations non-souhaitées*. Pour cela, nous avons demandé à des enseignants en informatique d'étiqueter les programmes validés par la plate-forme.

Parmi les 5690 programmes qui composent le corpus, 1304 ont été validés par la plate-forme (23%). Après suppression des doublons, l'ensemble S des exemples est circonscrit à 727 programmes. Chaque programme a reçu trois avis selon que l'enseignant juge sa validation non-souhaitable (il aurait voulu intervenir = exemple positif) ou souhaitable (il n'aurait pas jugé utile d'intervenir = exemple négatif). Une analyse rapide des étiquetages met en évidence à la fois :

- une expertise métier majoritairement partagée : les trois enseignants sont en parfait accord pour la majorité des exemples (54%),
- une différenciation pédagogique certes minoritaire mais réelle : sur 46% des exemples, la nécessité d'une intervention de l'enseignant n'est pas unanime.

Dans la suite de l'expérience nous avons considéré comme exemples positifs S^+ les programmes pour lesquels deux enseignants au moins jugent une intervention nécessaire, S^- étant de fait composé des programmes sur lesquels au moins deux d'entre eux ne jugent pas utile d'intervenir. S^+ représente 32% des exemples (234 programmes) et S^- renferme 68% des exemples (493 programmes). Nous avons ensuite cherché à apprendre un modèle de détection automatique des exemples positifs de sorte à utiliser ce modèle pour alerter l'enseignant sur des programmes susceptibles de nécessiter une intervention de sa part. S'agissant d'une tâche de classification supervisée (à deux classes), l'ensemble S a été

séparé en un ensemble d'apprentissage (80%) et un ensemble de test (20%). Pour représenter les programmes, nous avons utilisé un modèle d'*embeddings* appris avec l'algorithme *code2aes2vec* sur le corpus de 5690 programmes. Celui-ci a été paramétré de sorte à générer des vecteurs de représentation de taille 100. Nous avons considéré deux manières de représenter les programmes :

1. Sans recentrage : chaque programme $x \in S$ est représenté directement par son vecteur d'*embedding* $\phi(x) \in \mathbb{R}^{100}$ inféré par le modèle appris
2. Avec recentrage : les programmes sont recentrés par exercice tel que si s_i désigne la solution enseignant pour l'exercice $i \in \{1, \dots, n\}$ et $E(x) : S \rightarrow \{1, \dots, n\}$ la fonction associant chaque programme x à son exercice d'origine, la représentation $\phi'(x)$ d'un programme est donnée par $\phi(x) - \phi(s_{E(x)})$.

En considérant que l'espace des *embeddings* est organisé en groupes de programmes associés à un même exercice (hypothèse vérifiée dans [4]), l'opération de recentrage consiste à repositionner chacun des groupes *via* une translation de la solution enseignant pour l'exercice correspondant. Avec recentrage, la tâche de classification vise à apprendre un modèle générique indépendant de l'exercice.

Tableau 1. Taux de bonne classification calculé sur l'ensemble de test selon le type de représentation et de noyau considéré.

Classifieur	Sans recentrage	Avec recentrage
SVM linéaire	0.712	0.753
SVM polynomial	0.753	0.760
SVM RBF	0.719	0.829

Nous avons utilisé plusieurs algorithmes de classification supervisée et nous présentons en particulier les résultats obtenus par la méthode SVM consistant à apprendre des frontières de décision maximisant la marge entre les exemples des deux classes. La forme des frontières de décision diffère selon le type de noyau utilisé pour le SVM. Les modèles de classification ont été appris sur les données d'entraînement uniquement. Les taux de bonne classification obtenus sur les données de test montrent (Tableau 1) qu'il est possible de réaliser cette tâche de détection avec une précision supérieure à 80%, à condition d'une part de procéder effectivement au recentrage des données et d'autre part d'utiliser un noyau approprié. En l'occurrence, le noyau RBF semble adapté car il permet de construire des frontières correspondant à des zones convexes centrées sur des exemples (points de supports) positifs et positionnées à l'écart des exemples négatifs, situés quant à eux majoritairement au centre du nuage de points.

La Figure 5 confirme ces explications en présentant une visualisation des frontières de décision apprises sur une projection bi-dimensionnelle⁹ (obtenue par l'algorithme t-SNE [7]) des *embeddings* de programmes.

⁹ Les taux de bonne classification reportés sur la figure sont inférieurs à ceux du Tableau 1 du fait de cette projection.

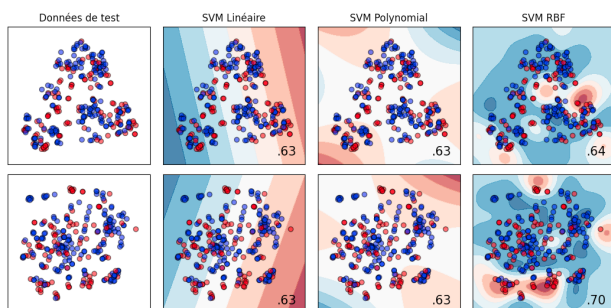


Fig. 5. Visualisation des frontières de décision apprises en fonction du type de noyau considéré et du type de représentation utilisé : sans recentrage des *embeddings* (ligne du haut) ou avec recentrage (ligne du bas). Positionnement des données de test : exemples positifs (rouge)/négatifs (bleu), et taux de bonne classification (coin inf. droit).

Au-delà du score de classification, il est utile pour l'enjeu applicatif, de détailler la validité intrinsèque du meilleur modèle appris (SVM RBF avec recentrage des *embeddings*). Le taux de bonne classification obtenu (0.829) correspond à la répartition suivante :

- 25 vrais positifs : programmes correctement identifiés comme nécessitant une intervention,
- 2 faux positifs : programmes identifiés à tort comme nécessitant une intervention,
- 23 faux négatifs : programmes nécessitant une intervention mais non détectés,
- 96 vrais négatifs : programmes ne nécessitant pas d'intervention et identifiés comme tel.

Il en résulte un modèle ayant une forte capacité à ne soumettre à l'enseignant que des programmes nécessitant son intervention (taux de spécificité de 0.925). Ce résultat très encourageant valide cette preuve de concept visant à assister efficacement l'enseignant à partir d'un apprentissage de représentations fines des programmes de ses apprenants.

6 Conclusion

Ce travail montre l'intérêt de l'apprentissage de représentations (*embeddings*) pour proposer de nouveaux outils de suivi à l'enseignant, et ainsi le remettre au cœur des plates-formes d'entraînement à la programmation. L'approche présentée s'appuie sur la construction d'*embeddings* sémantiquement riches pour

identifier des solutions justes mais nécessitant l'intervention de l'enseignant. La précision des résultats obtenus à ce niveau (82.9% de bonne classification avec seulement 2 faux positifs) illustre le potentiel de cette approche pour assister l'enseignant dans ses interventions (gain en temps et précision).

Les perspectives de ce travail sont nombreuses. Tout d'abord, la collecte de données complémentaires et introduisant d'autres langages de programmation permettrait d'évaluer la robustesse de l'approche et de tester d'autres scénarios pour détecter les solutions "valides mais non souhaitées". Il pourrait s'agir de construire un modèle par exercice (au lieu d'un modèle pour tous les exercices). De manière plus générale, ces *embeddings* de programmes ouvrent de nombreuses perspectives pour l'aide à l'enseignement. Ils pourraient être utilisés pour identifier des typologies d'erreurs, des solutions alternatives, ou des étudiants en situation de *blocage* voire de *désapprentissage* via l'analyse de leurs trajectoires.

References

1. Allamanis, M., Barr, E.T., Devanbu, P., Sutton, C.: A survey of machine learning for big code and naturalness. *ACM Computing Surveys* **51**(4), 1–37 (2018)
2. Azcona, D., Arora, P., Hsiao, I.H., Smeaton, A.: user2code2vec: Embeddings for profiling students based on distributional representations of source code. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge. pp. 86–95 (2019)
3. Broisin, J., Herouard, C.: Soutien à l'apprentissage de la programmation: conception et évaluation d'un indicateur sémantique. In: 9e Conference Environnements Informatiques pour l'Apprentissage Humain (EIAH 2019). pp. 235–246 (2019)
4. Cleuziou, G., Flouvat, F.: Apprentissage d'embeddings de codes pour l'enseignement de la programmation : une approche fondée sur l'analyse des traces d'exécution. In: 21ème conférence Extraction et Gestion des Connaissances (2021)
5. Ihantola, P., Vihavainen, A., Ahadi, A., Butler, M., Börstler, J., Edwards, S.H., Isohanni, E., Korhonen, A., Petersen, A., Rivers, K., et al.: Educational data mining and learning analytics in programming: Literature review and case studies. In: Proceedings of the 2015 ITiCSE on Working Group Reports, pp. 41–63 (2015)
6. Le, Q., Mikolov, T.: Distributed representations of sentences and documents. In: International conference on machine learning. pp. 1188–1196 (2014)
7. LJPvd, M., Hinton, G.: Visualizing high-dimensional data using t-sne. *J Mach Learn Res* **9**, 2579–2605 (2008)
8. Marzin-Janvier, P.: Des outils numériques pour apprendre les sciences (2020), <https://hal.archives-ouvertes.fr/hal-03096594>
9. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 (2013)
10. Panigrahi, R., Srivastava, P.R., Sharma, D.: Online learning: Adoption, continuance, and learning outcome—a review of literature. *International Journal of Information Management* **43**, 1–14 (2018)
11. Piech, C., Huang, J., Nguyen, A., Phulsuksombati, M., Sahami, M., Guibas, L.: Learning program embeddings to propagate feedback on student code. In: Proceedings of the 32nd International Conference on Machine Learning. p. 1093–1102. ICML'15, JMLR.org (2015)
12. Wang, K., Singh, R., Su, Z.: Dynamic neural program embeddings for program repair. In: International Conference on Learning Representations (2018)

Suivi longitudinal de la fluence en lecture par évaluation automatique de la parole

Erika Godde^{1,2}, Gérard Bailly¹, Marie-Line Bosse²

¹ Université Grenoble Alpes, CNRS, Grenoble INP, GIPSA-Lab, 38000 Grenoble, France
erika.godde@gipsa-lab.fr gerard.bailly@gipsa-lab.fr

² Université Grenoble Alpes, CNRS, LPNC, 38000 Grenoble, France
marie-line.bosse@univ-grenoble-alpes.fr

Résumé. La fluence en lecture à haute voix est couramment évaluée en mesurant le nombre de mots correctement lus par minute. Cette évaluation se fait au détriment de l'expressivité et du phrasé, pourtant nécessaire à la compréhension, mais complexes à évaluer car très subjectifs. Nous proposons ici d'évaluer automatiquement les quatre dimensions de la fluence en lecture : décodage, vitesse, phrasé et expressivité. Cet outil s'appuie sur une analyse automatique des paramètres acoustiques, notamment par comparaison à de multiples lecteurs experts, et sur une échelle d'évaluation subjective adaptée pour le français. Nous avons pu montrer que cet outil est plus précis et consistant que des juges humains. Il permet de plus de s'affranchir de la variabilité inter-juges. Il est donc particulièrement adapté à un suivi longitudinal des enfants. Cet outil a été utilisé pour évaluer les lectures de 49 enfants suivis du CE1 au CM1. Ces analyses ont permis de mettre en évidence les profils de développement de chaque dimension de la fluence en lecture.

Mots-clés : Fluence en lecture, prosodie en lecture, évaluation automatique, expressivité

Abstract. Reading fluency is usually assessed with the number of words correctly read in a minute. Expressivity and phrasing are then omitted, probably because these dimensions are more subjective and complex to assess. We propose here a new tool to assess the four dimensions of reading fluency : pace, smoothness, phrasing and expressivity. This tool use an automatic analysis of acoustic features and multiple expert references to predict the score to a subjective scale. We show that this tool is valid, precise and more consistent than human raters. This tool is then well adapted to multiple assessments. We used it to assess the reading fluency of 49 young readers three times in CE1, CE2 and CM1. These analyses enable us to characterize the development patterns of the four dimensions of reading fluency.

Keywords: Reading fluency, reading prosody, automatic assessment, expressivity

1 Introduction

La fluence en lecture est couramment évaluée en mesurant le nombre de mots correctement lus par minute (NMCLM). Cependant, cette méthode se concentre sur la mesure du décodage et de la vitesse uniquement, au détriment de la prosodie, expressivité et phrasé. Cette mesure restrictive entretient ainsi la confusion entre lecteur rapide et bon lecteur, créant des lecteurs rapides mais qui ne comprennent pas forcément le texte lu [1]. Cette prévalence du NMCLM dans l'évaluation de la fluence est en partie due à la difficulté d'évaluation des dimensions prosodiques, notamment leurs conséquences sur le phrasé et l'expressivité des lectures. En effet, s'il est aisé d'évaluer la justesse de la prononciation des mots, il y a de nombreuses façons licites de lire un texte avec expressivité.

La prosodie est pourtant une compétence dont l'importance dans l'acquisition de la lecture est de plus en plus reconnue [2][3]. Le lien entre compréhension et prosodie a été montré dans plusieurs langues et à tous les niveaux scolaires [4][5]. Pourtant le développement de la prosodie en lecture a été assez peu étudié [4]. Les études, menées essentiellement en anglais et en espagnol avec des enfants de primaire, montrent que le développement de la prosodie en lecture débute dès les premières années, en parallèle des compétences en décodage et automaticité [6][7]. Le phrasé se développe rapidement du CP au CE2. L'expressivité se met en place plus lentement que les autres dimensions de la fluence et son développement continue au-delà de l'école primaire. Le schéma semble similaire en français. Une étude transversale a montré le développement rapide des dimensions du CE1 au CM2 [8]. Il n'existe à notre connaissance pas d'étude longitudinale en français. Ce type de données permettraient de montrer les étapes de développement de la fluence en lecture et de concevoir des outils d'entraînement adaptés à ces étapes de développement.

Pour mener ces études longitudinales, il est nécessaire d'évaluer la fluence, et notamment ses dimensions prosodiques, de façon fiable à de multiples reprises, pour pouvoir mesurer des variations fines. On trouve deux types d'outils dans la littérature. Tout d'abord les outils d'évaluation subjective. Il s'agit d'échelle de notation sur lesquelles des juges s'appuient pour donner un score à diverses compétences. La plus utilisée dans la littérature est la *Multidimensional Fluency Scale* (MFS) [2]. Elle a été traduite et adaptée pour le français [8]. Cette échelle, que nous appelons *Échelle MultiDimensionnelle de Fluence* (EMDF) en français, permet d'évaluer les quatre dimensions de la fluence : expressivité, phrasé, décodage et vitesse. Les juges donnent un score de 1 à 4 à chacune de ces dimensions. C'est un outil valide, donnant de bons coefficients d'accord entre juges quand le protocole d'utilisation et d'entraînement des juges est strictement respecté [9]. La variabilité inter-juges reste cependant problématique quand on réalise plusieurs évaluations espacées dans le temps. Le deuxième type de mesure de la fluence en lecture est la mesure de paramètres acoustiques [6][7]. Ces mesures sont précises et fiables dans le temps. Elles ne permettent cependant pas de caractériser aisément chaque dimension de la fluence. Dans la littérature, les analyses acoustiques se concentrent essentiellement sur les marqueurs rythmiques, tels que les pauses. Les paramètres mélodiques et d'intensité sont peu exploités [4]. Quand ils le sont, c'est en comparaison avec une lecture

experte unique. Ces analyses se heurtent donc à la problématique de la variété des productions acceptables

Il existe donc un vrai besoin d'outils d'évaluation de la fluence incluant la prosodie permettant un suivi large et fiable des compétences en lecture dans le temps pour la recherche et l'enseignement. Un premier outil proposé par Bolanos et al. [10] utilise à la fois des paramètres verbaux (NMCLM, taux de phonation, taux d'articulation et nombre de répétitions) et acoustiques (localisation de l'accentuation, durée des mots et des syllabes, pauses silencieuses et non silencieuses et lien avec la ponctuation) pour prédire des scores d'enfant de CP au CE2 à une échelle subjective unidimensionnelle. Pour cela, les auteurs utilisent un apprentissage supervisé. Les lectures sont également évaluées à l'aide de cette échelle par deux juges experts. Les auteurs obtiennent de bons résultats avec un coefficient de corrélation entre notes calculées/jugées de 0.86. Ils concluent cependant sur la difficulté de l'évaluation de l'expressivité, dû à un manque d'objectifs simples à atteindre. On peut également noter qu'ils utilisent une échelle unidimensionnelle et ne séparent donc pas les dimensions de prosodie et d'automatisme. La multiplicité des productions acceptables est bien au cœur de la problématique de cette évaluation. Hirst et al. [11] proposent ainsi d'évaluer la prosodie des productions de synthétiseurs vocaux en utilisant une comparaison avec plusieurs références naturelles. Après plusieurs alignements (durée des segments, F0 et intensité), ils utilisent alors comme mesure de la qualité prosodique la valeur moyenne quadratique de la distance entre la production de synthèse et la production naturelle la plus proche. Cette approche ne permet cependant pas de prendre en compte la globalité des productions possibles en se limitant à peu de dimensions et à une seule production cible. Nous avons ici choisi d'utiliser une méthode déjà éprouvée en analyse du regard [12] : la mise à l'échelle multidimensionnelle. Les références sont projetées comme des points dans un espace multidimensionnel. La distance des productions à évaluer à chaque référence est ensuite calculée pour placer ces productions dans l'espace multidimensionnel et obtenir ses coordonnées, caractéristiques de la performance par rapport aux références. Cette méthode utilisée pour l'analyse du regard social peut être transposée pour la lecture.

Dans un premier temps nous présenterons et montrerons la validité d'un nouvel outil automatique de mesure de la fluence en lecture, reposant sur les paradigmes présentés ci-dessus. Cet outil permet de s'affranchir de la variabilité inter évaluateur et de mesurer plus finement chaque dimension de la fluence. Dans un deuxième temps, nous utiliserons cet outil dans le cadre d'une étude longitudinale visant à caractériser le développement des dimensions de la fluence chez les jeunes lecteurs français du CE1 au CM1.

2 Méthodologie

2.1 Participants

Les enfants participants à l'étude sont scolarisés dans deux écoles et un collège de la région grenobloise, soit 14 classes. Ces établissements ne sont pas situés en Réseau

d'Éducation Prioritaire et accueillent un public de milieu socio-économique diversifié. Nous avons obtenu l'aval de la Directrice Académique des Service de l'Éducation Nationale de l'Isère, des Inspectrices de l'Éducation Nationale des circonscriptions concernées, des directrices des écoles et du principal du collège, pour enregistrer les enfants dans les établissements pendant le temps scolaire. Tous les enregistrements ont été réalisés avec l'autorisation écrite des parents et l'accord verbal des enfants.

Corpus transversal (étalonnage de l'outil automatique). Les données ont été récoltées auprès d'enfants du CE1 à la 5e en fin d'année scolaire, en mai et juin. Le corpus comprend 646 lectures recueillies auprès de 323 enfants normo-lecteurs dont au moins une langue maternelle est le français. Vingt adultes lecteurs experts ont également été enregistrés. Deux textes ont été lus : un texte narratif et un dialogue.

Corpus longitudinal. Pour ce corpus 49 enfants ont été enregistrés en CE1, CE2 et CM1 sur la lecture des deux textes utilisés dans le corpus transversal.

Le Tableau 1 présente les caractéristiques des participants.

Tableau 1. *Caractéristiques des participants au corpus transversal et longitudinal. Age moyen (SD) en année:mois(mois).*

transversale	CE1	CE2	CM1	CM2	6 ^e	5 ^e	adultes
sujets	61	47	54	63	33	37	20
filles	27	25	26	23	15	21	10
age	7:11 (3.8)	8:11 (3.4)	9:11 (5.1)	10:11 (4.5)	11:11 (5.4)	12:8 (5.4)	29:6 (85)
longitudinale		CE1	CE2	CM1			
sujets		49	49	49			
filles		23	23	23			
Age		7:11(3.8)	8:7(3.8)	9:1(3.8)			

2.2 Évaluation subjective

Les 744 lectures du corpus transversal ont été évaluées subjectivement en utilisant l'EMDF [8]. Dans cette étude, vingt-sept juges, 13 enseignants, 4 chercheurs et 10 étudiants, ont évalué 744 lectures au total. Nous disposons de 6853 évaluations au total.

2.3 Évaluation objective

L'évaluation objective des lectures se basent sur des paramètres verbaux et non-verbaux. Pour ces analyses acoustiques, l'enregistrement audio est aligné automatiquement avec le texte en utilisant un dictionnaire mis à jour au fil des nouvelles lectures contenant toutes les formes des mots rencontrés, correctes et incorrectes.

Caractérisation des paramètres verbaux. Les paramètres verbaux suivants sont extraits de chaque lecture du corpus transversal :

- nombre de mots incorrects par minute (NMI)
- nombre de mots lus par minute (NMLM)
- nombre de répétition par minute (NR)
- nombre de pauses intra-mots par minute (NPIM)
- nombre de noyaux vocaliques par minute (NV)

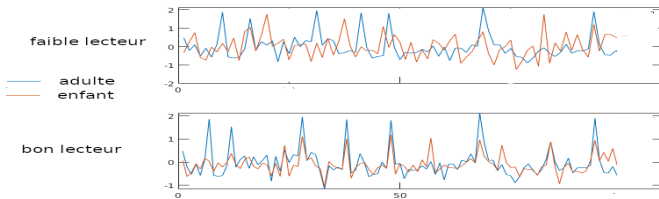


Fig. 1. Alignement de lecture d'enfants (rouge) sur une lecture de référence adulte (bleu) pour l'allongement syllabique. Les pics d'allongements sont dus aux pauses.

Caractérisation des schémas prosodiques. Pour prendre en compte la multiplicité des productions acceptables en termes de prosodie, nous avons choisi de comparer les lectures d'enfants à plusieurs productions de lecteurs experts : les lectures de références sont projetées comme des points dans un espace tridimensionnel. Deux performances sont prises en compte : les schémas mélodiques (basés sur les variations de F0) et les schémas rythmiques (basés sur l'allongement des syllabes). Les lectures des enfants sont ensuite projetées dans l'espace de référence par calcul de leur distance à chaque lecture de référence. Pour cela, les lectures sont alignées (seuls les mots correctement lus sont considérés) et la distance cumulée entre les motifs mélodiques (vs. rythmiques) des syllabes de l'énoncé est calculée (la moyenne des motifs est soustraite (ce sont des échelles logarithmiques) afin de s'affranchir du registre vs. de la vitesse de lecture). La matrice des distances entre la lecture de l'enfant aux lectures adultes et alors projetée dans l'espace de référence. La Fig. 1 présente ces alignements pour un bon lecteur et un lecteur moyen. Finalement les coordonnées des lectures d'enfants dans ces espaces sont utilisées pour caractériser les performances prosodiques des enfants : F0x pour l'espace mélodique et COEX pour l'espace rythmique. La Fig. 2 présente les projections dans le premier plan factoriel pour chacun des schémas prosodiques. On peut voir sur la Fig. 2 les clusters correspondant à différents profils de lecteurs, évalués avec l'EMDF. On constate que chaque profil est placé comme attendu par rapport aux lecteurs de référence : on retrouve bien le groupe des bons lecteurs proches des lectures de références et les groupe des faibles lecteurs plus éloignés. Ces faibles lecteurs sont également plus proches de ref0), lecteur fictif avec une prosodie « plate » (sans variation de F0 ni de durée syllabique).

2.4 Liens entre paramètres objectifs et évaluations subjectives.

Nous avons réalisé des régressions linéaires multiples pour déterminer les relations entre scores subjectifs et paramètres verbaux, puis paramètres verbaux et prosodiques (cf. Tableau 2).

On peut constater que les corrélations entre scores subjectifs et paramètres acoustiques augmentent sensiblement avec l'introduction des paramètres prosodiques dans les régressions, particulièrement pour la dimension expressivité par exemple (.62 à .79 pour le texte 1) et pour la dimension phrasé (.77 à .83 pour le texte 1). L'introduction des paramètres prosodiques apporte donc une information supplémentaire pertinente pour la modélisation des scores subjectifs de fluence. On constate également que les paramètres les plus significatifs pour l'expressivité sont les coefficients du schémas mélodiques et pour le phrasé les coefficients du schéma rythmique. Ces résultats conforte le choix de ces schémas pour la caractérisation de la qualité prosodique des lectures. Finalement, les corrélations sont élevées (entre .8 et .9) ce qui confirme la qualité des modèles présentés en table 2.

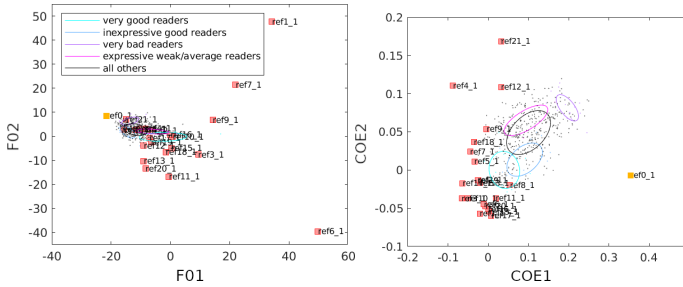


Fig. 2. Projection des lectures d'enfants (points noirs) dans les espaces des lectures de référence (points rouges). Les structures mélodiques et rythmiques sont projetées respectivement à gauche et à droite. Notez la plus grande dispersion des structures rythmiques réalisées par les enfants, assez éloignées de la réalisation monotone (ref0_1).

2.5 Dépendance au texte.

L'évaluation subjective a été réalisée sur les deux textes enregistrés. On constate que les modèles obtenus (cf. 2) pour les deux textes sont proches pour les paramètres verbaux, les différences proviennent principalement des paramètres prosodiques. On peut expliquer cela par la nature différente des textes. Le texte 2 est un dialogue, induisant plus d'expressivité et de rythme qu'un texte narratif comme le texte 1. C'est bien ce qui apparaît dans les modèles présentés dans le 2. Ainsi les modèles peuvent donc différer d'un texte à l'autre mais restent cohérents dans leur structure globale combinant paramètres verbaux et prosodiques.

Tableau 2. Prédicteurs principaux des scores aux évaluations subjectives et corrélation (pseudo R2 de McFaden) pour les régresseurs avec paramètres verbaux et paramètres verbaux et prosodiques. Seuls les régresseurs significatifs sont présentés dans l'ordre d'importance de leur contribution au modèle.

	Paramètres verbaux		Paramètres verbaux + prosodiques	
Texte 1				
expressivité	.62	1+ NV	.79	1+ F01 + COE1 + NMLM
phrasé	.77	1+ NMI + NV	.83	1+ NMLM + COE1
décodage	.75	1+ NMI + NMLM	.76	1+ NMLM + NMI
vitesse	.82	1+ NMLM	.84	1+ NMLM
Texte 2				
expressivité	.65	1+ NV	.83	1+ COE1 + NMLM + F02 + F01
phrasé	.81	1+ NMI + NV	.91	1+ COE1 + NV + COE3 + NMI + F03
décodage	.86	1+ NMI + NV	.88	1+ NV + NMI + COE1
vitesse	.89	1+ NV	.91	1+ NV + COE1

La Fig. 3 compare les deux textes pour les juges humains et les scores prédits par les modèles, pour chacune des dimensions. Les juges humains tendent à mieux noter le texte 2 que le texte 1. Les scores prédits sont plus proches avec un léger avantage pour le texte 1. Avec un R2 significativement plus élevé, on observe également que les scores prédits sont plus consistants que les scores des juges humains.

Ainsi la méthode de prédiction des scores présentés ici peut l'être avec différents types de texte. Les textes peuvent induire quelques différences, notamment d'expressivité, ces différences étant cependant moins importantes qu'avec des juges humains.

Le processus décrit ci-dessus (alignement, extraction des paramètres verbaux et prosodiques, calculs des scores à chaque dimension à partir des modèles) a été ensuite automatisé pour évaluer directement les lectures à partir des données audio brutes.

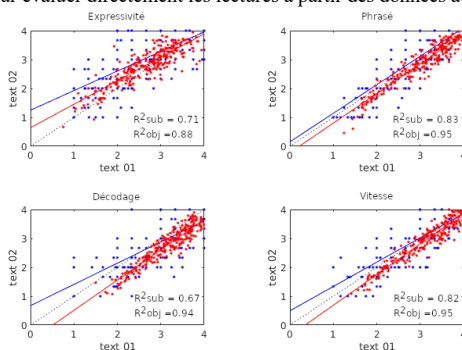


Fig. 3. Scores au texte 2 en fonction des scores au texte 1 pour les dimensions de phrasé et d'expressivité. Scores subjectifs (bleu) et scores calculés (rouge).

3 Application à une étude longitudinale

3.1 Analyses statistiques

Nous analysons ici les données du corpus longitudinal. Les analyses ont été réalisées avec le logiciel R. L'effet du niveau a été testé en utilisant des modèles linéaires multiniveaux de croissance, en utilisant le sujet comme paramètre aléatoire de niveau 2 et le parcours comme paramètre aléatoire de niveau 3, avec la fonction lmer du package lme4 (Bates et al., 2015). Les participants sont répartis sur plusieurs classes chaque année (quatre classes en CE1, cinq en CE2 et six en CM1). Cette répartition évolue chaque année. Les sujets suivent donc des parcours différents pendant ces trois années. Au total, les participants se répartissent sur 12 parcours du CE1 au CE2 (année 2) et sur 25 parcours du CE1 au CM1 (année 3). Pour chaque modèle, l'effet aléatoire de niveau 3 est significatif. Il est donc conservé pour la suite des analyses. Quand l'effet du niveau est significatif, une analyse post-hoc a été menée pour déterminer les différences entre les niveaux deux à deux. Nous avons utilisé le test HSD de Tukey. Pour cette analyse nous rapportons la p-value la plus élevée parmi les tests deux à deux. Nous avons également testé l'effet du texte et son interaction avec le niveau.

Le tableau 3 présente les scores obtenus pour chaque niveau et chaque texte les scores aux dimensions de l'EMDF. La 4 présente les modèles de croissance pour chacune des dimensions de la fluence pour chacun des textes évalués.

Tableau 3. Scores moyens (écart-type) prédits à l'EMDF pour le texte 1 par l'outil automatique pour chaque année du corpus longitudinal.

année	Texte 1			Texte2		
	1	2	3	1	2	3
niveau	CE1	CE2	CM1	CE1	CE2	CM1
Expressivité	2.16(0.53)	2.35(0.53)	2.57(0.52)	2.22(0.37)	2.55(0.42)	2.82(0.46)
Phrasé	2.35(0.49)	2.71(0.50)	2.99(0.50)	2.21(0.44)	2.71(0.43)	3.03(0.42)
Décodage	2.62(0.43)	2.92(0.39)	3.10(0.40)	2.19(0.33)	2.60(0.35)	2.91(0.37)
Vitesse	2.53(0.49)	2.98(0.49)	3.24(0.48)	2.32(0.41)	2.83(0.44)	3.14(0.40)

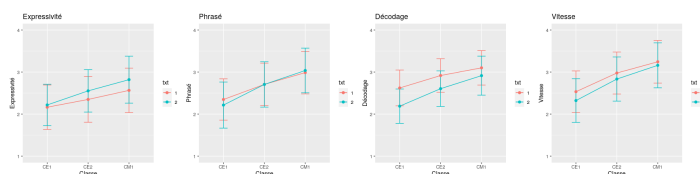


Fig. 4. Modèle de croissance du CE1 au CM1 pour chacune des quatre dimensions de la fluence pour un texte narratif (rouge) et un dialogue (bleu)

3.2 Expressivité

Le modèle de croissance montre un effet fixe positif du niveau sur le score à la dimension expressivité. L'ajout du niveau comme paramètre aléatoire de niveau 2 n'améliore pas notre modèle. Il n'y a donc pas de variance des pentes entre les élèves. On observe par contre un effet d'interaction entre niveau et texte. Le texte 2 a un taux de croissance légèrement plus élevé pour le texte 2, 0,08 points en plus par an par rapport au texte 1. L'analyse post-hoc de Tukey confirme la croissance progressive de l'expressivité avec des scores significativement différents pour chaque année ($p < .001$). Les scores à la dimension expressivité augmentent en moyenne de 0,25 points par an.

3.3 Phrasé

Le modèle de croissance montre un effet fixe positif du niveau sur le score à la dimension phrasé. L'ajout du niveau comme paramètre aléatoire de niveau 2 améliore notre modèle. On observe donc un effet du niveau et une variance des pentes entre élèves. Le rythme de progression est donc différent entre les élèves. L'interaction entre niveau et texte améliore également notre modèle. On constate à la fois un effet du texte sur les scores, significativement plus faibles pour le texte 2 (-0,23 points), et sur le taux de croissance légèrement plus élevé pour le texte 2 (0,07 points en plus par an). Comme on peut le constater sur la figure 4, la croissance des scores à la dimension phrasé semble assez linéaire du CE1 au CM1. L'analyse post-hoc de Tukey confirme la croissance progressive de l'expressivité avec des scores significativement différents pour chaque année ($p < .001$). Les scores à la dimension expressivité augmentent en moyenne de 0,37 points par an.

3.4 Décodage

Le modèle de croissance montre un effet fixe positif du niveau sur le score à la dimension décodage. L'ajout du niveau comme paramètre aléatoire de niveau 2 n'améliore pas notre modèle. On observe donc un effet du niveau mais pas de différence entre élèves dans les rythmes d'acquisition. L'ajout de l'interaction du niveau avec le texte améliore grandement notre modèle. Les scores du texte 2 sont significativement inférieurs à ceux du texte 1 (-0,57 points) et son taux de croissance significativement supérieur (0,10 points de plus par an). Comme on peut le constater sur la figure 4, la croissance des scores à la dimension décodage semble changer du CE2 au CM1. Les scores, compris entre 1,83 et 3,85, pour le texte 1, ne plafonnent pas en CM1. L'analyse post-hoc de Tukey confirme la croissance progressive du décodage avec des scores significativement différents pour chaque année ($p < .001$). Les scores à la dimension décodage augmentent en moyenne de 0,31 points par an.

3.5 Vitesse

Le modèle de croissance montre un effet fixe positif du niveau sur le score à la dimension vitesse. L'ajout du niveau comme paramètre aléatoire de niveau 2 améliore notre modèle. On observe donc un effet fixe du niveau et une variance des pentes au niveau inter-élèves. Tous les élèves ne progressent donc pas au même rythme. L'ajout de l'interaction entre niveau et texte améliore le modèle. Le score à la dimension vitesse est significativement plus faible pour le texte 2 (-0.30 points) et le taux de croissance légèrement plus élevé. La différence reste cependant faible (+ 0.04 points par an). Comme on peut le constater sur la figure 3, la croissance des scores à la dimension vitesse semble ralentir du CE2 au CM1. Les scores, compris entre 1.7 et 4 plafonnent en CM1. L'analyse post-hoc HSD de Tukey confirme la croissance progressive de la vitesse avec des scores significativement différents pour chaque année ($p < .001$). Les scores à la dimension vitesse augmentent en moyenne de 0.39 points par an.

4 Discussion

4.1 Modèle d'évaluation utilisé

Le modèle d'évaluation automatique proposé utilise des paramètres objectifs verbaux et prosodiques issus d'une comparaison avec de multiples références. Ce modèle permet de prédire les scores subjectifs à l'EMDF pour chacune des dimensions à partir de données acoustiques objectives. Les scores obtenus sont très corrélés à ceux donnés par des juges humains, tout en étant plus simple et rapide à mettre en œuvre. Ce modèle permet ainsi de s'abstraire de la variabilité inter-évaluateurs pouvant poser problème dans une analyse avec de multiples point de suivis. Cet outil est également moins dépendant du texte que l'évaluation humaine. Les scores entre deux textes différents sont plus consistants. De plus les juges humains sont soumis à plusieurs biais : les attendus selon l'âge supposés des lecteurs, l'interdépendance perceptive des dimensions de la fluence, la tendance à utiliser l'intégralité de l'échelle (y compris pour des élèves proches). Finalement, les modèles présentés s'appuyant sur des données objectives, donnent des scores continus. Ces scores sont donc beaucoup plus précis que ceux obtenus avec des juges humains, scores discrets. Cet outil permet donc de capter des variations plus fines de fluence chez les jeunes lecteurs.

Ainsi cet outil est un outil objectif, valide et précis tout à fait utilisable en recherche. Il est par ailleurs totalement adapté aux études longitudinales. Son objectivité permet de réaliser des évaluations multiples sans se soucier de la variabilité inter-évaluateurs. Les scores aux évaluations répétées sont directement comparables et les variations fines facilement détectables. Ainsi il pourrait être utilisé pour le suivi des progrès dans le cadre de la mise en place d'atelier spécifique ou, ans le cadre d'un entraînement, pour donner un feedback aux élèves et de visualiser leur progrès.

4.2 Développement de la prosodie

Les modèles de croissance des quatre dimensions de la fluence montrent des scores en augmentation du CE1 au CM1 dans toutes les dimensions. Cependant les profils sont différents d'une dimension à l'autre. En effet, les scores en CE1 sont plus élevés dans les dimensions d'automatisme (vitesse et décodage) que dans les dimensions prosodiques (expressivité et phrasé). Le rythme d'acquisition moyen dépend aussi de la dimension étudiée. Les scores de phrasé et de vitesse augmentent plus rapidement que ceux de décodage et d'expressivité. On peut supposer que les enfants de CE1 ayant déjà acquis un bon niveau de décodage, les scores plafonnent très rapidement. L'expressivité est une dimension qui s'acquiert plus tardivement et plus lentement.

Ces modèles nous montrent également des différences entre les enfants dans leurs rythmes d'acquisition des compétences de phrasé et de vitesse. On pourrait expliquer cette différence de rythme pour la vitesse par une forte hétérogénéité en CE1 dans cette compétence, avec des élèves plus faibles en CE1 qui « rattrapent » leurs camarades en CM1. Les compétences en expressivité sont moins hétérogènes en CE1 et progressent de façon similaire chez tous les enfants entre le CE1 et le CM1. Il serait intéressant de tester d'autres compétences pouvant expliquer ces différences (vocabulaire ou compréhension par exemple).

Finalement, on constate une différence entre les rythmes de progression sur les deux textes. Les scores progressent plus vite pour le dialogue, dans toutes les dimensions. Ces différences sont particulièrement fortes pour l'expressivité et le décodage. Il est probable que les enfants de CM1 soient plus à l'aise pour mettre en voix le dialogue et obtiennent donc des scores rapidement plus élevés que pour le texte narratif. Il est également possible que cet exercice soit peu familier aux enfants de CE1 et beaucoup plus en CM1. Il serait intéressant de prolonger cette étude longitudinale pour mieux cerner le profil d'acquisition spécifique à chaque dimension et particulièrement pour l'expressivité dont le développement se prolonge au-delà du CM1.

Ce type d'étude longitudinale donne des outils pour un meilleur diagnostic des difficultés rencontrées par les enfants. D'autre part, elle permet de mieux cibler les dimensions à entraîner à quel moment de l'apprentissage

5 Conclusion

Nous avons montré ici qu'en français, les dimensions prosodiques de la fluence se développent plus lentement et plus tardivement que les dimensions d'automatisme. Leur suivi est indispensable pour mieux comprendre l'acquisition des diverses compétences en lecture et repérer les élèves en difficulté. L'outil objectif proposé et utilisé ici permet des mesures répétées fiables, il est particulièrement adapté à ce type d'études longitudinales. L'intérêt pour l'enseignement de la lecture de cet outil réside dans sa simplicité et rapidité d'utilisation. Son utilisation simple en classe permet un suivi régulier et un feedback aux élèves. Son implémentation dans des applications d'entraînement de la fluence pourrait également permettre de faire le suivi de leur efficacité en classe.

Références

1. Meisinger, E., Bradley, B. Schwanenflugel, P., Kuhn, M. Morris, R. : Myth and reality of the word caller : The relation between nominations and prevalence among elementary school children. *School Psychology Quarterly* 24(3), 147(2009).
2. Rasinski, T. Assessing Reading Fluency. Pacific Resources for Education and Learning (PREL), Honolulu (2004)
3. Kuhn, M., Schwanenflugel, P.: Aligning theory and assessment of reading fluency : Automaticity, prosody and definitions of fluency. *Reading Research Quarterly* 45(2), 230-251 (2010).
4. Godde, E., Bosse, M.L., Bailly, G., : A Review of Reading Prosody Acquisition and Development. *Reading and Writing*, 33, 399-426 (2020)
5. Wolters A., Kim YS, Szura J.: Is reading Prosody Related to Reading Comprehension ? A Meta-analysis. *Scientific Studies of Reading* (2020).
6. Álvarez-Cañizo, M., Suárez-Coalla, P. et Cuetos, F. : Reading prosody development in spanish children. *Reading and Writing* 31(1), 35–52 (2018).
7. Miller, J. et Schwanenflugel, P. J. : A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading research quarterly* 43(4), 336–354(2008).
8. Godde, E., Bosse, M.L., Bailly, G., Echelle Multi-Dimensionnelle de Fluence : nouvel outil d'évaluation de la fluence en lecture prenant en compte la prosodie, étalonné du CE1 à la 5ème. *L'Année Psychologique*. (Sous presse)
9. Moser, G. P., Sudweeks, R. R., Morrison, T. G. et Wilcox, B. : Reliability of ratings of children's expressive reading. *Reading Psychology* 35(1), 58–79(2014).
10. Bolanos, D., Cole, R. A., Ward, W. H., Tindal, G. A., Schwanenflugel, P. J. et Kuhn, M. R. : Automatic assessment of expressive oral reading. *Speech Communication* 55(2), 221–236(2013).
11. Hirst, D., Rilliard, A. et Aubergé, V. Comparison of subjective evaluation and an objective evaluation metric for prosody in text-to-speech synthesis. In *The Third ESCA/COCOSDA Workshop (ETRW) on Speech Synthesis*. (1998)
12. Bailly G., Raidt S. et Elisei, F.: "Gaze, conversational agents and face-to-face communication", *Speech Communication - special issue on Speech and Face-to-Face Communication*, 52(3): 598–612(2010).

Rétroactions épistémiques dans une plateforme d'entraînement aux mathématiques : apports croisés de l'informatique et de la didactique

Sébastien Jolivet^{1,3}, Amel Yessad¹, Mathieu Muratet^{1,4}, Vanda Luengo¹, Benjamin Reiter¹, Brigitte Grugeon-Allys², Elann Lesnes-Cuisiniez³

¹ Sorbonne Université, LIP6

[sebastien.jolivet][amel.yessad] [mathieu.muratet]

[vanda.luengo]@lip6.fr

reiter.benjamin.rb@gmail.com

² Université Paris Est Créteil, LDAR

brigitte.grugeon-allys@u-pec.fr

³ Université de Paris, LDAR

elann.lesnes@gmail.com

⁴ INS HEA

Résumé. Cet article porte sur la conception de rétroactions épistémiques pour une plateforme d'entraînement aux mathématiques. Nous explicitons les fondements didactiques (incluant une modélisation du savoir à enseigner et des profils cognitifs prédéfinis d'apprenants), la modélisation informatique des rétroactions et un processus de décision de ces rétroactions, pour lequel nous détaillons les facteurs pris en compte. Nous montrons alors l'algorithme qui formalise ce processus. Enfin, nous discutons les apports et limites du travail réalisé et présentons la suite du travail engagé.

Mots-clés : rétroaction, rétroaction épistémique, modèle de rétroaction, algorithme de décision, didactique des mathématiques.

Abstract. We present the design and the modeling of epistemic feedbacks for a mathematics training platform. We explain the didactic basis (including the modeling of didactic knowledge and predefined learner's profiles), the computer modeling of feedbacks and a decision process for these feedbacks, a process for which we detail the factors taken into account. We then present the algorithm formalizing this process. Finally, we then discuss the contributions and limitations of the work carried out and present the continuation of the work undertaken.

Keywords: feedback, epistemic feedback, feedback modeling, decision algorithm, didactic of mathematics.

1 Introduction et état de l'art

Le projet MindMath¹ vise à la construction d'une plateforme d'entraînement aux mathématiques du collège (algèbre et géométrie). Il regroupe deux laboratoires de recherche (LDAR et LIP6) et trois entreprises (Tralalère, CabriLog et Domoscio). Une spécificité du projet est d'articuler les expertises, en didactique du LDAR, en IA du LIP6 et de Domoscio et éditoriale de Tralalère et CabriLog. Dans la plateforme l'élève réalise des exercices, organisés au sein de parcours d'entraînement adaptatifs, durant lesquels il bénéficie de rétroactions. La production des exercices, des parcours, le contenu et la décision des rétroactions sont fondés épistémologiquement et didactiquement. C'est sur la dimension rétroaction du projet que nous centrons cette contribution.

Les rétroactions sont largement identifiées comme un moyen important pour l'apprentissage et de nombreux travaux interrogent les conditions de leur efficacité. Ainsi, Wisniewski & al (2020) signalent, dans une revisite de l'article « *The power of Feedback* » (Hattie & Timperley, 2007), que les rétroactions ont un impact plus important sur l'apprentissage que sur la motivation et le comportement. Nous ne visons pas à agir sur ces deux derniers éléments dans le travail que nous réalisons. Nous considérons des rétroactions formatives telles que définies dans Shute « (...) [an] information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning » (Shute, 2008, p.154). Encore plus précisément, nous nous intéressons à la dimension épistémique des rétroactions, au sens proposé par Luengo (2009), c'est-à-dire qui porte sur le savoir enjeu d'apprentissage.

Il existe de multiples approches pour décrire le contenu et catégoriser les finalités des rétroactions. Leibold & Schwarz (2015) signalent quatre types de rétroactions selon la nature des informations qu'elles contiennent (rétroaction corrective, rétroaction épistémique, rétroaction suggestive, rétroaction épistémique et suggestive). Shute (2008, p.160), synthétisant de nombreux travaux, propose de multiples critères qui l'amène à une typologie des rétroactions ordonnées par complexité. Nous retenons les critères relatifs à l'objectif des rétroactions (*directive feedback* et *facilitative feedback*) et à la caractérisation du contenu en termes de niveaux d'information (*feedback specificity*).

Du point de vue de la modélisation informatique, tel qu'introduit précédemment, les connaissances didactiques (liées aux connaissances à enseigner et à la modélisation des connaissances de l'apprenant) et pédagogiques (temps, intention pédagogique, forme...) semblent nécessaires pour la conception des rétroactions épistémiques. Cependant les connaissances sur la rétroaction la plus pertinente à donner pour un apprenant donné à un moment donné restent un verrou. Une piste est donc de faire évoluer ce type de modèle centré experts vers un modèle capable d'apprendre des interactions avec les apprenants pour adapter ses rétroactions. Les techniques informatiques pour produire des rétroactions adaptées à partir des théories cognitives, des connaissances sur le domaine ou sur l'apprenant sont variées : production des règles, réseaux bayésiens, traitement automatique de la langue, etc. (Bimba et al., 2017). Cependant, l'évolution de ces systèmes peut être difficile ou fastidieuse à mettre en place (Luengo, 2009). Il est donc nécessaire de concevoir un modèle capable de produire des rétroactions à partir

¹ Projet lauréat du Fond Unique Interministériel 23. <http://mindmath.education/>

des connaissances expertes mais également capable d'évoluer, de façon informée, au fur et à mesure des interactions avec les apprenants. Une approche permettant de représenter le processus de décision de la rétroaction (Murray et al 2001), et le type de rétroaction, de façon explicite et découplée peut permettre ces deux types d'évolutions.

Nous visons, dans le projet, à proposer de manière automatique des rétroactions avec un contenu épistémique et adaptées à l'apprenant. Notre problématique est donc de concevoir une modélisation informatique et de déterminer un processus permettant au système de décider les rétroactions, avec prise en compte de l'apprenant, tout en fondant didactiquement l'ensemble de ces éléments. Pour y répondre nous avons :

- défini le contenu des rétroactions et leur modélisation en appui sur une modélisation didactique du savoir et de l'apprenant.
- identifié les facteurs de décision et produit l'algorithme permettant leur mise en œuvre.

Dans cette contribution nous présentons ces éléments et les fondements didactiques retenus pour les aborder. Enfin, nous discutons certaines limites et présentons les prolongements, engagés ou à venir, de ce travail.

2 Modèle de rétroactions

Dans cette section nous présentons les principes qui ont orienté notre travail, les fondements didactiques utilisés et enfin le modèle de rétroaction.

2.1 Principes du modèle et fondements didactiques

Pour construire le modèle et le contenu des rétroactions, nous avons fixé cinq principes qui nous permettent de prendre en considération différentes observations de Shute (2008) qui insiste sur la qualité du contenu et sur l'intérêt de l'adaptation des rétroactions au niveau de compétence de l'apprenant (« *For instance, instead of feedback complexity, a more salient facet of feedback may be the nature and quality of the content, such as providing information about learning goals and how to attain them* » (ibid. p. 161) « *For instance, a struggling student may require greater support and structure from a formative feedback message compared to a proficient student* » (ibid. p. 154)) :

1. Les rétroactions s'appuient sur le savoir en jeu et les moyens de sa mise en œuvre.
2. Les rétroactions sont mathématiquement correctes².
3. Certaines rétroactions prennent en compte des erreurs fréquentes.
4. Il existe des rétroactions adaptées à différents profils d'apprenants.

² Les notations sont rigoureuses et il n'y a pas dans les feedbacks, même si on les rencontre parfois dans les classes, des formulations approximatives (« segments égaux » à la place de « segments de même longueur », « angle » à la place de « mesure de l'angle », etc.) ou relevant de « recettes » (énoncé du type « si on change de côté alors on change de signe »). Ce principe est pris en charge essentiellement lors de la production effective des feedbacks.

5. Le modèle présente un niveau de généralité qui lui permet de décrire des rétroactions pour les différents domaines mathématiques présents dans la plateforme.

Pour mettre en œuvre ces principes nous exploitons les apports de la didactique au travers d'une modélisation des savoirs visés, des erreurs associées et de leur cause, et d'une définition du niveau de compétence de l'apprenant. La modélisation du savoir est construite à partir de l'approche praxéologique de la théorie anthropologique du didactique (Chevallard, 1992), de ces développements avec le cadre T4TEL (Chaa-choua, 2018) et des apports réalisés dans le cadre du projet MindMath (Jolivet et al., à paraître). Le savoir est structuré en domaines (algèbre, géométrie, etc.) dans lesquels sont travaillés des thèmes (les équations du 1^{er} degré, les triangles, etc.) au sein desquels sont définies des types de tâches à réaliser (résoudre une équation du 1^{er} degré, modéliser un problème à l'aide d'une équation du 1^{er} degré, etc.). Cette structuration est le fruit d'un travail épistémologique et didactique. Ce travail porte aussi sur l'identification des erreurs habituelles en lien avec le savoir modélisé (voir par exemple Grugeon-Allys et al. (2012) pour le cas de l'algèbre). Ces éléments permettent de prendre en compte les principes 1 et 3. De plus, comme cette modélisation du savoir ne dépend pas du domaine mathématique, cette approche nous permet de prendre en compte le cinquième principe ci-dessus. Nous présentons les éléments de cette modélisation utilisés pour les rétroactions.

- Les tâches à réaliser : elles sont regroupées dans des familles de tâches qui sont définies à partir d'un verbe d'action (par exemple résoudre), d'un objet mathématique sur lequel s'exerce l'action (par exemple une équation du premier degré) qui permettent de définir un grand ensemble de tâches (dans notre exemple : résoudre une équation du premier degré). On définit alors un ensemble de variables, et de valeurs associées, qui permettent de caractériser didactiquement des sous-ensembles de tâches (par exemple, la forme de l'équation, le nombre de solutions, la nature des coefficients, la nature de la solution, etc.). Les variables et leurs valeurs sont choisies pour pouvoir : 1) définir le niveau scolaire auquel la tâche est adaptée ; 2) caractériser les procédures qui permettent de réaliser, ou non, les tâches ; 3) déterminer la complexité des tâches. L'instanciation de ces différentes variables définit une famille de tâches, c'est-à-dire un ensemble de tâches qui peuvent toutes être résolues par au moins une procédure commune et qui sont de complexité identique (au regard des valeurs choisies pour les variables de complexité).
- Les procédures : il s'agit des moyens existants pour réaliser les tâches d'une, ou plusieurs, famille(s) de tâches. Par exemple : regrouper les termes contenant l'inconnue dans le même membre ; regrouper les termes constants dans l'autre membre ; calculer la valeur de l'inconnue.
- Les justifications : il s'agit des propriétés, définitions, etc., qui permettent de justifier la validité mathématique des procédures de résolution. Par exemple : « une égalité est préservée quand on réalise la même opération dans les deux membres » ; « le produit d'un nombre et de son inverse est égal à un ».
- Les erreurs classiques : il s'agit des erreurs susceptibles de se manifester lors de la réalisation des tâches. Elles sont identifiées à partir des travaux de la didactique.

En ce qui concerne la modélisation de l'apprenant on définit le niveau de compétence à l'aide du mode d'activité (Grugeon-Allys, 2016). Il permet de caractériser le rapport de l'apprenant, construit au fil de ses apprentissages, à différents domaines des mathématiques sur plusieurs aspects de l'activité mathématique (calculer, modéliser, représenter, construire, raisonner). Dans le cadre du projet MindMath, ce mode est initialisé lors de la réalisation d'un ensemble de tâches au début du travail sur une notion, il est régulièrement actualisé en fonction de l'activité de l'apprenant. Il est exploité lors de la décision des rétroactions permettant ainsi de prendre en compte le quatrième principe présenté au début de la section. Nous disposons maintenant des principes et des fondements didactiques qui nous permettent de présenter notre modèle de rétroactions.

2.2 Patron de rétroactions

Nous distinguons deux types de rétroactions selon leur finalité :

- Amener l'apprenant à interroger la validité de sa réponse de manière à lui permettre d'établir qu'elle est erronée et de s'engager dans une reprise de sa production.
- Fournir des éléments qui lui permette de retravailler la tâche. Ceux-ci peuvent davantage porter sur des éléments de justification ou sur des éléments de procédure. Ils peuvent être illustrés sur une tâche proche où sur la tâche à traiter elle-même.

Le premier type est à rapprocher de la question de la *vérification* proposée par Shute (2008), qui signale par ailleurs que cela peut être fait de manière explicite ou implicite. L'aspect explicite de la vérification est pris en charge dans la plateforme par des coches rouges et vertes qui indiquent la validité ou non de la réponse. Notre proposition est plus implicite car il s'agit d'indiquer à l'apprenant une propriété, une définition, etc. et l'inviter à vérifier sa bonne mise en œuvre, charge à lui d'en déduire l'aspect erroné de sa réponse et la nécessité d'un retravail sans qu'une orientation particulière soit proposée. Pour le deuxième type, l'élève est informé que sa réponse est erronée et il s'agit de lui fournir des éléments permettant de retravailler la tâche. Il contient à la fois des rétroactions plutôt directives (*directive feedback*) et d'autres plutôt facilitatrices (*facilitative feedback*), au sens proposé par Shute (2008), selon que l'accent est mis sur la procédure ou les justifications. Ces contenus peuvent être illustrés sur un exemple proche de la tâche à réaliser ou sur la tâche à réaliser elle-même (par exemple, illustrer avec une équation qui présente une complexité identique sans avoir les mêmes coefficients ou utiliser la même équation que celle à résoudre par l'apprenant). Pour des raisons liées à la plateforme nous avons concentré notre travail sur le 2^e type de rétroactions. Afin de les définir nous avons identifié les paramètres et valeurs possibles nous permettant de définir le patron d'une rétroaction à l'aide du vecteur suivant :

$$FB = \langle J | P ; G | IFt | It ; C | NC ; \{Er\}_i | null \rangle.$$

- Première composante : la rétroaction porte sur justification (J) ou une procédure (P).
- Deuxième composante : la rétroaction est rédigée de manière générique (G), instanciée à la famille de tâche (IFt) ou instanciée à la tâche (It).

- Troisième composante : quand la rétroaction porte sur une procédure, indique si elle est commentée (C) ou non commentée (NC).
- Quatrième composante : prise en compte d'une erreur usuelle (Er_i) ou non (null).

Nous illustrons ce modèle sur deux exemples présentés dans la **Fig. 1**

Fig. 1. Exemples de rétroactions. Type < J ; G ; C ; null > et < P ; I ; C ; Er_oublCofacteur1 >

La rétroaction de gauche est du type < J ; G ; C ; null > : elle présente la formule à utiliser de manière générique (les coefficients sont des lettres). Des commentaires sont proposés pour guider l'utilisation de la formule (cadres grisés) et pour rappeler différents éléments de vocabulaire. Il ne se réfère pas à une erreur particulière. Celle de droite est du type < P ; I ; C ; Er_oublCofacteur1 > : elle présente une procédure de résolution instanciée, de manière commentée et évoque une erreur usuelle qui est l'oubli du cofacteur 1. Un commentaire signale une manière de détecter une telle erreur.

Les différentes combinaisons possibles des valeurs des composantes du vecteur permettent de définir de nombreuses rétroactions permettant d'explorer le champ des possibles. D'une part on a des rétroactions avec une information essentiellement centrée sur les éléments de justifications (*facilitative feedback*), ce qui est un objectif d'apprentissage fondamental pour avoir une activité mathématique réfléchie et contrôlée. D'autre part, on a des rétroactions avec une information spécifique et procédurale qui permettent, elles, d'améliorer la réussite immédiate de la tâche (*directive feedback*). Il nous semble difficile d'associer une rétroaction avec un unique effet visé, cependant, nous avons formulé, à partir de l'expertise didactique, certaines hypothèses sur :

- l'effet le plus probable, apprentissage ou réussite, des différentes rétroactions,
- une relation de pertinence entre les différents rétroactions et le profil d'apprenant.

Pour cela, nous distinguons deux critères de pertinence : permettre la réussite de l'apprenant sur la tâche en cours et permettre l'apprentissage par l'apprenant des connaissances mobilisées par la tâche. Ces hypothèses sont synthétisées dans la **Fig. 2**. Les lignes du tableau correspondent à des contenus portant sur les procédures ou les justifications, ses colonnes précisent si la rétroaction est générique ou instanciée et commentée ou non commentée tandis que les cases indiquent la prise en compte ou non d'une erreur usuelle dans la rétroaction (Er_i pour la prise en compte d'une erreur ; null

si non utilisation d'une erreur ; X indique que ce type de rétroactions n'existe pas³). L'hypothèse sur l'effet de la rétroaction est que, plus on se déplace vers la droite, plus il est probable de favoriser la réussite de la tâche. Plus on descend plus l'effet probable est de favoriser l'apprentissage. Enfin, le lien avec les profils d'apprenant est explicité par les couleurs selon la légende indiquée à droite du tableau.

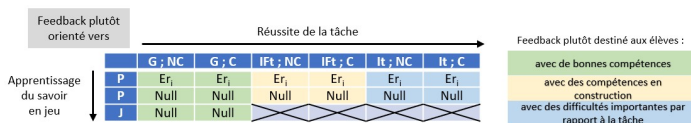


Fig. 2. Hypothèses sur la fonction et la cible des différents types de rétroactions

Par exemple nous faisons l'hypothèse qu'une rétroaction du type < P ; It ; C ; null >, qui présente une procédure de résolution commentée (au moins partielle) de la tâche exacte à réaliser par l'apprenant, doit favoriser grandement sa réussite immédiate, mais sans pour autant favoriser la prise de recul nécessaire à l'apprentissage. Une telle rétroaction est donc plutôt destinée à un élève en difficulté. *A contrario*, une rétroaction du type < J ; G ; NC ; null > qui propose un rappel de cours sans commentaire, ne sera sans doute pas très efficace pour un apprenant avec de faibles compétences.

3 Décision des rétroactions

Maintenant que nous avons défini les rétroactions pouvant être proposés nous présentons les facteurs utilisés par le système pour décider les rétroactions, puis l'arbre de décision. Le peuplement de ses feuilles est appuyé sur les hypothèses synthétisées dans la Fig. 2. Pour éviter d'éventuelles ambiguïté nous précisons deux éléments : 1) dans la plateforme les exercices sont constitués d'une unique tâche, il n'y a donc pas à décider relativement à quelle tâche de l'exercice la rétroaction est la plus opportune ; 2) après une rétroaction un apprenant peut revenir sur la tâche et tenter une nouvelle résolution et bénéficier éventuellement d'une autre rétroaction en cas de nouvelle réponse erronée. Nous nous assurons qu'il n'a pas deux fois la même rétroaction lors de la résolution d'un exercice.

3.1 Critères pris en compte pour la décision

Pour décider les rétroactions, deux catégories de critères sont définies : 1) la première relative à la tâche et aux erreurs liées, 2) la seconde relative aux apprenants. Les critères sont synthétisés dans le **Tableau 1**. On distingue l'activité immédiate de l'apprenant qui réalise la tâche et l'historique des activités de l'ensemble des apprenants dans la

³ Il s'agit en effet des feedbacks portant sur une propriété, un théorème, une définition. Les instancier à une famille de tâches ou une tâche revient à mettre en œuvre la propriété, et donc à présenter la procédure. On n'est donc plus dans les feedbacks portant sur la justification.

plateforme. Cet historique permet l'actualisation du mode d'activité de l'apprenant et la détermination du rapport de l'ensemble des apprenants à la tâche en cours.

Tableau 1. Facteurs de décision de la rétroaction

Sujet	Attribut	Valeurs	Contexte
Apprenant	Modalité de déclenchement de la rétroaction	- Validation de la réponse - Demande d'aide - Inactivité longue	Activité immédiate de l'apprenant
Apprenant	Réponse à la tâche	- Correcte - Incorrecte	Activité immédiate de l'apprenant
Apprenant	Erreur identifiée	- Oui - Non	Activité immédiate de l'apprenant
Apprenant	Nombre de tentatives sur la tâche	- Première tentative - Plusieurs tentatives	Activité immédiate de l'apprenant
Apprenant	Fréquence de l'erreur identifiée chez l'apprenant	- Très fréquente - Fréquente - Peu fréquente - Non fréquente	Historique de l'activité dans la plateforme
Apprenant	Niveau de compétence de l'apprenant dans le domaine de la tâche	- 0 (faible) - 1 (moyen) - 2 (bon)	Historique de l'activité dans la plateforme
Ensemble des apprenants	Fréquence des erreurs possible sur la tâche pour l'ensemble des apprenants	- Erreur la plus fréquente - 2e erreur - ...	Historique de l'activité dans la plateforme

3.2 Algorithme de décision

Le processus de décision des rétroactions s'appuie sur un arbre de décision (**Fig. 3**) dont les nœuds sont les attributs décrits dans le **Tableau 1**, et considérés par les experts dans un ordre bien précis. Chaque branche de cet arbre de décision est la conjonction de valeurs d'un sous-ensemble de ces attributs. À une feuille (ou un nœud terminal) de l'arbre de décision sont associés une ou plusieurs rétroactions pondérées. La pondération d'une rétroaction dans une feuille représente, du point de vue de l'expert, la pertinence de la rétroaction pour l'apprenant et le contexte. Les rétroactions associées à une même feuille de l'arbre de décision peuvent être pondérées différemment. La **Fig. 3** présente un extrait de l'arbre de décision utilisé pour décider les rétroactions. Le peuplement des feuilles est réalisé avec les différents rétroactions produites et de manière à être cohérents avec les hypothèses présentées dans la **Fig. 2**. Ainsi à chaque branche de l'arbre sont associées un certain nombre de rétroactions, chacune de ces rétroactions étant pondérée pour rendre compte de l'avis expert sur sa pertinence dans la situation. L'algorithme de décision consiste donc à utiliser cet arbre de décision avec les pondérations définies par les experts et justifiées par nos hypothèses didactiques.

Dans le **Tableau 2** nous présentons le peuplement réalisé et les pondérations définies pour les feuilles 5 et 11 qui correspondent respectivement aux cas :

- réponse erronée, sans erreur type identifiée, apprenant avec de bonnes compétences

- réponse erronée, associée à une erreur très fréquente chez l'apprenant, niveau faible

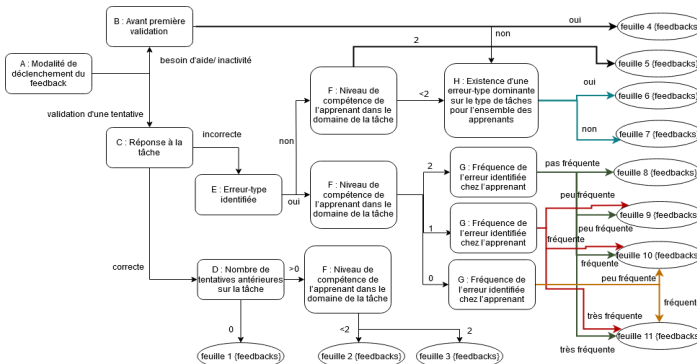


Fig. 3. Extrait de l'arbre de décision des rétroactions

Tableau 2. Peuplement, avec pondération, des feuilles 5 et 11 de l'arbre

Feuille	< Rétroaction > ; pondération			
5	<J ; G ; NC ; Null> ; 0,3	<J ; G ; C ; Null> ; 0,3	<P ; G ; NC ; Er ⁴ > ; 0,2	<P ; G ; NC ; null> ; 0,2
11	<P ; IFt ; C ; Er> ; 0,3	<P ; IFt ; C ; Null> ; 0,3	<P ; It ; C ; Er ⁵ > ; 0,2	<P ; It ; C ; null> ; 0,2

4 Discussion et perspectives

4.1 Apports et limites

Dans cette contribution nous avons proposé un patron de rétroaction fondé épistémologiquement et didactiquement en termes de contenus. Pour permettre au système de les décider, dans une plateforme d'entraînement aux mathématiques, nous prenons en compte le savoir en jeu, l'activité et le profil de l'apprenant ainsi que l'histoire de l'activité, pour l'ensemble des apprenants, dans la plateforme. Le patron de rétroaction et l'arbre de décision ne sont pas spécifiques à un domaine des mathématiques et peuvent être utilisés dès lors qu'une modélisation du savoir, telle que présentée dans la section 2.1⁶, est disponible. A l'heure où nous écrivons cet article 142 rétroactions sont déjà produites pour cinq thèmes du projet (4 en algèbre et 1 en géométrie). La réalisation de

⁴ erreur la plus fréquemment réalisée sur la famille de tâche par l'ensemble des apprenants

⁵ erreur identifiée dans la réponse de l'apprenant

⁶ De telles modélisations sont déjà disponibles dans de nombreux domaines qui concernent l'école élémentaire et le collège (Jolivet, 2018; Pilet, 2015; Vu & Tchounikine, 2020).

ce projet avec des partenaires industriels a permis cette production importante mais a aussi nécessité la prise en compte de certaines contraintes éditoriales, techniques et économiques. En particulier sur la forme des rétroactions qui contiennent beaucoup d'informations, comme le montre les exemples proposés sur la **Fig. 1**, et sont, de plus, proposées d'un seul bloc.

La suite est la réalisation d'expérimentations et leur analyse pour évaluer les hypothèses qui ont fondé notre travail et, en particulier, l'effet des rétroactions sur l'apprenant, selon les deux dimensions réussites et apprentissage proposées dans cette contribution. La préparation des expérimentations est l'occasion de pointer certaines limites du travail réalisé pour anticiper leur prise en compte. L'approche choisie pour notre travail est une approche experte ; cependant, pour tester certaines hypothèses que nous avons faites sur des limites de notre travail, nous avons réalisé une enquête par questionnaire auprès de 17 enseignants de profils divers (expérience, niveaux d'enseignement, exercice en éducation prioritaire ou non). Elle a été réalisée en utilisant des maquettes de rétroactions et non pas lors d'une utilisation effective de la plateforme. Nous ne la détaillons pas⁷ par manque de place, mais nous retenons trois éléments importants.

Le premier enseignement de l'enquête confirme la remarque faite plus haut sur la forme des rétroactions. Un enseignant signale ainsi « Il y a beaucoup de texte, cela peut bloquer certains élèves en très grande difficulté », et un autre « Globalement j'ai trouvé que les propositions étaient assez inaccessibles pour des élèves en REP car elles sont trop formelles et ont trop de texte ». Ces commentaires vont dans le sens de Shute (2008) qui signale « (...) *if feedback is too long or too complicated, many learners will simply not pay attention to it, rendering it useless.* ». Lors de l'expérimentation nous devons donc, avant d'analyser l'effet des rétroactions, évaluer leur prise en considération. L'analyse des traces, en particulier le temps de consultation des rétroactions, sera un indicateur de cette considération ou non.

Le deuxième enseignement de l'enquête est que, sur la base des données recueillies, on ne peut pas mettre en évidence que les hypothèses issues de l'expertise didactique sont largement reprises dans les conceptions des enseignants qui ont répondu. Il nous semble donc important d'avoir à l'esprit, lors de l'analyse des résultats de l'expérimentation, que tous les élèves n'ont pas, antérieurement à l'utilisation de la plateforme, construit le même rapport aux rétroactions et que ceci peut impacter leur effet dans la plateforme. L'enquête suggère aussi que l'on ne peut pas associer le profil de l'enseignant à un rapport particulier aux rétroactions, ainsi le contrôle des lieux d'expérimentation (établissement plus ou moins favorisé, rural ou urbain, etc.), ne peut, *a priori*, suffire à contrôler « l'effet enseignant ». Nous prévoyons donc, en cas de disparités importantes d'une classe à l'autre d'un même établissement, d'interroger le rapport de l'enseignant aux rétroactions et ses pratiques usuelles à ce sujet. Pour cela nous envisageons de croiser ses réponses à un questionnaire, des productions (copies corrigées d'élèves, corrections produites en classe) et les résultats d'observation in-situ.

Enfin, certains retours d'enseignants, renvoient à des éléments que nous avons envisagés mais non mis en œuvre au regard des possibilités actuelles de la plateforme. Par

⁷ Le questionnaire utilisé est disponible à ce https://link.infini.fr/enquete_fb

exemple l'un d'eux demande « pourquoi ne pas proposer à l'élève de contrôler sa réponse comme premier feedback », ce qui renvoie au premier type de rétroaction que nous avons évoqué dans la section 2.2 mais qu'il n'a pas été possible d'implémenter pour le moment. D'autres suggèrent de mieux associer rétroaction et activité de l'élève lors de la résolution de la tâche. Ceci nécessite une finesse du diagnostic de l'activité de l'élève qui est pour le moment limitée dans la plateforme. Pour l'expérimentation cela va nous amener à apporter une attention particulière à la corrélation entre l'erreur remontée et la réalité de la production de l'élève. En effet, si l'erreur remontée n'est pas celle effectivement réalisés par l'élève on peut penser qu'une rétroaction qui prend en compte cette erreur ne sera pas nécessairement très efficace.

Ces différentes limites, en plus de leur prise en compte dans notre protocole expérimental, nous amènent à deux propositions d'améliorations à court terme. Première, pour prendre en compte le décalage qu'il peut y avoir entre le contenu des rétroactions et certains usages dans les classes de mathématique (notamment en raison de l'application du principe de rigueur mathématique), nous avons déjà augmenté les rétroactions d'un glossaire qui permet une reformulation en des termes plus habituels, pour certains élèves, de certains mots utilisés dans les rétroactions. Deuxièmement, pour remédier aux limites liées à la forme des rétroactions, diverses pistes sont explorées, notamment des améliorations de la médiatisation avec, en particulier, un affichage dynamique et temporisé des différents éléments de la rétroaction et l'ajout de commentaires audio qui permettraient d'apporter des éléments de métadiscours difficiles à transcrire mais importants pour l'explicitation des contenus de certaines rétroactions.

Au-delà de ces améliorations, en court ou prochaines, nous avons aussi identifié de nouveaux défis que nous abordons dans la section suivante.

4.2 Vers un algorithme d'apprentissage par renforcement

Nous avons rappelé dans la section précédente que les décisions s'appuient d'une part sur un diagnostic et d'autre part sur une expertise didactique. Nous avons aussi pu constater que certaines hypothèses utilisées pour construire l'algorithme de décision ne reflètent pas nécessairement les pratiques des enseignants et que cela peut donc influencer sur la manière dont les rétroactions sont reçues par les apprenants, et, *in fine*, sur leur efficacité. Enfin, si l'expertise didactique est importante pour la description du savoir, pour l'identification des erreurs et pour l'évaluation des compétences de l'apprenant ; il n'en est pas de même concernant les effets des rétroactions sur l'apprenant. Il apparaît donc pertinent, pour pallier ces limites, d'augmenter le dispositif par les apports de l'IA. Pour cela nous développons actuellement un algorithme d'apprentissage par renforcement (Kaelbling et al., 1996) pour la décision des rétroactions. L'approche experte déjà réalisée va permettre de répondre au problème de démarrage à froid des algorithmes d'IA. Un premier enjeu, à la confluence entre didactique et IA, est de définir les récompenses permettant l'apprentissage de l'algorithme et donc une évaluation des effets des rétroactions sur l'apprenant. Dans un deuxième temps, une fois l'algorithme d'apprentissage par renforcement déployé et expérimenté avec les élèves, le défi suivant va être de créer les synergies entre didactique et informatique pour :

- Comparer les résultats entre l'utilisation uniquement basée sur l'expertise et l'utilisation en mode apprentissage. Une fois les éventuelles différences identifiées il faudra se donner les moyens théoriques, du point de vue didactique, pour comprendre les causes des différences et éventuellement explorer de nouvelles approches.
- Analyser avec un point de vue didactique les résultats issus de l'algorithme d'apprentissage et exploiter les résultats de ces analyses pour l'améliorer.

Références

1. Bimba, A. T., Idris, N., Al-Hunaiyyan, A., Mahmud, R. B., & Shuib, N. L. B. M. (2017). Adaptive feedback in computer-based learning environments : A review. *Adaptive Behavior*, 25(5), 217-234. <https://doi.org/10.1177/1059712317727590>
2. Chaachoua, H. (2018). T4TEL, un cadre de référence didactique pour la conception des EIAH. In J. Pilet & C. Vendaïra (Éds.), *Actes du séminaire de didactique des mathématiques 2018* (p. 8-25). IREM de Paris - Université Paris Diderot; hal-02421410.
3. Chevallard, Y. (1992). Concepts fondamentaux de la didactique : Perspectives apportées par une approche anthropologique. *Recherche en didactique des mathématiques*, 12(1), 83-121.
4. Grugeon-Allys, B. (2016). 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.
5. Grugeon-Allys, B., Pilet, J., Chenevotot-Quentin, F., & Delozanne, E. (2012). 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.
6. Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
7. Jolivet, S. (2018). *Modèle de description didactique de ressources d'apprentissage en mathématiques, pour l'indexation et des services EIAH* [Thèse de doctorat, Université Grenoble Alpes]. <https://tel.archives-ouvertes.fr/tel-02079412/document>
8. Jolivet, S., Lesnes-Cuisiniez, E., & Grugeon-Allys, B. (à paraître). Conception d'une plateforme d'apprentissage en ligne en algèbre et en géométrie : Prise en comptes et apports de modèles didactiques. *Annales de didactique et de sciences cognitives*.
9. Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning : A survey. *Journal of artificial intelligence research*, 4, 237-285.
10. Leibold, N., & Schwarz, L. M. (2015). The art of giving online feedback. *Journal of Effective Teaching*, 15(1), 34-46.
11. Luengo, V. (2009). *Les rétroactions épistémiques dans les Environnements Informatiques pour l'Apprentissage Humain* [HDR, Université Joseph Fourier - Grenoble 1]. https://hal.archives-ouvertes.fr/docs/00/69/98/02/PDF/Luengo_2009.pdf
12. Pilet, J. (2015). Réguler l'enseignement en algèbre élémentaire par des parcours d'enseignement différencié. *Recherches en Didactique des Mathématiques*, 35(3), 271-312.
13. Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.
14. Vu, T. M. H., & Tchounikine, P. (2020). Supporting teacher scripting with an ontological model of task-technique content knowledge. *Computers & Education*, 163, 104098.
15. Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited : A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10, 3087. <https://doi.org/10.3389/fpsyg.2019.03087>

Le cursus académique personnalisé dans une approche par compétences avec érosion : étude d'un nouveau problème fondamental

Alexis Lebis, Jérémie Humeau, Anthony Fleury et Mathieu Vermeulen

IMT Lille Douai, Institut Mines-Télécom, Univ. Lille, Centre for Digital Systems
F-59000 Lille, France
`alexis.lebis@imt-lille-douai.fr`

Résumé. De plus en plus, les universités et les écoles d'ingénieurs tendent à adopter une approche par compétences et à proposer aux étudiants la personnalisation de leur cursus pour correspondre à leurs attentes de carrière, au détriment de cursus prédéfinis. Une telle personnalisation de cursus soulève d'importantes problématiques, comme une répartition convenable des compétences en son sein, l'assurance que les étudiants conservent bien leurs acquis, l'ordonnancement des cours entre eux ou encore gérer la non validation d'un cours dans le cursus. Ce papier pose les fondements théoriques de ce nouveau problème qu'est la personnalisation d'un cursus académique avec érosion des compétences, en propose les toutes premières modélisation et tentative de résolution.

Mots-clés : Modélisation, Cursus, Personnalisation, Compétences

Abstract. Nowadays in education, more and more universities and engineering schools tend to remove predefined specialization curriculum in order to promote student ability to construct their own personalized training, matching their career goal. This leads to several major problems such as knowledge dissemination, courses ordering to satisfy knowledge prerequisite, courses congestion and recommendation, or even handling student's non-validation of courses in the curriculum. In this paper, we introduce an ecological definition of the personalized university curriculum organization problem.

Keywords: Modelisation, Curriculum, Personnalisation, Competencies

1 Introduction et contexte

Les structures académiques modifient leurs pratiques d'enseignement afin d'exploiter les nouvelles technologies et les avancées dans le domaine des EIAH, et plus généralement pour répondre aux besoins de la société. Ainsi, il est primordial

d'étudier ces changements pour mieux les comprendre et assister correctement les acteurs pédagogiques (*e.g.* étudiants, enseignants, institutionnels).

Une transformation majeure de ces institutions est le remplacement progressif de cursus pré-définis en faveur d'un modulaires [13]. Dans ce paradigme, un étudiant doit choisir une certaine quantité de cours chaque semestre pour construire son propre parcours, lui donnant la liberté d'être en accord avec ses attentes (*e.g.* objectifs professionnels). Cela introduit de nombreuses problématiques pour tous les acteurs pédagogiques. En effet, l'accroissement du volume des étudiants, combiné aux nouvelles structurations des enseignements [4], requiert d'apporter des outils automatisés afin d'aider les acteurs académiques à proposer un support personnalisé et adapté aux étudiants [5].

Du point de vue des étudiants, la personnalisation du cursus nécessite une capacité de planification importante sur plusieurs années ; et une phase de formation en amont qui l'est tout autant afin de s'assurer que les cours choisis correspondent à leurs attentes et leur apportent les compétences nécessaires. Lors d'un échec ou d'un nouvel objectif de carrière, ils doivent également s'assurer que leur objectif de formation reste atteignable, que les cours choisis sont adéquats, et que les compétences qu'ils ont su acquérir le leur permettent. De plus, cette personnalisation subit une explosion combinatoire plus le nombre de cours à prendre et le catalogue de cours sont volumineux.

Du point de vue des enseignants, une telle variété de cursus existants rend certaines pratiques difficilement exploitables, comme les rattrapages. Cela conduit aussi à des recommandations extrêmement personnalisées, nécessitant de fait que les compétences acquises par les étudiants soient correctement analysées et évaluées [9]. De plus, outre que la notion de maîtrise d'une compétence soit difficile à estimer, cette maîtrise est susceptible de décroître dans le temps lorsque la compétence n'est pas sollicitée [6] : c'est cela que l'érosion. Cette érosion rend alors la tâche de remédiation encore plus complexe puisque chaque cursus, autrement dit l'enchaînement des cours, est par essence différent.

Du point de vue institutionnel, cette érosion de la compétence dans le temps vient renforcer la difficulté de valider le cursus personnalisé d'un étudiant. En effet, un *pensum* de l'institution est de prémunir les étudiants d'un éventuel cursus mal-formé : l'enchaînement des cours doit être possible, notamment en terme d'apports de ces derniers et de leur pré-requis. Or, dans des catalogues de cours importants, cette tâche devient manuellement peu, voire non, réalisable. Aussi, dans ce paradigme, les problèmes d'emploi du temps et de logistique (*e.g.* salles, enseignants) sont naturellement renforcés. De plus, il est du devoir de l'institution de garantir l'égalité entre tous les étudiants, notamment en certifiant que tous les objectifs de formations professionnels proposés au sein de la structure sont atteignables et que les compétences requises sont enseignées dans le parcours de l'étudiant. La conséquence de cela est la nécessité, avec l'aide du corps enseignant, de correctement cataloguer les compétences apportées par chacun des cours et les prérequis de ces derniers. Ce catalogage est également complexifié par le fait que certaines compétences sont transverses par nature [8].

De tout cela, il apparaît clairement que permettre à l'étudiant de personnaliser entièrement son cursus afin d'atteindre ses objectifs professionnels, tout en tenant compte de son profil (*e.g.* compétences initiales, propension de démobilitation des compétences), est certes extrêmement complexe, mais crucial à comprendre et à résoudre. Ce papier, à notre connaissance, fait la toute première contribution à ce problème en éducation. Nous apportons les fondamentaux théoriques quant à sa modélisation et sa compréhension, et nous investiguons également l'effet de certaines propriétés éducationnelles dans sa résolution et les recommandations qui en découlent.

Toutefois, bien que ce problème du cursus personnalisé avec érosion des compétences n'ait pas été traité dans la littérature, il existe certains travaux allant dans ce sens. Ebbinghaus, et des travaux subséquents [6] arrivent à quantifier la manifestation de l'oubli des acquis des étudiants, qui survient notamment lorsque les étudiants ne les mobilisent pas. Récemment, l'on observe des travaux proposant d'exploiter des techniques d'apprentissages semi-supervisés ou non supervisés sur des corpus de données d'étudiants et de formations, afin de produire des recommandations de cursus – notamment sous la forme de graphes probabilistes [2,14]. Bien que de tels travaux ne résolvent pas notre problème, ils mettent en valeur l'existence de propriétés exploitables et discriminantes pour l'élaboration de cursus personnalisés. Aussi, le travail réalisé par Molontay, *et al.* [11] mettent en évidence l'importance et l'effet des prérequis dans un cursus, les caractérisent et proposent différentes métriques pour quantifier leurs impacts.

Dans la section 2 suivante, nous formalisons le problème du cursus personnalisé avec érosion des compétences comme un problème de satisfaction de contraintes. Dans la section 3, nous proposons une application de ce modèle pour trouver une solution valide à ce problème. La section 4 présente l'implémentation de cette proposition et les résultats expérimentaux. À partir de ces résultats, nous suggérons dans la section 5 des recommandations prescriptives et des pistes de réflexions sur la mise en place de tel cursus.

2 Formalisation

Afin de correctement définir le problème du cursus personnalisé, nous faisons la double hypothèse que les aspects logistiques (*e.g.* disponibilité des salles) sont indépendants et n'influencent pas la qualité d'un cursus personnalisé, et que le catalogue de cours est correctement formé. Un cursus personnalisé \mathcal{P} est donc défini comme un ensemble temporellement ordonné de cours C qui aboutit aux objectifs \mathcal{O} des étudiants (*e.g.* objectif professionnel). Les cours le constituant sont prélevés au sein d'un catalogue de cours¹ et ne peuvent être repris par la suite. Ces cours ont des prérequis \mathcal{K}^- afin de les suivre correctement, et apportent aux étudiants des compétences \mathcal{K} . Ces mêmes compétences font face à une démobilitation δ dans le temps lorsque apprises par les étudiants.

1. Même s'il s'agit d'une tâche extrêmement difficile, il est du ressort de l'institution et des enseignants d'établir les catalogues à la fois de cours, et de compétences.

2.1 Cours, apport, prérequis et ordonnancement

Un cours c peut être considéré comme le principal vecteur de diffusion d'un ensemble de compétences \mathcal{K}_c . Chaque compétence k enseignée dans un cours c ($k \in \mathcal{K}_c$) est associée à une valeur de maîtrise m_k . Chaque compétence k est acquise par l'étudiant à un degré m_k lorsqu'il réussit le cours c (cf. subsection 2.3).

De plus, un cours peut également nécessiter des prérequis pour pouvoir être suivi. Ces prérequis forment l'ensemble \mathcal{K}_c^- constitué des compétences $k \in \mathcal{K}$ dont la maîtrise attendue m_k est supérieure au seuil $\lambda_{(c,k)}$:

$$\mathcal{K}_c^- = \begin{cases} \emptyset \\ \{k \mid k \in \mathcal{K}, \mu(m_k) \geq \lambda_{(c,k)}\} \end{cases} \quad (1)$$

où $\mu(x)$ est une fonction de *fuzzification* (cf. subsection 2.2) pour tenir compte de la difficulté d'évaluation de ces compétences.

À cela se rajoute les contraintes temporelles dans la modélisation d'un cours, afin de savoir à quelles périodes \mathcal{T}_c il peut être choisi (e.g. semestre 2 et 4). Nous définissons la disponibilité d'un cours c comme $\mathcal{T}_c = \{t_i \mid t_i \in T, i \in [0, n], 1 \leq n \leq |T|\}$, où T représente l'ensemble ordonné² de toute les périodes t du cursus et où $|T|$ vaut la durée maximale du cursus (e.g. 5 ans). Une année académique est déterminée par un nombre τ de périodes consécutives : dans une année académique classique, il y a deux semestres $\tau = 2$ (e.g. $\{t_0, t_1\}$).

En sus, nous considérons que les crédits ECTS conditionnent aussi la validité d'une solution. Dans ce standard, lorsqu'un cours est validé, il rapporte des points (crédit ECTS) à l'étudiant ; une année correspond à 60 crédit ECTS. Aussi, chaque cours c se voit attribué une contribution en crédit ECTS $\epsilon \in \mathbb{N}$. Une année académique dans un cursus personnalisé est donc considérée valide lorsque l'affectation des cours est telle que :

$$\sum_{i=0}^{\tau} \sum_{c_j=0}^{|\mathcal{C}_{(\mathcal{P}, t+i)}|} \epsilon_{c_j} \geq 60 \quad (2)$$

s.t. $t = \min(Y)$

Finalement, un cours c est un 4-tuples défini comme suit :

$$c = \langle \mathcal{K}_c, \mathcal{K}_c^-, \mathcal{T}_c, \epsilon \rangle \quad (3)$$

Comme contrainte supplémentaire, notons qu'un cours s'étend sur toute la période dans laquelle il a été sélectionné. De fait, un étudiant est limité par le nombre maximal de cours γ qu'il peut suivre dans une période t , tel que dans l'absolue $|C_t| = \gamma$, avec $C_t = \{c \mid t \in \mathcal{T}_c\}$ les cours sélectionnés en t . Incidemment, cela englobe un ensemble de sous-problèmes intéressants, comme le cursus possédant le moins de cours possible. En conclusion, un cursus \mathcal{P} est un ensemble de cours C exprimés par un T -uplets tel que $\mathcal{P} = (C_{t_0}, \dots, C_{t_T})$.

2. Une période t_i est avant une autre t_j ssi. $\forall i, j \in \mathbb{N}, i < j \iff t_i \cap t_j = \emptyset, \forall x \in t_i, \forall y \in t_j, x < y$.

2.2 La compétence et sa maîtrise

Dans notre problème du cursus personnalisé, les compétences k sont les entités utilisées pour identifier si un étudiant respecte les attendus de ses objectifs \mathcal{O} ainsi que ceux des cours *via* les prérequis (cf. Equation 1). Pour ne pas souffrir d'une définition de la compétence – sujet controversé – peu inclusive des différentes institutions, nous ne considérerons k que comme un artefact informatique générique et facilement dérivable pour introduire de nouveaux concepts (*e.g.* savoir-faire). Néanmoins, même s'il est possible d'attacher explicitement une sémantique à une compétence (*e.g.* la gestion de projet), évaluer son degré de maîtrise par un étudiant n'est pas un processus formel.

Les acteurs pédagogiques en charge de cette quantification réalise souvent une projection des notes des étudiants dans des échelles de représentation de maîtrise des compétences. Un exemple est l'échelle *MAPI* [10], utilisé dans certains établissements français comme des écoles d'ingénieurs. Elle permet de définir une compétence comme étant **Maîtrisé**, **Acquis**, **Partiellement acquis** ou **Insuffisant**. Ces échelles, d'une manière générale, sont un outil de décision important pour ces acteurs, mais aussi pour les institutionnels, grâce à la sémantique inhérentes qu'elles véhiculent : elles permettent l'encodage de la maîtrise m_k d'une compétence dans des espaces flous interprétable par l'humain.

Ainsi, nous définissons une fonction $\mu(m_k) : [0; 1] \rightarrow \mathbb{S}$ projetant le score de maîtrise d'une compétence m_k dans un ensemble ordonné \mathbb{S} d'échelles. Par exemple, l'échelle de notation *MAPI* se définit comme un ensemble ordonné $\{\mathbf{i}, \mathbf{p}, \mathbf{a}, \mathbf{m}\}$ où $\mathbf{i} < \mathbf{p} < \mathbf{a} < \mathbf{m}$. Une instanciation d'une telle échelle de notation pourrait être telle que $\mu(m_k) = \mathbf{i}$ pour $0 \leq m_k < 0.25$, $\mu(m_k) = \mathbf{p}$ pour $0.25 \leq m_k < 0.5$, $\mu(m_k) = \mathbf{a}$ pour $0.5 \leq m_k < 0.75$ et $\mu(m_k) = \mathbf{m}$ pour $0.75 \leq m_k \leq 1$.

Considérant cela, chaque compétence $k \in \mathcal{K}_c$ apporte un score de maîtrise $m_k \in [0; 1]$. Formellement, le maîtrise d'une compétence k par un étudiant à un instant t se définit par l'accumulation des maîtrises $\mathcal{M}^k = \{m_k \mid k \in \mathcal{K}_c, c \in \mathcal{C}\}$ érodées au fil du cursus d'une quantité \mathcal{Q}_k (cf. subsection 2.4), tel que :

$$\forall k \in \mathcal{K}, t \in T, m_{(k,t)} = \sum_{i=0}^{|\mathcal{M}^k|} m_i - \sum_{j=0}^t \mathcal{Q}_k \quad (4)$$

2.3 Profil de l'étudiant

Outre la manière dont les compétences d'un étudiant s'érodent dans le temps – que nous traitons séparément dans la section 2.4 suivante – les éléments ci-après nous permettent de définir le profil de l'étudiant, qui intervient inexorablement dans la résolution du problème pour consolider la personnalisation du cursus.

Un étudiant possède, à une période t_i , un ensemble de compétences apprises $\mathcal{I}_i \subseteq \mathcal{K}$ qui, lorsqu'il débute son cursus (*i.e.* $i = 0$) représente de fait ses compétences initiales – celles acquises en amont de son entrée dans le cursus. Le profil est aussi défini par les objectifs $\mathcal{O} \subseteq \mathcal{K}$ de l'étudiant. Chacune de ces compétences $k \in (\mathcal{I}_i \cup \mathcal{O})$ est quantifiée par le niveau de maîtrise m_k . Aussi, lorsqu'un

étudiant suit un cours c , son ensemble de compétences apprises évoluent en accord avec les compétences apportées, tel que $\mathcal{I}_{i+1} = \mathcal{I}_i \cup \{\mathcal{K}_c \mid c \in C_{t_i}\}$. Ceci est important, en cela que si l'étudiant change ses objectifs \mathcal{O} sans avoir terminé son cursus (*e.g.* redoublement, reconversion) il devient alors possible de recalculer un nouveau cursus optimal en accord avec ses compétences acquises.

Finalement, un cursus personnalisé \mathcal{P} peut être considéré comme solution lorsque l'ensemble des compétences acquises par l'étudiant à la fin du cursus $\mathcal{I}_{|T|}$ contient les compétences requises par l'objectif \mathcal{O} de l'étudiant, avec au moins le niveau de maîtrise m_k attendu :

$$\forall i \in \mathcal{O}, \exists i \in I_{|T|}, m_i^{k\mathcal{O}} \leq m_i^{k\mathcal{I}_{|T|}} \quad (5)$$

Ainsi, deux étudiants ayant un même objectif \mathcal{O} mais un profil différent pourront possiblement se voir recommander deux cursus différents $\mathcal{P}_1 \neq \mathcal{P}_2$. La raison étant que l'ensemble \mathcal{I}_i est propre à un étudiant : il est vraisemblable que $\mathcal{I}_0^1 \neq \mathcal{I}_0^2$, ou qu'au moins le niveau de maîtrise des compétences soit différent.

2.4 Modélisation de l'érosion de la compétence

Il est captieux de penser la compétence comme binaire (*i.e.* possédée ou perdue). Comme le montre plusieurs travaux [6], la maîtrise d'une compétence peut changer au fil du temps : diminuer lorsqu'elle n'est pas sollicitée, et inversement, indépendamment du contexte pédagogique. Prédire cette évolution dans le temps est complexe mais important pour favoriser l'apprentissage, comme l'illustre les travaux dans les systèmes de répétitions espacées (SRS) [12].

De ce fait, il est nécessaire de prendre en compte cette érosion de la compétence dans la recherche d'un cursus personnalisé pour l'étudiant. Autrement, le risque majeur est de lui fournir des cursus extrêmement difficiles ou irréalisables au vue de ses capacités. En effet, même si une compétence est acquise par l'étudiant à un instant t , sa maîtrise m_k peut décroître au point où l'on peut considérer qu'il ne la maîtrise plus suffisamment pour respecter les prérequis $\mu(m_c)$ des cours planifiés, tel que $\mu(m_k) < \mu(m_c)$. Pour cela, nous définissons un modèle prédictif de l'érosion de la compétence k d'un étudiant tel que :

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto \delta_k(t) \end{aligned} \quad (6)$$

δ_k prédit la valeur de la perte de maîtrise de la compétence k à une période donnée $t_s \in T$ dans le cursus, depuis sa dernière mobilisation en $t_r \in T$. Ainsi, si un cours possède une compétence k en prérequis, et qu'il est planifié à un instant t_s dans le cursus, et que la compétence k n'a pas été renforcée (ou même nouvellement acquise) depuis t_r , alors la perte à ce moment t_s précis est prédit par $\delta(\Delta_t)$, avec $\Delta_t = t_s - t_r$ et $t_r \leq t_s$. Par conséquence, et parce que nous discrétisons le temps, la quantité \mathcal{Q}_k de perte de maîtrise (*i.e.* l'érosion) d'une compétence k à un instant t_s se définit comme suit :

$$Q_k = \sum_{t=0}^{\Delta_t} \delta_k(t) \quad (7)$$

Pour conclure cette section sur la formalisation du problème, il faut noter que nous avons opéré certaines simplifications. Parmi celles-ci, nous considérons un cours comme une entité holistique, en cela que les compétences qu'elle apporte est l'union des compétences enseignées par ces différents constituants (e.g. cours magistraux, travaux dirigés, pratiques, etc.) ; nous considérons aussi que la fonction d'érosion est la même tout au long du cursus (i.e. $\delta_{(k,t)} = \delta_k$), or il est envisageable que ce ne soit pas le cas plus l'étudiant apprend à apprendre.

3 Méthode de résolution

Dans cette section, nous présentons une méthode de résolution du problème de cursus académique personnalisé, comme décrit dans la section 2. De par la morphologie du problème, adopter une approche par satisfaction de contraintes est une première étape nécessaire pour explorer sa topologie et le comprendre correctement [3]. Néanmoins, ce problème revêt une complexité importante³ et peut vite devenir non solvable avec des approches exactes. Pour palier à cela, nous faisons les hypothèses suivantes :

Hypothèse 1 *Une année académique est constituée de deux semestres (i.e. $\tau = 2$). La somme des crédits ECTS d'un semestre doit toujours être supérieure ou égale à la moitié des crédits d'une année (i.e. 30) : cela permet de résoudre le problème pour des situations où l'étudiant arrive en cours d'année.*

Hypothèse 2 *L'érosion est exactement la même, quelque soit la compétence : $\forall k_i, k_j \in \mathcal{K}, \delta_{k_i}(x) = \delta_{k_j}(x)$. De plus, il n'existe pas d'érosion à l'intérieur d'un semestre, soit $\delta_k(0) = 0$, puisque $t_s = t_r$. Enfin, inspiré par les travaux sur l'oubli [1] – et notamment ceux d'Ebbinghaus dans le domaine des neurosciences [6] – nous définissons la fonction d'érosion comme :*

$$\delta(x) = \frac{e^{\frac{x}{s}} + 5}{100}, 1 \leq x \leq 10 \quad (8)$$

Cette équation traduit le fait que moins un étudiant utilise sa compétence, plus il en perd la maîtrise. Attention toutefois, cette fonction est définie principalement pour un cursus de 5 ans. Pour une autre durée, il convient d'en modifier le coefficient de stabilité mémoriel s proposé par Ebbinghaus (ici, $s = 2$).

Hypothèse 3 *Il n'est pas possible d'avoir deux fois le même cours dans \mathcal{P} .*

3. Ce problème est très probablement au moins NP-Complet ; nous finalisons la preuve de complexité réalisée *a posteriori* des premières expérimentations.

Pour résoudre notre problème, prenons $\mathcal{D} = \{c_1, \dots, c_n\}$ l'ensemble des tous les cours disponibles (*i.e.* le catalogue de cours). Prenons \mathcal{O} l'ensemble des compétences nécessaires pour l'objectif de l'étudiant. Soit un ensemble de m variables $X = \{x_1, \dots, x_m\}$, où $x \in \mathcal{D}$, et l'ensemble \mathcal{C} de contraintes suivant :

- Chaque compétence de l'objectif doit être présent dans le profil de l'étudiant final, et sa maîtrise supérieure ou égale à celle attendue ;
- Chaque semestre doit apporter au moins 30 crédits ECTS ;
- Chaque cours dans le cursus doit être unique ;
- Un cours ne peut être sélectionné que si tout ses prérequis sont satisfaits ;
- Un cours ne peut être sélectionné que lorsqu'il est disponible dans la période en question ;
- Le nombre de cours d'un semestre doit être strictement égal au nombre maximum de autorisé par semestre.

Ainsi, le problème du cursus personnalisé avec érosion des compétences consiste à trouver une affectation totale $\alpha : X \rightarrow \mathcal{D}$ qui satisfasse toutes les contraintes dans \mathcal{C} , dénoté $\alpha(\mathcal{C}) = 1$. Le problème est non solvable lorsque, quelque soit les affectations α , $\alpha(\mathcal{C}) = 0$.

4 Implémentation et résultats expérimentaux

En accord avec les propositions précédentes, nous avons développé deux solveurs de contraintes *prolog*⁴ pour résoudre le problème du cursus personnalisé avec érosion des compétences. Le *Solver 1* utilise des heuristiques de résolution issues du modèle prédictif induit par l'érosion pour conditionner les cours à choisir en priorité, et le *Solver 2* qui évalue la validité d'une solution une fois l'affectation réalisée. L'intérêt de ce solveur est d'être similaire à une démarche utilisateur : l'affectation étant vérifiée *a posteriori*. Les objectifs principaux de cette expérimentation étaient d'une part d'étudier la faisabilité de trouver une solution à ce problème, et d'autre part d'étudier l'effet de ses propriétés (*e.g.* érosion, nombre de cours). Avant de présenter les résultats globaux, précisons que ces expérimentations ont été réalisées avec un processeur Intel i5-8250 à 2.0 GHz et 8 Go de RAM. Dans la suite, nous considérons plutôt les inférences logiques (LI) réalisées par les solveurs plutôt que le temps pris pour trouver une solution⁵. De plus, en raison de l'application prévue de notre problème dans le monde académique, nous avons défini une limite de 12h de calcul au delà de la quelle le problème est considéré comme non solvable en un temps raisonnable.

Définition des jeux de test (i.e. catalogue de cours). Pour palier l'absence de catalogues de cours publiquement disponibles, nous avons été obligés de générer ces catalogues de manières aléatoires⁶ pour chaque instance du problème, tout

4. Le code source, les jeux de tests utilisés et leur configuration sont disponibles sur Software Heritage à l'adresse suivante : archive.softwareheritage.org/swh:1:dir:1eba2c7a6dbec16774a59a0988e6101f3e99851f

5. $9 \times 10^5 \text{LI} \approx 1\text{s}$.

6. Il n'y a donc aucune assurance *a priori* quant à l'existence d'une solution.

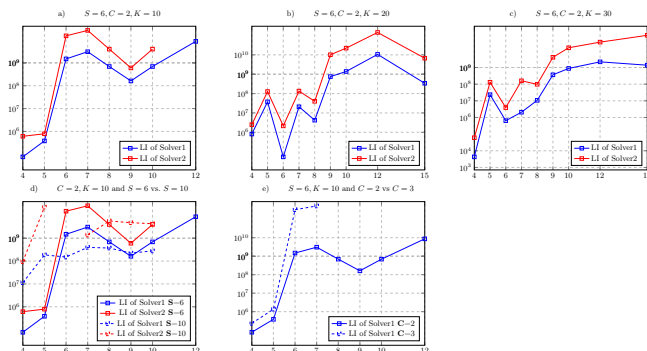


FIG. 1 – Comparaison des inférences logiques (LI) (ordonnée) réalisés par nos deux solveurs lorsque le nombre de cours augmente de $N = 4$ à $N = 15$ (abscisse) dans des configurations différentes du problème.

comme l'objectif \mathcal{O} à trouver. Dans ces instances du problème, nous avons changé le nombre de semestres S , le nombre de cours disponibles à chaque semestre N , le nombre de compétences totales existantes K ou le nombre de cours C qu'un étudiant doit choisir à chaque semestre. Indépendamment de l'instance du problème, le nombre de prérequis pour chaque cours fut tiré aléatoirement dans l'intervalle $[0; 4]$ et, pour chacun, leur valeur le fut dans l'ensemble flou $\{\emptyset, i, p, a, m\}$. De même, le nombre de compétences \mathcal{K}_c apportées par un cours fut tiré aléatoirement dans l'intervalle $[1; 4]$, leur valeur m pour chacune l'étant dans l'intervalle $[10; 90]$. Pour l'objectif, nous avons fixé $|\mathcal{O}| = 2$; leur valeur se prenant aléatoirement dans $\{i, p, a, m\}$. L'apport d'un cours concernant les crédits ECTS se prenant dans l'intervalle $[2; 6]$.

Nous avons autant que possible tenté de nous rapprocher de situations réelles, mais la difficulté du problème nous a obligé à restreindre l'espace explorable, notamment pour les cours par semestre C . Dans ces expérimentations, nous commençons au premier semestre $S = 0$ et l'étudiant n'a aucune compétence initiale ($\mathcal{I}_0 = \emptyset$), ceci afin de diminuer les biais de parcours. Les résultats expérimentaux sont présentés dans la Figure 1.

Résultats généraux Notons tout d'abord qu'au regard des inférences logiques (LI) réalisées, le solveur 1 utilisant l'érosion pour guider la recherche présente de meilleur résultat que le solveur 2 dans tous les cas testés. Ce résultat montre qu'un modèle pédagogique prédictif peut être efficacement utiliser pour guider la recherche lors d'un problème de satisfaction par contraintes.

Dans la Figure 1 a), b) et c), nous tentons de trouver un cursus de licence (L1, L2 et L3) personnalisé ($S = 6$) où deux cours doivent être suivis par semestre

($C = 2$). Nous y augmentons le nombre de cours N disponible à chaque semestre (en abscisse) : aux alentours de $N = 15$, le problème devient généralement non solvable en un temps raisonnable. Nous y faisons aussi varier le nombre de compétences existantes K à 10 pour a), 20 pour b) et 30 pour c). On observe que le problème a) devient insolvable à partir de $N = 10$ pour le solveur 2 et à partir de $N = 12$ pour le solveur 1. Dans d), nous comparons la résolution d'un cursus de licence a) avec la résolution du problème concernant un cursus allant jusqu'au master dans les mêmes conditions – excepté $S = 10$ pour ce dernier. On observe que, dû à l'explosion combinatoire, le problème devient non solvable en temps raisonnable pour $N = 10$ (courbes en pointillées). Dans la Figure e), nous comparons l'effet du nombre de cours à prendre $C = 2$ et $C = 3$. Les résultats montrent clairement que lorsque $C = 3$, le problème devient non solvable en un temps raisonnable beaucoup plus tôt ($N = 7$).

Nos résultats expérimentaux sur ce nouveau problème mettent en exergue sa difficulté de résolution, le rendant non praticable pour un humain. Aussi, l'utilisation des méthodes exactes pour trouver une affectation α tel que $\alpha(C) = 1$ ne semble pas recommandée, tout du moins sans heuristiques efficaces pour explorer l'arbre de recherche. Néanmoins, grâce aux observations et aux résultats obtenus, il nous est possible d'énoncer de premières recommandations prescriptives, pouvant ainsi servir d'aides à la décision pour les différents acteurs concernés.

5 Recommandations prescriptives

Compétences et submersion. Une observation intéressante qui résulte de la Figure 1 a), b) et c) est l'évolution des inférences logiques nécessaires en fonction de la quantité de compétence K disponible. Lorsque $C < 8$, plus cette quantité est importante, plus le problème semble facile à résoudre pour le solveur. D'ailleurs, le problème est également solutionné lorsque $N = 15$ pour $K = 20$ et $K = 30$. Ce comportement, que nous appelons *submersion* et qui consiste en un accroissement important des chances de retrouver dans les cours les compétences nécessaires à l'objectif, peut avoir deux explications. La première tient à la distribution probabiliste des compétences au sein des cours lors de la création aléatoire du problème. La deuxième explication résonne avec un caractère d'avantage naturel : plus la granularité des compétences est faible – autrement dit moins elles sont nombreuses et donc précises – et plus ces dernières seront enclines à être associées à des cours, même si elles ne recouvrent pas entièrement la portée pédagogique de ces cours. De fait, les acteurs pédagogiques et institutionnels doivent faire montre d'une grande vigilance dans l'identification des compétences existantes dans leurs formations, au risque sinon de compromettre la pertinence des cursus solutions.

De l'effet de la quantité des cours et de la taille du cursus. Dans d), où nous nous intéressons à l'effet de la durée du cursus dans la résolution en comparant $S = 6$ et $S = 10$, on observe que le problème est plus difficile pour $S = 10$ – c'est d'autant plus flagrant avec le solveur 2. Aussi, ces résultats tendent à confirmer

que plus un cursus personnalisé est long, plus il devient difficile de trouver une affection α qui puissent satisfaire les attentes des acteurs : cela motive d'autant plus la nécessité de fournir un outil d'aide à la décision pour ce type de problème. Nous pensons donc qu'en absence d'outils efficaces il est préférable d'offrir des cursus personnalisables sur des périodes courtes, comme le master (où $S = 4$), afin de pouvoir assurer une éducation de qualité aux étudiants. La Figure 1 e) nous a permis d'observer que lorsque le nombre de cours à prendre augmente, plus le problème devient difficile à résoudre. Cela s'explique conjointement par l'augmentation du nombre de variables $x \in X$ affecté et l'augmentation de la probabilité d'enfreindre les contraintes. Aussi, au vue des résultats, nous pensons que les cours doivent être correctement équilibrer au sein des formations, quitte à devoir y proposer un même cours également plus tard lorsque possible : il n'est plus besoin de conditionner l'étudiant par son année d'étude, portant la réflexion selon nous sur l'établissement d'une hiérarchie de cours cohérente.

Consolider l'établissement de référentiels taxinomiques des compétences. Il n'est pas besoin de rappeler que l'établissement d'un référentiel de compétences pour une institution est une tâche extrêmement complexe, surtout si une notion taxinomique y intervient (*e.g.* dépendances, niveaux). De manière intéressante, nos premiers résultats expérimentaux montrent qu'il n'y a pas besoin d'une telle organisation pour pouvoir répondre au problème du cursus. Plus encore, et à partir de l'Equation 1, l'on remarque qu'il est possible de fournir un outil d'aide à ces institutions pour établir un CbKST (Competence-based Knowledge Space Theory) [7] à partir d'un catalogue de cours \mathcal{C} disponible de manière automatique une fois ce dernier construit. En effet, puisqu'il est possible d'organiser tous les cours les uns par rapport aux autres dans le cursus, une relation d'ordre partielle apparaît : il devient alors possible d'identifier automatiquement quelles compétences intervient avant quelles autres, et quelles sont leurs prérequis.

6 Conclusion et perspectives

Dans ce papier, nous avons proposé les premières formalisation et résolution du problème du cursus personnalisé avec érosion de compétences en milieu académique. Nous avons montré que ce problème est très difficile dû à sa combinatoire, et qu'il n'est pas envisageable de demander sa résolution par les acteurs académiques sans outils d'aides à la décision adéquates. C'est pour cela que nous avons présenté une première ébauche d'un tel outil, utilisant une méthode de résolution exacte. Bien que les premiers résultats expérimentaux montrent que ce type de méthode n'est pas adapté sans heuristiques plus performantes, nous avons pu identifier de précieuses recommandations pour l'établissement de ces cursus, et avoir une meilleure compréhension de la topologie du problème.

Nous continuons activement nos travaux concernant ce problème. Nous finalisons conjointement la preuve de complexité du problème et sa résolution par des méthodes métaheuristiques, afin de mieux le comprendre et trouver des solutions pertinentes. Ces travaux permettront d'aboutir un outil exploitable par

les étudiants, les acteurs pédagogiques et les institutions. Les étudiants pourront ainsi élaborer et personnaliser leur cursus de manière assistée; les enseignants gagneront un outil capable de les aider dans l'accompagnement de l'étudiant, dans ses efforts de rémédiation et de recommandation en temps réel; et les institutions pourront vérifier l'intégrité des parcours sélectionnés par les étudiants et les points de congestions, ainsi que vérifier la cohérence de leur catalogue de formation, et favoriser les collaborations avec d'autres établissements.

Remerciement

Cette contribution scientifique a été réalisée dans le cadre du projet de recherche APACHES, financé par l'i-SITE ULNE (FIPE18-007-VERMEULEN).

Références

1. Averell, L., Heathcote, A. : The form of the forgetting curve and the fate of memories. *Journal of Mathematical Psychology* **55**(1), 25 – 35 (2011), special Issue on Hierarchical Bayesian Models
2. Backenköhler, M., Scherzinger, F., Singla, A., Wolf, V. : Data-driven approach towards a personalized curriculum. *International EDM Society* (2018)
3. Brailsford, S.C., Potts, C.N., Smith, B.M. : Constraint satisfaction problems : Algorithms and applications. *European Journal of Operational Research* **119**(3), 557 – 581 (1999)
4. Cappellari, L., Lucifora, C. : The “bologna process” and college enrollment decisions. *Labour Economics* **16**(6), 638–647 (2009)
5. Desmarais, M.C., Baker, R.S. : A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction* **22**(1-2), 9–38 (2012)
6. Ebbinghaus, H. : Memory : A contribution to experimental psychology. *Annals of neurosciences* **20**(4), 155 (2013)
7. Heller, J., Steiner, C., Hockemeyer, C., Albert, D. : Competence-based knowledge structures for personalised learning. *Int. Journal on E-learning* **5**(1), 75–88 (2006)
8. Lebis, A., Prior, E., Mandran, N., Karami, A., Vermeulen, M. : Promouvoir et soutenir la Pédagogie Par Projet Centré Humain dans le supérieur : le projet APACHES. In : DIDAPRO 8 - DIDASTIC. Lille, France (Feb 2020)
9. Lefevre, M., Jean-Daubias, S., Guin, N. : An approach for unified personalization of learning. In : Perez-Marin, M.K.O.S.J.B.D. (ed.) *International Workshop on Personalization Approaches in Learning Environments - Conference User Modeling, Adaptation, and Personalization*. pp. 5–10. Montreal, Canada (Jul 2012)
10. Ministère de l'éducation nationale : Evaluer la maîtrise du socle commun aux cycles 2, 3 et 4. <https://eduscol.education.fr/cid103803/evaluer-la-maitrise-du-socle-commun-du-cycle-2-au-cycle-4.html> (2016), accessed : Jan. 19, 2020
11. Molontay, R., Horváth, N., Bergmann, J., Szekrényes, D.L., Szabó, M. : Characterizing curriculum prerequisite networks by a student flow approach. *IEEE Transactions on Learning Technologies* (2020)
12. Settles, B., Meeder, B. : A trainable spaced repetition model for language learning. In : *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*. pp. 1848–1858. Association for Computational Linguistics, Berlin, Germany (Aug 2016)
13. University of Reading : The modular system at reading explained. https://student.reading.ac.uk/essentials/_study/course-and-departments/the-modular-system-at-reading-explained.aspx (2019), accessed : Jan. 19, 2020
14. Wong, C. : Sequence based course recommender for personalized curriculum planning. In : *Int. Conference on A.I. in Education*. pp. 531–534. Springer (2018)

Intention d'utilisation des tablettes en formation initiale d'enseignants à Mayotte

Gaëlle Lefer Sauvage¹

¹ Centre Universitaire de Formations et de Recherches à Mayotte

Résumé. A Mayotte, la formation universitaire des enseignant.e.s propose depuis 2019 d'utiliser des tablettes tactiles dans la pratique enseignante. A partir des travaux sur l'acceptabilité [14, 30], 87 professeur.e.s des écoles stagiaires de Mayotte ont répondu à un questionnaire en ligne pour déterminer la manière dont se forment leur intention d'utilisation selon les contextes (en formation initiale / en classe devant les élèves). Les résultats montrent que les contextes participent grandement à l'intentionnalité et que le sentiment de compétences est médiateur de l'acceptabilité. L'impact de l'intention d'utilisation en formation universitaire sur les celles en classe est nuancé, sauf pour les étudiants possédant une tablette personnelle, pour qui les frontières d'usage entre les contextes sont plus perméables. Ainsi, le modèle de l'acceptabilité est questionné dans une perspective contextuelle et des réflexions sur la formation initiale sont envisagées.

Mots-clés : Acceptabilité, intention d'utilisation, tablettes, contextes, Mayotte.

Abstract. In Mayotte, since 2019, the university training of teachers suggest using tactile tablets in teaching practice. Based on technology acceptance model [14, 30], 87 trainee teachers in Mayotte schools answered an online questionnaire to determine context' effect in intention to use tactile tablets (in initial training / in class in front of children). The results demonstrate that contexts contribute to intentionality and self efficacy mediates acceptability. The impact of intentionality of use in university training on the classroom is nuanced, except for students with a personal tablet. In this case, the frontiers of use between contexts are more permeable. Thus, the model of acceptability is questioned from a contextual perspective and reflections on initial training are envisaged.

Keywords: Technology Acceptance Model, Mobile Tablets, Contexts, Mayotte.

1 Contexte de l'étude : les tablettes tactiles à l'école et à l'Université à Mayotte

Le déploiement d'outils numériques (ordinateur, TBI) au sens large à Mayotte est balbutiant. Dans le cadre d'une enquête pour le Cnesco en 2018 : « en 2016, le nombre d'ordinateurs pour 100 élèves dans les écoles élémentaires va de 1 à Mayotte à 30 en Lozère » [13, p.33]. Concernant les tablettes, certes, les circonscriptions du premier degré en acquièrent régulièrement, mais leur nombre est insuffisant pour doter chaque école. De plus, l'insécurité ambiante, liée à la possession de matériel, ne permet pas de laisser une grande autonomie de fonctionnement des enseignants. Les tablettes sont donc centralisées dans les inspections des circonscriptions et sont disponibles en prêt à la demande des enseignant.e.s pour un temps souvent limité. Quelques enquêtes [1,

3] ont permis de décrire le paysage en équipement. Plusieurs sources explicatives sont proposées pour comprendre les freins institutionnels et contextuels qui participent au non-usage (insécurité du fait de vols antérieurs de matériels dans les classes, manque de formation des enseignant.e.s sur l'insertion des outils informatiques en classe, manque de formateurs ou d'instituteurs.trices animateurs.trices informatiques -IAI-⁴ et géographie difficile à couvrir). À notre connaissance, deux enquêtes [20, 23] ont jusque-là été menées sur l'usage du numérique dans le domaine scolaire et universitaire à Mayotte. À partir d'un questionnaire, 250 étudiant.e.s mahorais.e.s toutes filières et tous niveaux universitaires confondus (dont 77 personnes en Master), sont interrogés sur les usages d'outils numériques [23]. Les résultats montrent que les étudiant.e.s se déclarent comme moyennement experts du numérique, avec une tendance des réponses allant vers la maîtrise (12,3 % pas du tout *versus* 21,7 % tout à fait). Les auteurs montrent aussi que les étudiant.e.s considèrent qu'ils se sont auto-formés à Internet (55,1%) ou formés par l'école (52,7%). Les chercheurs comparent ces résultats avec ceux des étudiant.e.s de Bretagne qui sont sensiblement les mêmes. Parmi les équipements des étudiant.e.s de Mayotte, les résultats de l'enquête montrent que 80 % des étudiant.e.s possèdent un ordinateur portable dans la sphère privée, 50 % possède un mobile avec internet et 25 % possède une tablette. Par contre, ces équipements ne sont pas utilisés dans la sphère universitaire (moins de 10% d'équipement pour chaque variable susmentionnée). Ces éléments doivent être soumis à plus de prudence pour la population des étudiant.e.s en formation en Master, qui sont financé.e.s pour leurs études, comparativement aux étudiant.e.s en Licence qui n'ont pas de financement [23]. Ce sont également des mesures déclaratives. Toutefois, ces éléments mettent en valeur le fait que les étudiant.e.s à Mayotte déclarent ne posséder que peu d'outils mobiles dans le cadre de leur formation et qu'ils se sentent relativement compétents pour utiliser le « numérique » dans leur quotidien.

2 Acceptabilité et intention d'utilisation des tablettes tactiles

Les caractéristiques des tablettes tactiles induisent des attitudes et des représentations dont les impacts pédagogiques sont forts, ce qui leur confère un statut particulier : elles seraient porteuses d'un changement de paradigme éducatif [24]. L'industrialisation est perçue comme la marque d'un changement social dont la technologie est le vecteur et dont l'éducation devient l'enjeu, comme le précise [17, p.97].

La pratique des tablettes tactiles (dans le contexte de Mayotte, mais aussi dans celui de l'enseignement scolaire) reste relativement marginale. Les débats scientifiques autour de l'acceptation des nouvelles technologies et de l'acceptabilité portent sur le lien existant entre l'intention d'utiliser une nouvelle technologie (ici, les tablettes tactiles) et l'utilisation réelle. Ce lien supposé entre une anticipation projetée de l'usage est au départ conçu comme déterministe, ayant un impact direct sur l'utilisation [14]. Mais cette conception évolue, et laisse place à deux paradigmes : celui du continuum possible entre l'intention d'utilisation et l'appropriation, médiée par un ensemble de variables psychologiques et sociodémographiques [3, 11, 25, 30, 31, 32] ; un autre modèle, celui de l'acceptation [6] qui considère que c'est au sein de l'appropriation que l'intention d'utilisation se modifie et qu'elle reste dépendante des contextes. L'intention d'utilisation *a priori* n'est pas déterministe puisqu'elle est re-négociée au moment de l'utilisation et de l'appropriation. Ce modèle met en valeur 5 dimensions à l'œuvre dans l'acceptation : la dimension intra-personnelle entendue dans une dimen-

sion plutôt cognitive (à travers une évaluation des coûts et des bénéfices que génère le dispositif), la dimension inter-personnelle qui met en valeur la manière dont les collectifs et les réseaux de travail se ré-équilibrent, la dimension socio-organisationnelle, qui prend en compte « la manière dont les organisations cherchent à modifier le système de contrôle et d'autonomie des salarié » (p.366), la dimension transpersonnelle, biographique et anthropologique, où la technologie met en exergue le pouvoir d'agir de la personne, et la dimension relative à la fiabilité technique (façon dont la personne accorde une confiance en la technicité du dispositif). Ce modèle contextuel est pertinent et riche mais, même s'il est envisagé dans pratique réelle de la technologie (et moins dans une conception *a priori*), il peut tout à fait être adapté à une intentionnalité préalable à l'utilisation. Autrement dit, la connaissance technique de l'outil mais surtout les variables psychologiques et symboliques attribuées à l'outil (valeur ajoutée de l'outil, usage potentiel dans l'apprentissage, utilité perçue de la tablette, sentiment de compétences) sont d'autant plus importantes qu'elles sont considérées comme un facteur participatif de l'utilisation [3, 11, 25, 30]. Lorsqu'un outil est considéré comme utile (par rapport aux gains et aux coûts cognitifs), facile et fiable au départ, il facilite (et ne détermine pas) le processus d'appropriation et d'intériorisation des compétences techniques et cognitives à l'œuvre dans l'usage. L'intention d'utilisation est alors évaluée à partir de deux indicateurs : l'utilité perçue de la nouvelle technologie et la facilité d'utilisation perçue. L'utilité renvoie « au degré auquel une personne croit que l'utilisation d'un système va améliorer ses performances. La facilité d'utilisation perçue fait référence au degré auquel une personne croit que l'utilisation du système lui demandera peu ou pas d'effort » [15, p.26].

3 Enjeux de l'usage et des pratiques (utilisations)

Les notions d'usage, d'utilisation ou de pratique laissent place à un débat scientifique qu'il est difficile de clarifier, tant les conceptions sont nombreuses, dépendantes de chercheurs et de paradigmes, mais aussi de disciplines scientifiques. D'ailleurs, même Rabardel [28] utilise parfois les deux notions de façon interchangeable : « Il manque encore à l'artefact de s'inscrire dans des usages, des utilisations, c'est-à-dire des activités où il constitue un moyen mis en œuvre pour atteindre les buts que se fixe l'utilisateur » (p.80). Jouët [19] évoque quant à lui une « individualisation de l'usage » (p.500). L'enjeu de ces notions est de comprendre la place de l'individu dans son rapport avec la technologie, et la place du collectif dans ce même rapport, dont le dispositif technique et les profils usagers sont au centre. Globalement, on considère que l'usage est un construit social [19], alors que l'utilisation est une construction individuelle, le premier étant le prolongement du second [19] -même si cela reste à confirmer-. L'usage est relativement stable dans le temps et reproductible. L'utilisation peut être stable (intégrant les expériences passées et projetées) mais aussi temporaire (dépendante de la situation) et pas nécessairement reproductible. L'usage et la pratique sont incorporés par individu ou le collectif, mais là encore, le focus temporel (en amont ou en aval) accordée à la technique est questionné. Jouët [19] rappelle aussi que « la recherche réfute le schéma causal selon lequel les usages seraient le simple produit de la diffusion et de l'adoption de ces technologies » (p.508) et que l'usage dépasse les frontières géographiques et de lieux.

4 Question de recherche

Les nombreux travaux issus de la Recherche EXTATE [33] ont permis de montrer qu'il y a des contraintes fortes dans l'usage des tablettes en classe. Les caractéristiques des tablettes amènent à des contraintes sur l'activité de l'enseignant qui sont étudiées dans le cadre du rapport Extate et repris ultérieurement [34] : difficultés techniques (besoin du réseau Wifi, manque d'espace de dépôt, etc.) et d'organisation cognitive (gestion de fenêtres multiples, organisation des données sous forme d'organigramme, absence de visibilité du processus d'élaboration cognitive des élèves), contraintes institutionnelles, etc. Malgré cela, l'utilisation des tablettes favorise l'entrée dans l'activité [21] notamment par le jeu [18], mais il n'y a pas de consensus entre les enseignant.e.s sur le fait qu'elle permette d'apprendre ou de mieux apprendre [18].

Dans le cadre de la formation continue, les recherches ont aussi montré que les étudiant.e.s en formation d'enseignant.e.s ont tendance à reproduire les pratiques des formatrices. C'est-à-dire que l'usage des TICE (Technologie de l'information et de la communication en éducation) en général est très faible en formation initiale et sur leurs lieux de stage, mais qu'il est souvent visible à long terme, après la formation initiale [3, 26, 27]. Des chercheurs ont aussi montré que, quand bien même les étudiant.e.s considèrent important d'utiliser les nouvelles technologies dans les apprentissages [2], la réalité de la classe et le manque de temps lié au manque d'expérience, amènent les jeunes professionnel.le.s à ne pas privilégier les nouvelles technologies [7, 27]. On sait aussi que les compétences développées à l'université en matière de maîtrise de numérique et celles ré-exploitées en classe par les enseignant.e.s sont très distinctes [4, 5]. La maîtrise de la technicité des outils, décontextualisée des apprentissages, n'est pas suffisante ; elle peut même être contre-productive, au risque de considérer la formation comme prescriptive et renforçant les écarts dans l'usage réel [4, 12, 27]. Les outils mobiles sont peu exploités, aussi bien dans la formation initiale des professeur.e.s stagiaires que sur le terrain par les enseignant.e.s, au bénéfice d'outils plus conventionnels tels que des ordinateurs ou vidéo-projecteurs [9]. Cet élément trouve une explication dans le fait que les technologies doivent s'insérer dans la pratique quotidienne des enseignants [4, p.58].

Face à ces contraintes, aux bénéfices accordés aux pratiques des tablettes particulièrement, et aux spécificités des populations enquêtées, cet article interroge la manière dont l'intention d'utilisation se construit. L'intention renvoie au fait que la pratique est projetée, imaginée, non acquise, et que la connaissance de l'utilisateur est emprunte d'informations générales, d'expériences succinctes et en partie construites psychiquement, et d'injonctions universitaires et/ou ministérielles. Les travaux de Jouët [19] ou encore de Rabardel [28] nous confortent dans l'utilisation des « intentions d'utilisation » plutôt que d'usage, même si certains auteurs confondent les deux. La plupart des recherches sur l'acceptabilité partent d'un principe prédictif selon lequel l'intention amène à la pratique puis ensuite à l'usage [6], ce qui n'est pas le parti pris ici : la question soulevée dans cet article est de comprendre le poids des contextes dans l'intention d'utilisation des PES¹ l'objectif étant de saisir les caractéristiques des organisations (structurelles, culturelles), de la technologie et les caractéristiques personnelles psychologiques issues du modèle l'acceptabilité [3, 14, 15] qui facilitent ou

¹ Professeurs des écoles stagiaires

freinent l'intention d'utilisation des tablettes. Nous faisons ainsi l'hypothèse que l'intention d'utilisation dans un certain contexte peut médiatiser l'intention d'utilisation dans un autre contexte. La méthodologie adoptée dans cette recherche, issue de précédents travaux [20], distincte de celle de Davis [14], permettra probablement d'extraire des variables différentes et complémentaires. Les expressions « CUFR² » ou « en formation initiale » seront utilisées dans le même sens, de même que « en classe auprès des élèves » ou « en classe », pour éviter les redondances.

5 Méthodologie

5.1 Outils

Un questionnaire en ligne est proposé. Le questionnaire est repris de travaux antérieurs [20, 32] composé de 23 questions, évaluant les craintes envers l'usage (4 items), le sentiment de compétences (8 items) et acceptabilité perçue (2 items), la capacité des élèves à l'usage (5 items). 3 items sont aussi relatifs à des informations socio-géographiques classiques (âge, sexe, temps passé à Mayotte) et 1 renvoie à la possession, dans son domicile, d'une tablette ou pas. Le contexte d'usage est contrôlé : en classe auprès des élèves (10 questions) *versus* au CUFR (10 questions). Toutes les questions évaluent en 4 points, allant de *Tout à fait d'accord* à *Pas du tout d'accord* : la facilité perçue, l'utilité perçue, le sentiment de compétences par rapport à l'usage, les craintes relatives à l'usage de la tablette. Pour chaque question et pour chaque contexte (CUFR / en classe), des questions ouvertes sont proposées pour approfondir leurs réponses. L'ordre des questions a été modifié pour éviter un effet d'habitation aux questions.

5.2 Population

Sur les 182 professeur.e.s des écoles stagiaires en Master 1, seul.e.s 87 ont accepté de répondre au questionnaire. L'échantillon représente 59 % de femmes (N=51) et 41 % d'hommes (N=36). Les enquêté.e.s ont entre 18 et 30 ans (50 %), puis entre 30 et 40 ans (38 %). Un peu moins de la moitié de l'échantillon vit à Mayotte depuis plus de 10 ans (43 %), 15 % des personnes interrogées vivent à Mayotte depuis quelques mois et 18 % vivent à Mayotte entre 4 et 10 ans. 38 % des personnes interrogées possèdent une tablette tactile à leur domicile.

5.3 Procédure

Au cours du second semestre 2020, le questionnaire a été déposé sur Moodle, une plateforme libre que les étudiant.e.s sont amené.e.s à utiliser dans le cadre de leurs enseignements. Les étudiant.e.s ont deux possibilités pour répondre au questionnaire : soit chez eux.elles, sur un temps libre, soit lors d'un enseignement d'informatique, dans le cadre de leur formation. Dans ce second cas, les étudiant.e.s sont amené.e.s à répondre au questionnaire, sans obligation de réponse, pendant le cours et sous couvert d'anonymat. Ces choix sont effectués pour faciliter les accès à internet (parce que le CUFR n'est pas doté d'un réseau wifi pour permettre aux étudiant.e.s de se connecter, seule la salle informatique est dotée d'une connexion câblée), et encourager les étudiant.e.s à se poser des questions sur leur propre usage (ces éléments sont ré-exploités dans le cadre de leur formation). Une des limites de ce choix est d'encourager le

² Centre Universitaire de Formations et de Recherches

contrat didactique avec des réponses attendues et un sentiment d'être obligé de répondre. Sur les 87 répondants, 35 réponses ont été faites hors enseignement, sur un temps personnel.

Par rapport à l'analyse des résultats, le logiciel SPSS 16 sera utilisé pour effectuer les descriptions d'analyses statistiques et des analyses de relations entre variables catégorielle (Khi^2 notamment). Les données qualitatives seront étudiées à partir d'analyses inter-juges (trois juges), croisées ensuite avec les grilles d'analyses de Villemonaix [33], sur les usages en classe (nature de la tâche, finalité pédagogique, modalités de mises en œuvre) et intégrées dans le logiciel Tropes pour enrichir les éléments lexico-sémantiques et les données quantitatives textuelles.

6 Résultats

Les résultats présentent d'abord une structure de l'intention d'utilisation selon les contextes d'usage (CUFR / école), pour comprendre les leviers de l'usage, mais aussi déterminer des profils d'usage en fonction des facteurs. Dans ce cadre, des tris croisés seront opérés pour déterminer des relations entre les variables, puis des profils sont dégagés pour tester la structure initiale. Dans un second temps, des analyses qualitatives avec Tropes permettront de déterminer les usages pédagogiques envisagés par les PES pour compléter les analyses précédentes.

6.1 intention d'utilisation (facilité et utilité d'utilisation) selon les contextes

En classe. La plupart des répondants considèrent que la tablette est utile en classe (78%) ou plutôt utile (21%). 86 % des PES considèrent que c'est facile (22 %) ou plutôt facile (51%) d'utiliser les tablettes en classe avec les élèves. Les tests de Khi^2 ne présentent pas de liens entre l'utilité perçue de la tablette et la facilité perçue de la tablette en classe. Les craintes d'usage en classe ne sont pas également pas significativement liées avec les deux facteurs d'intention d'utilisation (utilité, facilité). Mais le sentiment de compétences est lié à la facilité perçue ($Khi^2(9)=18,7, p<.03$) et l'utilité perçue ($Khi^2(6)=13, p<.04$).

Au CUFR. Concernant l'intention d'utilisation en formation initiale, la plupart des répondant.e.s considèrent que la tablette n'est pas vraiment utile en formation initiale (46%), mais que c'est plutôt facile (39 %) ou facile (28 %). 9 % des répondants considèrent que la tablette n'est pas du tout facile d'usage au CUFR et 27 % considère qu'elle n'est pas du tout utile au CUFR. Les tests de Khi^2 présentent un lien relativement élevé entre l'utilité perçue de la tablette et la facilité perçue de la tablette au CUFR ($Khi^2(9)=23,6, p<.005$). Les craintes d'usage au CUFR sont significativement liées à l'utilité perçue au CUFR ($Khi^2(9)=16,81, p<.05$) et la facilité perçue ($Khi^2(9)=18,9, p<.03$). Le sentiment de compétences est lié à la facilité perçue ($Khi^2(9)=96,8, p<.001$) et l'utilité perçue ($Khi^2(9)=17,4, p<.05$).

Articulation des contextes classe / CUFR. Lorsque l'on croise les différentes variables pour déterminer le poids du contexte sur les intentions d'usage, il semble que certaines frontières des contextes puissent dépasser les intentions d'usages. Certes, si l'on cherche à comprendre l'intention d'utilisation en classe (utilité et facilité) à travers des intentions d'usages du CUFR, mais aussi l'ensemble des variables psychologiques explicatives, aucun lien significatif n'apparaît. Mais le sentiment de compétences en classe ($Khi^2(8)=16,6, p<.05$) et l'utilité perçue de la tablette en classe

($\text{Khi}^2(8)=11,5, p<.10$) participent à la genèse de l'utilité perçue au CUFR. Ainsi, malgré les frontières entre les contextes, des variables psychologiques impactent l'utilité perçue.

Dès lors, est-ce que ces éléments participent à des profils de réponses chez les étudiants ?

Profil facilitant l'intention d'utilisation : la possession personnelle de tablettes. Le seul profil qui se dégage de l'ensemble des variables est la possession ou pas d'une tablette tactile personnelle (Fig 1). La plupart des PES questionné.e.s n'ont pas de tablettes. Parmi les PES qui possèdent une tablette ($N=33$), 35 % ne disent utiliser aucune application précise, ni n'ont une idée d'application à tester auprès des élèves (85%) et 15 % a cité des éléments non pertinents (« Android », « internet », « ne sais pas »). A la question « quelle application aimeriez tenter auprès de vos élèves », la plupart des répondant.es de l'enquête souhaitent approfondir des applications déjà connues (80%), rares sont ceux.celles qui proposent d'autres applications que celles qu'il.elles utilisent déjà (10%). Les personnes qui possèdent une tablette personnelle considèrent que c'est un outil plutôt utile ou utile en classe (99%) et que c'est plutôt facile ou facile à utiliser en classe (85%). La plupart (60%) n'ont pas de craintes par rapport à la pratique en classe et se sentent compétent.es ou plutôt compétent.es (95%). La moitié de cet échantillon considère que c'est utile ou que ce n'est pas utile au CUFR. La plupart considère que ce n'est pas facile à utiliser au CUFR (60%). Parmi les PES qui ne possèdent pas une tablette ($N=54$), comparativement aux précédentes réponses, la plupart des étudiants enquêtés (69%) ont des craintes par rapport à la pratique en classe et se sentent compétent.e.s ou plutôt compétent.e.s (75%).

En conclusion de ces premiers résultats, le fait de posséder une tablette tactile chez soi pourrait agir comme un objet transitionnel entre les contextes d'usage [8] qui facilite les liens entre les intentions d'usage au CUFR et ceux en classe. La non possession personnelle d'une tablette agit sur augmentation des craintes par rapport à l'intention d'utilisation en classe (mais pas sur une meilleure connaissance de l'outil tablette, des applications existantes). Le sentiment de compétences reste très lié à l'utilité perçue de l'usage, indépendamment des contextes d'exercice d'usage, même s'il ne permet pas, au regard de ces premiers résultats, de discriminer des profils d'intention d'utilisation. L'intention d'utilisation au CUFR semble plus explicatif des *intentions d'usage en classe plutôt que l'inverse*.

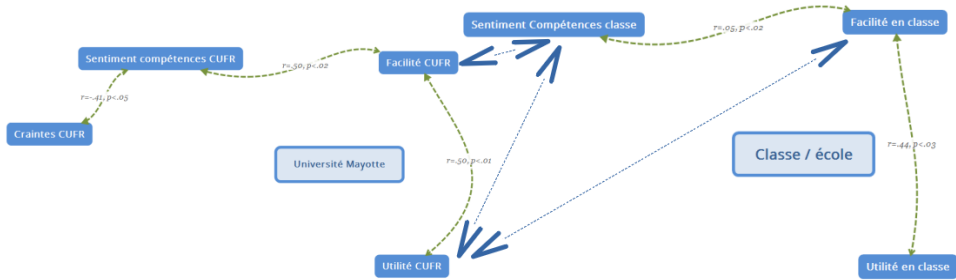


Fig. 1. Profil d'étudiants possédant une tablette personnelle et relation entre variables psychologiques de l'intention d'utilisation

6.2 Poids des systèmes technique et technologique dans l'intention d'utilisation : articulées ou en concurrence ?

Les réponses des PES mentionnent des éléments qui articulent l'outil tablette avec un ensemble d'autres outils. Le contexte d'usage (classe / CUFR) prend de l'importance car il est un facteur explicatif des intentions d'utilisation ou des non intentions d'utilisation. Les éléments ergonomiques liés à la spécificité des tablettes tactiles sont ici précisés et mis en lien avec le reste du contexte technologique disponible, qui freine, ou participe à, l'intention d'utilisation.

Au CUFR, lorsque les personnes considèrent que la tablette n'est pas utile dans le cadre de leurs enseignements au CUFR, les réponses sont relatives au fait que la formation universitaire implique essentiellement une prise de notes. Or, les répondants précisent que l'ordinateur (39%) ou la prise de note manuelle sur papier (21%) est plus efficace, rapide, aisée, et permet de mieux mémoriser. Lorsque les personnes considèrent que la tablette n'est pas facile d'usage, parmi les 10 réponses apportées, 7 sont relatives à l'infrastructure (« connexion non stable », « pas d'endroit de chargement des tablettes », « surface pas adaptée », « pas de prise USB pour recharger les tablettes ») et 3 renvoient à l'objet technique en lui-même par manque de connaissance technique (« taper sur la tablette sans clavier n'est pas facile »). D'ailleurs, les éléments explicatifs des craintes de pratique de la tablette au CUFR renvoient à l'insécurité matérielle (vol de la tablette, peur de perdre les données puisqu'il n'y a pas d'enregistrement automatique sur un cloud en l'absence de wifi).

Lorsque les personnes considèrent que la tablette est utile dans le cadre de leurs enseignements au CUFR, 31 % des réponses des PES sont relatives au stockage des cours, 26 % des réponses des PES sont relatives à des aspects ergonomiques (plus léger, performant, transportable), 10 % des réponses des PES sont relatives à la synchronisation possible des enseignements (« accéder aux cours tout le temps »), mais aussi à l'usage comme multimédia (« faire des photos des cours, enregistrer les cours

sous format audio, rechercher des informations sur le web ») ou encore aux aspects écologiques (« *on utilise moins de papier* »). Lorsque les personnes considèrent que la tablette est facile d'usage, parmi les 3 réponses apportées, toutes sont relatives au fait que cela permettrait d'enregistrer vocalement les enseignements.

En classe, les tablettes présentent des intérêts pédagogiques diversifiés (différenciation pédagogique, suivi individualisé, multiplicité des ressources existantes et spécifiques disciplinaires et gestion de classe), longuement mentionné par les PES enquêtés. La diversité des réponses et la quantité de réponses sur les fonctions pédagogiques expliquent probablement le fort taux de réponses quant à l'utilité perçue des tablettes en classe. Les appréhensions de l'écosystème technologiques apparaissent moins dans les réponses (présence d'une seule occurrence pour le TBI et le vidéo-projecteur), mais les spécificités multimédias de l'outil (présenter des vidéos, du son, des images sur un même support) sont imaginées comme média intéressant aux apprentissages.

Les réponses des PES enquêtés sont ainsi de nature d'usage divers : les PES répondant considèrent que la tablette peut être utile en formation initiale sur des aspects techniques (multimédia, ergonomique, stockage des données) et inutiles sur des aspects pédagogiques et infrastructurels (prise de notes, apprentissage / accès à internet). La mise en valeur de l'absence de facilité est relative aux aspects infrastructurels (recharge des batteries, accès à internet).

7 Discussion

L'objectif de cette recherche exploratoire est de comprendre le poids des contextes dans l'intention d'utilisation des étudiant.e.s en formation de futur.e.s enseignant.e.s. La méthodologie a une visée réflexive : interroger les étudiant.e.s sur leur propre intention d'utilisation est un premier pas vers une attitude réflexive de la part des étudiant.e.s PES par rapport à leur pratique [10]. Les résultats de cette enquête ont été mis à disposition des PES, et les données issues de ces analyses sont intégrées dans le dispositif de formation, pour améliorer la formation par la recherche. 87 étudiant.e.s sont enquêtés à partir d'un questionnaire, disponible en ligne, qui a été proposé lors d'un enseignement en informatique dans le cadre de leur formation initiale. Cette étude reste exploratoire et les premières analyses permettent non seulement de décrire les intentions d'usage des tablettes par les étudiant.e.s enquêtés, mais aussi d'envisager des pistes prescriptives d'utilisation.

Un des premiers éléments importants de cette recherche est que l'outil tablette tactile n'est pas considéré de la même façon selon le contexte université *versus* école. Une grande partie des étudiant.e.s ne possède pas de tablettes tactiles et n'en fait pas usage dans un quotidien. Les tablettes tactiles sont ainsi globalement considérées comme peu utiles et peu faciles d'usage dans le contexte universitaire, contrairement à un usage scolaire, auprès des élèves. En effet, pour les PES enquêtés, les besoins d'outils numériques en formation sont essentiellement réduits à une prise de notes. La tablette ne permet pas d'y répondre aisément. Des usages liés à des problématiques techno-ergonomiques sont majoritairement mentionnés (prise de note, léger, stockage des données, ne plus utiliser du papier). Dans ce cadre, les tablettes tactiles sont directement en concurrence avec d'autres outils considérés comme plus performants par les personnes enquêtées, en évitant les freins susmentionnés. Seul l'usage multimédia

de la tablette (effectuer en même temps, une vidéo des enseignements, une prise de notes des enseignements, enregistrement audio) est considéré comme utile (mais là encore, en concurrence avec les outils déjà utilisés). Le rapport coût / bénéfice impliqué par l'usage des tablettes est trop important dans ce contexte. On sait que l'usage des tablettes passe par une technostucture forte [33] ce qui n'est pas le cas au CUFR, ce qui rend alors très difficile une pratique efficiente.

Toute la difficulté reste dans le transfert de compétences du métier d'étudiant au métier d'enseignant qui passe par un changement d'approche pédagogique universitaire et de dispositif de formation [12, 22]. Parmi les variables explicatives prises en compte dans cette recherche pour comprendre la formation de l'intention d'utilisation selon les contextes d'usage, il semble que la possession de matériel (tablette) favorise l'intention d'utilisation en classe auprès des élèves ou au CUFR. Les résultats tendent à montrer que cette variable agit comme facilitateur des liens entre les contextes d'intention d'utilisation. Toutefois, la possession d'une tablette tactile ne crée pas d'attentes en termes de formation, sur des applications particulières. Les étudiant.e.s conçoivent que les terminologies utilisées dans les enseignements (Web, internet, applications, logiciels, onglets, etc.) ne sont pas nécessairement connues ni comprises.

Ainsi, les tablettes sont considérées comme utiles en classe auprès des élèves, porteuses d'un idéal d'apprendre différemment [24] et peu utilisables, alors qu'elles sont considérées comme peu utiles et moyennement utilisables au CUFR. Bien que les analyses doivent être complétées pour aboutir à une modélisation, le modèle de l'acceptabilité [14, 30, 32], semble prendre une forme distincte à travers ces données. L'utilité perçue prendrait plus de poids dans l'acceptabilité que la facilité perçue, qui se confond avec le sentiment de compétences. Les contextes pondèrent différemment l'intention d'utilisation et les frontières entre les contextes [16] ralentiraient la genèse de l'acceptabilité.

En conclusion, ces éléments doivent être complétés pour aller dans le sens d'un meilleur accompagnement pédagogique universitaire, au plus proche des besoins des étudiant.e.s. Les espaces de frontières entre les identités estudiantines et professionnelles semblent difficilement communiquer, sans pour autant créer un espace « d'inter-métier » [29]. Soutenir les PES dans leur capacité méta-reflexive quant à leur pratique de la tablette en situation didactique semble alors être une phase essentielle de leur construction identitaire et de maîtrise des TICE. Les formateurs ont un rôle essentiels dans un rapport d'imitation (de modelage) qui les dépassent [27]. Un des risques fort quant à l'usage de la tablette tactile est de considérer que cet outil « miracle » va aider les élèves en difficultés puisqu'il va les ré-inscrire dans une dynamique d'apprentissage. Pour les élèves les plus fragiles, la dimension relationnelle et psychologisante est mise sur le devant des explications des enseignant.e.s, au détriment de l'analyse didactique et d'une prise en compte du collectif dans l'apprentissage.

References

1. Assoumani, N.: Les enseignant.e.s mahorais face aux tablettes numériques (Mémoire de Master), Centre Universitaire de Formation et de Recherche, CUFR de Mayotte, France (2018).
2. Attenoukon, S. A., Karsenti, T., Lepage, M.: L'apprentissage avec des supports mobiles dans l'enseignement supérieur au Bénin : analyse des usages des apprenantes. *Revue inter-*

- Intention d'utilisation des tablettes en formation initiale d'enseignants à Mayotte
- nationale des technologies en pédagogie universitaire / *International Journal of Technologies in Higher Education* 12(3),62–74 (2015).
3. Betrancourt, M.: L'ergonomie des TICE : quelles recherches pour quels usages sur le terrain ? In: Charlier, B., Peraya, D. (eds). *Regards croisés sur la recherche en technologie de l'éducation* (pp. 77-89), De Boeck, Bruxelles (2007).
 4. Béziat, J.: Former aux TICE : entre compétences techniques et modèles pédagogiques. *Revue internationale des technologies en pédagogie universitaire / International Journal of Technologies in Higher Education* 9(1-2), 53-62 (2012). <https://doi.org/10.7202/1012902ar>
 5. Béziat, J., Villemonteix, F.: Les TICE au quotidien : suffit-il d'en faire ? Le cas de l'école primaire en France. *Education & Formation* 304(2), 41-52 (2016).
 6. Bobillier Chaumon, M.E., Oprea Ciobanu, R.: Les nouvelles technologies au service des personnes âgées : entre promesses et interrogations – Une revue de questions. *Psychologie française* 54, 271-285 (2009).
 7. Boulton, H., Hramiak, A.: Cascading the use of Web2.0 technology in secondary schools in the United Kingdom: Identifying the barriers beyond pre-service training. *Technology, Pedagogy and Education* 23(2), 151-165 (2014). doi:10.1080/1475939X.2013.802994
 8. Bourdon, P., Lefèr Sauvage, G., Mercier, C., Teutsch, P., Lopez-Cazaux, S.: Le rôle de l'imitation dans l'appropriation des outils numériques chez les enfants avec autisme: Étude de l'usage des tablettes tactiles en ULIS – école TED et IME. *Enfance* 1(1), 147-168 (2018). doi:10.3917/enf2.181.0147.
 9. Caneva, C.: Facteurs d'adoption ou de rejet des technologies chez les professeurs-formateurs d'enseignants : le cas de l'université costaricienne. *Revue internationale des technologies en pédagogie universitaire* 16(1), 15–35 (2019). <https://www.erudit.org/fr/revues/ritpu/2019-v16-n1-ritpu05061/1066592ar.pdf>
 10. Caneva, C. et Akkari, A.: Polyvalence des usages des TIC par les futurs enseignants au Costa Rica. *Revue internationale des technologies en pédagogie universitaire* 3(2), 16-33 (2018).
 11. Carugati, F., Tomasetto, C: Le corps enseignant face aux technologies de l'information et de la communication : un défi incontournable. *Revue des sciences de l'éducation* 28 (2), 305–324 (2002). <https://doi.org/10.7202/007356ar>
 12. Charlier, B., Daele, A., Deschryver, N: Vers une approche intégrée des technologies de l'information et de la communication dans les pratiques d'enseignement. *Revue des sciences de l'éducation* 28(2), 345–365 (2002). <https://doi.org/10.7202/007358ar>
 13. Cnesco, Blanchard, A., Botton,H., Miletto,V., Caro, P.: *Panorama des Inégalités scolaires d'origine territoriale en France*. Paris, CNESCO (2018).
 14. Davis, F.: Perceived Usefulness, Perceived Ease of Use, and User Acceptation of Information Technology. *MIS Quarterly* 32(4), 319-340 (1989).
 15. 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. Thèse de doctorat, Université de Rennes 2, Rennes (2011).
 16. Fluckiger, C.: L'école à l'épreuve de la culture numérique des élèves. *Revue française de pédagogie* n°163, 51-61 (2008).
 17. Guillemet, P.: L'industrialisation de la formation, la fin d'un paradigme ? *Distances et savoirs*, 2(1), 93-118 (2004). doi:10.3166/ds.2.93-118.
 18. Hamon, D., Villemonteix, F.: Le rapport des élèves et des enseignants aux tablettes numériques à l'école primaire: vers une évolution de la forme scolaire ? *Distances et médiations des savoirs*, 3(11) (2015). <https://journals.openedition.org/dms/1143>
 19. Jouët, J.: Retour critique sur la sociologie des usages. *Réseaux*, 18, 100, 487-521 (2000).

20. Lefer Sauvage, G., Bachelot, S.: Étude exploratoire de l'intention d'utilisation des tablettes tactiles par des enseignants.e.s d'ULIS-école à Mayotte. In: Priolet, M. (eds). *L'école à Mayotte*. Paris, l'Harmattan (2021, à paraître).
21. Karsenti, T., Fievez, A.: L'iPad à l'école: usages, avantages et défis : résultats d'une enquête auprès de 6057 élèves et 302 enseignants du Québec (Canada). Montréal, QC : CRIFPE (2013).
22. Karsenti, T., Peraya, D., Viens, J.: Conclusion : bilan et perspectives de la recherche sur la formation des maîtres à l'intégration pédagogique des TIC. *Revue des sciences de l'éducation* 28(2), 459-470 (2002). <https://doi.org/10.7202/007363ar>
23. Kerneis, J., Delamotte, E., Priolet, M., Cheik-Ahamed, A.K.: Pratiques numériques des étudiant.e.s mahorais.e.s : approches contextuelles entre insularités et mobilités. In: Bonfils, P., Dumas, P., Massou, L. (eds.). *Pédagogie et numérique : l'enseignement supérieur au défi de la mondialisation ?* (Actes de la 11ème édition du colloque international TICE-MED), Questions de communication, Série Actes, 40 (2020, à paraître).
24. Moeglin, P.: Ardoises numériques, changement de paradigme ? Communication orale retranscrite dans le cadre du colloque Ecritech'3 (2012). Consulté le 20-01-2020 : http://www.ecriture-technologie.com/wp-content/uploads/2012/09/ecritech3_avril2012_transcription_P_MOEGLIN.pdf
25. Nogry, S., Sort, C.: Le temps de l'appropriation d'une classe mobile par les enseignants à l'école primaire. *Distances et médiations des savoirs*, 16 (2016). <https://doi.org/10.4000/dms.1655>
26. Peraya, D.: Technologies, innovation et niveaux de changement : les technologies peuvent-elles modifier la forme universitaire ?. *Distances et médiations des savoirs* [En ligne], 21 | 2018, mis en ligne le 10 mars 2018, consulté le 21 mars 2019 (2018). URL : <http://journals.openedition.org/dms/2111>.
27. Peraya, D., Lombard, F., Betrancourt, M.: De la culture du paradoxe à la cohérence pédagogique. Bilan de 10 années de formation à l'intégration des TICE pour les futur-e-s enseignants du primaire à Genève. *Formation et pratiques d'enseignement en questions* 7, 11-28 (2008).
28. Rabardel, P.: Qu'est-ce qu'un instrument ? Outils pour le calcul et le traçage de courbes, CNDP-DIE, p.61-65 (1995). http://tecfa-labs.unige.ch/mitic/articles/rabardel_1995_quest-ce_quun_instrument.pdf
29. Thomazet, S., Merini, C.: L'école inclusive comme objet frontière, *La nouvelle revue de l'adaptation et de la scolarisation*, vol. 70-71, no. 2-3, pp. 137-148 (2015).
30. Tricot, A., Plégat-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. *Sticf*, 391-402 (2003). <https://edutice.archives-ouvertes.fr/edutice-00000154/document>
31. Venkatesh, V.: Determinants of perceived ease of use : integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information system research* 11(4), 342-365 (2000).
32. Venkatesh, V., Davis, F. D.: A Model of the Antecedents of Perceived Ease of Use: Development and Test. *Decision Sciences* 27, 451-481 (1996). <http://dx.doi.org/10.1111/j.1540-5915.1996.tb01822.x>
33. Villemonteix, F., Hamon, D., Nogry, S., Séjourné, A., Hubert, B., Gélis, J.-M.: *Extate : Expérience tablettes tactiles à l'école primaire* (p. 86). Université Cergy-Pontoise: Ministère de l'Éducation Nationale (2014). <https://hal.archives-ouvertes.fr/hal-01026077v2/document>
34. Villemonteix, F., Nogry, S.: Usages de tablettes à l'école primaire : quelles contraintes sur l'activité pédagogique ? *Recherche & formation* 1(81), 79-92 (2016).

Expérimentation à grande échelle d'applications pour tablettes pour favoriser l'apprentissage de la lecture et de l'anglais oral

Sonia Mandin¹, Ahmed Zaher¹, Svetlana Meyer², Mathieu Loiseau³, Gérard Bailly⁴,
Coralie Payre-Ficout⁵, Julien Diard¹, Fluence-Group⁶, Sylviane Valdois¹

¹ CNRS, Université Grenoble-Alpes, Univ. Savoie Mont Blanc, LPNC, UMR 5105, F-38000
Grenoble, France
{sonia.mandin, ahmed.zaher, julien.diard, sylviane.valdois}@univ-
grenoble-alpes.fr

² Université Grenoble-Alpes, Univ. Savoie Mont Blanc, LPNC, UMR 5105, F-38000 Grenoble,
France
svtln.meyer@gmail.com

³ Univ. Lyon, INSA-Lyon, CNRS, LIRIS, UMR5205, F-69621, France
mathieu.loiseau@liris.cnrs.fr

⁴ GIPSA-Lab, Université Grenoble-Alpes, CNRS UMR 5216, 11 rue des mathématiques,
38400 Grenoble, France
gerard.bailly@univ-grenoble-alpes.fr

⁵ Univ. Grenoble Alpes, LIDILEM, F-38000 Grenoble, France
coralie.payre-ficout@univ-grenoble-alpes.fr

⁶ Anne Blavot, Marie-Line Bosse, Yaël Briswalter, Nathalie Chalon, Erika Godde, Sandrine
Ingremeau, Marie-Pierre Jouannaud, Christine Lequette, Emilie Magnat, Monica Masperi,
Anne-Laure Piat-Marchand, Adrea Soh, Marc Zaroni

Résumé. Le projet FLUENCE (e-Fran, PIA2) a permis la réalisation de trois applications mobiles visant à améliorer les performances des élèves en lecture (EVASION et ELARGIR) et en compréhension orale de l'anglais (LUCIOLE). Les applications ont été déployées dans les classes de l'académie de Grenoble (France) auprès de 722 élèves suivis du CP au CE2. L'entraînement d'une durée totale de 10h s'étalait sur 10 semaines. Chaque année, des groupes d'élèves appariés utilisaient l'une des deux applications lecture ou l'application d'anglais. Les résultats montrent que s'entraîner avec EVASION en CP puis ELARGIR en CE1 améliore la vitesse de lecture en fin de CE1. L'entraînement avec LUCIOLE conduit à améliorer la compréhension de l'anglais oral des élèves tant en CP qu'en CE1.

Mots-clés : jeu, attention visuelle, prosodie, lecture, anglais

Abstract. The purpose of the FLUENCE Project (e-Fran, PIA2) was to develop three original tablet computer games aiming to improve reading skills (EVASION and ELARGIR) and oral comprehension in English as a second language (LUCIOLE). The effectiveness of the applications was checked on a large

sample of pupils from the academy of Grenoble (France) through a three-year follow-up study. The 722 participants were followed from Grade 1 to Grade 3 (ages 6-8). They were trained during 10 hours using either one of the two reading applications or the English application. The results show that reading skills at the end of Grade 2 improved for the group that received both EVASION training in Grade 1 and ELARGIR training in Grade 2. After one year of training (in either Grade 1 or Grade 2), the students using LUCIOLE showed better oral English comprehension than their peers.

Keywords: Tablet game training, visual attention, prosody, reading, English

1 Introduction

Cet article présente les recherches menées dans le cadre du projet FLUENCE, un des 22 programmes du volet e-FRAN financé par le PIA2 (Programme d'Investissement d'Avenir). Le projet regroupe les équipes de trois laboratoires de recherche de l'Université Grenoble-Alpes (le LPNC-UMR5105, Le GIPSA-Lab-UMR5216 et le LIDILEM-EA609) et celles des Rectorats des Académies de Grenoble et de Mayotte. FLUENCE a pour objectif la conception et la validation d'outils numériques favorisant les apprentissages. Deux apprentissages sont ciblés : la lecture et l'anglais.

Trois applications numériques¹ ont été conçues : « EVASION » pour développer l'empan visuo-attentionnel (VA), une compétence de base de l'apprentissage de la lecture, par le biais de 4 mini-jeux d'action ; « ELARGIR » pour entraîner la fluence en lecture via la lecture répétée assistée de textes ; « LUCIOLE » pour améliorer la compréhension orale en anglais par le biais d'exercices ludiques. Les trois applications relèvent de cadres théoriques différents qui ont conduit à une expérimentation commune dans laquelle des entraînements intensifs sont réalisés en classe afin d'améliorer les apprentissages visés et des évaluations sont effectuées chaque année afin d'évaluer l'efficacité des entraînements proposés.

La problématique générale et les cadres théoriques pour les deux applications de lecture (EVASION et ELARGIR) et pour l'application d'anglais (LUCIOLE) sont décrits en amont de la description des applications elles-mêmes. Le protocole et les résultats présentés se limitent aux expérimentations menées dans l'académie de Grenoble sur les deux premières années de suivi.

2 Problématique

Le projet Fluence s'inscrit dans les problématiques d'efficacité de l'enseignement et de différenciation pédagogique, ici appliquées à l'apprentissage de la lecture et de l'anglais, langue étrangère. Il vise à contribuer à l'efficacité de l'enseignement en proposant des outils pédagogiques fondés sur les avancées de la recherche et en appliquant une méthodologie expérimentale rigoureuse de façon à démontrer dans quelle mesure l'utilisation en classe de ces outils a un réel impact sur les apprentissages visés. Les outils

¹ Voir site web et vidéos de présentation : <https://fluence.cnrs.fr>

sont proposés sur support numérique selon une progression qui s'adapte aux besoins de chaque élève pour permettre un apprentissage différencié. En lecture, nous partons du constat que la plupart des entraînements, proposés en classe, visent l'apprentissage des relations graphème-phonème et l'entraînement de la conscience phonémique. Ces apprentissages, bien qu'absolument nécessaires à l'apprentissage de la lecture, ne sont pas suffisants [9, 19]. Ils permettent un décodage réussi des mots écrits mais ont peu d'effets sur la fluence en lecture. Le terme « fluence » recouvre ici à la fois une lecture efficace (rapide et sans erreur) et une prosodie adaptée. La fluence en lecture dépend à la fois de la capacité à reconnaître rapidement les mots écrits et du traitement des dimensions textuelles : prosodie, groupes syntaxiques et intégration sémantique [14]. Les enseignants sont relativement démunis quant aux moyens d'améliorer la fluence en lecture et ne disposent notamment pas d'outils pédagogiques permettant un apprentissage différencié dont l'efficacité aurait été démontrée.

De la même façon, concernant l'apprentissage des langues, la plupart des recherches insistent sur la nécessité d'entraîner la compréhension orale en préalable de toute exigence de production et sur l'importance d'exposer les enfants à des énoncés réalistes. Si les enseignants de début de primaire disposent de nombreuses ressources pédagogiques, notamment pour l'apprentissage de l'anglais, leur formation initiale généraliste couplée à une certaine insécurité linguistique rend leur exploitation difficile [5]. L'utilisation en classe de logiciels axés sur l'écoute d'énoncés produits par des locuteurs natifs et l'utilisation de ces énoncés en contexte ludique pourrait répondre à leurs besoins si cela s'avère efficace pour favoriser l'apprentissage de l'anglais.

3 Cadres théoriques

3.1 Reconnaissance rapide des mots et empan visuo-attentionnel (VA)

L'apprentissage de la lecture se caractérise par le passage d'une lecture laborieuse quasiment lettre-à-lettre en début d'apprentissage à la reconnaissance immédiate du mot entier par le lecteur expert. Le traitement efficace des mots écrits dépend des ressources attentionnelles disponibles lors du traitement [15]. On peut estimer les ressources VA déployées lors du traitement de séquences de lettres à travers des épreuves mesurant l'empan VA. Des capacités élevées d'empan VA en grande section de maternelle sont prédictives de la qualité et de la vitesse de lecture en fin de CP [18]. Un empan VA plus élevé en élémentaire ou au secondaire prédit une lecture plus rapide [7, 20]. Le lien entre empan VA et vitesse de lecture est retrouvé chez l'adulte et un déficit de l'empan VA est associé à des troubles dyslexiques [6]. L'ensemble de ces données témoigne d'un lien très fort entre capacités visuo-attentionnelles et apprentissage de la lecture, suggérant notamment qu'un entraînement de l'attention visuelle en début de scolarité pourrait faciliter l'apprentissage de la lecture. Par ailleurs, on sait que la pratique encadrée des jeux vidéo d'action améliore les capacités attentionnelles en général [4] et l'empan VA en particulier [1]. Un entraînement précoce de l'attention visuelle basé sur les principes des jeux vidéo d'action est à l'origine du développement de l'application EVASION décrite en section 4.1.

3.2 Importance de la Fluence et lecture répétée assistée de textes

Dans l'apprentissage de la lecture, l'accent est essentiellement mis sur le décodage et la précision. Pourtant un élève qui lit sans erreur n'est pas forcément un bon lecteur. Plusieurs auteurs soulignent l'importance de la prosodie en lecture [14], et donc, la prise en compte des dimensions de phrasé et d'expressivité [12].

La lecture et l'écoute simultanée sont un bon moyen d'améliorer la fluidité de la lecture à voix haute. Dans cette méthode, l'apprenant écoute quelqu'un d'autre lire un texte tout en suivant des yeux le texte lu. Puis, il essaie d'imiter le lecteur référent qu'il vient d'entendre. Cette méthode combinée à la lecture répétée des mêmes textes améliore rapidement la vitesse et la précision de la lecture. Plusieurs études montrent également une amélioration de la fluence grâce à la méthode de karaoké [11]. Utilisant les effets d'entraînement implicite de la lecture simultanée, une version améliorée par stimulation audiovisuelle, de type karaoké, a été proposée [2].

L'ensemble de ces données indique que l'utilisation intensive d'un karaoké de lecture avec surlignage simultané d'unités linguistiques de plus en plus larges (syllabes, mots, groupes de mots) pourrait favoriser la mise en place du phrasé et de l'expressivité. C'est cette hypothèse qui a présidé au développement de l'application ELARGIR (cf. section 4.2).

3.3 Enseignement de l'anglais langue étrangère en primaire

Les Instructions Officielles font du cycle 2 le point de départ de l'apprentissage des langues vivantes pour tous les élèves [17]. Elles mettent en exergue le savoir-être central à l'apprentissage d'une langue étrangère : « curiosité, écoute, attention, mémorisation, confiance en soi » (*ibid.*, p. 29), mais aussi « l'ouverture à d'autres sons, d'autres mots [...] un état d'esprit fait de curiosité et d'audace [...] accepter de ne pas (tout) comprendre et de s'exprimer imparfaitement » (*ibid.*, p. 32).

Cette conception de l'apprentissage des langues n'affranchit pas des objectifs plus linguistiques (*ibid.*, p. 34). Cependant, à l'heure actuelle, les acquis d'un apprentissage des langues à l'école élémentaire ne semblent pas attestés [16]. Ce constat, assorti à l'insécurité linguistique de beaucoup d'enseignants conduit à s'interroger sur l'efficacité des outils dont disposent les enseignants pour améliorer la compréhension orale en anglais en proposant des activités explicites d'écoute portant sur des éléments linguistiques spécifiques de la langue anglaise. Une exposition importante à l'anglais non-seulement à travers les activités, mais aussi à travers la narration devrait favoriser l'apprentissage. Ces objectifs ont conduit au développement de l'application LUCIOLE (cf. section 4.2).

4 Description des applications

4.1 EVASION

Le logiciel EVASION a pour but d'entraîner l'attention visuelle mobilisée pour l'identification des mots en lecture. Il s'agit notamment d'entraîner les enfants à déployer rapidement leur attention visuelle sur des séquences de lettres et à mobiliser progressivement davantage de ressources attentionnelles afin de traiter des séquences de plus en plus longues et complexes. Le logiciel propose quatre mini-jeux qui reprennent les principes des jeux vidéo d'action [1, 4] et reposent sur la manipulation de caractéristiques, telles que la longueur ou la complexité des séquences, qui ont permis d'améliorer l'empan VA en contexte dyslexique [22]. La difficulté des jeux dépend de plusieurs critères : la longueur des séquences à traiter, la rapidité de défilement des séquences et leur temps de traitement, le nombre de cibles simultanément présentées à l'écran et la similarité visuelle entre cibles et distracteurs. Un algorithme intégré au jeu permet une progression personnalisée. Sur la base des réponses aux essais précédents, l'algorithme estime la probabilité que l'enfant réussisse chaque combinaison de ces différents critères et sélectionne une combinaison que l'enfant a 75 % de chance de réussir. Une probabilité de 75 % laisse une marge de progression suffisante tout en évitant de mettre trop souvent l'enfant en échec, ce qui est censé optimiser l'apprentissage.

4.2 ELARGIR

Le logiciel ELARGIR a pour but d'entraîner la fluence sur des textes entiers, dont le contenu et la stratégie de surlignage sont adaptés à la performance du lecteur. Les lectures sont effectuées de manière répétée, à voix haute et majoritairement en Karaoké. Lors de 30 sessions, l'enfant est exposé à 10 textes, qu'il lit plusieurs fois à haute voix simultanément avec un adulte (dont la voix a été préenregistrée avec 4 styles de phrasé et d'expressivité) puis de manière autonome pour finalement écouter et évaluer les lectures d'autres camarades. Ces lectures pour soi, ensemble et pour l'autre permettent à la prosodie de jouer son rôle de pivot entre décodage et compréhension [21].

Afin d'accompagner la progression, les textes de chaque parcours (CE1, CE2, 6^{ème} et 5^{ème}) sont ordonnés en difficulté croissante et le niveau de difficulté (longueur et complexité du texte, style de lecture adulte et politique de mise en surbrillance) est adapté en fonction de l'évaluation de la performance effectuée par un camarade ou par l'enseignant (à terme de manière automatique, [3]). De plus, ELARGIR offre la possibilité de ralentir/accélérer la lecture adulte de manière à faciliter la synchronisation.

4.3 LUCIOLE

LUCIOLE² est un jeu sérieux intégrant une métaphore ludique (enfants agents secrets qui doivent retrouver des animaux kidnappés). Le jeu propose aux élèves une

² LUdique au service de la Compréhension Orale en Langue Étrangère

progression alternant narration et tâches réalistes pour un apprenant des cycles 2 et 3. Les mini-jeux font uniquement intervenir des stimuli audio et graphiques. L'apprenant y répond par la sélection d'items graphiques (par toucher ou glisser-déposer).

Certaines modalités ciblent la mémorisation du lexique (*cf.* objectifs des IO, *op. cit.*) en prenant en compte les recommandations du Cnesco [13]. Les autres activités ciblent différentes compétences (lexique et culture, phonologie, grammaire, stratégies de compréhension orale).

En effet, chaque mini-jeu est justifié par la narration et intègre des énoncés de locuteurs anglophones natifs, qui conservent leurs accents régionaux y compris pour les énoncés en français (satisfaisant ainsi la recommandation des IO d'accepter de ne pas tout comprendre et de commettre des erreurs). En effet, la plupart des personnages ne parlent que l'anglais, mais le mentor du joueur s'exprime en français quand il doit donner des instructions non-ambiguës, fournir des éclairages culturels ou linguistiques pour susciter la curiosité à l'égard de la culture cible. Le jeu redonne ainsi une place à l'enseignement explicite des langues en proposant des entraînements systématiques et répétés sur le lexique, la grammaire et la phonologie et propose des éléments de réflexion sur le fonctionnement de la langue étrangère et de la langue française [13].

5 Méthode

5.1 Participants

Pour évaluer l'efficacité des applications, les entraînements ont été proposés en classe à des élèves de l'académie de Grenoble suivis du CP au CE2. 37 écoles volontaires ont participé correspondant à un échantillon de 722 élèves de CP représentatif de la population des 5 départements de l'académie. Les critères retenus étaient : appartenance à une zone prioritaire ou pas (REP/REP+), à une zone rurale vs urbaine, nombre d'élèves dans l'école et dans le département.

5.2 Protocole longitudinal

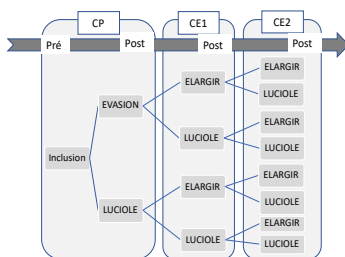


Fig. 1. Protocole de suivi longitudinal sur trois ans (CP-CE1-CE2). Une évaluation initiale est proposée à l'inclusion (Pré = pré-entraînement) en début de CP puis des évaluations (Post = post-entraînement) sont proposées à chaque fin d'année.

Le protocole de suivi longitudinal sur trois ans est schématisé dans la Figure 1. L'évaluation initiale (Pré-CP) a permis de définir deux groupes d'élèves appariés (*cf.* section 5.3), dont l'un sera entraîné sur EVASION et l'autre sur LUCIOLE. Tous les élèves d'une même classe utilisent le même dispositif d'entraînement au cours d'une même année scolaire. En CE1 et CE2, chaque groupe est scindé en deux groupes équivalents utilisant ELARGIR ou LUCIOLE. Donc, chaque année, des groupes d'élèves appariés, utilisent soit un des logiciels de lecture (EVASION ou ELARGIR) soit le logiciel LUCIOLE. Chaque groupe sert de contrôle à l'autre. LUCIOLE a notamment été conçu de façon à ne pas inclure d'information écrite et à mobiliser minimalement l'attention visuelle, qui sont au cœur des entraînements ELARGIR et EVASION. Ainsi, un niveau de lecture plus élevé chez les élèves des groupes EVASION ou ELARGIR comparativement aux élèves du groupe LUCIOLE, permettra de conclure à l'efficacité des premiers pour entraîner la lecture, indépendamment des aspects ludiques ou motivationnels induits par le contexte d'entraînement. Le protocole permet ainsi d'évaluer spécifiquement l'effet de chacun des entraînements. Il permet également d'évaluer si certains parcours (successions d'entraînements) conduisent à de meilleurs apprentissages.

5.3 Entraînements et évaluations

Les entraînements étaient proposés à raison de trois sessions de 20 minutes par semaine entre janvier et mai, pour un total de 10 h. Les enseignants organisaient les sessions par petits groupes. Une interface web a été développée pour permettre à l'enseignant de vérifier la progression de chaque élève au cours de l'entraînement. Il leur était conseillé de se connecter à l'interface au moins une fois par semaine afin de vérifier que les temps de jeu étaient respectés et, pour ELARGIR, qu'il n'y ait pas besoin de nuancer la progression dans le parcours de lecture suite à l'évaluation par les pairs.

Afin de vérifier l'efficacité de chacun des entraînements, une évaluation des aptitudes cognitives des élèves a été proposée avant et après chaque période d'entraînement. Les tests en début de CP évaluaient le niveau de performance de chaque élève (score et temps) sur les dimensions prédictives de l'apprentissage de la lecture (conscience phonémique, connaissance du nom et du son des lettres, langage oral, dénomination rapide, mémoire verbale, QI, capacités d'empan VA) ainsi que les connaissances précoces en lecture et la familiarité avec la langue anglaise. Les post-tests réévaluaient la plupart de ses compétences ainsi que les niveaux de lecture et d'orthographe.

Les évaluations comportaient des épreuves individuelles et des épreuves collectives se répartissant sur plusieurs sessions. La partie individuelle (45 min à 1 h) était administrée par un vacataire en face à face avec l'élève. Les tests collectifs (1 h à 2 h) étaient dispensés par les enseignants et/ou des vacataires. Dans la mesure du possible, les tests individuels et collectifs étaient réalisés sur deux jours différents de sorte à ne pas engendrer trop de fatigue chez les élèves. Si tel n'était pas le cas, ils étaient au moins proposés sur deux demi-journées.

Le protocole a reçu l'agrément du Comité d'Ethique de l'université Grenoble-Alpes (CERNI : IRB00010290-2017-07-04-22) et de la CNIL (déclaration n° 0979176). Les parents ont été informés des objectifs et du déroulé de l'étude. Seuls les élèves ayant une autorisation parentale dûment signée ont pu participer.

5.4 Formation des enseignants

Les enseignants ont tous suivi une formation d'une ou deux journées en amont des tests et des entraînements. La formation comprenait un volet théorique portant sur les mécanismes de base de l'apprentissage de la lecture (notamment sur le rôle de l'Empan VA), sur l'intérêt d'un entraînement à la fluence en lecture et sur l'importance de développer des aptitudes de compréhension orale en début d'apprentissage de l'anglais langue seconde. La formation était complétée d'un volet appliqué où le protocole expérimental était présenté ainsi que les trois applications. La formation comprenait de nombreuses périodes d'échanges et des travaux pratiques (utilisation des logiciels).

La logistique des formations a été facilitée par leur prise en compte par le Rectorat de Grenoble dans le plan de formation académique ce qui a permis de libérer les enseignants, assurer leur remplacement et le remboursement de leurs frais de déplacement. Outre les enseignants directement concernés par l'étude, d'autres professionnels de l'Education Nationale en ont également bénéficié (IEN et conseillers pédagogiques).

5.5 Hypothèses

Les hypothèses opérationnelles sont les suivantes :

Hypothèse 1. En fin de CP, les élèves du groupe EVASION devraient présenter de meilleurs scores aux tests d'empan VA et de lecture que les élèves du groupe LUCIOLE (H1a). Ces derniers devraient avoir de meilleurs scores au test d'anglais que les élèves du groupe EVASION (H1b).

Hypothèse 2. En fin de CE1, les élèves du groupe ELARGIR devraient avoir un meilleur niveau de lecture que les élèves du groupe LUCIOLE (H2a). De meilleurs scores en anglais sont attendus dans le groupe LUCIOLE (H2b).

Hypothèse 3. Les enfants entraînés deux années de suite sur des logiciels d'aide à l'apprentissage de la lecture (parcours EVASION-ELARGIR) devraient être meilleurs lecteurs en fin de CE1 que chacun des autres groupes (H3a). Par ailleurs, le parcours LUCIOLE-LUCIOLE devrait conduire à des scores plus élevés en anglais (H3b).

6 Résultats

La comparaison pré-post entraînement entre le début et la fin du CP montre des progrès significatifs en compréhension de l'anglais oral suite à l'utilisation de LUCIOLE. Comme illustré dans la Figure 2a, les élèves du groupe LUCIOLE ont de meilleurs scores en anglais en fin de CP que les élèves du groupe EVASION [$m_{\text{Luciole}} = 0.56 (0.17)$ vs $m_{\text{Evasion}} = 0.40 (0.17)$; $t_{(676)} = -10.59$, $p < .0001$]. Un résultat similaire est obtenu lorsqu'on compare les scores des élèves entre la fin CP et la fin CE1. Les élèves entraînés avec LUCIOLE progressent davantage en anglais que leurs pairs entraînés avec ELARGIR [$m_{\text{Luciole}} = 0.51 (0.19)$ vs $m_{\text{Elargir}} = 0.39 (0.17)$; $t_{(244)} = -5.43$, $p < .001$]. Les hypothèses H1b et H2b sont donc vérifiées. Par contre, les enfants qui ont utilisé EVASION n'ont pas un empan VA plus élevé et EVASION-CP pas plus qu'ELARGIR-CE1 ne conduisent à de meilleurs scores en lecture suite à

l'entraînement. Ces résultats ne permettent donc pas de valider les hypothèses H1a et H2a sur les bénéfices des logiciels de lecture, lorsqu'ils sont utilisés indépendamment.

L'analyse des parcours met en évidence deux effets supplémentaires. Concernant le logiciel d'anglais, les scores ne sont pas meilleurs chez les élèves entraînés sur LUCIOLE deux années consécutives que chez ceux qui n'ont été entraînés qu'en CP ou seulement en CE1 (H3b).

Par contre, un effet positif et significatif en lecture est démontré suite au parcours EVASION-ELARGIR (123 élèves) comparativement aux trois autres parcours (H3a). La Figure 2b illustre la comparaison avec le parcours LUCIOLE-LUCIOLE (128 élèves). Une ANOVA incluant la période de test en facteur intra-sujets et le parcours en facteur inter-sujets montre un effet d'interaction significatif [$F(1, 2885) = 7.20, p < .008$]. Alors qu'ils avaient des scores similaires en début et fin de CP, les deux groupes se distinguent significativement en fin de CE1 ; le nombre de mots lus par minute est alors significativement plus élevé pour le parcours EVASION-ELARGIR [$m_{EVA-ELA} = 38.79 (24.30)$ vs $m_{LUC-LUC} = 32.16 (17.30)$; $t(249) = 2.221, p = 0.027$].

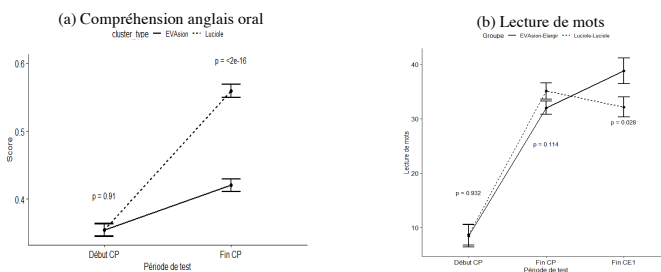


Fig. 2. (a) Évolution des scores en compréhension orale de l'anglais entre le début et la fin de CP pour les groupes EVASION et LUCIOLE ; (b) Évolution des scores en nombre de mots lus par minute entre le pré-test CP et le post-test CE1 selon le parcours d'entraînement (Luciole-Luciole versus Evasion-Elargir).

De plus, l'effet du parcours sur la lecture est modulé par le temps de jeu. Lorsqu'on oppose un groupe d'élèves dont l'entraînement effectif est plus intensif (temps de jeu sur EVASION > 6 h et Lecture des 10 textes ELARGIR) à un groupe d'élèves ayant un temps d'entraînement plus limité, l'effet du parcours n'est significatif que dans le premier groupe [$F(1,324) = 6.89, p = 0.009$ vs $F(1,236) = 3.20, p = 0.07$].

Les traces d'activité enregistrées lors des entraînements ont été analysées afin de mieux comprendre pourquoi les effets plus massifs attendus n'ont pas été obtenus. L'analyse montre notamment que le temps de jeu n'a le plus souvent pas été respecté. Pour le logiciel EVASION, le temps de jeu effectif est en moyenne de 6 h 13 min (E.T. = 2 h 15 min) au lieu des 10 h recommandées, les temps d'entraînement étant très variables d'un élève à l'autre (cf. Fig. 3a). Le problème de temps d'entraînement se pose également pour le logiciel ELARGIR (cf. Fig. 3b) puisqu'il est en moyenne de 3 h 53 min (ET = 2 h 36 min) avec, là encore, une forte variabilité inter-individuelle.

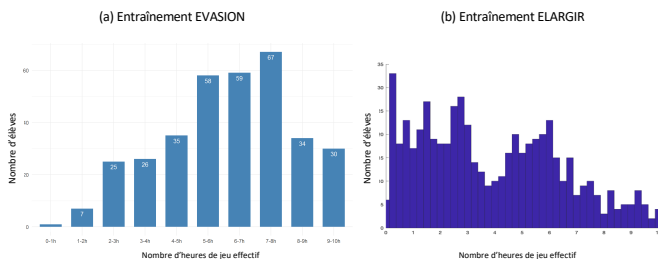


Fig. 3. Temps de jeu moyen des groupes EVASION (a) et ELARGIR (b).

D'autre part, les résultats de l'analyse des traces issues d'EVASION suggèrent un problème de calibration de l'échelle de difficulté des exercices. Environ 40 % des enfants ($N = 135$) obtiennent un taux de réussite d'au moins 80 % pour les exercices a priori conçus pour être les plus difficiles. L'entraînement ne leur offrait donc pas la marge de progression souhaitée. Ce constat nous a amené à développer une version révisée du logiciel qui a été utilisée avec succès auprès des élèves de l'académie de Mayotte, l'année suivante. L'analyse des enregistrements lors de l'utilisation d'ELARGIR a montré que le niveau de bruit ambiant était le plus souvent trop élevé pour permettre une bonne concentration des élèves sur les exercices proposés.

7 Discussion

Le projet FLUENCE s'inscrit dans la catégorie des expérimentations écologiques avec passage à l'échelle. A ce titre, notre objectif était de valider les logiciels d'entraînements en condition réelles d'utilisation, soit en classe sous la supervision des enseignants. L'étude a porté sur un large échantillon d'élèves de l'académie de Grenoble et sur un grand nombre de classes. Les aptitudes cognitives des élèves étaient mesurées avant et après chaque phase d'entraînement pour évaluer les progrès comparativement à un groupe contrôle actif qui utilisait un entraînement ne ciblant pas les mêmes aptitudes. Le suivi longitudinal sur trois ans devait permettre d'évaluer l'effet de chaque entraînement et des parcours à court et moyen terme sur les apprentissages ciblés.

Les résultats montrent que l'entraînement LUCIOLE est efficace en CP. Un effet positif de LUCIOLE est également objectivé en fin de CE1 mais uniquement chez les

élèves qui n'avaient pas été entraînés avec LUCIOLE l'année précédente. L'absence d'effet cumulatif lors de deux années successives d'entraînement avec LUCIOLE est contraire à nos hypothèses mais pourrait s'expliquer par des problèmes de logistique (certains élèves de CE1 ayant repris les exercices de CP).

Aucune amélioration de l'empan VA n'est observée en CP suite à l'entraînement avec EVASION, ce qui excluait d'observer un effet positif en lecture. Des temps d'entraînement souvent trop courts et des problèmes de calibration du logiciel pourraient contribuer à expliquer cette absence d'effet.

De la même façon, les enfants entraînés sur ELARGIR en CE1 ne lisent pas plus efficacement en fin d'année que leurs homologues entraînés avec LUCIOLE. En fait, à peine la moitié des utilisateurs s'est entraînée sur la totalité des textes, témoignant là encore d'un non-respect des temps recommandés.

Nous avons également comparé les différents parcours sur deux années consécutives : le parcours EVASION-ELARGIR correspondant à deux années d'entraînement à la lecture se solde par une amélioration de la fluence en fin de CE1, comparativement aux trois autres parcours. Il y a donc un effet positif cumulé des deux entraînements alors qu'aucun effet significatif n'était démontré pour chacun des entraînements indépendamment. Cet effet peut être très directement interprété comme la conséquence de l'entraînement puisque seuls les enfants qui ont le mieux respecté les temps d'entraînement recommandés améliorent significativement leur temps de lecture.

Deux conclusions s'imposent. D'une part, la mise en évidence d'effets significatifs lorsque l'entraînement est suffisamment intensif est très encourageante. LUCIOLE, ELARGIR et EVASION devraient constituer à terme des outils d'accompagnement pédagogiques efficaces pour faciliter l'apprentissage de la lecture et de la compréhension de l'anglais parlé. D'autre part, l'étude soulève la difficulté de coupler les exigences d'une méthodologie scientifique rigoureuse aux contraintes de terrain (calendrier scolaire, gestion de la classe, environnement bruité, couverture numérique...). Un effort particulier doit être mené pour sensibiliser les enseignants et les parents d'élèves aux nécessaires contraintes de la recherche et accroître la motivation des élèves.

Remerciements

Le projet FLUENCE (2017-2022) est une opération soutenue par l'État dans le cadre du volet e-FRAN du Programme d'Investissement d'Avenir (PIA2) opéré par la Caisse des Dépôts et Consignations (CDC). Nous remercions tout particulièrement les enseignants et les élèves pour leur engagement et leur participation à cette étude.

References

1. Antzaka, A., Lallier, M., Meyer, S., Diard, J., Carreiras, M., Valdois, S.: Enhancing reading performance through action video games: the role of visual attention span. *Scientific Reports* 7, 14563 (2017)
2. Bailly, G., Barbour, W.: Synchronous Reading: Learning French Orthography by Audiovisual Training. In: *Interspeech*, pp. 1153-1156. (2011)

3. Bailly, G., Godde, E., Piat-Marchand, A.-L., Bosse, M.-L.: Predicting multidimensional subjective ratings of children' readings from the speech signals for the automatic assessment of fluency. International Conference on Language Resources and Evaluation (LREC), Marseille, France (2020)
4. Bavelier, D., Green, C.S.: Enhancing attentional control: Lessons from action video games. *Neuron Review* 104, 147-163 (2019)
5. Behra, S.: Quelles sont les compétences nécessaires pour enseigner les langues à l'école élémentaires ? Conférence de consensus du Cnesco. De la découverte à l'appropriation des langues vivantes étrangères : comment mieux accompagner les élèves ?, Lyon (2019)
6. Bosse, M.L., Tainturier, M.J., Valdois, S.: Developmental dyslexia: the visual attention span deficit hypothesis. *Cognition* 104, (2007)
7. Bosse, M.L., Valdois, S.: Influence of the visual attention span on child reading performance: a cross-sectional study. *Journal of Research in Reading* 32, 230-253 (2009)
8. Candelier, M.: L'éveil aux langues à l'école primaire: Evlang : bilan d'une innovation européenne. . De Boeck Supérieur, Louvain-la-Neuve, Belgique (2003)
9. Castles, A., Rastle, K., Nation, K.: Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest* 19, 5-51 (2018)
10. Genot, S.: L'enseignement des langues vivantes à l'école élémentaire : éléments d'évaluation des effets au collège. *Revue française de pédagogie* 118, 27-42 (1997)
11. Gerbier, E., Bailly, G., Bosse, M.L.: Audio-visual synchronization in reading while listening to texts: Effects on visual behavior and verbal learning. *Computer Speech & Language* 47, 74-92 (2018)
12. Godde, E., Bosse, M.-L., Bailly, G.: Échelle Multi-Dimensionnelle de Fluence : nouvel outil d'évaluation de la fluence en lecture prenant en compte la prosodie, étalonné du CE1 à la 5ème. *L'année Psychologique / Trends in Cognitive Psychology* (2020)
13. Gruson, B.: Recommandations du jury. Conférence de consensus CNESCO. De la découverte à l'appropriation des langues vivantes étrangères : Comment l'école peut-elle mieux accompagner les élèves ?, (2019)
14. Kuhn, M.R., Schwanenflugel, P.J., Meisinger, E.B.: Aligning theory and assessment of reading fluency: automaticity, prosody and definitions of fluency. *Reading Research Quarterly* 45, 230-251 (2010)
15. Laberge, D., Samuels, S.J.: Toward a theory of automatic information processing in reading. *Cognitive Psychology* 3, 293-323 (1974)
16. Manoïlov, P.: Les acquis des élèves en langues vivantes étrangères. (2019)
17. MEN: Programmes d'enseignement de l'école élémentaire et du collège (Ministère de l'Éducation Nationale). vol. 11, pp. 386 (2015)
18. Valdois, S., Roulin, J.L., Bosse, M.L.: Visual attention modulates reading acquisition. *Vision Research* 165, 152-161 (2019)
19. van Bon, W., Bokseveld, L.M., Font Freide, T., van den Hurk, A.: A comparison of three methods of reading-while-listening. *Journal of Learning Disabilities* 24, 471-476 (1991)
20. Van den Boer, M., de Jong, P.F., Haentjens-van Meeteren, M.M.: Modeling the Length Effect: Specifying the Relation with Visual and Phonological Correlates of Reading. *Scientific Studies of Reading* 17, 243-256 (2013)
21. Veenendaal, N.J., Groen, M.A., Verhoeven, L.: Bidirectional relations between text reading prosody and reading comprehension in the upper primary school grades: A longitudinal perspective. *Scientific Studies of Reading* 20, 189-202 (2016)
22. Zoubinetsky, R., Collet, G.M., N'Guyen-Morel, M.A., Valdois, S., Serniclaes, W.: Remediation of allophonic perception and visual attention span in developmental dyslexia: a joint assay. *Frontiers in Psychology* 10, 1502 (2019)

Comment guider les doctorants dans l'utilisation du Design-Based Research ?

Nadine Mandran¹, Mathieu Vermeulen² et Estelle Prior²

¹ Grenoble Alpes Université, CNRS, France

² IMT Lille Douai, Institut Mines-Télécom, Univ. Lille, Centre for Digital Systems,
F-59000 Lille, France

`Nadine.Mandran@univ-grenoble-alpes.fr`

Résumé. Face aux problèmes rencontrés par les doctorants pour mener des recherches dans les EIAH, la méthode THEDRE a été développée pour les aider. Cet article présente cette méthode et ses guides pour mettre en œuvre le Design-Based Research qui est de plus en plus utilisé dans les recherches en TEL. Les auteurs proposent aussi un principe pour suivre le travail doctoral par des indicateurs. La méthode et les guides ont été évalués avec des doctorants et des encadrants lors de sessions de formation en 2019 et 2020.

Mots-clés : *Design-Based Research*, doctorant, outils pour la recherche, informatique centrée humain

Abstract. Faced with the problems encountered by PhD students in conducting research in TEL, the THEDRE method has been developed to support them. This article presents this method and its guides for implementing Design-Based Research which is increasingly used in TEL research. The authors also propose a principle to monitoring doctoral work through indicators. The method and the guides were evaluated with PhD students and supervisors during training sessions in 2019 and 2020.

Keywords: Design-Based Research, PhD Student, tools for Research, Human-Centered Computer Science

1 Introduction

Aujourd'hui, le monde de la recherche en informatique est en pleine interrogation face aux problématiques de reproductibilité [3]. Le problème est encore plus criant pour la Recherche en Informatique Centrée Humain (RICH) qui nécessite d'impliquer l'humain, par exemple, pour modéliser et outiller une activité humaine tel l'enseignement [15]. La RICH se déroulant très souvent en contexte réel (e.g. en salle de classe), la reproductibilité des résultats est difficile voire

impossible à atteindre. Il est cependant essentiel de pouvoir rendre compte de la qualité du travail de terrain, ce dernier contribuant à la création de connaissances scientifiques valables. Pour cela, il est nécessaire de disposer de cadres de travail et d'outils qui explicitent comment la recherche est conduite afin de décrire les expérimentations et les données produites et ainsi tracer ces recherches. Dans le cadre de la recherche en Environnement Informatique pour l'Apprentissage Humain (EIAH), une des méthodes de recherche le plus fréquemment citée est le *Design-Based Research* (DBR) [24]. Cette méthode pose des principes et des propriétés, mais reste sujette à de nombreuses interprétations car elle n'est pas outillée et les processus de mise en œuvre ne sont pas décrits. Des chercheurs du domaine font souvent part de cette situation telle cette enseignante-chercheuse lors d'une conférence en 2019 : "on fait tous du DBR car l'on mixe les méthodes de production de données, on travaille in situ, on fait des itérations, on intègre tous les utilisateurs mais en dehors de ça ?". De notre point de vue, cette méthode est importante car elle s'inscrit dans une approche systémique et pragmatique qui sont les conditions de réalisation de certaines recherches en EIAH (e.g. faire de la recherche dans un contexte réel avec des acteurs du terrain) [22]. Ce problème est d'autant plus difficile pour des doctorants, ces derniers n'ayant pas toujours les compétences pour intégrer l'humain dans la conduite de leur recherche [14]. De fait, ils rencontrent souvent des difficultés pour concevoir et faire des expérimentations afin de valider leur contribution scientifique [15]. La première difficulté concerne la production de leur travail, ils identifient mal ce que sera leur production finale. La seconde concerne la faible formation des informaticiens aux méthodes de production et d'analyse des données en SHS pour intégrer l'humain dans la conception et l'évaluation de manière rigoureuse. Par rapport aux difficultés des doctorants dans la mise oeuvre du DBR, nous nous posons alors les questions suivantes : (1) Comment outiller le DBR pour faciliter le travail de recherche des doctorants et favoriser la collaboration avec les encadrants ? (2) Comment tracer un processus de recherche dans le DBR et ainsi en assurer la qualité ? Dans le domaine de la recherche en Système d'Information (SI), un modèle de méthode de conduite de la recherche nommé THEDRE (*Traceable Human Experiment Design Research*) offre un langage pour concevoir des méthodes de conduite de la recherche en fonction des différents contextes de travail des chercheurs RICH [16]. THEDRE se positionne dans les sciences de l'artificiel [23,1] et dans le paradigme épistémologique du constructivisme pragmatique [2]. Il s'agit d'un paradigme épistémologique pour créer ou améliorer des connaissances scientifiques dans un contexte systémique qui mobilisent les situations réelles avec les acteurs du terrain. Il autorise l'utilisation de multiples sources de données pour construire et évaluer les connaissances scientifiques. Concrètement, avec THEDRE chaque doctorant peut créer sa propre méthode de conduite de la recherche traçable en relation avec son encadrant et en l'adaptant à son contexte. Un ensemble de guides est proposé pour accompagner les doctorants, leur permettre de tracer leur activité et d'échanger avec leurs encadrants. Par son côté adaptable, ce modèle peut être utilisé dans le cadre de recherche en EIAH conduite avec le DBR. Cet article vise de fait à étudier l'adap-

tabilité de THEDRE au travers de l'analyse de trois guides utilisés et évalués lors de formations avec des doctorants et des chercheurs (au total 69 personnes de 2019 à 2020).

La première section de cet article propose un état de l'art en présentant les principes du DBR, comment il a été outillé et les avantages et limites de cette méthode. La seconde présente la méthode THEDRE et les trois guides sur lesquels nous focalisons cet article. Nous présentons ensuite notre méthode d'évaluation et les résultats. Enfin, nous concluons sur les avantages et les limites de ces guides pour le DBR et nous ferons une ouverture sur la mise en oeuvre des principes du DBR.

Avant de commencer cet article, nous souhaitons apporter un élément sur la terminologie employée dans ce texte à propos des termes de méthodologie et de méthode qui sont souvent confondus : "*Using methodology as a synonym for method or set of methods leads to confusion and misinterpretation and undermines the proper analysis that should go into designing research*" [10]; "*Methodology is the systematic, theoretical analysis of the methods applied to a field of study. It comprises the theoretical analysis of the body of methods and principles associated with a branch of knowledge. Typically, it encompasses concepts such as paradigm, theoretical model, phases and quantitative or qualitative techniques*" [8]; "*Method is "an integrated collection of procedures, techniques, product descriptions, and tools, for effective, efficient, and consistent support of the engineering process*" [11]. Dans cet article, nous utilisons le terme *méthode* et nous le qualifions avec d'autres noms : méthode de processus de recherche, méthode de production de données, méthode d'analyse de données.

2 État de l'art

Pour aider le travail de recherche en EIAH, le DBR est une méthode de conduite de la recherche envisageable [24] [6]. Il encadre des recherches qui produisent des outils utilisables par l'humain, mis en oeuvre dans un contexte réel. Cet ancrage dans la réalité du terrain lui permet d'être aujourd'hui de plus en plus souvent utilisé [22]. Cependant, cette méthode est employée de manière différente selon les auteurs. Ce point est positif car c'est une méthode qui s'adapte à différents contextes de recherche, mais cela pose un problème sur la comparaison des recherches conduites avec cette méthode. McKenney et Reeves [18], dans leur ouvrage *Conducting Educational Design Research* proposent de découper le processus du DBR et des outils pour guider cette méthode. Ils proposent aussi un schéma pour le guider [21]. Dans cette partie nous présenterons les fondements et principes du DBR ainsi que les propositions de McKenney et Reeves.

La méthode de conduite de la recherche DBR trouve son origine dans les écrits de Collins [7]. Ce dernier, dans les travaux du Center for Technology in Education (New York), fait une synthèse des recherches technologiques en lien avec l'éducation et il propose de développer une méthode de conduite des expérimentations en classes avec des élèves pour étudier si de nouvelles technologies peuvent être intéressantes pour l'apprentissage. D'après le collectif pour le dé-

veloppement de la méthode DBR, "*Design-based Research can help create and extend knowledge about developing, enacting and sustaining innovative learning environments*".

Wang et Hannafin en 2005 [24] sur la base d'expérimentations réalisées dans le cadre des EIAH ont proposé neuf principes pour mener des recherches dans ce domaine. Ces principes concernent l'organisation de la recherche ; l'implication des acteurs et la prise en compte du terrain ; la production et l'analyse des données ; la documentation ; et la généralisation. Le tableau 1 présente la liste de ces principes. Au delà de ces principes, des propositions ont été faites pour structurer un processus de conduite de la recherche. Comme évoqué précédemment, Reeves [21] propose un schéma régulièrement repris dans la littérature. Mc Kenney et Reeves [18] proposent une structuration du DBR et un guidage par des outils, qui prend ses origines dans un projet mené pendant quatre ans pour étudier l'impact d'un ordinateur pour soutenir l'enseignement des sciences et des mathématiques [17], le programme CASCADE-SEA. Ce projet a été documenté pour "*carefully documenting the iterative process of analysis, prototype design, evaluation, and revision*". L'objet de la recherche comportait des volets techniques et humains, "*we sought insight into the characteristics of a valid and practical computer-based tool that possesses the potential to affect the performance of its users*". Ce projet a été aussi l'occasion de s'interroger sur le processus de recherche : "*while they also learn from the development process. Further, this research has contributed to the articulation of design principles and related developmental research methods*". McKenney et Van Den Akker [17] ont également proposé plusieurs guides facilitateurs pour la collaboration et la traçabilité mais ces guides proposés ne couvrent pas tout le processus de recherche.

Les avantages du DBR concernent les neuf principes pour justifier d'une recherche de terrain itérative menée en collaboration avec les acteurs. Ces principes sont intéressants car ils posent les jalons d'une épistémologie de la recherche ancrée dans le constructivisme. Néanmoins les critères de valeur et de validité pour poser un cadre épistémologique [13] ne sont pas mentionnés. Pour répondre à ce problème, des descriptions de processus de conduite de la recherche et des outils pour documenter le processus ont été proposés [18]. Mais de notre point de vue, ces processus restent schématiques et n'offrent pas de langage de modélisation. Les outils proposés ne couvrent pas tout le processus de recherche, restent proches de la gestion de projet et ne sont pas forcément formateurs pour les doctorants. Par ailleurs, si le DBR pose des principes généraux facilitateurs pour mener une recherche de terrain, certains de ces principes restent difficiles à mettre en œuvre. De plus, il n'existe pas de processus adaptable pouvant tracer un travail conduit avec le DBR, plus particulièrement pour les jeunes chercheurs. Enfin, si le principe itératif du DBR est pertinent pour faire progresser de manière interdépendante la connaissance et l'outil malgré la proposition de Collins et ses collègues [6], le DBR ne préconise pas d'utiliser d'indicateurs pour suivre ces cycles itératifs. Pour répondre à ces manques nous proposons de décliner la méthode de conduite de la recherche THEDRE (*Traceable Human Experiment*)

Design REsearch) et ses guides pour aider les doctorants à conduire une recherche selon les principes du DBR.

3 THEDRE pour outiller le DBR

Notre contribution porte sur la proposition d'un nouveau principe pour le DBR et d'un ensemble de guides pour conduire une recherche en EIAH inscrite dans le DBR et fournir des repères aux doctorants. Dans un premier temps, nous présentons succinctement la méthode THEDRE. Nous décrivons ensuite trois de ces guides en précisant la façon à laquelle ils répondent aux principes du DBR.

Comme évoqué précédemment, THEDRE est un modèle de méthode pour conduire, guider et tracer la RICH [15,16]. Elle a été construite pour faciliter le travail de recherche des doctorants et plus particulièrement pour les aider à concevoir des expérimentations avec des humains. Cette méthode a aussi pour objectif d'améliorer la communication entre doctorants et encadrants. D'un point de vue théorique, elle s'ancre dans les sciences de l'artificiel [23] dans lesquelles une recherche produit de la connaissance et un outil de manière interdépendante et en étudiant un terrain. D'un point de vue épistémologique, elle est basée sur le paradigme du constructivisme pragmatique [2]. Techniquement, elle s'inscrit dans une démarche qualité dite "d'amélioration continue" : le processus de recherche est structuré selon les quatre actions du cycle de DEMING *Plan-Do-Check-Act* [9] : il s'agit d'une méthode itérative. Elle inclut aussi la démarche centrée utilisateur [14] (*norme iso 9241-420*), qui permet d'intégrer l'utilisateur aux différents moments de la conception d'un outil numérique ou non. Comme les recherches impliquant l'humain en contexte réel sont difficilement reproductibles, THEDRE propose de tracer l'activité de recherche. En revanche, tout tracer, tout documenter pose un problème de profusion de données. Afin de minimiser ces problèmes, des guides sont proposés de manière à documenter l'essentiel du travail, seuls les éléments saillants et utiles au travail de recherche sont conservés. Ils facilitent la rédaction de certaines parties de la thèse (i.e. introduction, état de l'art, etc.) ou des articles de recherche. THEDRE étant un modèle de méthode, l'ensemble de ces guides sont adaptables à différents domaines ou contextes, ce qui permet de les proposer pour réaliser une recherche conduite dans le cadre du DBR. Ces guides sont accessibles en ligne sur <https://thedre.imag.fr/>.

Avant de présenter les apports de THEDRE pour le DBR, nous proposons d'ajouter un nouveau principe au DBR qui est "de poser des indicateurs d'objectifs pour piloter le processus" (P10 dans le tableau 1). Les indicateurs d'objectifs proposés par Collins [7] n'ont pas été relayés dans la description du DBR. Or identifier des indicateurs permet de se fixer des objectifs à atteindre et de prendre des décisions à certains moments du processus. Ils sont nécessaires pour piloter un cycle itératif afin de savoir si un nouveau cycle doit être lancé ou pas. Ce principe est un élément supplémentaire pour organiser le travail de recherche et le piloter. Le tableau 1 présente ces dix principes, organisés par catégories d'activités.

TABEAU 1. 5 catégories pour organiser les 10 principes du DBR.

Catégories d'activités	Principes du DBR
Organiser la recherche	P1 : Identifier les ressources pertinentes pour les besoins du projet en utilisant la littérature et les différents exemples de conception P2 : Fixer des objectifs pratiques pour développer la théorie et proposer un plan initial P3 : Faire évoluer régulièrement la contribution scientifique et les outils associés P10 : Poser des indicateurs d'objectifs pour piloter le processus
Impliquer des acteurs et la prise en compte du terrain	P4 : Conduire la recherche dans des conditions représentatives du monde réel P5 : Inclure les acteurs du terrain comme collaborateurs de la recherche
Produire et analyser des données	P6 : Mettre en oeuvre des méthodes de production des données de façon systématique et ciblée P7 : Analyser les données rapidement après la production
Documenter	P8 : Documenter le processus de recherche, les objectifs, la production et l'analyse des données et les résultats pour rendre compte du travail réalisé
Optimiser la conception locale versus généraliser	P9 : Optimiser la conception locale sans perdre de vue la généralisation

De manière globale, le principe de documentation du DBR (P8) est couvert par l'ensemble des guides qui visent à accompagner le doctorant, à garder une trace de son travail, de l'évolution de ses propositions et à documenter son travail de manière incrémentale. Par son ancrage dans la démarche d'amélioration continue, THEDRE est par nature cyclique. En effet, les quatre stades du cycle de DEMING (*Plan, Do, Check, Act*) rythment les itérations. A chacune des itérations, la contribution scientifique, les outils et les questions de recherche peuvent être revus. Il s'agit du principe itératif (P3) du DBR. Aujourd'hui, près de quinze guides sont disponibles dans THEDRE ; ils couvrent tout le processus de travail de thèse, de l'accueil du doctorant à la rédaction d'un article scientifique. Les trois guides que nous étudions sont propres à THEDRE, nous n'avons pas identifié ce type de guide par ailleurs (ils sont disponibles sur (cf. <https://thedre.imag.fr>). Ils sont importants car ils sont préconisés pour initier un travail de doctorat et poser des indicateurs.

TABEAU 2. Synthèse des trois guides et principes du DBR.

Principes	Organisation				Acteurs et terrain		Données		Documentation	Locale / Généralisation
	1	2	3	10	4	5	6	7	8	9
Guide 1	✓	✓	✓		✓				✓	
Guide 2		✓		✓			✓		✓	✓
Guide 3	✓	✓	✓	✓	✓	✓			✓	✓

Le guide 1 permet de "Spécifier la contribution scientifique et les outils actionnables" qui sont produits pendant le travail de doctorat. (Les outils actionnables [15] correspondent aux outils créés par le chercheur sur la base de modèles théoriques ; l'utilisateur s'en servira dans une activité ; ils peuvent être en format numérique ou non ; ils peuvent être créés ou améliorés). En début de thèse, un doctorant a très souvent des difficultés à identifier ce qui correspond à une contribution scientifique et aux outils actionnables associés. Ce guide aide l'encadrant à clarifier ce qu'il attend du doctorant (P2). Il facilite la coopération entre encadrant et doctorant. Il permet en effet de décrire la contribution scientifique, les outils actionnables, l'état de ces outils et la nécessité ou non de mobiliser des utilisateurs (P4). Il permet de fixer des objectifs pratiques pour développer la théorie et proposer un plan initial (P2) et de préciser quels sont les acteurs du terrain impliqués pour faire quoi (P5). En outre, les informations renseignées dans ce guide sont revues et actualisées tout au long du travail de thèse, ainsi la contribution et l'outil évoluent au fil de la recherche (P3).

Le guide 2, "Identifier des indicateurs d'objectif et les mesures expérimentales" (P10), répond à la proposition de Collins [7] sur la mise en place d'indicateurs d'objectifs et au besoin de critères pour savoir si le processus doit itérer ou non. Ce guide est composé d'un ensemble de questions pour créer d'une part des indicateurs d'activité. Ces derniers rendent compte du travail effectué pendant la thèse. D'autre part, il permet de créer des indicateurs d'objectif. La création d'indicateurs d'objectifs dans une activité de recherche n'est pas habituelle, mais dans le cas d'un processus de recherche itératif, ils sont indispensables pour orienter les décisions, par exemple de communiquer les résultats s'ils sont suffisamment aboutis ou de refaire un cycle de recherche (P3). Le guide incite aussi le doctorant à approfondir la définition des mesures prises pendant les expérimentations et ainsi affiner le protocole expérimental et le documenter (P6 et P8).

Le guide 3 "Bien commencer un travail de thèse, construire sa problématique et rédiger son introduction" facilite l'élaboration et le suivi de la problématique et il contribue à la rédaction de l'introduction de thèse. Il est composé de questions réparties dans huit rubriques pour expliciter : la problématique, les travaux précédents, les travaux techniques, les impacts sociétaux et la dimension éthique, le contexte du terrain, les méthodes expérimentales, les indicateurs d'objectifs et la contribution.

Le tableau 2 présente les trois guides de la méthode THEDRE étudiés et les principes auxquels ils répondent.

4 Terrain d'évaluation

L'objectif des expérimentations est d'évaluer THEDRE et ses guides. Les expérimentations répondront aux questions suivantes : Les guides issus de la méthode THEDRE facilitent-ils l'organisation du travail de recherche des doctorants ? Les guides sont-ils utilisables par rapport à leur travail de thèse ? Pour répondre à ces questions, nous mesurons, (1) avant la formation, les attentes des participants,

(2) pendant la formation, l'utilisabilité des guides et le niveau de satisfaction des participants, (3) trois mois après la formation, nous mesurons l'appropriation de la méthode et les manques éventuels. Notre recherche est conduite selon les principes du DBR. Pour l'évaluation des guides, nous utilisons principalement une démarche qualitative. La démarche qualitative [8] permet de comprendre et d'identifier des phénomènes. Des formations et séminaires THEDRE se sont déroulés entre Juillet 2019 et Juin 2020 à Lille, Laval, Paris et Grenoble. Au total 57 doctorants et 46 encadrants y ont pris part. Notre méthode de production des données et une méthode qualitative, nous ne pouvons pas parler de « représentativité », terme lié à des démarches quantitatives par échantillonnage. Le tableau 3 détaille les différentes formations. Elles ont été conçues comme des focus-groups pour évaluer les guides après utilisation. Nous utilisons aussi des questionnaires pour mesurer les attentes a priori des doctorants et les scores de satisfaction et d'utilisabilité des guides [5,12]. Les résultats ont contribué à l'évolution des guides entre chacune des formations. Lors des séminaires, nous avons pu faire évoluer les guides avec les remarques des participants. Notre objectif est d'identifier les éléments positifs et négatifs et des points d'amélioration. Pour l'analyse qualitative, nous avons utilisé une analyse thématique. Les thèmes issus de la réponse du participant à une question ouverte ou de son discours sont annotés. Des phrases extraites du discours, les verbatims, sont choisies pour illustrer ces thèmes [20]. Les réponses au questionnaire d'utilisabilité sont traitées selon le procédé de Bangor et al. [4].

TABLEAU 3. Formations (F) et séminaires (S) à la méthode THEDRE avec les guides.

Type	F	F	S	S	F	F	F
Version	V1	V2	V2	V3	V3	V3	V4
Place	CAPTE, Lille	LIUM, Laval	LIP6, Paris	LIG, Grenoble	ED, Grenoble	ED, Lille	Atelier à distance
Date	2019/07	2019/11	2019/11	2020/01	2020/01	2020/02	2020/06
Durée	12 h	12 h	2 h	2 h	12 h	12 h	3 h
Doctorants		6	3	3	10	11	24
Encadrants	12	7	8	9			11
Nb de guides	14	12	3	2	12	12	4

5 Résultats

Cette section expose les résultats obtenus sur la base des données recueillies lors des formations de Laval (A), Lille (L), Grenoble (G) et des post-entretiens effectués un mois après (E). Les verbatim, extraits des discours des participants sont notés en italique ; un code composé de la lettre du lieu de formation et un numéro est attribué au participant (e.g. G7 participant 7 de Grenoble).

Le guide 1 "Spécifier la contribution scientifique et les outils associés" est, de prime abord, difficile à utiliser car il aborde le problème de la recherche par la

production ce qui n'est pas habituel. La dualité de cette production, connaissance et outil, accentue encore les difficultés. La note d'utilisabilité est globalement faible (score SUS n=4 "mauvais", n=7 "acceptable"). En revanche, il permet de travailler sur la notion de connaissances scientifiques et d'outils et de fixer des objectifs sur ce qui sera produit "permet de subdiviser la perception d'une thèse en éléments compréhensibles" (L11). "Je pense que ça a vraiment consolidé le fait que l'apport n'était pas l'outil technique, c'était pas le prototype (...) là le fait de les avoir mis sous forme de tableau avec cette colonne apport scientifique, ça a fait en sorte de bien éclairer le fait que les apports scientifiques c'est vraiment les méthodes et les modèles qu'on publie qui sont importants". Il aide au travail d'explicitation "en le remplissant, j'ai eu l'impression que ça m'aidait à expliciter les choses qui avaient peut-être pas été écrites sous cette forme" (E1); et permet de guider (n=1) "ce sont des choses auxquelles je ne faisais pas attention (...) la décomposition en composants actionnables (...) ça permet de savoir où on va" (E3). Comme ce guide aborde l'implication de l'utilisateur, pour construire et évaluer l'outil, le doctorant avec ce guide prend conscience de la nécessité de faire appel à des utilisateurs. "Je pense que c'était très centré utilisateur final de l'outil ou du composant (...) c'est toujours pertinent parce qu'on oublie vite ça et ça permet dans certain type de recherche où on a besoin de faire valider par l'utilisateur final ou de le faire participer à la conception, de pas l'oublier" (E1). Ce guide est un support pour le travail entre encadrant et doctorant "j'avais trouvé ça très bien et en plus c'est arrivé juste après une grosse réunion avec mon doctorant pour formaliser justement ces éléments-là. C'est arrivé comme une forme de rappel. Je l'ai laissé faire et il a remis exactement ce qu'il fallait. Je me suis dit c'est bon, il a pigé ce que je voulais dire." (E2).

Bien qu'ayant un meilleur score d'utilisabilité (Score SUS n=5 "bon" et n=3 "excellent"), le guide 2 pour identifier des indicateurs d'objectifs et les mesures expérimentales a été difficile à prendre en main par les doctorants. "C'est difficile mais c'est parce que la question soulevée est difficile" (G8). Imaginer des indicateurs d'objectifs n'est pas usuel "Déposer les métriques d'évaluation dès le départ c'est essentiel. C'est rarement fait en pratique" (E4). Cependant, ce travail apporte une aide pour s'inscrire dans une logique expérimentale, d'anticiper et de construire les expérimentations "Ca permet d'y réfléchir en amont, de simplifier et de rendre la suite plus propre. Ça permet de définir ce qu'on va évaluer et de penser le protocole en ayant ça en tête (...). La métrique permet de définir ce qu'on va mesurer et comment" (E4).

Le guide 3 pour construire la problématique obtient un score d'utilisabilité supérieur au score moyen de référence (71 versus 68) (Score SUS n=8 "excellent" et n=4 «bon»). Pour les participants, il répond aux objectifs prévus, à savoir une aide à la rédaction de l'introduction du manuscrit, en offrant un plan synthétique et complet. Il permet de "clarifier sa pensée" (G8), présente un aspect "synthétique et complet" (L11), et permet d'approfondir des aspects précis de la thèse : "Très intéressant de se questionner sur des aspects précis, cela permet de réfléchir différemment et de voir qu'il y a besoin d'approfondir certains points." (G5). De plus, ce guide facilite le travail avec l'encadrant "Pratique et

peut-être plus facile à présenter à son directeur de thèse pour un temps de travail commun." (G4). Cependant, ce guide est sans doute plus facile à utiliser après quelques mois de thèse. G9 précise qu'il est facile s'il est utilisé en fin de thèse : *"Un peu facile quand on est déjà en 3ème année, mais je pense que c'est très bien pour ceux qui commencent leur thèse", "si je l'avais eu en début de thèse cela aurait été plus simple"*. Ces trois guides issus de la méthode THEDRE apparaissent comme des facilitateurs pour le travail des doctorants, ainsi ils peuvent contribuer à mettre en application les principes du DBR.

6 Discussion et Conclusion

Nous proposons un premier outillage pertinent pour les doctorants et leurs encadrants pour faciliter la mise en oeuvre des principes du DBR. Les trois guides proposés ont parfois été de prime abord difficiles à utiliser. De fait, les questions posées dans ces guides pour favoriser les réflexions ne sont pas habituelles pour les doctorants. Elles relèvent des pratiques du métier de chercheur. Cependant, après un temps d'utilisation les doctorants les ont trouvés utiles, utilisables et structurantes. Pour améliorer le DBR, nous avons proposé un nouveau principe sur l'identification d'indicateurs d'objectifs ainsi qu'un guide pour les créer. Bien que perçu comme difficile, cette tâche s'est avérée utile auprès des doctorants et des encadrants. Ce guide les a aussi sensibilisés à la *"logique expérimentale"* indispensable à tout travail de recherche, ce qui constitue un des résultats positifs non attendu de nos expérimentations. Ce nouveau principe renforce le principe sur la planification du travail. Cependant, malgré la présence d'indicateurs, la gestion des itérations reste un point délicat. À quel moment faut-il itérer, sur quels objets et pourquoi ? Le détail des questions proposées dans les guides de THEDRE semble être une valeur ajoutée pour le DBR afin de définir plus précisément les différents objectifs conduits lors d'une thèse (e.g. contribution, outils, expérimentations). Cette proposition est certes une évidence et relève de l'implicite. Cependant, cette étape est essentielle pour des doctorants non familiarisés avec la recherche mais aussi pour la collaboration entre encadrants et doctorants : *"savoir ce qui sera construit et comment"*. Sur le contenu de ces guides, une comparaison avec les outils de Mc Kenney et Reeves [18] pourrait les compléter.

Au niveau des principes du DBR, deux d'entre eux semblent difficiles à respecter pour un travail de thèse. Le principe P5 "Inclure les acteurs du terrain comme collaborateurs de la recherche" constitue un premier écueil. En effet, mobiliser des acteurs de terrain est déjà une tâche délicate pour conduire des expérimentations en situation réelle. Pour le principe P5, l'implication des collaborateurs est encore plus forte car il leur est demandé de participer à l'élaboration de la question de recherche, des plate-formes et à l'évaluation. Le profil des collaborateurs peut être très différent. En outre, ces personnes ne sont pas toujours aguerries au travail de recherche. Il convient alors de les initier à ce travail. Dans des conditions de doctorat, ce principe ne peut être mis en place par le doctorant contraint sur trois ans. Cependant, au niveau d'un projet de recherche, cette collaboration devrait être outillée pour tracer l'évolution de la

question de recherche. Dans ces conditions, faut-il concevoir un guide qui accompagne ce principe ou impliquer un médiateur de type "broker" [19]? Le principe P9 "les chercheurs doivent optimiser la conception locale sans perdre de vue la généralisation" semble le plus difficile à mettre en oeuvre et à outiller [24]. Ce principe concerne la nécessité de faire le lien entre les résultats et le contexte dans lequel les données et les résultats ont été produits. Plusieurs contextes pourront être étudiés et cette multiplicité de contextes garantit une certaine généralité des résultats et augmente leur niveau de validité. Les produits d'une recherche conduite dans le cadre du DBR doivent donc être testés et évalués dans des contextes différents. De notre point de vue, un travail de thèse devrait être évalué dans un minimum de deux contextes différents et répétés dans le temps pour étudier de manière comparative les situations. Pour ce faire, le doctorant doit être informé dès le début de sa thèse des terrains à mobiliser. Cette étape du travail est difficile et chronophage mais elle fait partie intégrante d'un travail inscrit dans le DBR. Cette mobilisation de terrain apparaît donc comme une compétence à acquérir par les doctorants. Ils doivent être formés à la mise en place de protocoles expérimentaux et aux méthodes de production de données des SHS.

De notre point de vue, disposer d'une méthode comme le DBR est une valeur ajoutée pour la communauté EIAH, car les principes proposés correspondent à nos contextes de travail en situation réelle. Pour autant, cette méthode demande à être outillée pour les doctorants, ce qui constitue un des objectifs des guides proposés par la méthode THEDRE. Ils sont aujourd'hui de plus en plus diffusés et adaptés en fonction dans différents contextes de travail. Il reste à travailler au niveau des cycles itératifs. Au niveau de la recherche en EIAH, Il nous semble nécessaire de travailler sur l'implication des acteurs du terrain au sein de la recherche. De même, le principe P9 sur la généralité demanderait à être discuté en détail.

Remerciements

Cette contribution scientifique a été réalisée dans le cadre du projet de recherche APACHES, financé par l'i-SITE ULNE (FIPE18-007-VERMEULEN). Nous remercions les participants aux formations qui ont largement contribué à l'amélioration des guides.

Références

1. Avenier, M.J. : Les Sciences de l'artificiel : une conceptualisation révolutionnaire de sciences fondamentales à parachever. *Projectics / Proyéctica / Projectique* **24**(3), 43 (2019)
2. Avenier, M.J., Thomas, C. : Finding one's way around various methodological guidelines for doing rigorous case studies : A comparison of four epistemological frameworks. *Systèmes d'information & management* **20**(1), 61–98 (2015)
3. Baker, M. : 1,500 scientists lift the lid on reproducibility. *Nature* **533**(7604), 452–454 (2016)

4. Bangor, A., Kortum, P., Miller, J. : An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction* **24**(6), 574–574–594 (Aug 2008)
5. Brooke, J. : SUS-A quick and dirty usability scale. *Usability evaluation in industry* **189**(194), 4–7 (1996)
6. Collective, D.B.R. : Design-based research : An emerging paradigm for educational inquiry. *Educational Researcher* **32**(1), 5–8 (2003)
7. Collins, A. : Toward a Design Science of Education. In : Scanlon, E., O'Shea, T. (eds.) *New Directions in Educational Technology*, pp. 15–22. No. 96 in NATO ASI Series, Springer Berlin Heidelberg (1992)
8. Creswell, J.W. : *Research design : Qualitative, quantitative, and mixed methods approaches*. Sage publications (2013)
9. Deming, W.E. : *The new economics for industry, government, education*. MIT press (2018)
10. Frankfurter, G.M. : *Theory and reality in financial economics : essays toward a new political finance*. World Scientific (2007)
11. Harmsen, A., Apers, P.M.G., Brinkkemper, J., Faculty of Electrical Engineering, M.C.S. : *Situational Method Engineering*. Moret Ernst & Young Management Consultants Utrecht (Jan 1997)
12. Lallemand, C., Gronier, G., Koenig, V. : User experience : A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior* **43**, 35–48 (Feb 2015)
13. Le Moigne, J.L. : *Les Epistemologies Constructivistes*. Les Presses Universitaires de France (1995)
14. Lebis, A., Prior, E., Mandran, N., Karami, A., Vermeulen, M. : *Promouvoir et soutenir la Pédagogie Par Projet Centré Humain dans le supérieur : le projet APACHES*. In : DIDAPRO 8 - DIDASTIC. Lille, France (Feb 2020)
15. Mandran, N. : *Traceable human experiment design research : theoretical model and practical guide*, vol. 9. Wiley, iste edn. (2018), oCLC : 1021214724
16. Mandran, N., Dupuy-Chessa, S. : *THEDRE : A Traceable Process for High Quality in Human Centred Computer Science Research*. International Conference on Information Systems Development (ISD) (Sep 2017)
17. McKenney, S., Van den Akker, J. : Computer-based support for curriculum designers : A case of developmental research. *Educational technology research and development* **53**(2), 41–41–66 (Jan 2005)
18. McKenney, S., Reeves, T.C. : *Conducting educational design research*. Routledge (2018)
19. NIZET, Isabelle; MONOD ANSALDI, R. : Construction de bénéfices mutuels en contexte collaboratif : pistes théoriques et méthodologiques. *Phronesis* **6**(1-2), 140–152 (2017). <https://doi.org/https://doi.org/10.7202/1040224ar>
20. Paille, P., Mucchielli, A. : *L'analyse qualitative en sciences humaine et sociales*. Sciences Humaines et Sociales, Armand Colin, 2ème edn. (2011)
21. Reeves, T.C., Herrington, J., Oliver, R. : Design research : A socially responsible approach to instructional technology research in higher education. *Journal of Computing in Higher Education* **16**(2), 96 (2005)
22. Sanchez, E., Monod-Ansaldi, R. : Recherche collaborative orientée par la conception. *Education & didactique* **9**(2), 73–94 (2015)
23. Simon, H. : *The sciences of the artificial*. MIT Press, Cambridge (1962)
24. Wang, F., Hammafin, M. : Design-based research and technology-enhanced learning environments. *Ed. Tech. Research and Development* **53**(4), 5–23 (2005)

Stratégies pour de nouvelles formes d'appropriation des ENT. Impact de l'école à la maison dans le développement des usages et pratiques numériques des enseignants

Christine Michel, Laëtitia Pierrot

Université de Poitiers, TECHNE EA 6316, F-86073 Poitiers, France
christine.michel@univ-poitiers.fr, laetitia.pierrot@univ-poitiers.fr

Résumé. Le confinement du printemps 2020 a contraint brutalement les enseignants à réaliser leur activité à distance. L'objectif de cet article est d'analyser, à partir d'un questionnaire diffusé de mai à juin 2020, les usages et pratiques numériques mises en œuvre par les enseignants du primaire et du secondaire ainsi que les processus d'appropriation de l'ENT. Notre analyse confirme pour partie les résultats des recherches antérieures (usages majoritaires à des fins de communication et transmission de ressources), mais on observe une montée en puissance des services de notification et d'alerte ainsi que de l'auto-formation via Internet. Cela permet aux enseignants de mettre en œuvre des pratiques pédagogiques plus actives et engageantes, qu'ils commencent à réinvestir via l'ENT. Trois stratégies de reconception de l'ENT sont proposées pour favoriser ces nouvelles pratiques.

Mots-clés : Intégration du numérique, pratiques pédagogiques, ENT, modèle d'appropriation, usages numériques.

Abstract. The 2020 spring lockdown brutally forced teachers to carry out their activity remotely. This paper aims to analyze how teachers from French primary and secondary schools implemented digital uses and practices, based on a survey carried out from May to June 2020. We paid more specific attention to the level of VLE integration into the teaching practices. Our analysis partly confirms the results from previous research (major uses for communicating and sharing resources), but the uses of notification, alert services, and online self-training are stepping up. It has enabled teachers to implement more active and engaging teaching practices, which they are starting to reinvest through VLE. Finally, we proposed three VLE redesign strategies to promote these new practices.

Keywords: Technology Integration, Teaching Practices, VLE, Appropriation Model, Digital Uses.

1 Introduction

L'épisode de pandémie lié à la Covid-19 a provoqué la fermeture des écoles françaises en mars 2020. Dans ce contexte, et dans le but d'assurer une continuité pédagogique, les activités d'enseignement et d'apprentissage, initialement prévues pour se dérouler en présentiel, ont basculé à distance. Les recherches sur la mise en œuvre de la modalité à distance révèlent qu'elle repose sur des fondements pédagogiques similaires à celle en présence, mais le fait de passer par le biais d'outils numériques transforme la formation [1–3]. Du côté des élèves, la distance tend à renforcer les inégalités entre eux, principalement lorsque l'activité d'apprentissage repose sur une autonomie forte [1, 2]. Pour les enseignants, la distance peut contribuer à une redéfinition de leur rôle (d'enseignant à accompagnant ou animateur) et surtout une diversification des moyens pédagogiques et techniques [3]. Cela rejoint le constat fait par Félix *et al.* [4], d'un « bricolage héroïque » avec les moyens disponibles observé chez les enseignants.

La disponibilité effective des moyens techniques est donc l'un des premiers éléments expliquant l'intégration du numérique dans les pratiques pédagogiques des enseignants. Il n'est pas le seul : le rapport du numérique à l'École de la Cour des comptes [5] montre qu'en dépit d'un déploiement croissant des services et ressources numériques, les usages des enseignants qui en découlent restent limités. En particulier, l'utilisation qui est faite de l'environnement de travail numérique (ENT), solution institutionnelle regroupant des outils et services dans un espace numérique unique à l'échelle d'un établissement scolaire [6], a fait l'objet d'études qui montrent des usages restreints à quelques services (la messagerie, le cahier de textes) pour quelques finalités (communiquer ou partager des ressources). L'un des principaux avantages identifiés par les usagers de l'ENT est sa capacité à faire le lien entre les différents acteurs susceptibles d'accompagner les enfants dans leurs apprentissages : les enseignants et autres personnels éducatifs d'une part, les parents d'autre part [7].

Nous avons réalisé une enquête de mai à juin 2020 dans le but de décrire l'expérience de continuité pédagogique des enseignants, élèves et parents concernant l'usage du numérique en général et de l'ENT en particulier. Au-delà de contribuer aux études longitudinales sur l'intégration du numérique dans et pour l'éducation, notre objectif est d'identifier comment cette injonction à faire de la formation à distance a joué sur les dynamiques d'appropriation et comment ces observations peuvent aider à formaliser des stratégies d'accompagnement adaptées aux besoins des enseignants. Nous avons collaboré avec l'entreprise Open Digital Education¹, qui déploie deux solutions d'ENT sur l'ensemble du territoire français (One pour l'école primaire, Neo pour le collège ou le lycée) de manière à diffuser l'enquête sur une large population. Plus de 5000 réponses ont été collectées. Dans cet article, nous proposons une analyse des 441 réponses des enseignants du 1^{er} et 2^d degré.

¹ <https://opendigitaleducation.com/>

2 L'intégration du numérique dans les pratiques pédagogiques des enseignants

2.1 Des usages numériques limités

Si en dehors de la classe, pour préparer les cours notamment, l'utilisation du numérique par les enseignants est répandue, elle demeure limitée dans la salle de classe [8]. Les causes sont multiples : des manques ou dysfonctionnements matériels dans les établissements qui ne permettent pas aux enseignants de réaliser complètement les activités numériques imaginées [9], un manque de formations techniques, techno-pédagogiques et des illustrations d'usage limitées [8], des inégalités d'accès ou sociales [10, 11]. Pour autant, la méta-analyse conduite par Tricot et Chesné [12] rend compte d'un effet plutôt positif du numérique sur les apprentissages, selon la fonction pédagogique qu'il remplit. C'est le cas de l'évaluation des apprentissages ou des activités d'enseignement à distance par exemple.

Parmi les outils à disposition des enseignants, l'ENT est la solution intégrée pensée pour remplir ces fonctions. Si la solution regroupe une variété de services et ressources, les usages qui se sont développés sont circonscrits à certaines activités ou fonctions : la messagerie et le cahier de textes restent les services globalement les plus utilisés parce qu'ils permettent de reproduire par un autre moyen des pratiques traditionnelles déjà existantes [13]. Le sentiment de charge de travail supplémentaire, le manque d'utilisabilité de certains services comme le forum [14, 15] ou la vision négative qu'en ont les enseignants peuvent expliquer des usages peu développés [16]. Cette solution vient aussi en concurrence des solutions « bricolées », préférées par les enseignants les plus *passionnés* [16], car existantes avant l'arrivée de l'ENT [17]. De plus les stratégies de déploiement de l'ENT suivent la volonté ministérielle d'homogénéiser des services, au détriment d'une mise en correspondance des besoins des usagers et des fonctionnalités de l'outil [6, 18]. Pour autant, de nouvelles pratiques émergent [19] et l'ENT construit de nouvelles formes de valorisation du travail des enseignants [20]. L'ENT semble donc pouvoir contribuer à mettre en place des activités dans et en dehors de la classe [11].

2.2 L'observation des usages numériques et des pratiques pédagogiques

L'observation des usages numériques des enseignants sur l'ENT vise différents objectifs : produire des études descriptives des utilisations, identifier et expliquer les facteurs qui conditionnent les usages/non-usages, ou des effets liés aux usages, ou bien modéliser ou formaliser les dynamiques d'appropriation.

Les enquêtes par questionnaire sont représentatives, car à grande échelle, des tendances générales pour une population. En ce sens, elles sont utiles pour produire des études descriptives ou des modèles de l'appropriation. Elles se composent généralement d'une partie déclarative des caractéristiques sociodémographiques de l'individu, son parcours professionnel, son environnement de travail (en intégrant son équipement informatique) et d'une partie déclarative des usages réalisés concernant tel ou tel type

de service ou de technologie. Des critères complémentaires, identifiés dans les études plus globales de l'acceptation et l'appropriation du numérique, peuvent être ajoutés. L'analyse de l'utilité, l'utilisabilité et l'acceptabilité permet d'identifier tous les aspects « pratiques » de l'utilisation [21]. Plusieurs des études citées dans la section précédente [14–16] utilisent ces facteurs pour expliquer les usages/non-usages des enseignants. Le fait que de nombreux tests standardisés existent concernant l'utilité et l'utilisabilité [22] facilite leur analyse par des questionnaires. Les analyses opérées dans ces études, surtout dans une perspective de modélisation, sont globalement inspirées des recherches en sociologie des usages sur la diffusion des innovations cherchant à modéliser le niveau d'adoption [23].

Ces méthodologies restent questionnables sur au moins trois aspects pour lesquels nous proposons des adaptations. (1) Ces études ne considèrent souvent que les services les plus utilisés ou les services les plus courants, laissant de côté les usages minoritaires, mais en cours d'émergence [19] ou les usages spécifiques à un service innovant présent uniquement sur quelques ENT. De plus, cette approche quantifiée des usages tend à écarter les finalités d'usage ou les motivations, qui contribuent pourtant à comprendre les objectifs des enseignants et justifier le choix d'un service plutôt qu'un autre [12]. Pour pallier cette limite, nous proposons d'analyser l'ensemble des services existant sur un ENT et de croiser ces analyses avec des questions fermées relatives aux tâches réalisées pour atteindre différents objectifs, aux bénéfices observés [20, 24] ou à la valeur construite dans l'usage [25] et des questions ouvertes de type « pourquoi » [26]. (2) Certains critères sur les niveaux de compétences des enseignants sont mal formalisés. Nous proposons d'utiliser le modèle *Technological Pedagogical Content Knowledge* (TPACK) [27] qui décrit les trois principaux champs de connaissance que les enseignants sont amenés à mobiliser lors de l'intégration d'une technologie (connaissances liées aux contenus à enseigner, celles sur la pédagogie et celles sur la technologie). (3) Le caractère linéaire de ces modèles diffusionnistes ne permet pas de comprendre la progression des usages qui ne sont pas unifiés pour tous les services ou applications considérés [13]. Pour pallier cette limite, nous proposons d'utiliser des processus de classification multidimensionnels potentiellement plus adaptés [26, 28].

De manière à pouvoir positionner les résultats de notre étude comme un prolongement longitudinal des résultats des autres études, nous avons choisi d'utiliser la même méthode de recueil (questionnaire), d'intégrer des questions sur les éléments classiquement observés et de les compléter par d'autres, pour fournir des diagnostics plus précis.

3 Étude Open Digital Education-Nunc

3.1 Contexte de l'étude

L'enquête, conçue dans une optique d'évaluation et d'analyse des conditions d'appropriation du numérique, a été proposée aux utilisateurs de la solution d'ENT fournie par l'éditeur Open Digital Education. Le questionnaire a été administré et diffusé à grande échelle (cet ENT étant présent dans des académies en France métropolitaine et ultramarine) entre les mois de mai et de juin 2020. L'invitation à répondre au questionnaire était proposée directement sur l'ENT. En complément du questionnaire, des entretiens

d'approfondissement et une collecte de récits d'expérience ont été organisés à distance entre juin et juillet 2020. Le questionnaire comprend 3 parties (voir table 1). La première porte sur le profil du répondant. La deuxième porte sur l'expérience du numérique et de l'ENT : exploitation du numérique/ENT pour la réalisation de l'activité (les tâches pédagogiques structurées en fonction de l'approche envisagée [28]), expérience de l'ENT (échelle d'utilisabilité UMUX [22], et services préférés). Des questions ouvertes sont posées pour illustrer les avis. La troisième porte sur l'expérience globale de la continuité pédagogique en termes de motivation, sentiment d'efficacité, compétence (caractérisé selon les dimensions TPACK [27]), liens sociaux et autonomie. Les ressentis exprimés pour le début et la fin du confinement sont utilisés pour calculer les bénéfices/dégradations sur le plan personnel.

Table 1. Structure du questionnaire

Section	Description
1 - Profil	Caractéristiques sociodémographiques (genre, âge, ancienneté, niveau d'enseignement, discipline) Caractéristiques de l'établissement (taille, privé/public, localisation, école/collège/lycée) Conditions de travail à la maison (nombre d'enfants en continuité pédagogique au foyer, aide pour gérer ces enfants)
2 – Numérique et ENT	Mode de réalisation des tâches pédagogiques ou de gestion (4 modes de réalisation, 24 tâches) Score UMUX ENT (4 items, échelle en 6 points) Services ENT préférés (classement sur 25 services)
3 – Continuité pédagogique	Ressentis globaux au début et à la fin du confinement (12 items, échelle en 7 points).

3.2 Participants

467 enseignants ont répondu au questionnaire dont 78 % de femmes. 88 % des répondants ont entre 25 et 55 ans, 62 % ont plus de 11 ans d'ancienneté dans l'Éducation nationale et une majorité (93 %) sont titulaires dans des établissements publics (97 %). 60 % des répondants sont des enseignants du 1^{er} degré. Les enseignants du 2^d degré (188 personnes) exercent principalement en collège (86 %). Près d'un quart (24 %) des enseignants signale que moins de 5 % des élèves ne réalisent pas les activités proposées, alors que pour 15 % d'entre eux, la moitié de leurs élèves ou plus sont dans cette situation. Les réponses proviennent d'enseignants de différents départements avec une prédominance pour la Somme (24 % des enseignants) et la Martinique (13% des enseignants). Au vu de leur faible représentation, nous avons éliminé les réponses des enseignants exerçant en Lycée et Collège/Lycée. Le corpus d'analyse est donc composé de 441 réponses d'enseignants (279 en primaire, 162 en collège). Toutes les analyses statistiques (tris à plat, analyse bivariée ou multivariée) ont été réalisées à l'aide des logiciels Excel, XLStat et Jmp.

4 Résultats

4.1 Modalité de réalisation des tâches pédagogiques ou de gestion

La figure 1 présente, pour le collège et l'école élémentaire, les modalités de réalisation de 24 tâches, réorganisées selon les objectifs visés à plus haut niveau : conception, transmission, animation, vérification, communication et auto-formation. Les enseignants ont indiqué si ces tâches étaient réalisées ou non, et si oui par quels moyens (ENT, outils numériques, outils non numériques).

Objectifs	Tâches	Collège-ENT	Ecole primaire-ENT	Collège-Num	Ecole primaire-Num	Collège-NonNum	Ecole primaire-NonNum	Collège-Non	Ecole primaire-Non
Conception	1-Adapter les activités pour certains élèves	41%	38%	40%	42%	7%	13%	11%	6%
	2-Créer des activités	14%	27%	74%	63%	8%	7%	4%	3%
	3-Mettre à disposition des cours, ressources	73%	79%	24%	18%	1%	1%	2%	2%
Transmission	4-Mettre à disposition les activités du jour	71%	78%	22%	16%	1%	2%	6%	4%
	5-Mettre à disposition des exercices	72%	75%	23%	21%	2%	3%	2%	2%
	6-Mettre à disposition des séquences d'activité	61%	62%	23%	21%	4%	4%	12%	13%
	7-Mettre à disposition des activités pédagogiques pour développer les interactions la collaboration entre les élèves	42%	51%	18%	19%	2%	4%	38%	26%
	8-Travailler avec la vidéo (youtube, tutoriels, etc.)	39%	57%	43%	33%	0%	1%	18%	10%
	9-Publier, écrire avec les élèves (journal, blog, pad, etc.)	33%	67%	15%	7%	2%	1%	49%	25%
Animation	10-Travailler à l'oral (fichiers audio, webradio, podcast, etc.)	24%	36%	36%	32%	1%	2%	40%	30%
	11-Enseigner avec des applications spécifiques dédiées (Excal, Edumedia, Sesamath, Quidos...)	7%	5%	22%	20%	1%	0%	70%	75%
	12-Organiser des classes virtuelles avec de la vidéo	6%	5%	43%	61%	1%	1%	51%	33%
	13-Organiser des classes virtuelles avec des forums	6%	4%	26%	25%	1%	0%	67%	71%
	14-Réceptionner les productions des élèves	80%	55%	10%	32%	6%	6%	4%	8%
	15-Vérifier que le travail a été fait	81%	42%	20%	39%	10%	8%	9%	12%
	16-Evaluer les élèves	49%	12%	29%	34%	10%	1%	12%	36%
Communication	17-Gérer les demandes individuelles	77%	59%	14%	33%	6%	5%	4%	3%
	18-Maintenir le lien entre les élèves malgré la distance	67%	58%	22%	36%	4%	2%	7%	4%
	19-Organiser des points de suivi réguliers avec les familles	44%	29%	24%	46%	14%	19%	18%	6%
	20-Rendre visibles les questions de chacun	31%	30%	9%	14%	5%	6%	54%	50%
	21-Coopérer entre enseignants de non établissement	49%	17%	38%	61%	7%	15%	6%	7%
Auto-formation	22-Me former sur les technologies à utiliser	11%	13%	67%	69%	4%	2%	18%	16%
	23-Coopérer entre enseignants d'autres établissements	15%	5%	39%	38%	3%	5%	43%	52%
	24-Chercher des informations sur les activités à faire faire aux élèves	6%	4%	77%	85%	10%	8%	6%	3%

Fig. 1. Modalité de réalisation des tâches par les enseignants de collège et école élémentaire

Les comportements des enseignants de collège et école primaire sont globalement assez similaires. Les tâches les plus réalisées relèvent de quatre objectifs (transmission, communication, conception et recherche d'information). Si l'on regarde plus en détail, la transmission des activités concerne des cours, ressources et exercices (T3, T6). Elle se fait plutôt avec l'ENT et unitairement (souvent au jour le jour) (T4). Les enseignants proposent moins souvent des séquences d'activité (T6) ou des activités à faire en groupe (T7). La communication est, de la même manière, plutôt réalisée avec l'ENT, pour maintenir le lien avec les élèves (T18) et gérer les demandes individuelles (T17). Elle se fait principalement de manière interpersonnelle, les enseignants ne partagent pas les questions et réponses de façon collective (T20). L'auto-formation se fait principalement par des recherches sur Internet (T24) et par des échanges entre enseignants du même établissement via l'ENT ou d'autres outils numériques (mails, téléphone) (T21). La conception d'activité (T2) se fait plutôt avec des outils accessibles sur Internet et peu avec l'ENT sauf pour quelques enseignants du primaire. La vérification du travail des élèves (T14, 15, 16) se fait de la même manière principalement avec l'ENT en collège et avec l'ENT et le numérique en primaire. Il n'y a pas de modalité privilégiée pour l'évaluation du travail des élèves. Les tâches les moins réalisées sont des tâches d'animation. Les enseignants animent principalement les cours en utilisant des ressources vidéo. Les autres formes d'animation (écriture collaborative, travail à l'oral, classes

virtuelles, utilisation d'outils spécifiques) restent moins développées, mais quand elles sont faites, l'ENT permet plutôt des animations d'écriture (T9) et les autres outils numériques des animations de classe virtuelle (T12, 13).

4.2 Niveau d'intégration de l'ENT et des autres outils numériques dans les pratiques des enseignants et facteurs d'explication

Nous avons identifié les niveaux d'intégration de l'ENT et des autres outils numériques en fonction des 6 objectifs précédents en faisant une analyse avec la méthode des *K-means*. Nous avons choisi 5 niveaux pour respecter l'échelle TIM [28], mais nous ne les qualifions pas à ce stade. La figure 2 présente les classes de pratiques en fonction de la moyenne de réalisation des tâches par moyen utilisé (ENT, autres outils numériques).

	Classe	Auto-formation	Conception	Transmission	Vérification	Communication	Animation
ENT	1	0,08	0,07	0,12	0,13	0,21	0,11
	2	0,07	0,00	0,77	0,10	0,31	0,27
	3	0,15	0,00	0,76	0,70	0,58	0,22
	4	0,13	0,63	0,83	0,17	0,48	0,34
	5	0,20	0,64	0,85	0,85	0,69	0,32
Numérique (Hors ENT)	1	0,32	0,07	0,04	0,06	0,09	0,13
	2	0,72	0,45	0,09	0,09	0,16	0,28
	3	0,60	1,00	0,14	0,16	0,22	0,33
	4	0,74	0,62	0,09	0,78	0,47	0,36
	5	0,68	0,85	0,84	0,67	0,57	0,50

Fig. 2. Moyenne de réalisation des tâches par moyen (ENT, autres outils numériques) et par objectif (les valeurs sont normalisées entre 0 et 1)

La stratégie d'intégration de l'ENT se fait en variant et intensifiant les modalités d'interaction avec les élèves dans un premier temps par la transmission de cours, ressources, activités (niveau 2). Les niveaux 3 et 4 sont assez proches. Ils sont caractérisés par l'ajout d'objectifs de communication. Nous avons affecté le niveau 3 aux usages intégrant la vérification et le niveau 4 aux usages intégrant la conception et l'animation. Le niveau 5 correspond à une intensification de ces 5 types d'objectifs. À tous les niveaux, les objectifs d'auto-formation sont faibles. À l'inverse, l'intégration des technologies numériques sert principalement (niveaux 1-2-3) des objectifs de productivité pour les enseignants (Auto-formation, Conception), puis elle augmente en fonction des interactions à destination des élèves (niveaux 4-5).

Nous avons cherché à expliquer ces *niveaux d'intégration* en procédant à des analyses de covariance (seuil de significativité à 0,05) avec les variables descriptives du profil, de l'avis et des usages des outils pour la réalisation des tâches et du ressenti liés à la réalisation de l'activité dans le contexte de la continuité pédagogique (voir tableau 1).

Intégration de l'ENT. Les variables contextuelles (caractéristiques de l'établissement) et sociodémographiques (âge, genre, ancienneté) ne contribuent pas de manière significative à expliquer le niveau d'intégration de l'ENT. Les variables qui contribuent significativement à une non-intégration de l'ENT sont un faible niveau d'intégration du numérique (0,62) et paradoxalement une bonne expérience d'utilisation de l'ENT (0,27) ou la présence d'une communauté d'enseignants dans l'établissement (0,15). En

effet, l'expérience utilisateur des enseignants est dans l'ensemble très bonne (score UMUX de 4,07 sur 6) car les enseignants le trouvent facile à utiliser et les fonctionnalités répondent à leurs exigences. Mais cette expérience est frustrante car ils rencontrent différents problèmes d'utilisation, en particulier au collège. Les problèmes d'utilisation sont principalement liés aux problèmes techniques d'accès aux plateformes observés pour tous les ENT au début du confinement. Les frustrations exprimées concernent les processus de travail avec les documents (taille de l'espace de stockage trop petite, processus de chargement des documents fastidieux, manque de notification ou d'accès direct au travail rendu par les élèves, contraintes sur le format du travail rendu en un seul fichier, absence d'une confirmation de lecture des ressources), l'ergonomie de l'accès aux applications (rationaliser/limiter les applications accessibles par défaut, faciliter l'accès, faciliter la bascule entre les applications) et le manque d'une application de classe virtuelle. Les variables qui contribuent positivement et significativement à l'intégration de l'ENT sont plus complexes, car combinées : savoir que cela ne demande pas d'efforts et ne pas avoir de difficultés pour construire les contenus (0,05), en cas de difficultés sur les contenus, savoir pouvoir compter sur une communauté (0,04) ou avoir le sentiment que l'utilisation de l'ENT fait évoluer sur le plan professionnel (0,02). L'importance de savoir construire des contenus se matérialise en termes d'intégration par les actions de transmission et conception qui sont les pratiques les plus représentatives des niveaux d'intégration élevée. Les freins, observables même en cas de bonne expérience préalable ou de présence de communauté, indiquent que les enseignants n'ont globalement pas l'habitude d'utiliser ces fonctionnalités d'écriture de l'ENT et ne peuvent donc pas s'entraider.

Intégration du numérique. Comme pour l'ENT, les variables contextuelles et socio-démographiques ne contribuent pas de manière significative à expliquer le niveau d'intégration du numérique. Les variables qui contribuent négativement et significativement à l'intégration du numérique sont une faible intégration de l'ENT (0,62), et, dans une moindre mesure, la peur de perdre le lien avec les élèves (0,05), le manque de motivation (0,02) ou un sentiment d'efficacité faible (0,05). Réciproquement, celles qui y contribuent positivement et significativement sont : le sentiment de ne pas avoir de difficultés avec les outils numériques (0,13), le fait de considérer que l'activité est utile aux élèves, combinés à la peur de perdre le lien avec eux (0,05), le sentiment d'un effort minime nécessaire (0,04), et enfin une expérience positive de l'ENT combinée au fait de pouvoir compter sur une communauté (0,04).

Pour résumer, les facteurs les plus critiques d'intégration des ENT sont l'intégration d'autres outils dans les pratiques numériques et l'expérience d'utilisation. Réciproquement l'intégration du numérique est favorisée par l'auto-formation et l'intégration de l'ENT. Ainsi, l'intégration du numérique et l'ENT se font de manière conjointe, la première poussant la seconde sur l'innovation pédagogique, la seconde poussant la première sur l'organisation collective et collaborative de l'activité. La section suivante a pour objectif d'identifier plus précisément les utilisations qui sont faites de l'ENT, de manière à identifier quelles stratégies mettre en œuvre pour mieux développer son intégration et ainsi stimuler l'intégration du numérique plus globalement.

4.3 Les utilisations de l'ENT

Nous proposons d'analyser les services préférés des enseignants de la classe 5 en fonction des objectifs visés (figure 3) dans le but de caractériser les formes d'intégration élevées de l'ENT. En vert sont mentionnées les valeurs supérieures au 3^e quartile (Q3=11) de la série de données.

Les résultats sont globalement cohérents avec les résultats des précédentes enquêtes concernant l'utilisation massive de la messagerie, du cahier de textes, du cahier multimédia et du blog [17, 19, 26]. On peut observer une utilisation en forte progression du gestionnaire de notification, de la version mobile de l'ENT et du fil de nouveauté ce qui confirme l'intérêt de l'ENT pour organiser l'activité collectivement. Les utilisations sont en cours de développement pour certains services : Exercice et évaluation, Compétences, Mur collaboratif, Présence et Réserve de ressources. 12 services sur 25 sont sous-utilisés.

Services	Utilisation Globale	Utilisation par objectif					
		Animation	Communication	Vérification	Transmission	Conception	Auto-formation
Messagerie	128	11	77	128	105	34	3
Cahier de textes	95	10	64	95	78	23	3
Gestionnaire de notification	74	7	46	74	62	22	2
Cahier multimédia	70	5	43	70	57	24	1
Blog	61	6	42	61	52	17	1
Espace documentaire	53	4	34	53	44	18	2
Version mobile	48	4	31	48	42	13	2
Fil de nouveauté	42	4	28	42	32	9	2
Exercice et évaluation	27	0	10	27	20	11	1
Compétences	20	3	15	20	14	4	0
Mur collaboratif	12	4	9	12	8	6	0
Présence	11	2	9	11	10	2	0
Réserve de ressources	11	1	8	11	9	0	0
Wiki	10	0	7	9	8	4	0
Agenda	9	1	4	10	9	2	0
Annuaire	8	0	4	8	7	4	0
Carte mentale	7	1	4	7	6	3	1
Posto-fichier	7	1	5	7	6	2	1
Pad collaboratif	6	1	6	6	5	2	0
Forum	6	1	3	6	6	3	0
Bibliothèque d'activités	6	1	2	6	6	2	0
Carnet de bord	4	0	3	4	4	2	0
Pages	4	1	2	3	3	2	0
Actualité	3	0	1	4	3	0	0
Frise chronologique	2	0	1	2	2	1	0

Fig. 3. Répartition des utilisations par objectif

5 Perspectives : des stratégies de reconception de l'ENT

Pour améliorer l'intégration de l'ENT dans les pratiques il nous semble utile d'explorer la piste de la reconception de l'ENT sur les services les plus attractifs ou générant le plus de freins et frustrations en termes d'ergonomie et de gestion documentaire.

Considérant les freins sur les aspects ergonomiques, il serait intéressant d'offrir aux enseignants des moyens de personnaliser les services de leur espace de travail personnel ou de classe [29] : en sélectionnant/désélectionnant les services (sur le modèle des *stores*), et en proposant un modèle « par défaut » intégrant uniquement les services les plus pertinents. En complément, les enseignants devraient pouvoir paramétrer la mise en œuvre des services, par exemple concernant les notifications ou la synchronisation du stockage.

Pour pallier les problèmes de complexité de certaines opérations, une reconception pourrait s'opérer sur la base de parcours utilisateurs pour repenser l'architecture de l'information des ENT : identifier les moyens de mieux articuler les services de gestion de document, de notification, de communication, de production d'activité et de contrôle/vérification en : rationalisant l'accès aux services dans les espaces de travail, redéfinissant les services par fonctionnalité (faire converger les multiples services d'alerte et de notification), mais aussi en optimisant certains traitements (synchronisation des ressources dans les espaces documentaires, type de notification/alerte orienté en fonction de l'activité). Les études complémentaires sur l'organisation de la continuité pédagogique montrent en effet que les enseignants, les élèves et les parents sont souvent perdus entre les différents moyens possibles d'identifier, de réaliser, de transmettre ou d'avoir un retour sur les activités pédagogiques [30, 31].

La troisième stratégie de reconception est de mieux identifier et développer les moyens d'auto-formation, en particulier sur les formes d'animation des formations. Si l'on observe les stratégies globales des enseignants (figure 1 et 2), l'auto-formation se fait principalement par la recherche de ressources documentaires, puis par de la collaboration, dans l'établissement et dans une moindre mesure à l'extérieur de l'établissement. Actuellement, les seuls moyens d'auto-formation que propose l'ENT sont la communication avec les pairs (messagerie) et la bibliothèque d'activité (service ouvert début mars 2020). Ce service nous semble extrêmement prometteur, mais peut être retravaillé pour accueillir, en plus des modèles d'activité réalisés par les enseignants, des ressources de formation ou des liens vers des ressources externes. Cette ouverture, concomitante avec une ouverture des profils utilisateur vers les acteurs académiques comme les enseignants référents, les IAN (Interlocuteurs académiques pour le numérique) ou les inspecteurs chargés de la mission numérique, permettrait d'agréger et de valoriser les actions de cette communauté autour des objectifs de formation, d'animation, d'innovation.

6 Conclusion et perspectives pour développer les usages de l'ENT

L'épisode de continuité pédagogique vécue au printemps 2020 a été l'occasion de traiter à nouveau la question de l'intégration du numérique dans les pratiques pédagogiques des enseignants français. Une étude réalisée auprès des usagers des ENT One et Neo a pu mettre en évidence des usages essentiellement à des fins de conception, transmission, communication, dans une plus faible mesure de vérification et auto-formation. Les usages d'animation restent moins développés. En explorant la manière dont l'ENT s'intègre dans les pratiques des enseignants, nous avons identifié deux logiques différentes : diversifier les canaux de communication avec les élèves, en particulier via les notifications, et proposer des formes d'apprentissage plus actif, collaboratif et engageant. La stratégie d'intégration du numérique est quant à elle principalement motivée par des besoins d'efficacité en termes de conception et d'auto-formation pour identifier des moyens complémentaires pour proposer des formes d'apprentissage plus actif, collaboratif et engageant. Les deux stratégies se rejoignent donc sur ce dernier objectif

avec plus de difficultés à identifier comment les mettre en œuvre dans l'ENT. L'ENT trouve donc une place complémentaire aux autres outils de par son statut d'environnement de travail partagé avec les élèves et les parents, mais nécessite d'être reconçu sur certains aspects : la personnalisation, la rationalisation du parcours utilisateur et les services d'auto-formation.

Remerciements. Cette étude a été financée par le projet FUI AAP21 « REPI ».

References

1. Charlier, B., Deschryver, N., Peraya, D.: Apprendre en présence et à distance. Distances et savoirs. Vol. 4, 469–496 (2006).
2. Maulini, O.: Que penser ... de l'enseignement à distance. Université de Genève, Faculté de psychologie et des sciences de l'éducation. (2020).
3. Gérin-Lajoie, S., Papi, C., Paradis, I.: De la formation en présentiel à la formation à distance : Comment s'y retrouver ? 13 (2019).
4. Félix, C., Filippi, P.-A., Martin, P., Gebeil, S.: École et famille en temps de confinement. Et après ?, <https://www.cahiers-pedagogiques.com/Ecole-et-famille-en-temps-de-confinement-Et-apres>, last accessed 2020/11/06.
5. Cour des comptes: Le service public numérique pour l'éducation : un concept sans stratégie, un déploiement inachevé. Cour des comptes, Paris (2019).
6. Bruillard, É.: Le déploiement des ENT dans l'enseignement secondaire : entre acteurs multiples, dénis et illusions. Revue française de pédagogie. Recherches en éducation. 101–130 (2011).
7. OECD: Strengthening online learning when schools are closed: The role of families and teachers in supporting students during the COVID-19 crisis. , Paris (2020).
8. MENJS: PROFETIC 2018 - Connaître les pratiques numériques des enseignants. Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Paris (2019).
9. Jalal, G., Lachand, V., Tabard, A., Michel, C.: How teachers prepare for the unexpected: bright spots and breakdowns in enacting pedagogical plans in class. In: 13th European Conference on Technology Enhanced Learning. pp. 59–73. LNCS, Springer, Leeds, United Kingdom (2018).
10. MENJS: EVALuENT Synthèse globale. Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche (2018).
11. MEN: Cycle de consultation des acteurs des ENT. (2013).
12. Tricot, A., Chesné, J.-F.: Numérique et apprentissages scolaires - Rapport de synthèse. Cnesco, Paris (2020).
13. Poyet, F., Genevois, S.: Vers un modèle compréhensif de la généralisation des usages des ENT dans l'enseignement secondaire. RFP. 83–98 (2012).
14. Prieur, M., Steck, P.: L'ENT : un levier de transformation des pratiques pédagogiques pour accompagner les apprentissages du socle commun ? 13 (2011).
15. Pacurar, E., Abbas, N.: Analysis of French secondary school teachers' intention to integrate digital work environments into their teaching practices. Educ Inf Technol. 20, 537–557 (2015).
16. Schneeweile, M.: Représentation sociale d'un ENT dans l'enseignement secondaire : une étude pour comprendre et analyser les usages. Carrefours de l'éducation. n° 37, 211–226 (2014).

17. Daguet, H., Voulgre, E.: Discours et pratiques autour des Environnements Numériques de Travail - Utopie ou réalité ? In: Conférence Environnements Informatiques pour l'Apprentissage Humain. pp. 231–241. , Mons, Belgium (2011).
18. Puimatto, G.: Les réseaux numériques éducatifs, régulateurs, acteurs et vecteurs de l'évolution des pratiques et de l'organisation des établissements et de l'institution scolaires, (2006).
19. Hanna, D., Charalampopoulou, C.: Travail collaboratif sur l'ENT et innovation dans les pratiques professionnelles des enseignants. *Spirale - Revue de recherches en éducation*. N° 63, 23–35 (2019).
20. Codreanu, E., Michel, C., Bobillier-Chaumon, M.-E., Vigneau, O.: L'acceptation et l'appropriation des ENT (Espaces Numériques de Travail) par les enseignants du primaire. *STICEF*. 24, 13–49 (2017). <https://doi.org/10.3406/stice.2017.1724>.
21. Tricot, A., Plégat-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: Conférence Environnements informatiques pour l'apprentissage humain 2003. pp. 391–402 (2003).
22. Finstad, K.: The usability metric for user experience. *Interacting with Computers*. 22, 323–327 (2010).
23. Depover, C., Strebelle, A.: Une modélisation du processus d'innovation s'articulant sur une dynamique de réseaux d'acteurs. Presented at the January 1 (2007).
24. DeLone, W., McLean, E.: The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*. 19, 9–30 (2003). <https://doi.org/10.1080/07421222.2003.11045748>.
25. Michel, C., Bobillier-Chaumon, M.-E., Sarnin, P.: Prise en compte de la construction de valeur liée à l'expérience utilisateur dans la modélisation de l'acceptation technologique. In: Conférence conjointe de la 24ème édition d'IHM et de la 13ème édition de ERGO'IA. p. 8p. ACM (2012).
26. Poyet, F.: Généralisation des usages des ENT dans l'enseignement secondaire en France : analyse diachronique (2009 - 2014). *STICEF*. 23, 9–32 (2016).
27. Mishra, P., Koehler, M.J.: Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Rec*. 108, 1017–1054 (2006).
28. Kozdras, D., Welsh, J.: Enter the Matrix: A Pedagogy for Infusing Technology. In: Society for Information Technology & Teacher Education International Conference. pp. 536–541. Association for the Advancement of Computing in Education (AACE) (2018).
29. Paquienseguy, F.: L'usage, de l'appropriation au design. *Ocula*. 20, (2019).
30. Genevois, S., Lefer-Sauvage, G., Wallian, N.: Questionnaire d'enquête auprès des enseignants - Confinement et continuité pédagogique. *ICARE* (2020).
31. Rakotomalala Harisoa, N.A.: Pratiques, confinement et besoins en formation des enseignants. Réseau Canopé, Chasseneuil du Poitou (2020).

Réalité virtuelle et formation à distance : impact de la représentation des avatars sur la coprésence des apprenants dans un environnement virtuel collaboratif

Richard Ngu Leoubou^{1,2}, Marc Trestini², and Benoit Crespin¹

¹ University of Limoges - CNRS, XLIM UMR 7252 France
<https://www.xlim.fr/>

² University of Strasbourg - LISEC (EA 2310) France
<http://www.lisec-recherche.eu/>

Résumé. La réalité virtuelle est de plus en plus utilisée dans le domaine de l'éducation et de la formation. Les applications de réalité virtuelle fournissent aux apprenants l'occasion de vivre des expériences qu'ils ne peuvent pas trouver (totalement ou partiellement) dans le monde réel et qui favorisent les apprentissages. Dans le secteur de la formation à distance, il ressort des travaux existants que la réalité virtuelle offre de nouveaux supports d'apprentissage et une nouvelle façon de collaborer et de se socialiser. Elle permet de simuler des situations d'apprentissage semblables à celles réalisées dans une formation présentielle (cours magistral, manipulation des objets dans un laboratoire, etc.) en offrant aux utilisateurs, une grande liberté d'interactions dans des systèmes immersifs. L'objectif de notre article consiste à étudier un nouvel apport de la réalité virtuelle sur l'expérience utilisateur des apprenants dans un dispositif de formation à distance. Nous montrons notamment comment la représentation des avatars peut influencer le sentiment de coprésence et rapprocher les apprenants dans un environnement virtuel collaboratif.

Mots-clefs : Réalité virtuelle, Expérience utilisateur, Avatar, Coprésence, Formation à distance.

Abstract. Virtual reality is increasingly used in the field of education and training. VR applications provide learners with experiences that they cannot find (totally or partially) in the real world and that promote active learning. For distance learning, existing works indicate that VR offers a new way to learn, collaborate and socialize. It allows to simulate learning situations similar to those carried out in a face-to-face training (lectures, manipulation of objects in a laboratory, etc.) and gives users the freedom of interactivity with immersive features. The goal of this paper is to study a new contribution to the user experience of learners in a distance learning environment. More specifically, we show how the representation of avatars can influence the feeling of co-presence and bring learners closer together in a collaborative virtual environment.

Keywords: VR, User experience, Avatars, Copresence, Distance learning.

1 Introduction

Au cours de ces dernières années, les chercheurs en Environnements Informatiques pour l'Apprentissage Humain (EIAH) se sont de plus en plus intéressés à l'expérience utilisateur des apprenants dans des situations d'apprentissage informatisées. L'enjeu majeur consiste à concevoir des environnements numériques permettant de manipuler et améliorer la motivation, le flow, l'engagement, la satisfaction, les émotions positives des apprenants, etc. Parmi les technologies utilisées, force est de constater que la réalité virtuelle (RV) suscite un intérêt croissant dans la communauté des chercheurs. Elle offre en effet le moyen de créer un monde tridimensionnel simulé représentant un environnement d'apprentissage réaliste pour les étudiants en termes de vision, d'audition et de toucher, en faisant en sorte que les apprenants se sentent comme des participants de l'environnement virtuel (EV) [8]. Les technologies mobilisant la RV pour l'apprentissage sont connues sous le terme d'Environnements Virtuels Educatifs (EVE). Elles offrent de nouveaux supports d'apprentissage et une nouvelle façon de collaborer et de se socialiser [16]; elles permettent de simuler des situations d'apprentissage réalisées dans une formation présentielle (cours magistral, manipulation des objets dans un laboratoire, etc) [16, 8] et donnent aux utilisateurs une liberté d'interactivité avec la caractéristique d'immersion [1].

L'utilisation de la RV dans le domaine de l'apprentissage à distance est en encore un sujet assez peu investi par les chercheurs en EIAH. Elle est trop souvent considérée comme un simple moyen technique visant à faire communiquer les personnes entre-elles sans tenir compte de leur ressentis, et leurs impressions. De fait, elle tend généralement à oublier les besoins des apprenants, leur satisfaction à l'égard de l'expérience d'apprentissage [1] et les émotions vécues entre apprenants dans les différentes interactions. Ainsi nous proposons dans cet article d'analyser l'impact de la RV sur l'expérience utilisateur des apprenants. Il s'agit en particulier de montrer comment les différentes représentations des avatars utilisées dans un environnement virtuel peuvent agir sur le sentiment de coprésence entre apprenants et ainsi rapprocher ces derniers. Une expérimentation menée sur un groupe d'étudiants sera détaillée pour montrer ce résultat.

Cet article est organisé comme suit. Dans la section 2, nous présentons le cadre conceptuel de la RV. La section 3 est consacrée à l'état de l'art et à notre problématique de recherche. La section 4 décrit la méthodologie de recherche utilisée, puis nous discutons dans la section 5 des résultats de notre expérimentation avant de conclure.

2 Réalité virtuelle, coprésence et avatars

La RV est un domaine scientifique et technique exploitant l'informatique à travers des interfaces comportementales en vue de simuler dans un environnement virtuel le comportement d'entités 3D, qui sont en interaction en temps-réel entre elles et avec une (ou plusieurs) personne(s) en immersion pseudo-naturelle [11]. Les interfaces matérielles de la RV sont dites comportementales

car elles exploitent les comportements humains. Elles se composent d'interfaces sensorielles, motrices et sensorimotrices, permettant aux humains d'échanger avec l'ordinateur [11, 23]. L'immersion en RV représente l'exposition d'un utilisateur dans un EV à l'aide de dispositifs occultant en partie la perception de son environnement réel, pour afficher en lieu et place une image du monde virtuel [7]. Dans ce contexte, la présence en RV est le sentiment authentique d'exister dans un monde autre que le monde physique où notre corps se trouve [6]. Cette présence peut être évaluée selon trois dimensions [18, 15] : il y a d'abord la présence physique ou environnementale, dans laquelle l'utilisateur va être amené à percevoir comme réels les objets du monde virtuel durant son expérience ; ensuite, la présence de soi ou présence personnelle qui intervient lorsque l'utilisateur perçoit son soi comme faisant partie du monde virtuel ; enfin, la présence sociale ou *coprésence* qui représente le sentiment d'être ensemble avec d'autres dans un EV collaboratif, même lorsque les participants sont physiquement situés dans différents sites [25]. Dans ce contexte, les avatars sont utilisés pour mettre en place l'interaction sociale entre les utilisateurs. Un avatar est une représentation visuelle d'un utilisateur [11] ou d'un personnage virtuel dans un EV. De manière générale, il offre de multiples possibilités de représentations plus ou moins fidèles de l'utilisateur [5]. Il existe trois composantes principales pour caractériser la fidélité visuelle de l'avatar dans un EV [12] : (1) la fidélité de la forme visuelle qui décrit le respect de la forme de l'utilisateur représenté par l'avatar ; (2) la fidélité de l'identité qui décrit à quel point l'identité de l'utilisateur représenté par l'avatar est reconnaissable par les autres utilisateurs (visage, taille, vêtements, accessoires, etc.) ; (3) la fidélité cinématique qui décrit la fidélité de reproduction des mouvements de l'utilisateur par son avatar. Pour les besoins de notre étude, nous nous sommes limités à la première composante qui décrit la fidélité de la forme.

3 Comment la réalité virtuelle impacte-t-elle la formation à distance ?

3.1 État de l'art

Les premiers travaux dans le domaine de la RV appliquée à la FAD datent de moins d'une décennie [1, 17, 8]. Une grande partie de ces travaux utilisent les possibilités offertes par la 3D et la RV pour reproduire (dans la limite des possibilités) les conditions d'une formation présentielle. Nous pouvons par exemple citer Chang *et al.* [8] qui décrivent un système simulant la vie réelle sur un campus où les étudiants peuvent participer à des activités comme avoir des cours, faire des exercices et se faire des amis ; les enseignants peuvent proposer des conférences, revoir les travaux des étudiants et mener des examens ; enfin, les gestionnaires gèrent les activités quotidiennes d'enseignement et les affaires des étudiants. Suivant le même principe, le projet CLEV-R (a collaborative learning environment with virtual-reality) offre un environnement qui imite le monde réel où les utilisateurs, représentés par des avatars, assument divers rôles dans

le monde virtuel [16]. Le système VRMed (Virtual reality in Medical education) [9] est lui plutôt destiné à l'enseignement à distance en médecine, en proposant une salle de cours virtuelle reproduisant les conditions d'une vraie salle de classe universitaire, un laboratoire virtuel permettant aux étudiants de manipuler des objets tels que des microscopes, une bibliothèque virtuelle où les étudiants peuvent naviguer pour rechercher des connaissances, etc.

D'autres travaux ont permis de montrer l'impact positif des avatars sur l'apprentissage dans les EV, notamment sur l'engagement des apprenants et des opportunités d'interaction de qualité entre apprenants à distance [4]. Baker *et al.* [3] ont constaté que les avatars permettent de développer un sentiment de communauté entre apprenants dans un contexte de formation à distance. Plus récemment les travaux de Gorisse [13] montrent un impact positif de la fidélité et de la véracité des avatars sur l'expérience utilisateur et le sentiment de présence des apprenants dans un environnement de réalité virtuelle.

3.2 Problématique

Contrairement aux formations traditionnelles dites *en présentiel* qui réunissent en un même lieu et en même temps apprenants et enseignants, l'une des caractéristiques principales des modalités propres à la FAD est précisément la distance. Elle peut être considérée comme fondement de la relation pédagogique [14]. Cette distance se manifeste souvent chez l'apprenant selon plusieurs dimensions [10] : spatio-temporelle, technologique, socioculturelle, économique, socioaffective, etc. Du fait de l'éloignement géographique, la distance spatio-temporelle peut créer par exemple chez certains apprenants un sentiment d'isolement qui lorsqu'il est mal vécu peut se transformer en solitude. La dimension socioaffective de la distance, c'est-à-dire l'ensemble des sentiments et émotions véhiculés dans et par les relations sociales entre les acteurs du dispositif de formation, fait que la confiance en l'autre sera plus difficile à mettre en place, car la régularité des interactions est plus faible que dans une formation présentielle [22]. Il est donc important de prendre en compte les composantes de la distance dans la conception des dispositifs de FAD. D'après Puimatto [20], la capacité de ces dispositifs à apprivoiser la distance conditionne leur efficacité et leur acceptabilité.

Les travaux présentés dans la littérature scientifique répondent partiellement à cette problématique en proposant des possibilités de collaboration et d'interaction. Cependant, les dimensions socioaffective et spatio-temporelle de la distance restent très peu ou pas traitées. De plus, les aspects didactiques liés à la présentation des contenus et visuels graphiques (design émotionnel), ont montré par le passé leurs importances dans l'apprentissage. ils ne sont donc pas à négliger car ils permettent d'induire des émotions positives et facilitent la compréhension chez les apprenants [19].

Nous proposons donc d'étudier comment la RV peut prendre en compte ces dimensions en se focalisant sur le sentiment de coprésence et le rapprochement entre apprenants dans les dispositifs de FAD. Notre questionnement de recherche porte plus précisément sur l'impact du choix des avatars sur le sentiment de

coprésence des apprenants. Nous formulons l'hypothèse que dans un environnement de RV en contexte de FAD, le sentiment de coprésence des apprenants varie en fonction du type d'avatar utilisé. Notre expérimentation se focalise en outre sur le degré d'anthropomorphisme des avatars, en utilisant d'abord des formes géométriques simples (sphères) pour les représenter puis des personnages imaginaires type robots, et enfin des personnages 3D modélisés à partir des photos des apprenants. Le degré d'anthropomorphisme représente le degré de ressemblance à une forme humaine.

4 Méthodologie de recherche

4.1 Évaluation de la coprésence en réalité virtuelle

Deux types de méthodes sont généralement utilisées pour mesurer le sentiment de présence dans un environnement de réalité virtuelle: les méthodes subjectives et objectives [6]. Les méthodes subjectives sont, pour la plupart, basées sur un principe d'évaluation par questionnaire. Il existe des questionnaires de mesure de la présence physique [24, 21] et des questionnaires de mesure de la présence sociale et de la coprésence [2, 18]. Les méthodes objectives sont quant à elles basées sur l'évaluation des réponses physiologiques et l'évaluation des réponses comportementales [6].

Selon Riva *et al.* [6], différents critères permettent de s'assurer de la validité de la méthode utilisée. Il s'agit de la fiabilité (la mesure doit être stable et cohérente au cours du temps); la validité (l'expérimentation mesure le facteur escompté); la sensibilité (la mesure doit pouvoir distinguer les niveaux de présence); la robustesse (la méthode employée pour la mesure doit pouvoir être transposée à différents systèmes de réalité virtuelle et différents contenus et contextes); la non-intrusivité (la méthode utilisée ne doit pas perturber le déroulement de l'expérience); la commodité (la méthode doit être simple à comprendre, à mettre en place et à administrer). Pour vérifier notre hypothèse, nous avons utilisé la méthodologie basée sur des évaluations subjectives par questionnaire en adaptant le questionnaire de Nowak (voir Table 1) traitant de la coprésence en environnement virtuel [18].

Ce questionnaire permet d'évaluer la présence perçue de l'autre et la coprésence perçue par soi-même. Pour les besoins de notre étude, nous avons extrait et traduit 9 questions en les contextualisant dans le domaine de la FAD (le terme *partenaire d'interaction* a par exemple été remplacé par *autres apprenants*). La coprésence est ainsi évaluée à partir de 6 variables dépendantes :

- l'implication dans l'activité (i)
- la sensation (chaleur/froideur) perçue entre apprenants (ii) ;
- l'échange des informations personnelles (iii) ;
- le sentiment de proximité entre apprenants (iv) ;
- les communications vocales ou verbales entre apprenants (v) ;
- la motivation à réaliser une activité en groupe (vi).

Chaque affirmation est évaluée sur une échelle de Likert (pas du tout d'accord, pas d'accord, ni d'accord ni en désaccord, d'accord, tout à fait d'accord).

Table 1. Adaptation du questionnaire subjectif de Nowak

Coprésence perçue par les autres participants		
Code	Variable	Items
COO1	(i)	Les autres apprenants ont été fortement impliqués dans l'activité.
COO2	(ii)	Les autres apprenants ont communiqué la froideur plutôt que la chaleur.
COO3	(iii)	Les autres apprenants n'étaient pas disposés à partager des informations personnelles avec moi.
COO4	(iv)	Les autres apprenants ont créé un sentiment de proximité entre nous.
COO5	(v)	Les autres apprenants étaient intéressés à me parler.
Coprésence perçue par soi-même		
COO8	(iii)	Je n'étais pas disposé à partager des informations personnelles avec les autres apprenants.
COO9	(iv)	J'ai essayé de créer un sentiment de proximité entre nous.
COO10	(v)	Je souhaitais parler aux autres apprenants.
COO11	(vi)	L'activité d'apprentissage avec les autres apprenants m'a particulièrement motivé.

4.2 Caractéristiques de l'environnement d'expérimentation

Le public de notre expérience est composé de 22 apprenants tous inscrits en formation à distance dans des cursus proposés par les universités de Strasbourg et de Limoges. L'expérimentation se déroule dans une salle virtuelle, dont l'un des murs contient un tableau destiné à la projection de contenus (images, fichiers PDF, etc). Les utilisateurs connectés peuvent se déplacer dans l'environnement virtuel par téléportation. La communication se fait à l'aide d'un chat vocal utilisant le son spatialisé (i.e. permettant aux apprenants de localiser l'origine des sons dans la salle).

4.3 Activité groupale proposée

Le domaine d'application de l'activité groupale proposée aux apprenants est l'entrepreneuriat. L'objectif final est de trouver une idée innovante pour la création d'une entreprise. L'expérience se décompose en une série de 3 séquences de 20 min en moyenne. Chaque séquence est caractérisée par l'utilisation d'une forme différente d'avatar : soit un avatar avec un faible degré d'anthropomorphisme représenté sous forme d'une sphère (Fig. 1), soit un personnage imaginaire représenté sous la forme d'un robot (Fig. 2), soit un avatar avec un haut degré d'anthropomorphisme (Fig. 3) modélisé à partir de la photo de l'apprenant. Chaque forme d'avatar est complétée par le prénom des apprenants.

L'activité est réalisée en groupes de 4 ou 5 participants (5 groupes). La première séquence rassemble les participants d'un même groupe ainsi qu'un tuteur dans une classe de réalité virtuelle. Dans cette séquence, le tuteur se présente et montre un diaporama contenant les instructions sur l'activité à réaliser en groupe. Dans la deuxième séquence, les apprenants se retrouvent

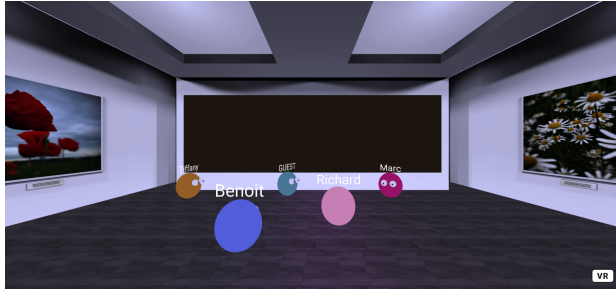


Fig. 1. Avatars avec une forme de sphère



Fig. 2. Avatar avec une forme de robot



Fig. 3. Avatars avec un haut degré d'anthropomorphisme

dans l'environnement virtuel pour réaliser l'activité. Le scénario consiste en un tour de table où chacun se présente. Ensuite le groupe réalise une forme de brainstorming pour recueillir les idées, puis choisit l'idée qui sera présentée sur un document partagé. La dernière séquence se réalise avec le groupe en présence du tuteur. Chaque groupe désigne un rapporteur qui présente les membres du groupe et l'idée innovante retenue.

Les 3 séquences sont réalisées sur une période de 2 semaines. A la fin de la série, notre questionnaire est soumis aux participants pour évaluer la coprésence en fonction des différentes variables.

5 Résultats et analyse

Si nous prenons à titre d'exemple la variable (iv) qui concerne la motivation à réaliser une activité en groupe, un premier tri à plat nous permet d'observer que les apprenants semblent motivés pour réaliser des activités groupales avec leurs pairs dans l'environnement RV. Cette observation se monte à 72,7% (soit 50% "d'accord" et 22,7% "tout à fait d'accord") quand l'activité se déroule avec un avatar ayant la forme d'une sphère, à 81,8% (59,1% "d'accord" et 22,7% "tout à fait d'accord") avec un avatar ayant la forme d'un robot, et 95,4% (31,8% "d'accord" et 63,6% "tout à fait d'accord") avec un avatar ayant une forme humaine. Si nous croisons maintenant la variable avatar avec la variable (iv) portant sur la motivation des apprenants à travailler ensemble dans l'environnement virtuel, nous obtenons les résultats reportés dans la Table 2.

Table 2. Tri croisé récapitulant la motivation à travailler ensemble (COO11) en fonction de la forme d'avatar

AVATAR /C0011	Pas du tout d'accord	Pas d'accord	Ni d'accord ni en désaccord	D'accord	Tout à fait d'accord	Total
AVATAR SPHERE	0,0%	0,0%	27,3%	50,0%	22,7%	100%
AVATAR ROBOT	4,6%	9,1%	4,6%	59,1%	22,7%	100%
AVATAR HUMAIN	0,0%	0,0%	4,6%	31,8%	63,6%	100%
TOTAL	1,5%	3,0%	12,1%	47,0%	36,4%	100%

Dans ce tableau croisé, nous remarquons d'abord que la forme des avatars semble avoir une influence non négligeable sur le degré de motivation d'un apprenant à travailler avec d'autres apprenants dans une activité groupale. Un test de χ^2 permet de le vérifier : en effet, avec un nombre de degrés de liberté ddl égal à 8, on obtient $\chi^2 = 20,91$ avec $1 - p = 99,23\%$. Par ailleurs, le V de Cramer est égal à 28,15 : cette valeur témoigne, de manière complémentaire, de la force de la relation étudiée. On en déduit que les différences observées entre nos valeurs expérimentales et nos valeurs théoriques ne sont pas dues au hasard, autrement dit que la dépendance entre la forme des avatars et la motivation des apprenants à travailler ensemble dans l'environnement virtuel est avérée et très significative.

Si l'on se focalise maintenant sur ces écarts, on constate que les variables en bleu dans la table 2 (4,6% "pas du tout d'accord" pour l'avatar robot, 9,1% "pas d'accord" pour l'avatar robot, 27,3% "ni d'accord ni en désaccord" pour l'avatar sphère, et 63,6% "tout à fait d'accord" pour l'avatar humain) sont celles dont l'effectif réel (données expérimentales) est nettement supérieur à l'effectif théorique. Par conséquent, ces modalités particulières sont précisément celles qui contribuent le plus à la dépendance de la relation. Une analyse similaire a été menée pour chaque croisement de variables envisagé, permettant d'évaluer la dépendance entre chacune d'elles et la forme d'avatar utilisée. Les résultats sur l'ensemble des variables évaluées sont présentés dans la Table 3.

Table 3. Résultat du tri croisé pour l'ensemble des variables

Variable	Question	χ^2	ddl	1-p	V de Cramer	Dépendance
i	COO1	18,25	6	99,44%	37,18	Très significative
ii	COO2	21,80	8	99,47%	40,64	Très significative
iii	COO3	6,87	8	44,89%	22,81	Pas significative
	COO8	5,24	6	48,61%	19,92	Pas significative
iv	COO4	14,81	8	93,69%	33,49	Peu significative
	COO9	16,95	8	96,93%	35,83	Significative
v	COO5	12,14	6	94,10%	30,32	Peu significative
	COO10	12,12	6	94,07%	30,30	Peu significative
vi	COO11	20,91	8	99,23%	39,70	Très significative

Les résultats obtenus nous permettent de faire un certain nombre d'observations sur ces variables particulières et sur les relations qui les lient :

1. En contexte de FAD, l'environnement de RV semble contribuer à l'implication des apprenants dans la réalisation d'activités groupales. Le niveau d'implication semble varier significativement en fonction de la représentation des avatars, avec un niveau d'implication plus élevé pour les avatars présentant un haut degré d'anthropomorphisme.
2. En environnement RV, les apprenants semblent communiquer de manière chaleureuse. Le degré de chaleur ressenti dans la communication interpersonnelle semble varier en fonction de la représentation des avatars, avec un degré plus élevé pour les avatars présentant un haut degré d'anthropomorphisme. Ce résultat peut se justifier par l'effet d'immersion dans l'environnement virtuel conduisant d'une part au sentiment d'être avec les autres, et d'autre part au sentiment de réconfort provoqué par le réalisme des avatars (visages connus et reconnus).
3. Les apprenants ne semblent pas très disposés à partager des informations personnelles dans un environnement RV conçu pour la formation à distance, quel que soit le type d'avatar. Ce résultat pourrait se justifier par le fait que notre expérience n'a duré que 60 minutes au total. De fait, les apprenants n'ont pas eu le temps de s'y sentir à l'aise et de faire confiance à leurs pairs.

Ce constat mériterait néanmoins d'être à nouveau testé sur une période plus longue.

4. Etant éloignés physiquement les uns des autres dans un contexte de FAD (distance spatiale), les apprenants semblent chercher dans cet environnement expérimental à recréer un sentiment de proximité (i.e. à réduire la distance perçue entre eux) mis à mal par la distance physique. Ce sentiment de proximité est perçu différemment selon le sens des relations interpersonnelles considérées, de soi par rapport aux autres apprenants ou des autres par rapport à soi. La perception de la proximité de soi par rapport aux autres semble varier en fonction de la représentation des avatars, avec un sentiment plus élevé pour les avatars représentant un haut degré d'anthropomorphisme. En revanche, la perception de la proximité des autres par rapport à soi semble moins varier en fonction de l'avatar. Une explication logique à cette différence peut être le fait que dans l'environnement RV, les apprenants ne voient pas leur propre avatar mais ceux des autres. Les avatars utilisés n'étant pas expressifs (pas de mouvement de la bouche lorsqu'on parle, pas de modification du regard, etc), il devient plus difficile de percevoir le sentiment des autres.
5. Les échanges vocaux semblent susciter un très grand intérêt chez les apprenants immergés dans l'environnement RV. Ce résultat reste vrai dans les deux sens indiqués précédemment (l'intérêt de soi à parler aux autres et l'intérêt des autres à parler à soi), quel que soit le type d'avatar. Cet intérêt pourrait se justifier par la difficulté à effectuer des échanges textuels dans l'environnement.
6. Réaliser une activité groupale avec ses pairs semble provoquer chez les apprenants une réelle motivation. Ce sentiment augmente en fonction de la forme de l'avatar utilisé, et en particulier pour les avatars présentant un haut degré d'anthropomorphisme.

Les premiers résultats obtenus dans cette étude sont très encourageants et semblent montrer que la représentation des avatars pourrait avoir un impact positif sur certains facteurs de coprésence. Cependant, avant de se prononcer plus sûrement sur cette question de recherche, il serait utile de reproduire l'expérience avec un public beaucoup plus large, le nôtre étant réduit à 22 participants. De plus, la variable avatar est de type intra-sujet. Dans ce contexte, chaque participant voit plusieurs modalités de la variable indépendante (les différents types d'avatars). L'ordre dans lequel ces conditions expérimentales sont présentées peut avoir une influence sur les résultats. Un des avantages de ce type d'expérimentation est le fait d'éviter le problème des variations inter-individuelles, puisque l'on compare la perception du sujet à propos d'un traitement (un type d'avatar) à la perception du même sujet à propos d'un autre type de traitement (un autre type d'avatar). Cependant, il serait nécessaire d'équilibrer l'ordre de chaque type d'avatar en les contrebalançant dans différents groupes pour conforter les résultats. En d'autres termes, pour les 3 types d'avatars il faudrait prévoir 6 groupes (i.e. le nombre de permutations possibles de 3 types d'avatars).

Une autre amélioration consisterait, au lieu de soumettre un seul questionnaire à la fin des 3 séquences d'expérimentation, à proposer un questionnaire aux apprenants après chaque séquence.

6 Conclusion

L'utilisation de la réalité virtuelle dans la formation à distance est récente, mais elle semble très prometteuse. Les possibilités offertes par cette technologie font émerger de nouveaux moyens d'apprentissage.

L'expérience réalisée dans le cadre de cette étude nous permet de constater que la réalité virtuelle pourrait contribuer à améliorer l'expérience des utilisateurs à travers la coprésence des apprenants dans l'environnement virtuel. Nous avons pu observer que certains facteurs de la coprésence liés à l'expérience utilisateur pourraient être favorablement influencés par la RV. Il s'agit de l'implication des apprenants dans la réalisation d'une activité groupale, la motivation à réaliser une activité avec d'autres apprenants, ou les relations chaleureuses que ces derniers ont ressenties dans leurs échanges. Ces trois premiers facteurs semblent être fortement influencés par la représentation des avatars dans l'environnement virtuel, avec une influence plus marquée pour les avatars qui possèdent un haut degré d'anthropomorphisme. Par ailleurs, bien que peu ou pas influencés par la représentation des avatars, les facteurs tels que le sentiment de proximité entre apprenants et les communications verbales semblent avoir contribué au sentiment de coprésence dans l'environnement. Enfin, notre étude montre que les apprenants sont peu disposés à partager les informations personnelles dans l'environnement, quelle que soit la forme d'avatar.

Il est important de noter que notre expérience a été réalisée sur un échantillon relativement faible (22 personnes). Néanmoins, nos premiers résultats montrent qu'un environnement de réalité virtuelle pour la formation à distance, conçu autour des préférences et des besoins des apprenants et utilisant des avatars représentés par une forme humaine, pourrait concourir de façon positive au bien-être et à la motivation des apprenants.

Bibliographie

1. Abidi, M., El-Tamimi, A., Al-Ahmari, A.: Virtual reality: Next generation tool for distance education. *International Journal of Advanced Science and Engineering Technology* **2**, 95–100 (2012)
2. Bailenson, J., Rex, C., Beall, A., Loomis, J.: Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators Virtual Environments* **10**, 583–598 (2001)
3. Baker, S.C., Wentz, R.K., Woods, M.M.: Using virtual worlds in education: Second life® as an educational tool. *Teaching of Psychology* **36**(1), 59–64 (2009)
4. Blake, A.M., Moseley, J.L.: The emerging technology of avatars: Some educational considerations. *Educational Technology* pp. 13–20 (2010)
5. Bonfils, P.: Environnements immersifs: spectacle, avatars et corps virtuel, entre addiction et dialectique sociales. *Hermès, La Revue* (1), 53–58 (2012)

6. Bouvier, P.: La présence en réalité virtuelle, une approche centrée utilisateur. Ph.D. thesis, Université Paris-Est (2009)
7. Burkhardt, J.M., Bardy, B., Lourdeaux, D.: Immersion, réalisme et présence dans la conception et l'évaluation des environnements virtuels. *Psychologie Française* **48**, 35–42 (2003)
8. Chang, X.q., Zhang, D.h., Jin, X.x.: Application of virtual reality technology in distance learning. *International Journal of Emerging Technologies in Learning (IJET)* **11**(11) (2016)
9. Dimitropoulos, Kosmas, M.A.M.I.: Building virtual reality environments for distance education on the web: A case study in medical education. *International Journal of Social Sciences* **2** (2008)
10. Dussarps, C.: L'abandon en formation à distance. *Distances et Médiations des Savoirs* (10) (2015)
11. Fuchs, P.: Théorie de la réalité virtuelle : les véritables usages. Collection Mathématiques et informatique, Mines Paristech PSL, Paris (2018)
12. Gamelin, G., Chellali, A., Dumas, C., Otmame, S.: Fidélité de l'avatar du partenaire distant dans un environnement virtuel immersif: effets sur les interactions spatiales. In: 30eme conférence francophone sur l'interaction homme-machine. pp. 227–233 (2018)
13. Gorisse, G.: Impact du point de vue et de la fidélité des avatars sur les sentiments de présence et d'incarnation en environnement virtuel immersif. Ph.D. thesis, Paris, ENSAM (2019)
14. Jacquinot, G.: Apprivoiser la distance et supprimer l'absence ? ou les défis de la formation à distance. *Revue Française De Psychanalyse* **102**, 55–67 (1993)
15. Lee, K.: Presence, explicated. *Communication Theory* **14**, 27 – 50 (2004)
16. McArdle, G., Monahan, T., Bertolotto, M., Mangina, E.: A web-based multimedia virtual reality environment for e-learning. *Proceedings of Eurographics* **4**, 9–13 (2004)
17. Monahan, T., McArdle, G., Bertolotto, M.: Virtual reality for collaborative e-learning. *Computers Education* **50**, 1339–1353 (2008)
18. Nowak, K., Biocca, F.: The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments* **12**, 481–494 (2003)
19. Plass, J.L., Heidig, S., Hayward, E.O., Homer, B.D., Um, E.: Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction* **29**, 128–140 (2014)
20. Puimatto, G.: Apprivoiser la distance ? Distances et médiations des savoirs (2019)
21. Schubert, T., Friedmann, F., Regenbrecht, H.: The experience of presence: Factor analytic insights. *Presence* **10**, 266–281 (2001)
22. Siméone, A., Eneau, J.: Trust and credibility in an online computer-supported collaborative learning (cscl) task. In: 9th World Conference on Computers in Education - IFIP (2009)
23. Sternberger, L.: Interaction 3-D en réalité virtuelle. Ph.D. thesis, Université Louis Pasteur - Strasbourg I (2006)
24. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. *Presence* **7**(3), 225–240 (1998)
25. Youngblut, C.: Experience of presence in virtual environments. *Institute for Defense Analyses* p. 149 (2003)

Combiner les leviers de l'approche par compétences et de l'auto-régulation pour accompagner le travail en autonomie à l'université. Analyse du service COMPER

Laëtitia Pierrot¹, Christine Michel¹, Julien Broisin², Nathalie Guin³, Marie Lefevre³ et Rémi Venant⁴

¹ Université de Poitiers, TECHNE EA 6316, F-86073 Poitiers, France

christine.michel,laetitia.pierrot@univ-poitiers.fr

² Université Toulouse 3 Paul Sabatier, IRIT, UMR 5505, F-31062 Toulouse, France

julien.broisin@irit.fr

³ Université de Lyon, CNRS, LIRIS UMR 5205, F-69622 Villeurbanne, France

nathalie.guin,marie.lefevre@liris.univ-lyon1.fr

⁴ Le Mans Université, LIUM EA 4023, F-72085 Le Mans, France

remi.venant@univ-lemans.fr

Résumé. Combinant approche par compétences et auto-régulation, le service COMPER associe la présentation graphique d'un profil de compétences à un système de recommandations personnalisées d'exercices d'entraînement à utiliser en complément des cours. Destiné aux étudiants, le service a été expérimenté auprès de 181 étudiants de DUT1 (Informatique), en complément de Travaux pratiques réalisés à distance, de manière à en évaluer l'utilité et l'utilisabilité. L'analyse des traces d'activité et des avis exprimés par les étudiants montre que le service est globalement jugé utile, en dépit d'une utilisation restreinte à la moitié d'entre eux, en particulier pour les profils d'étudiant travaillant individuellement.

Mots-clés : Auto-régulation, utilisabilité, utilité, autonomie, approche par compétences.

Abstract. Combining a competence-based approach and self-regulation, COMPER consists of a graphical presentation of the student's competence profile and a personalized resources recommendation system to be used and the courses. We tested COMPER service with 181 first-year students (DUT1 in Computer Science), in addition to practical work carried out remotely, in order to assess its usefulness and usability. Based on the survey (usage data and scales) analysis, our study shows that students globally perceive the service as useful, especially for those who work individually.

Keywords: Self-Regulated Learning, Usability, Utility, Autonomous Learning, Competency-based Approach

1 Introduction

Étudier à l'université demande, de la part des étudiants, de disposer de certaines compétences en termes d'organisation et de réalisation du travail. Or, la plupart des étudiants entrant à l'université ne sont pas formés à la mise en œuvre de ces méthodes. En effet, une étude sur les habitudes de futurs étudiants [1] montre que la stratégie d'apprentissage la plus courante (70,7 % de l'échantillon de lycéens interrogés) est l'organisation et la transformation de contenu pédagogique par la création de résumés ou de cartes conceptuelles. Les stratégies de recherche d'information (6,7 %), de définition d'objectifs et planification (4 %), ou de demande d'aide aux enseignants (4 %) sont marginales. Seul un participant a déclaré demander l'aide de ses pairs. D'autres stratégies (auto-évaluation, suivi, structuration de l'environnement) sont inexistantes. Il est donc nécessaire de fournir aux apprenants des formes de soutien à l'autonomie dans leurs apprentissages. Ce problème est encore plus critique si l'on considère qu'avec la crise sanitaire les étudiants doivent réaliser la majorité de leurs activités d'apprentissage à distance.

Différentes stratégies ont été utilisées pour pallier ce problème, dont celles visant à aider les étudiants à s'auto-réguler [2] et les approches par compétences (APC). L'auto-régulation de l'apprentissage peut être définie comme « un ensemble de processus par lesquels les sujets activent et maintiennent des cognitions, des affects et des conduites systématiquement orientées vers l'atteinte d'un but » [3].

Notre objectif consiste à évaluer l'utilisabilité et la pertinence de combiner deux outils, utilisables en présence ou à distance, et conçus pour soutenir les stratégies d'auto-régulation des étudiants en s'appuyant sur l'APC : l'accès à des exercices de renforcement des compétences et une visualisation du profil des compétences développées. Pour répondre à cet objectif nous avons observé l'utilisation et les avis de 181 étudiants de DUT 1^e année engagés dans une activité d'apprentissage de la programmation Shell et disposant de ces outils. Nous les avons croisés avec leur profil d'apprentissage auto-régulé (SRL). Cet article présente les résultats de cette étude et des recommandations pour la conception et la mise en œuvre de services articulant auto-régulation et APC.

2 Accompagner le travail en autonomie à l'université

Les capacités d'auto-régulation ont été identifiées comme un facteur de performance des apprenants. Elles leur permettent de contrôler et modifier de manière efficace leurs démarches pour atteindre des buts fixés [4]. Ils ne se bornent donc pas à les encourager à répondre aux *stimuli* de l'environnement, ils l'organisent. Cette démarche s'appuie sur différentes stratégies que l'apprenant peut mettre en œuvre : s'auto-évaluer, étudier les ressources de formation, demander de l'aide ou définir des objectifs et planifier son travail pour n'en citer que quelques-unes [5].

2.1 Soutien des stratégies SRL

Les outils proposés pour mesurer et promouvoir l'auto-régulation ont évolué en trois vagues [6]. Les premiers étaient fondés sur des données auto-rapportées fournies par

les étudiants à partir de questionnaires, les deuxièmes sur les traces d'interactions des étudiants avec les plateformes d'apprentissage. Dans ces deux cas, l'exploitation des données contribuait à fournir à l'apprenant ou aux enseignants des vues globales utiles pour faire des diagnostics et adapter leurs comportements. La troisième vague d'outils exploite aussi les traces d'interactions et ajoute des fonctionnalités stimulant la construction des compétences d'auto-régulation chez les apprenants. Les fonctionnalités les plus répandues dans les outils actuels (2008-2018) correspondent à des outils des vagues 2 et 3 et comprennent des tableaux de bord alimentés à partir des *Learning Analytics* (LA), des agents intelligents et des feedbacks/prompts personnalisés proposés aux apprenants. Cependant, des études [6, 7] ont pointé la nécessité d'aller plus loin dans l'exploitation des LA et le manque de modèles permettant la mise en œuvre des outils de type 3^e vague.

Schumacher et Ifenthaler [8] ont analysé l'intérêt pour le SRL de fonctionnalités différentes réalisables à partir des traces d'activité. Ils ont en parallèle évalué leur utilité et acceptabilité par les étudiants. La fonctionnalité jugée la plus utile correspond aux invitations à l'auto-évaluation avec retour immédiat qui peuvent soutenir les phases de réflexion et performance. Cette fonction est en effet jugée nécessaire pour évaluer l'état de connaissance et planifier les étapes suivantes. Les deux fonctionnalités jugées ensuite les plus utiles sont la recommandation de thématiques ou de cours à travailler en fonction, par exemple, du contenu sur lequel l'apprenant a déjà travaillé. Quant aux fonctionnalités plutôt orientées vers la planification ou le suivi de l'activité, elles ne sont pas jugées utiles ou acceptables.

Le fait de s'auto-évaluer avec retour immédiat nécessite de disposer d'activités ou de fonctionnalités permettant à l'étudiant de se situer par rapport aux objectifs de la formation. Les exercices sont un moyen de faire rapidement le diagnostic des compétences acquises, et l'affichage du profil de compétences est un moyen de se situer par rapport au référentiel de la formation.

2.2 Exerciceurs, visualisation de l'activité et SRL

Les générateurs d'exercices de type exerciceur visent deux fonctions principales : la mémorisation de performance et le développement de compétences, à travers un apprentissage par essais et erreurs fondé sur la répétition [9]. L'exerciceur conservant des traces de l'activité d'apprentissage et fournissant un *feedback* immédiat, il facilite la régulation chez l'apprenant et permet une réflexion explicite sur les compétences travaillées [10]. Depover *et al.* [11] considèrent l'exerciceur utile pour pratiquer et consolider des savoirs ou savoir-faire, ou tester un degré de maîtrise des savoirs (évaluation formative), ce qui en fait un outil complémentaire à une situation pédagogique existante. Cela suppose que l'enseignant identifie au préalable les notions de référence à travailler sur l'exerciceur. Sans cette condition, les auteurs signalent le risque pour les apprenants de ne pas faire ce lien dans un contexte de travail en autonomie. Le recours à un outil de visualisation des traces d'activité peut alors faciliter la mise en activité de l'apprenant.

De tels outils, qui prennent la forme de tableaux de bord, ont montré leur efficacité en particulier dans un contexte d'enseignement en ligne ou hybride [12]. La visualisation des objectifs fournit en effet aux apprenants des normes par rapport auxquelles se

positionner [13]. Cependant ces stratégies ont des limites en termes d'efficacité. Seuls les apprenants les plus âgés disposent des compétences métacognitives permettant de mettre en œuvre un processus d'auto-régulation de l'apprentissage [12]. De plus, ces stratégies doivent être conjointes avec des évaluations pour que les apprenants puissent vérifier l'atteinte des objectifs. Enfin, Sedrakyan *et al.* [13] suggèrent qu'il faut compléter les approches statistiques, actuellement largement appliquées, par des approches qui intègrent le niveau de préparation et les dépendances entre les objectifs et sous-objectifs d'apprentissage. Ce type de formalisation est en particulier décrit dans les démarches d'APC et ,dans un contexte de soutien au SRL, trois grandes catégories de visualisation, appelées profils [14], peuvent être proposées aux apprenants en exploitant cette déclinaison : les profils de planification, de suivi et d'adaptation.

3 Soutenir l'auto-régulation dans un cours d'initiation à la programmation Shell

Dans l'objectif de soutenir l'auto-régulation dans un cours proposé à des étudiants de 1^{re} année, plusieurs outils ont été mobilisés dans ce projet : un référentiel de compétences, un outil de génération d'exercices (ASKER) [15] et un outil de visualisation des profils de compétences des étudiants (module OLM)¹. Ils permettent de couvrir les trois étapes du cycle du SRL, correspondent aux préférences des apprenants [8] et respectent les recommandations de conception identifiées précédemment [13].

3.1 Les outils observés

Un référentiel de compétences « Programmer en langage Shell », regroupant 55 compétences a été conçu à partir d'un méta-modèle produit dans le cadre du projet COMPER. Ce référentiel permet de plus de décrire les ressources pédagogiques qui peuvent être exploitées pour travailler les compétences (comme des cours ou des exercices de la plateforme ASKER). Il permet enfin de fournir une structuration des données du profil qui sera affiché par le module OLM.

En effet, ASKER [15] est une plateforme permettant à un enseignant de créer et diffuser des exercices d'auto-évaluation pour ses apprenants. La réponse de l'apprenant est automatiquement et instantanément évaluée par le système. Il reçoit ainsi un retour d'information immédiat lui permettant de s'auto-évaluer. Dans le cadre du projet COMPER, les exercices proposés par ASKER ont été reliés aux compétences décrites dans le référentiel. Le module OLM permet de présenter aux apprenants des profils de compétences décrivant leur maîtrise des savoirs, savoir-faire et compétences en fonction des objectifs définis dans le référentiel. Les taux de maîtrise sont estimés en fon-

¹ En complément, un système de recommandation de ressources personnalisées a été conçu pour orienter automatiquement les étudiants vers les ressources, à partir d'une page de recommandations personnalisées. L'analyse de ce service n'est pas présentée dans cet article.

tion de la réussite des apprenants aux activités d'apprentissage proposées par l'enseignant, à partir d'une analyse des traces ou d'avis rapportés par les enseignants. Quatre vues de ces profils peuvent être proposées aux étudiants (cf. figure 1).

Ces outils proposent des moyens de soutenir les processus d'auto-régulation. ASKER est utilisé depuis plusieurs années en formation universitaire, mais le module OLM n'a jamais été testé en conditions écologiques. Les outils ont été mis en œuvre dans le cadre d'une initiation à la programmation Shell de DUT 1^{re} année réalisée à distance à partir de novembre 2020.

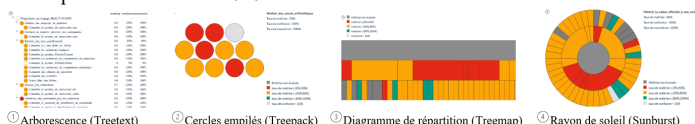


Fig. 1. Visualisations des profils de compétences dans le projet COMPER

3.2 Mise en œuvre des outils dans une formation à distance à la programmation

L'initiation à la programmation Shell est traditionnellement réalisée en présentiel en utilisant Lab4CE [16], une plateforme web de télé-TPs qui dote les apprenants d'une infrastructure de machines et de réseaux virtuels adaptée selon les besoins du TP. Du fait de la crise sanitaire, la formation a été réalisée à distance et les outils du projet COMPER mobilisés pour accompagner les étudiants dans leur travail en autonomie, lors des 5 séances de TP (une par semaine), suivies de 3 semaines de révision et d'un examen en présentiel. Avant le cours l'enseignant a défini le référentiel de compétences, les travaux pratiques à réaliser dans Lab4CE et les exercices à réaliser dans ASKER.

Pendant les séances de TP, l'enseignant accompagnait à distance les étudiants sur la réalisation du TP, corrigé en fin de semaine. Cette évaluation permettait de mettre à jour les profils de compétences. Entre les séances, une liste de ressources triées en fonction des compétences à travailler et donnant accès aux supports de cours et à des exercices ASKER était proposée aux étudiants. Cette liste s'incrémentait au cours du temps pour toujours afficher les ressources non travaillées par l'étudiant, qui pouvait ensuite filtrer les compétences qu'il jugeait prioritaires à travailler. L'accès à l'ensemble des outils était libre durant la phase d'entraînement (durant les TP) et la phase de révision.

3.3 Problématique

Outre le fait de vérifier la faisabilité de l'intégration des outils dans une formation existante, cette mise en œuvre dans le cadre du cours de programmation visait à répondre à différentes questions de recherche.

QRI : Est-ce que les outils sont utilisables par des étudiants de 1^{re} année de cursus universitaire guidés ponctuellement, mais travaillant globalement en autonomie ? Plus spécifiquement (**QRI.1**) est ce que les visualisations proposées dans le profil de compétences sont faciles à comprendre et à exploiter par les étudiants ?

QR2 : Quelle est l'utilité de cet accompagnement en termes d'auto-régulation ? Plus spécifiquement (**QR2.1**) est-il utile tout au long du cours ou juste pendant une phase donnée ? (**QR2.2**) Est-il utile à tous les étudiants ou bien juste à ceux ayant des problèmes à s'auto-réguler ? (**QR2.3**) Est-il utile pour construire des compétences SRL ?

4 Étude

4.1 Participants et tâche

181 participants ont été sollicités pour réaliser cette expérimentation. Tous sont des étudiants inscrits en première année d'une formation universitaire technologique (Diplôme Universitaire de Technologie français [DUT]), âgés de 17 à 19 ans, majoritairement de genre masculin (87% hommes pour 13% femmes). Les étudiants ont suivi la formation telle qu'elle a été décrite précédemment en travaillant seuls et à distance. Pour les besoins de l'expérimentation, ils ont été répartis de manière équilibrée en fonction de leur niveau académique dans 4 groupes dits expérimentaux (45 à 56 étudiants), ayant accès lors des deux premières semaines de la phase d'entraînement à une seule visualisation de profils et aux 4 visualisations ensuite.

4.2 Données collectées

Le premier questionnaire, administré en ligne et proposé au début et à la fin de l'expérimentation, est basé sur l'échelle EAREL [17] qui mesure le niveau d'apprentissage auto-régulé en ligne des étudiants et qui correspond ainsi à la situation vécue par les étudiants. L'échelle EAREL est composée de 24 items structurés selon 4 dimensions : la procrastination (PROC), le contrôle du contexte d'apprentissage (CTXT), les stratégies cognitives/métacognitives mobilisées pour l'apprentissage (COGN) et le soutien des pairs (PAIRS). Les étudiants se positionnent selon une échelle de Likert à 7 niveaux.

En fin d'expérimentation, les étudiants ont été invités à compléter aussi un questionnaire basé sur le *System Usability Scale* (SUS) [18]. Le SUS permet d'évaluer l'utilisabilité d'un dispositif à travers 10 items et a un score normalisé sur 100. L'utilisabilité d'un système est considérée comme défailante si le score est inférieur à 50. De 50 à 70 elle est correcte mais nécessite d'améliorer des défauts d'utilisabilité, entre 71 et 85 elle est bonne et au-delà de 85, elle est excellente.

5 Résultats

5.1 Capacité des étudiants à mettre en œuvre des stratégies SRL

Les profils de SRL des étudiants sont caractérisés par de bonnes capacités à organiser leur contexte de travail. Ils ont une forte propension à la procrastination et font preuve de peu de communication et de stratégies de travail. Les écarts-types et les coefficients de variation sont élevés, il y a donc une forte disparité dans la population (cf. table 1).

Par ailleurs, on n'observe que peu d'écarts entre les valeurs recueillies avant et après l'expérimentation.

Table 1. Score moyen (m), écart-type (σ) et coefficient de variation (CV) aux dimensions EAREL en début et fin d'expérimentation.

Pratiques SRL	CTXT	PAIRS	PROC	STRAT
m avant (n=94)	4,84	4,13	4,07	3,24
σ avant	1,00	1,56	1,40	1,33
CV avant	21 %	38 %	34 %	41 %
m après (n=59)	4,72	3,84	4,40	3,20
σ après	0,97	1,34	1,45	1,37
CV après	21 %	35 %	33 %	43 %

De manière à mieux comprendre la variété des profils SRL, nous avons procédé à une analyse par composante principale (ACP) et une partition par *K-means*. L'ACP a montré que les variables caractérisant les profils sont le fait de procrastiner et de mettre en œuvre des stratégies de travail (inversement proportionnels sur l'Axe 1 de la figure 1), et de communiquer avec ses pairs (Axe 2). Les comportements de contrôle du contexte ne sont pas discriminants. En complément, la méthode des *K-means* a permis d'identifier 4 groupes d'apprenants structurés selon leur niveau de mise en œuvre des stratégies SRL. Les *décrocheurs* (classe1-bleu, n=28) n'utilisent pas de stratégie, procrastinent et communiquent peu avec ses pairs. Les *suiveurs* (classe2-jaune, n=19) utilisent quelques stratégies, mais ont tendance à procrastiner et se resituer en communiquant avec leurs pairs. Les *appliqués solitaires* (classe3-vert, n=35) utilisent des stratégies, ne procrastinent pas et ne communiquent pas avec leurs pairs. Les *efficaces* (classe4-orange, n=23) utilisent des stratégies, ne procrastinent pas, et communiquent avec leurs pairs. La figure 2 montre la répartition des étudiants selon les groupes *K-means* et dans le plan vectoriel composé des axes 1 et 2 de l'ACP.

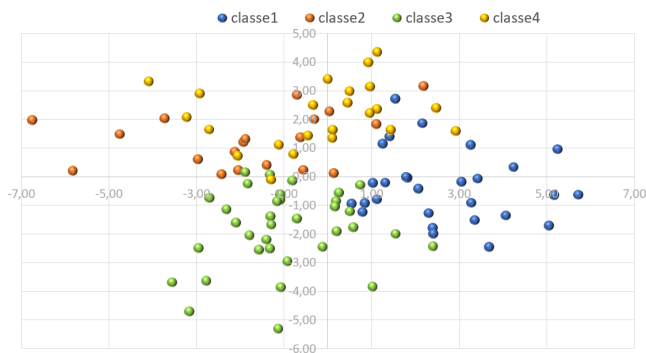


Fig. 2. Distribution des apprenants par profil selon le niveau de procrastination, d'utilisation de stratégies et de demande de soutien des pairs.

5.2 Utilisation des outils proposés

Analyse de l'utilisation d'ASKER. L'analyse des traces d'activité (cf. figure 3a) montre que la grande majorité des étudiants n'a pas utilisé ASKER (83/181). L'exercice est pourtant jugé utile (à 83,56 %) et facile à utiliser (à 67,12 %) (cf. figure 4a). La principale raison évoquée par les étudiants est un manque de temps. Les principaux utilisateurs sont ceux qui communiquent le moins avec leurs pairs soit la classe 1 (décrocheurs) pour 13 % et la classe 3 (appliqués) pour 14 %. Les étudiants ont utilisé ASKER à la fois dans les phases d'entraînement (pendant les TP) et en phase de révision (entre la fin des TP et l'examen) (cf. figure 5a). Des pics d'utilisation sont observés les jours correspondants aux séances de TP et une nouvelle dynamique se met en place à la fin des TP, pendant la période de révision avant l'examen final. Cette utilisation est motivée à la fois par la volonté pour les étudiants de se positionner dans leur parcours d'apprentissage (vérifier s'ils ont compris, savoir où ils en sont) que pour cibler et travailler des compétences spécifiques.

Analyse de l'utilisation du module OLM. L'analyse des traces d'activité montre qu'une proportion encore plus large d'étudiants n'a pas utilisé le module OLM (102/181, cf. figure 3b). L'outil de visualisation est en effet jugé non utile (à 45,83 %) et difficile à utiliser (à 55,55 %) (cf. figure 5b). De plus, les étudiants ont indiqué ne pas connaître ce service et se fier à d'autres ressources, comme les retours directs de l'enseignant, pour évaluer leur progression. Quand il a été utilisé, c'était faiblement (à 92 %, cf. figure 3b). Certains ne s'y sont connectés qu'une à cinq fois, les jours du TP, et seuls 10 d'entre eux l'ont utilisé lors des TP et une fois les séances de TP terminées. Sachant que l'évaluation du TP de la semaine précédente effectuée par les enseignants était intégrée au profil de compétence, les étudiants ont utilisé l'outil pour voir leur progression globale sur le cours. Pour mieux comprendre ce qui a freiné l'utilisation de ce service, nous en avons fait une analyse approfondie.



Fig. 3. (a) Nombre d'utilisations de ASKER et (b) du profil de compétences

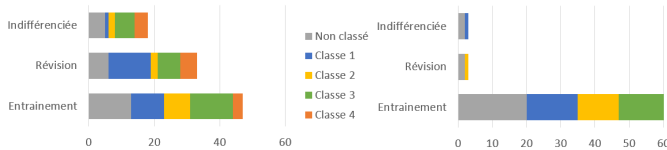


Fig. 4. (a) Période d'utilisation de ASKER et (b) du profil de compétences

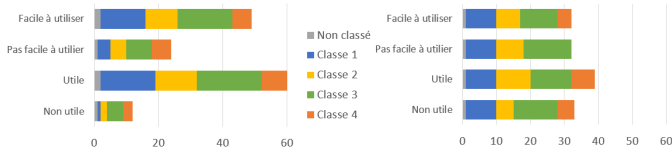


Fig. 5. (a) Avis sur ASKER et (b) avis sur le profil de compétences

5.3 Analyse des difficultés rencontrées avec le service de visualisation des profils de compétences

Analyse de l'utilisabilité. Le score SUS obtenu pour le profil de compétences est de 59,49/100. Le profil de compétences peut donc être considéré comme ayant une utilisabilité correcte mais limitée. Ces limites sont la nécessité de recourir au support technique (score moyen de 3,167/4) et des difficultés à utiliser le module (items « J'ai trouvé qu'il y a trop d'incohérences dans ce service », score moyen de 3,056/4, et « je trouve ce service très lourd à utiliser », score moyen de 3,022/4). Pour aller plus loin dans la compréhension des difficultés rencontrées, nous avons aussi pris en compte leurs traces d'activité sur le module OLM. Elles montrent pourtant que seule une minorité des traces concerne des erreurs techniques rencontrées par les étudiants, soit à l'affichage d'une compétence, soit au changement de visualisation. Le type de difficultés rencontrées concerne donc plutôt le type de visualisation que des problèmes techniques.

Type de visualisation. D'après les traces d'activité, la présentation arborescente est la plus consultée par les étudiants, 46% des actions sur le profil sont avec cette vue. Quand les étudiants ont eu le choix entre les 4 vues, la vue « rayon de soleil » a systématiquement été remplacée par une autre, principalement le *treemap* (à 52%). Les avis exprimés dans le questionnaire le confirment et montrent que la vue préférée est l'arborescence (pour 37 des étudiants). Pour eux, cette visualisation est facile à lire et à comprendre, elle permet de saisir rapidement les compétences présentées. Ils sont minoritaires (5 sur 47) à préférer la présentation *treemap*. Quand ils disent préférer les autres visualisations, c'est parce qu'il s'agit des premières auxquelles ils ont eu accès.

6 Discussion

6.1 Utilisabilité et mise en œuvre des outils

Le premier objectif de l'étude était d'identifier l'utilisabilité des services proposés par des étudiants de 1^e année (QRI). Dans l'ensemble, les outils du service COMPER sont faciles à utiliser. Dans le détail, l'exerciseur ASKER a été jugé plus utilisable et a été plus utilisé que le module OLM. Le profil de compétences a obtenu un score d'utilisabilité moyen (59/100) qui témoigne des difficultés de compréhension rencontrées. Pourtant, son utilisation, qui n'a pas été totalement abandonnée, et la prédominance de la visualisation sous la forme d'une arborescence, jugée facile à lire et comprendre,

invitent à poursuivre son développement. En effet, les résultats confirment le lien entre utilisabilité perçue et utilité réelle des services : les étudiants qui ont évalué le plus favorablement les outils sont aussi ceux qui en ont eu l'utilisation la plus importante et la plus longue dans la durée. Une intégration plus complète du module OLM dans ASKER est une piste intéressante à suivre pour faciliter l'accès au module et limiter la perception qu'ASKER représente une charge de travail supplémentaire au TP.

On observe des différences entre étudiants dans la consultation du profil de compétences (**QR1.1**) : si la présentation sous la forme d'une arborescence est préférée par beaucoup, d'autres ont privilégié la vue en *treemap*, cette visualisation se rapprochant de celles auxquelles ils disent être habitués. L'expérience des étudiants en matière de visualisation (dans des contextes de formation ou non) devra donc être explorée dans des expérimentations futures.

6.2 Utilité des outils

Le second objectif de l'étude consistait à comprendre l'utilité de l'accompagnement proposé par COMPER en termes d'auto-régulation (**QR2**). Trois dynamiques ont pu être observées dans l'utilisation du dispositif COMPER, en fonction de périodes d'utilisation distinctes (**QR2.1**) : la moitié des étudiants a utilisé les services pendant les TP, pour s'entraîner, un quart les ont utilisés surtout en vue de l'examen final, pour réviser, et le dernier quart les ont utilisés de façon indifférenciée. L'utilisation à des fins de révision va dans le sens d'une régulation de la charge de travail de l'étudiant, et du choix de faire les exercices de renforcement une fois les TP finis. Par ailleurs, si la principale raison motivant l'utilisation des services est d'y trouver une aide pour se repérer dans le parcours d'apprentissage, les résultats du questionnaire et de l'analyse des traces montrent que ce repérage ne s'est pas toujours personnalisé. La plupart des étudiants ont travaillé toutes les compétences abordées pendant les TP plutôt que d'en cibler certaines.

Le profil SRL des étudiants a joué un rôle important dans l'utilisation des services (**QR2.2**). Compte tenu des différences observées (cf. figure 6), on peut en conclure que la mise à disposition d'exercices de renforcement semble principalement venir pallier le manque de stratégies qui caractérise les étudiants « décrocheurs ». La visualisation des compétences n'est pour l'instant pas exploitable par eux. Les étudiants « suiveurs » et « efficaces » se sont moins emparés des services COMPER. Les « efficaces », ayant déjà des stratégies SRL, jugent ces services comme peu utiles pour une aide complémentaire. L'utilisation des profils de compétences par les « suiveurs » en période d'entraînement laisse entendre qu'elle n'est liée qu'à l'intégration du service pendant les séances de TP. On observe pour ces deux classes une utilisation d'ASKER surtout en période de révision. L'approche de l'examen est un facteur motivationnel externe. Enfin, la classe des « appliqués solitaires » se distingue par une utilisation majoritaire des services pendant la période d'entraînement et pas en phase de révision. Comme pour les « efficaces », ils ont mobilisé leurs propres stratégies de SRL. Des entretiens d'approfondissement permettront de comprendre comment ces outils, ou leurs exploitations, peuvent être repensés pour amorcer la mise en œuvre de stratégies SRL et compléter des stratégies SRL existantes. La **QR2.2** a donc partiellement été traitée au cours de

cette étude, le dispositif COMPER semblant répondre aux besoins des étudiants isolés et qui ne mettent pas ou peu de stratégies SRL en œuvre.

	« Décrocheurs »	« Suiveurs »	« Appliqués solitaires »	« Efficaces »
ASKER	Utilisation importante en période d'entraînement	Utilisation modérée en période de révision	Utilisation polarisée (faible ou intense) en période d'entraînement	Utilisation faible à modérée en période de révision
OLM	Utilisation faible en période d'entraînement Score SUS parmi les moins élevés	Utilisation faible en période d'entraînement Score SUS parmi les plus élevés	Utilisation modérée en période d'entraînement Score SUS correct	Utilisation faible à modérée en période d'entraînement Score SUS parmi les moins élevés

Fig. 6. Principales caractéristiques de l'utilisation des services COMPER par les profils SRL

Le dernier élément traité concerne la manière dont les compétences d'auto-régulation ont pu être construites par les étudiants en expérimentant le scénario pédagogique présenté (QR2.3). L'analyse des écarts entre les échelles EAREL1 et EAREL2 montre que 3 des 4 construits de l'échelle ont diminué, le contrôle du contexte, du soutien des pairs et de la mobilisation de stratégies SRL, en moyenne de $-0,472$. À l'inverse, la perception de leur procrastination a augmenté ($+0,295$), ce qui laisse surtout penser que les étudiants ont pris conscience de la manière dont ils travaillent et apprennent. Les effets de la période sanitaire n'ayant pas été mesurés dans cette étude, il est aussi possible que cette baisse aille de pair avec leur état général.

7 Conclusion

Cet article décrit l'expérimentation, par des étudiants de DUT1 en Informatique, de services articulant approche par compétences et auto-régulation et conçus pour accompagner le travail en autonomie : l'accès à des exercices de renforcement et à un profil de visualisation de compétences. L'expérimentation a montré qu'ils étaient globalement utiles et utilisables, mais de manières différentes en fonction des profils SRL des étudiants. Les services viennent principalement répondre aux besoins des étudiants qui ont peu développé de stratégies SRL. En fournissant à ces étudiants des outils pour planifier, contrôler et évaluer leurs apprentissages, ils ont su les mobiliser en phases d'entraînement et de révision. L'étude montre aussi que l'ensemble des services a moins été utilisé par les autres profils SRL. L'exploitation des données issues des questionnaires et des traces d'activité des outils ont permis de confirmer la pertinence de la démarche globale auprès d'étudiants en situation de travail en autonomie, mais des entretiens complémentaires sont prévus afin d'améliorer la conception de chaque outil et leur exploitation en formation.

Remerciements. Cette expérimentation et la conception des outils ont été réalisées dans le cadre du projet ANR-18-CE38-0012.

References

1. Panadero, E., García-Pérez, D., Fernández-Ruiz, J.: A Transitional Year Level to Higher Education: Challenges, Experiences and Self-regulatory Strategies during the Final Year of the University Preparatory Level. *Estud. Sobre Educ.* 39, 109–133 (2020).
2. Zimmerman, B.J.: A social-cognitive view of self-regulated academic learning. *J. Educ. Psychol.* 81, 329–339 (1989).
3. Schunk, D.: Self-regulation of self-efficacy and attributions in academic settings. In: *Self-regulation of learning and performance*. pp. 75–99. Hillsdale, NJ (1994).
4. Pintrich, P.R.: The role of goal orientation in self-regulated learning. In: *Handbook of self-regulation*. pp. 451–502. Academic Press, San Diego (2000).
5. Zimmerman, B.J., Pons, M.M.: Development of a Structured Interview for Assessing Student Use of Self-Regulated Learning Strategies. *Am. Educ. Res. J.* 23, 614–628 (1986).
6. Panadero, E.: A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Front. Psychol.* 8, 422 (2017).
7. Araka, E., Maina, E., Gitonga, R.: Research trends in measurement and intervention tools for self-regulated learning for e-learning environments—systematic review (2008–2018). *Res. Pract. Technol. Enhanc. Learn.* 15, 6 (2020).
8. Schumacher, C., Ifenthaler, D.: Features students really expect from LA. *Comput. Hum. Behav.* 78, 397–407 (2018).
9. Lemerrier, C., Tricot, A., Chénierie, I., Marty Dessus, D.: Quels apprentissages sont-ils possibles avec des exercices multimédia en classe ? (2001).
10. Steffens, K.: Self-Regulated Learning in Technology-Enhanced Learning Environments: lessons of a European peer review. *Eur. J. Educ.* 41, 353–379 (2006).
11. Depover, C., Karsenti, T., Komis, V.: Les environnements et les logiciels conçus pour l'enseignement. In: *Enseigner avec les technologies* (2007).
12. Kizilcec, R.F., Pérez-Sanagustín, M.: Self-regulated learning strategies predict learner behavior and goal attainment in MOOC. *Comput. Educ.* 104, 18–33 (2017).
13. Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S.: Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Comput. Hum. Behav.* 107, 105512 (2020).
14. Sedrakyan, G., Mannens, E., Verbert, K.: Guiding the choice of learning dashboard visualizations: Linking dashboard design and data visualization concepts. *J. Comput. Lang.* 50, 19–38 (2019).
15. Lefevre, M., Guin, N., Cablé, B., Buffa, B.: ASKER : un outil auteur pour la création d'exercices d'auto-évaluation. In: *Conférence EIAH 2015*. Agadir (2015).
16. Broisin, J., Venant, R., Vidal, P.: Awareness and Reflection in Virtual and Remote Laboratories: the case of Computer Education. *Int. J. Technol. Enhanc. Learn.* 9, 1 (2017).
17. Cosnefroy, L., Fenouillet, F., Heutte, J.: Construction et validation de l'Échelle d'auto-régulation des apprentissages en ligne (EAREL). *Can. J. Behav. Sci. Rev. Can. Sci. Comport.* 52, 255–260 (2020).
18. Brooke, J.: SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189, (1996).

Génération d'un profil dynamique du stress pour l'entraînement à la gestion de situations de crise

Luca Pelissero-Witoslawski, Domitile Lourdeaux et Dominique Lenne

Alliance Sorbonne Université, Université de technologie de Compiègne, CNRS, Heudiasyc UMR 7253 57 avenue de Landshut – 60203 Compiègne Cedex, France
luca.pelissero-witoslawski@hds.utc.fr; domitile.lourdeaux@hds.utc.fr;
dominique.lenne@hds.utc.fr

Résumé. Nos travaux portent sur l'entraînement à la gestion de situations stressantes en environnement virtuel. Nous faisons l'hypothèse que l'apprenant peut améliorer la gestion de ces situations en étant confronté à une grande variété de situations. Nous souhaitons générer des situations personnalisées mettant en jeu différents stresseurs. Nous nous intéressons pour cela à la représentation et au diagnostic des capacités de l'apprenant à faire face à ces stresseurs. Le diagnostic d'un profil de stress est réalisé à partir de capteurs de stress qui fournissent des mesures entachées d'incertitudes. Par ailleurs, les situations que nous générons sont des situations complexes mettant en jeu des stresseurs pouvant influencer différemment chaque apprenant. Le diagnostic est donc empreint de nombreuses incertitudes. Pour prendre en compte ces incertitudes, nous nous appuyons sur le modèle des croyances transférables, et nous proposons de représenter les stresseurs sous forme d'une taxonomie paramétrable par le formateur afin de lui permettre d'explorer différents niveaux de granularité.

Mots-clé: Stress · Profil · Fonctions de croyance · Génération dynamique · Scénario

Abstract.

Our work focuses on improving the management of stressful situations using virtual environments. We hypothesize that learners can improve how they deal with these situations by being confronted with a wide variety of scenarios. We want to create customized situations using different stressors. Which is why we are interested in the representation and diagnosis of the learner's ability to cope with said stressors. The diagnosis of a stress profil is carried out using stress sensors which provide measurements. However these measurement aren't quite spotless and uncertainties are present. Moreover, the situations we generate are complex and involve stressors that can impact each learner differently. Hence, the diagnosis is also impacted by many uncertainties. In order to take these uncertainties into account, we rely on the transferable belief model, and we came up with stressors depicted in the form of a taxonomy that can be configured by the trainer. This will allow him to explore different levels of granularity.

Keywords: Stress · Profile · Belief functions · Dynamic generation · Scenario

1 Introduction

Les entraînements en environnement virtuel permettent la mise en place de situations complexes qu'il serait difficile de reproduire dans le cadre d'un entraînement classique, comme les situations de crise (attaque terroriste, situation de conflit armé, catastrophe naturelle). De plus, la réalité virtuelle pour l'entraînement permet de confronter l'apprenant à une grande diversité de situations.

Les situations de crise sont particulièrement stressantes pour les personnes devant oeuvrer dans des conditions réelles similaires. Nous faisons l'hypothèse que multiplier les confrontations à des situations de plus en plus complexes et/ou stressantes permettrait aux apprenants de mieux réguler leur stress.

Cependant, chaque personne est unique. Une situation de développement potentiel pour un apprenant pourrait se révéler être une situation difficilement surmontable pour un autre, le mettant par conséquent en échec. De la même manière, toutes les personnes ne sont pas stressées par les mêmes stimuli.

Nous proposons un système permettant de définir un profil dynamique et personnalisé pour chaque apprenant. Ce profil, reflétant les capacités de l'apprenant à gérer des éléments stressants, est généré grâce à des données issues de capteurs physiologiques. Cependant, ces données sont empreintes d'incertitudes. En plus de l'incertitude liée aux erreurs potentielles de mesure, les signaux analysables en temps réel permettant de connaître l'état de stress d'une personne sont influencés par de nombreux facteurs, phénomènes ou actions que l'on retrouve fréquemment lors de l'utilisation d'un casque de réalité virtuelle (fatigue, mouvement, mal du simulateur, etc.). Il est également difficile d'isoler le facteur à l'origine de ce stress, les causes pouvant même se révéler être multiples. C'est pourquoi ce système s'appuie sur la théorie des croyances transférables ainsi que sur un modèle de connaissances pour permettre d'effectuer des raisonnements sémantiques dans l'incertain.

À cela s'ajoutent des besoins spécifiques de travailler à différents niveaux de granularité pour chaque formateur, et de prendre en compte leurs différentes stratégies pour explorer et travailler des compétences, ce que notre système propose au travers d'un transfert de croyance entre ces différents niveaux.

Ce profil est utilisé conjointement à un planificateur dynamique s'appuyant sur des théories narratives pour générer des situations d'entraînements personnalisés. Le profil est utilisé pour déterminer quels stressseurs seront joués dans l'environnement virtuel. Nous nous plaçons alors dans le cadre du courant constructiviste.

Ces scénarios devant être adaptés à chaque apprenant, il est important de pouvoir modifier dynamiquement les situations que nous lui présentons lors d'une session d'entraînement. Nous souhaitons par ce procédé diagnostiquer les effets d'une situation proposée durant son exécution, mais également pour les futures sessions. Nous nous intéressons donc à une adaptation du scénario aussi bien au niveau microscénaristique que mésoscénaristique.

2 Stress et apprentissage

2.1 Définition du stress

Le stress fait l'objet de nombreuses études. Hans Selye décrit le stress comme "l'ensemble des modifications qui permettent à un organisme de supporter les conséquences d'un traumatisme naturel ou opératoire." [9]. Ses recherches le conduisent également à définir la notion d'*eustress*, signifiant "bon stress". Selon sa théorie, l'eustress serait à l'origine d'un sentiment de satisfaction et de bien-être. En opposition, il décrit également l'existence d'un mauvais stress.

Ces notions de bon stress et de mauvais stress seront reprises quelques années plus tard par Susan Folkman et Richard Lazarus. Selon leur définition (Définition 1.), cette transaction est mesurable par une estimation de l'effet produit par un événement sur l'individu ainsi que par la façon dont l'individu va faire face à cette situation mettant en danger son bien-être. Ces situations seront qualifiées de situations stressantes.

Définition 1. *Le stress est "une transaction particulière entre un individu et une situation dans laquelle celle-ci est évaluée comme débordant ses ressources et pouvant mettre en danger son bien-être."* [4]

Ce "faire-face" est une notion importante dans le travail des deux chercheurs. Ils la décrivent comme un mélange entre deux mécanismes distincts, l'un comportemental et l'autre cognitif. Ils développeront cette notion et la nommeront *coping* (voir Définition 2.)

Définition 2. *Le coping est "l'ensemble des efforts déployés volontairement par un individu pour faire face à une situation qu'il évalue comme stressante, impliquant que cette situation est perçue comme difficile à surmonter et menaçante pour son bien-être."* [4]

De plus, Lazarus fait le lien entre stress et émotion, et va ainsi montrer qu'un état émotionnel positif facilite les capacités de prise de décisions, même les plus complexes, et permet d'augmenter la capacité d'assimilation de l'information. Le stress est donc un élément essentiel dans les capacités d'assimilation et d'apprentissage d'une personne.

2.2 Apprentissage en situation stressante

Comme avancé par Selye puis par Lazarus, le stress et l'apprentissage sont étroitement liés. Le cortisol engendré par le stress a un impact multiple sur les différentes phases de l'apprentissage et sur le fonctionnement de la mémoire [6]. Un taux moyen de cortisol est bénéfique pour l'apprentissage, augmentant les facultés de mémorisation. Cependant, si ce taux est trop important, le cortisol va avoir un effet néfaste sur les capacités de l'apprenant durant les trois étapes fondamentales de la mémorisation : l'étape d'encodage, l'étape de consolidation et l'étape de remémoration. Cette dernière est l'étape sollicitée lors

de la réalisation d'un entraînement, et celle qui nous intéresse particulièrement dans nos travaux.

Un contrôle sur le stress d'un individu pourrait lui permettre d'atteindre un état d'apprentissage optimal. C'est pourquoi dans le cadre d'entraînements en environnement virtuel, et plus particulièrement en situation stressante, il est important de ne pas engendrer un stress trop important, qui aurait un effet non désiré sur l'apprentissage de l'apprenant, mais qui pourrait aussi avoir un effet sur sa santé.

3 Contribution

Nous avons élaboré un système permettant de générer dynamiquement un profil du stress d'un apprenant, dans le but d'être utilisé dans le cadre de l'entraînement à la gestion du stress en situation de crise grâce à l'utilisation conjointe d'un second système de scénarisation dynamique et adaptatif. Ce profil est établi à partir de données physiologiques issues de capteurs et construit selon une taxonomie de stressseurs.

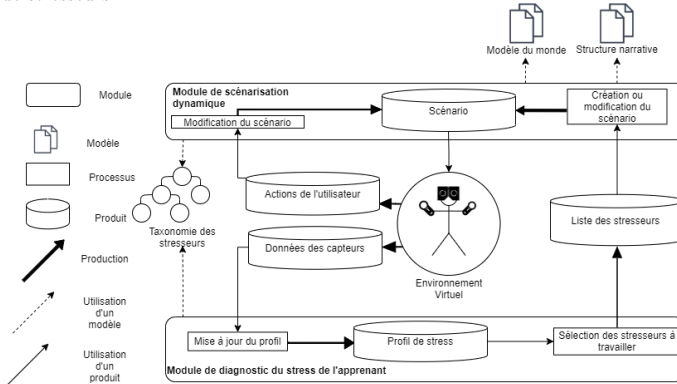


Fig. 1. Architecture générale des systèmes de diagnostic du stress et de scénarisation dynamique

3.1 Données

Afin d'établir le profil, le système a besoin d'une source d'informations sur les réactions physiques de l'utilisateur lorsque celui-ci est soumis à un stressseur. Il existe de nombreux signaux analysables. Cependant, le système doit répondre à des contraintes fortes en terme de vitesse d'exécution, et doit opérer l'ensemble de ses actions en moins d'une seconde, valeur correspondant au temps interactif.

Nous devons également exclure les capteurs qui seraient difficilement utilisables avec un casque de réalité virtuelle.

Deux capteurs répondent à nos critères. Ils nous permettent d'obtenir respectivement des informations sur l'activité électrodermale (EDA) et sur l'activité électrique du cœur (ECG) [8], tous deux variant lors de l'exposition à un stresser et régis par le système nerveux sympathique. L'analyse de la variation du potentiel électrique, de la résistance électrique ainsi que de la variabilité de la fréquence cardiaque nous permet dès lors de quantifier pour un individu la façon dont son corps va répondre à un élément stressant. Nous considérons que moins la réponse au stimulus est importante, plus l'apprenant a réussi à faire face à la situation stressante.

Les signaux issus de l'analyse des EDA et des ECG peuvent être parasités par d'autres stimuli que des stressers. Par exemple, l'activité physique du corps entraîne une hausse de l'activité cardiaque et électrodermale. Le phénomène du mal du simulateur, trouble se manifestant lors d'une situation présentant une discordance entre le système vestibulaire et la perception visuelle, peut également agir sur les signaux que nous analysons. Ces problèmes, très fréquents lors de l'utilisation d'un dispositif de réalité virtuelle, créent de l'incertitude dans nos mesures. De plus, les capteurs peuvent produire des erreurs de mesure, selon un taux d'erreur que nous devons prendre en compte. C'est pourquoi nous avons recours à la fusion de données multi-capteurs afin d'améliorer la fiabilité de nos données.

D'autres problèmes peuvent se poser lors du diagnostic de l'apprenant au cours du temps et des différents entraînements auxquels il prend part. Il est notamment possible que les analyses de données soient totalement contradictoires entre plusieurs entraînements. Il est également possible que l'analyse des données soit différentes entre deux entraînements présentant un scénario identique. La mesure en temps interactif du stress est encore à ce jour un problème de recherche. Chaque personne ayant des signaux physiologiques différents, il n'existe pas de seuil universel permettant de définir l'état de stress. La durée de réponse à un stimulus est également variable entre les personnes, et le temps pour retrouver un niveau basal dans les signaux que nous analysons peut être assez long. Toutes ces contraintes apportent de l'incertitude supplémentaire à nos mesures.

Enfin, il est difficilement envisageable qu'un apprenant puisse réaliser un entraînement sur chaque situation qu'il est possible de voir. Nous devons donc également prendre en compte dans la réalisation de notre diagnostic ce manque d'informations existant sur les situations qui ne seront jamais explorées.

Pour prendre en compte ces problèmes lors de notre diagnostic, nous devons utiliser un modèle mathématique permettant de caractériser la connaissance issue de l'analyse des données en provenance de plusieurs sources d'informations, tout en permettant de matérialiser des incertitudes ainsi que des contradictions et le manque d'informations.

3.2 Modélisation de croyances incertaines

Nous avons examiné plusieurs théories dans la littérature permettant de modéliser des croyances empreintes d'incertitudes : théorie des ensembles; théorie des ensembles flous [14]; théorie des possibilités [15,3]; théorie des réseaux bayésiens [7]; théorie des fonctions de croyance [10]; théorie des croyances transférables [11].

Notre choix s'est porté sur la théorie des croyances transférables. Cette théorie englobe l'ensemble des autres théories que nous avons énumérées, et permet également de représenter des conflits entre deux sources d'informations contradictoires, ainsi que le manque de connaissance. Ce cadre formel propose également des outils afin de fusionner des informations de plusieurs sources, ainsi que des outils permettant d'effectuer des inférences sur les valeurs d'une variable présentant un manque de connaissance (Voir 3.4 Profil).

3.3 Représentation des stressseurs

Dougall et Baum [2] proposent un modèle permettant de catégoriser des stressseurs, représentables sous forme de taxonomie. Selon cette catégorisation, ces stressseurs peuvent être organisationnels (propres à un cadre ou une activité) ou aigus (éphémères et soudains, pouvant ne pas être liés à l'individu subissant le stressseur).

Notre objectif étant de confronter l'apprenant à des stressseurs proches de ceux auxquels il sait faire face selon notre système, nous souhaitons utiliser cette taxonomie pour définir des distances sémantiques entre les stressseurs. La prise en main et la modification d'une taxonomie étant aisées, nous mettons l'expert ou le formateur au coeur de l'élaboration de la hiérarchie.

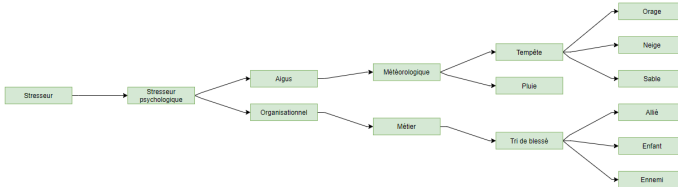


Fig. 2. Représentation taxonomique de stressseurs aigus et métiers

Nous proposons un ordonnancement entre les stressseurs. Nous pourrions intuitivement penser qu'il est plus complexe de faire face à certains stressseurs. Cependant, nous considérons qu'il est impossible de trouver un ordonnancement universel des stressseurs à cause du caractère unique de chaque apprenant.

Exemple 1. Intuitivement, on peut penser que les cris d'un adulte sont moins stressants que les cris d'un enfant. Mais selon le contexte, cette supposition peut être altérée. Les cris d'un adulte pourraient chez un apprenant particulier lui rappeler un épisode traumatisant, engendrant chez lui un stress plus important.

Pour lier les concepts de la taxonomie entre eux, nous instaurons une mesure de similarité sémantique entre deux nœuds. Nous nous appuyons sur la mesure de la similarité sémantique de Wu et Palmer [12], qui a l'avantage de ne s'appuyer que sur la distance entre les nœuds, sans prendre en compte l'information liée aux nœuds. On note $h(X)$ la distance séparant le nœud X de la racine de la taxonomie ($h(\text{racine}) = 1$). Soient N et M deux nœuds de la taxonomie. On note P le premier nœud parent commun à N et M ($N \neq M$). On note sim la similarité sémantique entre N et M tel que :

$$\text{sim}(N, M) = \frac{2 \times h(P)}{h(N) + h(M)}$$

La similarité sémantique est d'autant plus grande (i.e. plus proche de 1) que les nœuds sont profonds et proche dans la taxonomie. Ainsi, la structure de la taxonomie à une grande importance dès son élaboration, et doit être créée avec la prise en compte du calcul à réaliser. Cette similarité sémantique est utilisée dans la propagation de croyance au sein du profil dynamique de l'apprenant (voir 3.4 Profil).

3.4 Profil

Nous nous intéressons à l'entraînement des compétences métier au travers de situations stressantes. L'un de nos objectifs est de proposer des scénarios d'entraînement personnalisés dans lesquels les situations d'entraînement mettent en jeu des compétences sous conditions de stress. Comme la capacité à gérer le stress est propre à chaque apprenant, notre second objectif est d'étendre progressivement la zone de développement du profil de chaque apprenant, généré par notre système.

Nous pouvons donc assimiler nos objectifs à la représentation et à l'expansion de la Zone Proximale de Développement (ZPD) de Vygotsky (voir Définition 3.).

Définition 3. *"Distance entre le niveau de développement actuel (ZDA), tel qu'on pourrait le déterminer par les capacités de l'enfant à résoudre seul des problèmes, et le niveau de développement potentiel, tel qu'on pourrait le déterminer à travers la résolution de problèmes par cet enfant, lorsqu'il est aidé par des adultes ou collabore avec des pairs initiés."*, [13]

Au travers du profil dynamique du stress, nous souhaitons pouvoir rendre compte du niveau de développement présent des capacités d'un apprenant à faire face à des stressseurs, de son niveau de développement potentiel, mais également pouvoir conseiller au formateur de nouveaux stressseurs à travailler. De plus, nous voulons aussi laisser au formateur un certain contrôle sur le paramétrage des entraînements, ainsi que sur la visualisation du profil de l'apprenant.

Carpentier propose une représentation de la ZPD nommée espace-zpd. Cette représentation s'appuie elle-même sur le modèle des croyances transférables [1] afin d'établir un diagnostic de l'apprenant.

Le domaine d'apprentissage est représenté grâce à l'utilisation de classes de situations (voir Définition 4.)

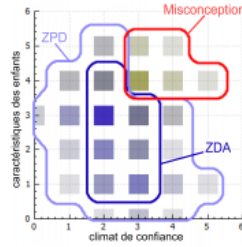


Fig. 3. Représentation de la ZPD selon Carpentier

Définition 4. Une classe de situations est une "représentation abstraite d'un ensemble de situations présentant des propriétés communes nommées descripteurs. Un descripteur est un élément permettant de décrire un variant". On associe à la classe de situations S le vecteur σ tel que :

$$\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$$

avec σ_i la valeur correspondante au descripteur i .

Ces descripteurs permettent de différencier les classes de situations entre elles.

En appliquant la théorie des fonctions de croyance, Carpentier associe à chaque classe de situations S_i une distribution de masse de croyance m selon un cadre de discernement Ω représentant l'ensemble des hypothèses qui sont émises sur l'état des compétences de l'apprenant.

Nous proposons d'étendre ce modèle¹ pour diagnostiquer l'évolution de l'apprenant à différents niveaux hiérarchiques de compétence, en parallèle. Ce diagnostic s'appuie en partie sur les données physiologiques issues de nos capteurs. Nous voulons également que le système soit paramétrable selon les besoins du formateur.

Espace-zpd par stresser

Nous voulons représenter la capacité de l'apprenant à faire face à chaque stresser qui compose notre taxonomie. L'objectif est de permettre à l'apprenant de travailler ses compétences déjà acquises en lui proposant des situations stressantes les illustrant.

À chaque classe de situations est associée une distribution de masse de croyance $(a(S_i), d(S_i), c(S_i), i(S_i))$. Ces masses sont calculées à partir des données

¹ Voir Calcul des masses de croyances : <https://tinyurl.com/dtfebhte>

des capteurs physiologiques que l'on utilise ainsi que les observables permettant de déterminer la réussite de la tâche confiée. A l'initialisation, $i(S_i) = 1$ car il n'y a aucune information exploitée, ce qui correspond à une ignorance.

Dans notre représentation de la ZPD sous forme d'espace-zpd, un axe (ou dimension) peut représenter :

- Une compétence métier. Les descripteurs de cet axe correspondent à des variables didactiques ou des habiletés.
- Une feuille de la taxonomie des stressseurs, c'est-à-dire un stressseur. Les descripteurs de cet axe sont les trois niveaux d'intensité possibles d'un stressseur (faible, moyen et intense). Ils sont ordonnés.
- Un nœud de la taxonomie qui n'est pas une feuille, c'est-à-dire un ensemble de stressseurs. Les descripteurs de cet axe sont les enfants du nœud. Ils sont non-ordonnés.

Notre objectif est d'entraîner des compétences métier au travers de situations stressantes.

Chaque espace-zpd contient n dimensions ($n \in N^*$). Une dimension représente soit une compétence métier, soit un stressseur, soit un ensemble de stressseurs. Chaque élément de l'espace-zpd représente une classe de situations, dont la couleur varie en fonction de sa masse de croyance. Il est ainsi possible de représenter autant d'espaces-zpd que de combinaisons possibles de noeuds de notre taxonomie.

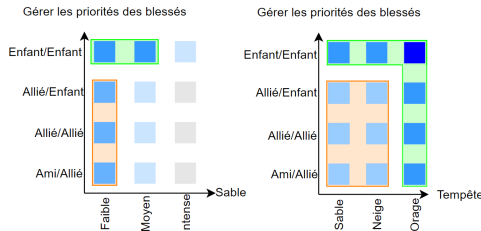


Fig. 4. Exemple d'espace-zpd à deux dimensions généré à partir d'une feuille (Sable) ou d'un nœud non-feuille (Tempête) de la taxonomie et d'une compétence métier.

Ces croyances sur une classe de situations sont par la suite étendues aux autres classes de situations proches, et permettent d'apporter de l'information sur les concepts proches de la taxonomie.

3.5 Maîtriser de nouvelles situations

Le système propose avant chaque session d'entraînement une situation permettant d'étendre la ZDA et la ZPD de l'apprenant.

Pour chaque situation S , si $a(S) > 0.85$ et $d(S) < 0.1$, alors la situation intègre la ZDA de l'apprenant : il maîtrise la situation. Si $a(S) > 0.65$ alors la

situation intègre la ZPD : la compétence est proche d'être acquise, soit d'après des entraînements, soit d'après la propagation des croyances depuis des situations proches. Dans la Figure 4 la ZDA est représentée par la zone verte, tandis que la ZPD est représentée en orange. Le gradient de couleur des classes de situations varie ici en fonction de $a(S)$.

Nous nous intéressons aux algorithmes permettant à notre système de proposer la prochaine situation à mettre en place au sein d'un entraînement.

L'objectif est de maximiser l'expansion de la ZDA (et intrinsèquement de la ZPD) en un minimum d'entraînements.

Nous devons utiliser un algorithme parcourant l'ensemble de nos classes de situations et sélectionnant celle qui est la plus à même de faire l'objet d'un entraînement. Nous avons étudié plusieurs algorithmes de recherches applicables à une structure arborescente répondant à ces besoins et avons retenu les algorithmes suivants² : sélection arbitraire dans la ZPD; algorithme glouton; algorithme des k plus proches voisins; algorithme des k plus proches voisins avec critère de maximisation. Leur complexité algorithmique simple et leur capacité à être facilement paramétrable par le formateur ont été les critères sur lesquels nous avons retenu ces algorithmes.

3.6 Filtre

En plus de proposer au formateur plusieurs stratégies de sélection du prochain stresser sur lequel travailler, nous proposons un système de filtre permettant de réduire l'espace de recherche des algorithmes.

Lorsque le formateur sélectionne un nœud de la taxonomie afin de générer son espace-zpd, le système va automatiquement restreindre l'espace de recherche au sous-ensemble constitué des feuilles ayant pour parent le nœud sélectionné. En d'autres termes, le système restreint la recherche à l'ensemble des stresser contenus dans le concept sélectionné. Ainsi, le formateur est capable de contrôler la granularité du système.

4 Exemple d'application

Nous nous plaçons dans le cas des entraînements proposés dans le cadre de VICTEAMS [5], un environnement virtuel pour l'entraînement de leaders d'équipes médicales à des compétences non-techniques en situation de crise.

Dans cet exemple, le formateur souhaite entraîner l'apprenant à des situations stressantes mettant en jeu des tris de blessés en présence de divers stresser, selon la taxonomie réduite présentée en Figure 2 qu'il a au préalable filtrée grâce aux outils proposés. Le profil actuel de l'apprenant est tel qu'illustré dans la Figure 4. Nous nous plaçons dans un cas simple où le profil ne comporte pas de conflit.

² Voir Algorithmes : <https://tinyurl.com/nt4bjr5p>

Pour le prochain entraînement, le système détermine la situation permettant d'étendre le plus efficacement la ZDA de l'apprenant. Les algorithmes déterminent que le candidat idéal est la classe de situations correspondant à un tri entre un allié et un ami (noté Am/A) dans un contexte de tempête de sable (noté S) intense (noté I). Cette tempête se caractérise par une diminution du champ de vision dans l'environnement virtuel, l'augmentation de bruits parasites et l'impossibilité de demander une évacuation d'urgence.

Au terme de la session d'entraînement, le système collecte et analyse les données issues des capteurs, ainsi que les observables indiquant la réussite ou l'échec de la tâche. Il détermine une masse de croyance reflétant le résultat de la session qui est propagée au reste des classes de situations. Analysons deux résultats possibles :

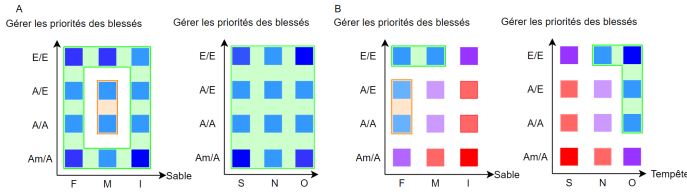


Fig. 5. Illustration des conséquences sur le profil en cas de réussite (A) ou d'échec (B) de l'entraînement à la classe de situations $\sigma = ("Tri\ Ami/Allié", "Tempête\ de\ Sable\ Intense")$

- l'entraînement est une réussite (cas A). La classe de situations, ainsi que les classes proches, intègrent la ZDA, grâce à la propagation des croyances.
- l'entraînement est un échec (cas B). La masse de croyance de la classe de situations est mise à jour (en rouge), et sa propagation génère du conflit (en violet) avec les masses déjà existantes. Il faut donc lever ce conflit grâce à de nouveaux entraînements. La ZDA est affectée par ce conflit.

5 Conclusion et perspectives

Nous avons présenté une approche permettant de générer un profil dynamique du stress d'un apprenant, en nous basant sur des données physiologiques et en utilisant le modèle des Croyances Transférables afin d'émettre des suppositions quant à la maîtrise de situations stressantes variées. Notre approche a l'avantage d'être paramétrable et adaptable par le formateur.

Ce système est couplé à un planificateur ne faisant pas l'objet de cet article, afin de générer des scénarios d'entraînements personnalisés en fonction de chaque profil de stress.

Notre travail est très fortement impacté par la crise sanitaire mondiale actuelle. Nos évaluations requièrent la présence de personnes, ainsi que le port d'un ensemble de capteurs et du port d'un casque de réalité virtuelle. Cet équipement

nécessite en temps normal un protocole sanitaire lourd et strict, et est désormais des plus complexes à mettre en place. Ainsi notre système n'a pu être testé qu'avec des données simulées. Notre contribution fera l'objet d'une évaluation permettant de démontrer la qualité de l'élaboration du profil. Nous voulons également évaluer la capacité de notre système à émettre des croyances correctes sur les capacités d'un apprenant et à proposer un parcours optimal pour étendre sa ZPD en le moins d'entraînements possible.

Remerciements

Ces travaux sont financés par la Région Hauts-de-France et le Fonds Européen de Développement Régional (FEDER) 2014/2020 ainsi que le Labex MS2T.

References

1. Carpentier, K., Lourdeaux, D.: Generation of learning situations according to the learner's profile within a virtual environment. In: international conference on agents and artificial intelligence. pp. 245–260. Springer (2013)
2. Dougall, A.L., Baum, A.: Stress, health, and illness. (2012)
3. Dubois, D., Prade, H.: Possibility theory: an approach to computerized processing of uncertainty. Springer Science & Business Media (2012)
4. Lazarus, R.S., Folkman, S.: Stress, appraisal, and coping. Springer publishing company (1984)
5. Lourdeaux, D., Afoutni, Z., Ferrer, M.H., Sabouret, N., Demulier, V., Martin, J.C., Bolot, L., Boccara, V., Lelong, R.: Victeams: a virtual environment to train medical team leaders to interact with virtual subordinates. In: Proceedings of the 19th ACM International Conference on IVA. pp. 241–243 (2019)
6. Marin, M.F., Raymond, C., Lupien, S.J.: Memory and stress. In: Stress: Physiology, Biochemistry, and Pathology, pp. 69–78. Elsevier (2019)
7. Pearl, J.: Probabilistic reasoning in intelligent systems: networks of plausible inference. Elsevier (2014)
8. Schmidt, P., Reiss, A., Duerichen, R., Marberger, C., Van Laerhoven, K.: Introducing wesad, a multimodal dataset for wearable stress and affect detection. In: Proceedings of the 20th ACM international conference on multimodal interaction. pp. 400–408 (2018)
9. Selye, H.: The stress of life (1956)
10. Shafer, G.: A mathematical theory of evidence, vol. 42. Princeton university press (1976)
11. Smets, P., Kennes, R.: The transferable belief model. *Artificial intelligence* **66**(2), 191–234 (1994)
12. Tchechmedjiev, A.: State of the art: Local semantic similarity measures and global algorithms for knowledge-based word sense disambiguation. In: Proceedings of the Joint Conference JEP-TALN-RECITAL 2012, volume 3. pp. 295–308 (2012)
13. Vygotsky, L.S.: Mind in society: The development of higher psychological processes. Harvard university press (1980)
14. Zadeh, L.A.: Fuzzy sets. In: Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh, pp. 394–432. World Scientific (1996)
15. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems* **1**(1), 3–28 (1978)

Soutien aux pédagogies actives et transformations impulsées par une plateforme numérique dédiée aux sciences expérimentales

Maëlle Planche^{1,2}, Cédric d'Ham², Christian Hoffmann², Nadine Mandran², Isabelle Girault², Claire Wajeman², Patricia Marzin³

¹ Centre des Nouvelles Pédagogies, F-38000 Grenoble, France

² Université Grenoble Alpes, CNRS, Grenoble INP, LIG, F-38000 Grenoble, France

³ Université de Bretagne Occidentale, CREAD, F-29200 Brest, France

maelle.planche@grenoble-inp.fr

Résumé. Cet article s'attache à analyser le soutien aux pédagogies actives et les transformations pédagogiques impulsées par une plateforme numérique dédiée à l'enseignement des sciences expérimentales : la plateforme LabNbook. Une évaluation longitudinale conduite auprès d'enseignants du supérieur a permis de mettre en évidence les intentions d'utilisation de LabNbook, puis d'en étudier les usages réels. Les résultats, après plusieurs périodes d'utilisation, pointent le rôle joué par la plateforme pour soutenir les pédagogies actives déjà en place et pour impulser des transformations dans les activités pédagogiques proposées par les enseignants. Il apparaît que l'usage de la plateforme conduit les enseignants à interroger l'alignement pédagogique de leur enseignement au niveau de la cohérence entre objectifs d'apprentissage et activités proposées aux étudiants, avec pour conséquence de modifier les objectifs d'apprentissage disciplinaires.

Mots-clés : plateforme numérique, pédagogies actives, usages, transformation pédagogique, alignement pédagogique

Summary. This article examines how the platform LabNbook, a digital environment dedicated to the teaching of the experimental sciences, supports active pedagogies and promotes pedagogical transformations. We conducted a longitudinal study with higher education teachers in order to identify the platform usage intentions and its actual usage. After several periods of use, the results highlight the role played by the platform to support the active pedagogies already in use and to stimulate changes in the pedagogical activities. It seems that the uses of the platform lead teachers to question the instructional alignment between the learning objectives and the activities offered to the students, resulting in the modification of their disciplinary learning objectives.

Keywords: digital environment, active learning, usage, pedagogical transformation, instructional alignment

1 Introduction

Dans l'enseignement des sciences, une place importante est donnée aux « pédagogies actives » qui favorisent l'autonomie des étudiants et leur implication dans le processus d'apprentissage. Des études ont montré les avantages de ces pédagogies, que ce soit en termes d'apprentissage de concepts [1] ou d'acquisition de nouvelles compétences scientifiques [2]. Les pédagogies actives impliquent que les étudiants et les enseignants gèrent de nouvelles tâches. Les étudiants doivent devenir plus autonomes dans la production de solutions et travaillent souvent en collaboration lors de séquences étendues avec des échanges en présentiel et à distance, en synchrone et en asynchrone. Les enseignants doivent structurer, suivre 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 de supports adaptés [3]. C'est dans ce contexte que la plateforme LabNbook a vu le jour. Elle est le fruit de recherches menées par des didacticiens sur les environnements informatiques pour l'apprentissage des sciences. La conception de la plateforme a été inspirée par le cadre des théories socio-constructivistes de l'apprentissage : la construction de connaissances résulte de la confrontation entre les connaissances de l'apprenant, ses perceptions du monde réel et les interactions qu'il entretient avec son environnement social. Suivant ces principes, LabNbook a été conçue comme un support pour aider les apprenants à mener des démarches expérimentales en interaction avec d'autres étudiants et sous la supervision de leurs enseignants [4].

Conçue spécifiquement pour l'enseignement des sciences expérimentales au lycée ou dans le supérieur, LabNbook offre aux étudiants des espaces de travail partagés pour leurs travaux d'écriture scientifique : rapports, comptes rendus, cahiers de laboratoire... Elle est utilisée dans des situations variées : en travaux pratiques (TP), en projet ou en apprentissage par problème (APP), en présentiel ou à distance, en synchrone ou en asynchrone. Quatre outils sont fournis aux étudiants pour produire des écrits scientifiques : textes, schémas, protocoles expérimentaux et jeux de données expérimentales ajustés par des modèles mathématiques. Des outils de communication (messagerie, tchat, etc.) sont aussi inclus dans la plateforme.

Bien que conçue à partir d'un modèle didactique [4] et avec l'intention de soutenir les pédagogies actives, LabNbook n'impose pas de démarche pédagogique préétablie et laisse aux enseignants une totale liberté pour définir leurs objectifs d'apprentissage et leurs stratégies d'enseignement. Dans LabNbook, une « mission » désigne une activité dévolue à des équipes d'étudiants qui partagent un espace de travail commun. L'enseignant compose cet espace en le structurant en différentes parties accompagnées de consignes de travail et en donnant accès à certains outils et ressources. L'enseignant visualise les productions des étudiants et accède au résumé de leur activité sous forme de tableau de bord. LabNbook permet aux enseignants de scénariser librement leur enseignement de façon plus ou moins élaborée, cela allant du remplacement des comptes-rendus de TP papier, au guidage de démarches expérimentales complexes comme, par exemple, lors de séances d'APP réparties sur plusieurs semaines [5].

En 2016, la plateforme a obtenu un financement de l'université Grenoble-Alpes (UGA) suite à un appel à projet fléché « Transformations pédagogiques et plateformes

learning-by-doing ». Un volet d'évaluation conséquent a été mis en place pour déterminer si LabNbook soutient effectivement les pédagogies actives, voire impulse des transformations dans les pédagogies mises en place par les enseignants. Cet article propose de répondre aux deux questions suivantes : un outil numérique comme LabNbook, spécifiquement pensé pour les pédagogies actives mais qui n'impose aucune démarche pédagogique spécifique, agit-il comme vecteur de transformation de la pédagogie des enseignants ? Si oui, quelle est la nature des transformations ?

2 La transformation pédagogique

Biggs [6] a développé le modèle de l'alignement constructif où il décrit la nécessaire mise en cohérence de trois ingrédients lors de la construction d'une séquence pédagogique : les objectifs d'apprentissage visés, les activités pédagogiques proposées et l'évaluation des acquis à l'issue de la séquence. Nous avons choisi de caractériser la transformation d'une séquence pédagogique en évaluant les modifications opérées par les enseignants sur deux de ces ingrédients :

- Les objectifs d'apprentissages, différenciés selon deux catégories :
 - Les compétences disciplinaires, entendues comme l'acquisition de savoirs et de savoirs-faires (concepts, méthodes de résolution de problème, techniques de laboratoire ou encore étapes de la démarche expérimentale) propres à chaque discipline (biologie, chimie et physique notamment).
 - Les compétences transversales comme, par exemple, apprendre à communiquer avec les membres de son équipe, à s'organiser, à agir face aux imprévus, à faire preuve d'esprit critique, etc. Ces compétences ne sont pas propres aux sciences expérimentales mais leur acquisition dépend fortement des activités pédagogiques proposées dans les situations d'apprentissage.
- Les activités pédagogiques définies comme l'ensemble des tâches proposées aux apprenants et la manière dont elles leur sont distribuées (aspects collaboratifs) et organisées temporellement (séquençage des activités, choix des modalités synchrone ou asynchrone) et spatialement (travail à distance ou en présentiel).

Le troisième ingrédient de l'alignement, l'évaluation des acquis, est absent dans notre caractérisation du fait que LabNbook n'inclut pas actuellement d'outil d'évaluation et que celle-ci se déroule donc généralement en dehors de la plateforme. Nous avons néanmoins eu accès à certaines pratiques en matière d'évaluation formative. Ces pratiques intervenant tout au long de l'enseignement, il nous a semblé naturel de les inclure dans les activités pédagogiques.

Le modèle développé par Biggs met l'accent sur un deuxième aspect : l'alignement doit être constructif. C'est-à-dire que chaque ingrédient est, dans sa conception, centré sur l'apprenant qui, en concordance avec les idées du constructivisme, doit lui-même construire activement ses connaissances pour qu'elles fassent sens pour lui. Biggs ne prescrit pas une méthode que l'enseignant devrait appliquer, mais décrit une attitude par rapport à l'enseignement qui implique une conscience de l'apprenant et de son monde. Il s'agit donc de proposer des « pédagogies centrées-étudiants » (Student-Centered Learning, SCL). Afin de caractériser les transformations au regard d'un

alignement constructif, nous nous sommes basés sur les travaux de Hannafin [7] [8] pour établir cinq critères d'évaluation du « degré de centration » d'une pédagogie :

- Les objectifs d'apprentissage et les compétences visées sont explicités.
- Les modalités de résolution de problèmes sont déléguées aux étudiants.
- Une organisation qui permet de travailler en dehors de la classe est proposée.
- Le travail collaboratif est encouragé.
- Les étudiants peuvent évaluer leurs apprentissages en cours d'activité.

3 Méthodes d'évaluation

Avec plusieurs milliers d'utilisateurs par an dans des lycées et des établissements d'enseignement supérieur, LabNbook a pu être évaluée en milieu écologique selon une approche systémique. Ce type d'étude assure des résultats représentatifs d'une utilisation réelle et favorise l'amélioration continue de la plateforme. Nous présentons ici les grands principes du processus d'évaluation longitudinale (PEL) [9] [10] élaboré sur la base de la méthode Thèdre [11].

Un travail en étroite collaboration avec des enseignants-utilisateurs a été réalisé afin d'explorer l'impact de la plateforme sur leurs pratiques et sur celles des étudiants. L'originalité de cette évaluation repose avant tout sur sa dimension longitudinale qui suppose de suivre l'évolution de différentes mesures effectuées au cours de plusieurs années d'utilisation de la plateforme. L'évaluation longitudinale s'est déroulée sur trois années d'enseignement, de septembre 2017 à juin 2020. Selon les outils de mesure utilisés, les données ont été collectées auprès des étudiants et des enseignants à différents moments de l'utilisation : avant enseignement ou après enseignement lors de la première, deuxième ou troisième année d'utilisation. *In fine*, l'évaluation a concerné 23 unités d'enseignement, 157 enseignants et 1345 étudiants des établissements de l'UGA. Plusieurs points de mesures ont été retenus pour comprendre les attentes des enseignants, leurs usages de LabNbook, l'adéquation de la plateforme à leurs besoins, et les éventuelles transformations de leurs pratiques pédagogiques. Le PEL mobilise à la fois des approches quantitatives et qualitatives et des données de nature factuelle ou déclarative. Le tableau ci-après présente les différents outils de mesure utilisés (Table 1).

Table 1. Outils de mesure du PEL organisés par approche et par nature des données

	Approche quantitative	Approche qualitative
Données de nature factuelle	Traces d'activité (n=4,6M)	Analyse des missions (n=144) Débriefings, séminaires, etc.
Données de nature déclarative	Questionnaires baromètres enseignants (n=90/159 *) Questionnaire de satisfaction étudiants (n=1345)	Entretiens semi-directifs enseignants (n=35/47 *)

* données mobilisées pour cet article / total des données collectées

Pour cet article, nous mobilisons les données de nature déclarative récoltées auprès des enseignants : questionnaires baromètres et entretiens. Des questions spécifiques pour mesurer les indicateurs SCL ont été construites pour les questionnaires baromètres et pour les entretiens. Ces indicateurs ont permis de mesurer le « degré de centration sur l'étudiant » des pédagogies proposées par les enseignants. L'évolution des objectifs d'apprentissage a été appréhendée par le biais des entretiens et des retours informels des enseignants. Au total, 159 questionnaires baromètres ont été recueillis en trois vagues de passation : avant utilisation, après une période d'utilisation puis deux périodes d'utilisation. Afin de mesurer l'évolution des perceptions de l'utilité et de l'utilisabilité de LabNbook chez les enseignants, le protocole développé pour la passation des questionnaires impliquait que les mêmes enseignants répondent à chacune de ces trois vagues. Sur les 159 questionnaires recueillis au total, nous en exploitons 90 correspondant à 26 enseignants ayant répondu aux trois vagues de passation et 6 enseignants n'ayant répondu qu'aux deux premières vagues. Nous mobilisons 35 entretiens (15 avant utilisation et 20 après utilisation) d'enseignants ayant utilisé LabNbook sur toute la période d'évaluation. Ces entretiens ont été conduits auprès de 22 enseignants. Les entretiens ont été enregistrés, intégralement retranscrits et ont fait l'objet d'une analyse thématique [12] [13].

4 Avant utilisation de LabNbook : activités pédagogiques existantes et intentions d'usage de la plateforme

Pour mesurer les transformations pédagogiques au fil de l'utilisation de LabNbook, nos analyses se sont focalisées sur des thématiques différentes selon la période de passation des entretiens. Pour l'analyse des entretiens pré-utilisation, notre regard s'est porté sur deux thématiques : 1) les éléments permettant de caractériser les activités pédagogiques mises en place avant utilisation de LabNbook et notamment le niveau de « centration » sur l'étudiant et 2) les éléments relatifs à la perception de l'utilité de LabNbook par les enseignants ainsi que leurs intentions d'utilisation de la plateforme. Nous prenons appui ici sur le concept d'*utilité perçue* [14] [15] qui renvoie au degré avec lequel un individu considère que l'utilisation d'un système est susceptible d'améliorer sa performance, et sur celui d'*intentions d'usage* qui correspond à « la résultante de la perception de l'utilité et de l'utilisabilité des technologies par les utilisateurs » [16].

4.1 Caractérisation des enseignements avant utilisation de LabNbook

Les réponses aux questionnaires mettent en évidence des différences selon les indicateurs SCL choisis pour mesurer le niveau de centration sur l'étudiant des pédagogies en place dans les enseignements avant utilisation de LabNbook. C'est en particulier sur la demande de travail en dehors des séances présentiels et l'explicitation des objectifs d'apprentissage visés que les enseignants interrogés sont les plus unanimes : la majorité déclaraient demander du travail en dehors des séances à leurs étudiants (n=29/32) et expliciter leurs objectifs d'apprentissage (n=31/32) avant utilisation de la plateforme. Ils sont plus partagés concernant la possibilité de faire

travailler leurs étudiants de manière itérative (n=16/32), de leur laisser la liberté d'organiser leur temps pendant la séance (n=17/32), de choisir leur méthode de résolution des problèmes (n=20/32) et de valider le travail des étudiants en cours d'activité (n=22/32). Un peu moins d'un quart des répondants proposaient à leurs étudiants des outils d'auto-évaluation (n=7/32).

Les entretiens permettent de caractériser l'implication initiale des enseignants dans les pédagogies actives. Parmi les 19 enseignants rencontrés, 12 s'étaient investis dans des pédagogies actives dans le cadre d'autres UE. C'est le cas, par exemple, d'une enseignante de physique qui propose des modalités d'apprentissage par problème (APP) et explique qu'elle n'est « *pas du tout enseignante par défaut, [que] c'est un vrai choix [...]* [qu'elle a] toujours eu cette sensibilité à la démarche pédagogique » et que LabNbook l'a moins intéressé « *par son aspect technologique, que par son aspect outil pédagogique accompagnant la démarche APP* ».

Les résultats analysés mettent donc en évidence un élément important : les enseignants ayant fait le choix d'utiliser LabNbook avaient auparavant des pratiques pédagogiques centrées-étudiants. Ceci est un premier résultat sur l'acceptation de LabNbook : les enseignants les plus enclins à utiliser LabNbook semblent être ceux déjà engagés dans des pédagogies actives.

4.2 Intentions d'utilisation de LabNbook

Quatre types d'intentions d'utilisation de LabNbook ressortent de l'analyse des pré-entretiens. La première touche aux activités pédagogiques proposées dans les enseignements. Les deux suivantes ciblent les objectifs d'apprentissage au niveau des compétences disciplinaires ou des compétences transversales. La quatrième rassemble les enseignants ayant une intention d'utilisation floue ou aucune intention d'utilisation explicite. Un même enseignant peut avoir plusieurs intentions d'utilisation.

Intention d'usage pour l'acquisition de compétences disciplinaires

Dans les questionnaires, « l'aide apportée aux étudiants pour structurer leur rapport d'expérience » (n=27/32) et « l'accès à des outils de rédaction d'un rapport scientifique » (n=25/32) arrivent en tête parmi l'ensemble des propositions relatives à l'utilité perçue de LabNbook. L'intention d'utilisation ciblant les compétences disciplinaires avec LabNbook apparaît centrale pour ces répondants.

Les enseignants rencontrés en pré-entretiens, sont nettement moins nombreux (n=3) à décrire une intention d'utilisation ciblant l'acquisition de compétences disciplinaires. C'est le cas de deux enseignants de physique d'une même UE qui décrivent assez précisément leur intention pédagogique avec LabNbook : « *Le but ça serait qu'au début, c'est nous qui mettons vraiment les protocoles très stricts pour qu'ils aient une idée d'à quoi ça ressemble [...] et après, de relâcher au fur et à mesure pour qu'ils suivent en fait cette même stratégie de protocole dans leur rédaction. C'est là où on pensait que ça pourrait aider d'avoir LabNbook parce que, pour le moment, ce qu'on a fait, c'est d'imprimer quelques fiches clés avec les questions de base qu'ils doivent se poser à chaque fois. S'ils ont une trame avec ces questions clés mais qu'ils doivent rendre spécifique à la mesure qu'ils font, je pense que ce sera nettement plus facile* ».

Cet enseignant décrit un apprentissage disciplinaire très spécifique – l'apprentissage de la structuration d'un protocole – et projette d'utiliser les possibilités de pré-structuration de la plateforme pour permettre aux étudiants d'acquérir cette « *stratégie de protocole* ». C'était également le cas d'une responsable d'UE de biologie, dont les propos mettent en évidence l'adéquation entre les objectifs d'apprentissages – préparer la phase expérimentale en répondant à des questions précises – et l'atteinte de ces objectifs avec LabNbook : « *en fait, le premier tutorat, il est très théorique. C'est à dire qu'on discute vraiment de l'objectif de leur projet, de leurs questions, d'arriver à bien définir la question. [...]. Et ensuite, à la fin du 2ème tutorat, on leur distribue des fiches de manière à ce qu'ils préparent vraiment la suite, la partie expérimentale. Ces fiches leur permettent de répondre à des questions très précises en fait [...]. Et ça, c'est quelque chose qu'on peut faire de manière beaucoup plus optimisée avec LabNbook* ».

Intention d'usage pour l'acquisition de compétences transversales

Les deux-tiers des répondants aux questionnaires envisageaient LabNbook comme un support favorable pour permettre aux étudiants d'acquérir des compétences pour « organiser leur travail » (n=20/32), pour « partager des ressources entre eux » (n=18/32) et « travailler de façon collaborative » (n=18/32).

Les projections des enseignants rencontrés en entretien pré-utilisation se focalisaient principalement sur l'acquisition de compétences transversales par les étudiants, faisant abstraction des outils spécifiques à l'enseignement des sciences expérimentales. Pour ces enseignants, LabNbook était perçue comme « un plus », un outil favorable à l'acquisition de compétences supplémentaires, sans mobilisation de ses spécificités pour les sciences expérimentales : « *Moi je pense que c'est un réel outil pour arriver de plus en plus à faire travailler les étudiants en TP sous la forme de projet. [...]. Je pense que c'est des compétences malheureusement sur lesquelles on ne va pas assez par rapport au besoin professionnel. Et aussi de développer leur autonomie* ». Parmi l'ensemble des compétences transversales citées en entretien, l'apprentissage du travail collaboratif et du travail à distance sont revenus très fréquemment.

Intention d'usage pour transformer les activités pédagogiques

Les résultats issus des questionnaires et relatifs à la perception de l'utilité de LabNbook pour soutenir les activités pédagogiques proposées aux étudiants mettent en évidence une volonté de transformation de la part des enseignants. Les trois-quarts d'entre eux attendaient de la plateforme qu'elle leur permette de faire évoluer leurs pratiques pédagogiques (n=24/32). Leurs attentes en la matière ciblaient également les possibilités de suivi de l'activité des étudiants (n=26/32). De façon un peu moins unanimes, leur perception de l'utilité de LabNbook portait aussi sur la facilitation du travail entre les séances (n=20/32). Ils étaient nettement moins nombreux à envisager de travailler de manière plus collaborative entre collègues de l'UE (n=12/32).

Bien qu'un certain nombre d'enseignants aient évoqué des attentes relatives à une transformation des activités pédagogiques avec l'utilisation de LabNbook, ces attentes sont apparues secondaires et assez floues durant les entretiens. Certains craignaient les résistances probables de leurs collègues. D'autres évoquaient leur intention de faciliter le travail des étudiants entre les séances, mais sans avoir encore d'idée précise des

modalités d'organisation du travail à distance avec LabNbook.

Aucune intention précise d'usage

Parmi les 32 enseignants ayant répondu au questionnaire, ils étaient 9 à voir l'introduction de LabNbook dans leur enseignement comme une contrainte potentielle.

Parmi les 19 enseignants rencontrés en entretien, seuls deux ne percevaient aucune valeur ajoutée à l'utilisation de LabNbook. L'un de ces deux enseignants s'était conformé au choix du reste de l'équipe pédagogique. Il ne partageait pas l'ensemble des intentions pédagogiques de ses collègues et voyait comme principal intérêt à son utilisation la « *suppression des comptes-rendus papier* ». Pour l'autre, c'était principalement parce qu'il connaissait « depuis très longtemps » deux des concepteurs de la plateforme, qu'il avait décidé de tester LabNbook : « *ça, c'était ma motivation initiale parce que, s'il n'y avait pas eu ça, peut-être que je l'aurais laissé de côté [...]. J'ai d'autres choses à faire !* ».

Les intentions d'utilisation des enseignants rencontrés avant utilisation portent donc principalement sur l'acquisition de compétences transversales. Mais avant l'utilisation de LabNbook, les projections demeurent floues et n'ont pas encore impliqué de réflexion sur l'alignement pédagogique entre les objectifs d'apprentissage ciblés et les activités pédagogiques à mettre œuvre pour atteindre ces objectifs avec l'outil.

5 Après utilisation de LabNbook : pédagogies soutenues et impulsions de transformations

Avec l'utilisation de LabNbook, les enseignants sont entrés dans une phase d'appropriation durant laquelle ils ont été amenés à instrumenter leur activité et à instrumentaliser la plateforme [14] [15]. Ils ont parfois été accompagnés par l'équipe conceptrice au moment d'imaginer et de mettre en place leur scénario pédagogique sur LabNbook et parfois après une ou deux périodes d'expérimentation. Dans cet article, nous ne développons pas le processus d'appropriation – qui implique une analyse en profondeur des obstacles rencontrés et des ajustements didactiques et pédagogiques opérés – mais proposons d'analyser les formes de pédagogies actives finalement soutenues par LabNbook et les transformations pédagogiques impulsées par son utilisation. Nous mobilisons les 32 questionnaires baromètres recueillis après une ou deux périodes d'utilisation de la plateforme et les 20 entretiens réalisés à l'issue de la période d'évaluation. L'analyse des entretiens post-utilisation s'est focalisée sur la satisfaction ou les déceptions en lien avec les intentions pédagogiques de départ et les découvertes ayant permis d'impulser d'éventuelles transformations pédagogiques.

5.1 LabNbook soutient les pédagogies actives déjà en place

Les résultats des questionnaires à l'issue de la période d'utilisation mettent en évidence un renforcement, sur certains indicateurs SCL, des pédagogies actives déjà en place. C'est le cas, par exemple, concernant la demande de travail en dehors des séances (de $n=29$ à $n=32$) et la liberté laissée aux étudiants d'organiser leur temps pendant la séance

(de $n=17$ à $n=22$). L'étude des effets de LabNbook sur les pratiques enseignantes donne des résultats positifs sur les deux aspects évalués : le suivi des étudiants ($n=26$ à $n=27$) et l'évolution des pratiques pédagogiques ($n=28$ à $n=28$). Dans les deux cas, les attentes initiales en la matière étaient assez élevées et semblent être satisfaites à l'issue d'une première, puis d'une seconde période d'utilisation. À la question « *Utiliser LabNbook est une contrainte car c'est utiliser un outil numérique supplémentaire* », les enseignants interrogés répondent majoritairement par la négative pour toutes les vagues de passation. Trois enseignants qui se déclaraient en accord avec la proposition lors de la première vague, se déclarent néanmoins en désaccord à l'issue de la période d'utilisation, ce qui semble témoigner d'une forme d'appropriation de la plateforme. La proportion des répondants faisant état de l'utilité de LabNbook tend également à évoluer positivement au cours du temps.

En revanche, LabNbook n'impulse aucune évolution remarquable concernant la mise à disposition d'outils d'auto-évaluation et la possibilité de faire travailler les étudiants de manière itérative : ces pratiques, peu fréquentes, se maintiennent presque à l'identique après une période d'utilisation de la plateforme. Néanmoins, aucun indicateur ne « s'effondre » après une ou deux périodes d'utilisation ce qui laisse penser que si LabNbook n'impulse pas toujours de transformations sur les activités pédagogiques vers des pédagogies actives, elle soutient celles déjà mises en place.

5.2 LabNbook impulse des transformations pédagogiques non anticipées

Le rôle joué par la plateforme pour impulser des transformations pédagogiques non envisagées initialement est un autre résultat de l'évaluation. Si ce constat ne ressort pas des questionnaires qui mesurent les critères SCL, c'est dans les entretiens qu'on identifie le mieux ce phénomène. Ainsi, nous pouvons citer le cas évoqué précédemment de cet enseignant, peu engagé dans les pédagogies actives, qui avait pour intention initiale la « *suppression des comptes-rendus papier* ». En utilisant LabNbook, il semble découvrir à la fois de nouvelles activités pédagogiques – échanger régulièrement avec ses étudiants, suivre leur travail, proposer des modalités de travail de type itératif – et les effets de ce type de pédagogie sur le travail des étudiants : « *pour les TD, ils arrivaient en salle, si on leur avait donné le thème du TD, ils ne le travaillaient pas avant. Alors qu'avec LabNbook, ils sont obligés de travailler puisqu'on pose des questions auxquelles il faut qu'ils répondent avant de venir en TD. Donc ça, c'est bien.* ». Il définit la plateforme comme un outil « *tourné vers la relation enseignant et étudiants* » et fait le constat qu'en mobilisant la plateforme, il a finalement transformé les activités proposées aux étudiants pour coller davantage à ses objectifs pédagogiques en matière d'apprentissages disciplinaires.

Les transformations les plus notables touchent moins aux activités pédagogiques et aux objectifs d'apprentissage transversaux, qu'aux objectifs d'apprentissage disciplinaires. Ces transformations sont parfois impulsées par la découverte d'une fonctionnalité dédiée aux sciences, qui va permettre d'aborder différemment une notion importante pour la discipline. C'est ce que nous explique, par exemple, cet enseignant de physique au sujet de l'ajustement non automatique de données expérimentales par des modèles mathématiques qui est possible grâce à un outil de LabNbook :

« L'ajustement de courbes paramétrées, moi c'est un truc que j'aime beaucoup dans LabNbook. Ça n'existe pas à ma connaissance dans les tableurs. Ça rend vraiment concret ce qu'est la notion d'ajustement y compris, d'ailleurs, pour des ajustements simples par une droite. Ça rend vraiment concret ce qu'est la pente, ce qu'est l'ordonnée à l'origine. ». C'est également ce que nous raconte cet enseignant de chimie qui, en se familiarisant avec certains outils, a précisé ses objectifs d'apprentissage disciplinaires : *« Cette année, on a essayé d'aller un peu plus loin dans la simulation [...], on a insisté sur l'importance de faire de beaux graphiques, de belles courbes. [...] Ils doivent tracer les courbes. C'est LabNbook qui trace les courbes. Ils se rendent bien compte là où il y a des sauts plus importants. Et on ne fait pas plusieurs fois les manips [...]. Donc, je dirais qu'une première fois, c'est avec LabNbook pour qu'ils se rendent compte où il y a un saut, qu'il faut doser ou aller plus doucement »*.

Dans certains cas, ces transformations ont été impulsées au détour d'une phase d'instrumentation, lorsqu'il a fallu adapter le scénario pédagogique aux contraintes de la plateforme : *« c'est parti d'une contrainte technique mais finalement, ça s'est avéré une extrêmement bonne idée, le fait qu'on fasse deux missions distinctes [...]. Et ça a très bien marché, avec des structures de compte-rendus plus claires. Avec également [...] une clarification sur ce qu'est la phase de manipulation, de saisie de données et ce qu'est la phase de rédaction de compte-rendu [...]. Enfin vraiment, je dois dire, c'était un peu inattendu ! On voulait juste avoir des comptes-rendus plus agréables à lire et à corriger et on s'est aperçu que là, il y avait un vrai progrès pédagogique qu'on n'avait pas anticipé ! Donc voilà, ça fait très plaisir ce genre de chose parce qu'on découvre que les outils eux-mêmes peuvent être des facteurs de progrès pédagogiques ! »*. À l'instar des propos de cet enseignant, la satisfaction de ceux pour lesquels LabNbook a impulsé des transformations pédagogiques non anticipées est généralement élevée et les projections d'utilisation future semblent évidentes.

5.3 LabNbook impulse une réflexion autour de l'alignement pédagogique

Les entretiens post-utilisation mettent en évidence un constat assez général : les enseignants dont les intentions initiales étaient tournées principalement vers les compétences transversales, n'établissent pas de lien clair entre l'utilisation de LabNbook et l'acquisition attestée de ces compétences par les étudiants. L'explication repose probablement sur le fait qu'ils n'ont généralement pas mis en place les modalités d'évaluation de cette acquisition. Ce faisant, lorsque ces enseignants sont satisfaits après plusieurs périodes d'utilisation, cette satisfaction porte généralement plus sur les nouvelles activités pédagogiques qui ont pu être proposées sur LabNbook et/ou sur l'acquisition de compétences disciplinaires. C'est, par exemple, le cas de cet enseignant de chimie dont les attentes initiales visant l'acquisition de compétences transversales avec LabNbook étaient fortes – il parlait notamment de favoriser l'apprentissage du travail en mode projet, d'apprendre à s'organiser, à se répartir le travail et à échanger au sein d'une équipe. Voici le bilan qu'il tire après deux périodes d'utilisation : *« c'est un outil qui m'a intéressé [...] surtout pour la structuration du travail que l'on demande aux étudiants [compétences disciplinaires] et aussi l'échange que l'on peut avoir avec eux [activités pédagogiques]. [...] Mais au moins ça a eu le gros mérite aussi de faire*

réfléchir. [...] Ça a eu le gros intérêt de réinterroger ma pratique, mon savoir ». Cet enseignant soulève un point important, que l'on retrouve en filigrane dans de nombreux entretiens post-évaluation : LabNbook impulse une réflexion sur l'alignement pédagogique, c'est-à-dire la cohérence entre les objectifs pédagogiques, les activités proposées aux étudiants sur LabNbook et en dehors, et les formes d'évaluation mises en place. LabNbook peut donc impulser une réflexion – individuelle et parfois collective – qui va éventuellement conduire à une évolution pédagogique.

Lorsqu'émergent des désaccords au sein des équipes pédagogiques au sujet de l'alignement pédagogique, il en ressort un risque d'abandon de la plateforme. C'est ce qui semble se produire dans cette UE de biologie destinée à des étudiants de niveau Master, dans laquelle LabNbook était principalement utilisée pour soutenir la démarche en mode projet [activités pédagogiques]. Les intentions initiales des enseignants ciblaient également l'acquisition de compétences transversales, notamment l'autonomie et la professionnalisation des étudiants en les poussant à collaborer davantage entre eux, à s'organiser et à développer leur sens critique. À l'issue de deux années d'utilisation, les enseignants dressent le constat que *« ça ne fonctionne pas très bien »* parce que les étudiants *« ne jouent pas le jeu »*. Nous avons rencontré deux enseignants de cette UE. Ces deux enseignants dressent le même bilan. L'un envisage d'arrêter d'utiliser la plateforme, l'autre de travailler davantage l'ajustement entre les attentes pédagogiques et les activités – selon lui *« trop complexes »* – proposées aux étudiants sur LabNbook. Sans travail autour de cet alignement, il est probable que l'outil soit abandonné.

6 Conclusion

Finalement, un outil numérique comme LabNbook, agit-il comme vecteur de transformation de la pédagogie des enseignants ? Évidemment, introduire un nouvel EIAH dans une pratique enseignante est en soi une démarche de transformation pédagogique. Nous avons confirmé la capacité de LabNbook à soutenir certaines caractéristiques SCL de pédagogies déjà en place et notre évaluation met en évidence que les enseignants qui décident d'utiliser LabNbook ont initialement une appétence pour les pédagogies actives. Le deuxième résultat de notre étude concerne la nature des transformations impulsées par LabNbook. La plateforme semble faire évoluer les enseignants sur les objectifs pédagogiques situés au niveau des compétences disciplinaires. Ce résultat paraît étonnant car les intentions d'utilisation initiales des enseignants ciblaient peu ces savoirs et davantage l'acquisition de compétences transversales. La phase d'appropriation, qui suppose d'instrumentaliser l'outil numérique et d'instrumenter l'activité pédagogique, conduit inévitablement à un travail autour de l'alignement pédagogique. C'est le troisième résultat notable de l'évaluation conduite : LabNbook impulse une réflexion sur l'approche pédagogique et sur sa cohérence, ce qui peut éventuellement conduire à des transformations. Ce résultat plaide pour un accompagnement des enseignants afin de soutenir la réflexion « en continu » sur les questions d'alignement pédagogique lors de l'introduction d'un EIAH tel que LabNbook dans les enseignements.

Références

1. 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(23), 8410–8415 (2014).
2. 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(1), 54–98 (2010).
3. Land, S.M, Hannafin, M.J. & Oliver, K.M. : Student-Centered Learning Environments: Foundations, Assumptions, and Design. In: Land, S., Jonassen, D. (eds) *Theoretical Foundations of Learning Environments* (Chapter 1), Routledge (2012).
4. d’Ham, C., Wajeman, C., Girault, I., Marzin-Janvier, P. : Transposition de la démarche expérimentale dans un environnement numérique de support : LabNbook, de la caractérisation didactique à l’utilisation en situation écologique. In: *Actes des 11ème Rencontres Scientifiques de l’ARDIST*, Bruxelles, Belgique (2021).
5. 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: *Actes du 10e colloque Question de Pédagogies dans l’Enseignement Supérieur (QPES)*, 1424-1435, Brest (2019).
6. Biggs, J. : Enhancing teaching through constructive alignment. *Higher Education* 32, 347-364 (1996).
7. Hannafin, M.J. : Student-centered learning. In: *Encyclopedia of the Sciences of Learning*. Springer US, 3211–3214 (2012).
8. <http://edglossary.org/student-centered-learning/>
9. Mandran, N., Marzin-Janvier, P., Planche, M., Karoui, A., Girault, I. : Processus d’évaluation longitudinale (PEL) d’une plateforme pédagogique (LMS) : le cas de LabNbook. In: *Environnements Informatiques pour l’Apprentissage Humain*, Paris, France (2019).
10. Planche, M., Girault, I., Mandran, N., Marzin, P., d’Ham, C., Wajeman, C. : Contribution de différents outils de mesure à l’évaluation des usages d’une plateforme numérique par un processus longitudinal : cas du travail à distance. In : *Environnements Informatiques pour l’Apprentissage Humain*, Paris, France (2019).
11. Mandran, N. : *Traceable Human Experiment Design Research: Theoretical Model and Practical Guide*. John Wiley & Sons, Inc., Hoboken, NJ, USA (2018).
12. Beaud, S., Weber, F. : *Guide de l’enquête de terrain. La découverte*, Paris, (2010).
13. Paillé, P., Mucchielli, A. : *L’analyse qualitative en sciences humaines et sociales*. Armand Colin (2021).
14. Tricot, A., Plégat-Soutjis, F., Camps, J-F., Amiel, A., Lutz, G., Morcillo, A. : Utilité, utilisabilité, acceptabilité : interpréter les relations entre les trois dimensions de l’évaluation des EIAH. In : Desmoulins, C., Marquet, P., Bouhineau, D. (eds), *Environnements informatiques pour l’apprentissage humain*. 391-402 (2003).
15. Davis, F. D. : Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13(3), 318-339 (1989).
16. Poyet, F. : Perception de l’utilité et usages pédagogiques d’environnements numériques de travail par des enseignants du second degré. *STICEF – Recueil*. 45-64 (2015).
17. 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).
18. Béguin, P., Rabardel, P. : Concevoir des activités instrumentées. *Revue d’Intelligence Artificielle* 14, 35-54 (2000).

La prise en compte de l'incertitude dans le processus de diagnostic des compétences de l'apprenant

Nesrine Rahmouni¹, Domitile Lourdeaux², Azzeddine Benabbou³, Taher Bensebaa¹

¹ Université de Badji Mokhtar Annaba – Algérie

nesrinerahmouni.pro@gmail.com

t.g-bensebaa@gmail.com

² Alliance Sorbonne Université, Université de technologie de Compiègne, CNRS, Heudiasyc
UMR 7253

domitile.lourdeaux@hds.utc.fr

³ Université de Toulouse 3 - Institut de Recherche en Informatique de Toulouse

azzeddine.benabbou@irit.fr

Résumé. Nos travaux portent sur le processus de diagnostic dans les environnements informatiques pour l'apprentissage humain (EIAH). Dans cet article, nous proposons un module basé sur la théorie des fonctions de croyances pour le diagnostic des compétences de l'apprenant dans un EIAH. Ce module permet de prendre en compte l'imprécision, l'incertitude et la contradiction des informations récoltées sur l'apprenant. Nous visons par ce module la détermination de l'état des compétences évaluées, ainsi que la prédiction de l'état des compétences non évaluées de l'apprenant. Notre postulat principal est que ces résultats pourront servir d'appui à la rétroaction et à la remédiation rendues à l'apprenant, en lui offrant ainsi une meilleure personnalisation.

Mots-clés : diagnostic, compétences, apprenant, incertitude, fonction de croyances.

Abstract. Our work is related to the diagnosis process in intelligent tutoring systems (ITS). In this paper, we propose a diagnosis based on the theory of belief functions to take into account uncertainties in this process. The objective of our work is twofold: first, a dynamic diagnosis of the evaluated skills, then, the prediction of the state of the non-evaluated one. Our main premise is that these results will serve as a support to the feedbacks and to the remediation given to the learner, by providing her/him a proper personalization.

Keywords: diagnosis, skills, learner, uncertainty, belief functions.

1 Introduction

L'un des principaux atouts des environnements informatiques pour l'apprentissage humain (EIAH) est leur capacité à fournir un apprentissage sur mesure à chaque apprenant [1]. La personnalisation des apprentissages est en effet l'un des verrous majeurs des EIAH. Pour ce faire, il est indispensable de disposer d'une représentation informatique

de l'apprenant dans ces environnements [2]. C'est dans ce cadre de recherche que s'inscrit notre travail. Nous souhaitons proposer un module informatique pour le diagnostic des compétences de l'apprenant dans un EIAH dédié à l'apprentissage de l'algorithmique. Nous visons par ce module à la fois l'identification de l'état des compétences évaluées de l'apprenant et la prédiction de l'état des compétences non évaluées. Néanmoins, le diagnostic des compétences dans les EIAH est souvent un processus complexe [3] car il repose sur des évaluations souvent imparfaites tachées d'incertitudes, d'imprécisions et parfois de contradictions. Le problème que nous traitons ici est comment prendre en compte ces imperfections dans le processus de diagnostic afin d'améliorer la personnalisation des apprentissages ?

Nous faisons l'hypothèse que la théorie des fonctions de croyances permet de prendre en compte les imperfections du diagnostic et donc de l'améliorer. Dans la suite de cet article, nous présentons dans la section 2 quelques travaux connexes. Dans la section 3, nous détaillons l'architecture générale dans laquelle s'intègre notre module de diagnostic que nous décrivons dans les sections 4 et 5. Enfin, dans la section 6, nous exposons l'expérimentation réalisée afin d'évaluer le module de diagnostic.

2 État de l'art

Le Bayesian Knowledge Tracing (BKT)[4] représente l'une des techniques d'intelligence artificielle les plus utilisées pour le diagnostic des connaissances [3]. Le BKT s'appuie sur quatre paramètres : $p_0(x)$ la probabilité que l'apprenant maîtrise la compétence a priori, $p(guess)$ la probabilité que l'apprenant applique correctement la compétence par chance, $p(slip)$ la probabilité que l'apprenant applique incorrectement la compétence sans avoir une défaillance dans cette compétence, et $p(T)$ la probabilité de transition d'une compétence de l'état « non acquise » à « acquise » [4]. Ce modèle suppose que la réalisation de la compétence observée est binaire : correcte ou incorrecte [5]. Cependant, dans notre contexte, le processus de diagnostic de compétences est basé sur des évaluations numériques non binaires. De plus, le modèle BKT considère les concepts du domaine comme des éléments indépendants les uns des autres. Cela implique une perte d'informations utiles qui pourrait enrichir le processus de diagnostic. Les réseaux bayésiens, utilisés par exemple dans [6][7][8], permettent d'exploiter les interactions entre les concepts du domaine. En général, dans le cadre d'un ITS, les nœuds du réseau représentent les connaissances ou les compétences, et sont connectés par des relations probabilistes. Toutefois, les réseaux bayésiens ont l'inconvénient majeur de devoir être paramétrés [9]. Or, ce paramétrage est difficile si l'expert ne dispose pas d'une quantité de données suffisante. D'autres travaux de la littérature utilisent la logique floue pour la modélisation de l'apprenant [10][11]. Cette technique permet l'exploitation de variables linguistiques et la conception de modèles à la fois intelligibles et exploitables par des systèmes informatiques. À l'instar de [12], nous proposons un modèle de diagnostic reposant sur la théorie des fonctions de croyances. Une des originalités de cette proposition porte sur la possibilité de quantifier le potentiel conflit qui pourrait survenir des contradictions entre les différentes sources d'informations (ex. évaluations, croyances propagées). Cette proposition permet également, contrairement

aux approches probabilistes classiques, une meilleure représentation de l'état d'ignorance en le distinguant du cas où les hypothèses seraient effectivement équiprobables. Par ailleurs, contrairement à [12], ce modèle exploite les relations entre les compétences du domaine afin d'inférer l'état de celles qui n'ont pas encore été évaluées.

3 Architecture générale

Nous nous intéressons particulièrement au diagnostic des compétences de l'apprenant dans un EIAH dédié à l'apprentissage de l'algorithmique. Le module de diagnostic que nous proposons, s'intègre dans le cadre d'une architecture plus générale qui met en jeu différents modules à savoir : le module d'évaluation et le module de scénarisation. Les modules de cette architecture reposent sur une représentation ontologique du domaine qui permet de spécifier les différents concepts ainsi que les relations qu'ils ont les uns avec les autres. Dans cette section, nous présentons le modèle du domaine ainsi que les différents modules de l'architecture.

3.1 Le modèle du domaine

Dans le cadre de nos travaux, le domaine est modélisé grâce à l'ontologie *Algoskills* [13] qui apporte aux experts une meilleure visualisation et indexation des différents concepts de l'algorithmique. Cette ontologie est organisée autour de trois classes : la classe *Topics*, la classe *Skills* et la classe *Ressources*. La classe *Topics* représente la hiérarchie des notions de cours en algorithmique. La classe *Skills* représente quant à elle l'ensemble des compétences en algorithmique. Elle comporte deux types de compétences : les compétences disciplinaires et les compétences transversales. La classe *Topic* est reliée aux compétences disciplinaires par la relation sémantique « is-useful-to ». Par exemple, le *Topic* « Récursivité » is-useful-to « Identifier la structure récursive d'un problème ». Finalement, la classe *Ressources* regroupe l'ensemble des exercices utilisés pour évaluer les compétences. Cette classe est reliée à la classe *Skills* par la relation sémantique « has-for-objective ».

Afin de faciliter le processus de prédiction de l'état des compétences non évaluées de l'apprenant, nous étendons cette ontologie en ajoutant à la relation hiérarchique existante entre les compétences « is-a », une relation de précédence explicite « is-prerequisite ». Par exemple une compétence C_i est un prérequis de la compétence C_j , si l'acquisition de la compétence C_i précède forcément l'acquisition de la compétence C_j .

3.2 Module d'évaluation

Le module d'évaluation est en charge d'évaluer les productions de l'apprenant. Ce module, reçoit en entrée les productions de l'apprenant et les compare avec les productions attendues par l'expert du domaine. L'interprétation de cette comparaison s'exprime sous la forme de résultats numériques (notes d'évaluations) attribués aux compétences de l'apprenant. Ces notes d'évaluation sont transmises par la suite au module de diagnostic que nous présentons dans la section suivante.

3.3 Module de diagnostic

Dans cette architecture, le module de diagnostic reçoit les résultats retournés par le module d'évaluation. A partir de ces résultats, il génère des croyances sur l'état des compétences de l'apprenant. Une compétence peut être acquise, non acquise, probablement acquise, probablement non acquise ou indéterminé dans le cas où le système ignore l'état de la compétence. Une fois les croyances générées, elles sont transmises par la suite au module de scénarisation qui se charge de générer les décisions adaptées. Le module de diagnostic constitue la contribution majeure de cet article. Nous la détaillons dans les sections 4 et 5.

3.4 Module de scénarisation

Ce module est chargé de la scénarisation des sessions d'apprentissage pour chaque apprenant. Il reçoit les croyances sur l'état des compétences de l'apprenant établies par le module de diagnostic. A partir de ces croyances, il génère les scénarios d'apprentissage (exercices, cours) adéquats. Par exemple si l'état de la compétence A est « Probablement acquis », le module de scénarisation propose des exercices pour consolider cette compétence. En revanche, si l'état de la compétence est « Acquis » le module scénarisation peut proposer des exercices pour lesquels la compétence A est considérée comme prérequis

4 Le module de diagnostic

Les notes d'évaluation sont issues d'interprétations parfois subjectives d'un expert, ou d'un processus automatique d'évaluation. De ce fait, les données fournies sont entachées d'un certain nombre de défauts tels que :

- L'incertitude : un apprenant peut donner une réponse correcte sans pour autant maîtriser les compétences qui lui sont associées. A l'inverse, un apprenant peut donner une réponse erronée sans que cela ne soit nécessairement lié à une défaillance dans les compétences.
- L'imprécision : la note d'évaluation attribuée à la compétence peut ne pas être précise. Par exemple, un processus d'évaluation automatique pourrait estimer que la note d'évaluation attribuée à une compétence est comprise entre 15 et 20.
- La contradiction : les notes d'évaluations peuvent être déterminées à partir de différentes sources. Par conséquent, ces évaluations peuvent être contradictoires.

S'ajoute à ces défauts, l'ignorance de l'état des compétences qui n'ont pas encore été évaluées. Pour faire face à ces défauts, le module de diagnostic doit prendre en compte toutes ces imperfections. Par ailleurs, il doit être capable de fusionner des informations issues de sources hétérogènes tout en représentant l'éventuel conflit existant. Pour cela, nous avons considéré plusieurs cadres formels : la théorie des ensembles, la théorie des probabilités, la théorie des ensembles flous, la théorie des possibilités et la théorie des fonctions de croyances. Le cadre formel des fonctions de

croyanances nous parait le plus pertinent pour le diagnostic des compétences de l'apprenant. En effet, il généralise toutes les théories précédentes, et répond à tous nos besoins à savoir la représentation de l'incertitude, l'ignorance et la contradiction. Par ailleurs, cette théorie permet de fusionner et représenter des informations issues de sources hétérogènes. Plus précisément, nous utilisons le modèle des croyances transférables [14].

Dans le cadre des fonctions de croyances, l'état du système prend des valeurs discrètes w_n appelées hypothèses. L'ensemble de ces hypothèses $\Omega = \{w_1, w_2, \dots, w_n\}$ est appelé cadre de discernement. Une fonction de croyances peut être représentée de différentes manières. Nous utilisons la représentation par une distribution de masse de croyances notée m définie de 2^Ω dans $[0,1]$, et qui vérifie :

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (1)$$

Chaque sous-ensemble $A \subseteq \Omega$, tel $m(A) > 0$, est appelé élément focal de m . La masse $m(A)$ représente le degré de croyance attribué à la proposition A et qui, compte tenu de la connaissance, ne peut être affectée à un sous ensemble plus spécifique que A .

Dans notre contexte, une compétence peut être soit acquise soit non acquise. Par conséquent, l'application de ce cadre formel se fait tel que pour chaque compétence C , le cadre de discernement est $\Omega(C) = \{Acquise(C), \neg Acquise(C)\}$. Par ailleurs, pour chaque compétence C , nous définissons les masses et notations suivantes :

- $m(a) = m(\{Acquise(C)\})$ la masse de croyance attribuée au fait que la compétence soit acquise par l'apprenant ;
- $m(\neg a) = m(\{\neg Acquise(C)\})$ la masse de croyance attribuée au fait que la compétence soit non acquise par l'apprenant ;
- $m(i) = m(\Omega)$ la masse de croyance attribuée à l'ignorance qu'a le système quant à l'acquisition ou la non-acquisition de la compétence par l'apprenant ;
- $m(c) = m(\emptyset)$ la masse de croyance attribuée au conflit résultant de la contradiction entre les observations pour la même compétence.

$$\text{Avec la contrainte : } m(a) + m(\neg a) + m(i) + m(c) = 1 \quad (2)$$

5 Processus de diagnostic :

Notre processus de diagnostic s'appuie sur quatre fonctions essentielles (cf. figure 2) :

- **La transformation** des notes d'évaluation en masses de croyances,
- **La révision** des masses de croyance à l'arrivée de nouvelles informations,
- **La propagation** des croyances aux compétences voisines,
- **La fusion des masses** des croyances attribuées à la même compétence.

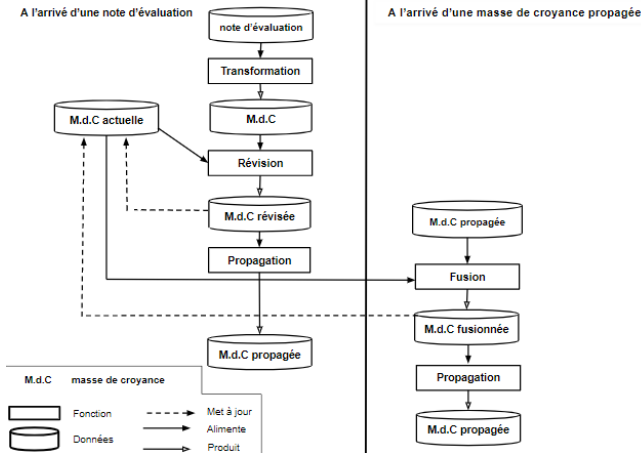


Fig. 1. Processus de diagnostic
 Nous décrivons chaque fonction en détail dans les sous-sections suivantes.

5.1 La fonction de transformation

Le module d'évaluation retourne des notes numériques pour chaque compétence. Ces notes ne sont pas directement exploitables par le module de diagnostic. Ce dernier utilise les masses de croyances pour déterminer l'état des compétences évaluées et prédire l'état des compétences non évaluées de l'apprenant. Par conséquent, il est nécessaire de transformer les évaluations numériques en croyances. Pour ce faire, nous définissons une fonction de transformation qui permet de définir l'opération permettant d'obtenir l'état de la compétence à partir d'une note d'évaluation. Cette fonction prend en entrée une valeur $ne \in [0, n]$, où 0 et n représentent respectivement la note minimale et maximale que l'apprenant peut avoir dans une évaluation. Elle prend également en entrée une valeur E qui correspond à la note minimale nécessaire afin de valider l'évaluation, et une valeur ε qui représente une marge de doute autour de E .

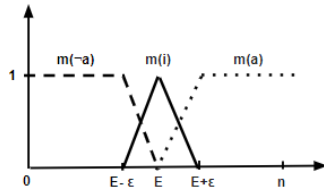


Fig. 2. Distribution de masses de croyance en fonction de la note d'évaluation

Nous pouvons distinguer 4 zones dans la figure :

- La zone 1 où $ne \in [0, E - \varepsilon[$. La distribution de masse est telle que : $m(a)=0$, $m(\neg a) = 1$, $m(i)=0$, $m(c)=0$. (3)

Nous considérons dans ce cas que la compétence est non acquise.

- La zone 2 où $ne \in [E - \varepsilon, E [$. La distribution de masse est telle que : $m(a) = 0$, $m(\neg a) = \frac{-ne}{\varepsilon} + \frac{E}{\varepsilon}$, $m(i) = \frac{ne}{\varepsilon} + 1 - \frac{E}{\varepsilon}$, $m(c) = 0$. (4)

Nous considérons dans ce cas que la compétence est probablement non acquise.

- La zone 3 où $ne \in] E, E + \varepsilon]$. Dans cette zone, la distribution de masse est telle que : $m(a) = \frac{ne}{\varepsilon} - \frac{E}{\varepsilon}$, $m(\neg a) = 0$, $m(i) = -\frac{ne}{\varepsilon} + 1 + \frac{E}{\varepsilon}$, $m(c) = 0$. (5)

Nous considérons dans ce cas que la compétence est probablement acquise.

- La zone 4 où $ne \in [E + \varepsilon, n [$. Dans cette zone, la distribution de masse est telle que : $m(a) = 1$, $m(\neg a) = 0$, $m(i) = 0$, $m(c) = 0$ (6)

Nous considérons dans ce cas que la compétence est acquise.

En plus de ces quatre zones, il existe le cas particulier où $ne = E$. Dans ce cas, la distribution est telle que : $m(a) = 0$, $m(\neg a) = 0$, $m(i) = 1$, $m(c) = 0$. (7)

Nous considérons dans ce cas que nous ignorons l'état de la compétence.

Prenons l'exemple d'un exercice sur 10 points ciblant une compétence particulière. L'enseignant considère que pour valider cet exercice, l'apprenant doit avoir au minimum la note de 5, avec une marge de doute de 2 points. On veut déterminer la croyance sur l'état de la compétence d'un apprenant qui a eu la note de 6 dans cette évaluation. La fonction de transformation indiquera dans ce cas la distribution de masse suivante :

$$m(a) = \frac{6}{2} + \frac{-5}{2} = \frac{1}{2}, m(\neg a) = 0, m(i) = -\frac{6}{2} + 1 + \frac{5}{2} = \frac{1}{2}, m(c) = 0$$

Nous considérons dans ce cas que la compétence est probablement acquise.

5.2 Fonction de révision

L'intérêt de cette fonction est de mettre à jour automatiquement l'état d'une compétence à l'arrivée d'une nouvelle information. Elle consiste à prendre en compte les nouvelles évaluations renvoyées par le module d'évaluation et qui viennent soit pour confirmer l'état de la compétence antérieure – donc l'apprenant a conservé son niveau – ou au contraire l'infirmer – le niveau de l'apprenant a évolué. On dit que l'apprenant a conservé son niveau si le nouvel état de la compétence est similaire à l'état actuel de cette compétence. A l'inverse, on dit que l'état de la compétence de l'apprenant a évolué si le nouvel état de la compétence est différent de son état actuel. De ce fait, la fonction de révision s'appuie sur deux mécanismes pour la révision des croyances que nous détaillons dans la suite.

Révision par une croyance similaire. Cela concerne le cas où la croyance actuelle est révisée par une croyance similaire. Dans ce cas, nous cherchons à maximiser la prise en compte de la nouvelle information tout en minimisant la perte de l'ancienne. Pour cela, nous utilisons la règle de révision interne [15]. Soit m_i et m_A deux distributions de masses distinctes pour la compétence C_p définies sur le même cadre de discernement

Ω avec m_I la distribution de masse de croyance initialement affectée à la compétence C_p , et m_A la distribution de masse de croyance affectée à la compétence C_p à partir de la nouvelle note d'évaluation, et \otimes l'opérateur de révision de la règle interne qui révisé m_I par m_A :

$$\begin{aligned} m_R(C_p) &= m_I(C_p) \otimes m_A(C_p) \\ &= \sum_{A \subseteq B} \sigma_A(A, B) * m_A(B) \\ \sigma_A(A, B) &= \left\{ \begin{array}{ll} \frac{m_I}{Bel(B)} & \text{pour } Bel(B) > 0 \\ 0 & \text{pour } Bel(B) > 0 \text{ et } A \neq B \\ 1 & \text{pour } Bel(B) = 0 \text{ et } A = B \end{array} \right\} \end{aligned} \quad (8)$$

Où $Bel(B)$ est la part de croyance spécifiquement allouée à B , avec $Bel(B) = \sum_{C \subseteq B, C \neq \emptyset} m(C)$.

L'application de cette règle permet de déterminer la distribution de masse de croyance révisée m_R :

$$\begin{aligned} m_R(a) &= \frac{m_I(a)}{m_I(a) + m_I(\neg a) + m_I(i)} * m_A(i) + m_A(a) \\ m_R(\neg a) &= \frac{m_I(\neg a)}{m_I(a) + m_I(\neg a) + m_I(i)} * m_A(i) + m_A(\neg a) \\ m_R(i) &= m_I(i) * m_A(i) \\ m_R(c) &= 1 - m_R(a) - m_R(\neg a) - m_R(i) \end{aligned} \quad (9)$$

Révision par une croyance différente. Cela concerne le cas où la croyance actuelle est révisée par une nouvelle croyance différente. Dans ce cas, nous proposons d'accorder la même importance à la nouvelle et l'ancienne masse de croyance. De ce fait, l'utilisation de la moyenne non associative nous paraît pertinente. Soit m_I et m_A deux distributions de masses distinctes pour la compétence C_p définies sur le même cadre de discernement Ω avec m_I est la distribution de masse de croyance initialement affectée à la compétence C_p , et m_A est la distribution de masse de croyance affectée à la compétence C_p à partir de la nouvelle note d'évaluation, et σ l'opérateur de révision de la moyenne non associative qui révisé m_I par m_A .

$$\begin{aligned} m_R(C_p) &= m_I(C_p) \sigma m_A(C_p) \\ &= \frac{1}{2}(m_I(C_p) + m_A(C_p)) \end{aligned} \quad (10)$$

L'application de cette règle permet de déterminer la distribution de masse de croyance révisée m_R :

$$\begin{aligned} m_R(a) &= \frac{m_I(a) + m_A(a)}{2} \\ m_R(\neg a) &= \frac{m_I(\neg a) + m_A(\neg a)}{2} \\ m_R(c) &= \frac{m_I(c) + m_A(c)}{2} \\ m_R(i) &= 1 - m_R(a) - m_R(\neg a) - m_R(c) \end{aligned} \quad (11)$$

5.3 La fonction de propagation de croyance :

L'objectif ultime du modèle de diagnostic est d'établir un diagnostic global des compétences de l'apprenant. Il ne s'agit pas seulement de diagnostiquer les compétences

directement évaluées, mais également de prédire l'état des compétences n'ayant pas été évaluées. Nous proposons alors une fonction de propagation des croyances entre les compétences du domaine. Nous avons évoqué dans la section précédente les relations de prérequis entre les compétences. Nous émettons l'hypothèse que la propagation des croyances entre ces compétences implique une perte d'informations sur la distribution de masses de croyances $m(a)$, $m(\neg a)$, $m(i)$ et $m(c)$ selon la distance qui les sépare. Afin de quantifier cette perte d'information, nous utilisons la règle d'affaiblissement de [16] :

$$\begin{aligned} m^\alpha(A) &= (1-\alpha) * m(A), \quad \forall A \subset \Omega \\ m^\alpha(\Omega) &= (1-\alpha) * m(\Omega) + \alpha \end{aligned} \quad (12)$$

Avec : $0 < \alpha < 1$

α Correspond au degré d'affaiblissement.

$1-\alpha$ Correspond au degré de fiabilité.

Nous définissons la fonction de propagation f qui prend en paramètres la compétence source et la compétence ciblée par la propagation. Cette fonction retourne une distribution de masse de croyance affaiblie. Nous proposons que l'affaiblissement soit calculé en fonction de la distance qui sépare les compétences. Ainsi, nous définissons la fonction comme suit :

$$\begin{aligned} m_2(a) &= (1-\alpha)^d \times m_1(a) \\ m_2(\neg a) &= (1-\alpha)^d \times m_1(\neg a) \\ m_2(i) &= (1-\alpha)^d \times m_1(i) \\ m_2(c) &= (1-\alpha)^d \times m_1(c) \end{aligned} \quad (13)$$

5.4 La fonction de fusion

Lorsqu'une propagation de croyances a lieu, les compétences concernées reçoivent en entrée une croyance propagée qui vient compléter la masse de croyance locale. Par conséquent, pour déterminer le nouvel état de ces compétences, il est nécessaire de fusionner ces masses de croyances. La fusion de ces croyances permet d'enrichir le modèle de diagnostic, et donc d'affiner la scénarisation. Par exemple, si la croyance propagée montre que cette compétence est acquise et si la croyance locale déterminée à partir d'une évaluation directe montre totalement le contraire, cela signifie qu'il existe une incohérence dans l'état de la compétence pour cet apprenant. Par conséquent, une évaluation supplémentaire serait souhaitable pour lever cette incohérence. Afin de réaliser cette fusion, nous utilisons la règle de combinaison conjonctive de Shafer [16], qui permet de fusionner les informations en maximisant la concordance entre les sources d'information.

Soit m_i et m_p deux croyances distinctes pour une compétence donnée définies sur le cadre de discernement Ω . \oplus est l'opérateur de fusion de la règle de combinaison conjonctive [16] :

$$\begin{aligned} m_f(A) &= m_i(A) \oplus m_p(A) \\ m_f(A) &= \sum_{B \cap C = A} m_i(B) * m_p(C) \end{aligned} \quad (14)$$

Nous définissons la fonction de fusion F qui prend en paramètre les différentes masses de croyances associées à une compétence (propagée ou locale) et retourne une masse de croyance fusionnée m_f . Où :

$$m_f(a) = m_i(a) * m_p(a) + m_i(i) * m_p(a) + m_p(i) * m_i(a)$$

$$\begin{aligned}
 m_f(\neg a) &= m_l(\neg a) * m_p(\neg a) + m_l(i) * m_p(\neg a) + m_p(i) * m_l(\neg a) \\
 m_f(i) &= m_l(i) * m_p(i) \\
 m_f(c) &= 1 - m_f(a) - m_f(\neg a) - m_f(i)
 \end{aligned}
 \tag{15}$$

Cet opérateur de fusion est commutatif, l'ordre dans lequel on fusionne les informations n'est pas important, il génère du conflit en donnant l'information sur la quantité de la contradiction entre les sources d'informations.

6 L'expérimentation

L'objectif de notre expérimentation est d'évaluer la pertinence des diagnostics réalisés par le module de diagnostic en les comparant avec les diagnostics réalisés par un expert du domaine. Cette expérimentation a été réalisée à l'Université de Technologie de Compiègne avec des étudiants en première année de préparation intégrée. Elle a rassemblé 28 étudiants guidés par un enseignant qui assure le cours d'algorithmique depuis 15 ans. Dans cette expérimentation, nous avons proposé des exercices afin d'évaluer séparément chacune des compétences suivantes : Rechercher un élément dans un tableau ; Parcourir un tableau ; Maîtriser la boucle For ; Maîtriser la boucle while ; Maîtriser la boucle do-while. Ces compétences sont reliées entre elles avec les relations de précedence telles qu'illustrées dans la figure 3. La compétence « Parcourir un tableau » est un prérequis de la compétence « Rechercher un élément dans un tableau ». Les compétences « maîtriser les boucles {for, while, do-while} » sont des prérequis de la compétence « parcourir un tableau ».

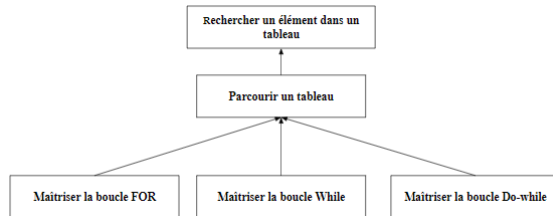


Fig. 3. Hiérarchie des compétences

Nous avons proposé de comparer la prédiction du module de diagnostic pour la compétence « Rechercher un élément dans un tableau » avec le diagnostic direct de cette compétence par l'expert du domaine. Pour ce faire, nous avons proposé aux étudiants deux séries d'exercices afin d'évaluer ces compétences. Le module de diagnostic reçoit les évaluations de chacune des compétences suivantes: « Parcourir un tableau », « Maîtriser la boucle For », « Maîtriser la boucle While », « Maîtriser la boucle Do-while », et à partir de ces évaluations et par sa fonction de propagation, il génère des prédictions sur l'état de la compétence « Rechercher un élément dans un tableau ». Ensuite nous faisons une comparaison entre ces prédictions et l'évaluation directe de l'expert. Pour effectuer cette comparaison, nous proposons la mesure de précision suivante:

$$\text{Mesure de précision} = \frac{\text{Nombre de prédictions pertinentes}}{\text{Nombre total de prédiction}}$$

Avec *Nombre de prédictions pertinentes* est le nombre des prédictions réalisées par le module de diagnostic pour la compétence « Rechercher un élément dans un tableau » similaires aux diagnostics de l'expert pour cette compétence.

Le Tableau 1 récapitule les résultats obtenus.

Table 1. Résultats de l'expérimentation

Nombre total de prédictions	Prédictions pertinente	Mesure de précision
28	23	0,82

6.1 Discussion :

Comme nous pouvons le voir dans le tableau ci-dessus, l'expérimentation révèle que pour un nombre total de 28 prédictions établies par notre module de diagnostic, 23 d'entre elles sont similaires à celles de l'expert. En revanche, 5 prédictions de le sont pas. En effet, les notes d'évaluations attribuées par l'expert, transformées par le module de diagnostic, montrent que la compétence « Rechercher un élément dans un tableau » n'est pas acquise pour ces cinq étudiants. Tandis que le module de diagnostic montre que cette compétence est probablement acquise. Selon notre expert, ces étudiants ont bien utilisé des boucles. Cependant ils ne les ont pas bien exploités pour résoudre l'exercice associé à la compétence « Rechercher un élément dans le tableau ».

Cette différence entre le diagnostic de l'expert et le module pourrait s'expliquer, entre autres, par les valeurs de la distance et/ou du taux d'affaiblissement sur lesquelles repose le module pour établir les prédictions.

7 Conclusion et perspectives :

Dans cet article, nous avons présenté notre proposition pour un module de diagnostic des compétences de l'apprenant dans un EIAH dédié à l'apprentissage de l'algorithmique. Ce module s'appuie sur la théorie des fonctions de croyances, pour quantifier l'incertitude, l'imprécision et la contradiction dans le processus de diagnostic. Ce processus repose sur quatre fonctions essentielles qui permettent à l'arrivée d'une nouvelle information, de mettre à jour, fusionner et propager les croyances sur l'état des compétences. Notre module a fait l'objet d'une première expérimentation auprès d'étudiants en première année d'études postbac, et les résultats sont encourageants. Une de perspectives de ce travail serait d'alléger encore plus la tâche de l'expert en calculant automatiquement les paramètres nécessaires à la fonction de propagation.

Références

1. M. Lefevre, J. Broisin, V. Butoianu, P. Daubias, L. Daubigney, et al. : Personnalisation de l'apprentissage : comparaison des besoins et approches à travers l'étude de quelques dispositifs. STICEF, ATIEF, 19, pp.353-387, (2012).
2. T. Carron and J.-C. Marty. : User Modelling in Learning Games BeIng Project View project Creating a Learning Lab View project. Proc. of the 3rd European Conference on Games Based Learning, Graz, Austria, pp. 86–94, (Oct. 2009).
3. A. Abyaa, M. Khalidi Idrissi, et S. Bennan. : Learner modelling: systematic review of the literature from the last 5 years. vol. 67, no. 5. Springer US,(2019).
4. A. T. Corbett and J. R. Anderson, Knowledge Tracing, User modeling and user-adapted interaction. vol. 4, no. 4. pp. 253–278,(1995).
5. M. V. Yudelson, K. R. Koedinger, and G. J. Gordon. : Individualized bayesian knowledge tracing models. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7926 LNAI, pp. 171–180, (2013).
6. C. J. Butz, S. Hua, and R. B. Maguire. : A web-based bayesian intelligent tutoring system for computer programming,. Web Intell. Agent Syst., vol. 4, no. 1, pp. 61–81, (2006).
7. G. Gogudze, S. Sosnovsky, S. Isotani, and B. M. McLaren. : Towards a Bayesian student model for detecting decimal misconceptions. Proc. 19th Int. Conf. Comput. Educ. ICCE 2011, no. January, pp. 34–41,(2011).
8. C. Y. Ting and S. Phon-Amnuaisuk. : Properties of Bayesian student model for INQPRO. Appl. Intell., vol. 36, no. 2, pp. 391–406, (2012).
9. A. Jameson. : Numerical uncertainty management in user and student modeling: An overview of systems and issues,. User Model. User-Adapted Interact. J. Pers. Res., vol. 5, no. 3, pp. 193–251, (1995).
10. D. Xu, H. Wang, and K. Su. : Intelligent student profiling with fuzzy models. In Proc. Annu. Hawaii Int. Conf. Syst. Sci., vol. 2002-Janua, no. February,(2002).
11. G. Tsaganou, M. Grigoriadou, T. Cavoura, and D. Koutra. : Evaluating an intelligent diagnosis system of historical text comprehension..Expert Systems with Applications, , pp. 493-502.(2003)
12. K. Carpentier. : Scénarisation personnalisée dynamique dans les environnements virtuels pour la formation., p. 271, (2015).
13. Belhaoues T.E., Bensebaa T., Abdessem M., Bey A. : AlgoSkills: An ontology of Algorithmic Skills for exercises description and organization. Journal of e-Learning and Knowledge Society, v.12, (2012).
14. P. Smets and R. Kennes. : The transferable belief model. Artif. Intell., vol. 66, no. 2, pp. 191–234, (1994).
15. J. Ma, W. Liu, D. Dubois, and H. Prade. : Revision rules in the theory of evidence. Proc. - Int. Conf. Tools with Artif. Intell. ICTAI, vol. 1, pp. 295–302, (2010).
16. G. Shafer. : A Mathematical Theory of Evidence. Princeton, New Jersey, USA: Princeton Univ. Press, (1976).

Les modèles apprenant pour soutenir l'apprentissage tout au long de la vie : revue de littérature

Sergio Iván Ramírez Luelmo, Nour El Mawas et Jean Heutte

CIREL - Centre Interuniversitaire de Recherche en Éducation de Lille, Université de Lille,
Villeneuve d'Ascq, France
{sergio.ramirez-luelmo, nour.el-mawas, jean.heutte}@univ-lille.fr

Résumé. De nos jours, l'analyse de l'apprentissage ou Learning Analytics (LA) est un sujet émergent dans les domaines des Environnements Informatiques pour l'Apprentissage Humain (EIAH) et de l'Apprentissage Tout au Long de la Vie (ATLV). Les modèles apprenant jouent un rôle essentiel dans l'utilisation et l'exploitation des données générées par les apprenants dans une variété d'EIAH dans le but de faciliter le suivi de l'apprenant et la personnalisation du contenu du cours. Dans cet article, nous nous concentrons sur la modélisation de l'apprenant dans les MOOC (Massive Open Online Course) dans un contexte d'ATLV. Cette étude nous permettra de comparer et de mettre en évidence des caractéristiques des modèles apprenant existants pour un MOOC dans une perspective d'ATLV. Ce travail est dédié aux concepteurs/fournisseurs de MOOC, aux ingénieurs pédagogiques et aux chercheurs qui souhaitent avoir un panorama des travaux existants sur la modélisation des apprenants des MOOCs sur la base des LA.

Mots-clés : apprentissage tout au long de la vie, modèle d'apprenant, revue de littérature, MOOC.

Abstract. Nowadays, Learning Analytics (LA) is an emerging topic in the fields of TEL and Lifelong Learning (LLL). Learner Models (LM) play an essential role in the use and exploiting of learners' generated data in a wide array of TELs, with the goal of facilitating the learners' follow-up, and course content personalization. In this research, we focus on learner modelling in MOOCs (Massive Open Online Course) in a LLL context. This study will allow us to compare and highlight characteristics of existing LMs for MOOC in a LLL perspective. This work is dedicated to MOOC designers/providers, and researchers wishing to have a panorama on MOOC learner modelling based on LA.

Keywords: Lifelong learning, learner model, literature review, MOOC.

1 Introduction

La popularité et portée mondiale des MOOCs découlent de leur concept original qui consiste à offrir des cours en libre accès à un grand nombre d'apprenants du monde entier et à leurs coûts, souvent faibles, même gratuits [1]. Pourtant, ils souffrent d'un taux d'abandon très élevé d'environ 93,5 % [2], [3]. Ainsi, l'amélioration de la réussite scolaire dans les MOOCs (en augmentant les résultats d'apprentissage et en diminuant le taux d'abandon moyen des apprenants) crée le besoin de personnaliser le contenu et les parcours d'apprentissage à travers la modélisation de l'apprenant [2], [4].

Dans le contexte actuel de Big Data, il est difficile de se faire une idée précise du paysage de la recherche sur les travaux relatifs aux modèles apprenant pour les MOOCs dans un contexte d'ATLV. Ainsi, la motivation derrière cette étude est de présenter au public cible de cet article un panorama plus adéquat et plus précis à ce sujet.

La suite de l'article est découpée comme suit. La section 2 de cet article fournit le cadre théorique applicable, notamment le concept des modèles apprenant, leur rôle dans les MOOC et l'importance de la dimension de l'ATLV comme contexte environnant. La section 3 détaille les étapes de la méthodologie et discute des résultats de cette revue de la littérature. Enfin, la section 4 conclut ce document et présente ses perspectives.

2 Cadre théorique : le modèle apprenant

Les modèles apprenant représentent les croyances du système sur les caractéristiques spécifiques de l'apprenant [5], [6]. Ils peuvent ou pas inclure les préférences personnelles, les états cognitifs, ou les préférences en matière d'apprentissage et de comportement [7]. La modélisation de l'apprenant a pour objectif ultime de permettre l'adaptation et la personnalisation des environnements et des activités d'apprentissage, y compris dans les MOOCs [8], [9]. De plus, la personnalisation de l'apprenant est l'un des concepts essentiels dans le contexte d'ATLV [10], [11].

Au cours du processus d'initialisation du modèle apprenant, il est possible de rencontrer la problématique de démarrage à froid (*cold-start problem*), où les données disponibles sur l'apprenant sont insuffisantes pour instancier correctement le modèle. Une situation similaire appelée « données parcimonieuses » (*data sparsity*) peut également se produire pendant la phase de mise à jour, empêchant la mise à jour correcte du modèle apprenant [12].

Les modèles ouverts de l'apprenant (*Open Learner Models* - OLM) sont un type de modèle apprenant où le modèle est explicitement communiqué à l'apprenant [9], [13]. Ils peuvent être classés en trois catégories, en fonction du mode d'édition et de communication du modèle : négociable, éditable ou visualisable [13]. Un OLM négociable demandera et vérifiera des preuves factuelles pour accepter toute modification, alors qu'un OLM éditable exigera plutôt un ensemble d'autorisations pour permettre des changements. Un OLM visualisable, en revanche, ne permet pas de modifications, laissant uniquement ses mécanismes de mise à jour au système d'hébergement [6]. Concernant le système d'hébergement lui-même, [14] soulève l'idée que la transparence du fonctionnement du système est une caractéristique souhaitable pour l'OLM car elle

contribue à « engendrer la confiance, à permettre la détection des erreurs et à favoriser l'apprentissage ». Les auteurs de [15] montrent l'importance des modèles apprenants indépendants de tout système et capables d'intégrer des données provenant de plusieurs sources, via des connecteurs standardisés. L'indépendance d'un modèle apprenant permet de prendre possession de ses propres données et de les contrôler [9]. Par ailleurs, nous reconnaissons que la manière dont les données sont représentées (la représentation des connaissances) dans le modèle apprenant est étroitement liée aux mécanismes de traitement de ces données (souvent, le système de recommandation/prédiction) [16]. Dans un contexte d'ATLV, le modèle apprenant doit être capable de s'adapter dynamiquement aux multiples sources (et types) de données qui vont s'y intégrer [17]. Ce dynamisme, joint à son degré d'interopérabilité, fait du modèle apprenant un portfolio de connaissances, compétences, préférences, accréditations, etc., comme l'exige le contexte d'ATLV. Ainsi, les OLM sont un moyen de rendre les apprenants autonomes tout en leur permettant de consulter leur modèle apprenant, de le partager et de le mettre à jour.

Grâce à l'étude du cadre théorique du modèle apprenant, nous avons identifié les caractéristiques (Cs) pertinentes pour le choix d'un modèle apprenant dans un MOOC pour l'ATLV: (C1) une approche de connexion à une plateforme, (C2) la gestion du démarrage à froid, (C3) la gestion des données parcimonieuses, (C4) la représentation des connaissances de l'apprenant, (C5) la méthode de recommandation/prédiction, (C6) l'ouverture du modèle apprenant, (C7) le dynamisme du modèle apprenant et, (C8) sa considération de l'ATLV.

Dans la section suivante, nous détaillons la méthodologie de revue de littérature pour repérer les modèles apprenant pour les MOOCs dans un contexte d'ATLV

3 Méthodologie de revue de littérature

Cette revue de la littérature suit la méthodologie décrite par [18], comporte les étapes suivantes : [A] Identification du besoin d'une revue de littérature, [B] Développement du protocole d'analyse, [C] Identification des questions de recherche, [D] Identification des bases de données de recherche, [E] Lancement de la recherche et enregistrement des citations, [F] Sélection des articles, [G] Résumé des articles sélectionnés et, [H] Interprétation des résultats.

3.1 Identification de la nécessité d'une revue de littérature [A] et élaboration du protocole d'analyse [B]

L'objectif de cette revue de la littérature est d'analyser les travaux les plus récents dans le domaine de modèles apprenant pour MOOCs dans un contexte d'ATLV. À notre connaissance, il n'existe actuellement aucun travail de recherche qui aborde la revue de littérature des modèles apprenant existants pour MOOCs dans ce contexte. D'après [18], le développement du protocole comprend « [...] tous les éléments de l'analyse plus quelques informations supplémentaires sur la planification [...] » que nous détaillons dans les étapes suivantes.

3.2 Identification des questions de recherche [C]

Cet article vise donc à répondre à la question de recherche suivante : Quels sont les modèles apprenant les plus appropriés pour les MOOCs dans un contexte d'ATLV ?

3.3 Critères de sélection et identification des bases de données de recherche [D]

Dans cette section, nous décrivons les critères d'inclusion et d'exclusion utilisés pour constituer le corpus de publications pour notre analyse. Nous détaillons et justifions également notre choix des termes de recherche, ainsi que des bases de données identifiées.

D'une part, nos critères d'inclusion sont : les travaux qui portent sur un modèle apprenant, ou qui présentent un nouveau modèle apprenant en le comparant à un autre modèle apprenant déjà existant (pour un MOOC, dans les deux cas). D'autre part, les critères d'exclusion que nous avons choisis sont les suivants : les travaux qui ne sont pas écrits en anglais, sous embargo, non publiés, et ceux qui ne traitent pas directement ou de façon périphérique les modèles apprenant. Nous ne conservons que la dernière contribution publiée (même auteur, même sujet, même année) et enfin, les travaux publiés dans des revues ont la priorité sur ceux des conférences.

Nous avons effectué ces recherches au cours des cinq dernières années (2015-2020) dans les bases de données scientifiques suivantes : Web of Science, Scopus et Google Scholar. Parmi les résultats de Google Scholar, nous sommes principalement intéressés par les résultats de Taylor & Francis Online, Science Direct, Sage Publications, Springer et IEEE Explore. L'idée derrière ces choix est d'avoir les travaux scientifiques les plus actuels et de qualité éprouvée dans le domaine.

Nos termes de recherche étaient "learner model" et "mooc", ce qui dans la plupart des moteurs de recherche se traduit par une requête booléenne de la forme $\{ (\text{learner} * \text{AND model} *) \text{AND} (\text{mooc} *) \}$. À noter que cette requête renvoie également les résultats au pluriel, au gérondif et les substantifs dérivés.

3.4 Lancement de la recherche [E]

Pour un processus de sélection des articles plus rationnel, nous avons conçu et développé un outil externe (publication en cours) qui récupère et organise les métadonnées d'une liste de sources académiques et les présente à l'examineur dans un contexte non biaisé (étape de collecte automatique de métadonnées, *automatic metadata collection*). Cet outil externe nous a facilité la visualisation, le stockage, le chargement et le triage des résultats à partir principalement des résumés des publications.

Le processus de sélection du papier est illustré dans la Figure 1 et s'est déroulé comme suit : tout d'abord, nous avons utilisé un fichier CSV comme pôle de concentrateur des données pour conserver les résultats de la recherche issus du moteur de recherche Google Scholar ; la base de données Scopus ; et la base de données Web of Science. Ensuite, nous avons automatiquement extrait les métadonnées pertinentes liées

aux résultats précédents (résumés et mots clés) à partir des pages Web ou des fichiers PDF des publications.

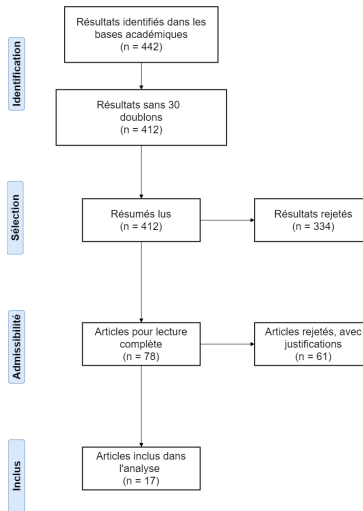


Fig. 1. : Vue d'ensemble du processus de sélection et de catégorisation des publications.

3.5 Sélection des articles [F]

Ultérieurement, nous avons lu tous les résumés extraits automatiquement et les avons catégorisés. Les publications dont le résumé était hors sujet du domaine de cet article ont été rejetées, tout en enregistrant leurs thématique principale¹. Sur les 419 résultats de Google Scholar, 342 publications ont été rejetées, car elles ne répondaient pas aux critères d'inclusion. 77 publications issues de Google Scholar ont été jugées « satisfaisantes » pour une lecture complète. Les deux autres moteurs (Scopus et Web of Science) ont fourni peu de résultats : un seul résultat « satisfaisant » issu de Scopus et aucun de Web of Science.

Ensuite, nous avons procédé à la lecture complète des publications « satisfaisantes » par le biais de notre abonnement institutionnel. Parmi ces 78 (77 Google Scholar + 1 Scopus + 0 Web of Science) articles, seuls 17 sont devenus des « propositions » (16 Google Scholar, 1 Scopus, 0 Web of Science). Le groupe d'articles rejetés était composé de publications de doctorat et de master, d'une publication non anglophone et d'articles ne répondant pas correctement aux critères d'inclusion.

¹ Par exemple: “ethical concerns of AI in education”, “panorama on open source LMS” or “evolution of higher education”.

Par ailleurs, nous avons pu identifier (et rejeter) 19 publications qui soit utilisaient les termes 'modèle apprenant' et 'profil apprenant' de manière interchangeable, soit erronée. Dans notre contexte, le profil apprenant et le modèle apprenant sont différents dans la mesure où le premier peut être considéré comme une simple information statique non interprétée sur l'apprenant [19].

Le processus de sélection s'est donc conclu par 17 « Propositions » de modèles apprenant, présentées dans l'Annexe. Dans la section suivante, nous considérons comment les modèles apprenants sélectionnés peuvent répondre aux enjeux et caractéristiques examinées dans le cadre théorique, à savoir pourquoi et comment elles se regroupent en différentes dimensions comparatives.

4 Résultats [G] et Interprétation [H]

Un tableau récapitulatif des 17 modèles apprenant trouvés et de leurs caractéristiques est présenté en Annexe. Dans cette section, nous examinons comment les caractéristiques présentes dans le cadre théorique pondèrent la composition des dimensions permettant de comparer les modèles apprenant sélectionnés.

Tableau 1. Dimensionnement des caractéristiques considérées.

Dim.	☐	Conditions d'octroi / rejet
I	✓	Si un connecteur de plateforme standardisé (C1) est spécifié.
	-	S'il y a mention uniquement de la plateforme d'implémentation du modèle apprenant.
	X	Si aucun des deux cas préalables.
D	✓	S'il existe une gestion détaillée de la gestion du démarrage à froid (C2) et des données parcimonieuses (C3).
	-	S'il existe une gestion détaillée d'au moins C2 ou de C3.
	X	Tous les autres cas.
K	✓	S'il y a mention détaillée de la représentation des connaissances de l'apprenant (C4) et dans la méthode de recommandation/prédiction (C5).
	-	Si C4 est détaillée.
	X	Tous les autres cas.
ATLV	✓	- Si une approche pour l'ATLV (C8) est explicitement détaillée, ou - Si un OLM est considéré (C6) et une mention détaillée du dynamisme (C7)
	-	Si un OLM est considéré (C6)
	X	Tous les autres cas.

La lecture du cadre théorique du modèle apprenant nous a permis d'identifier les huit caractéristiques (Cs) pertinentes pour le choix d'un modèle apprenant dans un MOOC pour l'ATLV (voir Section 2).

Nous avons agrégé ces huit caractéristiques en quatre dimensions : la dimension 'interopérabilité' (I) correspond à la caractéristiques C1, la dimension 'traitement des données parcimonieuses' (D) correspond aux C2 et C3, la 'représentation des connais-

sances' (K) correspond aux C4 et C5 et finalement la dimension 'ATLV' (ATLV) correspond aux C6, C7 et C8. La dimension de l'interopérabilité (I) indique si le modèle apprenant permet l'utilisation de connecteurs standard pour des systèmes d'hébergement externes.

Tableau 2. Comparaison des modèles apprenant sélectionnés selon les dimensions proposées.

Modèle apprenant	Référence	I	D	K	ATLV
TrueLearn	[20]	-	-	✓	✓
SBGf	[21]	✓	-	x	x
MOOCIm	[22]	✓	x	x	✓
STyLE-OLM	[23]	✓	-	✓	✓
SansNom-MOOC TAB	[24]	✓	-	✓	✓
SansNom-Tunis	[25]	x	x	✓	x
SansNom-Chine	[26]	-	-	x	x
EDUC8	[27]	-	-	✓	x
DiaCog	[28]	x	x	✓	x
SansNom-Chine	[29]	x	x	x	x
SansNom-ODALA	[30]	✓	-	✓	✓
SansNom-Tunis-France	[31]	-	x	✓	x
GAF	[32]	x	-	x	x
AUM (AeLF User Model)	[33]	✓	-	✓	✓
MLaaS	[34]	x	x	x	x
Logic-Muse	[35]	-	-	✓	✓
SansNom-Adaptive Hypermedia	[36]	x	x	x	x

La dimension Données parcimonieuses (D) indique si une approche donnée est envisagée en cas de données manquantes. La dimension de représentation des connaissances (K) concerne le niveau de détails pris en compte dans la représentation des connaissances de l'apprenant et/ou du domaine ainsi que les mécanismes utilisés pour mettre à jour le modèle apprenant ou pour recommander/personnaliser le contenu. La représentation des connaissances est une caractéristique importante car elle est étroitement liée à la façon dont le modèle apprenant conserve son intégrité et/ou prédit ou suggère des états intermédiaires. Enfin, la dimension d'ATLV illustre la façon dont le modèle apprenant est préparé à faire face aux exigences d'un contexte d'apprentissage tout au long de la vie. Ces dimensions sont des points clés à prendre en compte lors du choix d'un modèle apprenant pour un MOOC dans un contexte d'ATLV.

La façon dont les caractéristiques sont agrégées en dimensions dépend de la considération explicite² par les auteurs de la (des) méthode(s) utilisée(s) pour renforcer les caractéristiques proposées : voir le Tableau 1.

² Si la publication ne fournit pas une justification textuelle, explicitement citable traitant l'inclusion / exclusion de la caractéristique en question, nous l'avons considérée comme pas traitée du tout.

Une fois le dimensionnement défini, nous avons évalué les 17 publications sélectionnées et leur avons octroyé des [✓], [X] ou [-], d'après leurs caractéristiques. Le Tableau 2 présente les résultats de cette évaluation.

Nous poursuivons avec l'interprétation de résultats de nos travaux et donc les modèles apprenants les plus pertinents pour un MOOC dans un contexte d'ATLV.

Lorsque l'on examine les techniques mises en œuvre pour représenter les connaissances, nous avons constaté que les règles (ou d'autres méthodes similaires) sont l'approche privilégiée pour les systèmes de recommandation mentionnés dans [21], [22], [25], [27]–[30], [33]. Le choix de stratégies bayésiennes vient en 2ème position ([20], [24], [32], [35]). Les ontologies les suivent de près [25], [27], [30]. Les graphes conceptuels [23], l'apprentissage automatique [34], les corrélations de Pearson [26] et les méthodes de regroupement des k-means [29] sont peu utilisés, un seul article présentant chacune de ces techniques. À noter que certaines propositions utilisent une combinaison de ces techniques et d'autres techniques ad hoc.

En ce qui concerne l'interopérabilité des modèles apprenant examinés, l'utilisation des standards est limitée. La plupart des modèles apprenant ne mentionnent pas leur méthode de communication ou connecteur de plateforme, sauf pour le cas des ontologies pour la représentation des connaissances [25], [30], et pour xAPI comme protocole de communication [30], [33]. Mis à part les plateformes ad-hoc, edX a été utilisé deux fois [22], [24], Moodle, Coursera et Claroline étant mentionnés une fois chacun. Dans tous les cas, la dimension d'interopérabilité constitue un défi que la plupart des modèles apprenant évitent ou contournent via une solution ad hoc.

En outre, l'approche face au traitement des données parcimonieuses envisagée était généralement mal définie ; le problème du démarrage à froid était abordé via un 'questionnaire de départ', un 'enregistrement' ou une 'saisie de données externes'. De plus, aucun des articles ne prenait en considération le problème de données parcimonieuses.

Nous avons constaté que le contexte d'ATLV n'est pas encore pris en compte de manière explicite par la plupart des concepteurs de modèles apprenant : cinq publications abordent la question de manière très générale. Cependant, une publication [33] s'est distinguée du reste en fournissant des détails sur la mise en œuvre technique afin de remplir cette dimension (OpenID). Par ailleurs, un OLM n'est pas reconnu comme partie intégrante d'un contexte d'ATLV. [20], [22]–[24], [33] ont favorisé les OLM négociables et totalement ouverts. Ainsi, malheureusement, l'ATLV n'est pas une priorité pour les concepteurs de modèles apprenants, dont les propositions mettent l'accent sur l'application d'une nouvelle technique (machine learning) ou sur un moyen de diffusion de contenu spécifique (la vidéo pour l'apprentissage mobile).

Parmi ces 17 publications, sept [20], [22]–[24], [30], [33], [35] satisfont la dimension d'ATLV. Cinq de ces sept publications ont également pris en compte la dimension I. Néanmoins, seules quatre propositions de modèle apprenant finalement retenues ([23], [24], [30], [33]) fournissent les méthodes explicites de représentation des connaissances et de mise à jour du modèle apprenant nécessaires dans un contexte d'éducation et de formation tout au long de la vie. Nous pouvons en déduire que ces publications fournissent suffisamment d'éléments probants (dimensions I, D, K et ATLV) pour affirmer que leur proposition de LM est le choix le plus approprié pour un MOOC dans un con-

texte d'ATLV. Nous pouvons ainsi affirmer que la réponse à notre question de recherche (voir Section 3.2) est représentée dans les quatre modèles apprenants issus des quatre publications sélectionnées, dont les caractéristiques sont détaillées dans le Tableau 3, avec les plateformes utilisées ('?' indiquant une caractéristique non abordée).

Tableau 3. Modèles apprenant les plus appropriés pour un MOOC soutenant l'ATLV.

Caractéristique / Modèle apprenant	STyLE-OLM, [23]	SansNom-MOCTAB, [24]	SansNom-ODALA, [30]	AUM (AeLF User Model), [33]	
C1	Web Ontology Language	xAPI	OWL	xAPI	
Plateformes	Onto-AIMS, PORML, SAIC	edX	Ad-hoc	?	
C2	« géré »	« questionnaire de positionnement et accès au profil LinkedIn »	« Génération de profil initiale et résultats du test de styles d'apprentissage »	Indices d'un questionnaire de positionnement	
C3	?	?	?	?	
C4	Graphes conceptuels	con-	Modèle apprenant bayésien	Onto_TDM	LOM
C5	Graphes conceptuels	con-	Modèle apprenant bayésien	Formulas	Règles pédagogiques
C6	Ouvert	Négotiable	Édition sur Information générale, style d'apprentissage et préférences	?	
C7	?	Dynamique	États cognitifs	« croissance dynamique de la connaissance de l'apprenant »	
C8	Centrée formation des adultes	Abordé	?	OpenID	

5 Conclusion et perspectives

Cette revue de la littérature aborde les caractéristiques les plus pertinentes pour un modèle apprenant pour des MOOC dans un contexte d'ATLV : sur un ensemble de 442 publications, 17 articles ont été examinés, leurs caractéristiques ont été mises en évidence, agrégées et comparées, pour aboutir à quatre finalistes.

Cette étude nous a conduit à considérer huit caractéristiques comme des points clés à prendre en compte lors du choix d'un modèle apprenant pour un MOOC dans un contexte d'ATLV. Nous avons synthétisé ces huit caractéristiques en quatre dimensions, à savoir l'interopérabilité (I), le traitement des données parcimonieuses (D), la représentation des connaissances (K) et la prise en compte de l'ATLV (ATLV). Les quatre finalistes de cette étude [23], [24], [30], [33] ont rempli la plupart, mais pas la

totalité, de nos critères de comparaison. Nous avons conclu que leurs propositions de modèle apprenant étaient les plus appropriées pour un MOOC dans un contexte d'ATLV.

Notre prochaine étape de recherche consiste à mettre en place les quatre modèles apprenant finalistes dans le MOOC « Gestion de Projet »³ afin de comparer leurs résultats. Nous sommes convaincus que des acteurs autres que les concepteurs/fournisseurs de MOOC et des modèles apprenant, les ingénieurs pédagogiques et les chercheurs, peuvent tirer profit de cette étude pour les aider à évaluer les caractéristiques des modèles apprenant pour MOOC dans un contexte d'ATLV.

Remerciements

Ce projet a été réalisé avec le soutien financier de l'État français dans le cadre du Programme « Investissements d'avenir » (I-SITE ULNE / ANR-16-IDEX-0004 ULNE) géré par l'Agence Nationale de la Recherche française (ANR).

Références

1. A. M. F. Yousef, M. A. Chatti, U. Schroeder, M. Wosniza, et H. Jakobs, « MOOCs - A Review of the State-of-the-Art », in *Proceedings of the 6th International Conference on Computer Supported Education*, Barcelona, Spain, 2014, p. 9-20.
2. F. Clerc, M. Lefevre, N. Guin, et J.-C. Marty, « Mise en place de la personnalisation dans le cadre des MOOCs », in *7ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain - ELIAH'2015*, Agadir, Morocco, juin 2015, p. 291-300. Consulté le: janv. 19, 2021. [En ligne]. Disponible sur: <https://hal.archives-ouvertes.fr/hal-01177839>.
3. K. Jordan, « Initial trends in enrolment and completion of massive open online courses », *IRRODL*, vol. 15, n° 1, janv. 2014, doi: 10.19173/irrodl.v15i1.1651.
4. M. Lefèvre, N. Guin, J.-C. Marty, et F. Clerc, « Supporting Teaching Teams in Personalizing MOOCs Course Paths », in *Adaptive and Adaptable Learning*, Cham, 2016, p. 605-609, doi: 10.1007/978-3-319-45153-4_73.
5. L. Giannandrea et M. Sansoni, « A literature review on intelligent tutoring systems and on student profiling », *Learning & Teaching with Media & Technology*, vol. 287, p. 287-294, 2013.
6. J.-M. Gilliot, N. El Mawas, et S. Garlatti, « Towards Personal infrastructure to manage long term open learner models », présenté à 24th ACM Conference on User Modeling, Adaptation and Personalisation (UMAP), Halifax, Canada, juill. 2016, Consulté le: janv. 28, 2021. [En ligne]. Disponible sur: <http://ceur-ws.org/Vol-1618/PALE11.pdf>.
7. S. Jean-Daubias, « Ingénierie des profils d'apprenants », Habilitation à Diriger des Recherches, Université Claude Bernard - Lyon I, Lyon, FR, 2011.
8. P. Brusilovsky, « Adaptive and Intelligent Technologies for Web-based Education », in *C. Rollinger and C. Peylo (Eds.), Künstliche Intelligenz 4, Special Issue on Intelligent Systems and Teleteaching*, 1999, vol. 13, p. 19-25.
9. J.-M. Gilliot, « VERS UN ENVIRONNEMENT DE GESTION PERSONNELLE ET COLLECTIVE DES DONNÉES D'APPRENTISSAGE AU SERVICE DU POUVOIR D'AGIR DES APPRENANTS », présenté à Séminaire CIREL Lille 1er avril 2019, Lille,

³ <https://moocgdp.gestiondeprojet.pm>

- France, mars 31, 2019, Consulté le: janv. 29, 2021. [En ligne]. Disponible sur: <https://fr.slideshare.net/jm.gilliot/sedela-2019-environnement-de-gestion-de-donnees-personnelles-dapprentissage-tout-au-long-de-la-vie>.
10. P. Sloep *et al.*, « A European research agenda for lifelong learning », *International Journal of Technology Enhanced Learning*, vol. 3, n° 2, p. 204-228, janv. 2011, doi: 10.1504/IJTEL.2011.039403.
 11. J.-M. Gilliot, N. El Mawas, S. Garlatti, et E. Cousin, « Développer le pouvoir d'apprendre tout au long de la vie », présenté à Actes de l'atelier Personnalisation et adaptation dans les environnements d'apprentissage : un regard interdisciplinaire. Atelier des ORPHEE-RDV, Font-Romeu, 2017.
 12. A. Abyaa, M. Khalidi Idrissi, et S. Bennani, « Learner modelling: systematic review of the literature from the last 5 years », *Education Tech Research Dev*, vol. 67, n° 5, p. 1105-1143, oct. 2019, doi: 10.1007/s11423-018-09644-1.
 13. S. Bull et J. Kay, « Open Learner Models », in *Advances in Intelligent Tutoring Systems*, R. Nkambou, J. Bourdeau, et R. Mizoguchi, Éd. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, p. 301-322.
 14. S. Tanimoto, « Dimensions of Transparency in Open Learner Models », in *12th International Conference on Artificial Intelligence in Education*, Amsterdam, the Netherlands, 2005, p. 100-106.
 15. [S. Bull, T. J. Jackson, et M. J. Lancaster, « Students' Interest in Their Misconceptions in First-Year Electrical Circuits and Mathematics Courses », *International Journal of Electrical Engineering & Education*, vol. 47, n° 3, p. 307-318, juill. 2010, doi: 10.7227/IJEEE.47.3.6.
 16. S. Ramirez Luelmo, N. El Mawas, et J. Heutte, « A literature review on Learner Models for MOOC to support Lifelong Learning », in *Conference on Computer Supported Education (CSEdu)*, Prague, Czech Republic, mai 2020, vol. 1, p. 527-539.
 17. M. Lefèvre, « Dossier « IA & Éducation » », *Bulletin de l'AJIA*, n° 108, 2020.
 18. B. Kitchenham et S. Charters, « Guidelines for performing systematic literature reviews in software engineering », Keele University & University of Durham, Technical 2.3, juill. 2007.
 19. V. Vagale et L. Niedrite, « Learner Model's Utilization in the e-Learning Environments », *Databases and Information Systems Baltic DB&IS'2012*, p. 162-174, 2012.
 20. S. Bulathwela, M. Perez-Ortiz, E. Yilmaz, et J. Shawe-Taylor, « TrueLearn: A Family of Bayesian Algorithms to Match Lifelong Learners to Open Educational Resources », *Proceedings of AAAI Conference on Artificial Intelligence 2020*, vol. 34, n° 01, p. 565-573, nov. 2019.
 21. C. R. Calle-Archila et O. M. Drews, « Student-Based Gamification Framework for Online Courses », in *Advances in Computing*, vol. 735, A. Solano et H. Ordoñez, Éd. Cham: Springer International Publishing, 2017, p. 401-414.
 22. R. Cook, J. Kay, et B. Kummerfeld, « MOOCIm: User Modelling for MOOCs », in *User Modeling, Adaptation and Personalization*, vol. 9146, F. Ricci, K. Bontcheva, O. Conlan, et S. Lawless, Éd. Cham: Springer International Publishing, 2015, p. 80-91.
 23. V. Dimitrova et P. Brna, « From Interactive Open Learner Modelling to Intelligent Mentoring: STyLE-OLM and Beyond », *Int J Artif Intell Educ*, vol. 26, n° 1, p. 332-349, mars 2016, doi: 10.1007/s40593-015-0087-3.
 24. N. El Mawas, J.-M. Gilliot, S. Garlatti, R. Euler, et S. Pascual, « As One Size Doesn't Fit All, Personalized Massive Open Online Courses Are Required », in *Computer Supported Education*, Cham, 2019, vol. 1022, p. 470-488, doi: 10.1007/978-3-030-21151-6_22.

25. M. Harrathi, N. Touzani, et R. Braham, « A Hybrid Knowledge-Based Approach for Recommending Massive Learning Activities », in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, Hammamet, oct. 2017, p. 49-54, doi: 10.1109/AICCSA.2017.150.
26. X. He, P. Liu, et W. Zhang, « Design and Implementation of a Unified Mooc Recommendation System for Social Work Major: Experiences and Lessons », in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, Guangzhou, China, juill. 2017, p. 219-223, doi: 10.1109/CSE-EUC.2017.46.
27. O. Iatrellis, A. Kameas, et P. Fitsilis, « EDUC8 ontology: semantic modeling of multi-facet learning pathways », *Educ Inf Technol*, vol. 24, n° 4, p. 2371-2390, juill. 2019, doi: 10.1007/s10639-019-09877-4.
28. A. Karahoca, I. Yengin, et D. Karahoca, « Cognitive dialog games as cognitive assistants: Tracking and adapting knowledge and interactions in student's dialogs », *International Journal of Cognitive Research in Science, Engineering and Education*, vol. 6, n° 1, p. 45, 2018.
29. Y. Li, Y. Zheng, J. Kang, et H. Bao, « Designing a Learning Recommender System by Incorporating Resource Association Analysis and Social Interaction Computing », in *State-of-the-Art and Future Directions of Smart Learning*, Y. Li, M. Chang, M. Kravcik, E. Popescu, R. Huang, Kinshuk, et N.-S. Chen, Éd. Singapore: Springer Singapore, 2016, p. 137-143.
30. H. Lynda et F. Bouarab-Dahmani, « Gradual Learners' Assessment in Massive Open Online Courses Based on ODALA Approach: », *Journal of Information Technology Research*, vol. 12, n° 3, p. 21-43, juill. 2019, doi: 10.4018/JITR.2019070102.
31. W. Maalej, P. Pernelle, C. Ben Amar, T. Carron, et E. Kredens, « Modeling Skills in a Learner-Centred Approach Within MOOCs », in *Advances in Web-Based Learning – ICWL 2016*, vol. 10013, D. K. W. Chiu, I. Marenzi, U. Nanni, M. Spaniol, et M. Temperini, Éd. Cham: Springer International Publishing, 2016, p. 102-111.
32. M. Maravanyika, N. Dlodlo, et N. Jere, « An adaptive recommender-system based framework for personalised teaching and learning on e-learning platforms », in *2017 IST-Africa Week Conference (IST-Africa)*, Windhoek, mai 2017, p. 1-9, doi: 10.23919/ISTAFRICA.2017.8102297.
33. A. Qazdar, C. Cherkaoui, B. Er-Raha, et D. Mammass, « AeLF: Mixing Adaptive Learning System with Learning Management System », *IJCA*, vol. 119, n° 15, p. 1-8, juin 2015.
34. G. Sun, T. Cui, W. Guo, G. Beydoun, D. Xu, et J. Shen, « Micro Learning Adaptation in MOOC: A Software as a Service and a Personalized Learner Model », *Advances in Web-Based Learning -- ICWL 2015*, vol. 9412, p. 174-184, 2015.
35. A. Tato, R. Nkambou, J. Brisson, et S. Robert, « Predicting Learner's Deductive Reasoning Skills Using a Bayesian Network », in *Artificial Intelligence in Education*, vol. 10331, E. André, R. Baker, X. Hu, Ma. M. T. Rodrigo, et B. du Boulay, Éd. Cham: Springer International Publishing, 2017, p. 381-392.
36. M. Tmimi, M. Benslimane, M. Berrada, et K. Ouazzani, « A Proposed Conception of the Learner Model for Adaptive Hypermedia », *International Journal of Applied Engineering Research*, vol. 12, n° 24, p. 9, 2017.

Annexe

Le tableau récapitulatif des Modèles apprenant pour MOOC pour soutenir l'ATLV se trouve à l'adresse : <https://nextcloud.univ-lille.fr/index.php/s/4Zow3sNHqxc9czn>

Conception et évaluation d'un processus de personnalisation fondé sur des référentiels de compétences

Louis Sablayrolles^{1,2}, Nathalie Guin¹, Marie Lefevre¹ et Julien Broisin²

¹ Université de Lyon, CNRS, LIRIS UMR 5205, F-6922 Villeurbanne, France
{prenom.nom}@liris.cnrs.fr

² IRIT UMR 5505, 118 route de Narbonne, Toulouse, France
{prenom.nom}@irit.fr

Résumé. La recommandation d'activités pour les apprenants a pour objectif de leur fournir des ressources adaptées à leurs besoins, dans le but de faciliter leur apprentissage. Nous nous intéressons dans cet article à la recommandation d'activités aux apprenants à partir de leur profil cognitif exprimé en termes de compétences. Notre principale contribution concerne la formalisation d'un processus de personnalisation exploitant un méta-modèle de stratégie pédagogique à base de compétences. Ce processus de personnalisation est ensuite implémenté pour fournir des recommandations de ressources dans le cadre de l'apprentissage de la programmation Shell en 1^{ère} année de DUT Informatique. Enfin nous proposons une évaluation auprès des enseignants de la pertinence des recommandations fournies, ainsi que des pistes d'amélioration de notre processus.

Mots-clés : recommandation de ressources, méta-modèle de stratégie pédagogique, règles de personnalisation

Abstract. Recommending activities for learners enables them to facilitate their learning with adapted resources. In this article, we are interested in recommending activities to learners based on their cognitive profile expressed in terms of competencies. We propose the formalization of a competency-based personalization process and its pedagogic strategy's meta-model. This personalization process is then implemented in order to provide recommendations for learning Shell programming in a first-year computer science degree. We finally discuss the evaluation by teachers of the relevance of the recommendations.

Keywords: resource recommendation, pedagogic strategy meta-model, personalization rules

1 Introduction

Les systèmes de recommandation sont une part importante de la recherche en EIAH. Ils permettent de fournir des items de travail pertinents [1], sous la forme d'ensembles ou de séquences, tout en privilégiant certaines contraintes (items nouveaux, items les plus pertinents ou items liés à un chemin d'apprentissage).

Nous nous intéressons ici au contexte des approches par compétences et en particulier aux recommandations d'items à base de compétences. L'orchestration des ressources à utiliser n'étant pas l'objectif de nos travaux, nous étudions la recommandation de séquences de ressources pertinentes pour l'apprenant. Nos travaux s'inscrivent dans la continuité du modèle PERSUA2 [2] qui permet à une équipe enseignante d'agir sur le processus de personnalisation. À partir de celui-ci, nous avons cherché à concevoir un processus de personnalisation utilisant une approche par compétences formalisée par un référentiel contenant des compétences liées par des relations sémantiques. Dans l'objectif de fournir une recommandation personnalisée aux apprenants tout en respectant la pédagogie des enseignants, nous avons associé à ce processus de personnalisation un méta-modèle de stratégie pédagogique.

Dans cet article, nous commençons par présenter le projet ComPer dans lequel s'inscrivent nos travaux, ainsi que nos motivations, avant de faire état des travaux connexes existants. Nous poursuivons par la présentation de notre processus de personnalisation et du méta modèle de stratégie pédagogique associé, en prenant l'exemple de la stratégie pédagogique définie dans le cadre du projet. Nous présentons ensuite l'expérimentation de ce processus que nous avons implémenté pour fournir des recommandations d'activités à des apprenants inscrits en 1ère année de DUT Informatique, dans le contexte de l'apprentissage de la programmation Shell. Enfin nous concluons et exposons certaines perspectives d'évolution.

2 Projet et motivations

Nos travaux s'inscrivent dans le projet ComPer [13] dont le but est de concevoir des modèles et outils destinés à mettre en œuvre une approche par compétences pour accompagner l'apprentissage de manière personnalisée. Son environnement est illustré par la Figure 1. Un référentiel de compétences est un ensemble hiérarchisé de connaissances, savoir-faire et compétences, notés KSC (Knowledge, Skill, Competency) par la suite, liés entre eux par des relations sémantiques (telles que des relations de prérequis, de composition ou de complexification). Ce référentiel est défini par des experts du domaine (didacticiens) ou des enseignants. Les enseignants peuvent rattacher des ressources (i.e. cours, exercices) à un ou plusieurs nœuds d'un référentiel. Un profil de compétences de l'apprenant peut alors être calculé à partir des traces des activités qu'il a réalisées sur les ressources (e.g. réussite/échec aux exercices). Ce profil est exprimé par un taux de maîtrise des nœuds du référentiel de compétences, auquel sont associées deux autres métriques dont la confiance du système dans ce taux de maîtrise. Le contexte d'apprentissage est constitué des objectifs définis par l'apprenant (KSC qu'il veut

maîtriser) pour guider le processus ainsi que le temps maximal de la session. Le processus de personnalisation que nous proposons prend en compte ces différents éléments, et utilise l'expertise des enseignants au travers des stratégies pédagogiques.

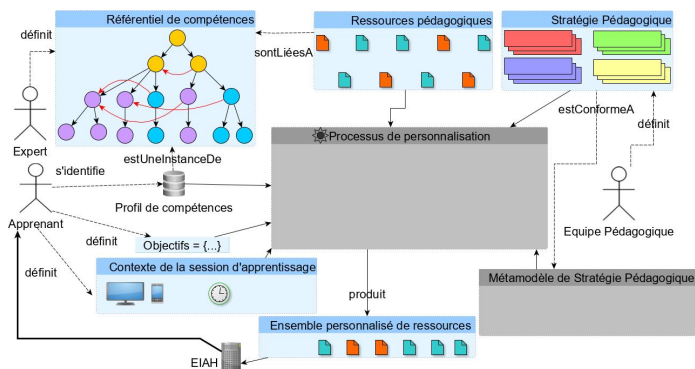


Fig. 1. Environnement du projet ComPer.

Selon les spécifications du projet, notre processus de personnalisation doit respecter les contraintes suivantes : (C1) applicabilité des composants de la stratégie par le processus de personnalisation ; (C2) explicabilité du processus de personnalisation aux apprenants et enseignants ; (C3) configuration et modification du processus par les enseignants.

Les travaux présentés dans cet article sont centrés sur la conception du processus de personnalisation et sur son évaluation dans un contexte d'apprentissage réel. Avant de détailler nos contributions, nous présentons dans la section suivante les approches existantes qui visent une personnalisation de l'apprentissage à base de compétences.

3 Personnalisation à base de compétences

Les processus de personnalisation fondés sur les compétences peuvent être divisés en trois grandes familles. Premièrement, il existe des algorithmes utilisant des approches d'apprentissage et raisonnement automatique. Deuxièmement, on trouve des approches statistiques. Pour finir, certains algorithmes sont fondés sur des modèles. Nous détaillerons dans un premier paragraphe les deux premières avant d'aborder la troisième. Pour finir, nous indiquerons des travaux parallèles utilisant l'expertise des enseignants.

Dans les approches d'apprentissage automatique, les modèles neuronaux peuvent être combinés ensemble [3] pour la génération et la validation de chemins d'apprentissage. Dans [4], les auteurs utilisent des probabilités pour prédire des relations entre les concepts, puis une matrice item/item construite par une approche bayésienne pour recommander les meilleures activités à l'apprenant. Il est également possible d'utiliser les parcours d'apprentissage des étudiants les plus performants, d'effectuer un *clustering*

des parcours d'apprentissage définis par des experts, ou d'utiliser l'*ant colony optimization* [3]. Cependant, toutes ces approches ne peuvent pas être utilisées dans notre contexte en raison de leur manque d'explicabilité (C2) et de paramétrage (C3).

Afin de se rapprocher de nos contraintes, nous avons trouvé les travaux similaires suivants fondés sur des modèles. Nous pouvons citer ceux de [8] qui utilisent l'approche Cb-KST sur un référentiel de compétences, liées entre elles par des relations de composition ou de prérequis. Un certain nombre de ressources sont recommandées parmi un ensemble de compétences choisies par l'enseignant (objectif) en fonction de paramètres de difficulté qu'il a définis. Cette recommandation peut être modulée à l'aide de plusieurs stratégies prédéfinies (Progression, Renforcement, Approfondissement). Cependant, celles-ci ne peuvent pas être configurées en fonction des besoins de l'enseignant, qui ne peut que choisir celle qui lui convient le mieux. Nous trouvons une structuration de compétences sous la forme d'un arbre de composition dans [9]. Utilisant un système multi-agent ils recommandent les ressources associées à l'élément le plus proche (défini par une mesure de distance) de celui sur lequel l'apprenant veut travailler. Les agents vont donc échanger pour se répartir les tâches de sélection d'activités. Dans [10], le domaine est modélisé selon une ontologie de différents éléments (concepts, notions, connaissances, et sujets) qui sont liés entre eux par une relation de composition pondérée. Leur modèle de l'apprenant est construit en utilisant une approche bayésienne sur les traces permettant de construire un profil. A partir de celui-ci, des tâches liées à un concept cible sont recommandées à l'apprenant. Cependant, il peut être compliqué pour les enseignants de mettre en place de tels dispositifs (C3).

Certains travaux tentent de faire appel à l'expertise des enseignants. Ainsi, dans [5], la modélisation du domaine est réalisée avec un arbre de chemins d'apprentissage possibles composé des concepts du cours liés par des relations de prérequis indiquant qu'un concept doit être travaillé avant un autre. En utilisant un concept de départ, les scores de l'apprenant pour chaque chemin d'apprentissage sont estimés et celui qui est le plus approprié et ayant le score estimé le plus élevé est sélectionné. En outre, certains travaux laissent la possibilité d'utiliser des intentions pour personnaliser la recommandation. Citons par exemple celle de « prérequis » [6], ou la formalisation d'un scénario pédagogique avec « découverte » et « renforcement » dans [7]. Cependant, dans le cadre de notre projet, cette présence de parcours d'apprentissage n'est pas utilisable car notre recommandation n'est valable qu'au moment de sa génération.

Ainsi, afin de recommander des ressources aux apprenants en fonction de leur profil de compétences et de leurs objectifs, nous identifions les verrous suivants : (i) comment reproduire l'expertise de l'enseignant au travers d'une stratégie pédagogique ; (ii) comment utiliser les relations sémantiques dans une stratégie pédagogique ; (iii) comment un processus de recommandation peut utiliser ces stratégies pédagogiques pour recommander des ressources à un apprenant. Notre question de recherche est donc (RQ) Comment proposer un processus de recommandation fondé sur un référentiel de compétences et correspondant à des pratiques pédagogiques ?

Dans notre contexte, il nous a semblé intéressant de réutiliser les concepts d'objectif de l'apprenant, d'intention et de stratégie pédagogique afin de respecter les contraintes (C1, C2, C3) nécessaires à notre démarche. Enfin, dans la continuité des travaux de [2], nous réutiliserons les définitions de "stratégie pédagogique", c'est-à-dire "un ensemble

de règles qui permettent de répartir les activités en fonction des activités disponibles et des profils des apprenants", mais aussi celui de "règles" (de la forme SI [condition sur les profils] ALORS [activités à proposer] SINON [autres activités à proposer]). Dans notre contexte, il est important que notre processus puisse s'appliquer sur différents terrains (informatique à l'université, physique/chimie au lycée, etc.). Donc pour répondre à la question de recherche, notre approche a consisté à (i) proposer un processus de personnalisation qui applique une stratégie pédagogique au profil de compétences d'un apprenant en fonction de ses objectifs ; (ii) définir une stratégie indépendante du niveau et du domaine enseigné ; (iii) définir, à partir de cette stratégie, un méta-modèle de stratégie pédagogique qui permettra la définition de stratégies expertes par les enseignants tout en assurant la compatibilité avec le processus.

4 Recommandation personnalisée d'activités

La conception de notre processus de recommandation a été réalisée en plusieurs étapes. À l'aide d'un questionnaire non standardisé, nous avons interrogé les enseignants du projet sur leurs besoins de personnalisation, que nous avons ensuite complété par des discussions avec les participants. Ceci nous a fourni les éléments pour élaborer une première proposition concernant les étapes du processus de recommandation. Suite à plusieurs modèles présentés aux membres du projet, nous avons pu faire valider les étapes du processus. Nous avons ensuite conçu une première stratégie pédagogique (nommée ComPer) qui nous a amené à améliorer les différentes étapes de notre processus de recommandation. Nous avons ainsi défini notre processus de recommandation et le métamodèle de stratégie sur lequel il s'appuie.

4.1 Processus de personnalisation

Après des interactions avec les enseignants du projet, nous avons choisi de traiter les KSC et les ressources en deux phases distinctes pour faciliter le paramétrage et l'explicabilité multi-niveaux de notre système. Comme entrée de notre processus de personnalisation, nous avons le profil de l'apprenant contenant, en plus, les relations du cadre de compétences, la stratégie pédagogique de l'enseignant et les éléments du contexte de la session tels que les objectifs de l'apprenant (comme une liste ordonnée) ou les paramètres de restriction. Notre objectif est de produire une séquence ordonnée de ressources pertinentes pour l'apprenant. Le méta-modèle du référentiel de compétences que nous utilisons propose une structure en graphe. En raison de la complexité élevée de l'exploration des graphes, nous décomposons la phase KSC en une partie de navigation/sélection et une partie d'ordonnement. Finalement, notre processus se déroule en 3 phases principales en utilisant une liste ordonnée d'objectifs pour chaque apprenant. Un objectif est une paire constituée d'un nœud du référentiel à travailler (par exemple, Manipuler des variables) et d'une intention pédagogique (par exemple, révision ou découverte). Les intentions sont utilisées pour qualifier la manière dont l'enseignant et/ou l'apprenant souhaite travailler un nœud (voir section 4.2).

La phase de Sélection permet de choisir les nœuds du référentiel que l'apprenant devra travailler en fonction de son profil de compétences et des objectifs qui ont été définis. Lors de cette phase, des tags sont apposés sur chaque nœud sélectionné afin de savoir pourquoi il doit être travaillé. La phase d'Ordonnancement permet de trier les nœuds sélectionnés afin de faire travailler l'apprenant prioritairement sur les nœuds les plus importants. Pour terminer, la phase de Ressources consiste, pour chacun des nœuds contenus dans la liste issue de la phase précédente, à récupérer de façon ordonnée les ressources (activité, cours, etc.) à proposer à l'apprenant. Ces trois phases sont appliquées indépendamment pour chaque objectif formulé par l'apprenant.

La phase de **Sélection** des KSC est elle-même constituée de trois étapes :

S.Initialisation : à partir du nœud correspondant à l'objectif traité, la phase d'initialisation permet de récupérer l'ensemble des sous-concepts (KSC) qui le composent.

S.Pre-Traitement : cette étape est facultative et permet, en fonction de l'intention de l'objectif et en suivant les relations du référentiel de compétences, d'explorer des sous-arbres de KSC liés à ceux récupérés lors de l'étape précédente. Ces sous-arbres peuvent être des pré-requis, des leviers facilitant la compréhension, ou des approfondissements des sous-concepts (KSC) issus de l'étape précédente.

S. Traitement: cette phase permet d'appliquer un traitement sur les nœuds issus de l'étape S.Pre-Traitement (ou S.Initialisation si l'étape de pré-traitement n'a pas eu lieu). Ce traitement comprend deux possibilités, indépendantes l'une de l'autre, qui sont appliquées en fonction d'une condition sur les valeurs de maîtrise, confiance et couverture du nœud : (i) sélectionner le nœud avec un tag explicitant l'intention avec laquelle ce nœud doit être travaillé ; (ii) ajouter le nœud à la liste des objectifs avec une nouvelle intention, dans le but de le traiter différemment. Nous avons laissé la possibilité d'effectuer l'un et/ou l'autre traitement sur les nœuds récupérés. Par exemple, lorsque l'objectif vise à travailler un nœud avec l'intention prérequis, notre approche consiste à parcourir les relations de prérequis du nœud objectif, puis à sélectionner les nœuds dont la maîtrise est inconnue ou insuffisante avec le tag `Pre_requis`. De plus, si ce nœud a une maîtrise inconnue, il est ajouté à la liste des objectifs avec l'intention découverte alors que si sa maîtrise est insuffisante, il sera ajouté avec l'intention renforcement.

Lors de la phase de Sélection, une formule de priorité est appliquée à chacun des nœuds rencontrés. La priorité d'un nœud est calculée en fonction de la priorité du père de ce nœud, de la/les relations qui a/ont été suivie(s), de la profondeur du nœud par rapport au nœud objectif, ainsi que de la règle de sélection appliquée. Cette priorité est utilisée durant la phase d'Ordonnancement, pour trier les nœuds sélectionnés.

La phase d'**Ordonnancement** comprend trois étapes pour ordonner les KSC sélectionnés lors de la phase de Sélection. (1) **O.Unicité** regroupe les sélections multiples d'un nœud au sein d'un objectif, en combinant leur priorité et leur tag de sélection. En effet, un même nœud peut être sélectionné plusieurs fois dans le cas où, par exemple, celui-ci est un sous-concept (KSC) du nœud objectif, mais aussi un prérequis d'un autre sous-concept. (2) **O.TriParPriorité** ordonne les KSC par priorité décroissante. Enfin, en cas de conflit entre deux KSC ayant la même priorité, (3) **O.TriParTag** ordonne les KSC selon leur tag de sélection et l'intention de l'objectif initial en exploitant une table de priorité indiquant les tags à exploiter prioritairement.

Enfin, à l'aide de l'ensemble ordonné des KSC à travailler issus de la phase précédente, la phase de Ressources permet, en deux temps, de retrouver les ressources associées à ces nœuds. (1) **R.Selection** permet de récupérer l'ensemble des ressources les plus adaptées liées aux nœuds sélectionnés en suivant les relations entre un Knowledge/Skill et une ressource (i.e. `hasLearning` et `hasTraining`). (2) **R.Ordonnement** trie les ressources sélectionnées pour fournir par exemple les ressources de type cours avant les ressources de type exercices.

4.2 Stratégie ComPer - une stratégie indépendante du niveau et du domaine enseigné

Comme décrit dans la section précédente, nos stratégies de personnalisation utilisent des intentions. Elles permettent de qualifier la manière dont l'enseignant et/ou l'apprenant désire travailler un nœud N du référentiel de compétences. Nous donnons ci-dessous celles que nous avons définies au sein du projet ComPer.

L'intention **Prérequis** consiste à travailler des thèmes/sous-thèmes prérequis à N. L'intention **Découverte** permet de travailler N pour la première fois. L'intention **Remédiation** fait travailler N, qui a déjà été travaillé mais qui est non maîtrisé. L'intention **Renforcement** a pour but de renforcer la maîtrise de N, qui a déjà été travaillé mais pour lequel la maîtrise est partielle. L'intention **Révision** fait travailler N dans le cadre d'une révision à un examen par exemple. L'intention **Approfondissement** consiste à travailler des nœuds qui sont des complexifications de N. Enfin, l'intention **Inconnue** est utilisée lorsque l'apprenant ignore l'intention avec laquelle il désire travailler N.

Comme nous allons le présenter dans notre méta-modèle, il est possible de définir des stratégies pédagogiques ayant des comportements différents pour chaque intention. Dans le cas de la stratégie ComPer, travailler un nœud objectif avec l'intention de **Prérequis** consiste à faire travailler les nœuds prérequis de l'objectif ou de ses descendants. Pour cela, on récupère l'ensemble des sous-nœuds (KSC) liés au nœud objectif. A partir de ces descendants, nous isolons les nœuds prérequis en suivant la relation *requires* dans le référentiel de compétences. L'ensemble des prérequis non maîtrisés sera ensuite sélectionné afin que l'apprenant puisse les travailler. Dans le cadre du travail d'un nœud objectif avec l'intention de **Découverte**, nous recherchons l'ensemble des descendants de l'objectif dont la maîtrise est inconnue, afin de proposer à l'apprenant des ressources qui y sont rattachées. Travailler un nœud objectif avec l'intention de **Renforcement** ou de **Remédiation** consiste à faire travailler les sous-nœuds qui ne sont pas assez maîtrisés. Cela peut correspondre à la réalisation d'exercices d'entraînement ou à revoir un cours dont une portion est mal comprise. Travailler un nœud avec l'intention de **Révision** a pour but de faire réviser le nœud à l'apprenant en lui proposant des activités déjà réalisées qu'il maîtrise déjà, mais en privilégiant le travail des nœuds dont la maîtrise est insuffisante. Le travail d'un nœud objectif avec l'intention d'**Approfondissement** consiste à vérifier tout d'abord que l'apprenant a une maîtrise suffisante de l'objectif et des sous-nœuds le composant, puis à aller chercher des nœuds qui permettent de travailler des concepts plus complexes. Enfin, si l'intention est **Inconnue**, le processus de personnalisation déterminera, à l'aide des valeurs de maîtrise du profil de compétences, l'intention qui conviendra le mieux à l'apprenant.

4.3 Méta-modèle de stratégie pédagogique

Nous avons proposé un méta-modèle permettant d'exprimer la stratégie ComPer mais également d'autres stratégies du même type. Il permet au processus de valider la stratégie en entrée. La grammaire de notre méta-modèle a été définie en utilisant un schéma JSON [14]. Cette grammaire comprend un ensemble de 5 classes d'éléments détaillés ci-dessous : la stratégie pédagogique, les procédures de sélection, les règles de sélection, les prédicats, et les paramètres. Chacun de ces éléments possède un identifiant, un nom, et un type (ComPer ou Expert). Cette dernière propriété indique si l'élément a été conçu par les membres de l'équipe du projet ComPer, ou par des experts/enseignants d'un domaine.

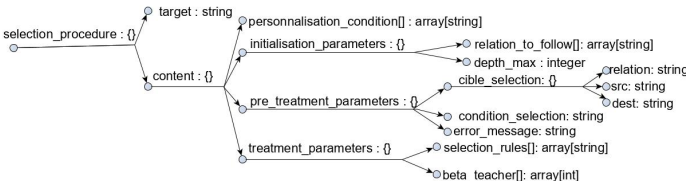


Fig. 2. Élément **selection_procedure** du méta-modèle de stratégie pédagogique.

L'élément *instructional_strategy* contient les informations nécessaires pour appliquer le processus de personnalisation : (i) Les paramètres pour la phase de Sélection, à savoir l'ensemble des éléments "procédures de sélection" (que nous détaillons dans la suite de cette section) à appliquer ainsi que la majeure partie des paramètres liés au calcul de la priorité d'un nœud. La priorité d'un nœud est calculée en utilisant les données récoltées lors du parcours du référentiel ainsi que les procédures de sélection qui lui sont appliquées. (ii) La table de priorité des tags de sélection en fonction de l'intention de l'objectif initial pour la phase d'Ordonnancement. Celle-ci est propre à chaque stratégie pédagogique et peut donc être adaptée aux besoins des équipes pédagogiques.

L'élément *selection_procedure* (cf. fig2) regroupe les paramètres de chacune des étapes décrites en Section 4.2. On trouve ainsi (dans *personnalisation_condition*) l'ensemble des intentions pour lesquelles cette procédure peut s'appliquer. L'élément *initialisation_parameters* permet de définir certains paramètres nécessaires à la phase d'initialisation telles que les relations sémantiques à suivre pour récupérer les sous-concepts (KSC) liés au nœud objectif, ainsi que la profondeur maximale à laquelle s'arrêter. Pour l'étape de S.Pre-Traitement, l'élément *pre_treatment_parameters* permet de définir les relations sémantiques qui doivent être suivies pour explorer les arbres de sous-concepts (KSC), ainsi que la condition nécessaire à leur exploration. Par exemple, dans le cas de l'intention d'approfondissement, les sous-arbres du nœud objectif ne seront explorés que si celui-ci est maîtrisé. Enfin, les paramètres de traitement proposent des paramètres spécifiques au calcul de la priorité d'un nœud, et aux règles de sélection applicables. Le premier arbre (fig3) illustre les valeurs de ces différents paramètres dans la stratégie ComPer pour la procédure de sélection traitant les objectifs

ayant l'intention prérequis. On retrouve ainsi la condition d'application de la procédure dans *personnalisation_condition*, les relations sémantiques de l'étape S.Initialisation à suivre (jusqu'à la profondeur maximale) pour obtenir les descendants de l'objectif, la relation sémantique *requires* pour l'étape S.Pre-Traitement, et l'ensemble des règles de sélection à appliquer pour l'étape S.Traitement dont la première est représentée sur le deuxième arbre (voir fig 3).

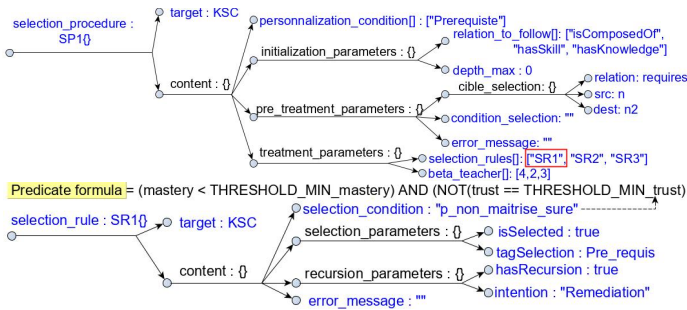


Fig. 3. Éléments *selection_procedure* (SP1) et *selection_rule* (RS1) de la stratégie ComPer.

L'élément *selection_rule* définit la manière de traiter un nœud lors de l'étape S.Traitement. La condition de sélection permet de savoir si la règle peut être appliquée pour le nœud traité (un exemple de sélection est donné par la formule en jaune sur le deuxième arbre de la figure 3). De plus, dans le cas de l'application de la règle à un nœud, on retrouve les paramètres qui concernent la sélection du nœud (élément *selection_parameters*, voir figure 3) ainsi que les paramètres liés à la récursivité du processus de sélection (élément *recursion_parameters* sur le deuxième arbre de la figure 3). Enfin, l'élément *predicate* définit une formule logique permettant de tester un ensemble de contraintes sur un nœud comme dans la condition des règles de sélection.

5 Expérimentation dans le cadre de l'apprentissage de la programmation Shell

La stratégie ComPer présentée plus haut a été conçue pour être appliquée indépendamment du niveau d'apprentissage et de la discipline enseignée. L'expérimentation présentée ici vise donc à évaluer cette stratégie dans un premier contexte d'apprentissage réel, celui de l'apprentissage de l'informatique dans l'enseignement supérieur.

5.1 Dispositif pédagogique

L'expérimentation que nous avons menée pour évaluer notre proposition utilise la plateforme Lab4CE dédiée à l'enseignement de l'informatique [11]. Pour notre expérimentation, cette plateforme a été utilisée dans le cadre d'un module d'apprentissage de la programmation Shell. Elle s'est déroulée pendant 6 semaines avec 180 étudiants inscrits en première année de DUT d'informatique. Le référentiel de compétences et les ressources associées ont été conçus par l'enseignant en charge du module. Il comprend un total de 96 nœuds (11 compétences, 24 connaissances, 19 savoir-faire et 42 ressources) liés entre eux par un total de 193 relations. La première semaine a été utilisée pour évaluer le profil des apprenants afin de déterminer leur maîtrise de certains KSC du référentiel de compétences. Chaque semaine, le lundi, les étudiants avaient une session d'apprentissage à distance de 90 minutes avec l'un des 6 enseignants du module, au cours de laquelle une série d'exercices devait être réalisée sur la plateforme ASKER [12]. La plateforme Lab4CE était accessible en ligne à tout moment, et les étudiants pouvaient poursuivre leur travail tout au long de la semaine. Chaque vendredi soir, les logs de la plateforme étaient extraits et deux enseignants les passaient en revue afin de mettre à jour les profils des apprenants en fonction de leurs productions. Chaque dimanche soir, notre processus de recommandation implémentant la stratégie ComPer était exécuté afin de générer, sur la base d'objectifs à maîtriser (en termes de KSC) définis par le superviseur du module, une recommandation pour chaque apprenant consistant en une liste ordonnée d'un nombre variable de ressources. Notons que les ressources ont été ordonnées en fonction de leur type (i.e. contenu très court des cours magistraux d'abord, et exercices). A la demande du responsable du module, nous avons également désactivé l'intention de découverte. Enfin, les apprenants étaient libres, à partir du lundi de la 2^{ème} semaine, de visiter ou non les ressources recommandées.

5.2 Protocole d'évaluation du processus de recommandation

Afin d'évaluer la pertinence de notre processus de recommandation, nous avons suivi le protocole suivant. Tout au long de l'expérimentation, nous avons sauvegardé l'ensemble des traces issues de la génération des recommandations ainsi que les profils des étudiants ayant donné lieu à ces recommandations.

Ensuite, en utilisant un regroupement manuel effectué sur les recommandations des semaines 2 et 3, un total de 4 profils par semaine a été sélectionné. Nous avons choisi les semaines 2 et 3 de l'expérience car elles sont situées au milieu de l'expérience, ce qui nous permet d'évaluer la précision après le début de l'expérience. Au moyen d'un questionnaire non standardisé [15], nous avons d'abord demandé aux six enseignants du module de fournir leurs recommandations pour deux de ces profils afin de les comparer aux recommandations de l'algorithme. Ensuite, nous leur avons demandé d'évaluer la recommandation fournie par notre système pour les deux autres profils en utilisant une échelle de Likert à 5 valeurs. Nous avons calculé la moyenne des évaluations des enseignants concernant la qualité des recommandations générées par notre système.

Nous avons ensuite effectué une analyse qualitative en représentant les recommandations de notre processus et celles des experts du référentiel de compétences afin de savoir si les recommandations étaient proches.

5.3 Résultats et discussion

En ce qui concerne les résultats de l'évaluation de la pertinence des recommandations par les enseignants, le score moyen pour les deux semaines était de 3,24/5 avec 3,83 pour la semaine 2 et 2,66 pour la semaine 3. Une analyse qualitative utilisant une représentation des recommandations du processus et les recommandations des experts a permis de montrer visuellement que les recommandations des experts étaient pertinentes. Nous avons pu remarquer que l'ensemble des nœuds recommandés par notre algorithme et les experts sont situés dans la même portion du référentiel. Cependant, l'ordre dans lequel les nœuds ont été sélectionnés était très différent entre notre processus et les experts, ainsi qu'entre les experts. Un processus itératif permettra de comprendre leur méthode d'ordonnement.

Ainsi nous avons pu constater que nos recommandations sont similaires à celles des enseignants. Nous avons de plus montré que chaque enseignant avait sa propre manière d'ordonner les nœuds à travailler ce qui peut expliquer les résultats obtenus lors de l'évaluation des recommandations. Une remarque laissée par un des enseignants permet d'illustrer ce problème : "L'ordre n'est pas pertinent !!! Il faut faire travailler les K et les S avant de faire travailler les C !!!". De plus, des enseignants nous ont fait remarquer que certains nœuds n'avaient pas été recommandés alors qu'ils auraient dû l'être. Après étude des traces, nous avons constaté que cela était dû au fait que la règle traitant ces cas-là (intention de découverte) avait été désactivée à la demande du responsable.

6 Conclusion et Perspectives

Nous avons dans cet article montré qu'il était possible de (i) concevoir un processus de personnalisation exploitant les objectifs de l'apprenant et appliquant une stratégie pédagogique (définie à partir d'un méta-modèle de stratégie) sur un profil de compétences de l'apprenant ; (ii) définir une stratégie pouvant s'appliquer à la recommandation d'activités dans le cadre de la programmation Shell ; (iii) pouvoir modifier/créer une stratégie pédagogique en fonction des besoins d'un enseignant ; (iv) fournir aux apprenants des recommandations qui ne sont pas aberrantes.

Pour la suite de nos travaux, nous exploiterons les réponses au questionnaire fourni aux enseignants et les traces d'activités des apprenants pour améliorer notre formule de priorité et ses paramètres. Puis, nous voulons continuer à évaluer notre stratégie ComPer et le processus de personnalisation au sein des différents terrains d'expérimentation du projet ComPer pour lesquels des référentiels plus complexes ont été définis. Nous espérons ainsi proposer des mécanismes d'amélioration de la stratégie en suggérant certaines modifications sur ses paramètres. Pour finir, nous souhaitons travailler sur la génération d'explications offrant la possibilité à l'apprenant et à l'enseignant de comprendre les recommandations formulées par le système.

7 Remerciements

Le travail présenté dans cet article a été financé par l'ANR au sein du projet ComPer ANR-18-CE38-0012.

References

1. Ricci, F., Lior, R., Bracha, S.: Introduction to Recommender Systems Hand-book. In Recommender Systems Handbook, pp 1-35, 1st ed, Behaviour & Information Technology, Springer, Boston MA (2018)
2. Lefevre, M., Jean-Daubias, S., Guin, N. An approach for unified personalization of learning. Workshop on Personalization Approaches in Learning Environments, User Modeling, Adaptation and Personalization 2012, Montreal, Canada, pp. 5-10
3. Diwan, C., Srinivasa, S., Ram, P. Automatic Generation of Coherent Learning Pathways for Open Educational Resources In European Conference on Technology Enhanced Learning - Technology-Enhanced Learning 2019, pp. 321-334, Springer, Netherlands (2019)
4. Desmarais, M.C., Meshkinfam, P., & Gagnon, M.. Learned student models with item to item knowledge structures. User Modeling and User-Adapted Interaction, 16(5), pp 403-434 (2006)
5. Nabizadeh A.H, Gonçalves D., Gama S., Jorge J., Rafsanjani H.N., Adaptive learning path recommender approach using auxiliary learning objects, Computers & Education, 147, 2020
6. Auzende, O., Giroire, H., Le Calvez, F.: *Séquences d'entraînement guidées par les erreurs*. In Environnements Informatiques pour l'Apprentissage Humain'2013, Toulouse, pp 105-116 (2013).
7. Pernin, J.F, Emin V., Guéraud V.: Intégration de la dimension utilisateur dans la conception de systèmes pour l'apprentissage. Scénarisation pédagogique dirigée par les intentions. Ingénierie des Systèmes d'Information. 14(3): 9-30 (2009)
8. El-Kechaï, N., Melero, J., and Labat, J.-M. *Adaptation de serious games selon la stratégie choisie par l'enseignant : approche fondée sur la Competence-based Knowledge Space Theory*. In Environnements Informatiques pour l'Apprentissage Humain'2015, pp 126-137, Agadir, Maroc (2015).
9. Paquette, G., Mariño, O., Rogozan, D., & Léonard, M. *Competency-based personalization for massive online learning*. In Smart Learning Environments, 2(1), 4 (2015).
10. Mediani C., Abel MH., Djoudi M. Towards a Recommendation System for the Learner from a Semantic Model of Knowledge in a Collaborative Environment. In: Computer Science and Its Applications. Conference on Computer Science and its Applications 2015. IFIP Advances in Information and Communication Technology, 456. Springer, Cham, 2015
11. Broisin, J., Venant, R., & Vidal, P., *Lab4CE: a Remote Laboratory for Computer Education*. In International Journal of Artificial Intelligence in Education , 27(1), pp 154–180 (2015).
12. Lefevre, M., Guin, N., Cablé, B., & Buffa, B. ASKER: un outil auteur pour la création d'exercices d'auto-évaluation. In Atelier Evaluation des Apprentissages et Environnements Informatiques - Conférence Environnements Informatiques pour l'Apprentissage Humain 2015, Agadir, Maroc (2015).
13. ComPer Project Homepage, <http://comper.fr/>, last accessed 2021/04/21.
14. Métamodèle, <https://cloud.irit.fr/index.php/s/1WBfslZgcjWscpi>, last accessed 2021/04/21.
15. Formulaire, <https://cloud.irit.fr/index.php/s/qba53PR4q0F6dwX>, last accessed 2021/04/21.

Conception et expérimentation de tableaux de bord d'apprentissage pour les enseignants et les apprenants

Yassine Safsouf^{1,3}, Khalifa Mansouri², Franck Poirier¹

¹ Laboratoire STICC, Université Bretagne Sud, Vannes, France
yassine.safsouf@isga.ma

franck.poirier@univ-ubs.fr

² Laboratoire SSDIA, Université Hassan 2, Casablanca, Maroc
khmansouri@hotmail.com

³ Laboratoire LIMIE, Groupe ISGA, Marrakech, Maroc

Résumé. Depuis le début de la pandémie de COVID-19, de nombreux pays ont adopté l'enseignement en ligne comme alternative aux cours en présentiel. Cette situation a accentué la prise de conscience de l'importance d'analyser les données d'apprentissage laissées par les élèves, pour améliorer et évaluer le processus d'apprentissage. Cet article expose les résultats d'une étude conduite sur une classe de 26 élèves de deuxième année du cycle ingénieur d'un établissement supérieur au Maroc, suivant tous un cours en modalité mixte sur la plateforme Moodle. Cette étude porte sur un outil destiné aux enseignants ainsi qu'aux élèves afin de faciliter le suivi et le contrôle du processus d'apprentissage, les résultats montrent que l'outil a permis d'améliorer l'engagement et le taux de réussite des élèves.

Mots-clés : Visualisation de l'information, analyse de l'apprentissage, tableau de bord, apprentissage autorégulé, environnements d'apprentissage en ligne.

Abstract. Since the beginning of coronavirus pandemic (COVID-19), many countries have adopted online education as an alternative to traditional classroom courses. As a result, a new awareness has emerged of the importance of analyzing the learning data collected from students to improve and evaluate the learning process. This article presents the results of a study conducted on a group of 26 second-year engineering students from a higher education institution in Morocco, all of them taking a blended-mode course on the Moodle platform. This study focuses on a tool for teachers and students to help track and monitor the learning process, the results show that this tool improved student engagement and success rates.

Keywords: Information visualization, learning analytics, learning analytical dashboard, self-regulated learning, online learning environments.

1 Introduction

Le domaine de l'éducation est stimulé dans ces dernières années par de nouvelles tendances visant à améliorer le processus d'apprentissage, L'analyse de l'apprentissage (Learning Analytics ou LA) qui permet de comprendre le parcours d'un apprenant s'est considérablement développée ces dernières années. Consacrée à l'analyse des données de formation, elle vise à exploiter le potentiel des quantités de plus en plus importantes de données décrivant les informations personnelles, les données d'interaction et les données académiques générées par l'utilisation des environnements d'apprentissage en ligne (EAL) [1]. Il s'agit d'un domaine relativement nouveau, défini en 2011, par Georges Siemens théoricien du connectivisme, dans la première conférence dédiée à ce sujet LAK1 (Learning Analytics and Knowledge), comme la détection, la collecte, l'analyse et l'exploitation des traces numériques laissées par les apprenants dans leurs activités quotidiennes. Afin de comprendre et d'optimiser l'apprentissage et les différents environnements dans lesquels il se produit [2].

Les outils du LA fournissent aux enseignants des informations pertinentes sur le processus d'apprentissage des élèves [3], améliorant ainsi l'environnement dans lequel il prend place. Ces outils sont également très bénéfiques pour les apprenants, car ils leur offrent la possibilité de suivre leur progression et performance en temps réel [4] [5]. Cependant, la diversité des EAL existants complique la tâche d'analyse des données d'apprentissage, cette situation est encore amplifiée par la nécessité de combiner des données provenant de diverses sources. Notre étude s'inscrit dans ce contexte, elle répond aux questions suivantes : sur quelle base peut-on concevoir un outil d'analyse capable de favoriser l'engagement et de prédire la réussite des apprenants en ligne ? comment rendre l'outil d'analyse interopérable ? En effet, nous proposons dans ce document un outil de communication visuelle interopérable, conçu sous forme de tableaux de bord pour les enseignants et les apprenants. Notre objectif est de permettre une analyse des données d'apprentissage de manière à faciliter le suivi et le contrôle du processus d'apprentissage, dans le but d'améliorer l'engagement et le taux de réussite des apprenants en ligne.

Après cette introduction, nous commencerons dans la section 2, avec une définition du Learning Analytics et une présentation des communautés de recherche dans ce domaine, nous proposerons les méthodes et approches principales avec quelques travaux de recherches. Nous présenterons ensuite notre outil dans la section 3. La section 4 sera dédiée à la méthodologie et au contexte de l'étude. L'analyse des données de l'étude avec une discussion sur les résultats obtenus sera détaillée dans la section 5, suivie d'une conclusion.

2 État de l'art

L'une des méthodes du LA les plus courantes est l'utilisation des tableaux de bord comme outil de soutien pour les parties prenantes, principalement les enseignants et les apprenants, dans le but de leur permettre d'avoir un aperçu du processus d'apprentissage. Schwendimann a défini ces outils comme un ensemble d'affichages uniques qui

regroupent beaucoup d'indicateurs sur le processus et/ou le contexte d'apprentissage en une ou plusieurs visualisations [4]. Ioana Jivet et al., ont proposé une étude bibliographique visant à mieux comprendre et décrire les fondements théoriques derrière l'utilisation des tableaux de bord dans des concepts éducatifs [6]. L'étude a révélé que le fondement le plus courant de la conception de tableaux de bord analytiques est la théorie de l'apprentissage autorégulé (SRL pour Self-Regulated Learning theory). Cette théorie est principalement utilisée dans la prise de conscience et au déclenchement de la réflexion, fournissant ainsi un certain soutien pour les phases de performance et d'autoréflexion du cycle SRL, telle que définie par Zimmerman et Moylan [7]. Dans une recherche menée par Nicholas Diana et ses collègues [8] pour voir comment les tableaux de bord seraient en mesure de prédire les résultats des étudiants à différents moments d'un cours. L'étude montre que les résultats des apprenants peuvent être prédits avec un algorithme d'apprentissage automatique supervisé. Ces prédictions ont été intégrées dans un tableau de bord de l'instructeur qui facilite la prise de décision pour les apprenants classés comme nécessitant une assistance.

Dès la première conférence LAK1, George Siemens a porté un projet ayant pour but de concevoir une plateforme ouverte intégrant une variété de modules d'analyse dédiés à l'apprentissage [9]. Cette plateforme avait trois défis à relever, le premier est que la généralisation et l'innovation des tableaux de bords exige que les traitements, les algorithmes et les technologies utilisés soient ouverts. Le deuxième est que la plateforme soit modulaire et quelle permette d'intégrer tous types d'outils d'analyse de l'apprentissage. Enfin, elle devait aussi offrir des fonctionnalités adaptées à chaque utilisateur (enseignants, apprenants, personnels d'encadrement, assistants d'éducation ou chercheurs et producteurs de contenus), en 2016 ce projet a été abandonné. En France, un autre projet nommé « ANR HUBBLE » (HUMAN OBSERVATORY BASED ON ANALYSIS OF E-LEARNING TRACES) a vu le jour en 2014 [10], ce projet avait pour but de créer un observatoire national et un référentiel de processus d'analyse, pour accompagner les différents acteurs (enseignants, apprenants, concepteurs, administrateurs ou politiques) afin d'analyser et d'expliquer les phénomènes d'enseignement et d'apprentissage. En 2019, une thèse faite par d'Inès Dabbebi dans le cadre de ce projet, ayant comme objectif de concevoir un processus de génération des tableaux de bord d'apprentissage adaptatifs, permettant de prendre en compte les besoins des utilisateurs : enseignants, responsables pédagogiques, etc. [11]. Dans cette thèse un prototype d'un outil intitulé « HubbleBoard Generator » a été développé, cet outil permet de produire automatiquement un tableau de bord d'apprentissage en respectant le contexte d'utilisation en prenant en compte les besoins des utilisateurs et leurs niveaux de décision.

Plusieurs solutions commerciales proposent également des tableaux de bord qui donnent aux enseignants une vue d'ensemble du processus éducatif de l'apprenant. Le plugin Inspire [12] de Moodle met en œuvre une analyse de l'apprentissage et fournit des prédictions sur la réussite de l'apprenant, des diagnostics et des conseils aux apprenants et aux enseignants. Malheureusement, ce plugin n'est fonctionnel que pour Moodle 3.4. Analytics [13] est un autre plugin créé par l'équipe de développement de Moodle. Ce plugin vise à prédire et à détecter des aspects inconnus du processus d'apprentissage, sur la base de données historiques et du comportement actuel. Ce plugin prend en

charge deux types de modèles : les modèles basés sur les algorithmes de Machine Learning et les modèles "statiques" permettant de détecter des situations préoccupantes. Mais comme pour le plugin Inspire, le plugin Analytics n'est fonctionnel qu'avec la version 3.4 ou supérieure.

3 Présentation de notre outil TaBAT

Dans l'enseignement supérieur, l'utilisation du LA est devenue une nécessité. Comme l'apprentissage se fait de plus en plus en ligne, les établissements d'enseignement supérieur sont confrontés à de nombreux défis, notamment en termes de taux d'abandon, de réussite et de réduction des coûts. En utilisant des outils d'analyse et de représentation des données pour générer des rapports utiles, certaines organisations transforment ces défis en atouts. Cependant, comme présenté dans la définition, la diversité des EAL, des contenus pédagogiques et des plateformes de communication collaborative, complique la tâche d'analyse de ces traces d'apprentissage. Ceci est encore amplifié par la nécessité de combiner des données provenant de sources diverses.

Dans cette section, nous proposons un nouvel outil, nommé TaBAT (pour Tableau de Bord d'Analyse des Traces d'apprentissage), créé pour fonctionner avec différents EAL, sous forme de tableaux de bord accessible en ligne (<https://safsouf.net/tabat>). TaBAT permet aux enseignants et aux apprenants de visualiser leur processus d'apprentissage. En effet, avec TaBAT, nous pouvons consulter ce qui s'est passé pendant le cours en ligne (aspect descriptif), voir qui sont les étudiants qui réussiront ou non le cours (aspect prédictif), savoir pourquoi des étudiants ont été déclarés comme décrocheurs (aspect diagnostique) et enfin obtenir des informations sur les actions à entreprendre pour améliorer la progression et le succès des étudiants dans le cours en ligne (aspect proactif).

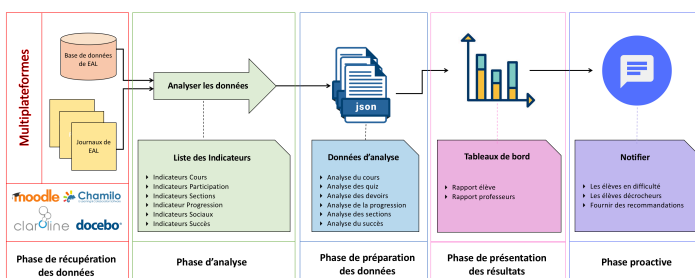


Fig. 1. Les phases du processus de fonctionnement de l'outil TaBAT.

Comme le montre la figure 1, le processus de fonctionnement de notre outil se compose de cinq phases, que nous décrivons dans la suite.

3.1 Phase de récupération des données

Le but de cette phase consiste à déterminer et récupérer les données (traces d'apprentissage) qui seront utilisées dans les phases suivantes. Les données sont localisées soit dans une base de données (MySQL, MariaDB ou PostgreSQL) dans des tables dites logstore (cas de la plateforme Moodle), dans des fichiers journaux ou bien les deux. Pour accéder aux traces d'apprentissage, l'utilisateur peut choisir (dans un fichier de configuration proposé) la plateforme utilisée, créer la chaîne de connexion avec base de données, saisir le préfixe des tables ainsi que le login et le mot de passe pour accéder à la base de données.

3.2 Phase d'analyse

Dans nos précédents travaux de recherche sur les facteurs qui peuvent influencer la réussite des apprenants en ligne [14, 15], nous avons proposé et validé statistiquement un modèle d'évaluation de la réussite des apprenants en ligne nommé e-LSAM (e-Learner Success Assessment Model). Cette deuxième phase a pour objectif de créer des algorithmes d'analyse des activités des élèves (basés sur les données préalablement collectées lors de la phase précédente) sur la base d'indicateurs découlant de notre modèle théorique. Nous classerons ces indicateurs en 6 catégories, à savoir :

1. **La catégorie cours** : les indicateurs choisis sont : le nombre d'étudiants inscrits, le nombre de sections prévues et le nombre d'activités/ressources créées.
2. **La catégorie participation** : cette catégorie est plus centrée sur les actions que l'on peut effectuer sur le cours qui font que l'élève est considéré comme actif sur ce dernier. Nous distinguons deux types d'actions : les actions de consultation et les actions de contribution. Quatre indicateurs sont choisis : la date de la dernière action effectuée, la durée totale des actions effectuées dans la dernière session, la durée totale des actions effectuées depuis le début du cours et le nombre d'actions effectuées par un élève pour chaque type d'action. Le niveau de participation est également calculé pour chaque étudiant.
3. **La catégorie section de cours** : deux indicateurs choisis dans cette catégorie, les activités/ressources (leçons, quiz, devoirs, url, chat, fichiers, etc.) consultées par l'élève, ainsi que le nombre total des activités/ressources contenues dans chaque section de cours. Ces deux indicateurs sont utilisés pour calculer le niveau de progression de l'étudiant dans chaque section de cours.
4. **La catégorie progression** : la progression d'un étudiant représente l'avancement de ce dernier au sein d'un cours. Les trois indicateurs choisis pour le calcul de la progression sont : le nombre d'activités que l'élève a déjà réalisées, le nombre d'activités non réalisées par rapport à une date limite et le nombre d'activités définies par l'enseignant au début. Ces trois indicateurs représentent directement la progression de l'élève et donc sa progression personnelle. Le niveau de progression est également calculé sur la base de ces indicateurs.

- 5. La catégorie sociale** : cette catégorie se concentre sur les interactions sociales qui peuvent avoir lieu pendant le cours qui font que l'élève est considéré comme socialement actif dans ce dernier. Deux indicateurs sont choisis, le nombre de messages postés par les utilisateurs dans les deux activités Chat et Wiki ou directement sur la plateforme, le nombre total de messages des mêmes activités et de la plateforme. Le niveau social est également calculé sur la base de ces indicateurs.
- 6. La catégorie succès** : le but est de donner une estimation du niveau de succès d'un élève dans le cours en ligne. Les deux indicateurs choisis, le premier est le niveau de progression avec succès, c'est-à-dire avec réussite (note supérieure à la moyenne) des activités (quiz, devoir et leçons), le second indicateur est plus global, il donne une idée générale du niveau de réussite d'un apprenant dans un cours en ligne. Les résultats de notre modèle e-LSAM, montre que la réussite s'explique (avec un taux de prédiction de 80,7%) par 24,1% d'autorégulation (représentée dans notre cas par le niveau de progression avec succès) et par 75,7% de continuité de l'utilisation du système. Cette dernière s'explique par 38,5 % du niveau d'interaction sociale et 61,5 % du niveau de participation au cours. Pour le calcul de l'indicateur, nous avons arrondi les valeurs pour le niveau d'interaction sociale à 40% et le niveau de participation au cours à 60%. L'équation permettant de déterminer la continuité de l'utilisation du système est la suivante:

$$\text{Continuité} = (0,4 \times \text{niveau social}) + (0,6 \times \text{niveau de participation})$$

Comme pour le calcul de la continuité d'utilisation du système, nous avons arrondi le niveau d'autorégulation à 30% et le niveau de continuité d'utilisation à 70%. Le niveau de réussite est représenté par l'équation suivante :

$$\text{Succès} = (0,3 \times \text{progression avec succès}) + (0,7 \times \text{continuité})$$

Les indicateurs présentés ci-dessus nous donnent une valeur numérique représentant les données correspondant à un étudiant donné. Dans la section présentation des résultats, nous symboliserons ces données numériques par des indicateurs en couleur. Trois couleurs sont choisies : la couleur verte signifie que l'étudiant participe activement, le jaune signifie que l'étudiant pourrait être plus impliqué enfin le rouge signifie que l'étudiant ne participe pas assez au cours en ligne et doit absolument changer sa façon de travailler. Bien entendu, chaque indicateur de couleur doit être analysé en fonction des critères qu'il représente.

3.3 Phase de préparation des données

Cette troisième phase joue le rôle de relais entre la phase d'analyse et la phase de présentation des résultats. C'est également une phase essentielle permettant d'assurer l'interopérabilité de notre outil. Le but est de permettre d'une part de regrouper, transformer et générer des fichiers JSON de nomenclature standardisée contenant les données essentielles à notre outil, cachant ainsi leurs provenances principales. D'autre part de donner la possibilité aux développeurs d'étendre l'utilisation de notre outil à d'autres EAL, en utilisant tout langage de programmation permettant la génération de ces mêmes fichiers (par exemple : PHP ou Python).

3.4 Phase de présentation des résultats

Pour cette phase, des rapports sous forme de tableaux de bord sont construits et visualisés par les parties prenantes. Ces rapports communiquent directement avec les fichiers JSON pour récupérer les données d'analyse nécessaires. Deux vues sont présentées indépendamment : le rapport enseignant et le rapport élève.

3.4.1 Rapport enseignant

Le rapport pour l'enseignant présente des données statistiques (provenant des indicateurs de la catégorie cours), ainsi que des graphiques de synthèse (provenant des indicateurs des catégories participation, section et sociale) sur la consultation et la contribution des étudiants dans la classe pour chaque section du cours en ligne, sans oublier le calcul du progrès et de la réussite (issus des indicateurs des catégories progression et succès). La figure 2 donne une vue générale des tableaux de bord du rapport enseignant. Pour une question d'éthique, nous avons caché les noms des élèves.

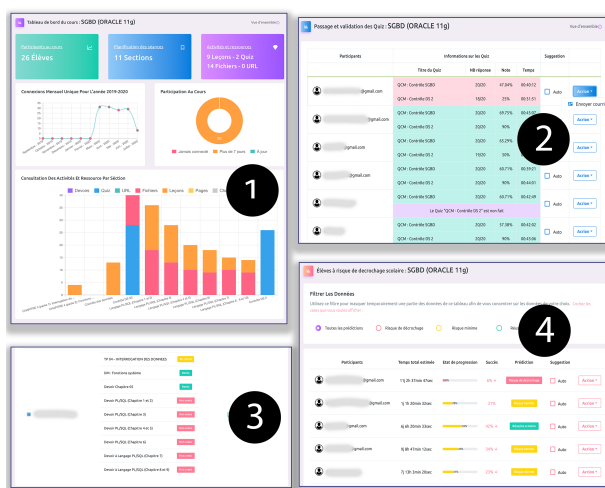


Fig. 2. Vue générale des tableaux de bord du rapport enseignant

La première page de ce rapport (1) donne un résumé du cours (le nombre d'étudiants inscrits, le nombre de sections, le nombre d'activités/ressources du cours, le nombre d'étudiants qui participent activement au cours), des statistiques sur les connexions mensuelles pour l'année en cours ainsi que des statistiques sur le nombre de fois que les étudiants consultent les activités et les ressources. La seconde page du rapport (2) fournit une analyse détaillée des quiz (liste des quiz passés ou non, nombre de questions

répondues sur le nombre total de questions, score final obtenu en pourcentage et le temps de passage de chaque élève). La troisième page (3) affiche une analyse des devoirs par titre et statut (rendu ou non ou bien en retard) de chaque élève. La dernière page (4) présente un tableau qui affiche pour chaque élève une estimation du temps total passé dans le cours, un état de progression avec succès, le niveau de réussite, une flèche pointant vers le haut ou vers le bas indiquant soit l'augmentation soit la diminution du niveau de réussite par rapport à la dernière valeur enregistrée (cette flèche n'est pas affichée si le niveau ne change pas), ainsi qu'un statut permettant d'indiquer le résultat de la prédiction : risque d'abandon, risque minimal ou réussite.

Un codage couleur permet de différencier visuellement les devoirs déposés par les élèves de ceux qui ne le sont pas, les quiz faits de ceux qui ne le sont pas. Ce codage est aussi utilisé pour le niveau de progression et le statut de prédiction.

3.4.2 Rapport élève

Le rapport pour l'élève présente une vision synthétique qui se veut épuré mais efficace. Les données affichées sont basées sur les indicateurs présents dans la catégorie participation, section, progression et sociale. La figure 3 regroupe les pages de ce rapport.

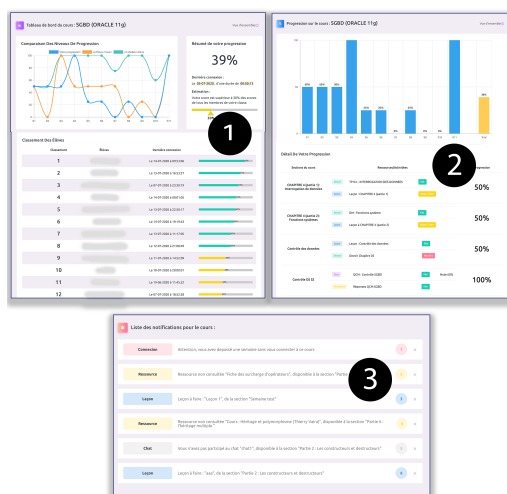


Fig. 3. Vue générale des tableaux de bord du rapport élève

La première page (1) donne un positionnement du niveau de progression de l'élève pour chaque section du cours avec celui du meilleur élève et de l'élève moyen. Cette page affiche aussi le niveau de progression global en pourcentage, la date et la durée de la dernière connexion ainsi qu'une estimation du classement de l'étudiant par rapport

aux autres élèves de la classe. Un tableau de classement de tous les élèves de la classe est affiché aussi, ce tableau vise à motiver et à soutenir les processus métacognitifs et d'autorégulation des élèves. Dans la deuxième page (2) l'élève peut suivre les détails de sa progression dans le cours en ligne, avec les détails de la progression dans chaque section du cours. La troisième page (3) affiche des notifications que l'élève peut consulter, cette page sera expliquer plus en détail dans la phase suivante.

3.5 Phase proactive

Cette dernière phase permet à l'enseignant de contacter les élèves manuellement ou de programmer des notifications automatiques. Les trois dernières pages du rapport enseignant (2, 3, 4) présenté dans la figure 2, donnent à l'enseignant la possibilité de choisir le ou les étudiants qui recevront automatiquement des suggestions (ou notifications) concernant leurs réalisations, des devoirs à soumettre ou des quiz à faire, des ressources à consulter ou même des leçons à visionner. Chaque page comporte également un bouton permettant de contacter l'élève par courrier électronique. La dernière page (3) du rapport élève présentée dans la figure 3, permet de consulter les notifications (marquées comme non lues) envoyées automatiquement. Ces notifications sont affichées par type (connexion, ressources, social, leçon, quiz et devoir). L'objectif est d'alerter l'élève par diverses suggestions et d'actions à entreprendre.

Un script Linux est programmé pour envoyer automatiquement les notifications deux fois par jour, à 8 heures et à 20 heures. Si la même notification a déjà été envoyée et n'a pas encore été lue, l'envoi n'est pas effectué.

4 Contexte de l'étude

4.1 Objectif de l'étude

Cette étude a pour objectif de tester l'outil et d'évaluer son impact sur l'autorégulation et la prédiction de la réussite des élèves d'une classe (deuxième année du cycle ingénieur) d'un institut supérieur (ISGA de Marrakech) au Maroc.

4.2 Participants

Cette étude est faite sur une population de 26 élèves suivant tous un cours organisé en modalité mixte. Afin de respecter le protocole sanitaire mis en vigueur pendant la pandémie de Covid-19, l'institut a réduit le nombre d'élèves de 10 à 14 élèves maximum, avec le maintien de la distanciation physique entre les élèves de groupes différents.

La classe cible est divisée en deux groupes de 13 élèves (8 de sexe féminin), âgés de 18 à 35 ans (24 entre 18 et 25 ans et 2 entre 26 et 35 ans). En termes de temps d'utilisation d'Internet et des appareils informatiques par jour, 6 élèves ont déclaré que leur temps d'utilisation étaient entre deux et cinq heures, 18 entre cinq et dix heures et 2 autres plus de dix-heures par jour.

4.3 Méthodologie de l'étude

Le cours a été disponible en modalité mixte (blended learning), qui combine formation en présence et en ligne. Les élèves des deux groupes, ont tous suivi en présentiel un cours intitulé « Base de données orienté objet », avec certains chapitres en ligne sur la plateforme Moodle 3.8, sur une période de huit semaines, finalisé par un examen en présentiel surveillé.

Afin d'évaluer l'impact de l'outil sur l'autorégulation et la prédiction de la réussite des élèves en ligne, un des deux groupes a pris connaissance de l'expérience visant à utiliser l'outil TaBAT (groupe exposé), tandis que le second groupe n'a pas eu accès au tableau de bord (groupe témoin).

5 Résultats de l'étude

La partie du cours suivie par les élèves en ligne comporte 7 sections (parties), avec 3 leçons, 12 fichiers à télécharger, 2 liens URL à visiter et 4 devoirs à rendre à des dates planifiées au début du cours. L'analyse des traces propres aux élèves de chaque groupe s'est faite à l'aide des tableaux de bord de l'outil TaBAT via le rapport enseignant. Le tableau 1 décrit le résultat de l'expérience menée sur les deux groupes.

Tableau 1. Statistique d'usage de l'outil TaBAT.

	Groupe exposé	Groupe témoin
Nb. d'utilisateurs actifs	13	11
Cumul de la durée totale à faire le cours	45 h 57 min 15 s	15 h 57 min 23 s
Score moyen de progression sur une échelle de 0 à 100	85,54	37,61
% des devoirs rendus à temps	44,23 %	11,54 %
% des devoirs rendus en retard	26,92 %	15,38 %
% des devoirs non rendus	28,85 %	73,08 %
Prédiction du succès (réussite en ligne)	13/13	8/13
Succès réel (validation de l'examen en présentiel)	13/13	9/13

Nous remarquons au départ que tous les élèves du groupe exposé se sont connectés au cours en ligne, tandis que pour le groupe témoin, 2 élèves n'ont pas suivi la partie du cours en ligne. Ceci s'explique par le fait que, pour le groupe témoin, la seule façon de communiquer avec le professeur était en présentiel. Tandis que pour le groupe exposé, l'enseignant avait la possibilité (à travers l'outil TaBAT) de contacter chaque élève par e-mail, ce qui a permis d'avoir un suivi individuel. La deuxième remarque concerne la durée totale passée à faire les activités du cours en ligne. Cette durée est représentée dans le tableau 1, cumulée pour chaque groupe. Les élèves du groupe exposé ont passé presque trois fois plus de temps que ceux du groupe témoin à suivre et à faire le cours en ligne. Cette augmentation traduit donc un fonctionnement autonome

et une résistance aux distractions, faisant ainsi du travail à la maison une forme particulière d'apprentissage autorégulé.

La troisième remarque porte sur la performance de chaque groupe, cette performance est représentée dans le tableau 1 par le score moyen de progression dans le cours en ligne. Rappelons que la progression de chaque élève représente le nombre d'activités ou ressources consultées ou bien accomplies sur le nombre d'activités ou ressources définies par l'enseignant au début du cours. Nous remarquons une augmentation significative de la performance du groupe exposé, expliqué principalement par les actions proactives faites manuellement par l'enseignant ou bien envoyé automatiquement par l'outil TaBAT (phase proactive), afin de rappeler aux élèves (à l'aide de notifications) s'il leur reste des ressources non consultées (fichier à télécharger ou url à visiter) ou des activités non accomplies (leçons, devoir à rendre, quiz à faire, etc.). Sans oublier le rôle important du rapport élève qui permet de s'autoévaluer et de suivre des stratégies métacognitives pour améliorer leurs performances en ligne.

La quatrième remarque concerne l'analyse du rendu des devoirs. En effet, le groupe exposé a un taux de rendu des devoirs (planifiés au début du cours) de 71,15 % (devoirs rendus à temps avec ceux rendus en retard), tandis que pour le groupe témoin, le même taux était de 26,92 %. Ce résultat s'explique tout particulièrement par les notifications reçues s'il y a un devoir à rendre, ou bien non remis dans le délai (en retard). La notification comporte la date et le nombre de jours pour rendre le devoir.

La dernière remarque touche le succès des élèves. Dans cette étude, le niveau de succès calculé par l'outil TaBAT est comparé par rapport à celui obtenu après l'examen terminal. Le tableau 1 montre que l'outil TaBAT a pu prédire à 100 % le succès des élèves du groupe exposé, et à 80 % le succès des élèves du groupe témoin. Ce résultat démontre que la performance et la réussite des élèves tient en partie à leur capacité à fournir un travail efficace à la maison, les préparant mieux aux contrôles en classe.

6 Conclusion et limites

Nous avons présenté dans cet article la conception du tableau de bord d'apprentissage multiplateforme TaBAT et son expérimentation dans un cours d'informatique au Maroc. TaBAT permet aux enseignants et aux apprenants de visualiser différents indicateurs pédagogiques afin de contrôler le processus d'apprentissage. Après une étude sur une classe cible divisée en deux groupes, nous pouvons conclure que l'outil a fourni des informations claires et pertinentes pour motiver et réguler l'effort des élèves, sans oublier l'aide qu'il a pu apporter à l'enseignant pour suivre les élèves et détecter les décrochages scolaires.

Nous terminerons en soulignant que cette première étude comporte plusieurs limites, la première est le faible nombre de participants. En effet, cette étude est limitée à une classe, l'intégration de plusieurs classes dans l'étude demanderait des moyens plus importants. En raison de cette limite de taille d'échantillon, des analyses statistiques plus poussées n'ont pas été réalisées. Enfin, étant donné que l'outil TaBAT est nouvellement développé comme outil externe aux EAL, le test dans d'autres organisations faisant partie du secteur de l'enseignement privé ou public, nécessiterait des autorisations et

du travail logistique. Ces premiers résultats sont toutefois très encourageants, ils devront être confirmés dans une étude de plus grande ampleur.

Références

1. Ferguson, R.: Learning analytics: Drivers, developments and challenges. *Int. J. Technol. Enhanc. Learn.* 4, 304–317 (2012). <https://doi.org/10.1504/IJTEL.2012.051816>.
2. Siemens, G.: Analytics in learning and education. In: 1st International Conference on Learning Analytics and Knowledge. pp. 31–40 (2011).
3. Biggs, J.: What the student does: Teaching for enhanced learning. *High. Educ. Res. Dev.* 31, 39–55 (2012). <https://doi.org/10.1080/07294360.2012.642839>.
4. Schwendimann, B.A., Rodriguez-Triana, M.J., Vozniuk, A., Prieto, L.P., Boroujeni, M.S., Holzer, A., Gillet, D., Dillenbourg, P.: Understanding learning at a glance: An overview of learning dashboard studies. *ACM Int. Conf. Proceeding Ser.* 25-29-April, 532–533 (2016). <https://doi.org/10.1145/2883851.2883930>.
5. Durall, E., Gros, B.: Learning analytics as a metacognitive tool. *CSEDU 2014 - Proc. 6th Int. Conf. Comput. Support. Educ.* 1, 380–384 (2014). <https://doi.org/10.5220/0004933203800384>.
6. Jivet, I., Scheffel, M., Drachler, H., Specht, M.: Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 10474 LNCS, 82–96 (2017). https://doi.org/10.1007/978-3-319-66610-5_7.
7. Zimmerman, B.J., Moylan, A.R.: Self-Regulation, Where Metacognition and Motivation Intersect. *Handb. METACOGNITION Educ.* 299–315 (2009). <https://doi.org/10.4324/9780203876428.ch16>.
8. Nicholas, D., Grover, S., Eagle, M., Bienkowski, M., Stamper, J., Basu, S.: An instructor dashboard for real-time analytics in interactive programming assignments. *ACM Int. Conf. Proceeding Ser.* 272–279 (2017). <https://doi.org/10.1145/3027385.3027441>.
9. Siemens, G., Dragan, G., Haythornthwaite, C., Dawson, S., Buckingham Shum, S., Ferguson, R., Duval, E., Verbert, K., Baker, R.S.J. d.: Open Learning Analytics: an integrated & modularized platform. 145–150 (2011).
10. Luengo, V.: *Projet ANR HUBBLE HUMAN OBSERVATORY BASED ON ANALYSIS OF E-LEARNING TRACES*. (2014).
11. Dabbebi, I.: *Conception et génération dynamique de tableaux de bord d'apprentissage contextuels* To cite this version : HAL Id : tel-02443909 Inès DABBEBI Conception et génération dynamique de tableaux de bord. (2020).
12. Monllaó, D., Dalton, E.: Moodle Inspire, https://moodle.org/plugins/tool_inspire.
13. Moodle Analytics, <https://docs.moodle.org/35/en/Analytics>.
14. Safsouf, Y., Mansouri, K., Poirier, F.: An analysis to understand the online learners' success in public higher education in morocco. *Journal of Information Technology Education: Research*. 19, 87–112 (2020). <https://doi.org/https://doi.org/10.28945/4518>.
15. Safsouf, Y., Mansouri, K., Poirier, F.: A new model of learner experience in online learning environments. *Int. Conf. Proceeding. Information Systems and Technologies to Support Learning*. 29-38 (2019). https://doi.org/10.1007/978-3-030-03577-8_4.

Détection de déficits d'auto-évaluation et d'auto-efficacité dans un logiciel enseignant la lecture et l'écriture

Thomas Sergent^{1,2}, François Bouchet¹, Morgane Daniel² et Thibault Carron¹

¹ Sorbonne Université, CNRS, LIP6, F-75005 Paris, France {prénom.nom}@lip6.fr
² Lalilo, Paris, France {prénom}@lalilo.com

Résumé. Plusieurs travaux montrent que la capacité à auto-réguler son apprentissage a un impact significatif sur les résultats scolaires. Nous présentons ici une étude visant à détecter les déficits d'auto-régulation de l'apprentissage liés à l'auto-évaluation et à l'auto-efficacité pour de jeunes (5-7 ans) élèves, dans le contexte d'une application web d'apprentissage de la lecture. Nous avons recueilli les réponses de plus de 15 000 enfants travaillant sur une telle application en classe. À partir de ces réponses, nous proposons une définition opérationnelle de différentes formes de déficits dont nous évaluons la prévalence auprès des élèves.

Mots-clé: Auto-régulation de l'apprentissage · Fouille de données éducatives · École primaire · Apprentissage de la lecture

Abstract. The ability to self-regulate one's learning is considered to have a significant impact on educational outcomes. We present here a study to detect self-regulated learning (SRL) deficits related to self-evaluation and self-efficacy for young (5-7 years old) students, in the context of a literacy web application. We collected answers to SRL statements of over 15,000 children working on such an application while in the classroom. From these SRL answers we propose an operational definition of different forms of deficits whose prevalence among students we assess.

Keywords: Self-regulated learning · Educational Data Mining · Primary school · Learning to read

1 Introduction

L'amélioration des compétences des enfants en matière d'apprentissage auto-régulé est critique pour améliorer les performances scolaires car les élèves auto-régulés savent globalement mieux comment apprendre, ce qui peut avoir un impact positif dans toutes les disciplines [16]. Plus tôt les enfants commencent à développer ces compétences, plus l'impact sur l'ensemble de leur scolarité peut se faire sentir, et des programmes de formation à l'auto-régulation pour les élèves de l'école primaire ont déjà été élaborés dans ce but [6]. Néanmoins, il peut être difficile pour les enseignants de se concentrer sur l'aide individualisée à apporter

à chaque élève, à la fois sur la tâche à accomplir (par exemple, apprendre à lire) et sur leurs compétences d'auto-régulation.

L'auto-régulation de l'apprentissage est un cycle en trois phases qui se répète à chaque nouvelle tâche à laquelle l'apprenant est confronté [16]. D'abord la phase d'anticipation pendant laquelle l'apprenant se prépare à la tâche (ex : choix d'objectifs d'apprentissage ou activation de connaissances antérieures relatives à la tâche), puis la phase de performance pendant laquelle l'apprenant exécute la tâche et où il peut suivre ses progrès vers son objectif d'apprentissage, et enfin la phase d'auto-réflexion qui consiste notamment à évaluer son efficacité d'apprentissage afin de tirer des conclusions pour l'apprentissage futur. Nous nous concentrons ici sur cette phase d'auto-réflexion qui permet de travailler les compétences d'auto-régulation sans interférer avec la tâche à accomplir. Dans ce contexte, nous visons plus particulièrement deux aspects : l'**auto-efficacité** qui concerne la perception de ses propres compétences à accomplir une tâche [3], et l'**auto-évaluation** qui concerne les jugements relatifs à sa propre performance et les réactions à ces jugements [12]. En effet, l'amélioration du sentiment d'auto-efficacité est corrélée avec des gains d'apprentissage accrus [7] et l'auto-évaluation est un processus clé de l'auto-régulation [13]. De plus l'auto-évaluation est une capacité qui se développe progressivement mais dont disposent de jeunes enfants de 5 ans et plus [14], tout comme l'auto-efficacité dans des domaines liés à l'apprentissage de l'écriture [8].

Une méta-analyse des dispositifs informatiques mis en place jusqu'à 2016 pour aider l'auto-régulation de l'apprentissage montre leur effet positif significatif sur la progression [15]. Cependant, ces dispositifs concernent uniquement des élèves plus âgés (au-delà du CM2), ils se concentrent sur la phase de performance, et on mesure la progression de l'élève grâce aux dispositifs favorisant l'auto-régulation plus que la progression des capacités d'auto-régulation elles-mêmes. En effet, l'auto-régulation est surtout vue comme soutenant l'apprentissage, plutôt que comme une compétence à évaluer et entraîner en tant que telle. Dans cet article nous envisageons un autre angle : celui de la mesure directe de capacités d'auto-régulation, en vue à terme d'améliorer celles-ci.

Lalilo est l'une des nombreuses applications web utilisées par les enseignants en classe pour les aider à mettre en place une pédagogie différenciée. Elle est actuellement utilisée par 40 000 classes de maternelle et élémentaire anglophones et francophones chaque semaine pour renforcer l'alphabétisation en proposant une série d'exercices adaptés au niveau des élèves, tout en offrant à l'enseignant un tableau de bord pour suivre les activités et les progrès des élèves. Il s'agit donc d'un terrain d'essai pertinent pour évaluer puis essayer de corriger les capacités d'auto-évaluation et d'auto-efficacité. Un défi supplémentaire est qu'il n'existe à notre connaissance pas d'études sur l'application de ces approches aux enfants de cet âge (5-7 ans), d'où le besoin de savoir si l'on peut identifier en contexte ce phénomène (pour ensuite tenter d'y remédier) et estimer correctement sa fréquence (pour savoir quels déficits viser en priorité). Plus précisément, nous examinerons deux questions de recherche : (QR1) Peut-on mesurer les capacités d'auto-évaluation et d'auto-efficacité des jeunes élèves qui apprennent avec une

application web ? (QR2) Les déficits d'auto-évaluation et d'auto-efficacité sont-ils des problèmes courants pour les jeunes élèves qui apprennent à lire ?

Dans la suite de cet article, nous commencerons par examiner les travaux connexes sur la mesure et l'entraînement des capacités d'auto-régulation dans le cas des jeunes enfants. En section 3, nous présenterons le fonctionnement de Lalilo et l'intégration des évaluations des compétences d'auto-régulation. La section 4 présentera les données recueillies, les déficits d'auto-régulation détectés et leur fréquence, avant de discuter des limites et perspectives en section 5.

2 Travaux connexes

Dans le contexte de l'apprentissage sur ordinateur, l'auto-régulation peut être soutenue par différents types d'étayages (*scaffolding*) [1] comme des invites (*prompts*) [5] ou des rétroactions (*feedback*) automatisées [4].

Pour mesurer les capacités d'auto-régulation, on a parfois recours (hors cadre informatique) à des journaux d'apprentissage (*learning diaries*) [11], car les traces de systèmes informatiques sont souvent difficile à interpréter en termes d'auto-régulation [9].

Des programmes d'entraînement à l'auto-régulation ont montré leur effet positif significatif chez des enfants de primaire, mais hors contexte informatique [6]. D'autres, comme MetaTutor, mesurent et entraînent l'auto-régulation mais pour des étudiants du supérieur [2]. De plus, une des conclusions de [2] était la nécessité d'un temps long pour mesurer un impact, d'où la pertinence de mesurer l'auto-régulation dans un logiciel tel que Lalilo utilisé pendant une voire plusieurs années scolaires (de la Grande Section au CE1). Molenaar et al. [10] vise à entraîner les capacités d'auto-régulation d'élèves de CM2 via des tableaux de bord. Ils montrent une amélioration des capacités d'auto-régulation des élèves ayant accès au tableau de bord (mieux régulés), notamment par l'usage de la forme des *Moment by Moment Learning Curves*, démontrant ainsi également qu'on peut mesurer l'auto-régulation chez des enfants assez jeunes avec une approche informatique. Dans notre cas, les élèves sont encore plus jeunes (5-7 ans vs. 9-10 ans) et notre métrique se base directement sur la réponse d'un élève à des questions d'auto-régulation.

3 Évaluer l'auto-évaluation et l'auto-efficacité des élèves

3.1 Contexte

Lalilo possède deux interfaces : une interface élève (cf. Figure 1) dans laquelle l'élève répond à des exercices et un tableau de bord enseignant sur lequel l'enseignant peut visualiser la progression de ses élèves. Cet outil étant destiné aux élèves de 5 à 7 ans, il dispose d'une interface simple. En général, les instructions sont uniquement lues (et non écrites) et réécoutables. L'application couvre une grande variété d'exercices sur un champ de difficulté étendu, allant de la maternelle avec des exercices d'association graphèmes-phonèmes jusqu'à des exercices de conjugaison et de vocabulaire pour les CE1-CE2.

Le déroulement typique d'une session (20 minutes en moyenne) est la réalisation par l'élève d'une quinzaine d'exercices courts de 3 à 7 questions chacun, choisis par un algorithme d'apprentissage adaptatif (non détaillé ici). Pour

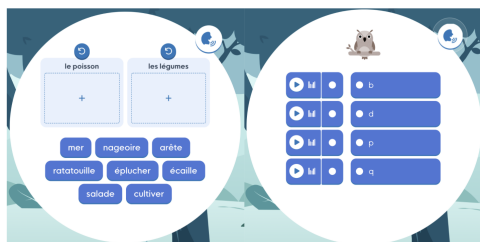


Fig. 1: Trier des mots en fonction de leur champ sémantique (gauche) et relier un son à la lettre correspondante (droite)

certains types d'exercices, l'élève peut essayer plusieurs fois une question. Les activités des élèves (par exemple, connexion, temps passé sur une question/un exercice, erreurs) sont tracés. Ici, nous nous concentrerons particulièrement sur les réponses des élèves à un exercice, nous appellerons donc désormais **trace** uniquement les réponses à cet ensemble de questions du même type.

3.2 Méthodes

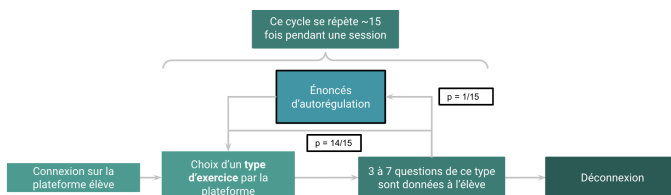


Fig. 2: Session type d'un élève sur Lalilo (durée moyenne : 20 minutes)



Fig. 3: Énoncé demandant la difficulté perçue (gauche) puis la difficulté voulue (droite). Une fois qu'une réponse est sélectionnée, un bouton de confirmation s'affiche dessous. Par exemple, à droite, l'élève a sélectionné "de même niveau".

Collecte des données, filtrage et nettoyage Pour évaluer certains aspects des compétences d'auto-régulation des élèves, nous avons introduit deux énoncés (cf. figure 3) affichés l'un après l'autre à la fin d'un exercice à la fréquence *Freq_{autoreg}* d'une fois tous les quinze exercices. Chaque élève y répond donc en moyenne une fois pendant une session d'apprentissage type (cf. Figure 2). Tout d'abord, l'énoncé de **difficulté perçue** demande à l'élève : "Quelle était la difficulté de cet exercice pour toi ? Ensuite, l'élève doit compléter l'énoncé de **dif-**

ficulté souhaitée "*Tu voudrais des exercices...*". L'énoncé de difficulté perçue vise à mesurer la capacité **d'auto-évaluation** des élèves, c'est-à-dire leur capacité à estimer correctement la difficulté des questions auxquelles ils viennent de répondre. L'énoncé de difficulté souhaitée vise à mesurer leur **sentiment d'auto-efficacité**, c'est-à-dire la façon dont ils réagiraient à leur représentation de la difficulté. Avant d'introduire les évaluations, nous avons vérifié qualitativement dans une classe utilisant Lalilo que les énoncés étaient compris par les élèves de CP. Les élèves interagissaient avec Lalilo normalement pendant qu'un expérimentateur était assis derrière eux et observait leur réaction devant les 2 questions. Ensuite une discussion permettait de tester leur compréhension du concept de "difficulté". Bien qu'informel et sur un échantillon réduit, ce travail a permis de vérifier que les énoncés compris par les élèves ne présentaient pas de décalage complet par rapport à l'intention sous-jacente. Elle a aussi permis de choisir la formulation la plus claire pour les élèves lorsque plusieurs options étaient envisagées (détails non présentés ici).

Nous avons recueilli des traces de classes de maternelle, de CP et de CE1 basées en France, au Canada et aux Etats-Unis apprenant le français (FR) ou l'anglais (EN) entre le 1er août et le 26 octobre 2020 sur la plateforme Lalilo. Cette période correspond à un moment où les élèves se trouvaient principalement dans les classes et non à la maison. Nous n'avons conservé que les traces pour lesquelles les élèves avaient répondu aux questions d'auto-régulation et nous allons désormais appeler *trace* l'ensemble regroupant les réponses à l'exercice avec les réponses associées aux énoncés d'auto-régulation.

Nous avons limité les sources potentielles de biais dans nos données en identifiant plusieurs phénomènes qui pourraient avoir un impact sur elles, notamment :

1. pas assez de réponses aux énoncés d'auto-régulation pour un élève donné
2. élèves ayant ignoré ou mal compris les questions, ce qui devrait entraîner une tendance à répondre presque au hasard, puisqu'il est impossible de sauter ou de ne pas répondre aux énoncés d'auto-régulation

Pour le premier point, nous avons filtrons les élèves ayant eu moins de $N_{min} = 12$ réponses aux énoncés d'auto-régulation. En effet, observer des déficits établis nécessite de mesurer suffisamment de données par élève. Ce seuil peut sembler élevé : il est probable qu'un élève présentant 5 fois un même déficit sur des questions puisse être raisonnablement considéré comme souffrant de ce déficit d'auto-régulation. Néanmoins ce choix conservateur renforce la certitude du diagnostic quand il est posé, et limite donc le nombre de faux positifs.

En ce qui concerne le deuxième point, nous avons supprimé les élèves semblant répondre au hasard. Pour déterminer si une réponse implique une part de hasard, nous considérons que deux combinaisons de réponses sont particulièrement incohérentes (cf. figure 4 - gauche). Nous considérons qu'un élève répond potentiellement au hasard si les deux combinaisons de réponses incohérentes apparaissent au moins une fois dans ses traces. Là encore, ce choix est conservateur (un élève peut faire une erreur une fois sans que ça soit significatif d'une tendance à répondre au hasard) et a vocation à limiter le nombre faux positifs. L'impact des 2 filtres est résumé sur la figure 4 (droite) : le filtre le plus important est le

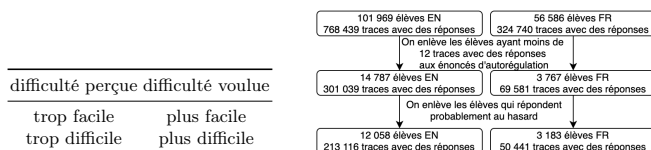


Fig. 4. Labellisation des réponses incohérentes (gauche) et filtrage des élèves (droite)

premier qui divise par plus de 10 la taille de l'échantillon initial. Toutefois cela ne signifie pas une incapacité à identifier des déficits chez ces élèves. On pourrait ainsi (1) soit abaisser le seuil de 12 (élevé, comme susmentionné), (2) soit étendre la période de collecte (2 mois actuellement) pour qu'ils aient répondu 12 fois aux 2 questions, ce qui devrait arriver à tout élève utilisant régulièrement Lalilo, (3) soit augmenter la fréquence de prompt ($Freq_{autoreg} = \frac{1}{15}$ actuellement).

3.3 Caractérisation des déficits

Au niveau des traces. Nous rappelons que dans notre contexte, les **traces** considérées sont celles produites lorsqu'un exercice (trois à sept questions du même type) est suivi des deux énoncés d'auto-régulation. Une trace enregistre donc les réponses à chaque question de l'exercice ainsi qu'aux deux énoncés d'auto-régulation. On peut alors calculer le taux de réussite d'une trace, défini comme le nombre de réponses correctes sur le nombre total de questions de la trace. Comme pour certains types d'exercices, l'élève peut répondre plusieurs fois (parfois avec des indications fournies entre les essais), nous ne prenons en compte que la première réponse pour le calcul du taux de réussite. À partir du taux de réussite, nous pouvons déterminer un **libellé de performance** d'une trace avec l'une des trois valeurs suivantes :

- excellente : si toutes les réponses étaient correctes, on note ce seuil $Perf_+$ (100% ici)
- mauvaise : si $Perf_- = 34\%$ ou moins des réponses étaient correctes
- moyenne : si $Perf \in]Perf_-, Perf_+[$.

Nous avons choisi un seuil de 34% pour $Perf_-$, de sorte que les traces qui n'ont qu'une seule bonne réponse sur 3 soient considérées comme faibles. En effet, pour un QCU avec 3 choix, la probabilité attendue de réussite aux questions est toujours d'au moins $1/3$, ce qui signifie que les élèves ayant un taux de réussite d' $1/3$ ou moins n'obtiennent pas de meilleurs résultats que le hasard. Il convient également de noter que le seuil $Perf_+$ est, ici encore, assez conservateur, car on pourrait estimer qu'un élève qui a répondu correctement à 6 questions sur 7 pourrait être considéré comme ayant également de très bonnes performances.

Notre objectif est de comparer la performance réelle d'un élève avec la difficulté qu'il a perçue, puis sa difficulté perçue - qui est une représentation subjective - avec la difficulté qu'il aimerait avoir pour les prochains exercices (cf. figure 5). En effet, la comparaison entre la difficulté souhaitée et la performance réelle peut ne pas être pertinente si les élèves sont biaisés dans leur perception de leur performance réelle ou dans la difficulté de la tâche. À partir de la différence

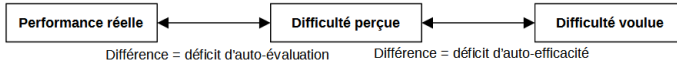


Fig. 5: Caractérisation des déficits d'auto-évaluation et d'auto-efficacité

entre leur performance et leur perception de la difficulté, on en déduit un **libellé d'auto-évaluation** (cf. tableau 1). Ensuite, à partir de la différence entre la difficulté perçue et la difficulté souhaitée, on en déduit un **libellé d'auto-efficacité** (cf. tableau 2). Notons que les deux couples de la figure 4 peuvent être considérés soit comme incohérents, soit comme étant les signes d'une auto-efficacité très élevée - pour (trop dur, plus dur) - ou très faible - pour (trop facile, plus facile).

Dans la perspective d'apporter une remédiation aux déficits présentés, il est nécessaire d'introduire une notion de priorité entre les déficits, car un élève peut avoir un déficit d'auto-évaluation et d'auto-efficacité. On introduit donc un libellé supplémentaire - le **libellé de déficit global** - dont l'intérêt est d'être orienté vers le feedback à donner (à l'élève ou l'enseignant). Nous pensons qu'il est d'abord nécessaire de résoudre les éventuels déficits d'auto-évaluation avant de s'attaquer aux déficits d'auto-efficacité. Ainsi, si le libellé d'auto-évaluation de la trace montre un déficit d'auto-évaluation, alors le déficit de la trace est ce déficit d'auto-évaluation. S'il n'y a pas de déficit d'auto-évaluation mais que le libellé d'auto-efficacité montre un certain déficit, alors deux nouveaux libellés peuvent apparaître : "évitant la difficulté" et "recherchant la difficulté".

Libellé de perf.	Diff. perçue	Libellé d'auto-évaluation
excellente	trop difficile	sous-évaluation
excellente	bien	légère sous-évaluation
excellente	trop facile	cohérent
mauvaise	trop difficile	cohérent
mauvaise	bien	légère sur-évaluation
mauvaise	trop facile	sur-évaluation

Tableau 1: Libellés d'auto-évaluation

Diff. perçue	Diff. voulue	Libellé d'auto-efficacité
trop difficile	plus facile	
trop difficile	de même niveau	élevé
trop difficile	plus difficile	très élevé/incohérent
bien	plus facile	faible
bien	de même niveau	
bien	plus difficile	élevé
trop facile	plus facile	très faible/incohérent
trop facile	de même niveau	faible
trop facile	plus difficile	

Tableau 2: Libellés d'auto-efficacité

Libellé de perf.	Diff. perçue	Diff. voulue	Libellé de déficit global
excellente	trop difficile	plus facile	sous-évaluation
excellente	trop difficile	de même niveau	sous-évaluation
excellente	trop difficile	plus difficile	sous-évaluation
excellente	bien	plus facile	légère sous-éval.
excellente	bien	de même niveau	légère sous-éval.
excellente	bien	plus difficile	légère sous-éval.
mauvaise	trop facile	plus facile	sur-évaluation
mauvaise	trop facile	de même niveau	sur-évaluation
mauvaise	trop facile	difficile	sur-évaluation
mauvaise	bien	plus facile	légère sur-éval.
mauvaise	bien	de même niveau	légère sur-éval.
mauvaise	bien	plus difficile	légère sur-éval.
excellente	trop facile	plus facile	évitant la difficulté
excellente	trop facile	de même niveau	évitant la difficulté
mauvaise	trop difficile	plus difficile	cherchant la difficulté
mauvaise	trop difficile	de même niveau	cherchant la difficulté
excellente	trop difficile	plus difficile	incohérent
mauvaise	trop facile	plus facile	incohérent

Tableau 3: Triplets de trace libellés comme ayant un déficit

Au niveau de l'élève Comme nous avons défini la labellisation des traces, nous pouvons maintenant considérer l'ensemble des N traces d'un élève telles que $N \geq N_{min}$. Notre objectif est alors de détecter certaines tendances dans les réponses de l'élève afin de caractériser globalement son profil d'auto-régulation.

L'algorithme 1 est utilisé pour déterminer si un élève donné sera labellisé comme ayant un déficit ou non. Il dépend de 2 paramètres: $FrDef_{min}$ la fréquence minimale d'un libellé de déficit dans les traces, et $NDef_{min}$ pour nombre minimum de traces devant présenter un déficit pour labelliser l'élève avec ce déficit. Ces deux paramètres sont nécessaires pour caractériser uniquement les élèves ayant un déficit marqué ($FrDef_{min}$) et limiter les faux positifs qui auraient le déficit dans leurs traces par hasard ($NDef_{min}$). Connaissant le nombre de libellés de déficit d'auto-régulation d'un élève avec le tableau 3 et le nombre de performances faibles et excellentes de l'élève, le **déficitRatio** est calculé comme le rapport entre le nombre de réponses labellisées comme ayant un déficit et le nombre de performances associées : "surévaluation" et "cherchant la difficulté" sont liées à de mauvaises performances tandis que "sous-évaluation" et "éviter la difficulté" sont liées à d'excellentes performances. Nous avons choisi les seuils du *déficitRatio* pour être significatif à 50% avec au moins 2 occurrences du déficit c'est-à-dire $NDef_{min} = 2$ et $FrDef_{min} = 50\%$. Ces valeurs sont bien supérieures au choix aléatoire qui est de 33 % (puisque 3 réponses sont proposées sur chaque énoncé) et nous voulions exclure les erreurs ponctuelles des élèves.

3.4 Résultats

Le tableau 4 résume nos résultats. Nous rappelons le choix de nos différents seuils ($Perf_+ = 100\%$, $Perf_- = 34\%$, $NDef_{min} = 2$ et $FrDef_{min} = 50\%$) est

Algorithm 1: Caractérisation des élèves

Result: Déficit de l'élève

```

for déficit in déficits do
  nbTracesAyantUnDéficit = count(traceSRLTags[déficit]);
  if déficit in [sous-évaluation, évitant la difficulté] then
    | déficitRatio = nbTracesAyantUnDéficit / nbExcellentesPerformances;
  else if déficit in [sur-évaluation, cherchant la difficulté] then
    | déficitRatio = nbTracesAyantUnDéficit / nbFaiblesPerformances;
  if déficitRatio  $\geq FrDef_{min}$  and nbTracesAyantUnDéficit  $\geq NDef_{min}$ 
    then
    | élèveADéficit[déficit]

```

Déficit(s)	EN % (N=12,058)	FR % (N=3,183)
pas de déficit détecté	71.2% (42.5%)	63.1% (38.4%)
sur-évaluation	8.9% (10.2%)	9.8% (13.6%)
évitant la difficulté	8.6%	12.4%
sous-évaluation	5.2% (32.6%)	5.3% (26.2%)
évitant la difficulté - sur-évaluation ¹	3.6%	4.4%
cherchant la difficulté	1.6%	3.2%
cherchant la difficulté - sous-évaluation ²	0.5%	1.0%
évitant la difficulté - cherchant la difficulté ³	0.2%	0.6%
sur-évaluation - sous-évaluation ³	0.2%	0.1%
cherchant la difficulté - sur-évaluation ³	0.0%	0.0%
évitant la difficulté - sous-évaluation ³	0.0%	0.1%
évitant la diff. - cherchant la diff. - sous-éval. ³	0.0%	0.0%
cherchant la diff. - sur-évaluation - sous-éval. ³	0.0%	0.0%

¹L'élève clique toujours sur "trop facile", quelle que soit sa performance

²L'élève clique toujours sur "trop difficile", quelle que soit sa performance

³L'élève répond probablement au hasard

Tableau 4: Pourcentage global d'élèves ayant un ou plusieurs déficits (un élève ne peut appartenir qu'à une seule ligne). Les valeurs entre parenthèses correspondent à l'inclusion des libellés *légère* sous-évaluation ou sur-évaluation

assez conservateur, de sorte que nous détectons probablement moins de déficits qu'il n'y en a en réalité. Nous avons défini deux possibilités de calcul des déficits de surévaluation et de sous-évaluation. Pour le calcul de la première valeur de ces déficits dans le tableau, seules les traces marquées comme "surévaluation" et "sous-évaluation" dans le tableau 3 sont incluses ; tandis que pour la valeur entre parenthèses, les traces marquées comme "légère sous-évaluation" et "légère surévaluation" sont également incluses. Cela nous permet d'avoir une estimation de la limite inférieure et supérieure de ces deux déficits. Les trois déficits les plus fréquents correspondent à trois des quatre déficits définis dans la section 3.3 : surévaluation, évitant la difficulté et sous-évaluation. L'ordre entre les trois varie si l'on tient compte des libellés "légère surévaluation" et "légère sous-évaluation". En effet, la limite supérieure de la prévalence de la sous-évaluation est supérieure à 25 % pour les élèves francophones (FR) et anglophones (EN), ce qui suggère qu'un nombre important d'élèves se sous-évaluent dans une certaine mesure. Le

quatrième schéma le plus fréquemment détecté concerne les élèves qui présentent à la fois des déficits de "évitant la difficulté" et de "surévaluation". Bien que contradictoires à première vue, en regardant les triplets de trace associés à ces deux déficits dans le tableau 3, on remarque qu'ils sont associés à une difficulté perçue "trop facile". Nous supposons donc que les élèves labellisés comme ayant ces deux déficits ne comprennent pas correctement les énoncés d'auto-régulation mais cliquent toujours sur "trop facile". De même, pour le sixième déficit le plus fréquemment détecté ("cherchant la difficulté - sous-évaluation") où nous supposons que ces élèves cliquent toujours sur "trop difficile".

Enfin, les combinaisons de déficit dans les six dernières lignes du tableau 4 représentent très peu d'élèves, nous pouvons donc probablement nous permettre de ne pas les considérer comme des phénomènes distincts. En regardant les triplets de réponses associés, nous remarquons que ces combinaisons de déficit sont incohérentes, ce qui pourrait correspondre à des élèves répondant de manière aléatoire mais non éliminés par le filtre initial.

L'ordre de prévalence des déficits est assez similaire chez les élèves FR et EN. La fréquence des réponses "évitant la difficulté" et "cherchant la difficulté" est plus élevée chez les élèves FR. Cela pourrait indiquer que les élèves FR et EN ont des capacités d'auto-évaluation comparables, mais que les élèves FR ont plus de difficultés à se positionner par rapport à la difficulté attendue. Bien que ce résultat puisse être lié à des approches pédagogiques différentes, nous restons cependant prudents car il pourrait y avoir des biais liés aux différences de didactique des langues et au profil des utilisateurs selon la langue utilisée.

Ces résultats montrent la possibilité de détecter de manière assez fiable des déficits d'auto-évaluation et d'auto-efficacité chez de jeunes élèves, et la fréquence des phénomènes détectés suggère qu'il serait nécessaire d'essayer de les traiter.

4 Discussion, limites et perspectives

Notre premier objectif était de vérifier la capacité à proposer un modèle permettant à un logiciel d'estimer des déficits d'auto-évaluation et d'auto-efficacité. Comme toute intervention, l'intervention proposée peut influencer les performances des élèves et leurs compétences réelles d'auto-régulation. Une première limite est donc liée au compromis nécessaire entre la valeur du seuil $Fr_{eqautoreg}$, fréquence des mesures et l'impact potentiel des mesures sur l'auto-régulation des élèves. L'équilibre trouvé ici garantit qu'en général, un élève ne devrait pas recevoir plus d'un énoncé par session d'apprentissage, mais nous n'avons pas évalué une éventuelle "fatigue de l'énoncé" qui peut conduire à des réponses peu fiables. Nous nous intéressons à l'auto-évaluation et l'auto-efficacité d'une part car il était possible de le faire techniquement avec Lalilo, et d'autre part parce qu'elles ont rarement été étudiées dans le contexte informatique pour des élèves aussi jeunes. D'autres aspects liés à l'auto-régulation comme la planification supposent une maîtrise par l'élève des compétences à travailler (ce qui n'est pas le cas dans Lalilo qui guide l'élève avec un algorithme d'apprentissage adaptatif).

Nous détectons des réponses aléatoires via des filtres de pré-analyse afin d'avoir des fréquences de déficit aussi proches que possible de la réalité. Cependant, nous avons pu observer *a posteriori* que des élèves ayant probablement

répondu au hasard sont encore présents (cf. tableau 4) et ont donc pu affecter la distribution des déficits des élèves. Nous avons également obtenu des déficits inattendus (répétition de réponses “trop facile” ou “trop difficile” à la difficulté perçue) : des analyses supplémentaires seraient nécessaires avant de classer les élèves dans une catégorie si une intervention automatique est ensuite déclenchée sur la base de ce classement.

Lors de la définition des déficits, nous avons délibérément choisi le libellé d'auto-évaluation pour qu'il ait la priorité sur le libellé d'auto-efficacité. Nous avons estimé que la difficulté souhaitée n'était pas pertinente si les élèves n'avaient pas une représentation correcte de la difficulté de l'exercice qu'ils venaient de résoudre. Ce choix limite notre intervention à essayer de résoudre un déficit à la fois, plutôt que de s'attaquer à deux déficits potentiels à la fois.

Une autre limite potentielle de ce travail est liée à la valeur des seuils. Dans notre modèle, nous avons identifié 5 (N_{min} , $Perf_+$, $Perf_-$, $FrDef_{min}$, $NDef_{min}$) paramètres qui correspondent à différentes valeurs seuils. Nous avons essayé différents seuils pour le marquage des élèves (fixés à 50 % et au moins deux réponses labellisées comme ayant le déficit dans l'algorithme 1) sans impact sur la fréquence relative des déficits, bien qu'elle varie en valeur absolue (résultats non affichés ici). Une adaptation précise des seuils devra garantir la sensibilité et la spécificité des déficits détectés, ainsi que le marquage des seuls déficits les plus saillants pour ne pas surcharger le tableau de bord de l'enseignant.

Bien que l'état de l'art le mentionnait, nous n'avons pas mobilisé la possibilité d'ajouter un tableau de bord. En effet, ici on s'intéresse uniquement à la mesure alors qu'un tableau de bord serait surtout adapté à la remédiation. De plus, les traces d'un élève liées à son auto-régulation sur un tableau de bord se heurtent à une forte difficulté d'interprétation, surtout chez de jeunes enfants. Nous essayons ici d'avoir une méthode la plus objective possible de mesure des déficits.

Enfin, les déficits détectés devront être confrontés aux opinions des enseignants sur les déficits de leurs élèves afin de valider *a posteriori* la manière dont nous les avons définis. Les capter en même temps aurait demandé des développements supplémentaires côté enseignant impossibles pour cette étude.

5 Conclusion et travaux futurs

Pour répondre aux QR1 et QR2, nous avons présenté une méthode de détection des déficits potentiels d'auto-efficacité et d'auto-régulation d'élèves de 5 à 7 ans en posant aux élèves deux questions sur leur difficulté perçue et la difficulté souhaitée pour les exercices suivants. En analysant et en caractérisant les schémas de réponses d'environ 12 000 élèves nord-américains et 3 000 élèves français, nous avons détecté des déficits près d'un tiers des élèves. Nous avons également proposé une définition opérationnelle des déficits permettant la labellisation.

Les travaux futurs comprennent la comparaison des données sur les déficits détectés avec la perception qu'en ont les enseignants pour leurs élèves afin de confirmer la pertinence des déficits détectés. On pourra notamment le faire en demandant à des enseignants volontaires s'ils sont d'accord avec les labellisations de sous- et sur-évaluation de leurs élèves en utilisant des échelles type Likert. On pourra également caractériser les sous-populations présentant un déficit partic-

ulier par de la fouille de séquence (*differential sequence mining*). Enfin, une fois ces déficits confirmés, une remédiation devrait être mise en œuvre, en étudiant l'évolution temporelle des déficits avec ou sans elle pour mesurer son efficacité.

Références

1. Azevedo, R., Hadwin, A.F.: Scaffolding Self-regulated Learning and Metacognition – Implications for the Design of Computer-based Scaffolds. *Instructional Science* **33**(5), 367–379 (Nov 2005)
2. Azevedo, R., Landis, R.S., Feyzi-Behnagh, R., Duffy, M., Trevors, G., Harley, J.M., Bouchet, F., Burlison, J., Taub, M., Pacampara, N., Yeasin, M., Rahman, A.K.M.M., Tanveer, M.I., Hossain, G.: The Effectiveness of Pedagogical Agents' Prompting and Feedback in Facilitating Co-adapted Learning with MetaTutor. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) *Intelligent Tutoring Systems*. pp. 212–221. LNCS, Springer, Berlin, Heidelberg (2012)
3. Bandura, A.: Self-Efficacy. In: *The Corsini Encyclopedia of Psychology*, pp. 1–3. American Cancer Society (2010)
4. Bimba, A.T., Idris, N., Al-Hunaiyyan, A., Mahmud, R.B., Shuib, N.L.B.M.: Adaptive feedback in computer-based learning environments: a review. *Adaptive Behavior* **25**(5), 217–234 (Oct 2017)
5. Bouchet, F., Harley, J.M., Azevedo, R.: Can Adaptive Pedagogical Agents' Prompting Strategies Improve Students' Learning and Self-Regulation? In: Micarelli, A., Stamper, J., Panourgia, K. (eds.) *Intelligent Tutoring Systems*. pp. 368–374. *Lecture Notes in Computer Science*, Springer International Publishing, Cham (2016)
6. Dignath, C., Buettner, G., Langfeldt, H.P.: How can primary school students learn self-regulated learning strategies most effectively? *Educational Research Review* **3**(2), 101–129 (Jan 2008)
7. Jackson, J.W.: Enhancing Self-Efficacy and Learning Performance. *The Journal of Experimental Education* **70**(3), 243–254 (Jan 2002)
8. Kim, J.A., Lorschbach, A.W.: Writing Self-Efficacy in Young Children: Issues for the Early Grades Environment. *Learning Environments Research* **8**(2), 157–175 (2005)
9. Molenaar, I., Horvers, A., Dijkstra, R., Baker, R.: Designing Dashboards to support learners' Self-Regulated Learning (Feb 2019)
10. Molenaar, I., Horvers, A., Dijkstra, R., Baker, R.S.: Personalized visualizations to promote young learners' SRL: the learning path app. In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. pp. 330–339 (2020)
11. Schmitz, B., Perels, F.: Self-monitoring of self-regulation during math homework behaviour using standardized diaries. *Metacognition and Learning* **6**(3), 255–273 (2011)
12. Schunk, D.H.: Goal and Self-Evaluative Influences During Children's Cognitive Skill Learning. *American Educational Research Journal* **33**(2), 359–382 (Jun 1996)
13. Schunk, D.H., Zimmerman, B.J.: Self-Regulation and Learning. In: *Handbook of Psychology*, Second Edition. American Cancer Society (2012)
14. Stipek, D., Recchia, S., McClintic, S.: Self-evaluation in young children. *Monographs of the Society for Research in Child Development* **57**(1), 1–98 (1992)
15. Zheng, L.: The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: a meta-analysis. *Asia Pacific Education Review* **17**(2), 187–202 (Jun 2016)
16. Zimmerman, B.J.: Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal* **45**(1), 166–183 (Mar 2008)

Analyse de remédiations proposées par des enseignants ciblant des déficits d’auto-régulation de jeunes élèves dans une application d’apprentissage de la lecture

Thomas Sergent^{1,2}, Morgane Daniel², François Bouchet¹ et Thibault Carron¹

¹ Sorbonne Université, CNRS, LIP6, F-75005 Paris, France {prénom.nom}@lip6.fr

² Lalilo, Paris, France {prénom}@lalilo.com

Résumé. La capacité à auto-réguler son apprentissage a un impact significatif sur les résultats scolaires. Nous avons interrogé 298 enseignants dont les élèves travaillaient avec une application web d’apprentissage de la lecture pour savoir comment ils aimeraient être informés d’éventuels déficits d’auto-régulation de leurs élèves et comment ils résoudreaient ces déficits afin que la remédiation des déficits dans le logiciel d’apprentissage en ligne puisse être conçue en collaboration avec eux. Les résultats montrent un fort intérêt des enseignants pour (1) remonter les déficits d’auto-régulation des élèves par le biais du tableau de bord de l’application, et (2) fournir une assistance automatisée via des encouragements et des retours formatifs, ainsi que des exercices adaptés en fonction des déficits.

Mots-clé: Co-conception · Auto-régulation · École primaire

Abstract. Self-regulated learning has a significant impact on academic performance. We surveyed 298 teachers whose students were working with a web-based reading learning application to find out how they would like to be informed of their students’ potential self-regulation deficits and how they would address them, in order to co-design these remediations with them. The results show a strong interest from teachers to (1) track students’ self-regulation deficits through the application’s dashboard, and (2) provide automated assistance through encouragement and formative feedback, as well as exercises tailored to their deficits.

Keywords: Co-design · Self-regulated learning · Primary school

1 Introduction

Adapter le retour formatif donné à un apprenant en fonction de sa réponse à un exercice permet un apprentissage plus efficace, ce que peuvent permettre certains systèmes informatiques de type tuteurs intelligents [12]. Par ailleurs, l’amélioration des compétences d’auto-régulation des enfants permet une amélioration de leurs performances scolaires, car les élèves auto-régulés savent globalement mieux comment apprendre [15]. Par conséquent, si l’on arrive à identifier

un déficit d'auto-régulation de l'apprentissage, on pourrait envisager - comme pour le retour formatif - de donner à l'apprenant un contenu différencié pour y remédier. De plus, on peut supposer qu'un enseignement précoce de ces compétences d'auto-régulation est préférable, et des programmes de formation pour des élèves de primaire ont déjà été élaborés dans ce sens [8].

Lalilo est l'une des nombreuses applications web utilisées par les enseignants en classe pour les aider à mettre en place une pédagogie différenciée. Début 2021, elle est utilisée par 40 000 classes de maternelle et élémentaire anglophones et francophones chaque semaine pour renforcer l'alphabétisation en proposant une série d'exercices adaptés au niveau des élèves, tout en offrant à l'enseignant un tableau de bord pour évaluer les activités et les progrès des élèves. Il s'agit donc d'un terrain d'essai pertinent pour évaluer d'abord et ensuite essayer de corriger les déficits d'auto-régulation des élèves. De plus nous avons précédemment montré qu'il était possible d'estimer la prévalence de déficits d'auto-évaluation et d'auto-efficacité chez les élèves utilisant Lalilo, avec au moins un tiers des élèves présentant un déficit [11]. Ce travail s'inscrit donc dans un contexte où des questions permettant d'estimer l'auto-évaluation et l'auto-efficacité, deux sous-compétences clés liées à l'auto-régulation, ont été introduites dans Lalilo. Ici, nous voulons étudier la prise de conscience par les enseignants de la prévalence de certains déficits d'auto-régulation de l'apprentissage chez leurs jeunes élèves et des approches (le cas échéant) qu'ils utilisent actuellement pour tenter de les réduire. Bien que la littérature sur l'auto-régulation soit riche en descriptions d'approches visant à en développer les compétences, il est essentiel - tant pour des raisons d'acceptabilité que d'efficacité - lorsque l'on introduit de nouvelles fonctionnalités, qu'elles soient en synergie avec les pratiques actuelles des enseignants dans leurs classes. Nous présentons donc les résultats d'une enquête menée auprès d'enseignants utilisant Lalilo dans leurs classes afin d'identifier les approches les plus prometteuses pour les aider à lutter contre certains déficits d'auto-régulation de l'apprentissage de leurs élèves. Nous pourrions ensuite utiliser ces approches suggérées pour concevoir des remédiations aux déficits d'auto-régulation ciblés.

Plus précisément, nous examinerons les questions de recherche suivantes : (QR1) Les enseignants sont-ils conscients des déficits d'auto-régulation de l'apprentissage de leurs élèves ? (QR2) Ont-ils des idées sur la manière de les résoudre ? (QR3) De quelles formes d'aides auraient-ils éventuellement besoin pour remédier à ces déficits ?

Le reste de cet article sera donc structuré comme suit. Dans la section 2, nous examinerons la manière dont les enseignants travaillent habituellement le développement des compétences d'auto-régulation de leurs élèves. La section 3 présentera les résultats de l'enquête auprès des enseignants. Nous discuterons des limites et des perspectives dans la section 4.

2 Travaux connexes

Dans le contexte de l'apprentissage sur ordinateur, l'auto-régulation peut être soutenue par différents types d'échafaudages (*scaffolding*) [1] notamment du *feedback* automatisé [4]. Savoir s'il est possible d'entraîner efficacement l'auto-régulation

chez les jeunes enfants, notamment via un outil informatique, reste une question partiellement ouverte. Dignath et al. [8] ont réalisé une méta-analyse de 30 études sur les programmes de formation à l'auto-régulation de l'apprentissage à l'école primaire, qui a montré un effet positif significatif. En outre, Zheng [14] a montré dans sa méta-analyse que les supports d'auto-régulation dans les environnements informatiques ont généralement un effet positif significatif. Enfin, Molenaar et al. [10] ont mis au point un tableau de bord pour l'apprenant qui a permis d'améliorer l'auto-régulation de l'apprentissage d'élèves de 10-11 ans.

Les enseignants semblent se sentir plus incertains dans l'enseignement de l'auto-régulation que dans la pratique générale en classe [7]. À partir d'observations vidéo en classe et d'entretiens avec des enseignants, Dignath et al. [9] ont souligné que les enseignants des écoles primaires et secondaires ont besoin d'une formation pour améliorer leur enseignement direct et indirect de l'auto-régulation de l'apprentissage. Ils pourraient notamment tirer profit d'un apprentissage de l'enseignement explicite des stratégies d'auto-régulation et de métacognition. De plus, comme nous l'avons déjà mentionné, peu de travaux ont examiné le cas de l'entraînement des capacités d'auto-régulation par le biais de logiciels ; MetaTutor étant une des exceptions mais ce logiciel est destiné à des étudiants du supérieur [2].

3 Connaissance de l'auto-régulation de l'apprentissage par les enseignants et techniques de remédiation

3.1 Méthode

Collecte des données. Nous avons envoyé une enquête à tous les enseignants utilisateurs de Lalilo d'écoles francophones (lien vers le questionnaire complet³). On a estimé, par l'introduction de deux questions posées ponctuellement aux élèves (cf. Figure 1), les éventuels déficits d'auto-évaluation et d'auto-efficacité [11]. L'estimation se fait en comparant les réponses de l'élève aux deux énoncés d'auto-régulation présentés Figure 1 avec sa performance réelle à l'exercice précédent. Pour ce faire, nous avons défini un algorithme étiquetant les réponses selon qu'elles semblent montrer un déficit ou non (par exemple la réponse d'un élève ayant 100% de réussite à l'exercice précédent et répondant "Trop difficile" à l'énoncé demandant "Quelle était la difficulté de cet exercice pour toi?" sera étiquetée "sous-évaluation"). Ensuite, en fonction des étiquettes de réponse d'un élève et de seuils déterminés (par exemple 50% de sous-évaluation quand l'élève a eu 100% de bonnes réponses à l'exercice précédent), on a étiqueté des déficits au niveau de l'élève le cas échéant. Pour mesurer le sentiment d'auto-efficacité, on utilise l'écart entre la difficulté perçue - représentation de la difficulté de l'exercice - et la difficulté souhaitée. Cela nous a permis d'extraire des statistiques sur les déficits d'auto-régulation des élèves que l'on a présenté aux enseignants pour leur demander une réaction par rapport à ces résultats en les confrontant à leur propre expérience. L'idée était aussi d'éviter une situation où des enseignants considéreraient, *a priori* à tort (à la fois d'après la littérature précédemment évoquée et ce que nous avons pu mesurer), que les problèmes

³ <https://anancus.ynh.fr/lstu/codesignautoregulationlalilo>

d'auto-régulation de leurs élèves sont marginaux, ce qui aurait pu limiter leurs réponses dans la suite du questionnaire. Nous avons rappelé aux enseignants ce



Fig. 1: Énoncé demandant la difficulté perçue (gauche) puis la difficulté voulue (droite). Ici l'élève a sélectionné "de même niveau", ce qui a affiché un bouton de confirmation dessous.

qu'est l'apprentissage auto-régulé en quelques mots et leur avons présenté les énoncés d'auto-régulation que Lalilo donne de temps en temps à leurs élèves et qui nous ont permis d'estimer la prépondérance de déficits d'auto-évaluation et d'auto-efficacité chez certains élèves. Nos objectifs étaient les suivants :

1. estimer le degré de familiarité des enseignants avec l'auto-régulation de l'apprentissage, l'alignement entre leur perception de la fréquence des déficits et la réalité, et leur éventuelle pratique actuelle pour entraîner leurs élèves à auto-réguler leur apprentissage – en lien avec la QR1,
2. comprendre comment un enseignant pourrait remédier au déficit potentiel d'un élève en général (hors Lalilo) – en lien avec la QR2,
3. savoir si les enseignants souhaitent être informés des déficits potentiels de leurs élèves en termes d'auto-régulation – en lien avec la QR3,
4. savoir si les enseignants acceptent, voire souhaitent, qu'un logiciel intervienne s'il détecte un déficit d'auto-régulation – en lien également avec la QR3.

Tableau de bord

Leçons assignées
 Leçons en cours
 Pas encore abordé
 Acquis
 En cours d'acquisition
 Non acquis
 Validée suite au positionnement

ÉLÈVE	Voyelles	Phonologie	Consonnes tenues	Consonnes non tenues
Lilo Demo	<div style="width: 100%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>	<div style="width: 20%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 10%; height: 10px; background-color: #27ae60;"></div>
Albertine	<div style="width: 80%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>	<div style="width: 0%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>
Gabriella	<div style="width: 10%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>	<div style="width: 0%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>
Kenzo	<div style="width: 0%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>	<div style="width: 0%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>
Nabil	<div style="width: 0%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>	<div style="width: 0%; height: 10px; background-color: #f4a460;"></div>	<div style="width: 0%; height: 10px; background-color: #27ae60;"></div>

Fig. 2: Tableau de bord dans Lalilo

Nous avons recueilli 298 réponses à des questions à choix unique, à choix multiple et à des questions ouvertes entre le 20 et le 23 octobre 2020. Il y avait cinq questions à choix unique, deux questions à choix multiples et huit questions ouvertes. Environ 40 000 personnes ont reçu l'enquête (l'ensemble des enseignants abonnés à la newsletter de Lalilo à cette date). On estime qu'un enseignant

pouvait répondre à l'enquête en 15 minutes environ. Plus précisément, les quatre questions ouvertes suivaient une mise en situation et visaient à comprendre la manière dont l'enseignant s'y prendrait pour résoudre le déficit d'un étudiant hypothétique. Elles ont été formulées comme suit : *Considérez le cas de Kenzo dans la classe ci-dessus [un tableau de bord hypothétique avec quelques étudiants dont Kenzo fait partie a été montré, voir Figure 2]. Il est dans la situation 3 : il a tendance à vouloir éviter les difficultés, c'est-à-dire que lorsque ses performances sont excellentes et qu'il perçoit correctement les exercices comme "faciles", il dit souvent qu'il veut des exercices "de même niveau" ou "plus faciles". Que feriez-vous pour l'aider ?* Notre raisonnement était que si nous voulons que les enseignants acceptent des remédiations d'auto-régulation pour leurs élèves dans Lalilo, il est essentiel de connaître à la fois les méthodes qu'ils emploient et jugeraient acceptables de déléguer à un logiciel pour pouvoir concevoir des retours (enseignant ou élève) qui leur soient réellement utiles dans leur pratique, et non seulement utiles d'un point de vue théorique. En l'absence de questionnaire standard à notre connaissance pour interroger les enseignants sur leurs pratiques de remédiation, nous avons optés pour des questions ouvertes. Notre travail reprend donc en partie des démarches inspirées de la co-conception, notamment le recueil de besoins des utilisateurs finaux.

Les réponses ouvertes ont été analysées manuellement, en listant les suggestions contenues dans chaque réponse et en regroupant les propositions similaires. Les suggestions dans les tableaux de résultats ont été déterminées *a posteriori*, en fonction des réponses données par les enseignants de notre échantillon. Nous avons constitué la grille de manière exploratoire en rapprochant les réponses similaires et en extrayant les suggestions principales, ce qui nous a donné les lignes du Tableau 2. Une fois la grille déterminée, les réponses ont été étiquetées, chaque réponse pouvant contenir plusieurs suggestions.

3.2 Résultats

Familiarité des enseignants avec les déficits d'apprentissage auto-régulés et pratiques de remédiation actuellement utilisées. Notre premier constat est que dans les classes ayant répondu, environ 1 enseignant sur 3 (37%) déclare travailler régulièrement sur les compétences d'auto-régulation de ses élèves. Après description des déficits potentiels, la plupart des enseignants estiment qu'au moins certains de leurs élèves présentent des déficits d'auto-régulation (voir Tableau 1 en haut), ce qui correspond à notre estimation d'un tiers [11].

Intérêt des enseignants à connaître les déficits d'apprentissage auto-régulés de leurs élèves lors de l'utilisation de Lalilo. Nous avons demandé aux enseignants s'ils souhaiteraient que Lalilo fasse état des déficits potentiels détectés en matière d'auto-régulation de l'apprentissage (tableau 1 milieu) et 54% d'entre eux se sont montrés très favorables à ce sujet (note de 5). Au contraire, 7% des enseignants se sont montrés défavorables ou très défavorables (note de 1 ou 2) à l'idée de connaître les déficits potentiels de leurs élèves. Cela souligne la nécessité de rendre cette information facultative pour l'enseignant ainsi que d'affiner la manière dont elle doit être communiquée.

Lorsqu'on leur demande explicitement de réfléchir aux inconvénients potentiels de la présence des déficits des élèves sur leur tableau de bord (voir tableau 1 en bas) par le biais d'une question à choix multiple spécifique, la principale préoccupation des enseignants (37% pour l'ensemble des enseignants et 50% des enseignants défavorables ou très défavorables) est la possibilité que les réponses des élèves soient aléatoires et donc peu fiables. Il est donc important de prendre en compte la détection de réponses aléatoires d'élèves qui ne comprendraient pas ou voudraient juste ignorer les questions de mesures d'auto-régulation, mais aussi d'expliquer aux enseignants que ces cas ont été exclus pour augmenter leur confiance dans les résultats qui leur seraient présentés. Certains enseignants (1 sur 4) craignent que l'affichage des déficits en matière d'auto-régulation dans leur tableau de bord ne le surcharge. Parmi les enseignants défavorables ou très défavorables, cette proportion monte à près de 60%. En effet, il est déjà rempli d'informations sur les progrès de leurs élèves et les informations sur l'auto-régulation ajoutent une dimension qui peut être moins exploitable que les résultats concrets des exercices. Il convient donc de veiller à ne pas surcharger le tableau de bord et à n'afficher les informations qu'à leur demande. Enfin, nous avons laissé un champ d'expression libre "Autre" dans cette question à choix multiple quand l'enseignant souhaitait lister d'autres inconvénients. Les réponses à ce champ ont été données à une fréquence trop faible pour apparaître.

Suggestions d'intervention pour résoudre les déficits d'apprentissage auto-régulés par les enseignants, et l'intérêt de les automatiser. Les suggestions des enseignants sont résumées dans le tableau 2. Certains enseignants ont fait plusieurs suggestions et figurent donc sur plusieurs lignes du tableau. Pour les quatre catégories de déficits considérées, 90 % des réponses peuvent facilement être associées et étiquetées avec un ensemble relativement restreint de cinq à huit suggestions. Dans chaque sous-tableau du tableau 2 apparaît la ligne "Autres". Cette ligne regroupe l'ensemble des suggestions qui ont été données dans moins de 1% des réponses des enseignants. En creusant dans les suggestions, nous avons remarqué que certaines d'entre elles apparaissaient dans plusieurs catégories de déficits :

- déficits lorsque les performances sont excellentes (sous-évaluation ou éviter les difficultés)
 - donner des exercices plus difficiles
 - expliquer aux élèves que l'on progresse par la confrontation avec la difficulté
 - encourager les étudiants à viser plus haut
 - encourager les étudiants à jouer un rôle de tuteur
- déficits lorsque la performance est mauvaise (sur-évaluation ou cherchant la difficulté)
 - donner les mêmes exercices ou des exercices plus faciles
- déficits d'auto-évaluation (sur-évaluation ou sous-évaluation)
 - demander aux élèves d'expliquer leur processus d'auto-évaluation
- déficit de l'auto-évaluation ou faible performance (sur-évaluation, sous-évaluation ou cherchant la difficulté)

Selon vous, quelle proportion des élèves de votre classe présente au moins un des quatre déficits d'auto-régulation de l'apprentissage listés ci-dessus ? [sur/sous-évaluation, évitant/cherchant la difficulté]	Pourcentage
Je ne sais pas	17%
Aucun	1%
Quelques-uns	27%
La moitié	25%
La plupart	28%
Tous	2%

Enfinement, souhaiteriez-vous savoir qu'un élève a peut-être un de ces quatre déficits d'auto-régulation ?	Pourcentage
1 (pas du tout)	3%
2	4%
3	13%
4	26%
5 (tout à fait)	54%

Quels seraient les inconvénients à remonter les déficits des élèves dans le tableau de bord ?	Pourcentage
<i>Aucun</i>	42%
Les réponses des élèves à ce genre de questions sont souvent aléatoires	37%
Cela risquerait de surcharger le tableau de bord	24%
Je ne préfère pas le savoir car cela pourrait biaiser mon comportement envers certains élèves	5%
Je ne suis pas d'accord avec la caractérisation des déficits	2%
<i>Autre</i>	10%

Tableau 1: Perception des enseignants des déficits d'auto-régulation dans leur classe (sur 298 enseignants)

- demander aux élèves de compter ou de visualiser les bonnes et les mauvaises réponses

Les autres suggestions étaient spécifiques au déficit visé. En particulier, pour la sur-évaluation des élèves, plusieurs enseignants ont suggéré de clarifier la signification des énoncés d'auto-régulation et la définition de "trop facile" et "trop difficile". En effet, certains enseignants ont fait remarquer que les élèves peuvent confondre "trop facile" et "vite fait", ce qui pourrait expliquer leur sur-évaluation, car les exercices sont conçus pour être rapides afin de maintenir l'intérêt des élèves. Enfin, plusieurs enseignants ont souligné que les élèves peuvent s'attendre à ce que "trop facile" soit la bonne réponse et ont donc suggéré d'expliquer aux élèves qu'il est normal d'avoir tort. Nous avons également classé les suggestions faites par les enseignants en trois catégories :

- ceux qui ont suggéré de faire faire quelque chose à l'élève (suggérer des activités et/ou adapter le choix de l'exercice non seulement à la performance de l'élève)

Suggestions des enseignants pour les élèves en sur-évaluation		P. (N=237)
Refaire l'exercice		37%
Demander de compter/visualiser les bonnes/mauvaises réponses		27%
Demander aux élèves d'expliquer leur processus d'auto-évaluation		18%
Expliquer les questions et les définitions relatives à l'auto-régulation		13%
Donner des exercices plus faciles		11%
Expliquer aux élèves qu'il est normal de se tromper, que "trop facile" n'est pas la réponse attendue		8%
Souligner le travail accompli		5%
Rappeler aux élèves qu'ils doivent essayer de répondre correctement et non au plus grand nombre de questions possible		3%
<i>Autres</i>		11%
Suggestions des enseignants pour les élèves en sous-évaluation		P. (N=240)
Leur donner confiance, souligner le travail accompli		42%
Demander de compter/visualiser les bonnes/mauvaises réponses		32%
Demander aux élèves d'expliquer leur processus d'auto-évaluation		19%
Donner des exercices plus difficiles		10%
Expliquer aux élèves que l'on progresse en se confrontant à la difficulté		6%
Encourager les élèves à viser plus haut		6%
Encourager les élèves à jouer un rôle de tutorat		4%
<i>L'exercice était difficile, même si l'élève l'a réussi</i>		3%
<i>Ne sait pas</i>		2%
<i>Autres</i>		7%
Suggestions des enseignants pour les élèves évitant la difficulté		P. (N=228)
Encourager les élèves à viser plus haut		28%
Donner des exercices plus difficiles		23%
Expliquer aux élèves que l'on progresse en se confrontant la difficulté		18%
Introduire explicitement certains exercices comme un défi		9%
Demander à l'élève de choisir entre des exercices faciles ou difficiles en sachant qu'il y a un bonus pour les exercices difficiles		9%
Demander de compter/visualiser les bonnes/mauvaises réponses		7%
Encourager les élèves à jouer un rôle de tutorat		2%
<i>Ne sait pas</i>		2%
<i>Autres</i>		10%
Suggestions des enseignants pour les élèves cherchant la difficulté		P. (N=217)
Expliquer qu'il faut d'abord réussir des exercices plus faciles avant de vouloir des exercices plus difficiles		33%
Donner des exercices de même niveau ou plus faciles		26%
Demander de compter/visualiser les bonnes/mauvaises réponses		18%
Mettre l'accent sur la volonté des étudiants de viser plus haut		6%
Souligner le travail accompli		5%
<i>Ne sait pas</i>		4%
<i>Autres</i>		12%

Tableau 2: Ce que les enseignants suggèrent de faire pour un élève présentant un déficit donné (N = 298 enseignants). Certaines réponses comportaient plusieurs suggestions. Le nombre de répondants est indiqué entre parenthèses.

- ceux qui ont suggéré d'encourager l'élève (motivation ou retour d'information émotionnel)
- ceux qui ont suggéré d'expliquer quelque chose à l'élève (retour formatif)

Le tableau 3 résume les pourcentages agrégés par déficit. On peut remarquer que la principale suggestion est toujours de faire faire quelque chose à l'élève, plutôt que d'être passif.

Déficit	Faire faire quelque chose à l'élève	Encourager	Expliquer
Sous-évaluation	100	8	32
Sur-évaluation	64	54	8
Évitant la difficulté	46	29	17
Cherchant la difficulté	51	18	36

Tableau 3: Pourcentages cumulés de suggestions d'un type donné par déficit

Enfin, la dernière question de l'enquête était : "Quelle que soit l'information rapportée dans le tableau de bord, accepteriez-vous que Lalilo intervienne automatiquement auprès des élèves pour les aider à corriger leurs déficits d'auto-régulation ?" Les réponses (93 % disent oui) soutiennent l'idée que la plupart des enseignants accepteraient l'intervention d'un logiciel pour tenter de remédier aux déficits d'auto-régulation. Ceux qui ne sont pas d'accord déclarent que les déficits d'auto-régulation ne peuvent pas ou ne doivent pas être corrigés par un logiciel mais plutôt par une personne. La distinction traditionnelle faite entre les communautés de *Learning Analytics* et de *Educational Data Mining* [3] montre que bien qu'une part des enseignants sont attachés à garder le contrôle, notamment grâce à un tableau de bord. Néanmoins dans notre contexte particulier de correction des déficits d'auto-régulation, ils semblent majoritairement disposés à déléguer une certaine prise de décision automatisée au logiciel utilisé par leurs élèves, même si ce chiffre est encore une fois à relativiser puisque les répondants sont un public particulier, usager d'une telle solution et suffisamment enthousiastes pour prendre le temps de répondre à ce questionnaire.

4 Discussion, limites et perspectives

Dans l'ensemble, les enseignants répondants sont déjà assez familiers avec l'auto-régulation et les déficits qui y sont liés : ils ont une assez bonne vue d'ensemble du problème, sont intéressés par une aide à cet égard et proposent des stratégies qui sont largement conformes à la littérature sur le sujet. Nous allons donc maintenant voir comment mettre en œuvre certaines de ces stratégies.

La plupart des encouragements (valoriser le travail effectué, encourager l'étudiant à viser plus haut, valoriser la volonté de viser plus haut) sont généralement déjà traités par la ludification de l'interface élève avec des badges, des diplômes ou des récompenses. Ce sont des recommandations applicables à tout EIAH avec des éléments de ludification. Les explications - qui peuvent également être considérées comme un retour formatif - seraient probablement traitées par le biais d'une instruction orale ou d'un exemple élaboré décrivant le comportement attendu, ce qui est applicable pour des logiciels posant des questions d'auto-régulation de la même manière que fait actuellement dans Lalilo.

Les réponses aux énoncés d'auto-régulation peuvent également devenir une entrée pour l'algorithme d'apprentissage adaptatif afin de prendre en compte les suggestions des enseignants concernant la difficulté de l'exercice à donner par la suite, ce qui est applicable à un logiciel qui est en mesure de caractériser le profil d'auto-régulation des élèves et utilisant un algorithme d'apprentissage adaptatif modulable en fonction d'autres paramètres que les seules réponses aux exercices. Pour les seuls étudiants qui évitent la difficulté, les enseignants ont donné des suggestions sur la façon dont les exercices devraient être présentés à l'étudiant : faire choisir à l'étudiant la difficulté de l'exercice avec un bonus potentiel pour le choix difficile, ou introduire explicitement certains exercices comme un défi. Le fait de demander aux élèves de compter ou de visualiser les

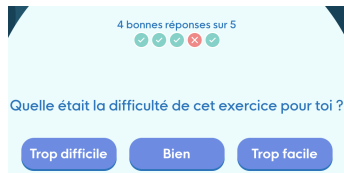


Fig. 3: Jauge permettant la visualisation des réponses.

bonnes et les mauvaises réponses plaide en faveur d'un ajout d'information sous forme de "mini-tableau de bord" de feedback immédiat pour les élèves (un exemple est affiché à la Figure 3), une approche couramment utilisée avec les modèles d'apprentissage ouvert, bien qu'elle soit généralement réservée aux enfants plus âgés [6]. Enfin, certaines suggestions ne semblent pas directement transposables à un logiciel existant, comme par exemple encourager les étudiants à jouer un rôle de tuteur [13] (bien que des logiciels dédiés comme Betty's Brain aient exploré cette approche [5]) et surtout faire en sorte que les étudiants expliquent leur processus d'auto-évaluation dans une discussion avec un adulte, ce qui est l'une des actions les plus recommandées pour les étudiants souffrant d'un déficit d'auto-évaluation.

Notre objectif était d'évaluer dans quelle mesure les déficits d'auto-régulation détectés étaient pertinents pour les enseignants et comment ils allaient les aborder pour leur proposer des solutions non seulement issues la littérature scientifique, mais correspondant à ce qu'ils font quotidiennement dans leur classe. Pour qu'une solution soit utilisée par les enseignants, il est en effet essentiel qu'elle corresponde à leurs besoins et puisse être associée à leurs pratiques.

Les résultats montrent également que les enseignants semblent très ouverts à l'idée d'avoir une intervention intégrée dans le logiciel, mais il ne faut pas surgénéraliser les résultats ici, en gardant à l'esprit qu'il y a un biais d'échantillonnage. En effet, nous pouvons supposer que les enseignants qui utilisent Lalilo sont plus ouverts à l'idée de recourir à la technologie pour une partie de leurs besoins pédagogiques, et il est probable que si nous interrogeons l'ensemble de la communauté enseignante, les résultats seraient différents.

Nous avons choisi de poser aux enseignants des questions ouvertes et non à choix multiples sur la manière dont ils aborderaient un élève présentant un déficit. Cette méthode a l'avantage de ne pas limiter l'imagination des enseignants ni de les influencer, mais l'analyse des réponses peut être moins précise. Nous avons décidé de ne pas créer de catégories de suggestions *a priori* afin de permettre une catégorisation aussi précise que possible. Dans un processus de co-conception, une façon de confirmer les catégories détectées serait de soumettre à nouveau aux enseignants, lors d'une prochaine enquête, la liste des catégories de remédiation que nous avons extraite de la première enquête et de leur demander leur accord ou désaccord avec chacune d'elles, ou de les classer.

5 Conclusion et travaux futurs

Nous avons recueilli les suggestions d'environ 300 enseignants sur la manière dont ils aborderaient les déficits d'auto-régulation. Les réponses collectées sont en phase avec ce que nous avons pu mesurer dans Lalilo, répondant ainsi positivement à la QR1 sur la conscience des déficits d'auto-régulation par les enseignants (du moins parmi les répondants). Pour répondre à la QR2, nous avons ensuite mis en évidence les similitudes et les différences dans la remédiation des déficits : une combinaison de mesures à prendre par l'élève, d'explications et d'encouragements, dont la quantité varie en fonction du déficit. Nous constatons que la plupart des suggestions semblent réalisables par un logiciel d'apprentissage en ligne et nous avons proposé des transpositions possibles dans Lalilo, répondant ainsi positivement à la QR3. La plupart des enseignants accueillent favorablement l'idée d'obtenir des informations sur les capacités d'auto-régulation de leurs élèves et de laisser un logiciel combler ces déficits. Certains enseignants craignent une surcharge du tableau de bord avec des informations pas toujours exploitables, et quelques enseignants préféreraient *a priori* que les déficits d'auto-régulation soient traités uniquement par une personne et non par un logiciel.

Cet article s'inscrit dans un contexte plus général dont le but est double : (a) améliorer certaines compétences d'auto-régulation des élèves, (b) utiliser les réponses aux énoncés d'auto-régulation pour informer la recommandation des exercices et la validation des compétences de lecture. Pour tendre vers l'amélioration de certaines compétences d'auto-régulation une fois ces déficits confirmés, la remédiation devrait être mise en œuvre en utilisant les suggestions des enseignants transférables à un logiciel et en étudiant l'évolution temporelle des déficits avec ou sans eux. Devant la diversité des réponses et l'incertitude d'une petite partie des enseignants quant à la pertinence des interventions, il semble dans tous les cas utiles de laisser à l'enseignant la possibilité d'activer tout ou partie des remédiations implémentées. Cela permettra par ailleurs, en analysant les données, d'estimer l'impact de chacune de ces interventions.

Références

1. Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., Landis, R.: Using Trace Data to Examine the Complex Roles of Cognitive, Metacognitive, and Emotional Self-Regulatory Processes During Learning with Multi-agent Systems. In: Azevedo, R., Alevan, V. (eds.) International Handbook

- of Metacognition and Learning Technologies, vol. 28, pp. 427–449. Springer New York, New York, NY (2013)
2. Azevedo, R., Landis, R.S., Feyzi-Behnagh, R., Duffy, M., Trevors, G., Harley, J.M., Bouchet, F., Burlison, J., Taub, M., Pacampara, N., Yeasin, M., Rahman, A.K.M.M., Tanveer, M.I., Hossain, G.: The Effectiveness of Pedagogical Agents' Prompting and Feedback in Facilitating Co-adapted Learning with MetaTutor. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) Intelligent Tutoring Systems. pp. 212–221. LNCS, Springer, Berlin, Heidelberg (2012)
 3. Baker, R.S., Inventado, P.S.: Educational data mining and learning analytics. In: Learning analytics, pp. 61–75. Springer (2014)
 4. Bimba, A.T., Idris, N., Al-Hunaiyyan, A., Mahmud, R.B., Shuib, N.L.B.M.: Adaptive feedback in computer-based learning environments: a review. *Adaptive Behavior* **25**(5), 217–234 (Oct 2017)
 5. Biswas, G., Jeong, H., Kinnebrew, J.S., Sulcer, B., Roscoe, R.: Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning* **05**(02), 123–152 (Jul 2010)
 6. Bodily, R., Kay, J., Alevan, V., Jivet, I., Davis, D., Xhakaj, F., Verbert, K.: Open learner models and learning analytics dashboards: a systematic review. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. pp. 41–50. LAK '18, Association for Computing Machinery, New York, NY, USA (Mar 2018)
 7. De Smul, M., Heirweg, S., Van Keer, H., Devos, G., Vandeveldt, S.: How competent do teachers feel instructing self-regulated learning strategies? Development and validation of the teacher self-efficacy scale to implement self-regulated learning. *Teaching and Teacher Education* **71**, 214–225 (Apr 2018)
 8. Dignath, C., Buettner, G., Langfeldt, H.P.: How can primary school students learn self-regulated learning strategies most effectively? *Educational Research Review* **3**(2), 101–129 (Jan 2008)
 9. Dignath, C., Büttner, G.: Teachers' direct and indirect promotion of self-regulated learning in primary and secondary school mathematics classes – insights from video-based classroom observations and teacher interviews. *Metacognition and Learning* **13**(2), 127–157 (Aug 2018)
 10. Molenaar, I., Horvers, A., Dijkstra, R., Baker, R.S.: Personalized visualizations to promote young learners' SRL: the learning path app. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge. pp. 330–339 (2020)
 11. Sergent, T., Bouchet, F., Daniel, M., Carron, T.: Détection de déficits d'auto-évaluation et d'auto-efficacité dans un logiciel enseignant la lecture et l'écriture. In: Environnements Informatiques pour l'Apprentissage Humain (Jun 2021)
 12. Shute, V.J.: Focus on Formative Feedback. *Review of Educational Research* **78**(1), 153–189 (Mar 2008)
 13. Topping, K.: Peer Tutoring and Paired Reading: Combining Two Powerful Techniques. *The Reading Teacher* **42**(7), 488–494 (1989), publisher: [Wiley, International Reading Association]
 14. Zheng, L.: The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: a meta-analysis. *Asia Pacific Education Review* **17**(2), 187–202 (Jun 2016)
 15. Zimmerman, B.J.: Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal* **45**(1), 166–183 (Mar 2008)

Articles courts

Fondements péda-go-didactiques d'un environnement informatique d'apprentissage destiné aux jeunes apprenants de l'école primaire

Lionel Alvarez¹, Aous Karoui², Mario Ramalho³, Thierry Geoffre⁴

¹ Haute École pédagogique Fribourg et Université de Fribourg, 1700 Fribourg, Suisse
Lionel.alvarez@unifr.ch [0000-0003-4708-251X]

² Université de Grenoble Alpes, 38031 Grenoble, France
aous.karoui@univ-grenoble-alpes.fr

³ École des Métiers, 1700 Fribourg, Suisse
Mario.Ramalho@eduf.fr

⁴ Haute École pédagogique Fribourg, 1700 Fribourg, Suisse
Thierry.Geoffre@eduf.fr

Résumé. Lorsqu'ils concernent l'apprentissage des jeunes élèves de la scolarité obligatoire, les Environnements Informatiques d'Apprentissage Humain (EIAH) méritent une attention particulière dans leur conception. Sortir des tâches pour parler de jeu, varier les niveaux taxonomiques, anticiper les (re)médiations pour éviter le décrochage, alterner activités numériques et analogiques, ou encore permettre la création de parcours adaptables, sont quelques exemples d'éléments péda-go-didactiques à anticiper afin de maintenir l'engagement et faciliter l'appren-tissage. Cet article expose et justifie les choix interdisciplinaires inscrits dans un Environnement Informatique d'Apprentissage construit pour des Enfants (EIAE) de la scolarité obligatoire, nommé *GamesHUB*.

Mots-clés : EIAH pour enfant, EIAE, fondements didactiques, pédagogie du jeu, parcours adaptables.

Abstract. When it comes to the learning of young students in compulsory school-ing, technology-enhanced learning environments deserve special attention in their design. Getting out of tasks to talk about games, varying taxonomic levels, anticipating (re)mediation to avoid dropping out, alternating digital and analog activity, or allowing the creation of adaptive learning paths, are some examples of pedagogical/didactic elements to be anticipated in order to maintain engage-ment and facilitate learning. This article exposes and justifies the interdiscipli-nary choices made in a TELE – called *GamesHUB* – built for children of com-pulsory schooling.

Keywords: TEL environment, TELE for children, didactic foundations, play-based learning, tailored learning paths.

1 Penser l'EIAH pour les jeunes apprenants

Lorsqu'il s'agit de penser l'apprentissage des enfants, quelques particularités s'ajoutent dans l'ingénierie pédago-didactique des environnements numériques pour l'apprentissage humain (EIAH). Ces particularités, non-exhaustives dans cet article, peuvent être classées en cinq points : la cognition des jeunes, l'interaction enfant-machine, la manipulation et les relations sociales, le ludique et les enjeux de l'école publique.

1.1 Spécificités de l'enseignement-apprentissage dans la scolarité obligatoire

Premièrement, il a été documenté qu'enfants et adultes ne sont pas égaux face au défi d'apprendre. Par exemple, il semble que les capacités à traiter l'information multimodale – propres aux EIAH – soient moins développées chez les jeunes [1] alors que leurs capacités à tirer profit des implicites seraient plus grandes [2].

Deuxièmement, l'interaction entre l'enfant et l'EIAH gagne à être prévue de manière intuitive, sans complexité de gestes comme un clic droit ou des sous-menus. Il semble que les tablettes soient plus généralement adoptées par les développeurs, les enseignants et les parents lorsqu'il s'agit de donner un appareil numérique aux jeunes [3].

Troisièmement, l'alternance numérique-analogique dans les EIAH à destination des enfants semble nécessaire, notamment suite aux résultats des recherches cognitives concernant l'impact des écrans dans leur compréhension large [4]. De plus, il a été proposé à de multiples reprises que l'enfant passe par la manipulation et les interactions sociales physiques pour apprendre (rôle du geste pour apprendre [5], rôle des pairs pour l'apprentissage vicariant [6], rôle du corps [7]...). Les EIAH devraient donc être conçus avec une attention particulière sur les alternances : numérique-analogique, individuelle-interactionnelle, manipulation-conceptualisation. On peut aussi rappeler la nécessité de varier les niveaux taxonomiques des activités [8].

Quatrièmement, le jeu semble être une porte d'entrée importante pour permettre l'apprentissage des jeunes [9]. Apprendre par le jeu a montré des effets sur les performances [10]. Mais la manière d'exploiter les mécaniques de jeu est déterminante [11].

Finalement, l'inscription d'un EIAH dans la scolarité obligatoire exige la création de liens explicites avec le curriculum officiel. Celui-ci liste les contenus disciplinaires visés, mais aussi des capacités transversales, et cette double finalité invite à repenser le défi didactique : l'EIAH doit non seulement permettre l'apprentissage des objets disciplinaires, mais aussi faciliter la collaboration, la créativité, la pensée critique, etc.

1.2 État des lieux de la littérature traitant des EIAE

Les opportunités d'apprentissage des enfants via des environnements en ligne semblent être en perpétuelle augmentation [12]. En général, la littérature traitant des environnements numériques d'apprentissage pour les enfants (EIAE) semble mettre l'accent sur les bénéfices [13]. De nombreux articles brossent l'importance du ludique [14] ou des mécanismes de gamification [15]. D'autres traitent de l'autonomie créée, du plaisir d'usage et de la réduction de la frustration [16]. Ainsi, les EIAE se popularisent et les objets étudiés sont variés.

Les EIAE sont envisagés comme plus prometteurs lorsqu'ils sont co-conçus par des praticiens et chercheurs [17]. Cette place laissée à l'enseignant permet de concevoir le retour d'information via notamment des analytiques de l'apprentissage dont l'intérêt a

été démontré pour les EIAE [18]. Ainsi, il semble qu'un EIAE gagnerait à permettre à la fois l'activité autonome de l'élève et son activité hétérodéterminée par l'enseignant.

Si la plupart des études identifiées s'intéressent à l'adoption de ces environnements par les enseignants, quelques articles explicitent l'adoption par des enfants. L'approche par expérience utilisateur semble bien adaptée à l'évaluation de l'adoption des EIAE [19]. Mais encore, la conception universelle de l'apprentissage (CUA) [20] permet de donner un cadre de conception des EIAE et tendre vers des environnements adaptés aux besoins de chacun grâce à une variation des médiations envisagées dès le départ.

Des dispositifs ont été proposés pour promouvoir un usage adapté des EIAE avec de jeunes apprenants. L'utilisation de cartouches – boîtiers physiques représentant des informations – a par exemple été proposée [21] ; ou l'ajout d'un jouet sérieux afin de permettre la manipulation [22]. Ces exemples illustrent le besoin de prévoir l'alternance entre activités analogiques et numériques dans les EIAE.

C'est sur ces bases qu'un projet EIAE a émergé en Suisse dans une collaboration entre le CRE/ATE de la Haute École pédagogique Fribourg et la section informatique de l'École des Métiers de Fribourg (Suisse) pour la création de GamesHUB [23].

2 Éléments constitutifs de l'EIAE développé : GamesHUB

GamesHUB est construit sur deux niveaux : un *framework* général et des jeux développés individuellement sur la base de fondements pédago-didactiques.

Le *framework* est contraignant pour le développement des jeux, afin de clarifier et d'unifier l'interaction élève-machine et d'éviter les éventuels écueils induits par une trop grande multimodalité de l'information. Chaque jeu doit avoir une zone de consignes, une zone de boutons et une zone interactive singulière. Le *framework* intègre *open dylsexic* et la synthèse vocale pour la lecture des consignes, ainsi qu'une variation de thèmes graphiques pour correspondre à l'attrait et aux besoins des apprenants, en accord avec les perspectives CUA. Les aspects techniques de GamesHUB sont présentés dans un autre article d'EIAH2021 intitulé « Parcours adaptatifs au sein de la plateforme européenne d'apprentissage personnalisé des langues ».

2.1 Fondements didactiques choisis : référentiels de compétences et d'erreurs

Les jeux sont développés singulièrement. Ils sont conçus pour mobiliser des compétences, capacités et connaissances qui correspondent à des attentes des programmes scolaires. La plateforme GamesHub est donc liée au référentiel de compétences officiel, le plan d'études romand (PER), via l'utilisation de l'*application programming interface* (API) du PER.

Un deuxième référentiel de compétences organise les jeux et oriente leur développement. Il a été construit dans le cadre d'un projet européen de plateforme d'apprentissage personnalisé des langues¹, à partir du référentiel COMPER². Dans notre cas, les objectifs ciblés sont les processus d'intégration [24], mobilisés dans la compétence de lecture-compréhension, pour des élèves de 8 à 10 ans de l'école ordinaire en français

¹ <https://blog.hepfr.ch/create/peapl/>

² Projet ANR-18-CE38-0012, accessible à <https://comper.fr/accueil>.

langue de scolarisation. Ce référentiel³ permet de représenter l'arborescence de la compétence de lecture-compréhension déclinée en *knowledges* et *skills* et leurs liens.

Chaque *skill* du référentiel est formalisé à partir d'une liste fermée de huit verbes d'action liés aux organisations praxéologiques linguistiques⁴ [25]. Ce choix permet de les rapprocher des techniques mobilisées par des types de tâches de lecture-compréhension dans la perspective praxéologique [26]. Il y a ici un double objectif : à court terme, favoriser l'association des ressources pédagogiques disponibles à ces types de tâches ; à moyen terme, favoriser la construction d'un référentiel d'erreurs.

Ainsi, avec ces fondements didactiques, la plateforme *GamesHUB* s'assure d'être à la fois exploitable dans le contexte institutionnel scolaire et construite sur une analyse détaillée de la langue de scolarisation et de son apprentissage.

2.2 Fondements pédagogiques choisis : jeux libres aux niveaux taxonomiques variés et parcours adaptables

L'utilisation première de *GamesHUB* invite l'apprenant à jouer sur des jeux conçus avec quatre niveaux taxonomiques préétablis. Ces modes de jeux sont : Explorer (découvrir et s'approprier les savoirs en jeu) ; S'entraîner (« driller » et obtenir des rétroactions) ; S'évaluer (laisser les traces de ses compétences à son enseignant) ; et Créer (concevoir un niveau de difficulté du jeu à soumettre à des camarades).

Ainsi, dès le *framework*, les jeux sont contraints de proposer une variété d'utilisations, invitant non seulement à apprendre, s'entraîner ou s'évaluer, mais aussi à créer des exercices pour d'autres utilisateurs, dans la classe ou non. L'ambition est de proposer à chaque enfant d'être d'abord utilisateur puis concepteur.

Dans le jeu « Par ici ou par-là » par exemple, explorer permet à l'apprenant de s'approprier les contenus (indications de déplacements sur une carte). Entraîner invite l'apprenant à s'exercer avec une diversité de difficultés et à recevoir des feedbacks détaillés. S'évaluer est similaire à Entraîner, mais les données sont captées et accessibles à l'enseignant. Finalement, Créer invite l'apprenant à créer une carte, un trajet et produire le texte associé.

2.3 Parcours adaptables conçus par l'enseignant

GamesHUB a été élaborée pour favoriser une utilisation autonome par les apprenants. L'importance des enseignants dans les EIAE (telle que rapportée dans la section 1.2) a appelé la nécessité d'une autre manière d'utiliser cette plateforme : la création de parcours adaptables via un portail [enseignant]. Ceci a été corroboré par les premiers tests⁵ de *GamesHUB* : le portail [élèves] génère une large satisfaction et permet une autonomie laissée aux apprenants, mais une prise en main des choix laissés aux élèves par les enseignants devrait être possible. Ainsi, la création de parcours à assigner à la classe, ou spécifiquement à certains apprenants, a donc été ajoutée.

Avec le portail [enseignant] de *GamesHUB*, les parcours d'apprentissage articulent différents jeux, modes de jeu et niveaux de jeu pour définir une progression et permettre

³ Publié sur <https://zenodo.org/deposit/4462850>

⁴ <https://zenodo.org/deposit/4001381>

⁵ Les tests ont été effectués auprès de 46 étudiants à l'enseignement : mise en pratique, puis *focus group* (ergonomie et intérêt pédagogique/didactique de *GamesHUB*) en 2020.

à l'élève de découvrir, s'approprier et mobiliser la compétence cible. Les enseignants peuvent assigner un parcours existant à des élèves ou à la classe, modifier un parcours existant en fonction des besoins de leurs élèves ou créer un nouveau parcours. Chaque parcours est lié à un objectif du PER et à des objectifs spécifiques du référentiel.

Les enseignants sont aussi invités à intégrer des « activités débranchées », pour lesquelles des consignes précises seront indiquées, afin de favoriser l'alternance numérique-analogique valorisée plus haut. Pour chaque jeu, Explorer et Créer sont activables pour varier les niveaux taxonomiques. Le parcours est alors conçu en étapes obligatoires et en étapes facultatives (illustration disponible dans l'autre article EIAH2021).

3 Conclusion

Ce parcours de la littérature concernant les EIAH à destination des apprenants plus jeunes permet d'identifier les thématiques importées des EIAH pour adultes, ainsi que les thématiques spécifiques. Les développements de *GamesHUB*, un environnement informatique d'apprentissage explicitement adressé à un public d'apprenants de l'école obligatoire, sont ainsi fondés sur des savoirs d'expérience, des savoirs académiques, ainsi que des compétences interdisciplinaires (pédagogie/didactique et informatique).

Les contraintes induites par le *framework* général ont déjà été éprouvées auprès d'étudiants à l'enseignement. Elles ont généré de la satisfaction et ont rappelé la nécessité de créer un portail pour les enseignants, étroitement lié au portail pour élèves déjà présent. *GamesHUB* est désormais intégré au projet PEAPL, via notamment l'ajout d'un référentiel de compétences créé à l'aide des organisations praxéologiques linguistiques. La suite du projet *GamesHUB* est alors de passer des parcours adaptables (générés par les enseignants) à des parcours adaptatifs (générés informatiquement grâce aux référentiels). Il s'agira d'éprouver systématiquement les propositions faites dans les classes pour évaluer si l'EIAE créé tend vers une CUA comme souhaité.

References

1. Lavi-Rotbain, O., Arnon, I.: Developmental Differences Between Children and Adults in the Use of Visual Cues for Segmentation. *Cogn. Sci.* 42, 606–620 (2018). <https://doi.org/10.1111/cogs.12528>.
2. Smalle, E.H.M., Muylle, M., Szmalec, A., Duyck, W.: The different time course of phonotactic constraint learning in children and adults: Evidence from speech errors. *J. Exp. Psychol. Learn. Mem. Cogn.* 43, 1821–1827 (2017). <https://doi.org/10.1037/xlm0000405>.
3. Major, L., Haßler, B., Hennessy, S.: Tablet Use in Schools: Impact, Affordances and Considerations. In: Marcus-Quinn, A. and Hourigan, T. (eds.) *Handbook on Digital Learning for K-12 Schools*. pp. 115–128. Springer, Cham, Switzerland (2017). https://doi.org/10.1007/978-3-319-33808-8_8.
4. Firth, J., Torous, J., Stubbs, B., Firth, J.A., Steiner, G.Z., Smith, L., Alvarez-Jimenez, M., Gleeson, J., Vancampfort, D., Armitage, C.J., Sarris, J.: The “online brain”: how the Internet may be changing our cognition. *World Psychiatry*. 18, 119–129 (2019). <https://doi.org/10.1002/wps.20617>.
5. Barrouillet, P.: Theories of cognitive development: From Piaget to today. *Dev. Rev.* 38, 1–12 (2015). <https://doi.org/10.1016/j.dr.2015.07.004>.
6. Carré, P.: Bandura : une psychologie pour le XXIe siècle? *Savoirs. Hors série*, 9–50 (2004). <https://doi.org/10.3917/savo.hs01.0009>.
7. Fugate, J.M.B., Macrine, S.L., Cipriano, C.: The role of embodied cognition for

- transforming learning. *Int. J. Sch. Educ. Psychol.* (2018). <https://doi.org/10.1080/21683603.2018.1443856>.
8. Anderson, L.W.: *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's*. Pearson (2013).
 9. Clerc-georgy, A., Maire, B.: *Penser la didactique des apprentissages fondamentaux en articulant jeu et curriculum*. *Forumlecture.ch*, 1, 1–17 (2020).
 10. Karakoç, B., Eryılmaz, K., Turan Özpolat, E., Yıldırım, İ.: *The Effect of Game-Based Learning on Student Achievement: A Meta-Analysis Study*. *Technol. Knowl. Learn.* (2020). <https://doi.org/10.1007/s10758-020-09471-5>.
 11. Ke, F.: *A Qualitative Meta-Analysis of Computer Games as Learning Tools*. *Gaming and Simulations*. 1619–1665 (2011). <https://doi.org/10.4018/978-1-60960-195-9.ch701>.
 12. Arnesen, K.T., Hveem, J., Short, C.R., West, R.E., Barbour, M.K.: *K-12 online learning journal articles: trends from two decades of scholarship*. *Distance Educ.* 40, 32–53 (2019).
 13. Cavanaugh, C.S., Barbour, M.K., Clark, T.: *Research and Practice in K-12 Online Learning: A Review of Open Access Literature*. *Int. Rev. Res. Open Distance Learn.* 10, (2009). <https://doi.org/10.32388/a1k3s4>.
 14. Szilas, N., Scherrer, F.: *Conception d'un jeu vidéo intégré pour l'apprentissage de la conjugaison des verbes français*. In: Guin, N., De Lièvre, B., Trestini, M., and Coulibaly, B. (eds.) *8ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain, EIAH 2017*. pp. 1–4. , Paris, France (2017).
 15. Nand, K., Baghaei, N., Casey, J., Barmada, B., Mehdi pour, F., Liang, H.-N.: *Engaging children with educational content via Gamification*. *Smart Learn. Environ.* 6, (2019).
 16. Jobert, M., Michel, G.: *Démonstration: autonomie et plaisir d'usage d'un EIAH*. In: *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine on - IHM '08*. p. 217. ACM Press, New York, New York, USA (2008). <https://doi.org/10.1145/1512714.1512759>.
 17. Capachi, F.M., Pierard, I., Moreaux, M.: *Intégration des EIAH en cyberclasse primaire*. In: Choquet, C., Dessus, P., Lefevre, M., Broisin, J., Catteau, O., and Vidal, P. (eds.) *6e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain*. Toulouse, France (2013).
 18. Baker, R.S., Koedinger, K.R.: *Towards Demonstrating the Value of Learning Analytics for K-12 Education*. In: Niemi, D., Pea, R.D., Saxberg, B., and Clark, R.E. (eds.) *Learning Analytics in Education*. pp. 49–62. Information Age Publishing, Charlotte, NC (2018).
 19. Michel, C., Sandoz-Guermond, F., Serna, A.: *Revue de littérature sur l'évaluation de l'usage de dispositifs mobiles et tactiles ludo-éducatifs pour les jeunes enfants*. *Atelier IHM avancées pour l'apprentissage*, Conférence EIAH. 11–18 (2011).
 20. Al-Azawei, A., Parslow, P., Lundqvist, K.: *The effect of Universal Design for Learning (UDL) application on e-learning acceptance: A structural equation model*. *Int. Rev. Res. Open Distance Learn.* 18, 54–87 (2017). <https://doi.org/10.19173/irrodl.v18i6.2880>.
 21. Riviere, G., Borgiel, K., Couture, N.: *Les Cartouches, un concept d'interfaces tangibles, appliquées à l'apprentissage des jeunes enfants*. In: *5e édition de la conférence Environnements Informatiques pour l'Apprentissage Humain*. pp. 53–60. Mons, France (2011).
 22. Marne, B.: *Éléments de réflexion autour de la conception du jouet pour un jeu sérieux*. In: *EIAH'2019*. Paris, France (2019).
 23. Alvarez, L., Ramalho, M.: *GamesHUB*, <https://hep3.emf-infopro.ch/>, (2021).
 24. Giasson, J.: *La lecture. De la théorie à la pratique*. De Boeck Education, Bruxelles, Belgique (2013).
 25. Geoffre, T., Colombier, N.: *Français langue de scolarisation : Réflexions sur les référentiels de compétences et l'adaptive learning*. In: *9th Conference on EIAH*. pp. 199–204. Paris, France (2019).
 26. Chevallard, Y.: *La transposition didactique : du savoir savant au savoir enseigné*. La Pensée sauvage, Grenoble, France (1991).

S'entraîner à l'asepsie avec la réalité virtuelle

Leonardo Angelini¹, Alexandre Favre¹, François Magnin², Mariateresa De Vito Woods², Omar Abou Khaled¹, Elena Mugellini¹

¹ Institut HumanTech, HES-SO, Fribourg, Suisse

² Haute École de Santé, HES-SO, Fribourg, Suisse

{leonardo.angelini, francois.magnin, mariateresa.devitowoods,
omar.aboukhaled, elena.mugellini}@hefr.ch,
alexandre.favre@master.hes-so.ch

Résumé. Dans une collaboration interdisciplinaire entre une Haute École d'Ingénierie et une Haute École de Santé, nous avons développé une plateforme de réalité virtuelle pour l'entraînement à l'asepsie dans un contexte de soins. La première application développée réplique soigneusement un scénario de prise de sang. La plateforme permet de visualiser les erreurs d'asepsie et les risques de contamination pendant l'exécution des gestes. Un système de points permet de ludifier l'entraînement et de comparer les scores des étudiants. Nous avons mené des tests préliminaires avec 9 étudiants avec un casque de réalité virtuelle Oculus Quest 2. Lors du test, nous avons comparé différentes modalités d'interaction (manettes ou suivi des mains) et différents degrés d'immersion (téléportation dans un environnement complètement virtuel, ou mappage d'une vraie salle de simulation, avec possibilité de toucher des éléments physiques). Ces expériences nous ont permis de tirer un bilan sur le niveau d'immersion atteint par les nouvelles technologies de réalité virtuelle et sur leur implication possible dans la formation pratique des étudiants.

Mots-clés : Réalité virtuelle, Asepsie, Simulation, Contaminations, Oculus Quest.

Abstract. In an interdisciplinary collaboration between a University of Applied Sciences of Engineering and a University of Applied Sciences of Nursing, we have developed a virtual reality platform for asepsis training in a healthcare context. The first application developed carefully replicates a blood sampling scenario. The platform makes it possible to visualize the asepsis errors and the risks of contamination during the execution of the gestures. A point system allows to gamify the training and compare the students' scores. We conducted preliminary tests with 9 students with an Oculus Quest 2 virtual reality helmet. During the test, we compared different modes of interaction (joysticks or hand gestures) and different degrees of immersion (teleportation in a completely virtual environment, or mapping of a real simulation room, with the possibility of touching physical elements). These experiments allowed us to evaluate the level of immersion achieved by the new virtual reality technologies and their possible involvement in the practical training of students.

Keywords: Virtual Reality, asepsis, simulation, contamination, Oculus Quest.

1 Introduction

Les erreurs médicales sont la troisième cause de décès aux États-Unis avec 250 000 morts par année après les maladies cardiaques et les cancers [1]. En particulier, les infections du site chirurgical (ISS) sont une des erreurs les plus fréquentes [2]. Afin d'éviter ces infections l'OMS propose une liste de 29 recommandations pour garder un environnement propre en salle opératoire et pour décontaminer les appareils médicaux et des instruments chirurgicaux [3].

La prévention des infections est donc un élément clé dans l'apprentissage des infirmiers et plusieurs activités pratiques en simulation sont menées lors de la formation des étudiants. À ce propos, des salles du centre de simulation de la Haute École de Santé de Fribourg ont été aménagées en guise de chambre de patient afin de simuler un environnement de soins. L'accès à ce centre est évidemment restreint par sa forte occupation et règlementé par un calendrier académique. Les nouvelles technologies, et en particulier la réalité virtuelle, pourraient apporter un bénéfice concret dans le déroulement des simulations, car elles permettent non seulement d'optimiser l'utilisation des salles de simulation (grâce à un temps réduit pour la mise en place des expériences) mais aussi d'en faire complètement abstraction, permettant à l'étudiant de s'entraîner directement à la maison ou en dehors des heures de cours. En effet, la diffusion des casques de Réalité virtuelle (RV) à un prix abordable ouvre des nouvelles perspectives dans l'entraînement pratique dans les soins infirmiers. En particulier, dans cet article nous nous référons au casque Oculus Quest 2, désormais disponible sur le marché à un prix de 350€ (au moment de l'écriture de cet article), qui permet d'expérimenter des environnements virtuels à haute-fidélité sans besoin d'une machine dédiée pour le rendu graphique.

Dans ce contexte, et dans le cadre d'une collaboration interdisciplinaire entre la Haute École de Santé (HEdS) et la Haute École d'Ingénierie, nous avons développé une plateforme de réalité virtuelle pour l'entraînement à l'asepsie, en particulier dans un scénario de prise de sang. Dans cet article, nous présentons le résultat d'une première expérience avec 9 étudiants, afin de comprendre sa pertinence et son utilisabilité pour un futur parcours d'apprentissage hybride, supporté par les outils numériques en plus de la simulation réelle. Notre travail s'inspire des précédentes recherches sur la simulation en RV dans le domaine de la santé [4]. Il s'inspire également des jeux sérieux en 3D tels que 3D Virtual Operating Room [5] et Safe Surgery Trainer [6] qui visent à entraîner le personnel périopératoire et à favoriser une meilleure communication entre eux. En particulier, pour la représentation des contaminations avec des points colorés, nous nous sommes inspirés de CatCare [7].

L'expérience et l'outil présentés dans cet article s'inscrivent dans un processus de développement itératif centré utilisateur. Nous présentons ici les résultats des tests exploratoires menés afin de combler au manque de réalisme et aux difficultés d'apprentissage de l'utilisation des manettes relevés lors du test de la première version de l'outil.

2 Application RV pour l'entraînement de l'asepsie

Pour la deuxième version de l'outil RV d'entraînement à l'asepsie, nous avons récréé une chambre de soins qui réplique celle du centre de simulation de la HEdS. Le degré de fidélité est très élevé (au centimètre près) afin d'améliorer le sentiment d'immersion et de permettre un mappage direct de la salle. De cette façon, nous avons pu tester une modalité d'interaction en virtualité augmentée [8], afin de pouvoir toucher des objets réels lors des déplacements dans la salle et des interactions avec l'environnement virtuel. Plusieurs modèles 3D ont été aussi mis en place pour les accessoires médicales (garrots, tubes Vacutainer®, tampons, aiguilles...) ainsi que pour le mobilier (charriot avec poubelles et désinfectant, boîte à gants...) et pour le patient (veine visible, simulation du saignement...). Le scénario agréé avec les professeurs de la HEdS et implémenté dans une application Oculus Quest 2 avec le logiciel Unity comprends 17 différentes étapes pour l'exécution correcte de la prise de sang. Plusieurs étapes sont dédiées à la désinfection des mains et à l'utilisation des gants pour éviter toute contamination.



Fig. 1. Environnement virtuel (à gauche dans chaque image) et environnement réel (à droite dans chaque image)

À chaque étape correcte, le système assigne des points. Des points additionnels sont assignés selon la rapidité de l'exécution du scénario, afin d'introduire un élément ultérieur de compétition entre les étudiants. La rapidité d'exécution a toutefois un poids inférieur sur le score final par rapport au respect de l'asepsie. Les points obtenus ainsi que le classement de chaque étudiant peuvent être visualisés dans une application web, accessibles aux étudiants et aux professeurs. L'application donne aussi un retour sur les performances avec mention particulière des erreurs effectuées lors du scénario de soin.

Un élément particulièrement important de l'application proposée est la possibilité de visualiser les contaminations sous forme de points colorés : violet si la source est le patient, par exemple quand l'infirmière touche le patient avec des gants et puis elle oublie de les jeter ; jaune si la source est l'infirmière, par exemple quand elle oublie de désinfecter les mains après être rentrée dans la salle. Les contaminations peuvent être visualisées pendant l'exécution du scénario (modalité facile) ou à la fin (difficile).

3 Tests préliminaires et résultats

Méthodes. Dans le test préliminaire de la deuxième version de l'application nous avons mené trois expériences, chacune visant à comparer différentes conditions d'immersion :

1. Modalité d'entrée (E) : suivi des mains (E_M) VS manettes (contrôleurs) (E_C)
2. Visualisation (V) des mains dans la condition (E_C) : « Mains uniquement » (V_M) VS « Mains et manettes » (V_MC) (la condition « manettes uniquement » a été écartée car moins immersive, comme montré dans des études précédentes [9]).
3. Mappage de la salle (M) : Mappage Complet (M_C) VS Mappage partiel (M_P). Dans le premier cas il s'agit de mapper complètement la salle afin de permettre à l'étudiant d'aller chercher un objet dans une vraie armoire ; dans le deuxième il s'agit de mapper seulement le lit et le chariot et de mettre à disposition tous les objets dans le chariot, sans nécessité de marcher dans la salle. Cette deuxième configuration serait utile afin de permettre la simulation dans une autre salle qui n'a pas la même configuration que celle mappée dans l'application.

Le test a eu lieu à la HEdS, dans une salle identique à celle mappée dans l'application. 9 utilisateurs ont participé à ce test, six femmes. Sept d'entre eux sont étudiants en soins infirmiers, deux sont étudiants en informatique. Huit utilisateurs avaient entre 20 et 24 ans et le dernier avait 52 ans. Les participants ont indiqué être assez à l'aise avec les nouvelles technologies (3,33 / 5,00) mais pas très familiers avec la RV et le suivi des mains en RV (1,56 / 5,00).



Fig. 2. Test utilisateur. À droite, prise d'écran de la scène de désinfection des mains.

Chaque utilisateur a testé les trois expériences (E, V et M), et les deux conditions pour chaque expérience (6 conditions différentes en total). La durée du test était d'environ 30 minutes. À la fin de chaque expérience, les utilisateurs devaient répondre à des questionnaires. L'ordre des expériences et des conditions a été croisé afin d'éviter des effets d'apprentissage. Pour chacune des trois expériences nous avons posé le questionnaire Self-Assessment Manikin (SAM)[10], qui vise à déterminer le ressenti émotionnel des utilisateurs en termes d'activation physiologique (A), de valence (V) et de dominance (D), et le questionnaire Igroup Presence Questionnaire (IPQ)[11], qui mesure le sentiment d'immersion selon les axes de présence spatiale (SP), d'implication (INV) et de réalisme perçu (REAL). Pour l'expérience sur les modalités d'entrées (E), qui répliquait la plus part du scénario décrit dans la Section 2, nous avons également posé aussi le questionnaire System Usability Scale (SUS)[12]. Les autres deux expériences étant plus courtes, nous avons considéré ce questionnaire comme non nécessaire.

Analyse. Vu le nombre réduit de participant, les données récoltées avaient une distribution non-normale. Nous avons donc comparé les deux conditions de chaque expérience avec un test Wilcoxon signed-rank. Nous reportons donc la médiane des résultats à la place de la moyenne.

Résultats. Le tableau suivant reporte les résultats (médianes) des questionnaires pour chaque expérience.

		SAM	IPO	SUS
Modalité (E)	Mains (E_M)	V: 8.00 A: 8.00 D: 7.00	SP: 2.20 INV: 1.25 REAL: 0.75	SUS: 82.50
	Manettes (E_C)	V: 7.00 A: 7.00 D: 8.00	SP: 2.40 INV: 1.25 REAL: 0.75	SUS: 85.00
Visualisation (V)	Mains (V_M)	V: 7.00 A: 7.00 D: 7.00	SP: 2.60 INV: 1.00 REAL: 0.25	
	Mains et Manettes (V_MC)	V: 7.00 A: 6.00 D: 6.00	SP: 2.00 INV: 1.00 REAL: 0.50	
Mappage (M)	Partial (M_P)	V: 7.00 A: 7.00 D: 6.00	SP: 2.00 INV: 0.75 REAL: 0.25	
	Compleat (M_C)	V: 8.00 A: 8.00 D: 6.00	SP: 2.00 INV: 1.25 REAL: 0.50	

L'analyse des résultats de l'expérience sur la modalité d'entrée (E) nous montre qu'il n'y a pas de différences significatives pour ce qui concerne les sentiments ressentis, l'immersion et l'utilisabilité entre les deux conditions testées (avec suivi des mains (E_M) ou avec les manettes de l'Oculus (E_C)). En effet, si d'un côté l'utilisation des mains pour interagir semblerait à première vue plus naturelle et immersive par rapport aux manettes, de l'autre côté, l'utilisateur est obligé de garder une position parfois innaturelle avec ses mains, afin d'assurer un bon suivi par le casque. Le geste de pincement pour saisir un objet virtuel est naturel, mais la perte du suivi des mains risque de briser l'immersion. Cela peut arriver en cas d'occlusion des mains, ou simplement quand elles se touchent, dans le cas où l'utilisateur tourne la tête et garde les mains devant lui, ou simplement pendant les déplacements, si le casque a de la peine à repérer les mains. Le suivi des manettes est beaucoup plus précis, sans risque d'occlusion, et permet facilement de faire abstraction de l'outil avec moins de risque de briser l'immersion par des comportements innaturels des mains virtuelles. Certes, pour quelques personnes l'apprentissage de l'utilisation des manettes risque de constituer une barrière, comme nous l'avons remarqué avec un utilisateur qui était plus âgé que les autres. Dans ce cas l'interaction mains libres est largement préférable. Pour la suite du projet nous avons donc décidé de garder les deux modalités d'interaction : il suffit de poser les manettes pour passer à la modalité main libre. Nous faisons noter que le score SUS pour les deux modalités s'approche de 85 points, qui est considéré comme « excellent » [13].

La deuxième expérience (V), comme la première, n'a pas fait relever des différences significatives entre les différentes visualisations des mains dans la modalité d'interaction avec manettes (E_C). En effet, si la visualisation des mains seules (V_M) pourrait

être plus immersive, pour les utilisateurs novices la possibilité de voir les manettes et le positionnement des boutons peut simplifier l'interaction. Comme pour l'expérience précédente, nous avons décidé de garder les deux modalités de visualisation.

L'expérience du mappage (M) a porté à des différences significatives entre les deux conditions. Les utilisateurs ont particulièrement apprécié la possibilité de marcher librement dans la salle (M_C), ce qui a généré une plus grande activation émotionnelle (A) et des émotions plus positives (V). Le réalisme perçu dans le scénario avec mappage complet en a aussi bénéficié par rapport au mappage partiel ($p=0.04$ avec le test Wilcoxon greater). Pour la suite, nous avons donc décidé d'abandonner l'idée de mapper seulement le lit et le chariot. De plus, pendant le test nous avons remarqué que des pertes de calibration de la salle pouvaient survenir. Dans le cas où la surface du chariot virtuel se trouve plus en bas par rapport à la surface du chariot réel, l'utilisateur pourrait se retrouver dans l'incapacité de récupérer les objets virtuels posés sur sa surface, car la surface réelle du chariot va empêcher physiquement le passage de ses mains ou des manettes. Cela peut affecter autant la condition de mappage complet que partiel. Afin d'éviter ce type de risque, nous avons décidé pour la suite du projet de ne pas mapper le chariot et le lit (qui ne devront pas se trouver donc dans la zone physique d'interaction), mais seulement les murs externes et les armoires. Finalement, pour permettre l'utilisation de l'application dans le cas où l'utilisateur se trouverait dans une salle différente de celle mappée, nous avons réintroduit la possibilité de se déplacer dans la pièce virtuelle par téléportation dans la version finale de l'application. Un nouveau test (par souci de bréveté non inclus dans cet article) a permis de valider le scénario final et de confirmer l'utilité de la plateforme de ludification et de retour sur les erreurs.

4 Conclusion

Dans cet article nous avons présenté une application de réalité virtuelle qui permet de simuler avec un haut degré d'immersion un scénario de prise de sang, avec la possibilité de visualiser les contaminations dues à des erreurs dans la désinfection et hygiène de mains. Nos premiers tests montrent que l'application est facile à utiliser et que les étudiants seraient motivés par la plateforme de ludification associée. L'étude menée avec le casque de réalité virtuelle Oculus Quest 2 montre que cet outil pourrait constituer un complément valide à la pratique dans le centre de simulation, à utiliser dans les salles de simulation en dehors des heures des cours ou même à domicile. Nous avons aussi montré que le mappage spatial d'une vraie salle dans l'environnement virtuel permet d'augmenter le niveau de réalisme perçu. Finalement, le suivi des mains introduit par le casque Oculus s'est avéré à présent pas assez robuste pour pouvoir être préféré aux manettes. Ceci est en accord avec les conclusions de [9].

Pour la suite de ce projet, afin de valider les bienfaits sur l'apprentissage des recommandations pour l'asepsie, nous mettrons à disposition d'un groupe d'étudiants des casques avec cette application et nous comparerons leurs acquis sur les bonnes pratiques pour l'asepsie avec ceux d'un groupe qui n'aura pas accès à cette application. De plus, d'autres scénarios pourraient être implémentés en plus de la prise de sang.

Références

1. Makary, M.A., Daniel, M.: Medical error—the third leading cause of death in the US. *Bmj* 353, (2016)
2. 9 Most Common Medical errors. <https://www.beckershospitalreview.com/quality/9-most-common-medical-errors.html> (2014)
3. Organization, W.H.: Global guidelines for the prevention of surgical site infection. World Health Organization (2016)
4. Moline, J.: Virtual reality for health care: a survey. *Studies in health technology and informatics* 3-34 (1997)
5. Lelardeux, C.P., Galaup, M., Panzoli, D., Boudier, V., Lagarrigue, P.: 3D Virtual Operating Room: Un learning game multi-joueurs 3D temps réel pour former à la gestion des risques. In: *Jeux & Enjeux*. (2017)
6. Kreutzer, C., Marks, M., Bowers, C., Murphy, C.: Enhancing surgical team performance with game-based training. *Int. J. Serious Games* 3, 43-52 (2016)
7. Klinker, K., Fries, V., Wiesche, M., Krcmar, H.: CatCare: Designing a serious game to foster hand hygiene compliance in health care facilities. In: *Designing the Digital Transformation: DESRIST 2017 Research in Progress Proceedings of the 12th International Conference on Design Science Research in Information Systems and Technology*. Karlsruhe, Germany. 30 May-1 Jun., pp. 20-28. Karlsruher Institut für Technologie (KIT), (2017)
8. Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Augmented reality: A class of displays on the reality-virtuality continuum. In: *Telemanipulator and telepresence technologies*, pp. 282-292. International Society for Optics and Photonics, (1995)
9. Voigt-Antons, J.-N., Kojic, T., Ali, D., Möller, S.: Influence of hand tracking as a way of interaction in virtual reality on user experience. In: *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1-4. IEEE, (2020)
10. Bynion, T.-M., Feldner, M.T.: Self-assessment manikin. *Encyclopedia of personality and individual differences* 1-3 (2017)
11. Schubert, T., Friedmann, F., Regenbrecht, H.: The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10, 266-281 (2001)
12. Lewis, J.R., Sauro, J.: The factor structure of the system usability scale. In: *International conference on human centered design*, pp. 94-103. Springer, (2009)
13. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 114-123 (2009)

Outils numériques pour l'apprentissage de la grammaire en français langue d'enseignement : Quels analyseurs pour les appréhender ?

Rosianne Arseneau¹, Thierry Geoffre²

¹ UQAM, Dépt de didactique des langues, 1205 rue Saint-Denis, Montréal (Québec) H2X 3R9
arseneau.rosianne@uqam.ca

² HEP Fribourg, UR DidaLang, 36 rue de Morat, CH-1700 Fribourg, Suisse
thierry.geoffre@eduf.fr

Résumé. Les nouvelles technologies permettent aux enseignants et aux élèves d'accéder à de nombreuses ressources pédagogiques et didactiques en ligne avec, entre autres spécificités, une grande variété d'organisation et de rétroactions. Ainsi peut-on observer un certain foisonnement de propositions promettant de s'attaquer à la dimension grammaticale de la langue. À notre connaissance, aucune contribution ne s'est intéressée aux connaissances scientifiques qui permettent de les appréhender. Nous présentons une typologie de six analyseurs qui s'appuie sur des connaissances issues de champs de recherche jugés éclairants (didactique de la grammaire et de l'écriture, psychologie cognitive, psycholinguistique, numérique en éducation). Quelques exemples d'outils en ligne sont fournis. Nous concluons sur les besoins qui nous semblent prioritaires, notamment une revue systématique éclairée par la typologie des analyseurs proposée.

Mots-clés : outils numériques, apprentissage de la grammaire, français langue d'enseignement, analyseurs, révision de texte

Abstract. New technologies allow teachers and students to access many online teaching and learning resources with, among other specificities, a wide variety of organization and feedback. Thus, we can observe a proliferation of numerical propositions promising to tackle the grammatical dimension of the language. What is the scientific knowledge that makes it possible to apprehend them? To our knowledge, no contribution has addressed this question. We present a typology of 6 analyzers based on knowledge from fields of research deemed enlightening (didactics of grammar and writing, cognitive psychology, psycholinguistics, digital learning). Examples of available tools are given. We conclude on the needs that seem to us as the most pressing such as drawing a global picture through a systematic review based on the typology of analyzers provided.

Keywords: digital tools, grammar learning, French as language of instruction, analyzers, text revision

1 Contexte et objectif

En avant-propos à un rapport PISA [1], Andreas Schleicher, directeur à la Direction de l'éducation et des compétences, soulignait les possibilités quasi-infinies offertes par le numérique : « La technologie est le seul moyen d'élargir au maximum l'accès à la connaissance. [...] les nouvelles technologies permettent aux enseignants et aux élèves d'accéder à des ressources spécialisées bien au-delà de ce que peuvent leur offrir les manuels scolaires, sous de multiples formats et sans aucune contrainte d'espace et de temps ou presque. » (p. 2).

Cet accès accru et élargi à des ressources via les technologies s'est avéré crucial pour la continuité des apprentissages avec l'arrivée en 2020 de la pandémie de Covid-19 qui a entraîné la fermeture de nombreux établissements d'enseignement.

Au-delà de cette accessibilité accrue, un des facteurs qui peuvent enrichir l'expérience d'apprentissage est la variété des formes de rétroaction offertes par les supports numériques, qu'elles concernent la motivation ou de réelles situations d'apprentissage, sous forme d'encouragements interactifs, de scores, de points gagnés, de personnalisation du parcours avec l'adaptive learning, etc.

Vu les difficultés des élèves de l'école obligatoire en français écrit, notamment pour la dimension grammaticale, il n'est pas surprenant d'observer un certain foisonnement de solutions numériques promettant de s'y attaquer. Quelles sont les connaissances scientifiques qui nous permettent de les appréhender ? À notre connaissance, aucune contribution ne s'est intéressée à cette question.

L'objectif de la contribution est de proposer des « analyseurs » permettant de classer les outils numériques et d'en évaluer la potentielle pertinence pédagogico-didactique en puisant à des concepts issus de différents champs (didactique de la grammaire et de l'écriture, psychologie cognitive, psycholinguistique, numérique en éducation). Cette liste d'analyseurs tracera, dans la lignée de la revue d'Alvarez, Cuko et Geoffre [2], la voie à une revue systématique des outils numériques.

2 Cadre conceptuel

Nous avons établi une grille qui nous permet d'appréhender les outils numériques selon six analyseurs synthétisés dans le tableau 1.

Tableau 1. Typologie des analyseurs des outils numériques.

Analyseurs	Catégories
1. Matériau écrit	phrases détachées / paragraphe / texte
	matériau préexistant / matériau entré par l'apprenant (« autoproduit »)
	sans erreurs / avec erreurs

2. Tâches cognitives sollicitées (adapté de Arseneau et Désilets, 2021 [3])	É-Étiqueter une unité <i>avec métalangage</i> (classe de mots, groupe ou fonction syntaxique, donneur/receveur d'accord, catégorie d'erreur) M-Manipuler des phrases pour vérifier une hypothèse R-Représenter l'organisation de la phrase de manière schématique D-Détecter une erreur C-Corriger une erreur J-Justifier une <i>graphie/construction/ponctuation</i> correcte ou une correction d'erreur <i>avec métalangage</i> Ch.-Choisir la bonne <i>graphie, construction ou ponctuation</i> P-Produire une <i>graphie/construction/ponctuation</i> correcte			
3. Contenus grammaticaux	orthographe grammaticale / syntaxe / ponctuation			
4. Rétroaction de l'outil	binaire (réponse correcte / erronée), quantitative (score), qualitative (explication, indice, représentation) ou hybride avec explications ou indices métalinguistiques / sans explications ni indices métalinguistiques			
5. Parcours articulés	<table border="1" style="width: 100%;"> <tr> <td data-bbox="400 760 780 806">à partir des ressources existantes présentes dans l'application</td> <td data-bbox="786 760 943 806" rowspan="2">possible / impossible</td> </tr> <tr> <td data-bbox="400 815 780 824">à partir de ressources produites par l'enseignant</td> </tr> </table>	à partir des ressources existantes présentes dans l'application	possible / impossible	à partir de ressources produites par l'enseignant
à partir des ressources existantes présentes dans l'application	possible / impossible			
à partir de ressources produites par l'enseignant				
6. Récupération des traces	possible / impossible			

2.1 Analyseur 1. Le matériau écrit.

Le matériau écrit utilisé par l'outil numérique apparaît comme un analyseur important d'abord en termes de longueur. En effet, les unités écrites plus longues (paragraphe, texte) requièrent de la part de l'utilisateur de l'outil une capacité de traitement en lecture plus avancée et un « empan perceptif » plus grand [4].

Ensuite, du point de vue de la composition du texte, la présence d'erreurs à détecter dans le matériau écrit saura complexifier la tâche de l'utilisateur en déclenchant un processus de révision exigeant au plan cognitif [5].

Enfin, le travail sur un matériau entré par l'apprenant (ou « autoproduit ») sera plus exigeant qu'un matériau préexistant fourni par l'outil : « [le scripteur] a tendance à percevoir ce qu'il croit avoir écrit plutôt que ce qu'il a écrit véritablement [...], ce qui entraîne qu'il est un correcteur plus efficace des textes d'autrui que des siens. » [6 : 151].

Dans cette perspective, un outil comme le Détecteur de faute¹, proposant à l'utilisateur une tâche de révision sur un ou des paragraphes donnés sera de niveau cognitif assez élevé, mais pas aussi élevé que si le texte à réviser avait été écrit par l'apprenant.

¹ <http://detecteur.ccdmd.qc.ca/#/exercices/p2/parcours2-sans-aide> utilisé au collégial québécois (17-20 ans).

2.2 Analyseur 2. Les tâches cognitives

Le concept de « tâche cognitive » utilisé dans différentes contributions en sciences de l'éducation s'appuyant sur la psychologie cognitive [7, 8] est ici articulé aux concepts psycholinguistiques d'« habiletés linguistiques / métalinguistiques ». « Les habiletés métalinguistiques sont généralement envisagées selon deux composantes, soit *l'analyse* des structures linguistiques, soit *le contrôle* de l'action posée sur le produit langagier [9, 10], ce qui fait écho à la notion de connaissances explicites [11]. » [12 : 53].

Aussi, les différentes tâches cognitives proposées dans cette contribution (cf. tableau 1) et adaptées de Arseneau et Désilets [3] font respectivement appel à une activité d'ordre plus ou moins « linguistique » ou « métalinguistique » induite par la consigne de l'outil numérique. Par exemple, les outils comme NetPhrases² sollicitant à la fois l'étiquetage d'unités linguistiques avec métalangage [É] (composante « analyse ») et la correction du matériau écrit [C] (composante « contrôle ») solliciteraient davantage d'habiletés métalinguistiques et prépareraient ainsi mieux à une mobilisation consciente de connaissances grammaticales en écriture [12] que des outils ne faisant pas appel à la dimension métalinguistique au moment de choisir une bonne graphie [Ch.]³.

2.3 Analyseur 3. Les contenus grammaticaux ciblés

La dimension grammaticale du français écrit, et plus spécifiquement de la grammaire de la phrase, renvoie à trois grands ensembles de contenus didactiques : l'orthographe grammaticale, la syntaxe (ou construction des phrases) et la ponctuation [13]. À la lumière d'autres contributions [3], nous faisons l'hypothèse que de plus nombreux outils numériques disponibles actuellement ciblent l'orthographe grammaticale, au détriment de la syntaxe et de la ponctuation qui méritent pourtant attention avec les nombreuses erreurs qui en découlent [14].

2.4 Analyseur 4. La rétroaction de l'outil

Les outils proposés en exemples fournissent des rétroactions de trois types : binaire (réponse correcte / erronée), quantitative (score), qualitative (explication, indice, représentation). Une combinaison de deux ou trois types de rétroaction serait qualifiée de « rétroaction hybride ».

2.5 Analyseur 5. Les parcours articulés

La possibilité de parcours articulés sous-entend la proposition, par l'outil, de parcours permettant d'aborder un ou plusieurs objets d'enseignement/apprentissage (objectifs spécifiques en relation à l'objectif général ciblé) de manière organisée et explicitement progressive (donc pas seulement une liste d'exercices similaires). Un prolongement est la possibilité pour l'enseignant de modifier ces parcours (par ajout ou suppression d'exercices), de les attribuer individuellement ou encore d'en créer à partir d'activités personnalisées (choix de mots, phrases, textes de base, par exemple). Un

² NetPhrases.net utilisé au secondaire, collégial et universitaire québécois (12 ans - adulte).

³ Par ex., <http://jaccorde.com/> utilisé au 3^e cycle du primaire et au secondaire québécois (10-17 ans).

outil comme la *famille Gram*⁴ permet à l'enseignant de construire ses propres phrases et de les intégrer dans des parcours personnalisés. Cette possibilité ouvre aussi sur celle de la création de phrases par l'apprenant, ce qui le place dans une tâche complexe de production [P].

2.6 Analyseur 6. La récupération des traces

Un dernier analyseur, plus technique que linguistique ou didactique, concerne la récupération des traces de l'apprenant. Quelques outils seulement semblent avoir anticipé cette question, par exemple *NetPhrases* et *la famille Gram*. Dans ces outils, l'enseignant peut en effet accéder à différentes informations telles que le nombre et la nature des exercices réalisés par un apprenant, le score par exercice et le score moyen par apprenant. Le tableau de bord enseignant de *la famille Gram* présente ainsi différentes informations : la date, le mode de jeu, le temps passé, le temps d'inactivité, le score et les différentes tentatives de l'élève avec indication de l'erreur et de son type. Il y a ici la possibilité d'un suivi fin du travail de l'élève et de ses réussites et erreurs.

3 Conclusion

Nous avons proposé ici une typologie d'analyseurs des outils numériques d'apprentissage de la grammaire en français langue d'enseignement en puisant à des disciplines contributives. Cette typologie doit maintenant être mise à l'épreuve d'une revue systématique des outils disponibles dans les moteurs de recherche. Les bases théoriques utilisées pour l'analyse pourraient contribuer à identifier les forces et lacunes des outils, puis à dégager différents résultats en termes d'enjeux didactiques, d'utilisation et de développements. Par exemple, nous aimerions vérifier à travers cette revue systématique l'hypothèse selon laquelle les traces d'activité des apprenants ne sont que très rarement collectées et que l'idée d'un enseignement-apprentissage avec des parcours auto-adaptatifs n'apparaît pas alors que la grammaire est un domaine suffisamment modélisable pour s'y prêter. L'existence ou non d'environnements de type EIAH capables, pour le français, de proposer des parcours d'enseignement/apprentissage de la grammaire en lien avec la compréhension et la production est une autre question.

Enfin, au-delà des descriptions et analyses plus théoriques s'impose la mise en place rapide d'études empiriques évaluant les impacts de ces outils sur l'apprentissage des contenus grammaticaux et également sur leur mobilisation en écriture. À ce sujet, des chantiers de recherches quasi-expérimentales (par ex., [15]) et d'autres types⁵ en cours donneront prochainement accès à des résultats et entraîneront dans plusieurs cas comme retombée seconde l'amélioration des outils étudiés. La démocratisation des apprentissages accélérée par le foisonnement d'outils souvent gratuits ne devrait pas se faire au détriment de la qualité des apprentissages qu'ils encourtent.

⁴ <https://www.lafamillegram.ch/#/> destiné à des élèves du primaire.

⁵ Par exemple avec le projet COMPER qui devrait s'intéresser à l'orthographe grammaticale dans la phrase (<https://comper.fr>).

Références

1. OCDE; Connectés pour apprendre ? Les élèves et les nouvelles technologies. PISA – Principaux résultats, Paris, France (2015). Repéré à <http://www.oecd.org/fr/education/scolaire/Connectes-pour-apprendre-les-eleves-et-les-nouvelles-technologies-principaux-resultats.pdf>
2. 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. *Revue systématiques du C·R·E/A·TE*, 2. <http://doc.rero.ch/record/324244>
3. Arseneau, R., & Désilets, M. : Comment améliorer l'intégration d'outils numériques d'apprentissage dans les cours de grammaire et de français écrit? Correspondance (avril 2021). Accessible à <https://correspo.ccdmd.qc.ca/index.php/document/comment-ameliorer-l-integration-doutils-numeriques-dapprentissage-dans-les-cours-de-grammaire-et-de-francais-ecrit/>
4. Vitu, F. & O'Regan, J.K. : Les mouvements oculaires comme indice « on-line » des processus cognitifs : rêve ou réalité ? *Psycholinguistique Cognitive*, 189-214 (2004). Accessible à <https://www-caim-info.proxy.bibliotheques.uqam.ca/psycholinguistique-cognitive--9782804146313-page-189.htm>
5. Roussey, J.-Y., & Piolat, A.: La révision du texte : une activité de contrôle et de réflexion. *Psychologie Française*, 50(3), 351-372 (2021)
6. Gombert, J. E. : Le rôle des capacités métalinguistiques dans l'acquisition de la langue écrite. *Repères*, 3, 143-156 (1991).
7. Barth, B.-M. : L'apprentissage de l'abstraction (2 éd.). Paris: Éditions Retz (2004).
8. Astolfi, J.-P. & Develay, M. Didactique des sciences et processus d'apprentissage. Dans *La didactique des sciences* (2016), p. 65 à 97 (2016).
9. Bialystok, E.: Factors in the growth of linguistic awareness. *Child Development*, 57(2), 498-510 (1986).
10. Simard, D., Foucambert, D. et Labelle, M.: Examining the contribution of syntactic and metasyntactic abilities to reading comprehension among native and non-native speakers. Dans K. Roehr et G. A. Gánem-Gutiérrez (dir.), *The metalinguistic dimension in instructed L2 learning*. Bloomsbury Academic (2013).
11. Nadeau, M. et Fisher, C.: Les connaissances implicites et explicites en grammaire: quelle importance pour l'enseignement ? Quelles conséquences ? *Bellaterra journal of teaching & learning language & literature*, 4(4), 1-31 (2011).
12. Arseneau, R. et Lefrançois, P.: « Auxquels ? C'est le pronom relatif » : susciter des interactions verbales pour développer la compétence métalinguistique des élèves du secondaire. *Revue des sciences de l'éducation*, 45(3), 47-79 (2019). doi:<https://doi.org/10.7202/1069640ar>
13. Riegel, M., Pellat, J.-C. et Rioul, R.: (Grammaire méthodique du français (6 éd.). Presses universitaires de France, Paris, France (2016).
14. Boivin, M.-C. et Pinsonneault, R.: Les erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale des élèves québécois en contexte de production écrite. *Revue canadienne de linguistique appliquée*, 21(1), 43-70 (2018).
15. Gauvin, I., Arseneau, R., Charland, P., Côté, M.-F., Cyr, S., Fontaine, M., L. Laplante, L., Plante, I. et Riopel M.: Quels effets de jeux éducatifs en ligne sur l'apprentissage et sur la motivation des élèves? Évaluation de ressources de Alloprof pour le français et les mathématiques. *Alloprof – Conseil de recherche en sciences humaines du Canada* (2020-2022).

Système de recommandation de ressources pédagogiques pour un apprentissage sur application mobile parascolaire

Anaëlle Badier, Mathieu Lefort et Marie Lefevre

Université Lyon 1, LIRIS, UMR 5205, F-69622, France
`prenom.nom@liris.cnrs.fr`

Résumé Nous présentons dans cet article les principes d'un système de recommandation de ressources pédagogiques dans le cadre d'une application mobile à usage parascolaire. Le système proposé permet de fournir des recommandations à travers différents niveaux scolaires et d'orienter les apprenants vers des ressources de remédiation, de continuité du programme ou d'approfondissement. Notre système s'appuie sur l'organisation des ressources pédagogiques disponibles dans l'application en un "arbre de connaissances", et sur l'évaluation du niveau réel d'un apprenant par la méthode de la théorie de réponse aux items (*IRT*). Les premières validations des experts pédagogiques semblent confirmer notre approche et nous encourageant à poursuivre par la validation du système en situation réelle ouvrant ainsi de nombreuses perspectives de recherche.

Mots-clé: Système de recommandation · Apprentissage adaptatif · Plateforme pédagogique mobile · *IRT*

Abstract. In this article, we present the principles of a learning resources recommendation system in a mobile application for extracurricular use. The system provides recommendations across different grade levels to direct learners toward resources for remediation, school curriculum continuity or in-depth study. Our system is based on the organisation of the pedagogical resources available in the application structured with a "knowledge tree", and on the assessment of the learner's actual level using the item response theory (*IRT*) method. The first feedback of pedagogical experts seem to validate the approach that will be next tested in real-life situations and open up numerous research perspectives.

Keywords: Recommendation system · Adaptive Learning · Mobile educational platform · *IRT*

1 Introduction

Les avancées technologiques de ces dernières années ont permis de proposer de nombreux systèmes d'apprentissage en ligne (MOOC, applications de e-learning,

etc.) [2]. La démocratisation de ces outils, jusqu'alors restreints au contexte du laboratoire, a augmenté la quantité de données utilisateurs en contexte réel disponibles et permet d'utiliser les techniques d'intelligence artificielle pour améliorer les systèmes d'apprentissage adaptatif [7].

Nous travaillons avec une application mobile qui propose du contenu éducatif pour tous les niveaux, du collège au supérieur et toutes les matières y compris les enseignements professionnels ou de spécialité. L'application est construite pour être un outil parascolaire volontaire, de révision ou de découverte, plus que d'apprentissage. L'enjeu est de pouvoir naviguer facilement entre les différents niveaux scolaires afin de compléter l'enseignement vu en classe.

Notre objectif est donc de proposer un système de recommandation qui identifie automatiquement les besoins de l'apprenant et propose le contenu de l'application le plus adapté. Nous nous focaliserons dans cet article sur les connaissances exploitées pour notre système de recommandation. Le modèle apprenant et l'interface d'interaction feront l'objet des travaux à venir.

Dans un premier temps, nous exposerons les travaux de recherche existants sur lesquels s'appuie notre projet pour contextualiser et justifier nos travaux (section 2) avant de présenter les principes sur lesquels s'appuie notre système (section 3). Nous détaillerons ensuite un protocole de validation du système proposé (section 4) pour conclure sur les perspectives de recherche (section 5).

2 État de l'art

Les systèmes de recommandation existants appliqués aux EIAH combinent plusieurs approches et peuvent être basés sur les profils des utilisateurs, sur la structure du contenu à recommander ou sur une approche axée compétences. Yago et al. proposent une revue des différents systèmes de recommandation [11].

Les systèmes de filtrage par contenu (*content-based filtering*) proposent de recommander un contenu proche du contenu déjà étudié donc plus susceptible de satisfaire les attentes de l'utilisateur. Les contenus, pour la plupart issus du web, sont structurés en ontologies [9] ou réseaux sémantiques. Les recommandations sont ensuite effectuées par un système de correspondance entre une requête (correspondant au contenu récemment étudié) et la base de ressources. L'association peut par exemple être réalisée en utilisant des techniques de web-sémantique qui s'appuient sur des métadonnées descriptives de la ressource [8]. L'utilisation de techniques de filtrage par contenu est particulièrement efficace lorsque l'on dispose d'une grande diversité de ressources qui traitent du même sujet.

Les méthodes de recommandation axées utilisateurs peuvent proposer des systèmes de filtrage collaboratif (*collaborative filtering*) en identifiant des similitudes entre utilisateurs pour proposer à un apprenant un contenu étudié par un autre apprenant au profil similaire, *via* différentes méthodes de construction de profils utilisateurs [7]. Les activités utilisateurs peuvent également être analysées par des algorithmes de fouille de données pour définir des règles d'associations entre un certain profil et une ressource à étudier [4]. Ces algorithmes rencontrent toutefois des limites computationnelles lorsque les jeux de données sont très im-

portants. De plus, dans le cas de notre application volontaire et parascolaire, il serait difficile de caractériser des profils utilisateurs sans avoir de suivi régulier dans le temps, ce qui limite l'efficacité de ces techniques dans notre contexte.

Les méthodes de recommandation par compétences sont liées aux mécanismes cognitifs de l'apprentissage. Ils introduisent une chronologie en proposant d'organiser les ressources à étudier en représentant leur accessibilité par compétences acquises et à acquérir. La structuration des ressources en compétences est utilisée comme paramètre des modèles de traçage de connaissances dans l'article de Vie et Kashima [10]. La prise en compte des précédentes tentatives de l'apprenant pour maîtriser un sujet permet une meilleure prédiction de réussite. Cette influence des cours déjà étudiés sur la réussite aux cours suivants est également mobilisée dans les travaux de Jiang et al [6], qui introduisent la notion de Zone de Développement Proximal (ZPD) développée par Lev Semyonovich Vygotsky. Clément et al. basent également leurs modèles sur la ZPD et propose de limiter les recommandations aux contenus de niveau légèrement supérieur au niveau réel de l'apprenant [3]. En effet, suggérer des ressources de niveau trop inférieur au niveau de l'apprenant ne lui permet pas d'augmenter ses connaissances et des ressources de niveau trop élevé risquent de provoquer une frustration.

La caractérisation du niveau de connaissance de l'apprenant est donc essentielle pour pouvoir utiliser une approche par acquisition de compétences. La méthode statistique de la théorie de réponse aux items (*Item Response Theory - IRT*) formalisée par Baker [1] permet d'évaluer la compétence d'un apprenant en s'appuyant sur les réponses des autres apprenants pour caractériser les items auxquels l'apprenant est confronté. L'application mobile sur laquelle nous travaillons étant utilisée par un million d'utilisateurs, nous disposons de nombreuses données qui permettent d'utiliser cette méthode statistique d'évaluation de compétences.

3 Combiner *Connaissances* et *IRT* pour recommander

L'approche que nous proposons est une méthode de type "Support de navigation adaptatif" [2]. Le système doit proposer à la fin de chaque ressource consultée la ressource qui serait la plus adaptée pour poursuivre l'apprentissage. Afin de proposer un contenu adapté au niveau de l'apprenant, nous proposons dans un premier temps une structuration des ressources en "arbre de connaissance" qui permet de les lier à travers différents niveaux scolaires (section 3.1), puis une évaluation du niveau de connaissance de l'apprenant en utilisant l'IRT (section 3.2) pour l'orienter vers une ressource de difficulté correspondant à son niveau réel, indépendamment de son niveau scolaire.

3.1 Arbre de connaissances

Organisation des ressources. Les ressources de l'application sont organisées par niveaux scolaires. Chaque niveau comporte des disciplines, structurées en sous-disciplines, elles-mêmes séparées en chapitres. Un chapitre comporte de 1 à 3 mini-cours et de 1 à 5 quizz de 10 questions en moyenne.

Annotation et structuration des ressources. Nous proposons dans un premier temps de faire la recommandation au niveau des chapitres. Afin de lier les chapitres entre eux, chacun est caractérisé par un ensemble de notions décrivant les connaissances mobilisées dans le chapitre. Ces notions font référence aux autres chapitres disponibles et aux relations pédagogiques qui existent entre ces chapitres. En l'absence de métadonnées descriptives du contenu de chaque chapitre, l'apposition des notions est effectuée par des professeurs, en suivant les descriptifs du programme de l'éducation nationale^[1]. Par exemple, le chapitre de mathématiques de Seconde traitant des *Fonctions de référence* est annoté en indiquant qu'il est nécessaire de maîtriser le chapitre *Fonctions affines* de niveau 3^e.

Le chemin pédagogique se construit par similarité entre les chapitres en fonction des connaissances mobilisées. Nous utilisons un algorithme de similarité basé sur les liens entre notions et sur la répartition des différentes notions au sein de nos ressources.

3.2 Évaluation du niveau de capacité de l'apprenant par IRT

Chaque quiz étant conçu par un professeur différent, les quizz n'ont pas tous le même niveau de difficulté. De plus, les quizz sont tous à choix multiples, ce qui rend possible une bonne réponse par sélection aléatoire. Nous disposons d'un grand volume de données (de 2000 à 40 000 réponses enregistrées pour les quizz de mathématiques de niveau collège-lycée), ce qui nous permet d'utiliser une méthode statistique pour évaluer le niveau réel de l'apprenant.

Nous utilisons la méthode de l'IRT pour avoir une valeur de ce niveau réel de compétence pour un item donné. Le modèle sélectionné est le modèle à 3 paramètres (IRT 3PL)^[1], qui estime le niveau de compétence réelle θ d'un apprenant par la formule :

$$P(\theta) = c + (1 - c) \frac{1}{1 + e^{-a(\theta - b)}}$$

(avec θ : compétence réelle de l'apprenant, a : paramètre de discrimination, b : paramètre de difficulté et c : paramètre de hasard).

Face à des profils d'apprenants très disparates, l'IRT présente l'avantage d'indépendance entre le niveau des apprenants ayant servi à estimer les paramètres d'un item et la capacité d'un apprenant. Les apprenants sont ensuite répartis en 3 groupes de niveau : G_{inf} pour les élèves présentant des difficultés sur le chapitre étudié, G_{med} pour le groupe de niveau moyen et G_{sup} pour les élèves maîtrisant bien le chapitre étudié. Les groupes sont définis selon les valeurs de θ : par les propriétés de l'IRT, la répartition des élèves n'influence pas la constitution des groupes. Cette répartition des apprenants permet de les orienter vers des ressources de remédiation (apprenants de niveau G_{inf}), de continuité du programme (niveau G_{med}), ou de découverte (apprenants de niveau G_{sup}), en fonction des notions et relations disponibles dans notre arbre de connaissances.

1. Exemple de l'Annexe 3 relative aux cours de 5e, 4e et 3e : <https://www.education.gouv.fr/bo/15/Special11/MENE1526483Aannexe3.htm>

4 Évaluation du prototype

Les propositions de recommandations de l'algorithme ont été soumises à validation auprès de 3 experts pour les mathématiques de Seconde générale. Nous avons choisi ce niveau scolaire pour correspondre d'une part à une population d'utilisateurs majoritaire et d'autre part pour tester l'algorithme dans un contexte de transition importante de niveau scolaire (passage collège-lycée), où les besoins en remédiation/approfondissement seront plus importants.

La validation a été présentée sous la forme d'un questionnaire en deux parties présentant des situations d'apprenants venant de terminer un chapitre donné et affecté à un groupe de niveau G_{inf} , G_{med} , G_{sup} d'après les résultats IRT. La première partie du questionnaire présente la liste des ressources disponibles dans l'application et demande aux experts le chapitre qu'ils auraient conseillé dans la situation présentée. La deuxième partie présente la recommandation de l'algorithme et demande aux experts d'évaluer cette recommandation.

Les retours d'experts sont plutôt favorables vis-à-vis des recommandations pour les situations des groupes G_{inf} . Pour le groupe G_{sup} , alors que notre modèle recommandait des ressources de niveau Première Générale, les experts préfèrent des exercices de même niveau scolaire mais de niveau de difficulté légèrement supérieur. Pour le groupe G_{med} , ils soulignent la nécessité d'une chronologie entre les chapitres d'un même niveau (d'après la citation d'un expert : "il faut s'assurer que le chapitre B a bien été étudié avant de suggérer le chapitre C ").

La prochaine étape de validation est de rendre la fonctionnalité de recommandation disponible au public pour approfondir la validation du système en étudiant l'utilisation effective des recommandations par les apprenants et leur impact sur leurs connaissances et sur leur utilisation de l'application. Le grand nombre d'utilisateurs permettra d'avoir des données pour faire une validation "en situation réelle" [\[5\]](#).

5 Conclusion et perspectives

Notre approche combinant "arbre de connaissances" et IRT nous a permis de proposer une première version d'un algorithme de recommandation de contenu "étape par étape". Contrairement au paradigme classique où de nombreuses ressources sont disponibles pour étudier un sujet précis, on se place ici dans la situation inverse où l'apprentissage se veut exploratoire avec la possibilité de transiter facilement d'un niveau à un autre.

Les évaluations par les experts semblent valider notre approche, en soulevant toutefois le problème du manque de ressources dans l'application. À partir de ces premiers résultats, nous travaillons actuellement sur une adaptation continue du processus d'Adaptive Learning, *via* une découverte automatique de connaissances et en interaction avec les acteurs du processus d'apprentissage. Cette adaptation automatique permettrait de s'adapter aux besoins de chaque apprenant et à leur comportement lors de l'utilisation de l'application.

Nous souhaitons pour cela exploiter les techniques d'apprentissage automatique et de fouille de données pour mener une analyse diagnostic permettant de construire un profil de connaissances acquises par l'étudiant, mais également une analyse comportementale pour comprendre la manière dont il exploite l'application et son niveau d'implication dans son processus d'apprentissage. Cette analyse comportementale passe par l'aspect motivation de l'élève et s'étudie également au travers de la ludification de l'application. Ces analyses exploiteront les traces d'apprentissage de l'apprenant, mais également celles de l'ensemble des utilisateurs pour faire émerger des comportements et situer l'apprenant par rapport à ces comportements. Ce profil de connaissances associé au profil comportemental permettra d'adapter les recommandations fournies et de choisir quand et comment lui faire des recommandations.

Références

1. Baker, F.B. : The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation, College Park, Md., 2nd ed edn. (2001)
2. Brusilovsky, P., Peylo, C. : Adaptive and Intelligent Web-based Educational Systems. *International Journal of Artificial Intelligence in Education* **13** (2003)
3. Clément, B., Roy, D., Oudeyer, P.Y., Lopes, M. : Multi-Armed Bandits for Intelligent Tutoring Systems. *Journal of Educational Data Mining* **7**(2) (2015)
4. Daher, J.B., Brun, A., Boyer, A. : Multi-source Data Mining for e-Learning. In : 7th International Symposium "From Data to Models and Back (DataMod)". Toulouse, France (2018)
5. Erdt, M., Fernandez, A., Rensing, C. : Evaluating Recommender Systems for Technology Enhanced Learning : A Quantitative Survey. *IEEE Transactions on Learning Technologies* **8**(4), 326-344 (2015)
6. Jiang, W., Pardos, Z.A., Wei, Q. : Goal-based Course Recommendation. In : Proceedings of the 9th International Conference on Learning Analytics & Knowledge. pp. 36-45. Association for Computing Machinery, New York, USA (2019)
7. Markowska-Kaczmar, U., Kwasnicka, H., Paradowski, M. : Intelligent Techniques in Personalization of Learning in e-Learning Systems. In : *Computational Intelligence for Technology Enhanced Learning* (2010)
8. Nguyen, C., Roussanaly, A., Boyer, A. : Learning Resource Recommendation : An Orchestration of Content-Based Filtering, Word Semantic Similarity and Page Ranking. In : 9th European Conference on Technology Enhanced Learning (2014)
9. Nguyen, C., Roussanaly, A., Boyer, A. : Studying Relations Between E-learning Resources to Improve the Quality of Searching and Recommendation. In : 7th International Conference on Computer Supported Education. vol. 1 (2015)
10. Vie, J.J., Kashima, H. : Knowledge Tracing Machines : Factorization Machines for Knowledge Tracing. *Conf. Artificial Intelligence AAAI* (2019)
11. Yago, H., Clemente, J., Rodriguez, D. : Competence-based recommender systems : A systematic literature review. *Behaviour & Information Technology* **37**, 1-20 (2018)

Conception d'un compagnon digital pour assister le maître du jeu dans un environnement mixte

Catherine Bonnat¹, Eric Sanchez¹, Matthieu Vallat¹, Omar Abou Khaled², Elena Muggellini²

¹ Université de Genève, TECFA, 1211 Genève 4, Suisse
Catherine.bonnat@unige.ch; Eric.sanchez@unige.ch;
matthieu.vallat@unifr.ch;

² Université de Fribourg, HumanTech (HELA), 1700 Fribourg, Suisse.
omar.aboukhaled@hefr.ch; elena.muggellini@hefr.ch

Résumé. Cet article porte sur la conception d'un *compagnon digital* pour aider à l'orchestration d'un jeu. Intégré dans un jeu (*Geome*) qui allie éléments tangibles et numériques, ce compagnon repose sur la collecte de traces multimodales mises à disposition du maître du jeu et des joueurs. Conçu selon une recherche de type orientée par la conception, le prototype présente une interface maître du jeu dédiée au paramétrage et au suivi de l'activité.

Mots-clés : Maître du jeu, assistant, traces multimodales, orchestration.

Abstract. This article deals with the design of a digital companion to help orchestrate a game. Integrated in a game (*Geome*) that combines tangible and digital elements. This companion is based on the collection of multimodal traces made available for the game master and the players. Designed according to design-based research, the prototype presents a game master's interface dedicated to the configuration and the monitoring of the activity.

Keywords: Game-master, assistant, multimodal tracking, orchestration.

1 Introduction

L'intégration des jeux dans un contexte d'éducation, bouscule la posture de l'enseignant et suscite encore de nombreuses questions. En effet, dans ce type d'activité, les postures des élèves et de l'enseignant évoluent, respectivement en joueurs et maître du jeu [1]. Bien que de nombreuses études aient été menées pour outiller l'enseignant dans la gestion, l'animation et le suivi de l'activité avec le développement de tableaux de bords [2] et d'outils d'orchestration [3], la question reste encore peu référencée dans une situation de jeu. L'accompagnement des joueurs par le maître du jeu pendant l'activité nécessite un suivi précis de l'activité en contexte de jeu hybride du fait de la multiplicité des interactions possibles avec les éléments tangibles et numériques du jeu.

Notre recherche questionne le traçage de l'activité des joueurs dans un environnement mixte, et les types d'informations à collecter pour aider le maître du jeu dans le suivi de l'activité. Nous contextualisons le développement de ce prototype au jeu *Geome* conçu dans le cadre d'un projet adjacent. *Geome* est un jeu d'équipe qui se joue sur tablette, en interaction avec les éléments tangibles (faune, flore) du musée de la Nature à Sion (Suisse). Ce musée se déploie sur plusieurs salles et sur deux étages. Le jeu se déroule en deux parties distinctes suivies de deux débriefings. L'univers et les mécaniques de jeu traduisent une métaphore d'un changement de rapport de l'Homme à la nature. Le joueur adopte deux postures successives, celle de chasseur puis d'enquêteur à travers la résolution d'énigmes qui peuvent varier selon les joueurs. Dans ce contexte, et relativement aux contraintes générées par le musée et le jeu hybride (*Geome*), nous présentons *DigitComp*¹, un prototype de compagnon digital qui accompagne le maître du jeu (médiateur du musée) dans l'orchestration et le suivi de l'activité. Ce prototype est le résultat d'un processus de co-conception entre chercheurs, praticiens et informaticiens.

2 Cadre théorique et questions de recherche

2.1 Traces multimodales et suivi de l'activité par le maître du jeu

L'importance de l'enseignant, devenu maître du jeu, est peu évoquée dans la littérature du domaine. On peut néanmoins relever les travaux de Molin [4] qui souligne que ce rôle ne consiste pas simplement à s'assurer que les règles du jeu sont respectées. En effet, il est attendu que l'enseignant orchestre le jeu ce qui suppose d'une part qu'il dispose d'outils lui permettant d'effectuer un suivi du joueur, et d'autre part d'interagir avec les joueurs.

Le suivi du joueur peut être basé sur la collecte de traces numériques d'interactions qui sont des enregistrements automatiques des actions effectuées par le joueur. La collecte, le traitement et l'analyse de ces traces sont parfois qualifiés d'analytique du jeu (*playing analytics*) [5]. Le traçage peut également concerner des paramètres physiologiques qui peuvent avoir un grand intérêt en tant qu'indicateurs d'état émotionnels dont on sait qu'ils influencent le comportement du joueur et l'apprentissage [6].

Ces traces d'interaction peuvent être collectées et synchronisées à partir de sources multiples (vidéos, logs, audio, gestuel, capteurs physiologiques...), et sont qualifiées de multimodales [7]. Les informations nécessaires au maître du jeu peuvent alors être fournies sous la forme d'un tableau de bord qui se présente comme une interface permettant d'afficher des informations pouvant être appréhendées rapidement [8].

Les travaux qui se sont attelés à modéliser le maître du jeu [9] insistent sur l'importance de points de passage (*waypoint*) qui, pour les jeux de rôle, sont des moments clefs de la narration au cours desquels le maître du jeu doit prendre des décisions. Cela suppose qu'il dispose, outre des outils de suivi, d'outils lui permettant d'interagir avec les

¹ Projet SPARK (12 mois) soutenu par le Fonds National Suisse

joueurs. Ainsi, les travaux que nous menons visent à définir les outils de suivi et d'interaction avec les joueurs qui doivent être mis à la disposition de l'enseignant afin qu'il soit en mesure, en tant que maître du jeu, d'orchestrer le jeu.

2.2 Question de recherche

L'objectif de cet article est de présenter comment le prototype *DigitComp* permet au maître du jeu, d'orchestrer le jeu dans un contexte complexe. En effet, *DigitComp* a été développé pour *Geome* [10], un jeu au cours duquel, les joueurs interagissent avec des éléments tangibles et numériques au sein d'un musée. Ce prototype est le résultat d'un processus de conception interdisciplinaire pour lequel nous cherchons à identifier les informations utiles au maître du jeu pour le suivi du joueur en temps réel et les interactions établies au cours du jeu. Nous nous interrogeons donc sur les informations du tableau de bord et souhaitons, dans ce contexte, identifier (1) les traces à collecter (2) la manière dont les indicateurs sont visualisés et (3) le mode d'utilisation de ces traces.

3 Méthode

La méthode de conduite de notre recherche utilisée dans le projet est de type ROC [11], et repose sur l'articulation de deux processus : la co-conception du prototype et sa co-évaluation en conditions écologiques c'est-à-dire testé au musée avec des élèves du cycle d'orientation (12-15 ans). Cette recherche itérative intègre chercheurs, informaticiens et praticiens qui participent, selon les objectifs aux différentes phases des processus. Dans cet article nous nous focalisons sur le résultat du processus de co-conception (Fig. 1) avec le développement du prototype *DigitComp*.

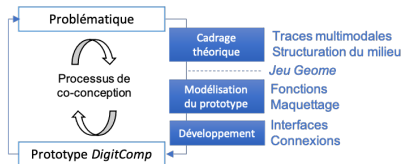


Fig. 1. Étapes de co-conception du prototype *DigitComp*

La conception a fait l'objet de 16 workshops organisés en trois phases.

La première phase consiste à positionner notre problématique de recherche selon les travaux existants. Cette étape aboutit au choix d'un cadre théorique [11] nécessaire à la modélisation du maître du jeu ainsi que d'un jeu. Le choix s'est porté sur *Geome* [10] qui présente l'intérêt que le joueur interagit avec des éléments tangibles et numériques. Intégrant les spécificités du *gameplay* (mécaniques, interactions, interfaces), et relativement à nos objectifs de recherche, la deuxième étape s'est focalisée sur la recherche et la description fine des fonctions du *DigitComp*, qui permettent le suivi des joueurs et

les interactions avec ces derniers. Cette étape permet l'identification des traces à récolter, la manière de les rendre accessibles au maître du jeu pour le suivi synchrone de l'activité des joueurs ainsi que la manière de les utiliser.

Enfin, la troisième phase, comprend à la fois le développement des interfaces pour le maître du jeu, mais aussi la connexion de l'ensemble des éléments du jeu, de manière à collecter et visualiser les traces d'interaction.

4 Résultats et discussion

Intégré dans l'univers et le scénario du jeu *Geome*, *DigitComp* offre des interfaces et modules dédiés au maître du jeu afin d'orchestrer la partie, depuis sa conception jusqu'au débriefing final. Nous commençons par présenter l'ensemble des fonctions qui lui sont dédiées, avant de détailler plus précisément le suivi de l'activité des joueurs.

4.1 Un compagnon digital pour orchestrer le jeu

DigitComp constitue une interface entre le maître du jeu et les joueurs selon le moment de jeu. En amont de la partie, il offre au maître du jeu une interface dédiée à la configuration du jeu. Basée sur une modélisation issue d'une description fine de l'activité proposée aux joueurs, cette interface permet de modifier les variables du jeu définies par défaut. Le maître du jeu peut par exemple, définir/modifier les objectifs de la partie, choisir sa durée, attribuer un type d'énigme par groupe de joueurs selon leur niveau, ainsi que configurer des feedbacks automatiques destinés à l'ensemble des joueurs. Par exemple, un message automatique peut être envoyé si l'état de santé de « l'arbre de vie », qui symbolise dans le jeu les points gagnés ou perdus, atteint un seuil que le maître du jeu considère comme critique. Actuellement, ces variables de jeu ne peuvent être modifiées au cours de la partie.

Pendant la partie, le maître du jeu dispose d'un module de clavardage de manière à accompagner des groupes de joueurs ou l'ensemble des joueurs représentés par l'avatar qu'ils ont configurés au début de la partie. Le module est intégré dans l'univers de jeu *Geome* si bien que le maître du jeu endosse le rôle d'un des personnages du jeu de manière à maintenir les joueurs dans l'univers du jeu. Ainsi, à tout moment, le maître du jeu peut apporter une aide différenciée selon les groupes de joueurs, en leur donnant par exemple un indice pour la résolution de leur enquête. Les joueurs, qui évoluent en équipe, sont avertis du message par l'intermédiaire d'une montre connectée, et prennent connaissance du contenu du message sur la tablette. Ils peuvent également à tout moment répondre et/ou solliciter le maître du jeu par l'intermédiaire de ce module. Ce type de communication répond à la contrainte spatiale des salles du musée qui se déploient sur deux étages, et permet une aide différenciée selon les équipes.

Afin d'accompagner au mieux les joueurs, le maître du jeu dispose d'une interface de suivi de l'activité en mode synchrone et, à l'issue de la partie, d'un tableau de bord sur lequel la phase de débriefing du jeu peut s'appuyer [12].

4.2 Un compagnon dédié au suivi de l'activité

DigitComp permet le suivi des joueurs. Chaque groupe de joueurs dispose d'une tablette et d'une montre connectée. Cette fonction d'observation repose sur une récolte de traces multimodales [7] stockées au format JSON sur une base de données créée spécifiquement pour le jeu. La communication entre les tablettes (joueurs, maître du jeu) et les montres connectées repose sur le réseau WIFI du musée.

Les traces de l'activité des joueurs réfèrent à l'ensemble des interactions des joueurs avec la tablette, mais aussi les interactions qu'ils peuvent avoir avec les éléments du musée. A cela, s'ajoutent des données physiologiques complémentaires, collectées par la montre connectée. Ces traces multimodales sont à la fois à disposition du chercheur (données brutes), mais aussi pour certaines d'entre elles à disposition du maître du jeu sur une interface dédiée au suivi et au bilan de l'activité [2].

Concernant le maître du jeu, il dispose de différents types de traces qui ont été, au préalable, traitées de manière à les rendre accessibles et visibles. Ces traces peuvent traduire des variables du jeu comme par exemple l'évolution de l'état de santé de « l'arbre de vie » selon les gains ou pertes de points. Elles peuvent également rendre compte des actions réalisées par les joueurs sur la tablette, qui peuvent être ou non en interaction avec les éléments tangibles du musée. Par exemple, au cours du jeu, les élèves sont amenés à enquêter sur des éléments du musée (animaux naturalisés). L'utilisation, dans le jeu, de la fonction « scan » permet au maître du jeu de connaître l'éléments avec lequel le joueur interagit, puisque chaque élément du musée est identifié par un QR code particulier. Enfin, relativement à la question des émotions qui peuvent être suscitées par le jeu [7], une donnée physiologique complémentaire (fréquence cardiaque) est également recueillie par l'intermédiaire de la montre connectée qui est portée par un des joueurs de chaque groupe.

Sur l'interface de suivi, le maître du jeu peut accéder aux traces d'activité d'un seul joueur, d'un groupe de joueurs ou de l'ensemble des groupes.

Les données qui lui sont accessibles sont actuellement le résultat d'un traitement élémentaire qui porte sur la nature et le nombre d'actions réalisées (par ex. nombre de « scan » du loup). Ces traces à destination du maître du jeu ont ainsi été sélectionnées et traitées de manière à rendre compte d'une partie de l'activité des joueurs, selon un modèle de traces conçu lors de la modélisation du prototype, à partir de l'analyse didactique de la situation de jeu. Nous ne le détaillons cependant pas dans cet article.

Enfin, du côté des joueurs, ces derniers disposent également d'un tableau de bord qui traduit le bilan de la partie (pour l'équipe), et ils peuvent suivre, à tout moment, sur la montre connectée, des variables de jeu telles que l'état de santé de leur arbre de vie.

La principale fonction du *DigitComp* repose ainsi sur d'observation de l'activité des joueurs à partir d'une collecte de traces multimodales. Cependant, ces traces traduisent une partie de l'activité des joueurs et pourrait être complétée par d'autres types de traces telles que celles laissées volontairement par les joueurs (par ex. des enregistrements audio ou textes) lors de la résolution de l'enquête dans *Geome*. Cependant, la combinaison des traces, réduisant certes la part interprétative sur ce que fait l'apprenant, reste complexe, si bien qu'une aide doit être apportée au maître du jeu de manière à lui fournir des éléments didactiques et comportementaux pertinents pour sa prise de décision.

Cette perspective d'aider le maître du jeu dans sa prise de décision fait l'objet d'un nouveau projet de recherche.

5 Perspectives

Dans cet article nous avons présenté les résultats d'un processus de co-conception à travers la description du compagnon digital qui vise à accompagner un maître du jeu à paramétrer et suivre l'activité des joueurs dans un contexte de jeu mixte en éducation.

DigitComp constitue ainsi un premier prototype qui a été testé en février 2021 au musée de la Nature (Sion, Valais Suisse) avec des élèves du cycle d'orientation (12-15 ans). Cette phase de co-évaluation permettra de réviser le prototype et les modèles théoriques sur lesquels la phase de co-conception s'appuie. Intégré au jeu *Geome* qui s'utilise spécifiquement dans ce musée, le modèle de prototypage sous-jacent peut être néanmoins être transposé à d'autres situations de jeu mixte.

References

1. Sanchez, E., Romero, M. : Apprendre en jouant. Paris : Retz (2020).
2. Dabbebi, I. : Conception et génération dynamique de tableaux de bord d'apprentissage contextuels. Thèse de doctorat, Université du Maine (2019).
3. Dillenbourg, P.: Design for classroom orchestration. *Computers & Education* 69, 485-492, (2013).
4. Molin, G.: The Role of the Teacher in Game-Based Learning: A Review and Outlook, in *Serious Games and Edutainment Applications.*, M. Ma and O. A, Editors. Springer, Cham (2017).
5. Peraya, D., Luengo, V. : Les Learning Analytics vus par Vanda Luengo, *Distances et médiations des savoirs* 27, (2019).
6. Lajoie, S. P., Pekrun, R., Azevedo, R., Leighton, J. P.: Understanding and measuring emotions in technology-rich learning environments. *Learning and Instruction* 70, (2020).
7. Blikstein, P.: Multimodal learning analytics. In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge* 102, (2013).
8. Few, S.: *Information Dashboard Design: Displaying data for at-a-glance monitoring.* Analytics Press (2013).
9. Tychsen, A., et al.: Modeling Game Master-based story facilitation in multi-player Role-Playing Games. In: *Proceedings of the 2009 AAAI Symposium on Intelligent Narrative Technologies*, pp. 24-32 (2009).
10. Bonnat, C., Oliveira, G., Sanchez, E. : Geome, un jeu pour comprendre l'Anthropocène lors de visites scolaires au musée. *Revista de la Asociacion para la Ensenanza de las Ciencias de la Terra*, 28(1), 89 – 98 (2020).
11. 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, *Éducation et didactique* 9(2), 73-94 (2015).
12. Bonnat, C., Sanchez, E., Paukovic, E., Kramar, N.: Didactic transposition and learning game design. Proposal of a model integrating ludicization, and test in a school visit context in a museum. In J. K. Almqvist, K (Ed.), *Didactics in a Changing World. European Perspectives on Teaching, Learning and the Curriculum: EERA Book Series* (sous-presse).

Vers une méthodologie de co-conception de *learning games* par les étudiants en soins infirmiers

Sebastian Gajewski, Nour El Mawas et Jean Heutte

Univ. Lille, ULR 4354 - CIREL - Centre Interuniversitaire de Recherche en Éducation de Lille,
F-59000 Lille, France
{sebastian.gajewski.etu, nour.el-mawas, jean.heutte}
@univ-lille.fr

Résumé. De nos jours, les erreurs médicales sont une des causes principales d'incidents et de décès liées aux faibles compétences des infirmières en raisonnement clinique. Les formateurs des Instituts de Formation en Soins Infirmiers (IFSI) cherchent alors à développer des méthodes pédagogiques permettant aux étudiants en soins infirmiers de renforcer ces compétences. De nombreuses études montrent l'impact des jeux sérieux sur l'apprentissage des étudiants, notamment depuis la disponibilité des outils de conception de jeu faciles à utiliser, même par ceux qui n'ont aucune compétence technique. L'apprentissage par le jeu comprend l'apprentissage par le simple fait de jouer et l'apprentissage par la conception de jeu. Cet article s'intéresse à l'apprentissage par la conception de jeu et identifie une méthodologie permettant aux étudiants de co-concevoir, eux-mêmes, un *learning game* dans le domaine du raisonnement clinique. Cette méthodologie est composée de 11 étapes, commençant par l'identification des objectifs pédagogiques jusqu'à l'évaluation du jeu.

Mots-clés : Apprentissage, Co-conception, *Learning game*, Étudiant en soins infirmiers, Motivation.

Abstract. Nowadays, medical errors are one of the leading causes of injuries and deaths related to poor clinical reasoning skills of nurses. Teachers in nursing schools look for designing new pedagogical methodologies, giving the nursing students the opportunity to improve these skills. Many studies highlight the effectiveness of serious games on students' learning, notably through the availability of game design tools, that are easy-to-use, even by people without any technical skill. Game-based learning includes gameplay-based learning and game design-based learning. This paper focuses on game design-based learning and identifies a methodology to enable students to codesign, themselves, a learning game in the clinical reasoning field. This methodology consists of 11 steps, starting from the identification of the pedagogical objectives until the game assessment.

Keywords: Learning, Co-design, Learning game, Nursing student, Motivation.

1 Introduction

De nos jours, plus de personnes meurent des conséquences d'une erreur médicale que d'un accident de la voie publique, d'un cancer du sein ou du SIDA [1]. Ces erreurs médicales sont liées aux faibles compétences des infirmières en raisonnement clinique [2]. Ainsi, il est nécessaire de renforcer l'apprentissage du raisonnement clinique en formation infirmière pour réduire la prévalence de ces erreurs médicales.

Selon Kafai [3], il existe 2 méthodes distinctes d'apprentissage par le jeu : l'instructionnisme (l'apprentissage par le simple fait de jouer) et le constructionnisme (l'apprentissage par la conception de jeu). Dans l'approche instructionniste, les apprenants jouent à un jeu pour apprendre. Tandis que dans l'approche constructionniste, les étudiants apprennent en concevant leur propre jeu.

Dans cet article, nous nous focalisons sur l'apprentissage par la co-conception de jeu en formation infirmière. Particulièrement, nous sommes intéressés par les étapes importantes dans l'apprentissage par la co-conception de jeu et l'évaluation de la motivation et de l'apprentissage des étudiants à travers cette co-conception.

Cet article est structuré comme suit. La section 2 présente des méthodologies de conception de jeu et décrit des travaux existants sur l'apprentissage par la conception de jeu. La section 3 conclut ce travail de recherche et propose des perspectives.

2 Apprentissage par la conception de jeu

Ce travail de recherche a pour objectif d'identifier une méthodologie avec des étapes claires pour permettre aux étudiants en soins infirmiers de concevoir, eux-mêmes, un *learning game* dans le domaine du raisonnement clinique. Pour cela, nous passons en revue des méthodologies et des travaux existants sur la conception de jeu.

2.1 Dispositifs de conception de jeu

Dans cette partie, nous présentons 3 dispositifs de conception de jeu.

Les 6 facettes du jeu sérieux est un outil d'analyse de la conception, facilitant l'équilibre entre motivation et apprentissage en articulant 2 formes d'expertise : une expertise pédagogique (experts du domaine, experts en pédagogie...) s'attachant à l'apprentissage, et une expertise ludique (*game designers*, *level designers*, scénaristes, graphistes, programmeurs...) s'intéressant à la motivation [4]. Cet outil a pour but d'analyser le *learning game* sous différents angles : (1) objectifs pédagogiques, (2) simulation du domaine, (3) interactions avec la simulation, (4) problèmes et progression, (5) décorum et (6) conditions d'utilisation.

LEGADEE (*LEarning Game DEsign Environment*) est un modèle de conception de *learning games* [5] qui s'appuie sur un processus de conception en 7 étapes : (1) Besoins client, (2) Spécification des objectifs pédagogiques, (3) Conception, (4) Contrôle qualité, (5) Réalisation, (6) Test sur public cible, (7) Utilisation et maintenance. Pour structurer les indicateurs en catégories facilement compréhensibles par les concepteurs, LEGADEE se base sur les 6 facettes du jeu sérieux mais propose une adap-

tation légèrement modifiée : (1) objectifs d'apprentissage, (2) interactions, (3) problèmes et progression, (4) décorum, (5) conditions d'utilisation, (6) coût prévisionnel.

ARGILE (*Architecture for Representations, Games, Interactions and Learning among experts*) est un guide de conception, simple d'utilisation, permettant de concevoir des jeux sérieux participatifs et intensifs en connaissances [6]. Elle permet l'addition, la modification et la discussion de nouveaux objets, connaissances et règles. Nous nous y sommes intéressés parce que les concepteurs ne sont pas des *game designers*. En effet, cette méthodologie compte 2 communautés : la communauté des concepteurs (experts du domaine) et la communauté des joueurs (apprenants).

Maintenant que nous avons présenté quelques exemples de dispositifs de conception de jeu existants, nous proposons de passer en revue des travaux sur la conception de jeu pour avoir un panorama complet des étapes de co-conception utilisées.

2.2 Travaux existants sur l'apprentissage par la conception de jeu

Dans notre travail de recherche, nous sommes intéressés par l'apprentissage par la conception de jeu. Ainsi, nous présentons des études existantes qui ont chacune pour objectif de proposer une approche de conception de jeu et d'évaluer leur efficacité sur l'apprentissage. Cette présentation nous permettra d'identifier et de détailler, dans la discussion, les étapes importantes de la co-conception de jeu.

Dans [7], 16 lycéens, inscrits au programme *Game Design through Mentoring and Collaboration* (GDTMC), ont utilisé *GameMaker* pour co-concevoir un jeu sur l'immunologie et la neurologie afin de les encourager à s'intéresser aux contenus de sciences, technologie, ingénierie et mathématiques. Les élèves étaient répartis en 4 groupes de 4 élèves. D'abord, les élèves, qui étaient inscrits à une formation durant l'année scolaire, avaient appris la modélisation et l'animation 3D, la programmation informatique, et la conception de jeu vidéo à raison d'1 session de 2 heures par semaine, le samedi, pendant 10 semaines. Ensuite, les élèves, souhaitant continuer, ont conçu un jeu, durant l'été, 4 heures par jour pendant 4 semaines. Durant la première semaine, un expert du domaine leur a transmis des notions en neurologie. Ensuite, les élèves ont travaillé sur leur projet de conception de jeu pendant 3 semaines. La classe a été dirigée par un formateur avec des connaissances et des compétences en programmation de jeux. Le formateur a aidé les élèves à se concentrer sur l'activité de conception de jeu. Trois mentors, qui étaient d'anciens participants du programme GDTMC et/ou qui étaient inscrits à l'Université dans une filière liée à la technologie, ont été recrutés pour aider les élèves sur les aspects techniques de leur jeu. Avant la programmation du jeu, les élèves ont créé un schéma de leur jeu. Les résultats ont montré que la conception de jeu favorisait l'apprentissage à travers 3 processus. D'abord, concevoir un jeu conduisait l'élève à mettre en lumière ses manques de connaissances et à chercher à combler les lacunes. Ensuite, concevoir un jeu contribuait à développer un sens élevé des responsabilités pour rendre le jeu attractif, engageant et scientifiquement fiable. Enfin, les élèves étaient capables d'exprimer verbalement leur compréhension des notions qu'ils avaient introduites dans le jeu.

Dans [8], 64 collégiens ont utilisé *Scratch* pour co-concevoir un jeu afin d'apprendre des opérations mathématiques telles que les multiplications et les divi-

sions. Les élèves étaient répartis en 10 groupes de 6 à 7 élèves. Avant de concevoir leur jeu, les élèves ont joué à une série de jeux de mathématiques. Ensuite, ils ont participé à 3 sessions de formation d'1 heure sur l'utilisation de *Scratch*. Les élèves étaient encouragés à jouer aux jeux disponibles sur le site internet communautaire de *Scratch*. Ainsi, ils pouvaient découvrir les fonctionnalités de *Scratch*. Les élèves ont travaillé sur leur projet de conception de jeu durant 2 sessions d'1 heure par semaine, pendant 6 semaines. Durant l'activité de conception de jeu, 5 étudiants, qui étaient inscrits à un cours de conception de *learning games*, ont apporté du soutien aux collégiens. Ils ont répondu aux questions liées aux notions de mathématiques et de conception de jeu. Ils leur ont apporté des *feedbacks* et les ont encouragés. Le processus de conception de jeu comprenait la réalisation d'un prototype papier. Durant le prototypage papier, les élèves travaillaient et menaient un *brainstorming* ensemble, partageaient leurs idées de conception de jeu, négociaient entre eux, et s'expliquaient les notions de mathématiques. Ils pouvaient programmer leur jeu, une fois le prototype papier réalisé. Après le prototypage papier, les membres des équipes se sont réparti les tâches. Des sous-groupes se sont créés dans lesquels des binômes d'élèves travaillaient ensemble sur leurs actions de conception de jeu, telles que la recherche de ressources par internet (telles que les images et la musique), la programmation de jeu, l'expérimentation et l'amélioration du jeu. Durant la dernière session de conception de jeu, chaque groupe a publié son jeu sur le site internet communautaire de *Scratch* et l'a présenté aux autres élèves, qui, en retour, ont apporté un *feedback*. 52% des élèves ont rapporté avoir appris des notions de mathématiques en concevant leur jeu. Pour évaluer l'attitude des élèves à l'égard des mathématiques, les élèves ont répondu au questionnaire *Attitudes Towards Math Inventory* (ATMI) avant et après l'activité de conception de jeu. Les résultats ont montré que la moyenne des scores de l'ATMI obtenus après l'activité de conception de jeu avait significativement augmenté comparativement à la moyenne avant l'activité de conception de jeu (Différence de moyenne = +2,56 ; $p = .01$).

Dans [9], 28 collégiens ont utilisé *BlockStudio* pour co-concevoir un jeu afin d'apprendre des opérations mathématiques telles que les puissances. Les élèves étaient répartis en 10 groupes de 2 à 3 élèves. D'abord, les élèves ont participé à une session préparatoire sur la conception de jeu, composée de 3 séances de 45 minutes permettant d'introduire le projet, d'expliquer comment concevoir un jeu et de présenter *BlockStudio*. Un mois plus tard, les élèves ont travaillé sur leur projet de conception de jeu pendant 11 sessions de 45 minutes. Chaque groupe devait concevoir son jeu en renseignant un *Game Design Document* (GDD). Un GDD est un document descriptif qui est renseigné durant l'activité de conception de jeu. Il comprend des caractéristiques telles que l'univers du jeu, l'histoire du jeu, les personnages, le *scenarrio*, les règles, les sons, la musique, l'interface utilisateur, les commandes de jeu, *etc.* Le dernier jour, chaque équipe a testé les jeux des autres élèves et leur a apporté des *feedbacks*. Ainsi, chaque équipe pouvait modifier et améliorer son jeu. Sur une échelle de *Likert* à 4 points, les élèves ont rapporté avoir compris les notions de mathématiques (Moyenne = 3,66). Les élèves étaient motivés par le projet. En effet, ils ont indiqué avoir travaillé pendant leur temps libre. Les enseignants ont également rapporté que les élèves étaient motivés (Moyenne = 3,5).

2.3 Discussion

Les dispositifs de conception de jeux, présentés dans la section 2.1, ne répondent pas aux besoins de notre recherche étant donné qu'ils ne sont pas orientés sur un concepteur et développeur qui seraient un apprenant. Néanmoins, nous nous en sommes inspirés pour identifier certaines étapes clés telles que l'identification des objectifs pédagogiques, la conception et les tests. Nous avons aussi identifié des acteurs importants tels que l'expert pédagogique, le *game designer* et l'apprenant.

Tableau 1. Méthodologie de co-conception de jeu.

Étapes	Actions	Acteurs
1	Identifier les objectifs pédagogiques	En
2	Identifier le logiciel de conception de jeu	G
3	Identifier des jeux dans un domaine similaire	G
4	Jouer aux jeux dans le domaine similaire pour s'en inspirer	Ét
5	Fournir le contenu pédagogique aux étudiants	En
6	Lire, regarder, écouter, comprendre le contenu pédagogique	Ét
7	Enseigner comment concevoir un jeu	G
8	Enseigner comment utiliser le logiciel de conception de jeu	G
9	Co-concevoir le jeu	Ét
10	Co-implémenter le jeu	Ét
11	Évaluer le jeu	C-En-Ét-G

C : Chercheur / En : Enseignant / Ét : Étudiant / G : *Game designer*

Les travaux de recherche examinés dans la section 2.2 nous ont permis d'identifier 11 étapes de co-conception de jeu (tableau 1). Cette méthodologie implique 4 acteurs différents : l'enseignant, le *game designer*, les étudiants et le chercheur. Chaque acteur réalise différentes actions. D'abord, dans l'étape 1, l'enseignant identifie les objectifs pédagogiques [4, 5, 7, 8, 9]. Dans l'étape 2, le *game designer* identifie le logiciel de conception de jeu selon différents critères (contraintes, caractéristiques des étudiants...) [7, 8, 9]. Dans l'étape 3, le *game designer* identifie des jeux dans un domaine similaire [8]. Dans l'étape 4, les étudiants jouent aux jeux dans le domaine similaire pour s'en inspirer pour leur propre jeu [8]. Dans l'étape 5, l'enseignant, qui est expert du domaine, fournit les contenus pédagogiques [7]. Dans l'étape 6, les étudiants lisent, regardent et écoutent les contenus pédagogiques [7, 9]. Ils partagent leur compréhension du contenu pédagogique. Dans l'étape 7, le *game designer* enseigne aux étudiants comment concevoir un jeu [9]. Dans l'étape 8, le *game designer* enseigne aux étudiants comment utiliser le logiciel de conception de jeu choisi par le *game designer* [8, 9]. Durant cette étape, le *game designer* explique, entre autres, l'importance de trouver un équilibre entre les contenus pédagogiques et les aspects ludiques à introduire dans le jeu. Il signale l'importance d'apporter des *feedbacks* immédiats aux joueurs car cela favorise l'état de *flow* [10]. Dans l'étape 9, les étudiants co-conçoivent le jeu [5, 7, 8, 9]. Dans l'étape 10, les étudiants co-implémentent

leur jeu [5, 7, 8, 9]. Enfin, dans l'étape 11, le jeu est évalué [5, 7, 8, 9]. Les étudiants testent les jeux des autres étudiants et leur apportent des *feedbacks*. Le *game designer* évalue les mécaniques de jeu introduites dans les jeux. L'enseignant évalue l'exactitude des contenus pédagogiques introduits dans les jeux. Le chercheur évalue la pertinence et l'efficacité de l'activité pédagogique.

3 Conclusion et perspectives

Dans cet article, nous avons identifié une méthodologie permettant aux étudiants en soins infirmiers de co-concevoir un *learning game*. Cette méthodologie, composée de 11 étapes, implique 4 acteurs (enseignant, *game designer*, étudiants et chercheur).

Cette méthodologie sera appliquée à IFSanté, un IFSI situé en France. La première évaluation concernera approximativement 135 étudiants en soins infirmiers de 2^{ème} année. Les étudiants seront répartis en 3 groupes. Un 1^{er} groupe (expérimental) co-concevra un jeu. Un 2^{ème} groupe (contrôle) recevra un enseignement classique. Un 3^{ème} groupe (contrôle) jouera au jeu développé. Notre objectif est d'évaluer la différence d'apprentissage et de motivation entre les étudiants qui co-conçoivent le jeu, ceux qui reçoivent un enseignement plus classique, et ceux qui y jouent.

Références

1. Kohn, L. T., Corrigan, J., Donaldson, M. S.: To err is human: Building a safer health system. National Academy Press (2000).
2. Levett-Jones, T., Hoffman, K., Dempsey, J., Jeong, S. Y.-S., Noble, D., Norton, C. A., Roche, J., Hickey, N.: The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients. *Nurse Education Today*, 30(6), 515-520 (2009). <https://doi.org/10.1016/j.nedt.2009.10.020>.
3. Kafai, Y. B.: Playing and making games for learning: Instructionist and constructionist: Perspectives for games studies. *Games and Culture Volume 1 Number 1*, 36-40 (2006).
4. Marne, B., Huynh-Kim-Bang, B., Labat, J.-M.: Articulier motivation et apprentissage grâce aux facettes du jeu sérieux. EIAH 2011 - Conférence sur les Environnements Informatiques pour l'Apprentissage Humain, May 2011, Mons, Belgique. pp.69-80.
5. Marfisi-Schottman, I.: Méthodologie, modèles et outils pour la conception de Learning Games (2012).
6. El Mawas, N.: Designing learning scenarios for serious games with ARGILE. *Knowledge Management & E-Learning*, 6(3), 227-249 (2014).
7. Khalili, N., Sheridan, K., Williams, A., Clark, K., Stegman, M.: Students designing video games about immunology: Insights for science learning. *Computers in the schools* 28, 228-240 (2011).
8. Ke, F., Im, T.: A case study on collective cognition and operation in team-based computer game design by middle-school children. *International Journal of Technology and Design Education* 24(2), 187-201 (2013).
9. Martins, A., Oliveira, L.: Educational video game design by 8th graders: investigating processes and outcomes. 12th European Conference on Games Based Learning (2018).
10. Heutte, J.: Les fondements de l'éducation positive : Perspective psychosociale et systémique de l'apprentissage. Dunod, Paris (2019).

Aide au suivi de la progression de groupes d'apprenants pour la mise en place d'une pédagogie différenciée

Fatima Harrak et François Bouchet

Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
fatima.harrak@lip6.fr
francois.bouchet@lip6.fr

Résumé. Ce travail se situe dans le cadre d'un assistant pédagogique à destination des professeurs des écoles pour accompagner leurs élèves de primaire en mathématiques. Pour proposer à l'enseignant une vision synthétique des activités de sa classe, nous utilisons des méthodes de clustering pour constituer automatiquement des groupes d'élèves de sa classe en fonction des activités réalisées dans l'application, où chaque élève suit un parcours individuel. Après une présentation des traces utilisées et de la qualité intrinsèque des clusters produits (coefficient de silhouette), nous analysons les retours d'une enquête auprès de 32 enseignants utilisateurs pendant au moins 6 semaines. Ces analyses ouvrent des perspectives quant aux améliorations possibles des clusters et de leur visualisation.

Mots-clé: Clustering · traces d'activités · tableau de bord.

Abstract. This work deals with a pedagogical assistant for primary school teachers to assist students in mathematics. To give the teacher a synthetic vision, we use clustering methods to automatically provide groups of students in their class based on their activities in the application, where each student follows a different path. We first present the traces used and the intrinsic quality of the clusters (Silhouette coefficient), and then the answers to a questionnaire filled by 32 teachers who have used the application for at least 6 weeks. These analyses open perspectives on how to improve the clusters and their visualization.

Keywords: Clustering · Students' activities · Dashboard

1 Introduction

La pédagogie différenciée est “une pédagogie fondée sur la différenciation des processus d'apprentissage” [8] dont la mise en oeuvre pour l'enseignant peut être complexe car elle suppose de différencier les contenus et de proposer des structures de travail dynamiques adaptées aux savoirs et lacunes de chacun [7].

Grâce aux algorithmes d'apprentissage adaptatif, de plus en plus d'applications proposent un parcours personnalisé à chaque élève. Cependant cela complique

pour l'enseignant le suivi des parcours individuels. Il faut donc lui proposer une vision synthétique (via un tableau de bord) du travail effectué par ses élèves, mais aussi lui permettre de déterminer des groupes d'élèves avec des lacunes communes auxquels proposer une activité de remédiation commune. Notre objectif est donc de répondre aux questions de recherche suivantes : (QR1) Comment regrouper des élèves ayant des parcours d'apprentissage hétérogènes dans des dispositifs d'apprentissage ? et (QR2) Comment restituer ces groupes de manière compréhensible et acceptable par les enseignants ?

2 Travaux connexes

Des recherches antérieures ont largement proposé des restitutions aux enseignants à l'échelle de groupes via des méthodes de clustering pour aider au suivi de progression des élèves. Ainsi [3] propose un système adaptatif pour constituer des groupes d'élèves de collège de différents niveaux de compétences en lecture. Les résultats du clustering (K-Means) ont montré des corrélations positives entre les niveaux de difficulté du texte et le temps de lecture pour les 4 groupes étudiés. Li et Yoo ont proposé une chaîne de Markov basée sur le clustering pour modéliser les comportements en ligne des élèves pendant le processus d'apprentissage pour un enseignement plus efficace et adaptatif [5].

D'autres approches utilisent le clustering pour regrouper les séquences d'interactions des élèves dans les environnements hybrides [1]. McBroom et al. [6] s'appuient sur un algorithme de clustering hiérarchique pour identifier les tendances comportementales d'élèves à partir de données temporelles, notamment en termes d'abandon par types d'exercices, information qui pourrait nous intéresser pour analyser les trajectoires individuelles.

Par ailleurs, K-Means a été utilisé dans un contexte différent pour faire des regroupements d'élèves selon la nature des questions posées [4]. Dans la suite de cet article, nous utilisons également K-Means pour faire des regroupements des élèves ayant des parcours d'apprentissage très hétérogènes.

3 Contexte

3.1 L'application Adaptiv'Math

Adaptiv'Math est un assistant pédagogique numérique en mathématiques pour les enseignants et élèves du cycle 2 (CP à CE2). Il associe deux algorithmes d'intelligence artificielle : (1) pour des parcours individualisés (ZPDES [2]) et (2) pour synthétiser les expériences individuelles de chaque élève. L'enseignant contrôle ces algorithmes via son tableau de bord. L'application permet aux élèves de travailler sur 5 modules de mathématiques. Chaque module est découpé en objectifs (15 à 20 par module) visant à développer une compétence particulière par des exercices de difficulté progressive, eux-mêmes découpés en 4 à 6 niveaux. Le fonctionnement de l'application au sein d'un module se résume en 5 étapes : (1) Un test diagnostic (40 questions) évalue le niveau initial de l'élève. (2) Un premier regroupement d'élèves par clustering est proposé à l'enseignant. (3) Chaque élève réalise une séquence d'exercices (cf. Figure 1) où chaque exercice

est choisit au fur et à mesure en fonction des compétences jugées acquises.

(4) Les algorithmes de clustering analysent les résultats de tous les élèves pour proposer de nouveaux regroupements.

(5) L'analyse des résultats, dont les clusters, est transmise aux enseignants via un tableau de bord. Les élèves ont également des retours d'information.

Encore en phase de déploiement, l'application a été utilisée d'octobre à décembre 2020 par 137 classes. 119 classes ont effectué le test initial sur au moins un module (étape 2 atteinte) et 102 ont travaillé sur l'application au-delà de ce test (étape 3 atteinte). Deux modules sur 5 étaient présents et nous nous concentrons dans la suite sur le module 1 (ayant donné lieu à plus d'interactions).

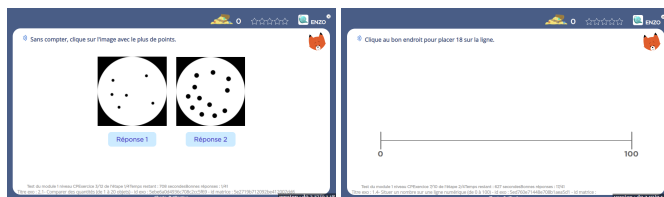


Fig. 1. Deux exemples d'exercices du module 1 sur le sens des nombres demandant de choisir le nuage avec le plus de points et de placer 18 sur une ligne allant de 0 à 100

3.2 Données collectées

Les traces d'activités de l'application sont étroitement liées aux "gameplays", c'est-à-dire un patron d'exercices (ex : QCM, ligne numérique ou glisser-déposer) portant sur différents objectifs. Au sein d'un gameplay, différents *événements* peuvent avoir lieu correspondant à des moments-clé de l'application (ex : un exercice est chargé) ou à une action de l'apprenant (ex : une réponse est sélectionnée). Chaque événement dispose d'un ensemble d'*attributs* associés. Par exemple, lorsqu'un élève termine un exercice, un événement `EvtExerciceEnded` est généré avec comme attributs associés la date et heure (timestamp), l'ID d'exercice, l'ID d'élève, le score sur l'exercice (normalisé entre 0 et 1) et le temps en millisecondes depuis l'événement `EvtExerciceStarted` associé.

Au sein du module 1, nous avons collecté 1 048 575 traces d'activités correspondant à 1 583 exercices différents réalisés par 1 534 élèves. Toutes les données remontées sont complètement anonymisées.

4 Constitution de groupes

Principes de l'algorithme. Bien qu'il soit possible d'appliquer un clustering aux données d'une classe, le nombre réduit de traces par élève limite la qualité des clusters obtenus. Nous proposons donc une approche en deux temps :

(1) phase d'*apprentissage* : on entraîne l'algorithme pour déterminer le meilleur partitionnement à partir des traces de toutes les classes à un instant T_{app}

(2) phase de *prédiction* : le partitionnement précédemment appris est appliqué pour prédire l'appartenance d'un nouvel élève à un cluster à un instant $T_{pred} > T_{app}$, selon la distance de celui-ci avec chacun des centroïdes des clusters. On utilise ainsi indirectement le clustering à des fins de prédictions, tel que fait dans [4]. La nature des exercices étant différente dans chaque module, il y a un clustering séparé pour chacun des modules de l'application.

Le premier clustering sur un module est donc réalisé exclusivement sur les données du test initial (cf. étape 2) avec les mêmes exercices travaillés par tous les élèves. Par la suite (cf. étape 4), les clusters produits dépendent de l'activité des élèves, et indirectement de la nature de ces exercices. Nous avons également intégré des heuristiques de sélection automatique du nombre de clusters et garantissant un nombre de clusters "raisonnable" pour l'enseignant (entre 2 et 5).

Prétraitement et application. Tout d'abord, il faut passer de la représentation événementielle des traces à une représentation matricielle où chaque ligne représente un élève et chaque colonne représente la valeur d'une variable donnée. Ensuite, nous avons calculé pour chaque objectif : le score moyen obtenu, la durée passée sur un exercice, le niveau de difficulté maximum atteint et le nombre d'exercices travaillés. Puis un ensemble de prétraitements est effectué : remplacement des données aberrantes, standardisation et réduction d'attributs par ACP. Enfin nous avons entraîné un algorithme K-Means++ (1000 fois pour éviter les optimums locaux) sur l'ensemble des données du module jusqu'à une date fixe en considérant les attributs ci-dessus, puis choisit le partitionnement par la méthode du coude (elbow method) en ayant fait varié le nombre de clusters de 2 à 20.

Résultats. Pour le module 1, nous avons trouvé 6 clusters globaux de référence à partir des traces d'activités de 1534 élèves (cf. Table 1). Nous avons utilisé le coefficient de silhouette (variant entre -1 et 1) pour évaluer la qualité des clusters. Quand le coefficient est proche de 0 mais positif, les clusters "se touchent" et des individus sont proches des centroïdes de deux clusters différents mais sans chevauchement. Les coefficients de silhouette des 6 clusters de référence variaient entre -0.027 et 0.21 (0.093 en moyenne). Ce score proche de 0 s'explique par le grand nombre d'attributs utilisés (40), même après l'ACP, qui rend plus difficile la séparation des clusters. On regroupe en effet des élèves effectuant des exercices différents ; sur le test initial uniquement (étape 2 - exercices identiques pour tous) les coefficients de silhouette sont logiquement supérieurs (0.23 en moyenne).

Nous avons considéré 5 variables descriptives pour caractériser les clusters (cf. Table 1) et 3 groupes distincts apparaissent : un groupe (1, 3 et 6) avec des scores et nombres d'exercices/objectifs plus élevé et moins rapides que les autres. Un autre groupe (2 et 4) à l'extrême des caractéristiques du premier groupe et un groupe (5) plus rapide mais entre les deux autres en termes de score, nombre d'exercices, etc. Le fait que 3 groupes ressortent et non 6 ne veut pas dire qu'un regroupement en 3 clusters serait plus pertinent, mais que les variables explicatives choisies ne permettent pas de faire ressortir ce qui distingue les clusters 2 et 4 par exemple. Ces variables doivent en effet être suffisamment

compréhensibles par les enseignants ce qui demande un équilibre. Dans le tableau de bord testé, seules 3 de ces 5 variables étaient présentées (cf. Figure 2).

Tableau 1. Moyennes des 6 clusters globaux du module 1 selon 5 variables descriptives

Cluster	N	Score	Durée	Nb objectifs	Nb exerc. faits	Nb exerc. réussis
cluster1	124	0.65	758.73	10.82	105.59	65.46
cluster2	503	0.55	153.04	4.64	41.56	22.47
cluster3	36	0.69	862.05	11.92	210.58	144.47
cluster4	504	0.58	127.84	5.67	49.49	28.02
cluster5	111	0.61	37.22	8.14	74.64	43.05
cluster6	256	0.66	848.76	9.05	103.11	66.11

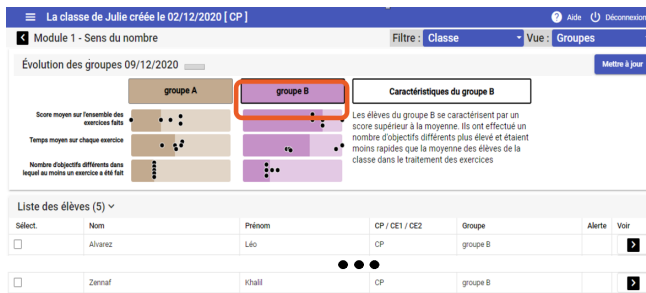


Fig. 2. Tableau de bord enseignant visualisant les clusters. Jauge = niveau moyen de chaque groupe d'élèves (ici A et B), point = un élève. Un clic sur un point le sélectionne dans la liste en-dessous, un clic sur un groupe affiche ses caractéristiques (ici le B).

5 Retour des enseignants

Pour répondre à la QR2, un questionnaire a été envoyé début décembre 2020 auquel 29 enseignants (sur 36) ont répondu afin d'évaluer l'accessibilité de la visualisation (cf. Figure 2), leur bonne compréhension de ce qui est montré et l'intérêt de cette information pour mettre en place une pédagogie différenciée.

Accessibilité. 18 enseignants déclarent être allés voir la visualisation de groupe générée par le clustering. 9 y sont parvenus facilement sans aide, 7 facilement avec les fiches fournies et 2 avec difficulté malgré l'aide. Il y a donc une difficulté d'accès à cette interface, qui a limité la visibilité des résultats proposés, d'autant que 7 des 11 non répondants déclarent avoir tenté d'y accéder sans succès.

Compréhensibilité de la visualisation des clusters. Parmi les 18 enseignants qui y ont eu accès, 5 l'ont comprise très facilement, 9 facilement avec les fiches fournies, 2 difficilement avec cette même aide, et 2 n'y sont pas parvenus faute de bien comprendre les critères retenus, soulignant le besoin d'explicitier davantage ces choix. Une majorité semble comprendre globalement la démarche et l'intérêt. Ce point ressort aussi quand les enseignants notent (de 0 à 10) l'utilité des critères de description fournis ($M = 6, \sigma = 2$) par rapport à leurs besoins.

Utilité pour une pédagogie différenciée. 25 enseignants répondent "oui" (1 "non", 3 "je ne sais pas") montrant qu'en dépit des problèmes d'accès et de lisibilité pour certains, la majorité perçoit l'intérêt du clustering et des groupes.

6 Conclusion

Nous avons présenté ici un modèle de clustering entraîné sur des sessions antérieures d'un module d'exercices de mathématiques, afin de prédire le cluster d'appartenance de nouveaux élèves à partir de leurs traces d'activités. Nous avons montré que les clusters obtenus, bien que proches dans l'espace, sont distincts et globalement explicables de manière compréhensible pour des enseignants, bien qu'il soit difficile de caractériser certains avec les variables simples utilisées. Ce choix de variables est un des points faibles actuellement mis en avant dans les retours d'enseignants pour la compréhension des groupes, et constitue donc un des points majeurs d'amélioration en collaboration avec eux. Une autre limite de notre approche est lié au démarrage à froid : il faut disposer d'un volume de données important et représentatif de l'ensemble des élèves pour constituer les clusters globaux de référence. Ces clusters peuvent ensuite être réappris au fil du temps quand les données augmentent. Actuellement, ce problème est contourné en proposant un clustering par classe (sur des données locales), non optimal, pendant les premières semaines d'ouverture d'un module.

7 Remerciements

Nous remercions EvidenceB pour la diffusion du questionnaire et le Partenariat d'Innovation sur l'IA (P2IA) de la DNE pour le financement de ce projet.

Références

1. Akpınar, N.J., Ramdas, A., Acar, U.: Analyzing student strategies in blended courses using clickstream data. arXiv preprint arXiv:2006.00421 (2020)
2. Clément, B., Roy, D., Oudeyer, P.Y., Lopes, M.: Online Optimization of Teaching Sequences with Multi-Armed Bandits (2014), <https://hal.inria.fr/hal-01016428>
3. Fadjljević, L., Maitz, K., Kowald, D., Pammer-Schindler, V., Gasteiger-Klicpera, B.: Slow is good: the effect of diligence on student performance in the case of an adaptive learning system for health literacy. In: Proc. Tenth Int. Conf. on Learning Analytics & Knowledge. pp. 112–117. LAK '20, ACM, New York, NY, USA (Mar 2020). <https://doi.org/10.1145/3375462.3375502>
4. Harrak, F., Bouchet, F., Luengo, V.: From Students' Questions to Students' Profiles in a Blended Learning Environment. *Journal of Learning Analytics* **6**(1), 54–84 (Apr 2019). <https://doi.org/10.18608/jla.2019.61.4>
5. Li, C., Yoo, J.: Modeling student online learning using clustering. In: Proceedings of the 44th annual Southeast regional conference. pp. 186–191 (2006)
6. McBroom, J., Yacef, K., Koprinska, I.: DETECT: A Hierarchical Clustering Algorithm for Behavioural Trends in Temporal Educational Data. In: Artificial Intelligence in Education. pp. 374–385. LNCS, Springer (2020). https://doi.org/10.1007/978-3-030-52237-7_30
7. Meirieu, P.: L'École, mode d'emploi: Des méthodes actives à la pédagogie différenciée. ESF éditeur, 15^{ème} edn. (2009)
8. Przesmycki, H.: Pédagogie différenciée. Hachette, Paris (1991)

Parcours adaptatifs au sein de la Plateforme Européenne d'Apprentissage Personnalisé des Langues PEAPL

Aous Karoui¹, Lionel Alvarez², Mario Ramalho³ et Thierry Geoffre⁴

¹ Université de Grenoble Alpes, 38031 Grenoble, France

Aous.karoui@univ-grenoble-alpes.fr

² Haute École pédagogique Fribourg et Université de Fribourg, 1700 Fribourg, Suisse

Lionel.alvarez@unifr.ch [0000-0003-4708-251X]

³ École des Métiers, 1700 Fribourg, Suisse

Mario.Ramalho@eduf.fr.ch

⁴ Haute École pédagogique Fribourg, 1700 Fribourg, Suisse

Thierry.Geoffre@eduf.fr.ch [0000-0003-0237-2310]

Résumé. La personnalisation des parcours d'apprentissage selon des profils de compétences et la pédagogie par le jeu sont de plus en plus adoptées par les acteurs de l'éducation grâce à leur potentiel pour optimiser l'efficacité des mesures pédaogo-didactiques. Elles se développent notamment à l'ère des plateformes numériques multisupports, enrichies par les nouvelles technologies. Cependant, l'apprentissage effectif peut varier chez les individus, notamment en fonction des besoins spécifiques (élèves allophones, dyslexiques, malentendants, etc.). Dans cet article, nous présentons tout d'abord *GamesHUB*, la plateforme de jeux pédagogiques destinés aux élèves de l'école primaire. Dans un deuxième temps, nous abordons la question de l'*adaptive learning*, en fonction des différents profils d'élèves, par l'intégration de ressources pédagogiques à des parcours adaptatifs dans le cadre du projet PEAPL.

Mots-clés : EIAH, plateforme numérique, pédagogie du jeu, parcours adaptatif.

Abstract. The customization of learning pathways based on competence profiles and game-based pedagogy is increasingly adopted by education stakeholders because of their potential to maximize the effectiveness of pedagogical measures. They are specially developing in the era of multimedia digital platforms, enriched by new technologies. However, actual learning can vary among individuals, particularly according to specific needs (L2 speakers, dyslexics, hearing impaired, etc.). In this article, we first present *GamesHUB*, the platform of learning games for elementary school pupils. Secondly, we address the issue of adaptive learning, according to the different pupils' profiles, through the integration of pedagogical resources into adaptive pathways within the PEAPL project.

Keywords: TEL environment, digital platforms, play-based learning, adaptive pathways.

1 Introduction

Les jeux pédagogiques sont de plus en plus utilisés par les enseignants grâce à leur potentiel pour favoriser l'engagement des apprenants [1, 2]. De plus, leur déploiement est aujourd'hui simplifié par la dissémination des outils numériques (tablettes, ordinateurs) dans les établissements scolaires à l'échelle européenne [3] et mondiale [4]. Ces plans d'équipement ont ainsi ouvert la voie à l'émergence de nombreuses plateformes numériques et ludiques [5] dans plusieurs contextes éducatifs. Cependant, bien que ces plateformes aient en général permis d'étendre les possibilités d'apprentissage et de le favoriser, notamment dans le contexte actuel de pandémie imposant le distanciel, leur utilisation s'accompagne aussi de questions sur les possibilités d'adaptation face aux profils hétérogènes d'apprenants. En effet, dans un contexte où l'interface numérique est le principal support d'accès à l'information, les capacités d'apprentissage peuvent plus ou moins varier en fonction des profils des apprenants et notamment de certains besoins spécifiques (dyslexie, allophonie, daltonisme, etc.). Pour toutes ces raisons, la question de l'apprentissage adaptatif est d'autant plus primordiale à l'ère du numérique et des Environnements Informatiques pour l'Apprentissage Humain (EIAH).

L'apprentissage adaptatif dans les EIAH ludiques varie de l'adaptation des scénarios pédagogiques en fonction des profils et des compétences des apprenants [6] à l'adaptation des mécanismes de jeu eux-mêmes [7] en fonction des motivations et des profils des joueurs [8]. Dans cet article, nous nous intéressons à la personnalisation des parcours d'apprentissage en fonction des profils de compétences des apprenants. Dans notre projet, ces parcours d'apprentissage seront autodiagnostiqués à partir des réponses des élèves, mais aussi à partir d'informations complémentaires de l'enseignant. Ainsi, des ressources pédagogiques spécifiques sont mises au point par les experts en didactique afin d'être assignées à ces parcours adaptatifs sur la base de référentiels de compétences. En conclusion, nous expliciterons les projets envisagés pour ajouter d'autres paramètres permettant l'élaboration de profils d'apprenants plus complets.

2 La plateforme *GamesHUB*

La plateforme *GamesHUB*¹ vise à promouvoir l'apprentissage de tous les élèves de l'école primaire ordinaire francophone en utilisant la pédagogie par le jeu et la personnalisation des parcours. La plateforme propose des jeux relatifs à des domaines d'apprentissage divers dans le cadre du Plan d'Études Romand² (PER). De plus, elle permet l'amélioration continue des scénarios d'apprentissage en enregistrant, dans le respect des Règles Générales de Protection des Données (RGPD), les traces d'activités des élèves. Ainsi, les données de chaque partie de chaque jeu sont enregistrées et peuvent ensuite être visualisées et analysées par l'enseignant afin de repérer les difficultés des

¹ La plateforme *GamesHUB* a été réalisée au sein du Centre de Recherche sur l'Enseignement/Apprentissage par les Technologies numériques (CRE/ATE) de la Haute École pédagogique Fribourg (HEP), en collaboration avec la section informatique de l'École des Métiers de Fribourg (EMF). Pour plus de détails : <https://hep3.emf-infopro-test.ch/>

² <https://www.plandetudes.ch/home>

élèves et avoir des pistes d'amélioration. L'ambition est de permettre aux élèves avec des besoins éducatifs particuliers (e.g. élèves allophones ou avec trouble du développement du langage) de pouvoir également développer des compétences via leur capacité d'interagir avec la machine.

La plateforme *GamesHUB* repose sur trois parties principales :

2.1 Le portail pédagogique

Le portail pédagogique permet aux enseignants de se connecter, de gérer leurs classes, de lancer les jeux pour les élèves et d'observer leurs traces d'activités. De plus, le portail permet aux enseignants de créer des parcours d'apprentissage personnalisés en fonction des différents besoins et profils des élèves. La fonctionnalité des parcours est actuellement en cours de développement dans le cadre du projet PEAPL que nous détaillons dans la section 3 de ce document.

2.2 Le Framework

Le Framework est un modèle de développement de jeux utilisant un *template* uniforme pour permettre l'intégration rapide de nouveaux jeux à la plateforme *GamesHUB*. Le Framework gère l'affichage et la mise à jour visuelle des éléments, ainsi que les traces d'utilisation. D'un point de vue technique, son utilisation nécessite que les nouveaux jeux développés respectent une certaine structure (inclusion d'un *canvas*, un *listener manager* et un *game mode* énuméré, relatif aux quatre modes de jeu détaillés dans la sous-section suivante). Ces jeux auront l'avantage de pouvoir être rapidement déployés sur la plateforme. Le Framework repose essentiellement sur le langage de programmation Type Script et sur la librairie open-source PixiJs³.

2.3 Les jeux

Plusieurs jeux sont proposés dans le cadre de la plateforme *GamesHUB*. Ces derniers sont tous développés en fonction du modèle décrit ci-dessus et sont liés à différentes thématiques scolaires (p. ex., mathématiques, français, sciences de la nature, etc.). Les jeux sont optimisés pour une exécution sur des tablettes numériques, mais aussi sur des ordinateurs. Le *gameplay* est réparti en quatre modes :

- Exploration : l'élève apprend le fonctionnement du jeu avec un mode guidé pas à pas ;
- Entraînement : les niveaux sont déjà créés et l'élève doit les compléter. Il n'est pas noté ;
- Évaluation : similaire au mode d'entraînement sauf que l'élève sera noté dans ce mode (0, 1 ou 2 ; équivalent à échoué, partiellement réussi ou réussi) ;
- Création : les élèves ou les enseignants créent des niveaux de jeux différents. Ces niveaux sont ensuite sauvegardés et validés par l'administrateur afin qu'ils puissent être réalisés par d'autres élèves.

³ <https://www.pixijs.com/>

Les jeux permettent à l'enseignant de retrouver les données relatives à la progression des élèves dans les exercices ainsi que leurs scores sur une période donnée. Conformément aux RGPD, seule l'information du prénom de l'élève est enregistrée par l'enseignant lors de la création de sa classe.

3 Le projet PEAPL

Dans une perspective de pédagogie universelle, et dans le cadre du projet européen Plateforme Européenne d'Apprentissage Personnalisé des Langues⁴ (PEAPL), notre équipe interdisciplinaire développe des parcours d'enseignement/apprentissage articulés, avec une interface adaptée aux différents besoins des apprenants. Ces parcours, créés à partir de référentiels de compétences, sont intégrés au tableau de bord de l'enseignant sur *GamesHUB*. Dans un deuxième temps, l'idée sera d'établir un diagnostic automatisé à partir de l'activité de l'élève afin de rendre les parcours adaptatifs (cf. section 3.2). Le but est donc d'apporter les outils pour aider les enseignants des classes ordinaires francophones à gérer l'hétérogénéité de leurs élèves et permettre à chaque élève d'avoir un parcours personnalisé adapté à ses besoins.

Comme expliqué dans un second article d'EIAH 2021 [9], les parcours d'apprentissage s'intègrent à la plateforme *GamesHUB* dans le tableau de bord de l'enseignant. Les parcours sont générés à partir de deux référentiels de compétences : le PER⁵ et le référentiel PEAPL qui concerne les compétences, savoir-faire et savoirs en lecture-compréhension au primaire. Celui-ci a été développé sur la base du métamodèle du projet COMPER⁶. Le référentiel PEAPL a été publié⁷, ainsi qu'un extrait de l'organisation praxéologique qui est à la base de ce référentiel⁸.

3.1 Des parcours adaptatifs aux parcours autoadaptatifs

En accédant et en modifiant les parcours relatifs aux deux référentiels évoqués ci-dessus, l'enseignant a la possibilité d'affecter les parcours à toute la classe ou bien à des élèves en particulier. Cependant, les parcours peuvent s'avérer plus ou moins adaptés aux différents profils des élèves, ce qui remet sur la table la question de l'*Adaptive Learning* et de son fondement. Celui-ci peut prendre plusieurs formes : de la définition manuelle des étapes et scénarios de remédiations par l'enseignant à la scénarisation assistée par un système de propositions sur la base du référentiel de compétences PEAPL, voire d'erreurs, et jusqu'à la suggestion automatisée des étapes en fonction du scénario, de la praxéologie et du profil de l'apprenant. Actuellement, *GamesHUB* affiche les parcours, permet leur personnalisation par l'enseignant et donc la création manuelle d'étapes de remédiation. Cependant, la tâche de définition des scénarios et des possibilités de remédiation peut constituer un travail fastidieux pour l'enseignant [10].

⁴ <https://blog.hepfr.ch/create/peapl/>

⁵ <https://bdper.plandetudes.ch/api/v1/>

⁶ <https://comper.fr/accueil>

⁷ <https://zenodo.org/record/4462850>

⁸ <https://zenodo.org/record/4001381>

Pour cette raison, nous développons sur *GamesHUB* un système de suggestions permettant d'assister l'enseignant dans cette tâche à partir du référentiel de praxéologies. Nous estimons également que cette assistance pourrait être renforcée par un profilage automatisé des apprenants à partir des traces recueillies par *GamesHUB*. Ainsi, à partir des traces d'activités recueillies sur *GamesHUB*, l'algorithme permettra d'identifier les élèves en difficulté et de proposer à son tour des étapes de remédiation.

3.2 Indices de difficultés

La création de parcours se fait en sélectionnant les jeux et niveaux à partir d'une série de filtres : difficulté, tags (associés par le créateur d'un niveau), élément du référentiel de compétence, langue du jeu (*GamesHUB* est pensée comme une plateforme multilingue). Pour chaque jeu, le mode *Exploration* et le mode *Création* sont activables pour des remédiations éventuelles ou des prolongements vers un niveau taxonomique plus élevé. Le parcours est alors conçu en étapes obligatoires et en étapes facultatives (remédiations), activées si l'apprenant échoue dans une étape obligatoire. Cependant, la scénarisation du parcours ne peut pas seulement être liée aux modes de jeu, elle doit aussi être corrélée à différentes variables didactiques et linguistiques : la longueur du texte ; sa mise en page ; la fréquence du vocabulaire ciblé (fréquence des connecteurs utilisés) ; la fréquence du vocabulaire global du texte ; l'utilisation ou non de genres et de nombres différents ; les informations spatio-temporelles linéaires ou non ; le nombre d'items ; la présence d'informations inutiles ou non ; la présence de perturbateurs linguistiques ou non.

Toutes ces variables sont associées à des valeurs (par exemple "1" pour le mode *Exploration* ; "2" pour le mode *Entraînement*, ...). Le tableau 1⁹ présente les valeurs possibles des différentes variables. Ces valeurs permettent de calculer un indice de difficulté nommé "D" pour chaque niveau du jeu en fonction de contraintes linguistiques et didactiques (tableau 2¹⁰). L'avancement dans le parcours implique donc l'augmentation de cet indice. Cela correspond à la progression ordinaire au sein d'une séquence didactique construite par l'enseignant avec différents supports traditionnels.

L'indice D est pour l'instant calculé en fonction de la somme des valeurs des différentes variables. La première phase d'expérimentation de la plateforme permettra de vérifier si toutes les variables sont pertinentes ou si d'autres manquent. Elle nous permettra également de vérifier si des coefficients devraient être ajoutés pour pondérer chaque variable en fonction de son importance.

4 Conclusion et perspectives du projet PEAPL

Dans cet article, nous avons présenté la plateforme de jeux pédagogiques *GamesHUB*, destinée aux élèves de l'école primaire. Nous avons présenté les différents modes de jeu possibles ainsi que le support d'action de l'enseignant, y compris l'analyse des traces d'activité et la création de parcours personnalisés. Ceci est réalisé avec des médiations

⁹ <https://doi.org/10.5281/zenodo.4673080>

¹⁰ <https://doi.org/10.5281/zenodo.4673080>

intégrées dans la plateforme et des remédiations basées sur des référentiels de compétences. Ainsi, les parcours d'apprentissage peuvent offrir un point de départ didactique qui peut être personnalisé en fonction des besoins individuels des apprenants, que ce soit pour favoriser l'accès à l'activité ou pour travailler sur des prérequis faibles ou manquants. La prochaine étape du projet nous permettra de réaliser des expériences dans des classes ordinaires de l'école primaire afin d'évaluer l'expérience utilisateur et de collecter des données sur l'utilisation par les apprenants. Enfin, l'évaluation des traces recueillies et leur analyse pour définir des trajectoires d'apprentissage en fonction des profils d'élèves font l'objet d'une demande de financement spécifique. Ces différentes étapes assureront le passage aux parcours autoadaptatifs.

5 Remerciements

Nos vifs remerciements s'adressent aux étudiants de l'EMF, ayant contribué à ce projet dans le cadre de leurs stages de fin d'étude : Florian Kolly et Vincent Audergon (concepteurs du Framework), Théo Texeira, Erwan Sturzenegger, Arnaud Kolly, Xavier Montasell, David Dumas.

Références

1. Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., Inal, Y., Kızılkaya, G.: The effects of computer games on primary school students' achievement and motivation in geography learning. *Computers & education*. 52, 68–77 (2009).
2. Hawlitschek, A., Joeckel, S.: Increasing the effectiveness of digital educational games: The effects of a learning instruction on students' learning, motivation and cognitive load. *Computers in Human Behavior*. 72, 79–86 (2017).
3. Ho, M.-H.: Le numérique éducatif: un portrait européen. (2014).
4. Karsenti, T., Fievez, A.: L'iPad à l'école: usages, avantages et défis. (2013).
5. Subhash, S., Cudney, E.A.: Gamified learning in higher education: A systematic review of the literature. *Computers in human behavior*. 87, 192–206 (2018).
6. Paramythis, A., Loidl-Reisinger, S.: Adaptive learning environments and e-learning standards. In: *Second european conference on e-learning*. pp. 369–379 (2003).
7. Monterrat, B., Lavoué, E., George, S., Desmarais, M.: Les effets d'une ludification adaptative sur l'engagement des apprenants. *STICEF (Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation)*. 24, 23p (2017).
8. Nacke, L.E., Bateman, C., Mandryk, R.L.: BrainHex: A neurobiological gamer typology survey. *Entertainment computing*. 5, 55–62 (2014).
9. Alvarez, L., Karoui, A., Ramalho, M., Geoffre, T.: Fondements pédagogo-didactiques d'un environnement informatique d'apprentissage destiné aux jeunes apprenants de l'école primaire. 10ème édition de la conférence sur les Environnements Informatiques pour l'Apprentissage Humain, Fribourg (en ligne) Juin 2021.
10. Vermeulen, M., Guigon, G., Mandran, N., Labat, J.-M.: Teachers at the heart of the learning games design: the DISC model. In: *17th International Conference on Advanced Learning Technologies (ICALT)*. pp. 145–149. IEEE (2017).

La chorégraphie des traces étudiantes : une nouvelle visualisation multidimensionnelle pour les acteurs pédagogiques

Alexis Lebis¹, Nadine Mandran² et Mathieu Vermeulen¹

¹ IMT Lille Douai, Institut Mines-Télécom, Univ. Lille, Centre for Digital Systems
F-59000 Lille, France

² Grenoble Alpes Université
`alexis.lebis@imt-lille-douai.fr`

Résumé. Les différentes visualisations en EIAH, comme les tableaux de bord, permettent de tenir compte du déroulé d'une activité pédagogique, de la comprendre et de fournir un retour aux différents acteurs pédagogiques. Cependant, ces visualisations sont souvent peu intégrées à l'activité en question. Dans ce papier, nous proposons le concept de chorégraphie. Il s'agit d'une visualisation permettant de rejouer les moments clefs d'une activité pédagogique réalisée dans un outil informatique, en plus de pouvoir les annoter sémantiquement. De par ses propriétés, la chorégraphie constitue un outil d'aide à la décision et à la remédiation important pour les étudiants et les enseignants, notamment par les possibilités de comparaisons et de retours d'informations qu'elle procure.

Mots-clé: Visualisation, traces, labélisation, learning analytics

Abstract. The different visualisations in TEL, such as the dashboards, allow to take into account the progress of a teaching activity, to understand it and to provide feedback to the different teaching actors. However, these visualisations are often poorly integrated into a such activity. In this paper, we propose the concept of choreography, a visualization that allows to replay the key moments of a pedagogical activity carried out in a computer tool, but also to annotate them semantically. Because of its properties, choreography is an important decision-making and remediation tool for students and teachers, particularly because of the possibilities for comparisons and feedbacks it provides.

Keywords: Visualisation, students' traces, labelling, learning analytic

1 Introduction

La définition d'indicateurs pédagogiques revêt une importance particulière, puisqu'ils renseignent avec une certaine précision l'état de l'activité pédagogique à un instant donné en fonction des traces d'activités de l'apprenant [1,3]. De fait, assembler ces indicateurs au sein de tableaux de bord permet de fournir aux acteurs pédagogiques (apprenants, enseignants, décisionnaires) une vue, plus ou

moins représentative, d'une situation d'apprentissage donnée [2]. Apodictiquement, ces indicateurs doivent donc être correctement interprétés par les acteurs, afin qu'ils puissent examiner correctement ladite situation. Néanmoins, cela peut s'avérer difficile de par la nature des indicateurs (*e.g.* trop technique, mauvais contexte, pas de corrélation avec la situation) et des biais peuvent ainsi survenir. Aussi, nous pensons qu'il est primordial de replacer les acteurs dans leur activité pédagogique lorsqu'ils consultent ces indicateurs. Il s'agit donc de proposer un autre paradigme que celui du tableau de bord pouvant être désincarné de l'activité ; proposer un paradigme de visualisation de l'activité pédagogique qui pourrait inclure directement les indicateurs dans son déroulé tout entier. La base d'un tel paradigme réside donc dans la possibilité de rejouer l'activité pédagogique, et de pouvoir la décrire. Cependant, et en accord avec la notion de scénario d'activités proposés par Gueraud *et al.* [4], il n'est pas besoin de rejouer l'intégralité de la trace d'activité d'un apprenant pour conserver la pertinence de l'activité réalisée et l'examiner : les moments signifiants du scénario comptent (*e.g.* objectifs, sous-objectifs, transitions), et certains états de l'outil utilisé.

Ce papier présente nos travaux sur la *chorégraphie*. Il s'agit de rejouer l'activité pédagogique réalisée par un apprenant de manière ciblée, ainsi que de la décrire par l'entremise d'annotation sémantique. Dans la section 2, nous définissons cette chorégraphie et ses apports. Dans la section 3, nous présentons le framework de cette chorégraphie pour sa mise en oeuvre au sein d'outils. Ensuite, dans la section 4, nous présentons les premiers retours d'expériences.

2 Chorégraphie

Nous définissons donc le principe de chorégraphie comme une succession temporelle des moments clés (que nous appelons dans la suite **keyframes**) d'une activité pédagogique, réalisée dans un outil informatique par un acteur – généralement l'apprenant : il ne s'agit donc pas d'un historique exhaustif de l'activité de l'apprenant (*i.e.* l'ensemble des traces d'activités), mais de son activité signifiante qui répond au scénario pédagogique de l'enseignant. Cette succession de keyframes, que nous appelons une **trame**, est rejouable au sein de l'outil en question, et permet d'obtenir une visualisation animée de l'activité de l'apprenant. De plus, la chorégraphie définit également l'annotation de la trame, ainsi que de ses keyframes et leur transition, par les différents acteurs pédagogiques (*e.g.* l'enseignant, d'autres apprenants) avec des éléments sémantiques forts, permettant d'y introduire de l'information pédagogique supplémentaire.

Le point central de la chorégraphie réside donc dans la définition des keyframes de l'activité pédagogique. L'identification de ces keyframes reposent conjointement sur une description du scénario pédagogique, afin de pouvoir spécifier les moments signifiants de l'activité, et l'utilisation d'un outil informatique capable de tracer cette activité (et la rejouer). C'est principalement cette trace d'activité qui doit pouvoir être analysée afin de repérer lorsque l'outil passe d'un état signifiant du scénario à un autre, dans l'objectif d'identifier les keyframes. Ainsi, certaines actions contribuent à la création d'une nouvelle keyframe. Par

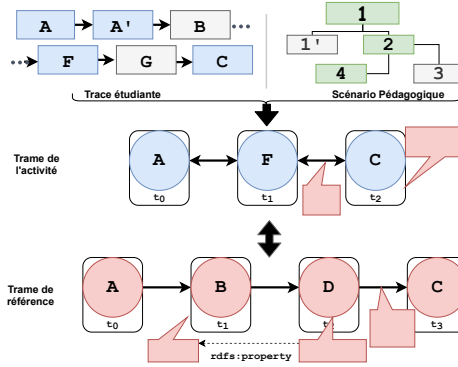


FIG. 1. Illustration de la chorégraphie, de l'identification des keyframes d'une trace d'activité à la comparaison à une trame de référence, en rouge.

exemple, dans un *taskboard*, le fait de déplacer une tâche donnée spécifiquement *après* une autre peut amener à la génération d'une nouvelle keyframe – si le scénario pédagogique prévoit un ordonnancement des tâches pour observer des compétences de gestion de projets.

De tout cela, il émerge de la chorégraphie un comportement intéressant : la définition d'une (ou plusieurs) trame(s) de référence pour l'activité pédagogique. Cette trame de référence – établit par exemple par un enseignant – va illustrer l'activité pédagogique attendue, et même permettre de la jouer au sein de l'outil concerné. Couplé au mécanisme d'annotation qui consiste à attacher des éléments et des propriétés sémantiques à la trame et à ses objets, l'on obtient donc une visualisation multidimensionnelle rendant possible l'explicitabilité de l'activité. En effet, il devient possible pour un étudiant de comparer sa propre trame avec une trame de référence, et ainsi avoir un corrigé détaillé immédiat.

La Figure 1 illustre notre principe de chorégraphie. À partir de la trace d'activité de l'étudiant et du scénario pédagogique, les éléments contributifs à la génération de keyframes sont identifiés³ (les rectangles bleus et verts). La trame de keyframes est ensuite créée, en prenant garde de les ordonner temporellement. Cette trame est ensuite annotée – ici par l'enseignant – pour indiquer certaines informations, comme des erreurs : par exemple, une tâche d'un *taskboard* déplacée au mauvais moment. Ensuite, cette trame de l'étudiant est mise en comparaison avec une trame de référence pour l'activité pédagogique réalisée, intensivement annotée. Les rectangles *B* et *G* sont des actions réalisées par l'étudiant qui n'ont pas d'effet notable concernant l'activité pédagogique.

3. Dans la section 3, nous proposons un moyen d'identifier ces keyframes.

3 Framework

Identification des keyframes. Pour discriminer efficacement les actions contenues dans les traces d'activités de l'étudiant comme étant des keyframes ou non, nous proposons d'enrichir la spécification xAPI. xAPI permet d'exprimer les activités d'apprentissages qui surviennent dans les dispositifs d'apprentissages de manière interopérable via un vocabulaire spécifique, des fois propre à un contexte (*i.e.* profil) – par exemple aux jeux sérieux. Plus précisément, il s'agit de proposer un nouveau profil xAPI, dédié à la chorégraphie des activités pédagogiques et enrichissant le standard avec des champs supplémentaires en lien avec les scénarios pédagogiques et l'état des outils utilisés. Ces nouveaux champs, au nombre de trois, viennent compléter le champ `context` d'xAPI. Tout d'abord, le champ `state` indique l'état de l'outil en question – encodable avec un vocabulaire dédié. L'état est enrichi par le champ `transition`, qui permet d'indiquer la transition entre cet état et un précédent, ainsi que de la qualifier. Enfin, le champ `scenario` permet de décrire pour quel élément du scénario pédagogique (*e.g.* `objective`, `sub-objective`) un élément d'une trace xAPI est pertinent en tant que keyframe. Conséquemment, produire une trame de référence n'est pas plus difficile ; l'activité doit simplement être réalisée par un enseignant.

Enrichissement. Les travaux réalisés dans l'annotation sémantique et sur la narration de processus d'analyse de traces d'apprentissages [5] offrent un cadre intéressant quant à l'enrichissement des trames de la chorégraphie. Il devient ainsi possible de définir un vocabulaire partagé des éléments sémantiques véhiculant un concept pédagogique particulier, par exemple *oublier une retenue*, et les attacher directement au niveau de la trame, de la keyframe ou même des transitions entre keyframes. Comme le précise [5], permettre la mise en relation de ces éléments sémantiques les uns par rapport aux autres, là encore par l'entremise de propriétés sémantiques, contribue à fournir un artefact informatique auto-suffisant, favorisant sa compréhension à la fois par la machine et par l'humain. Il émerge de cela la possibilité d'annoter la trame d'activité avec des indicateurs plus classiques, et ainsi créer une visualisation riche, contribuant à s'émanciper un peu plus des tableaux de bords. Couvrir les spécificités des différents contextes pédagogiques requiert selon nous une co-construction du vocabulaire.

Comparaison et instabilité chorégraphique. La comparaison d'une trame avec une trame de référence est un atout important de la chorégraphie. Bien que cette comparaison peut être manuellement effectuée par un acteur – par exemple un étudiant qui compare son activité avec celle de l'enseignant – elle peut aussi être effectuée automatiquement. Cela permet d'assister les acteurs dans l'identification des éléments qui diffèrent de la trame de référence et identifier leurs potentielles erreurs. Néanmoins, dans des scénarios complexes, il est rare qu'un état soit simplement substitué par un autre : en ensemble d'états non prévus peut manquer et/ou survenir par rapport à la trame de référence, rendant la comparaison complexe. Nous appelons cela l'instabilité chorégraphique. De tels ensembles sont généralement porteur d'informations à propos de l'étudiant : il est

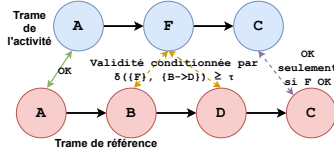


FIG. 2. Illustration de la comparaison et de l'instabilité chorégraphique.

donc nécessaire de conditionner la validité de la trame si elle ne correspond pas exactement à celle de référence, comme le montre la Figure 2. Pour réaliser cela, nous définissons une fonction de similitude d'états δ , permettant de qualifier le degré d'équivalence d'un ensemble d'états de la trame de l'étudiant par rapport à un ensemble d'états de la trame de référence⁴, comme illustré dans la Figure 2 avec les ensembles $\{F\}$ et $\{B, D\}$. En définissant un seuil τ au delà duquel l'équivalence est considérée comme valide, il devient possible de conditionner la validité des prochains états de la trame. Ainsi, fixer $\tau = 1$ revient à attendre de l'étudiant qu'il réalise l'activité exactement comme le souhaite l'enseignant.

Il ressort qu'il est possible de définir une métrique \mathcal{M} traduisant la qualité de l'activité de l'apprenant – sa trame t – par rapport aux attendus pédagogiques – la trame de référence r – et le nombre de fois n qu'ils ont été enfreints, tel que :

$$\mathcal{M}(c, r) = \frac{\sum_{i=0}^n \delta(E_i, R_i)}{n} \quad (1)$$

Plus la valeur tend vers 1 et plus l'activité se rapproche de la trame de référence et plus l'on peut la considérer comme correcte, d'après les critères de l'enseignant.

4 Expérimentation

Nous avons expérimenté ces travaux sur la chorégraphie au travers d'un cours de gestion de base de données (BDD) dispensé d'après l'approche ALPES [6], d'une durée de 16 heures et dans une classe de 32 étudiants, répartis en binôme. Plus particulièrement, nous étions intéressés par l'identification et l'évaluation de compétences transversales en gestion de projet appliqué à la BDD. Nous les avons rendu mobilisables à travers un outil informatique de type *taskboard* développé pour l'occasion, et nous avons tracé son utilisation par les binômes. Malheureusement frappés par la crise sanitaire de la COVID et donc par la modification drastique du contexte d'expérimentation invalidant de fait toute la démarche, nous ne pouvons pas réaliser une étude formelle des résultats expérimentaux. Néanmoins, informellement, les premiers retours expérimentaux sont encourageants. Les trames générées permettent de rejouer correctement le déroulé de l'activité pédagogique des étudiants (*e.g.* les moments d'hésitations, des

4. Le degré d'équivalence peut être défini de différente manière, notamment manuellement ou par la comparaison de différentes trames de références.

erreurs) et ainsi offrir un moyen d'évaluation supplémentaire aux enseignants. À ce titre, nous avons présenté les activités des étudiants sous forme de chorégraphie à un enseignant qui nous a indiqué l'utilité d'avoir à disposition une telle visualisation pour "mieux comprendre le cheminement de l'étudiant".

5 Conclusion et perspectives

Dans ce papier, nous avons présenté la chorégraphie : une visualisation riche, issue des traces d'apprentissage, permettant de rejouer l'activité pédagogique d'un apprenant dans un outil d'apprentissage en fonction du scénario pédagogique. À cette rejouabilité vient s'ajouter la possibilité d'annoter l'entièreté de la chorégraphie avec des éléments sémantiques, offrant de nouvelles perspectives de retours d'expériences aux étudiants et aux enseignants – par exemple en permettant de se comparer à une activité de référence ou en y adjoignant d'autres indicateurs. Enfin, nous travaillons sur la prédiction chorégraphique pour assister encore un peu plus les enseignants, en nous inspirant des travaux sur le transport optimal en image, pour prédire les comportements plausibles d'un étudiant. Nous travaillons aussi sur la capitalisation des trames de références, permettant, d'une part, aux acteurs d'avoir toujours plus de contenus pour réaliser ou concevoir une activité et, d'autre part, envisager la conception automatique d'exercices – par exemple en enfrenant volontairement des étapes de la chorégraphie.

Remerciement

Cette contribution scientifique a été réalisée dans le cadre du projet de recherche APACHES, financé par l'i-SITE ULNE (FIPE18-007-VERMEULEN).

Références

1. Choquet, C., Iksal, S. : Usage tracking language : a meta language for modelling tracks in tel systems. In : Proceedings of International Conference on Software Technologies. pp. 133–138. INSTICC (2006)
2. Dabbebi, I., Iksal, S., Gilliot, J.M., May, M., Garlatti, S. : Towards Adaptive Dashboards for Learning Analytic : An Approach for Conceptual Design and implementation. In : 9th International Conference on Computer Supported Education (CSEDU 2017). pp. 120–131. Porto, Portugal (Apr 2017)
3. Dimitrakopoulou, A. : State of the art on Interaction and Collaboration Analysis (2004), (D26.1.1) EU Sixth Framework programme priority 2, Information society technology, Network of Excellence Kaleidoscope, (contract NoE IST-507838), project ICALTS : Interaction & Collaboration Analysis
4. Guéraud, V. : Une approche auteur pour les scénarios d'activités. Scénariser l'enseignement et l'apprentissage : une nouvelle compétence pour le praticien pp. 33–38 (2006)
5. Lebis, A., Lefevre, M., Luengo, V., Guin, N. : Capitalisation of Analysis Processes : Enabling Reproducibility, Openness and Adaptability thanks to Narration. In : LAK '18 - 8th International Conference on Learning Analytics and Knowledge. pp. 245–254. ACM, Sydney, Australia (2018)
6. Vermeulen, M., Fleury, A., Fronton, K., Laval, J. : LES A.L.P.E.S. : APPROCHES AGILES POUR L'ENSEIGNEMENT SUPERIEUR. In : Colloque Questions de Pédagogie pour l'Enseignement Supérieur (QPES 2015). Brest, France (Jun 2015)

Les IPMC : une plongée au cœur de la mise en pratique des compétences

Alexis Lebis¹, Estelle Prior^{1,2}, Nadine Mandran², and Mathieu Vermeulen¹

¹ IMT Lille Douai, Institut Mines-Télécom, Univ. Lille, Centre for Digital Systems
F-59000 Lille, France

[alexis.lebis, mathieu.vermeulen]@imt-lille-douai.fr

² Grenoble Alpes Université

Nadine.Mandran@univ-grenoble-alpes.fr

Résumé. Dans les formations, et *a fortiori* dans les approches par compétences, inventorier les compétences que les étudiants acquerront est une tâche complexe et, conséquemment, l'observation de la maîtrise de ces dernières par les étudiants l'est encore plus. Sans définitions formelles, ces difficultés ne seront pas résolues et l'évaluation des compétences dans des dispositifs informatiques restera complexe. Dans cet article de positionnement, nous présentons nos travaux sur les Indicateurs de Perception de Mobilisation de Compétences (IPMC). Il s'agit non pas de définir formellement une compétence, mais la manière de la caractériser, de l'observer et de mesurer ses propriétés. En nous émancipant du côté nébuleux de la compétence, nous montrons comment il devient alors possible d'assister les acteurs pédagogiques pour identifier quels outils peuvent être pertinents pour évaluer une compétence.

Mots-Clefs : Approche par compétence, Indicateur, Évaluation

Abstract. In training courses, and *a fortiori* in competency-based approaches, making an inventory of the competences that students will acquire is a complex task. Thus, observing the mastery of these competences by students is even more so. Without formal definitions, such issues will not be resolved, especially in a TEL context. In this position paper, we propose a novel notion called the Perception Indicator of Competences Mobilisation (PICM). It is not related to a formal definition of competency, but to how to characterise it, observe it and measure its properties. By emancipating ourselves from the fuzzy side of the competency, we illustrate how it is possible to assist educational stakeholders in identifying which tools may be relevant to assess students' competencies.

Keywords: Competency-based approach, Indicator, Evaluation

1 Introduction

Malgré des modèles plus ou moins riches tentant de décrire les compétences [74], l'état de l'art de ces vingt dernières années laisse finalement transparaître un certain insuccès à incorporer efficacement les compétences dans des dispositifs

informatiques afin de les exploiter, de les évaluer et d'assister les différents acteurs [8]. Preuve en est avec le financement récent de projets de recherche autour de cette thématique, comme APACHES ou COMPER, qui s'intéressent à la mise en œuvre d'approches par compétences. Cette difficulté de représenter la compétence comme un artefact informatique mesurable objectivement tient principalement à la genèse de ses définitions, éloignées *a priori* des systèmes d'informations [1] – constat appuyé par Paquette lorsqu'il propose un modèle ontologique de la compétence [5]. À cela s'ajoute l'intrication d'un biais d'observation généré par les dispositifs informatiques utilisés pour évaluer les compétences avec la capacité d'un étudiant à les mettre en pratique. Il devient alors complexe d'identifier dans quel état d'appropriation se trouve un étudiant vis-à-vis de ses compétences.

Ce papier de positionnement présente le concept d'Indicateur de Perception de Mobilisation de Compétences (IPMC). Il s'agit des résultats préliminaires de nos travaux de recherche obtenus dans le cadre du projet de recherche APACHES, lors de la tentative d'élaboration d'un catalogue de compétences concernant la pédagogie par projet en informatique centrée humain [3] et les analyses associées pour observer ces compétences. Nous pensons que les IPMC constituent une piste d'innovation et de discussion intéressante pour toute la communauté à propos de la manière de traiter la compétence – de sa caractérisation jusqu'à son observation. Sans se placer en rupture avec les propositions admises [7], ils envisagent une inclusion plus opérante des compétences dans les dispositifs informatiques d'apprentissage, ainsi que des mécanismes d'assistances renforcés (*e.g.* partage, choix des outils, aide aux enseignants pour l'évaluation, visualisation). Dans la [section 2](#), nous présentons ce nouveau paradigme et, brièvement, son fondement théorique. La [section 3](#) nous permet d'illustrer notre proposition et ses multiples avantages, avant de conclure dans la [section 4](#).

2 Les IPMC ¶

2.1 Intuition, hypothèses et définition

L'état de l'art sur les compétences met en exergue leur aspect riche et complexe, et la difficulté à évaluer correctement leur degré d'acquisition par les étudiants [9]. Cette évaluation passe généralement par l'observation de la réalisation d'une ou plusieurs tâches qui impliquent des éléments discriminants de la compétence, et non pas la compétence en elle-même [6]. Par conséquent, il doit exister une corrélation forte entre ces éléments et ladite compétence, à tel point qu'ils la caractérisent suffisamment.

Nous faisons donc l'hypothèse que, s'il était possible d'évaluer directement la compétence d'un étudiant, ce serait déjà ce que nous ferions au sein des différentes communautés pédagogiques et de recherche. Cette évaluation ne peut se faire qu'à travers l'observation des éléments discriminants de la compétence : ses propriétés. Cela implique qu'elles doivent être observables et quantifiables en fonction de critères pédagogiques objectifs. Ainsi, nous pensons qu'observer, *via* des dispositifs adéquats, suffisamment de propriétés d'une compétence permet de traduire sa maîtrise par l'étudiant. De plus, puisqu'il n'est pas possible d'affirmer

avec certitude que la compétence n'est issue que du processus de cognition d'un individu et non pas aussi de sa conscience phénoménale, nous faisons également l'hypothèse que toute tentative d'observation d'une compétence [\[4\]](#) d'un étudiant entraîne inexorablement sa modification – infime ou non.

Conséquemment, il est possible d'exprimer, à un instant donné, l'appropriation de chacun des observables d'une compétence par un indicateur dédié. Ces indicateurs sont les Indicateurs de Perception de Mobilisation de Compétences (IPMC), que nous définissons comme étant *"une métrique permettant de quantifier, d'après les attentes pédagogiques, la valeur d'appropriation d'un aspect observable d'une compétence, par un individu, à un instant donné, au travers de dispositifs aptes à les mobiliser"*. Ainsi, un ensemble d'IPMC fournit à l'enseignant un faisceau de preuve quant à la maîtrise d'une compétence par un étudiant : il s'agit d'une vue, plus ou moins précise, de la compétence acquise par l'étudiant. Plus cet ensemble est important, plus la couverture est importante – puisque plus de propriétés sont observées – et plus l'observation de la compétence est précise. De plus, les compétences ne sont pas nécessairement mutuellement disjointes et peuvent donc partager certaines propriétés ; un seul IPMC peut ainsi contribuer à l'observation de multiples compétences.

Il est cependant important de noter qu'une propriété ne peut être observée que si le dispositif qui l'observe (*i.e.* l'EIAH) y est sensible. Par exemple, il faut que les traces d'activités que ce dispositif génère permettent d'identifier cette propriété – éventuellement après analyse des traces. La qualité de l'observation à donc un effet direct sur la qualité des IPMC. Néanmoins, propriété et observation sont toutes deux des notions bien distinctes et ne sauraient se confondre.

En conclusion, grâce aux IPMC et leur implication, nous définissons une chaîne de traitement complète et formelle, de la compétence jusqu'à l'opérationnalisation des techniques nécessaires à son observation, afin de fournir à n'importe quel acteur pédagogique un faisceau de preuve concernant la maîtrise de cette compétence. Tout l'enjeu de ce papier est donc d'engager une discussion autour de cette chaîne de traitement, de ses différentes étapes, notamment la définition de l'espace des états, et des IPMC. La réification opérationnelle de la compétence que nous défendons ici nous émancipe certes des modèles traditionnels de la compétence, mais il est tout à fait possible de définir une fonction f comme étant une injection d'un modèle plus classique, comme Paquette [\[5\]](#), vers notre proposition, grâce aux observables et espaces d'états (cf. [subsection 2.2](#)).

2.2 Prémisses théoriques et formalisation

Au fur et à mesure de nos travaux, nous avons constaté une analogie marquée avec les grands principes de la physique quantique, notamment la notion de système, d'observation, d'espace d'états et d'incertitude (notamment liée à l'évaluation de la compétence). Cela nous a permis de construire notre proposition dans un premier cadre théorique. Sans rentrer dans les détails, voici les prémisses fondamentales qui guident notre définition et sa mise en oeuvre.

¹ Observer une compétence revient à observer suffisamment de ses propriétés.

Ainsi, l'on caractérise la compétence K comme un ensemble de propriétés internes, où chaque propriété \hat{A} possède un espace complexe de Hilbert $\mathcal{H}_{\hat{A}}$ des états et de dimension n . On définit $|\Psi\rangle$, un état quelconque, comme un vecteur exprimé dans la base vectorielle E de $\mathcal{H}_{\hat{A}}$, ce qui nous permet de représenter les observations sur cet espace de Hilbert des états. Par conséquent, la valeur de \hat{A} , noté $\langle\hat{A}\rangle$ assimilable à la quantité statistique moyenne des observations, est mesurable et se définit par $\langle\Psi|\hat{A}|\Psi\rangle$. De plus, il devient possible de prédire l'évolution du système par l'action d'opérateurs linéaires L sur les vecteurs de l'espace de Hilbert.

Finalement, un IPMC ² se traduit par la distance entre l'état observé $|\Psi\rangle$ et l'état $|\Phi\rangle$ le plus proche qui représente une attente pédagogique, défini par l'enseignant, tel que $\mathfrak{F} = \|\Psi - \Phi\|$. L'une des questions que nous nous posons encore est si la compétence K peut être envisagée comme un système issu de la superposition linéaire des propriétés \hat{A} . Pour le moment, si plusieurs propriétés interviennent pour la définition de K , nous considérons que le faisceau de preuve s'exprime comme la somme pondérée des IPMC calculables : $K = \sum_{i=0}^n \omega_i \mathfrak{F}_i$.

3 Illustration et tests

La formalisation de notre proposition nous permet de fournir un procédé rigoureux de définition des IPMC associés à une compétence. Il consiste en cinq étapes majeures, représentées par les flèches de la [Figure 1](#). Nous illustrerons ce procédé grâce à un exemple, portant sur la compétence de *gestion de projet*.

La première étape de ce procédé consiste à **caractériser** la compétence, autrement dit d'en définir ses propriétés observables. Un observable \hat{A} de notre exemple serait *organiser ses tâches*, puisque s'il nous est possible d'observer cette propriété, alors nous renforçons notre connaissance sur l'appropriation de la compétence *gestion de projet*. Cette caractérisation n'a pas besoin d'être exhaustive : le faisceau de preuve alors obtenu ne sera que moins précis, nécessitant de l'enseignant une plus grande vigilance quant à l'interprétation des résultats. De plus, cette étape de caractérisation laisse transparaître la possibilité de partager et de co-construire la liste des observables de la compétence. La deuxième étape consiste à **déplier** l'observable en un espace d'états possibles \mathcal{H} respectant les contraintes énoncées dans la [subsection 2.2](#). Cette étape est peut être la plus délicate, puisque la base E de \mathcal{H} va conditionner la définition des états de la propriété. Par exemple, $|c_1\rangle$ – dans la [Figure 1](#) – pourrait correspondre à la notion de *tâches indiquées dans le bon état*, orthogonale à $|c_2\rangle$ *tâches indiquées dans le mauvais état*, tel que $\langle c_1|c_2\rangle = 0$. Ces deux premières étapes sont génériques aux activités pédagogiques, ce qui n'est pas le cas des trois étapes suivantes. L'étape d'**élection** consiste à indiquer quel(s) état(s) sont valides du point de vue de l'objectif pédagogique, et ainsi former un sous espace d'états valides. S'ensuit l'étape d'**observation** ³ de l'apprenant dans son activité pédagogique grâce à un dispositif capable d'observer la propriété en question. Dans notre exemple, cela

² Ce symbole tibétain, que l'on peut lire *Yig*, peut se traduire par : nouvelle page.

³ Cette étape est interchangeable avec l'étape d'élection.

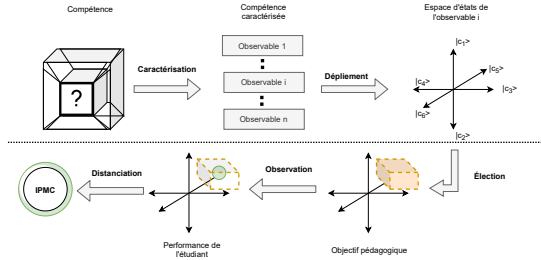


Figure 1. Illustration de la procédure en cinq étapes d'obtention des IPMC.

peut se faire par l'entremise d'un outil informatique de type *taskboard*, traçant les actions de l'apprenant lorsqu'il déplace ses tâches, et l'analyse de ces traces.

Enfin, l'étape de **distanciation** parachève l'IPMC \mathfrak{P}_i , en comparant l'état de l'apprenant observé à ceux attendus par l'enseignant (cf. [subsection 2.2](#)).

Implications. Trois implications importantes émergent de notre proposition. Tout d'abord, puisque seuls les outils capables d'observer une (ou plusieurs) propriété(s) contribuent à l'obtention de (ou des) IPMC associé(s), une aide à la décision pour les enseignants s'esquisse. En effet, pour une compétence donnée et caractérisée, ses IPMC sont connus et, par conséquent, tous les outils capables de les obtenir aussi : il devient donc possible d'indiquer aux enseignants quels outils sont adaptés pour observer une compétence, ceux qui ne le sont pas, et pourquoi. Il est aussi possible d'aider les enseignants à choisir le minimum d'outils nécessaire à déployer. Ensuite, d'après les principes énoncés par la capitalisation [2], il est possible de partager et de réutiliser les espaces d'états, ainsi que les IPMC dans des contextes différents – en élisant et destituant éventuellement des états valides. De plus, en définissant un vocabulaire commun, la métrique obtenue par distanciation peut être sémantiquement caractérisée, offrant, pour les acteurs, une meilleure interprétation (*e.g.* partiellement acquis). Enfin, l'élection des états valides d'après un objectif pédagogique peut aussi être partageable et réutilisable. Cela aboutit à la définition d'une banque d'objectifs pédagogiques, fonctions de contextes pédagogiques et émancipés des outils utilisés.

4 Conclusion et perspective

Ce papier de positionnement présente le concept d'Indicateur de Perception de Mobilisation de Compétences, brièvement sa formalisation et la procédure pour les obtenir. Les IPMC constituent un outil mathématique intéressant en permettant de caractériser formellement l'appropriation d'une compétence par un individu via un faisceau de preuve. De plus, ils font apparaître des possibilités de réutilisation et de partage efficaces de la compétence et de son analyse,

jusqu' alors difficile à mettre en place. Nous espérons que les prémisses de nos travaux susciteront de nombreux débats au sein de la communauté, pour changer la manière dont la compétence est actuellement considérée en terme d' artefact informatique. En plus de l' effort de finalisation de notre proposition, nous travaillons sur l' approfondissement de l' association entre les IPMC et les dispositifs d' observations. Nous avons en effet l' intuition qu' il est possible de considérer l' exercice comme élément contributif de la phase d' observation. Si c' est le cas, il deviendrait donc possible d' indiquer, pour une compétence donnée, tous les exercices qui permettraient de l' observer plus ou moins partiellement.

La première version de ce modèle a permis d' élaborer une grille d' IPMC associée à des compétences dans le cadre d' un enseignement en informatique centrée humain. Cette grille a ensuite été évaluée et améliorée par des enseignants en informatique (six au total) lors d' un séminaire. La notion d' IPMC a rapidement été adoptée par ces derniers. En effet, lors des réunions suivantes, portant sur d' autres sujets, ils ont ré-utilisé de manière spontanée ce concept. Cette expérimentation et les suivantes feront l' objet d' articles à venir.

Remerciement

Cette contribution scientifique a été réalisée dans le cadre du projet de recherche APACHES, financé par l' i-SITE ULNE (FIPE18-007-VERMEULEN).

Références

1. Bloom, B.S., et al.: Taxonomy of educational objectives. vol. 1: Cognitive domain. New York: McKay **20**, 24 (1956)
2. Lebis, A., Lefevre, M., Luengo, V., Guin, N.: Capitalisation of Analysis Processes : Enabling Reproducibility, Openess and Adaptability thanks to Narration. In: LAK '18 - 8th International Conference on Learning Analytics and Knowledge. pp. 245–254. ACM, Sydney, Australia (2018)
3. Lebis, A., Prior, E., Mandran, N., Karami, A., Vermeulen, M.: Promouvoir et soutenir la Pédagogie Par Projet Centré Humain dans le supérieur : le projet APACHES. In: DIDAPRO 8 - DIDASTIC. Lille, France (Feb 2020)
4. Mandin, S., Guin, N.: Basing learner modelling on an ontology of knowledge and skills. In: Sampson, D.G., Spector, J.M., Chen, N.S., Huang, R., Kinshuk (eds.) IEEE International Conference on Advanced Learning Technologies. pp. 321–323. IEEE Computer Society, Athènes, Greece (Jul 2014)
5. Paquette, G.: Modélisation des connaissances et des compétences. Presses de l' Université du Québec (2002)
6. Perrenoud, P.: Évaluer des compétences. L' éducateur pp. 8–11 (2004)
7. Poumay, M., Tardif, J., Georges, F., Scallon, G.: Organiser la formation à partir des compétences: un pari gagnant pour l' apprentissage dans le supérieur. De Boeck, Bruxelles (2017)
8. Ramandalahy, T., Vidal, P., Broisin, J.: An intelligent tutoring system supporting metacognition and sharing learners' experiences. In: International Conference on Intelligent Tutoring Systems. pp. 402–404. Springer (2010)
9. Scallon, G.: L' évaluation des apprentissages dans une approche par compétences. Pédagogies en développement, De Boeck, Bruxelles (2004)

Confinement et distanciel forcés chez les enseignant.e.s des 1er et 2nd degré : genèse d'une hybridation des pratiques ?

Gaëlle Lefer Sauvage¹, Sylvain Genevois², Jacques Kerneis¹, Nathalie Wallian²

¹ Centre Universitaire de Formations et de Recherches à Mayotte

² Université de La Réunion

Résumé. L'hybridation des dispositifs et des ingénieries de formation fait l'objet d'une littérature scientifique importante dans le champ de la pédagogie universitaire, alors qu'elle est peu développée dans l'enseignement primaire et secondaire. A partir d'un ancrage théorique de l'hybridation dans le contexte universitaire, nous émettons des hypothèses quant à une contextualisation possible dans le premier et le second degré. La crise pandémique COVID-19 et le confinement qui ont eu lieu en France entre mars et mai 2020 ont déclenché de nouvelles interrogations quant aux formes possibles d'hybridation dans le premier et le second degrés [4] à travers les bouleversements rapides qu'a connu l'espace-temps scolaire. Des analyses sont encore en cours, mais des premiers éléments de réponses amènent à revisiter les connaissances actuelles, quant aux manières dont il serait possible d'articuler présence-distance à l'école et au collège-lycée. Les enjeux de la genèse de l'hybridation des pratiques enseignantes dans le premier degré sont particulièrement questionnés ici.

Mots-clés : hybride, distanciel, confinement, COVID

Abstract. Blended-learn training and pedagogical scenarios is an extensive research literature. But in Grade/Middle/High school, it could take other forms. Based on a theoretical anchoring of Blended-learning in the university context, we put forward hypotheses as to the possible contextualization in elementary-secondary education. The context of confinement that occurred in France in March-June 2020 triggers new questions about the possible new blend-learning forms, in Grade/Middle/High school [4] through changes in space and time. Analyses are still in progress but first elements lead to revisit current knowledge about the ways in which it would be possible articulation presence-distance at school. Thus, genesis of Blended-learn in teaching practices in primary education is questioned here.

Keywords: blend-learning, distance teacher training, lockdown, COVID

1 Introduction

Pendant le confinement, de nombreux rapports de recherches nationaux [4, 8, 13, 14] et de synthèses [18, 22, 23, 25] sur les pratiques enseignantes ont vu le jour à l'échelle

nationale ou internationale. La plupart mettent en évidence des bouleversements dans l'organisation institutionnelle, dans la forme scolaire, dans les processus d'enseignement-apprentissage, dans les liens avec les familles. Les changements dans les pratiques conduisent les enseignants, dans le cadre d'un enseignement distanciel « forcé », à bricoler, innover, adapter ou approfondir par rapport à leurs pratiques antérieures. Selon PERAYA [18], cette « ingénierie de crise » renforce la « continuité du modèle pédagogique », globalement transmissif et bricolé. En octobre 2020, au sein d'un rapport pour le CNET, TRICOT [23] souligne les prudenances nécessaires dans la mise en place du distanciel en classe. Il signale que le distanciel passe par des pratiques pédagogiques renforcées pour éviter le sentiment d'abandon des élèves, notamment par une explicitation des consignes, de l'anticipation, des feedbacks bien définis. Ce qui suppose de mettre en place des stratégies métacognitives de l'activité en cours, une réflexion sur la régulation pédagogique par exemple lors de travaux restitués par les élèves et un changement de posture enseignante dans la manière d'apporter de l'aide (posture d'étayage) plutôt que de formuler des demandes (posture prescriptive).

Pendant la première phase de confinement en France continentale, relative à la COVID-19 entre mars et mai 2020, on observe une utilisation quotidienne des outils numériques dans le cadre d'un usage scolaire professionnel, tels l'ENT ou les messageries [3, 21], bien que ces données soient à relativiser selon que les enseignants travaillent dans le premier ou le second degré, mais aussi selon les contextes plus ou moins technologiques en métropole ou dans les départements d'outre-mer [3]. Des innovations par rapport aux usages antérieurs, allant au-delà de la transposition du cours présentiel sous la forme de classe virtuelle ou de cours enregistré sur YouTube, ont vu le jour. Ne pouvant parler de travail en distanciel mais plutôt à distance [1, 3], la question est de comprendre comment le contexte de pandémie a pu (ou non) favoriser une hybridation des pratiques enseignantes instrumentées. Nous prendrons comme point de départ les représentations des enseignants concernant le distanciel pendant le confinement. L'objectif est d'envisager des pistes de recherches sur la genèse des pratiques hybrides chez les enseignants travaillant dans le premier et le second degré.

Les ambiguïtés concernant le terme même d'« hybridation » nous ont amené, dans la période de formalisation qui a suivi le déconfinement, à nous intéresser à plusieurs numéros de revues concernant l'enseignement en distanciel et/ou l'hybridation des pratiques (trois numéros de la revue *Distances et médiations des savoirs* - 30, 31 et numéro à venir – ainsi que deux numéros -17 et à venir- dans RIPTU), qui renvoient à une pluralité de projets ou de dispositifs (par exemple, projet européen HySup 2009-2012 avec le numéro 301 de revue de 2014 dans *Éducation & Formation*).

2 Articulation présence-distance dans la pédagogie universitaire

Dans l'enseignement supérieur, l'analyse de l'hybridation des pratiques enseignantes remonte aux années 1980-90, avec une conceptualisation de l'hybridation et une formalisation autour des années 2000-2010 [7], puis une évolution des typologies des dispositifs de l'hybridation notamment à travers le projet européen HySup [5, 19, 20]. L'hybri-

dation passe d'abord par une alternance présence / distance en incluant (et non pas intégrant) un dispositif technique au sein de la relation enseignement-apprentissage. Cela suppose une prise de conscience de ces phases par l'enseignant « dans le scénario de son dispositif de formation par l'intermédiaire de l'environnement technopédagogique [avec] cette volonté consciente de scénariser les activités à distance » [7, p.113]. La difficulté de la mise en place de l'hybridation réside dans la diversité des formes d'hybridation : pas simplement asynchrone en distance et synchrone en présence, mais aussi des phases d'interaction synchrone avec des personnes en distanciel, ou encore, des phases comodales ou multimodales [15]. La complexité des liens qu'elle engendre en termes de médiation/médiatisation est très spécifique [7]. Elle nécessite une scénarisation des activités pédagogiques [6] et une scénarisation des interactions sociales [17]. Les relations de médiation-médiatisation se complexifient et les auteurs sont amenés à concevoir que les relations portent sur des objets différents : pour les auteurs, la médiatisation renvoie à une conception et une mise en dispositif médiatique du dispositif de formation « processus dans lequel le choix des médias les plus adaptés ainsi que la scénarisation occupent une place importante » [7, p.476]. Leur définition renvoie à un processus de production et de transformation, par un dispositif technique instrumenté, d'un changement de comportement cognitif et relationnel. La complexité de la médiation porte sur la multiplicité des dimensions auxquelles elle se réfère (une médiation technologie, mais aussi sensori-motrice, sociocognitive, relationnelle, praxéologique et réflexive) et son articulation avec la médiatisation qui transforme l'ensemble du dispositif pédagogique. L'hybridation prend aussi plusieurs formes : elle entre en rupture en modifiant (ou pas) les pratiques antérieures, ou s'ancre dans les pratiques ordinaires revêtant alors une nouvelle dimension innovante [7]). Plus tard [19], ces formes sont plutôt analysées sous l'angle d'une centration sur les apprentissages ou les enseignements (vers l'acquisition de connaissances, la mise à disposition de ressources ou d'outils d'interaction).

3 Contextualisation de la question de recherche

Ces recherches sont effectives dans l'enseignement supérieur mais les formes scolaires et universitaires ne sont pas équivalentes : à l'école, les publics sont jeunes, l'enseignant est polyvalent, l'organisation institutionnelle est hiérarchique, les liens avec les familles sont rapprochés, l'école n'a pas d'autonomie financière ni administrative et travaille en partenariat avec les collectivités territoriales, etc. [12, 24]. L'ensemble de ces paramètres contribue à modérer des pratiques et des usages, des intentions d'usage ou non usages des outils numériques [2] et de l'hybridation des pratiques instrumentées. Les synthèses de TRICOT [22, 23] et dans une moindre mesure, celle de FLUCKIGER [11, 12], montrent que la spécificité des populations (des enfants moins autonomes dans leur apprentissage) doit faire l'objet d'un effort sans précédent pour soutenir une auto-régulation de l'apprentissage à distance ; les consignes métacognitives (permettant de transférer les connaissances et savoir-faire dans d'autres contextes) doivent être imaginées en distanciel, et d'autres chercheurs complètent : l'externalisation de la forme scolaire hors

de l'école est dangereuse et repose beaucoup sur l'accompagnement parental [9], d'autant plus que le risque d'abandon est réel dans le domaine universitaire [10]. Potentiellement, certaines compétences en littératie numérique chez les élèves sont à envisager et de nouvelles compétences transversales dans la professionnalité enseignante peuvent voir le jour, sans pour autant qu'on puisse parler d'une nouvelle profession [6]. Toutefois, l'enjeu est fort car le risque est de sous-estimer la diversité des compétences professionnelles [1], et de ne pas prendre en considération les inégalités territoriales en termes d'infrastructures numériques [9].

L'articulation de la présence-distance dans le premier degré peut être étudiée à travers une multitude de recherches qui montrent une grande diversité dans les pratiques enseignantes instrumentées et dans des dispositifs d'ingénieries pédagogiques. Par exemple, parce qu'elles permettent de dépasser les frontières dans et hors classe, les tablettes tactiles peuvent jouer un rôle important dans la socialisation des enfants présentant des besoins éducatifs particuliers [16]. Elles permettent aussi de faire dialoguer les espaces (familles/éducateurs) et de mieux répondre aux besoins éducatifs des élèves. Dans le cadre de l'usage des tablettes tactiles, VILLEMONTÉIX et al. [24] rappellent que les spécificités techniques –affordance- de la tablette (support mobile, multifonctionnalités communicationnelles accessibles à travers les images, le son, la vidéo), sont des facteurs facilitant l'intégration dans l'apprentissage. L'usage des tablettes favorise les liens entre les espaces (dans les dispositifs spécialisés, en classe ordinaire, en famille) et participe à la complexification des pratiques et des ressources à plusieurs niveaux (écosystémique, institutionnelles, pédagogiques) : les ingénieries pédagogiques passent par un bricolage infrastructurel pour détourner l'ensemble des contraintes imposées par les caractéristiques de la tablette (exemple cité des Ipad qui ne permettent pas des affichages simultanés de plusieurs applications).

Quelques travaux de recherches sur la notion de médiation-médiatisation à l'école et dans les pratiques pédagogiques des enseignants permettent de recontextualiser ce qui a pu être formalisé au niveau universitaire. On entend par médiation l'action faite par un intermédiaire qui rend ainsi possible le déploiement d'activités et l'inclusion dans une culture. C'est un système complexe qui passe par des négociations interpersonnelles des postures (aider l'autre à aller au-delà de ce qu'il sait faire seul), des formes interactionnelles (pour permettre à l'enseignant de faire accepter à l'élève la responsabilité de la tâche -dévolution), des valeurs (regarder l'autre dans ses potentialités), des gestes professionnels pour mettre en situations les conditions qui rendent favorables l'apprentissage et son appropriation, des outils (notamment le langage), des ressources, des mises en place de situations pédagogiques qui dépassent la transmission [26]. Les changements de comportement ou les changements cognitifs ne sont pas des buts en soi puisqu'ils participent au processus de médiation. C'est l'artefact technique ou le dispositif médiatisé qui prend un sens psychologique pour les acteurs et la situation d'usage au sein du système scolaire dans lequel il se trouve.

Ainsi, questionner la manière dont les enseignants se représentent le distanciel est un premier outil d'accès à la conceptualisation de l'hybridation dans un sens contextuel et culturel, adapté au premier degré.

4 Conclusions et perspectives

La spécificité du contexte de confinement et du distanciel « forcé » a modifié profondément les pratiques des enseignants (1re et 2nd degré) vers une hybridation des contextes ainsi que leurs pratiques instrumentées, considérant que le confinement est le déclencheur d'une nouvelle approche de « l'hybridation », pour le premier et le second degré. Ce constat peut être discuté, et semble envisagé par d'autres chercheurs [1, 3, 4, 21]. La construction de l'hybridation des pratiques enseignantes, des outils, des ressources, dans les premier et second degré, prend alors une tournure très spécifique, dont il s'agit de mesurer les impacts. L'objectif est ainsi de questionner, dans les contextes hors université, de nouvelles formes d'hybridation.

Dans un rapport de recherche [4], les premières données décrivent une forte résistance au distanciel dans les pratiques quotidiennes déclarées (le numérique ne remplace pas l'enseignant pour 79,7 % des répondants) et des représentations hétérogènes (le numérique sert à communiquer -50,6 % de l'échantillon-, il peut être un levier dans les changements de pratiques -16,8 % de l'échantillon-). Cette première forme de représentation du distanciel, avec des effacements de soi, mais aussi, un éparpillement des outils et des ressources ainsi que des fortes dispersions des usages, sera à approfondir. Ces premières descriptions permettent d'envisager une méthodologie de recherche au plus près des pratiques quotidiennes dans et hors la classe, dans un esprit collaboratif, pour favoriser le mieux-être de l'enseignant et de l'élève, et lever en partie les résistances actuelles face au distanciel forcé. La prise de conscience des pratiques hybrides ordinaires, par un travail sur l'épistémologie pratique de l'enseignant [3], semble être une piste à privilégier pour le premier degré.

Références

1. Mercier, C., Lefer Sauvage, G.: L'école à l'heure de la COVID-19 : des constats de pratiques vers une nouvelle forme scolaire ? Formation et profession (2021, à paraître).
2. Lefer Sauvage, G.: Intention d'utilisation des tablettes en formation initiale d'enseignants à Mayotte. Communication orale pour EIAH (2021, juin).
3. Lefer Sauvage, G., Mercier, C.: Articuler présence et distance à la suite du confinement : genèse des pratiques hybrides ? Communication orale pour EIAH (2021, juin)
4. Genevois, S., Lefer Sauvage, G., Wallian, N.: Questionnaire d'enquête auprès des enseignants "Confinement et continuité pédagogique", Rapport de recherche (2020). Hal02934483.
5. Burton, R., Borrat, S., Charlier, B. et al.: Vers une typologie des dispositifs hybrides de formation en enseignement supérieur. Distances et savoirs, 1(9), 69-96 (2011).
6. Caron, A., Gauthier, M.-C., Provencher, J., Robillard, R.: L'enseignement hybride ou à distance au primaire. Québec : Clothilde Bariteau (2020). <http://plusinteractif.s3.amazonaws.com/django-assets/media/pedagogie/PAD-Primaire.pdf>
7. Charlier, B., Deschryver, N., Peraya, D. : Apprendre en présence et à distance : une définition des dispositifs hybrides. Distances et savoirs, 4(4), 469-496 (2006).
8. DEPP.: Note d'information n°20-26. [En ligne] <https://www.education.gouv.fr/crise-sanitaire-de-2020-et-continuite-pedagogique-les-eleves-ont-appris-de-maniere-satisfaisante305214>.

9. Devauchelle, B.: Peut-on imposer l'enseignement à distance? Café pédagogique (novembre 2020), [En ligne]. <http://www.cafepedagogique.net/lexpresso/Pages/2020/11/27112020Article637420575509707183.aspx>
10. Dussarps, C.: Étude communicationnelle de l'abandon en formation à distance universitaire. Adjectif.net, 2017 [En ligne] <http://www.adjectif.net/spip/spip.php?article422>
11. Fluckiger, C.: L'école à l'épreuve de la culture numérique des élèves. Revue française de pédagogie, n°163, 51-61 (2008).
12. Fluckiger, C.: La forme universitaire comme analyseur des « effets » de la technologie : perspective critique, Distances et médiations des savoirs, 22, 2020 [En ligne] <http://journals.openedition.org/dms/2329>
13. IFE: Les effets du confinement sur l'activité des enseignants du primaires et secondaire. [Rapport de recherche]. [http://ife.ens-lyon.fr/ife/recherche/groupes-de-travail/documentation-confinement-et-enseignement/synthese-enseignants\(2020\)](http://ife.ens-lyon.fr/ife/recherche/groupes-de-travail/documentation-confinement-et-enseignement/synthese-enseignants(2020)).
14. IGEN.: Les usages pédagogiques du numérique au service de la réussite des élèves, n°130. [En ligne] <https://www.education.gouv.fr/les-usages-pedagogiques-du-numerique-en-situation-pandemique-durant-la-période-de-mars-juin-2020-308421>.
15. Guillot, J-M.: Quelles modalités pédagogiques pour la rentrée 2020 : hybride, multi ou comodalité ? <https://tipes.wordpress.com/2020/08/04/quelles-modalites-pedagogiques-pour-la-entree-2020-hybride-multi-ou-comodalite/>
16. Mercier, C.: La construction et les effets de l'appropriation d'un outil numérique auprès des enfants avec autisme en IME : interactions en situation d'apprentissage en lien avec l'utilisation d'un agenda numérique. Thèse de doctorat, Université de Nantes, Nantes (2017).
17. Nissen, E.: Mise à l'épreuve de paramètres pour une articulation réussie du distanciel et du présentiel aux yeux des étudiants. Distances et médiations des savoirs [En ligne], 30, 2020. <http://journals.openedition.org/dms/5007>
18. Peraya, D.: L'ingénierie pédagogique en 2020. Au-delà de la crise sanitaire faire une place à l'apprenant. Distances et médiations des savoirs, 30 (2020).
19. Peraya, D., Charlier, B., Deschryver, D.: Une première approche de l'hybridation. Education & formation, 301, 15-34 (2014).
20. Peraya, D., Peltier, C.: Typologie des dispositifs hybrides : configurations et types. Dans N. Deschryver, et B. Charlier, Dispositifs hybrides, nouvelle perspective pour une pédagogie renouvelée de l'enseignement supérieur. Rapport final [En ligne], 54-86 (2012). <http://archive-ouverte.unige.ch/unige:23091>
21. Ria, L., Rayou, P.: La forme scolaire en confinement. Enseignants et parents à l'épreuve de l'enseignement à distance. Dossier Pandémie. Formation et Profession, 28(4), 1-11 (2020). doi : 10.18162/fp.2020.675
22. Tricot, A.: Se former : quelles nouvelles complémentarités entre présence et distance ? Communication orale pour Canopé. https://www.youtube.com/watch?v=N0fr1pclbpw&feature=emb_logo (2020a).
23. Tricot, A.: Numérique et apprentissages scolaires, Rapport pour le Cnesco, Paris : Cnesco (2020b). <http://www.cnesco.fr/fr/numerique-et-apprentissages-scolaires/>
24. Villemonteix, F., Hamon, D., Nogry, S., Séjourné, A., Hubert, B., Gélis, J.-M.: Extate : Expérience tablettes tactiles à l'école primaire (p.86). Université Cergy-Pontoise, Ministère de l'Éducation Nationale (2014).
25. Vidal, M.: « L'enseignement à distance, trait d'union en temps de pandémie : un regard sur les publications scientifiques sur l'EAD, de février à novembre 2020 ». Distances et médiations des savoirs, 30 (2020). <http://journals.openedition.org/dms/5721>
26. Vinatier, I., Laurent, J-M.: Médiation, enseignement-apprentissage. Présentation du dossier, La nouvelle revue de l'adaptation et de la scolarisation, 2(42), p. 5-14 (2008).

Vers une méthode de méta-design de jeux sérieux : Application pour l'apprentissage de la programmation à travers Blockly Maze

Bertrand Marne¹, Karim Sehaba^{2,3,4}, and Mathieu Muratet^{3,4}

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

² Université de Lyon, CNRS, Université Lyon 2, LIRIS, UMR5205, F-69676, France
karim.sehaba@liris.cnrs.fr

³ Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

⁴ INS HEA, 92150 Suresnes, France
mathieu.muratet@lip6.fr

Résumé. Pour faciliter l'appropriation des jeux sérieux d'apprentissage par les enseignants, nous étudions les méthodes de conception qui leur laissent une grande place à la fois dans les phases de conception et d'utilisation : les méthodes de méta-design. Dans ce cadre, nous nous intéressons au jeu sérieux Blockly Maze destiné à l'apprentissage de la programmation. Notre objectif est ainsi de proposer des modèles et outils de conception de niveaux et de scénarios pour ce jeu et de fournir à l'enseignant des indicateurs de suivi lui permettant de s'approprier et d'adapter le jeu en fonction des usages observés. Dans cet article, nous détaillons les premiers apports de cette étude (modèles, outils, indicateurs de suivi) et les premiers résultats de nos observations sur ces contributions.

Mots-clés : jeux sérieux · méta-design · pensée informatique · suivi · outils auteurs.

Abstract. To support the appropriation of serious learning games by teachers, we are studying design methods that centers them in both the design and use phases: meta-design methods. Our objective is to propose models and tools for designing levels and scenarios for Blockly Maze and to provide the teacher with monitoring indicators allowing him/her to appropriate and adapt the game according to the observed uses. In this article, we detail the first contributions of this study (models, tools) and the first results of our observations on these contributions.

Keywords: Serious Games · meta-design · computational thinking · monitoring · authoring tools.

1 Introduction

Les jeux d'apprentissage proposent de nombreux avantages par rapport aux outils d'enseignement classique. En effet, de nombreux auteurs considèrent les

jeux sérieux d'apprentissage comme prometteurs, notamment pour augmenter l'engagement et la motivation des apprenants [1], d'autres envisagent ces outils pour favoriser un apprentissage plus constructiviste [7]. Pourtant, leur adoption, notamment par les enseignants, reste rare [5]. Pour favoriser cette adoption, nous faisons l'hypothèse qu'une méthode de conception participative telle que le méta-design pourrait convenir. Le méta-design est une méthode de conception participative avancée, dans laquelle les utilisateurs finaux (« *owners of problems* ») sont impliqués de façon centrale dans les phases initiales de conception. Mais, ils doivent aussi avoir les moyens de continuer à être concepteurs durant les phases d'utilisation des artefacts, grâce à l'*underdesign* que Fischer et al. [3] définissent comme : « [...] *underdesign aims to provide social and technical instruments for the owners of problems to create the solutions [of their problems] themselves at use time.* »

Nos travaux visent à s'interroger sur cette approche dans le contexte de l'utilisation des jeux sérieux par les enseignants en tant que *owners of problems*. Plus précisément, les travaux présentés ici rapportent l'amorçage d'un projet dans lequel nous nous intéressons à l'enseignement de la pensée informatique, et à l'étude de la mise en place d'une démarche de méta-design avec des jeux sérieux. L'objectif du projet est d'exploiter un jeu sérieux existant et d'étudier dans quelle mesure il est possible de le rendre « appropriable » par des enseignants (*underdesign*), notamment en leur permettant d'en modifier le contenu en fonction de leurs besoins pédagogiques et ses usages observés ou envisagés. Nous avons choisi Blockly Maze⁵ (BM) notamment parce que c'est un logiciel libre et qu'il est réutilisé dans de nombreux travaux dérivés, comme l'*Heure de Code*⁶, *Algorea*⁷, et plus de 400 dérivés sur la forge logicielle Github⁸. Les limites de BM sont les mêmes limites que celles que l'on retrouve dans beaucoup de jeux sérieux : (1) **nombre de niveaux limité**, BM propose 10 niveaux abordant peu de notions de programmation ; (2) **scénarisation non modifiable**, le parcours de ces 10 niveaux est une succession imposée que l'enseignant ne peut modifier ; (3) **pas de collecte et analyse de traces sur l'activité des joueurs**, l'enseignant ne récupère pas de traces des travaux des apprenants/joueurs. Ainsi, il est difficile pour l'enseignant de comprendre les pratiques et difficultés de ses apprenants et ainsi de s'approprier et adapter le jeu.

Notre objectif est de proposer des outils et des méthodes permettant à des enseignants, plus ou moins à l'aise avec l'enseignement de l'informatique, de prendre en main BM, de construire des niveaux complétant ceux qui sont disponibles, et des successions de niveaux correspondant à leurs besoins. Nous voulons donc, à terme, instaurer un contexte de méta-design autour d'un modèle pour la conception et le développement de niveaux. Dans cet article, nous présentons BM et les problématiques scientifiques qui émergent des différents verrous issus de la littérature et de la pratique de l'enseignement de la pensée informa-

⁵ développé par Google : <https://blockly.games/maze> consulté le 28/04/2021

⁶ <https://code.org/> consulté le 28/04/2021

⁷ <https://algorea.org/> consulté le 28/04/2021

⁸ <https://github.com/google/blockly-games/network/members> (28/04/2021)

tique. Ensuite, nous présentons notre contexte méthodologique et nos premiers résultats. L'article se termine par une conclusion et des perspectives.

2 Présentation du jeu sérieux et de nos problématiques

Le principe de BM est d'enseigner la programmation par le biais d'activités de guidage d'un avatar dans un labyrinthe. Le guidage se fait par une programmation par blocs d'instructions inspirée de Scratch ⁹ Cependant, BM ne propose que 10 niveaux. Ils entraînent notamment à l'utilisation de séquences d'instructions, des conditions (`if/else`), des boucles (`repeat/until`), y compris imbriquées, et des rudiments d'algorithmique. C'est assez peu compte tenu des besoins variés dans l'enseignement de la pensée informatique [2]. Ainsi, de nombreux dérivés (*forks*) et jeux sérieux inspirés ont fleuri, proposant de nombreux nouveaux niveaux. Nous avons déjà cité l'*Heure de Code* qui propose de très nombreux niveaux complémentaires abordant de multiples aspects de la pensée informatique répartis dans divers « cours », et parcours d'apprentissages. L'offre de scénarios clés en main proposée par ce jeu est donc très diversifiée (mêlant d'autres types de jeux sérieux que BM), si bien qu'un choix éclairé de la ressource ou du parcours à utiliser demande d'acquérir une vision globale du panorama de ce qui est proposé et donc requiert un temps d'appropriation de la plateforme important. *Algoréa* [4] propose aussi des jeux proches de BM. Ces jeux peuvent tous utiliser des blocs comme Blockly (en complément du Python et du Scratch), mais peu utilisent Maze (sortie d'un labyrinthe), et dans ce cas ils portent seulement sur certains aspects de la pensée informatique : séquences d'instruction, appels de fonctions, boucles simples ou imbriquées, instructions conditionnelles.

En tout état de cause, il est difficile pour un enseignant connaissant ces ressources de constituer un parcours d'apprentissage pour ses apprenants en piochant ce dont il a besoin parmi elles. Elles ne sont pas conçues pour cette modularité, mais plutôt pour être utilisées sous forme de séquences assez longues et indivisibles de plusieurs niveaux. Afin de permettre une appropriation d'un jeu sérieux par la maîtrise et l'adaptation, nous nous sommes appuyés sur nos travaux antérieurs : l'outil auteur APPLiQ, comme nous l'avons montré précédemment [6], peut être utilisé pour scénariser BM. APPLiQ permet aux enseignants/utilisateurs de préparer et proposer aux apprenants/joueurs des successions de niveaux dans un jeu sérieux en tenant compte des objectifs pédagogiques prérequis et travaillés à chacune de ces étapes. Cette succession de niveaux n'est pas forcément linéaire, et peut donc dépendre des actions et des performances de l'apprenant/joueur (objectifs travaillés ou non dans les activités).

Cependant, APPLiQ a plusieurs limites. Dans le contexte de notre problématique, la limite la plus importante est qu'il ne permet que de modifier l'organisation des niveaux *entre eux*. Ainsi, si les enseignants/utilisateurs peuvent en modifier l'ordre et la succession, ils ne peuvent pas modifier les niveaux de

⁹ <https://scratch.mit.edu> consulté le 28/04/2021

BM eux-mêmes dans APPLiQ. APPLiQ prévoit bien la possibilité d'ajouter de nouveaux niveaux dans la description du modèle de scénario d'un jeu sérieux donné. Mais, il incombe à l'utilisateur qui fait cette modification dans APPLiQ d'opérer les adaptations du côté du jeu sérieux lui-même. Par conséquent, choisir APPLiQ comme outil pour l'instrumentalisation de BM par des enseignants suppose d'une part de le compléter par un autre outil dédié à l'édition de niveaux, et d'autre part de faire une mise à jour de BM pour l'y adapter. Pour ces travaux de modélisation de jeux sérieux et d'outillage des enseignants, nous avons inscrit notre démarche dans le contexte d'une recherche collaborative orientée par la conception avec 3 axes principaux décrits dans la section suivante.

3 Méthode et méthodologie

Pour réaliser notre objectif, notre démarche consiste, dans un premier temps, avec des enseignants, à réfléchir et conceptualiser la pensée informatique, pour des besoins pédagogiques, sous forme d'une carte conceptuelle. Pour mettre en place cette carte¹⁰, nous avons mené des entretiens avec deux enseignantes : l'une est enseignante chevronnée de mathématiques et enseigne aux enfants du secondaire hospitalisés et l'autre est la directrice de l'enseignement aux enfants de ce même hôpital.

La construction de cette carte a surtout été l'outil discursif principal que nous avons utilisé pour aborder de nombreux thèmes avec les enseignantes avec lesquelles nous avons travaillé. La carte n'a donc pas vocation à modéliser la pensée informatique de manière exhaustive ou consensuelle, mais nous permet de comprendre les démarches adoptées par les enseignantes pour construire des niveaux/scénarios et d'identifier d'éventuelles difficultés auxquelles elles pourraient faire face. La synthèse des différents entretiens effectués autour de la carte conceptuelle avec chaque enseignante nous a permis d'identifier en particulier :

- Les notions enseignées et la méthodologie adoptée par les enseignantes.
- L'intérêt pour la narration et les jeux comme moyen favorisant l'apprentissage.
- Le besoin de personnaliser les niveaux de BM pour, notamment, adapter la reformulation de l'énoncé, les blocs, les dialogues, rajouter des fonctionnalités (variables, timer, etc.).
- Le besoin d'indicateurs de suivi pédagogique permettant aux enseignantes de mesurer la progression de leurs apprenants et d'identifier leur appropriation des notions étudiées, et d'indicateurs de méta-design leur facilitant la mise à jour des niveaux/scénarios en fonction des usages observés.

Cette première étude nous a conduits à spécifier 3 axes regroupant les principales fonctionnalités de l'outil auteur dédié à la conception et au méta-design : (1) **Outil de conception de niveaux** ; (2) **Outil de conception de scénarios** ; (3) **Outils de suivi**. Ces outils sont en cours de développement en lien avec des enseignants sur le terrain.

¹⁰ Voir la carte conceptuelle complète : <https://mycore.corecloud.net/index.php/s/NsntsseDxGUILM4/download>. Consultée le 28/04/2021.

4 Premiers travaux et résultats

L'outil de conception de niveaux permettra à l'enseignant de mettre en place de nouveaux niveaux ou de réutiliser des niveaux existants selon un modèle et une méthodologie simples et intuitifs. Il s'agit donc, pour chaque niveau, de pouvoir (re)définir les labyrinthes eux-mêmes, les blocs mis à disposition de l'apprenant, des dialogues et les conditions de leurs déclenchements, ainsi que d'autres contraintes : nombre de blocs utilisables, durée de résolution, blocs déjà posés, etc. Techniquement, la modélisation des niveaux de BM s'est faite en deux étapes : (1) isolation dans le code source de BM de tous les éléments qui relevaient de la description des niveaux, et (2) conceptualisation de ces extraits de code source sous forme d'un modèle UML.

L'outil de conception de scénarios permet à l'enseignant de définir l'orchestration des niveaux et les conditions de leurs déclenchements. Il doit permettre de définir une succession de niveaux spécifiques pour chaque apprenant en fonction de ses performances. Pour cet usage, APPLiQ nous a semblé pertinent [6]. Il propose un modèle XML de la scénarisation (*MoPPLiQ*) dans lequel les niveaux (*activités*) sont considérés comme des boîtes noires, permettant de travailler ou non certains objectifs pédagogiques et ludiques. Chaque activité peut avoir plusieurs états de sortie en fonction de la combinaison des objectifs travaillés. Chaque activité peut aussi avoir plusieurs états d'entrée restreignant les possibilités de branchement de l'activité au reste du scénario en fonction d'objectifs prérequis. Le scénario (*parcours*) est donc décrit comme une succession de toutes les combinaisons de relations états de sortie-états d'entrée désirées.

L'outil de suivi devra permettre de collecter les données représentant les interactions entre l'apprenant et les niveaux/scénarios et de calculer des indicateurs portant sur l'apprentissage et la qualité des ressources pédagogiques. Il s'agit d'indicateurs de haut niveau comme le taux de réussite, les niveaux terminés ou non, les temps de réalisation, ou plus finement comme le nombre de blocs utilisés ou le nombre d'aides affichées. Pour fonder le système de suivi, nous avons choisi d'utiliser une version de *xAPI* adaptée aux jeux sérieux. *xAPI* permet de définir les actions de suivi sous forme de *statements* (triplets {**acteur**, **verbe**, **objet**}) qui sont stockés dans un *Learning Record Store* (LRS). Nous avons identifié deux niveaux différents de suivi : le suivi *dans* le niveau, et le suivi *entre* les niveaux.

5 Conclusions et perspectives

Pour en faciliter l'adoption et l'appropriation, notre objectif est de permettre le méta-design [3] de jeux sérieux visant l'apprentissage de la pensée informatique comme BM. C'est-à-dire de permettre aux enseignants/utilisateurs d'être concepteurs de ces jeux sérieux à la fois dans les temps de conception initiale que dans les temps d'utilisation. Nous avons décrit ici un état de nos travaux en cours. Il s'agit de la déconstruction de BM, à la fois d'un point de vue conceptuel et d'un point de vue logiciel. Notre objectif étant d'obtenir, avec l'aide

d’enseignantes participantes, des modèles explicites des niveaux et contenus pédagogiques. À terme, il s’agit de fournir un BM utilisable, mais proposant surtout l’essentiel des éléments nécessaires pour offrir une appropriation facilitée aux enseignants. Nous avons exploré une méthode de recherche, en collaboration avec les enseignantes/contributrices, centrée sur de la conception d’une carte conceptuelle d’un de leur domaine d’enseignement (la pensée informatique). La conception de cette carte a créé l’occasion de faire émerger de nombreuses problématiques liées à leurs enseignements ainsi que de faire émerger des besoins autour de BM. À partir de ces besoins, et toujours en collaboration avec ces enseignantes, nous avons mis en place 3 modèles décrivant BM. Puis, nous les avons implémentés dans une nouvelle version de BM dont nous proposerons les modifications aux auteurs (*pull request*). Le **modèle de niveau**, implémenté en JSON, s’est montré capable de décrire l’intégralité des niveaux actuels de BM, ainsi que d’être utilisé pour en concevoir de nouveaux avec les enseignantes. Pour le **modèle de scénario**, nous choisissons APPLiQ et envisageons de le faire évoluer pour répondre à nos questions de recherche [6]. Nous avons amorcé le travail sur le **modèle de suivi** autour de *xAPI*. Nos travaux actuels et futurs consistent à travailler avec d’autres enseignants à (1) l’élaboration de niveaux afin d’enrichir et d’éprouver notre modèle de niveaux de BM ; (2) l’amélioration de la scénarisation, et à son expérimentation ; (3) la mise en place du suivi des apprenants/joueurs pour permettre aux enseignants de valider leurs modifications de BM.

Références

1. Bouvier, P., Sehaba, K., Elise, L.: A trace-based approach to identifying users’ engagement and qualifying their engaged-behaviours in interactive systems. Application to a social game. *User Modeling and User-Adapted Interaction (UMUAI’14)* **45**, 413–451 (Dec 2014)
2. Brunet, O., Yessad, A., Muratet, M., Carron, T.: Vers un modèle de scénarisation pour l’enseignement de la pensée informatique à l’école primaire. In: *Didapro 8 – DidaSTIC*. Lille, France (Feb 2020)
3. Fischer, G., Giaccardi, E., Ye, Y., Sutcliffe, A.G., Mehandjiev, N.: Meta-design: A manifesto for end-user development. *Communications of the ACM* **47**(9), 33–37 (2004)
4. Léonard, M.: Score et chronomètre sur l’interface: Quels effets sur les résultats et la perception d’un concours en ligne? In: *8e Rencontres Jeunes Chercheurs En EIAH*, 2020. Poitiers (Jun 2020)
5. Li, Q.: Enactivism and teacher instructional game building: An inquiry of theory adoption and design consideration. *Educational Technology Research and Development* **66**(6), 1339–1358 (2018)
6. Marne, B.: Modèles et outils pour la conception de jeux sérieux : une approche meta-design. Thèse de Doctorat en Informatique, Université Pierre et Marie Curie (UPMC), Paris (Jul 2014)
7. Marne, B.: Éléments de réflexion autour de la conception du jouet pour un jeu sérieux. In: *Actes de la 9ème Conférence sur les Environnements Informatiques pour l’Apprentissage Humain*. pp. 1–12. Paris, France (Jun 2019)

Vers l'utilisation de pictogrammes affichables sur des tablettes à picots rétractables pour l'illustration de livres tactiles : importance de la résolution de la tablette

Carolane Mascle^{1,2}, Gwenaël Kaminski¹, Florence Bara¹, Christophe Jouffrais²

¹ Université Toulouse Jean Jaures, CNRS, CLLE UMR 5263, Toulouse, France
prenom.nom@univ-tlse2.fr

² Université Paul Sabatier, CNRS, IRIT UMR 5505, Toulouse, France
prenom.nom@irit.fr

Résumé. Les livres illustrés occupent une place particulière dans le développement de l'enfant. Cependant peu de livres adaptés sont disponibles pour les enfants avec une déficience visuelle. De nouvelles technologies telles que les tablettes à picots rétractables offrent la possibilité de créer facilement des illustrations dynamiques pour de nombreux livres. Les illustrations seraient basées sur un ensemble de points surélevés, représentant une forme. Dans cette étude, nous avons évalué la reconnaissance de stimuli réalisés avec des points par 28 enfants voyants et non-voyants. Nous avons observé que des pictogrammes en pointillés « haute résolution » peuvent être bien reconnus par des enfants non-voyants. En comparant notre étude avec la littérature, nous mettons en évidence l'importance de la taille et de la proximité des points, avec des taux de reconnaissance plus élevés pour des pictogrammes illustrés avec une meilleure résolution.

Mots-clés : Déficience visuelle, tactile, haptique, technologie, interaction homme-machine.

Abstract. Illustrated books hold a special place in the development of children. However, few suitable books are available for children with visual impairments. New technologies such as retractable pin tablets make it possible to easily create dynamic illustrations for many books. The illustrations would be based on a set of raised points, representing a shape. In this study, we evaluated the recognition of stimuli achieved with raised dots by 28 sighted and blind children. We have observed that "high resolution" dotted pictograms can be well recognized by blind children. Comparing our study with the literature, we highlight the importance of the size and proximity of the dots, with higher recognition rates for pictograms illustrated with better resolution.

Keywords: Visual impairment, tactile, haptic, technology, human-machine interaction.

1 Introduction

Les livres illustrés occupent une place particulière dans le développement de la littéracie précoce des enfants [1]. Cependant, la création d'images tactiles prend du temps et repose sur une expertise spécifique qui rend les livres tactiles particulièrement coûteux. Les nouvelles technologies telles que les tablettes à picots rétractables pourraient offrir la possibilité de créer des illustrations pour plusieurs livres avec le même dispositif. Ces tablettes sont basées sur des matrices de picots qui montent et descendent dynamiquement en fonction du contenu à afficher. La tablette se mettrait à jour lorsque l'enfant tourne les pages et les personnages pourraient apparaître, disparaître ou se déplacer au fil des pages et de l'histoire. De plus, les enfants, les enseignants ou les parents pourraient facilement créer leurs propres images avec un logiciel dédié.

Différentes tablettes dynamiques à picots rétractables ont été développées. Cependant, du fait de contraintes technologiques, ces tablettes présentent des résolutions différentes (taille et espacement des picots ; une tablette avec des picots de 1 mm espacés de 1 mm serait une tablette « haute résolution » tandis qu'une tablette avec des picots de 4 mm espacés de 4 mm serait de « basse résolution »).

Différentes tablettes ont été évaluées pour l'affichage de formes géométriques simples avec des taux de reconnaissance très variables (de 60% à 80%) [2][3][4]. Les différences de taux de reconnaissance entre ces études pourraient s'expliquer par la résolution des tablettes et en particulier l'espace entre les points. En effet, dans les études de Leo et al. [2] et Bellik et Clavel [3] qui observent des taux de reconnaissances élevés chez des participants voyants (resp. 90% et 80%), l'écart entre les picots des dispositifs utilisés était de 1 à 1,5mm. Tandis que dans l'étude de Velazquez et al. [4] qui ont observé un taux de reconnaissance plus faible (60%) l'écart était de 2,5mm.

Il semble que ce type de dispositif puisse permettre de reconnaître la majorité des formes affichées. Cependant, si les tablettes à picots rétractables ouvrent des perspectives intéressantes, rien ne permet pour l'instant d'affirmer qu'elles pourraient être efficaces pour créer des illustrations tactiles. Une étude récente chez des enfants voyants et non-voyants de 6 à 10 ans, évalue la reconnaissance de pictogrammes en pointillés [5]. La taille et l'écartement des points étaient basés sur la tablette Blindpad [6] (points de 4 mm espacés de 4 mm). Les auteurs ont observé un taux de reconnaissance moyen de 50% (pas de différence significative entre voyants et non-voyants). Ces résultats sont très inférieurs aux performances observées dans les études présentées précédemment [2][3][4]. Nous proposons plusieurs hypothèses pour expliquer ce résultat : l'âge des participants, l'utilisation de pictogrammes plutôt que des dessins classiques, l'affichage de ces pictogrammes à l'aide de points (plutôt que des lignes continues) et l'écartement entre les points.

Concernant l'âge des participants, les études présentées précédemment portaient sur des adultes et non sur des enfants. Or, les performances haptiques des enfants sont inférieures à celles des adultes [7]. Il est également possible que l'utilisation de pictogrammes ne soit pas adaptée pour la perception tactile des enfants. En effet, les enfants ont tendance à mettre en place des stratégies de « devinette » lors de l'identification haptique d'images en relief en se basant sur une seule caractéristique de l'objet, comme par exemple les cornes de la girafe [8]. Les pictogrammes ne reprennent que certaines

caractéristiques spécifiques de l'objet représenté (*e.g.* les ailes pour l'oiseau qui n'a alors ni bec ni pattes). Il est possible que les pictogrammes ne soient pas assez détaillés pour permettre aux enfants d'utiliser des caractéristiques locales de l'objet. En d'autres termes, si l'enfant ne comprend pas que ce sont les ailes de l'oiseau, il n'a pas d'autres indices locaux lui permettant de reconnaître l'image.

L'affichage des formes en pointillés (plutôt que des lignes continues) imposé par l'utilisation de tablettes à picots pourrait être un problème. En effet, il est nécessaire de percevoir les points distincts comme formant un tout. Ce principe de proximité de la Gestalt est applicable au toucher [9], bien qu'il soit moins sensible que la vision à ce principe [10]. Il est donc possible que l'utilisation de pictogrammes en pointillés soit un frein à la reconnaissance.

De plus, ces études ne mentionnent pas de seuil critique à partir duquel l'écart entre les points peut rendre difficile la perception de formes distinctes comme appartenant à un tout. L'écart entre les points pourrait donc expliquer les différences de reconnaissance tactile. Dans l'étude de Mascle et al. [5], l'écart entre les points était de 4 mm, soit la borne maximale des recommandations. Pour les autres études utilisant des tablettes à picots rétractables [2][3][4], l'écart entre les points était compris entre 1.5 et 2.5 mm.

Enfin, dans ces études aucun effet du statut visuel (voyants, non-voyants) n'a été observé. Cependant, il est possible que des pictogrammes affichés sur des tablettes proches du format braille (points de 1 mm espacés de 1 mm) soient mieux reconnus par les enfants non-voyants en plein apprentissage de la lecture en braille.

Dans cette étude, nous avons comparé les taux de reconnaissance de pictogrammes en pointillés et de pictogrammes en lignes chez des enfants voyants et non-voyants pour répondre à la question de recherche suivante : des pictogrammes tactiles réalisés en traits continus ou avec des pointillés peuvent-ils être utilisés pour afficher des illustrations adaptées pour des enfants ? Nous avons fait l'hypothèse supplémentaire que les enfants non-voyants devraient mieux reconnaître les pictogrammes en pointillés que les voyants puisque la résolution utilisée est proche de celle du Braille. Finalement, en comparant nos résultats à ceux de la littérature, nous avons aussi pu évaluer que le taux de reconnaissance dépend de la résolution avec laquelle sont affichés ces pictogrammes en pointillés.

2 Matériel et méthode

2.1 Participants et matériel utilisé

Cette étude inclut 20 enfants voyants travaillant sans voir de 6 à 9 ans pour un âge moyen de 8 ans (97 mois, $SD = 6$) et 8 enfants non-voyants de 6 à 10 ans pour un âge moyen de 7,6 ans (90 mois, $SD=9$).

Nous avons repris les pictogrammes en pointillés de l'étude de Mascle et al. [5]. Nous avons formé des pictogrammes en lignes continues en reliant tous les points des pictogrammes et avons lissé les courbes pour éviter les aspérités (Fig 1). Nous avons aussi réalisé des pictogrammes basés sur le format de tablettes dites « haute résolution » (*i.e.* point de 1 mm espacés de 1mm) ; c'est à dire proche du format du braille (Fig 1).

Pictogrammes liste 1			Pictogrammes liste 2			Pictogrammes liste 3		
Mot	Points	Lignes	Mot	Points	Lignes	Mot	Points	Lignes
Chat			Oiseau			Mouton		
Fraise			Poisson			Fleur		
Sac			Lunettes			Chapeau		
Collier			Ceinture			Crayon		
Poupée			Ballon			Sapin		
Bourdon			Hot dog			Python		
Chou-fleur			Espadrille			Manette		
Poignard			Nectarine			Jogging		
Cintre			Enceinte			Balayette		
Sparadrap			Caddie			Epuisette		

Fig 1. Matériel utilisé

2.2 Méthode

La procédure comprenait deux étapes : (i) une phase d'apprentissage dans laquelle les 10 pictogrammes ont été présentés dans un ordre aléatoire. Le mot était indiqué à l'oral aux enfants qui étaient ensuite libres d'explorer le pictogramme ; (ii) une phase de rappel dans laquelle les pictogrammes appris lors de la première phase ont été proposés aléatoirement un par un. Les enfants devaient alors explorer le pictogramme et l'identifier aussi rapidement et précisément que possible.

Le choix des listes et l'ordre de présentation des conditions ont été contrebalancés parmi tous les participants. Nous avons comparé la précision des réponses pour les deux conditions à l'aide d'un modèle linéaire généralisé à effets mixtes. Le statut visuel et la condition ont été définis comme effets fixes, et les participants et les items comme effets aléatoires.

3 Résultats

Les analyses montrent un effet principal du statut visuel $\chi^2(1, N=28) = 5.502, p=.019$ et de la condition $\chi^2(1, N=28) = 18.688, p <.001$, mais pas d'effet d'interaction entre ces deux variables $\chi^2(1, N=28) = 0.176, p=.067$.

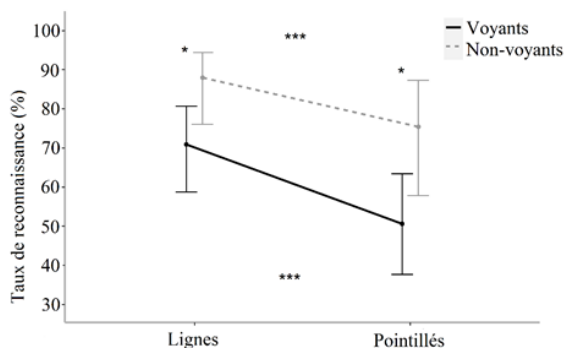


Fig 2 : Taux de reconnaissance et $IC_{95\%}$ en fonction de la condition et du statut visuel

Les pictogrammes en lignes sont mieux reconnus que les pictogrammes en pointillés par les deux groupes (Fig 2). De plus, les enfants non-voyants ont de meilleures performances que les enfants voyants pour les pictogrammes en lignes et les pictogrammes en pointillés (Fig 2). Ce résultat peut démontrer une certaine expertise des enfants non-voyants pour l'exploration et l'identification de contenus tactiles et ne serait pas uniquement lié à l'utilisation du braille.

4 Discussion

Le taux de reconnaissance des pictogrammes en ligne est particulièrement élevé chez les enfants voyants et non-voyants (Fig 2). Des pictogrammes tactiles peuvent donc être bien reconnus par des enfants et pourraient de ce fait être une solution d'illustration intéressante pour les livres tactiles.

Les taux de reconnaissance des pictogrammes en pointillés sont plus faibles que ceux des pictogrammes en lignes. Il semble donc que l'utilisation de pictogrammes en pointillés soit moins adaptée. Cependant, ce résultat n'est pas aussi tranché chez les enfants non-voyants. En effet, le taux de reconnaissance des pictogrammes en pointillés reste assez élevé (non-voyants : 78%, $IC_{95\%}$ [64 ;95]). Il semble même plus élevé que celui observé chez des enfants non-voyants avec des pictogrammes en pointillés à basse résolution (points de 4 mm espacés de 4 mm) dans l'étude de Mascle et al. [6] (55%, $IC_{95\%}$ = [46, 62]). Il semble donc que la résolution des pictogrammes ait un impact sur la reconnaissance.

5 Conclusion

Des pictogrammes tactiles et des pictogrammes en points peuvent être bien reconnus par des enfants non-voyants. Cependant, il semble que la résolution des pictogrammes

ait un impact sur la reconnaissance des pictogrammes en pointillés. La résolution de l'écran de la tablette à picots pourrait donc avoir un impact important sur la reconnaissance des illustrations affichées. Les tablettes à faible résolution comme le Blindpad [6] (picots de 4 mm espacés de 4mm) coutent entre 400 et 3000€. Cependant, les pictogrammes en pointillés basés sur cette tablette seraient probablement moins bien reconnus par les enfants que les pictogrammes basés sur des tablettes à plus haute résolution. L'utilisation de tablettes à « haute résolution » (picots de 1 mm espacés de 1mm) serait donc indispensable pour l'affichage d'illustrations tactiles. Cependant, le prix de ces dispositifs est actuellement dissuasif en raison des composants utilisés pour les produire. Par exemple, une tablette Hyperbraille© (HyperBraille, GE) sur une matrice de 76x48 picots est actuellement vendue plus de 10 000 €. Il est donc important de trouver un compromis entre la résolution proposée et le coût de fabrication et de commercialisation.

References

1. Neuman, SB.: Books make a difference: A study of access to literacy, *Read. Res. Q.*, vol. 34, no. 3, pp. 286–311 (1999)
2. Leo, F., Baccelliere, C., Waszkielewicz, A., Cocchi, E., & Brayda, L.: Tactile symbol discrimination on a small pin-array display. In Proceedings of the 2018 Workshop on Multimedia for Accessible Human Computer Interface, pp. 9–15, (2018)
3. Bellik, Y., & Clavel, C.: Geometrical shapes rendering on a dot-matrix display. In International Conference on Intelligent Human Computer Interaction, pp. 8–18 (2017).
4. Velazquez, R., Pissaloux, E. E., Hafez, M., & Szewczyk, J.: Tactile rendering with shape-memory-alloy pin-matrix. *IEEE Transactions on Instrumentation and Measurement*, 57(5), 1051-1057, (2008).
5. Mascle, C., Jouffrais, C., Kaminski, G., & Bara, F.: Vers l'Utilisation de Tablettes à Picots pour l'Illustration de Livres Tactiles. *Handicap*, (2020)
6. Zarate, J. J., Gudozhnik, O., Ruch, A. S., & Shea, H.: Keep in touch: portable haptic display with 192 high speed taxels. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 349–352, (2017).
7. Mazella, A., Albaret, J. M., & Picard, D.: Haptic-2D: A new haptic test battery assessing the tactual abilities of sighted and visually impaired children and adolescents with two-dimensional raised materials. *Research in developmental disabilities*, 48, 103-123, (2016).
8. Overvliet, K. E., & Krampe, R. T.: Haptic two-dimensional shape identification in children, adolescents, and young adults. *Journal of Experimental Child Psychology*, 166, 567–580 (2018)
9. Gallace, A., & Spence, C.: To what extent do Gestalt grouping principles influence tactile perception? *Psychological Bulletin*, 137(4), 538 (2011).
10. Hatwell, Y. : *Psychologie cognitive de la cécité précoce*. Dunod. (2003).

Représentations sociales de l’analytique des apprentissages avec le numérique

Camila Morais Canellas¹, François Bouchet¹ et Vanda Luengo¹

Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
{camila.canellas, francois.bouchet, vanda.luengo}@lip6.fr

Résumé. Cette étude vise à identifier les représentations sociales de l’analytique des apprentissages avec le numérique chez différents acteurs de l’éducation. Nous avons analysé les réponses de 286 participants à un questionnaire en utilisant la technique d’Évocation Hiérarchique et les résultats montrent que pour la majorité d’entre eux, les aspects les plus importants pour déterminer le domaine sont : apprentissage, données, analyse, suivi et traces. De plus, nous avons observé des variations entre différents sous-groupes de parties prenantes.

Mots-clé: analytique des apprentissages · représentations sociales · évocation · analyse de similarité · Test d’Indépendance de Contexte (TCI).

Abstract. This study aims to identify the social representations of the Learning Analytics domain among different education stakeholders. We analyzed the responses of 286 participants to a questionnaire using the Hierarchical Evocation technique. Results show that for the majority of them the most important aspects to determine the domain are: learning, data, analysis, tracking and traces. In addition, we observed variations between different subgroups of stakeholders.

Keywords: learning analytics · social representations · evocation · similarity analysis · Test of Context Independence (TCI).

1 Introduction et travaux connexes

L’analytique des apprentissages avec le numérique (traduction française de *Learning Analytics* (LA) d’après [4]) est un domaine scientifique récent lié aux contextes éducatifs dont la première conférence internationale scientifique, a eu lieu en 2011¹. Parfois, les connaissances produites dans un cadre scientifique entrent en conflit avec les croyances et habitudes antérieures de la société. Avec le temps et les nouveautés produites par le terrain, différents acteurs ou groupes sociaux forment une compréhension commune - une représentation -, de ce qu’est le domaine d’étude, de ce qu’il produit, de la façon dont il pourrait être utile ou néfaste, etc. Le processus d’internalisation (sociale) de nouvelles informations (scientifiques) a fait l’objet d’études dans de nombreux domaines par chercheurs intéressés par la théorie des Représentations Sociales (RS). Par conséquent, lorsque

¹ LAK 2011 : <https://tekri.athabascau.ca/analytics/>

l'on tente de mettre en œuvre les résultats des LA dans un cadre éducatif, il est important de comprendre à l'avance la nature de ces représentations. Nos questions de recherche sont : 1/ Quelle est la représentation sociale des LA parmi les parties prenantes ? 2/ Quelle est sa structure ? et 3/ Y a-t-il des différences entre les profils des différentes parties prenantes ?

En tant que nouveau domaine d'étude, au cours de ses premières années, la communauté de LA a produit des recherches sur le niveau de compréhension et les attentes envers le domaine, et un certain nombre de réflexions basées sur la bibliométrie ou des revues systématiques. En nous concentrant sur les articles qui s'appuient sur des mots-clés pour tenter d'identifier les limites du domaine, nous pouvons citer : (1) Balacheff et al. [3] qui ont analysé des mots-clés liés aux LA dans des textes produits pendant le *Stellar Grand Challenge* ; (2) Griffiths et al. [7] qui dans le cadre de *Learning Analytics Community Exchange (LACE)* ont identifié 13 thèmes qui ont émergé de leur analyse ; (3) Labarthe et al. [9] qui comparent les LA à des champs connexes, notamment en termes de mots-clés ; et (4) [12] où les auteurs ont utilisé des articles publiés entre 2011 et 2019 afin d'identifier la structure conceptuelle et les progrès thématiques dans les LA en utilisant notamment l'occurrence de mots.

Néanmoins, la définition formelle d'un champ, la pratique réelle de ses chercheurs et les représentations avérées des personnes impliquées d'une manière ou d'une autre sont trois choses potentiellement différentes à considérer.

2 Cadre d'étude et méthode

La théorie des RS se concentre sur l'individu situé dans son contexte historique et sociologique. Les RS sont des images mentales construites à partir d'expériences avec l'objet et de la pratique sociale et du langage [8]. De plus, ils peuvent avoir une dimension prédictive : en connaissant la RS d'un individu ou d'un groupe d'individus, on peut en déduire une tendance à réaliser certaines actions [6].

La théorie du noyau central. Nous représentons des objets sociaux en organisant des réseaux d'informations autour de connaissances ou de croyances, et c'est cette organisation qui est au cœur des préoccupations de la théorie complémentaire du noyau central [11]. Flament et Moliner [5] ont démontré l'existence de deux éléments dans les RS qui jouent des rôles différents : le noyau central et les éléments périphériques. Par conséquent, pour définir une RS, il est essentiel de connaître son contenu mais aussi sa structure.

Ainsi, cette étude est à la fois quantitative et exploratoire ; elle vise à identifier la RS des LA et sa structure parmi les personnes impliquées dans le domaine de l'éducation. Nous utilisons la technique d'Évocation Hiérarchique [2], basée sur des associations libres, et les résultats sont également examinés via une analyse de similarité. Lors de la collecte de données, un premier moment est dédié aux associations libres et un second à la hiérarchisation de ces associations, permettant d'accéder à une RS et à son hypothèse de centralité. L'interprétation repose sur deux prémisses : 1/ la *saillance* d'un élément, vue comme une fréquence élevée par rapport à la fréquence moyenne de tous les éléments évoqués, qui traduit

un fort consensus des participants et 2/ le *rang* de l'élément, qui dépend du fait que le mot ou l'expression soit classé comme parmi les plus importants. En combinant les deux indicateurs quantitatifs, les quatre cases décrites ci-dessous sont obtenues. L'analyse peut être effectuée à l'aide d'un logiciel [13].

La zone du *noyau central* est considérée comme liée à la mémoire collective et à l'histoire du groupe, elle est cohérente, rigide et stable dans le temps. Elle résiste au changement et n'est pas sensible au contexte immédiat, a tendance à avoir peu d'éléments, à être consensuelle et à définir l'homogénéité du groupe [1]. La *première périphérie* rassemble les éléments périphériques les plus importants. La case des *éléments contrastés* contient des éléments évoqués par peu de personnes, mais qui les considèrent très importants [2]. La *deuxième périphérie* est constituée d'éléments à basse fréquence et qui sont considérés comme peu importants dans le champ sémantique de la représentation.

Nous avons créé un questionnaire² diffusé en ligne au sein de groupes liés au domaine de l'éducation. Les réponses ont été recueillies entre mars et juillet 2019. Le questionnaire comprenait deux sections. La première consistait en deux questions permettant d'accéder à la RS de LA et à l'importance relative (classement) de chaque évocation. Les questions de la deuxième section permettaient d'établir des profils basés sur les rôles des participants et sur leur implication et expériences antérieures avec les LA. Avant l'analyse, nous avons préparé les données en suivant plusieurs étapes recommandées dans [13].

Afin de confirmer l'hypothèse de centralité découverte lors de la première phase [10] nous avons utilisé la méthode TCI. Un deuxième questionnaire (disponible sur le même lien : LAFR2_726112.pdf) a été envoyé aux mêmes listes et groupes que le premier, entre juillet et septembre 2020.

3 Résultats

Nous avons obtenu un total de $N = 286$ réponses valides, formant un corpus de 1843 mots ou expressions (598 uniques). La fréquence minimale (visant à réduire le corpus en enlevant des mots très peu cités) a été établie à 10 évocations, réduisant le corpus à 37 mots uniques (cf. Tableau 1). Le rang moyen est de 3, la médiane du nombre de mots demandés par le questionnaire. Compte tenu des caractéristiques de nos données et d'après [14], nous avons fixé la fréquence intermédiaire de coupure comme étant la moyenne des fréquences (29).

Le Tableau 1 affiche les évocations de toutes les parties prenantes participant à notre enquête. La fréquence (Fq.) est le nombre total de fois qu'un mot/une expression a été mentionné et le rang moyen (Rg.) la moyenne de sa position dans la classification faite par les participants du plus (1) au moins (5) important. Certains mots (e.g. "big-data") ont été évoqués en anglais et gardés ainsi.

La case du noyau central contient les éléments les plus fréquents et les plus importants. Il est composé de cinq mots, le plus mentionné étant "apprentissage" (86 fois) et le rang moyen le plus bas dans cette case est pour le mot "traces".

² LAFR_783534.pdf : <https://github.com/cmc-lip6/socialrepresentations>

Tableau 1: Analyse des évocations hiérarchiques de tous les participants

Freq.	Rang Moyen < 3		Rang Moyen \geq 3			
	mot	Fq.	Rg.	mot	Fq.	Rg.
≥ 29	apprentissage	86	2,4	évaluation	44	3,3
	données	64	2,7	statistiques	35	3,3
	analyse	62	2,5			
	suivi	49	2,6			
	traces	38	2,3			
[10, 29[apprenant	26	2,8	numérique	25	3,2
	big-data	24	3,0	enseignement	18	3,2
	personnalisation	22	2,9	adaptation	17	3,5
	parcours	19	2,1	intelligence-artificielle	16	3,7
	recherche	17	2,5	compétences	13	3,0
	formation	15	2,5	amélioration	11	3,2
	activité	15	2,6	contrôle	11	3,4
	compréhension	14	2,8	informatique	10	3,3
	indicateurs	12	2,4	tableau-de-bord	10	3,5
	mémoire	12	2,4	visualisation	10	3,9
	pédagogie	12	2,4			
	résultats	12	2,5			
	individualisation	11	2,2			
	remédiation	11	2,4			
	étudiant	11	2,4			
	connaissance	11	2,8			
	éthique	10	2,3			
	data	10	2,6			
	élèves	10	2,6			
	feedback	10	2,7			

Le mot "données", évoqué 64 fois, est présent dans le noyau même s'il n'est pas agrégé avec le même mot mentionné à l'origine en anglais (10 fois) et en "big-data" (24 fois). Les éléments de la deuxième périphérie ont encore une importance pour la représentation. Seuls deux mots sont présents, "évaluation" et "statistiques". La case des éléments contrastés est composée de 20 éléments, dont la fréquence varie de 26 à 10 et avec un rang moyen de 2,1 à 2,9.

Nous avons également effectué une analyse de similarité, basée sur la théorie des graphes, afin de mieux comprendre visuellement comment les mots sont évoqués ensemble (voir Figure 1). Cette analyse a été effectuée en utilisant Iramuteq³ avec un seuil d'au moins 12 occurrences pour une meilleure visibilité. Nous avons choisi l'index "arbre maximum" où la taille des nœuds représente la fréquence des mots et l'épaisseur des lignes représente les associations entre eux.

Ensuite, nous avons essayé d'identifier des profils de participants pour vérifier si des différences étaient observables entre des sous-groupes, selon le rôle, la discipline, l'implication et l'expérience avec les LA.

L'analyse des ces sous-groupes nous a permis d'observer que la zone de la première périphérie est souvent peu peuplée (voire vide), ce qui pourrait être le signe d'un fort consensus sur le noyau et, en même temps, d'une opinion assez dispersée des autres éléments formant les RS des LA. De plus, deux sous-groupes, "enseignants" et "sans expérience", semblent avoir des représentations assez différentes par rapport aux autres sous-groupes ainsi que par rapport à la représentation principale du total des participants.

³ <http://www.iramuteq.org/>

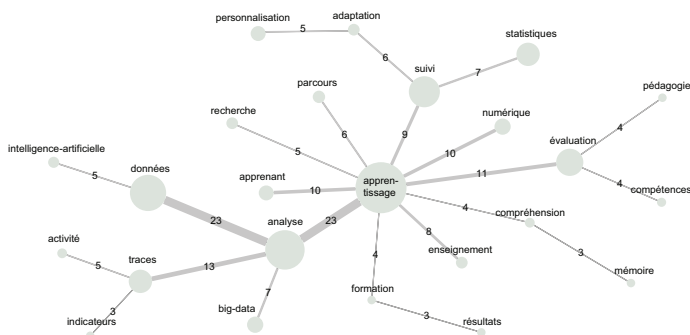


Fig. 1: Analyse de similarité des termes co-occurents

Le deuxième questionnaire, visant à confirmer ces résultats, a obtenu 61 réponses complètes ($N = 61$). Après avoir calculé les fréquences "oui", nous avons utilisé le test de Kolmogorov-Smirnov pour confirmer que la distribution était normalement distribuée ($p = 0,764$). Nous avons ensuite utilisé un test du Chi-2 pour identifier les évocations statistiquement significatives ($p < .05$) formant potentiellement le noyau central - où l'hypothèse nulle est que l'évocation testée est indépendante du contexte. Les évocations ainsi identifiées sont : "**apprentissage**", "statistiques", "indicateurs", "activité", "**data**", "**analyse**", "**données** et "**traces**". Les résultats montrent que toutes les évocations apparaissant précédemment dans le noyau y figurent à nouveau, sauf "suivi" (situé juste sous le seuil de significativité), avec quelques nouveaux ajouts (mots non en gras ci-dessus).

4 Discussion et limitations

Même s'il n'y a pas de consensus sur la taille optimale des échantillons pour étudier les RS, plus le groupe de participants est grand, plus les résultats sont stables. La taille de notre échantillon est plus grande ($N = 286$) que ce qui est parfois recommandé [14], mais ce n'est pas le cas de certains sous-groupes que nous avons analysés pour lesquels les résultats doivent être considérés avec prudence et nous envisageons donc de reconduire l'enquête auprès d'autres acteurs, notamment parmi des apprenants.

Dans cette étude, nous avons mis en lumière les représentations sociales de différents acteurs de l'éducation au sujet des LA. Nous pensons que la connaissance des RS des LA pourra aider les décideurs dans leurs résolutions et favoriser le dialogue entre les différentes parties prenantes pouvant être impactées par le domaine. Les variations entre les différentes parties prenantes, à confirmer,

pourraient potentiellement signifier qu'il est difficile de comprendre ce qu'est le domaine sans en faire l'expérience.

Références

1. Abric, J.C.: L'organisation interne des représentations sociales : système central et système périphérique. In: Guimelli, C. (ed.) Structures et transformations des représentations sociales, chap. 2, pp. 73–84. Delachaux et Niestlé, Paris (1994)
2. Abric, J.C.: La recherche du noyau central et de la zone muette des représentations sociales. In: Abric, J.C. (ed.) Méthodes d'étude des représentations sociales, chap. 3, pp. 59–80. Erès, Saint-Agne (2003)
3. Balacheff, N., Lund, K.: Multidisciplinarity vs. multivocality, the case of "learning analytics". In: LAK '13: Proceedings of the Third International Conference on Learning Analytics and Knowledge. pp. 5–13. No. April, Association for Computing Machinery, Leuven, Belgium (2013)
4. Cherigny, F., Kechai, H.E., Iksal, S., Lefevre, M., Labarthe, H., Luengo, V.: L'analytique des apprentissages avec le numérique - Groupes thématiques de la Direction du numérique pour l'Éducation (DNE-TN2). Tech. rep. (2020)
5. Flament, C.: Structure et dynamique des représentations sociales. In: Jodelet, D. (ed.) Les représentations sociales, chap. 9, pp. 224–239. Presses Universitaires de France, Paris, 7 edn. (1989)
6. Franco, M.L.P.B.: Representações Sociais, Ideologia e Desenvolvimento da Consciência. *Cadernos de Pesquisa* **34**(121), 169–186 (2004)
7. Griffiths, D.: Visions of the Future, Horizon Report. Tech. rep., Learning Analytics Community Exchange (2015)
8. Jodelet, D.: Représentations sociales : phénomènes, concept et théorie. In: Psychologie sociale, chap. 13, pp. 361–382. Presses Universitaires de France, Paris, 7 edn. (1998)
9. Labarthe, H., Luengo, V., Bouchet, F.: Analyzing the relationships between learning analytics, educational data mining and AI for education. In: S. Béland, M. C. Desmarais, N.L. (ed.) Proceedings of Learning Analytics Workshop at ITS. pp. 10–19. [Electronic medium], Montréal (2018)
10. Lo Monaco, G., Piermattéo, A., Rateau, P., Tavani, J.L.: Methods for Studying the Structure of Social Representations: A Critical Review and Agenda for Future Research. *Journal for the Theory of Social Behaviour* **47**(3), 306–331 (2017)
11. Moliner, P., Abric, J.C.: Central core theory. In: Sammut, G., Andreouli, E., Gaskell, G., Valsiner, J. (eds.) The Cambridge Handbook of Social Representations, chap. 6, pp. 83–95. Cambridge University Press, Cambridge (2015)
12. Papamitsiou, Z., Giannakos, M.N., Ochoa, X.: From Childhood to Maturity: Are We There yet? Mapping the Intellectual Progress in Learning Analytics during the Past Decade. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, LAK'20. pp. 559–568. Association for Computing Machinery, Frankfurt, Germany (2020)
13. Scano, S., Junique, C., Pierre Vergès: Ensemble de programmes permettant l'analyse des évocations. EVOC2000. (2002)
14. Wachelke, J., Wolter, R.: Critérios de construção e relato da análise prototípica para representações sociais. *Psicologia: Teoria e Pesquisa* **27**(4), 521–526 (12 2011)

Faciliter le recueil de traces par la numérisation d'un outil tangible de co-design : application à la conception de tableaux de bord d'apprentissage

Katia Oliver-Queleennec^{1,2,3}, François Bouchet¹, Thibault Carron¹, and Claire Pinçon²

¹ Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
{francois.bouchet, thibault.carron}@lip6.fr

² Univ. Lille, CHU Lille, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

³ Univ. Lille, GIVRE, DIP, France
{katia.queleennec, claire.pincon}@univ-lille.fr

Résumé Le co-design est une approche de recueil de besoins qui s'appuie souvent sur des outils sous forme de cartes, mais ce format tangible, limite le recueil des données et leur exploitation. Cet article présente les facteurs clés et points de vigilance identifiés dans le cadre d'une adaptation d'une méthode de co-design basée sur des cartes en une version numérique, à travers l'exemple de la méthode PADDLE (Participative Design of Dashboard for Learning in Education) pour concevoir des tableaux de bord d'apprentissage (TBA). Cette adaptation numérique et l'outil associé, appelé ePADDLE, ont été testés auprès d'étudiants de première année d'université (18 groupes, $N = 58$). Les participants ont évalué ePADDLE comme un peu moins adapté que l'original mais tout en restant positifs. Enfin, chaque groupe a réussi à concevoir un TBA, démontrant la validité du protocole expérimental d'ePADDLE.

Mots-clés : Tableaux de bord d'apprentissage, co-design, design participatif, cartes.

Abstract Card-based tools are a common co-design method to collect users' needs, but this tangible format limits data collection and usage. This article presents key factors and points of attention identified in adapting a card-based co-design method into a digital version, using the example of the PADDLE (Participatory Design of Dashboard for Learning in Education) method for designing learning dashboards. This digital adaptation and the associated tool, ePADDLE, were tested with first year university students (18 groups, $N = 58$). All groups have successfully designed a dashboard and users have evaluated ePADDLE as being almost as suitable as the original method. Moreover, each group has managed to design a learning dashboard, indicating the validity of the ePADDLE method.

Keywords: Learning analytic dashboard, co-design, participatory design, cards.

1 Introduction

Les méthodes de design basées sur la conception centrée utilisateur, et plus particulièrement le design participatif, permettent de créer des outils qui répondent aux besoins réels des utilisateurs [9]. Les sessions de co-design impliquent généralement des ateliers avec des participants manipulant des outils tangibles (cartes, plateau...). Bien qu'efficaces en termes de recueil de besoins, ces méthodes ne permettent pas forcément une récupération facile des traces des interactions, du processus de conception et des résultats. Cet article présente une démarche à vocation générique de transposition d'une méthode de conception participative basée sur des cartes, en une version numérique aidant à la récupération des traces (données d'observation de l'utilisation de la méthode pour faciliter son évolution, données intermédiaires pour pouvoir comprendre les conditions d'une collaboration efficace et résultats des productions). Le format numérique permet aussi de réaliser ces séances selon d'autres modalités, notamment à distance. Cette démarche est ensuite appliquée à une méthode de co-design nommée PADDLE (PArTicipative Design of Dashboard for Learning in Education) [6]. Nous présentons d'abord les travaux précédents concernant le design participatif dont PADDLE, puis nous abordons sa déclinaison numérique ePADDLE avant de décrire les ateliers réalisés et les données collectées. Ensuite, nous exposons nos premiers résultats avant de conclure par une discussion des gains et pertes liés à cette transposition numérique et en présentant de futures pistes de travaux.

2 Travaux précédents

2.1 Design participatif

Une revue de littérature de 155 outils de conception basés sur des cartes [8] montre que ce format est largement utilisé : leur forme tangible permet de combiner plusieurs idées, avec une quantité limitée d'information, proposant ainsi un bon intermédiaire entre les outils sans structure (post-it) et les outils complets (manuel d'instruction). Cependant, cette revue souligne aussi que seuls "certains des outils sont [...] également disponibles sous forme d'applications ou en ligne". Par ailleurs, l'existence d'une version numérique ne garantit pas que des traces soient enregistrées, la plupart des méthodes s'intéressant plus à la production finale qu'à une analyse des processus mis en œuvre pour les obtenir. Dans le domaine de l'analyse de l'apprentissage, certaines approches basées sur des cartes comme LA-DECK [7] ont montré leur efficacité. Ces ateliers sont souvent accompagnés d'un facilitateur qui introduit la séance, présente les cartes, répond aux éventuelles questions et conclut la séance [7]. Toutes les méthodes susmentionnées ont été utilisées uniquement lors d'ateliers en présentiel avec des cartes tangibles, sans aucun outil numérique. Certains travaux ont démontré que l'utilisation d'objets tangibles peut faciliter la collaboration [10] et pourraient expliquer l'utilisation récurrente de ce format pour le co-design. Les travaux de recherche de Jeong et Hmelo-Silver [4] sur la collaboration ont démontré que l'utilisation du numérique permet d'engager les participants dans des processus

productifs et dans un cadre de co-construction. Si l'outil peut contraindre les utilisateurs pour les guider et les aider à rester concentrés sur leur tâche principale, il peut aussi limiter la créativité s'il est trop directif : il faut donc trouver le juste milieu dans sa conception.

2.2 Méthode PADDLE

Pour concevoir des tableaux de bord d'apprentissage (TBA), Dabbebi et al. [2] ont développé un outil de conception participative utilisant également un ensemble de cartes qui a été évalué positivement par ses utilisateurs. Nous avons déjà proposé une adaptation de cette méthode pour l'enseignement supérieur [6] en modifiant notamment le vocabulaire, la durée pour la cible étudiante et l'ajout de questions pour les participants. Nous nous sommes appuyés sur cette adaptation, intitulée PADDLE, pour la transposer en format numérique et améliorer le recueil des traces. PADDLE permet de co-concevoir des TBA en petits groupes de deux à cinq personnes. Les ateliers durent 90 à 150 minutes selon les profils (étudiant ou enseignant) et comportent cinq phases : introduction, choix de l'objectif du TBA, définition du contexte, désignation des données et de leurs visualisations et enfin, élaboration du TBA. L'atelier est encadré par un animateur qui facilite les débats et aide les participants à formaliser leurs idées tout en leur laissant une certaine autonomie. PADDLE se présente sous forme de cartes plastifiées (pour pouvoir les réutiliser) et magnétiques (pour les manipuler sur un tableau blanc). Le recueil des traces des ateliers se fait à l'aide d'enregistrements audio des groupes, de photos de résultats d'étapes intermédiaires, de photos des TBA produits et d'un questionnaire d'évaluation papier ou en ligne.

3 Co-concevoir avec ePADDLE

3.1 Méthode ePADDLE

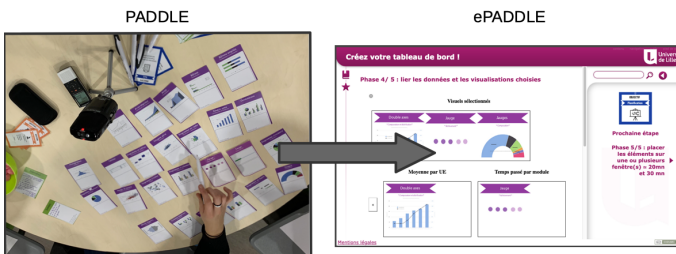


FIGURE 1. Illustration de PADDLE et ePADDLE.

Nous avons identifié 8 propriétés dans la méthode PADDLE (Table 1) et réfléchi à leurs possibles transpositions numériques pour améliorer le recueil des traces sans perdre leur fonctionnalité première :

- Explications initiales : l'utilisation d'une vidéo facilite le recueil de traces (temps passé à écouter la vidéo, nombre de visionnages), mais l'absence d'un humain pour lancer l'activité peut être préjudiciable à la motivation du groupe.
- Collaboration entre les participants : l'ajout de puces RFID permet d'identifier les cartes manipulées [3] mais limite la possibilité de sessions à distance. Proposer des champs individuels à remplir selon le nombre de participants oblige tous les membres du groupe à participer, renforçant l'engagement dans le processus de production et permettant un recueil de traces par participant. Mais cette obligation doit pouvoir être ponctuellement levée pour éviter des productions "forcées" générant du bruit dans les traces.
- Scénarisation de la collaboration : scénariser le dispositif numérique permet de reproduire l'ordre des cartes proposé initialement et bloque la possibilité de mélanger les cartes par inadvertance. Le format numérique permet de recueillir des informations sur les cartes les plus discutées.
- Réponses aux questions des participants : un agent conversationnel pour traiter les questions simples permet de tracer les problèmes récurrents et standardiser les réponses, mais risque de trop guider les participants dans la méthode.
- Régulation de la séance : forcer le passage d'une phase à l'autre est envisageable mais risque de contraindre des échanges riches, et il est recommandé que les participants soient acteurs dans la régulation de leurs activités de collaboration [4]. Le format numérique permet de tracer le temps passé par phase.
- Visibilité des décisions antérieures : afficher un récapitulatif de chaque phase permet au groupe de vérifier ses choix et de changer d'avis. On peut tracer les accès à cet écran pour estimer la coordination des décisions avec des choix passés.
- Production graphique : utiliser des outils numériques permet d'avoir des résultats potentiellement plus lisibles qu'une production manuscrite. Dans le cas d'utilisation de logiciels en ligne, le lien des productions est récupéré.
- Évaluation de la méthode et des résultats obtenus : la méthode d'évaluation initialement développée par questionnaire est adaptée au format numérique.

La Table 1 présente les transpositions retenues dans ePADDLE pour les 8 propriétés de la méthode PADDLE, dont le résultat est illustré par la Figure 1.

3.2 Évaluation des transpositions des propriétés

Chaque implémentation de propriété de PADDLE a été soit conservée, soit transposée par une implémentation permettant un meilleur traçage. Pour vérifier que ces changements n'ont pas d'impact négatif, nous avons organisé des séances de co-conception avec ePADDLE et comparé avec des séances PADDLE [6]. Ainsi, 4 ateliers ont été organisés à distance, impliquant au total $N = 58$ étudiants universitaires de première année, répartis au hasard dans 18 groupes de 2 à 4 étudiants. Les sessions ePADDLE ont permis de concevoir 18 TBA et de recueillir facilement les données saisies pendant l'atelier, en plus des traces recueillies habituellement. En réutilisant le questionnaire d'évaluation d'origine[2] (cf. Table

TABLE 1. Transposition numérique des propriétés de PADDLE dans ePADDLE

Propriété	PADDLE	ePADDLE	Traces supplémentaires
Explications initiales	Diaporama présenté par l'animateur	Diaporama présenté par l'animateur	-
Collaboration entre participants	Choix de cartes et annotations des cartes par les participants	Participant avec rôle de scribe et champs à remplir soit individuellement, soit collectivement	Réponses individuelles et collectives des participants
Scénarisation de la collaboration	Cartes ordonnées et interventions régulières de l'animateur	Interface en ligne pour une collaboration scénarisée	Temps passé par phase
Réponses aux questions	Animateur séance	Animateur par groupe ou animateur pour la séance	Trace des questions dans le chat
Régulation temporelle	Animateur séance	Participant avec rôle de maître du temps et suivi de l'animateur séance	-
Visibilité des décisions antérieures	Cartes magnétiques sur le tableau blanc et annotation des participants	Récapitulatif des choix faits entre chaque phase et via le menu du dispositif numérique	Récapitulatif des choix des participants et nombre de consultations du récapitulatif
Production graphique du résultat	Matériel de papeterie (papier, crayon, scotch, ...)	Au choix : logiciel maîtrisé par l'un des participants du groupe avec partage d'écran, logiciel d'édition en ligne collaboratif ou papier en cas de difficulté technique	Lien si production en ligne
Évaluation	Questionnaire	Questionnaire avec lien dans le dispositif numérique	-

2), ePADDLE a été jugé un peu moins adapté que PADDLE, principalement en ce qui concerne l'aide à la créativité où l'outil ne pouvait pas rivaliser avec une conception en présentiel sur papier. Le suivi des questions a été géré avec succès avec un seul animateur par séance. La régulation peut être considérée comme adaptée car la durée des ateliers est restée majoritairement dans les temps prévus (de 63 à 115 min, $M = 92$ min, $SD = 18$ min). Enfin, la visibilité des décisions antérieures et la production graphique peuvent être validées par les productions réalisées, qui sont en concordance avec les choix des participants. Mais ces résultats sont à confirmer avec des échantillons plus importants.

TABLE 2. Évaluation des transpositions des propriétés de PADDLE dans ePADDLE *degré de significativité du test de Mann-Whitney après correction de Bonferroni

Propriété	Opinion selon une échelle de Likert à 5 points	PADDLE ($N = 8$)	ePADDLE ($N = 58$)	p*
Collaboration	L'outil vous a aidé à avoir une bonne dynamique de groupe.	M=4.57, SD=0.79	M=4.14, SD=0.98	0.14
	L'outil vous a aidé à converger vers une solution.	M=4.29, SD=0.76	M=4.14, SD=0.87	0.27
Scénarisation	L'outil est facile de prise en main.	M=4.63, SD=0.74	M=4.17, SD=0.75	0.15
	Les cartes sont facilement compréhensibles.	M=4.25, SD=0.89	M=3.48, SD=0.90	0.09
	L'outil vous a permis de mieux préciser votre besoin.	M=4.63, SD=0.52	M=3.64, SD=1	0.01
	L'outil vous a permis de trouver des solutions originales.	M=4.75, SD=0.46	M=3.69, SD=0.98	0.005
	Et cette solution vous semble pertinente.	M=4.71, SD=0.49	M=3.97, SD=0.90	0.01

4 Conclusion et perspectives

Si l'avantage principal de cette transposition numérique est d'obtenir des données pré-formatées permettant un traitement rapide de celles-ci, nous pouvons observer d'autres bénéfices, comme la possibilité de l'utiliser à distance. Il reste cependant des points faibles sur lesquels cette transposition peut évoluer, comme outiller la production graphique pour faciliter la collecte des traces encore pauvre pour cette propriété, tout en facilitant cette activité pour les participants. Enfin, cette transposition nous a apporté des réflexions qui peuvent enrichir la méthode initiale, comme la définition de rôles. Ce premier travail devrait être approfondi avec d'autres transpositions des mêmes propriétés dans d'autres contextes afin de mieux déterminer les règles à suivre pour obtenir des transpositions adaptées. Les prochaines étapes de notre travail porteront sur les dynamiques de groupe, l'identification des conditions favorables à la collaboration avec une analyse des actes de langage basée sur les travaux de McNair et al. [5] et en essayant de trouver des liens avec le profil de collaboration de Belbin [1].

Remerciements Ce travail s'appuie sur le projet P3, développé par l'Université de Lille et co-financé par l'iSite Université Lille Nord-Europe.

Références

1. Belbin, R.M. : Management Teams. Taylor & Francis (2010)
2. Dabbebi, I., Gilliot, J.M., Iksal, S. : User Centered Approach for Learning analytics Dashboard Generation. In : CSEDU 2019. pp. 260–267. Heraklion, Greece (2019)
3. Jacob, R.J.K., Ishii, H., Pangaro, G., Patten, J. : A tangible interface for organizing information using a grid. In : Proc. CHI '02 (2002)
4. Jeong, H., Hmelo-Silver, C.E. : 7 affordances of computer-supported collaborative learning : How to support collaborative learning? how can technologies help? (2016)
5. McNair, L.D., Paretto, M.C., Davitt, M. : Towards a Pedagogy of Relational Space and Trust : Analyzing Distributed Collaboration Using Discourse and Speech Act Analysis. IEEE Transactions on Professional Communication **53**, 233–248 (2010)
6. Oliver-Queleennec, K. : Vers une conception participative de tableaux de bord étudiants adaptatifs avec PADDLE. In : Actes des 8èmes RJC-EIAH 2020 (2020)
7. Prieto Alvarez, C.G., Martinez-Maldonado, R., Buckingham Shum, S. : LA-DECK : A Card-Based Learning Analytics Co-Design Tool. In : Proceedings of the 10th International Conference on Learning Analytics and Knowledge. Frankfurt (2020)
8. Roy, R., Warren, J.P. : Card-based design tools : a review and analysis of 155 card decks for designers and designing. Design Studies **63**, 125–154 (Jul 2019)
9. Sanders, E. : From user-centered to participatory design approaches. In : Design and the Social Sciences : Making Connections., pp. 1–7. CRC Press (2002)
10. Schneider, B., Jermann, P., Zufferey, G., Dillenbourg, P. : Benefits of a Tangible Interface for Collaborative Learning and Interaction. IEEE Transactions on Learning Technologies (2011)

Faciliter l'intégration des critères RGPD et éthique lors de la co-conception des jeux épistémiques numériques

Estelle Prior¹, Mariem Jaouadi¹, Nadine Mandran², Eric Sanchez¹

¹ Université de Genève, TECFA, Genève, Suisse
[estelle.prior, mariem.jaouadi, eric.sanchez]@unige.ch
² Université de Grenoble, CNRS, LIG, Saint Martin d'Hères, France
nadine.mandran@univ-grenoble-alpes.fr

Résumé. Cet article présente une solution pour accompagner les concepteurs de jeux épistémiques numériques travaillant en contexte de recherche afin de respecter et intégrer les principes du règlement général de protection des données (RGPD) et les valeurs issues de l'approche éthique du *Value-Sensitive Design*. Ce travail est conduit dans le cadre du projet co.LAB.

Mots-clés : Règlement Général sur la Protection des Données, éthique, *Value-Sensitive Design*, jeux épistémiques numériques, projet co.LAB.

Abstract. This article presents a solution to support the designers of digital epistemic games working in a research context in order to respect and integrate the principles of the General Data Protection Regulation (GDPR) and the values stemming from the ethical approach of Value-Sensitive Design. This work is conducted in the framework of the co.LAB project.

Keywords: General Data Protection Regulation, ethics, Value-sensitive Design, digital epistemic games, co.LAB project.

1 Introduction

Le Règlement Général sur la Protection des Données (RGPD), “s’applique à toute organisation, publique et privée, qui traite des données personnelles pour son compte ou non” [1]. La communauté des Environnements Informatiques pour l’Apprentissage Humain (EIAH), dont les travaux de recherche comprennent la conception d’artefacts informatiques et le recueil de traces d’usage de ces derniers, est particulièrement concernée par cette réglementation. Pourtant, peu de travaux abordent cette question [2]. De plus, il est nécessaire de prendre en compte des questions éthiques pour inscrire la recherche dans “une réflexion morale des chercheurs sur leur activité” [3], partiellement prise en compte dans certains projets de recherche (e.g. projet HUBBLE). Les acteurs et institutions du monde éducatif sont, de fait, concernés par ces questions et soumis à

un degré d'exigence élevé dans le respect des obligations RGPD, avec les mesures techniques et organisationnelles qui en découlent.

L'accompagnement des équipes de recherche à la prise en compte de ces questions est un des volets du projet co.LAB, *A Digital Lab for the co-Design, co-Development and co-Evaluation of Digital Learning Games*. Ce projet vise à favoriser l'articulation entre pratique et recherche dans le cadre d'une recherche orientée par la conception [4]. Il porte sur le développement d'un processus et d'une plateforme permettant la co-conception, le co-développement et la co-évaluation de jeux épistémiques numériques (JEN) dans un contexte de recherche. Sur la plateforme, chaque étape du processus est associée à une "carte", proposant un espace de travail collaboratif, et des ressources (e.g. guides, références, exemples).

Le RGPD et l'éthique sont complexes à mettre en œuvre dans les étapes de la conception de JEN, et ce sont deux questions qu'il est nécessaire de prendre en compte au plus tôt pour anticiper au mieux les impacts du JEN et de l'analyse des données collectées sur les utilisateurs. Ainsi, nous nous intéressons à la manière de faciliter la rédaction des documents RGPD et la prise en compte des critères éthiques par l'équipe de conception. Dans ce qui suit, nous présentons d'abord une synthèse des éléments du RGPD et des modèles intégrant l'éthique dans la conception puis nous proposons une solution pour la plateforme co.LAB.

2 Etat de l'art

2.1 Réglementation pour la protection des données

Une recherche dans la revue Sticef de 2010 à 2020 avec les termes "RGPD" et "protection des données" montre qu'aucun article ne mentionne le RGPD et que seuls deux articles abordent la protection des données sans pour autant la mentionner en tant que telle. En effet, elle est abordée sous le terme de "respect de la vie privée" et de "données de navigation et d'inscription" dans le cadre d'analyse de traces [5] ou de "protection des libertés individuelles et publiques" et de "confidentialité des données" [6].

Les données à caractère personnel renvoient à "toute information se rapportant à une personne physique identifiée ou identifiable". Elle "peut être identifiée, directement ou indirectement, notamment par référence à un identifiant" (e.g. identifiant en ligne) "ou à un ou plusieurs éléments spécifiques propres à son identité" (e.g. identité physique) [7]. En Europe, le traitement de ces données est encadré par le RGPD. Il renforce le contrôle, par les citoyens, de l'utilisation qui peut en être faite [1]. Le RGPD protège "les libertés et droits fondamentaux des personnes physiques, et en particulier leur droit à la protection des données à caractère personnel" (art. 1). Le RGPD (art. 25, §1) consacre le principe de la protection des données dès la conception par le *Privacy by Design (PbD)* [8]. En effet, la législation impose que le responsable de la collecte et du traitement des données prenne des mesures techniques et organisationnelles pour mettre en œuvre des garanties nécessaires pour protéger les droits et libertés des personnes. La transparence, la clarté, le respect, le contrôle des utilisateurs, le consentement, l'accès et la responsabilité sont principalement mis en avant dans la littérature [9, 10]. Ces

concepts sont repris dans l'article 5 du RGPD [7] sous la forme de sept principes relatifs au traitement des données à caractère personnel. 1) *Egalité, équité et transparence* : La collecte de données ne doit pas violer la loi et ne cache rien aux personnes, 2) *Limitation des finalités* : Collecter des données à des fins spécifiées, explicites et légitimes, 3) *Minimisation des données* : Traiter les données personnelles nécessaire pour atteindre l'objectif, 4) *Exactitude/qualité* : S'assurer que les données sont exactes et les tenir à jour, 5) *Limitation de la conservation* : Supprimer les données lorsqu'elles ne sont plus nécessaires au traitement, 6) *Intégrité de la conservation* : Prendre des mesures de sécurité sur l'accès aux données, leur stockage et leur archivage, 7) *Responsabilité/accountability* : Démontrer la conformité : tenir les registres de traitement et mener les études d'impact sur la vie privée, les analyses de licéité et d'intérêt légitime.

2.2 Ethique intégrée au processus de conception

Une recherche dans la revue *Sticef* de 2010 à 2020 avec le terme "éthique", montre qu'il est employé dans certains articles dans les réflexions à avoir lors de l'utilisation d'outils numériques proposés pour l'apprentissage [5, 11], les réflexions dans le processus de recherche [5], ou encore dans les études soumises et validées par le comité d'éthique de l'institution [12]. Cependant, ces articles ne présentent pas de définitions univoques du terme éthique et ne proposent pas d'accompagnement pour l'intégration de ces réflexions dans le processus de conception. Néanmoins, Chan [13] précise que "l'éthique appliquée à la conception permet de traiter les questions ou conflits éthiques".

L'éthique par la conception ou *Ethics by Design*, comprend les méthodes et outils qui permettent de raisonner sur les aspects éthiques de la conception [14]. Pour prendre en compte l'éthique par la conception, différentes approches offrent des lignes directrices, comme le *Value-Sensitive Design (VSD)* [14], le *Value-based Engineering (VbE)* ou encore le *situated Cognitive Engineering (sCE) with values*. Ces approches intègrent la notion de valeur qui est à comprendre au sens de Friedman *et al.* [15] comme se référant à ce qu'une personne ou un groupe considère important. La technologie pouvant avoir un impact sur ces valeurs, elles doivent être identifiées par les concepteurs. Ainsi, l'approche du *VSD*, propose une théorie, des méthodes et des applications à toutes les étapes du processus de conception impliquant des valeurs [15]. L'étape 1 (*conceptual investigation*) permet de clarifier les questions fondamentales soulevées par le projet (*e.g.* acteurs et valeurs impactés). L'étape 2 (*empirical investigation*), est une étude empirique du contexte d'usage. L'étape 3 (*technical investigation*), se centre sur la façon dont les propriétés et les mécanismes liés soutiennent ou entravent les valeurs identifiées. L'ensemble constitue un processus itératif [15]. Critiquée pour son manque de clarté et la difficulté à être opérationnalisée, le *VSD* comporte des outils d'accompagnement [16]. L'approche a été améliorée à travers le *VbE* et le *sCE with values*. Le *VbE* s'applique plutôt aux entreprises en s'appuyant sur des "value experts" [17]. Le *sCE with values* intègre les valeurs du *VSD* et des méthodes de conception visant à améliorer l'interaction homme-machine [18]. Ces deux approches présentent des contextes d'application différents du notre c'est pourquoi nous choisissons l'approche du *VSD*.

3 Contribution scientifique

Notre contribution consiste à accompagner la mise en œuvre des critères de protection des données et des critères éthiques pendant la conception de JEN. Pour cela, nous spécifions les cartes du projet co.LAB par deux onglets "RGPD" et "Éthique".

3.1 Onglet "RGPD" dans les cartes du projet co.LAB

Nous proposons de formaliser l'onglet "RGPD" en s'appuyant sur le Modèle de registre simplifié RGPD proposé par la CNIL [19], pour renseigner les rubriques habituelles suivantes : la responsabilité, les objectifs et la finalité, les personnes concernées, les catégories des données traitées et des modes de collecte, les lieux et durée de conservation, les catégories de destinataires des données, et les mesures de sécurité. Pour nourrir les réflexions sur les sept principes de protection des données issus de l'article 5 du RGPD [7], nous proposons un ensemble de questions qui seront discutées entre l'équipe de conception du JEN. Ces éléments contribueront à la rédaction du document de dépôt RGPD. (1) *Egalité, équité et transparence* : En quoi la pratique de collecte de données peut impacter et nuire à la vie des personnes ? Comment les personnes sont informées des données récoltées ? (2) *Limitation de la finalité* : Pour quelles raisons les données sont-elles collectées ? Dans quelles mesures sont-elles récoltées à des fins légitimes ? (3) *Minimisation des données* : À partir de la liste des données récoltées, pouvez-vous indiquer celles nécessaires et suffisantes pour atteindre votre objectif ? (4) *Exactitude* : Quelles sont les procédures mises en œuvre pour assurer de la qualité des données ? Comment assurez-vous la mise à jour et la cohérence des données ? (5) *Limitation de la conservation* : Combien de temps est-il nécessaire de conserver ces données ? Quelles sont les procédures que vous avez prévues pour supprimer les données après utilisation ? (6) *Intégrité de la conservation* : Comment les données sont-elles protégées ? Dans quel lieu les données sont-elles conservées ? Quels acteurs ont accès à ces données ? (7) *Responsabilité/accountability* : Avez-vous mis en place un registre de traitement ? Quelles études d'impact avez-vous choisies de réaliser ? Quelles bases légales avez-vous choisi de suivre pour traiter les données collectées ? Avez-vous réalisé une analyse de licéité et une analyse d'intérêt légitime ? Malgré ces questions, les notions de responsabilité, de licéité, ou encore d'intérêt légitime demandent une expertise juridique qui passe par l'intégration d'un tel expert, et ce, dès le début du projet.

3.2 Onglet "Éthique" dans les cartes du projet co.LAB

L'onglet "Éthique" facilite la prise en compte du *VSD*. Nous intégrons ses trois étapes et les formalisons comme ceci : (1) *Liste des valeurs concernées* : Il s'agit d'identifier toutes les valeurs concernées par l'artefact conçu. Les concepteurs peuvent s'appuyer sur les valeurs proposées dans le *VSD* ou en proposer de nouvelles. Ils peuvent aussi s'appuyer sur l'outil *Envisioning Cards* qui aide à les identifier [16]. (2) *Définition des valeurs* : Il s'agit de définir les valeurs identifiées pour leur donner du sens. (3) *Niveau de l'impact de la technologie sur les utilisateurs* : Il s'agit de préciser si la valeur de

l'utilisateur est soutenue (impact positif) ou mise en conflit (impact négatif) par l'artefact conçu. (4) *Acteurs concernés* : Il s'agit d'indiquer les acteurs dont les valeurs sont impactées. (5) *Explicitation de l'impact ou du non impact* : Il s'agit d'énoncer comment l'artefact est responsable de l'impact. (6) *Conséquences sur les utilisateurs* : Il s'agit d'expliquer la nature de l'impact engendré. (7) *Proposition pour réduire le conflit sur les valeurs* : Il s'agit de mentionner des changements au niveau de la conception pour réduire l'impact négatif de l'artefact. (8) *Proposition alternative pour soutenir davantage les valeurs* : Il s'agit de présenter d'autres options de conception pour réduire le conflit sur les valeurs identifiées. Ces rubriques seront adaptées en fonction des étapes du processus de conception.

Nous développons un exemple pour la formulation des objectifs pédagogiques. Ces derniers précisent "les intentions générales et les buts poursuivis" [20] par le JEN. (1) *Liste des valeurs concernées* : bien-être. (2) *Définition des valeurs* : désigne les conditions physiques, matérielles et le bien-être psychologique [15]. (3) *Niveau de l'impact de la technologie sur les utilisateurs* : négatif. (4) *Acteurs concernés* : étudiants de L3 de psychologie. (5) *Explicitation de l'impact ou du non impact* : le JEN favorise les étudiants ayant choisi l'option ergonomie. (6) *Conséquences sur les utilisateurs* : les étudiants n'ayant pas choisi cette option peuvent se sentir placés en échec dans le jeu. (7) *Proposition pour réduire le conflit sur les valeurs* : l'objectif pédagogique est reformulé en intégrant des notions provenant d'une option commune aux étudiants. (8) *Proposition alternative pour soutenir les valeurs* : le suivi d'un MOOC sur les notions du JEN pour limiter le sentiment de mise en échec.

4 Perspectives

Dans cet article, nous avons proposé une solution pour faciliter les échanges entre concepteurs de JEN autour des principes du RGPD et des valeurs du *VSD*. Cette solution en cours d'implémentation sur la plateforme co.LAB, sera testée à l'automne 2021. Pour évaluer et améliorer notre proposition, nous envisageons de rencontrer des spécialistes de la protection des données, d'observer les utilisateurs de plateformes et de recueillir des traces d'activités de ces échanges.

Remerciement : le projet co.LAB est soutenu par le Fonds national suisse de la recherche scientifique (PNR 77).

Références

1. RGPD : de quoi parle-t-on ? | CNIL, <https://www.cnil.fr/fr/rgpd-de-quoi-parle-t-on>, last accessed 2020/12/16.
2. Tamburri, D.A.: Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation. *Information Systems*. 91, 101469 (2020).
3. Brun-Wauthier, A.-S., Vergès, E., Vial, G.: L'éthique scientifique comme outil de régulation : Enjeux et dérivés du contrôle des protocoles de recherche dans une perspective comparative. *Droit, sciences et techniques, quelles responsabilités ?* 61–83 (2011).

4. 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. *Éducation et didactique*. 9, 73–94 (2015).
5. Cisel, M.: Le choix d'un MOOC, un processus influencé par l'organisation des cours en catalogues ? *Sticef*. 25, 121–150 (2018).
6. Sanchez, E., Emin-Martinez, V., Mandran, N.: Jeu-game, jeu-play, vers une modélisation du jeu. Une étude empirique à partir des traces numériques d'interaction du jeu Tamagocours. *Sticef*. 22, 9–45 (2015).
7. Parlement Européen et du Conseil: RGPD, <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32016R0679#d1e40-1-1>, last accessed 2020/12/02.
8. Cavoukian, A.: Understanding How to Implement Privacy by Design, One Step at a Time. *IEEE Consumer Electronics Magazine*. 9, 78–82 (2020).
9. Prinsloo, P.: Guidelines on the Ethical Use of Student Data: A Draft Narrative Framework. (2017).
10. Sclater, N.: Developing a Code of Practice for Learning Analytics. *Learning Analytics*. 3, 16–42 (2016).
11. Bruillard, É.: L'écriture inclusive ouvre des liens surprenants. *Réflexions en didactique de l'informatique*. *Sticef*. 27, (2020).
12. Avry, S., Chanel, G., Bétrancourt, M., Pun, T., Molinari, G.: Effet des antécédents émotionnels de contrôle et de valeur sur la résolution de problème dans un jeu vidéo collaboratif. *Sticef*. 25, 195–220 (2018).
13. Chan, J.K.H.: Design ethics: Reflecting on the ethical dimensions of technology, sustainability, and responsibility in the Anthropocene. *Design Studies*. 54, 184–200 (2018).
14. Dignum, V., Baldoni, M., Baroglio, C., Caon, M., Chatila, R., Dennis, L., Génova, G., Haim, G., Kließ, M.S., Lopez-Sanchez, M., Micalizio, R., Pavón, J., Slavkovik, M., Smakman, M., van Steenberg, M., Tedeschi, S., van der Toree, L., Villata, S., de Wildt, T.: Ethics by Design: Necessity or Curse? In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. pp. 60–66. ACM, New Orleans LA USA (2018).
15. Friedman, B., Kahn, P.H., Borning, A.: Value Sensitive Design and Information Systems. In: *The Handbook of Information and Computer Ethics*. pp. 69–101. John Wiley & Sons, Ltd (2008).
16. Friedman, B., Hendry, D.: The envisioning cards: a toolkit for catalyzing humanistic and technical imaginations. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 1145–1148. Association for Computing Machinery, New York, NY, USA (2012).
17. Spiekermann, S., Winkler, T.: Value-based Engineering for Ethics by Design. *SSRN Journal*. (2020).
18. Harbers, M., Neerinx, M.A.: Value sensitive design of a virtual assistant for workload harmonization in teams. *Cogn Tech Work*. 19, 329–343 (2017).
19. France Num: Modèle de registre simplifié RGPD par la CNIL, <https://www.france-num.gouv.fr/comprendre-le-numerique/modele-de-registre-simplifie-rgpd-par-la-cnil>, last accessed 2021/04/13.
20. Colet, N.R.: Chapitre 6. Les projets d'enseignement interdisciplinaire : fondements et objectifs pédagogiques. In: *Enseignement universitaire et interdisciplinarité*. pp. 71–89. De Boeck Supérieur, Louvain-la-Neuve (2002).

Posters et démonstrations

Visualisation automatique de graphes pour l'apprentissage des algorithmes

Pierre Baron¹, Laurent Roche¹ et Thibault Raffailac^{1,2}

¹ École Centrale de Lyon

`pierre.baron@ec119.ec-lyon.fr`, `laurent.roche@ec119.ec-lyon.fr`

² LIRIS, CNRS UMR5205, F-69621, France

`thibault.raffailac@ec-lyon.fr`

Résumé. Lors de l'apprentissage des algorithmes en informatique, les visualisations sont généralement appréciées pour aider à raisonner sur des structures de données abstraites. Cependant les outils dédiés nécessitent souvent un investissement conséquent, qui freine leur adoption et réduit l'autonomie des étudiants. Cet article présente AutoGraph, un module pour Python permettant de visualiser toute implémentation d'un graphe, quel que soit son format, sans borner la pratique au sein d'un logiciel spécifique. Nous présentons ici un prototype fonctionnel, les usages auxquels il s'adresse et notre démarche de dissémination auprès de lycées et de formations post-bac.

Mots-clé: apprentissage de l'informatique · programmation · algorithmes de graphes · visualisation interactive · autonomie

Abstract. When learning algorithms in computer science, visualizations are generally appreciated to help reason about abstract data structures. However, dedicated tools often require a significant investment, which hinders their adoption and reduces the autonomy of students. This paper presents AutoGraph, a module for Python that allows visualizing any implementation of a graph, whatever its format, without restricting the practice within a specific software. We present a functional prototype, its intended uses and our approach to disseminate it to high schools and undergraduate courses.

Keywords: computer learning · programming · graph algorithms · interactive visualization · autonomy

1 Introduction

Depuis de nombreuses décennies, les algorithmes sont devenus des outils essentiels dans de nombreux domaines externes à l'informatique : génie civil, bioinformatique, économie, sociologie, etc. En particulier, les algorithmes de graphes sont très populaires et répandus aujourd'hui. Ils servent à traiter par exemple des cartes routières, des chaînes de protéines, des réseaux de change entre devises

ou encore des réseaux sociaux. Une conséquence de cette importance croissante est un enseignement de plus en plus précoce : autrefois réservés aux cursus universitaires, les graphes sont enseignés en terminale aujourd'hui [5].

Pourtant, la structure de données *graphe* et les algorithmes liés restent des notions difficiles à enseigner. En effet, contrairement à d'autres structures fondamentales (ex. tableaux, dictionnaires, piles), les langages de programmation comme Python n'ont pas de types natifs pour les graphes, ce qui rend leur implémentation particulièrement verbeuse. De plus leur représentation en code est souvent peu lisible (figure 1), ce qui impose un effort d'abstraction lors de l'implémentation d'algorithmes. Face à ces limites, nous cherchons à étudier **en quoi la visualisation de graphes lors de leur programmation peut améliorer l'apprentissage des algorithmes de graphes.**

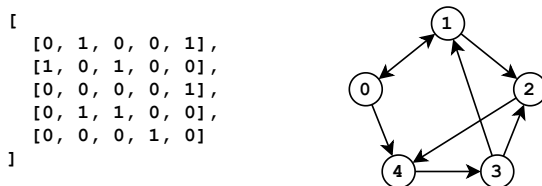


Fig. 1. Représentation d'un même graphe en code (matrice d'adjacence), et image.

De nombreux travaux de recherche ont proposé des plateformes complètes incluant éditeur de code, syntaxe de spécification de graphes et outil de visualisation intégré [6, 2, 1]. D'autres se sont basés sur des représentations textuelles interopérables (ex. DOT ou Newick) en facilitant les transformations vers et depuis ces formats [3]. Cependant en dépit de résultats positifs liés à l'engagement actif des étudiants dans l'apprentissage des algorithmes, ces outils ont été relativement peu utilisés en pratique [4].

Nous proposons ici une nouvelle alternative : la visualisation automatique de graphes, par détection d'implémentation lors de l'exécution du programme (voir <https://github.com/sical/autograph>). Ce travail en cours permet aux enseignants et étudiants d'utiliser toutes représentations plus ou moins complexes, sans être restreints à une syntaxe donnée. Nous présentons ici notre méthode de recherche, ainsi qu'une démonstration fonctionnelle en cours de validation.

2 Méthode de recherche

Notre méthode est itérative et centrée utilisateurs : nous interrogeons les enseignants et étudiants tout au long du projet, afin d'affiner notre compréhension des besoins et faire évoluer le prototype de recherche en conséquence. Ainsi une preuve de concept a été produite initialement, puis présentée à des enseignants

de 3 établissements d'enseignement supérieur locaux. Nous avons principalement relevé des environnements de développement en Python hétérogènes, ce qui a conditionné le choix d'une interface au format HTML, qui puisse être intégrée à un éditeur de code en ligne, un notebook Jupyter, ou un éditeur hors ligne avec le module `pywebview`. Le second prototype est celui présenté dans cet article. Il doit être testé directement par les enseignants, pour recueillir des besoins spécifiques à l'utilisation de l'outil. Une troisième version sera produite ultérieurement pour être utilisée en conditions écologiques par les étudiants.

Nous prévoyons de diffuser AutoGraph auprès de lycées et formations post-bac selon le protocole suivant :

- Pour chaque classe dans laquelle AutoGraph est introduit, nous organisons un workshop de 30min à 1h avec des enseignants, pour les former et recueillir des idées d'améliorations dans une démarche de co-conception.
- Durant toute la période de mise à disposition de l'outil, sous réserve d'accord des enseignants nous enregistrons les utilisations qui en sont faites, en vue d'analyser à quel point les étudiants se l'approprient.
- À l'issue de cette période, nous prévoyons des entretiens avec les étudiants et enseignants, pour recueillir en particulier (i) les problèmes observés avec l'utilisation d'AutoGraph, (ii) les fonctionnalités jugées utiles en pratique, et (iii) les différences perçues dans l'enseignement des graphes par rapport aux années précédentes.

3 AutoGraph, un module de visualisation de graphes

Ainsi qu'évoqué en introduction, la faible utilisation des environnements éducatifs de programmation de graphes nous a poussés à concevoir un outil facile à adopter. Celui-ci doit s'intégrer aux outils des utilisateurs avec le moins de contraintes possibles, et être utilisable avec le moins de connaissances préalables. Notre hypothèse est que la simplicité et la versatilité d'un tel outil favoriseront son utilisation sans supervision par les étudiants, et pourraient ainsi améliorer son impact sur l'apprentissage.

Nous avons donc conçu un module pour Python, qui puisse être utilisé aussi bien depuis un environnement Web (ex. Jupyter, tableau de bord) qu'hors ligne (ex. Spyder, PyCharm). Ce module contient une unique fonction `autograph`, qui prend en argument un objet quelconque (ex. matrice 2D, liste de paires d'entiers) et génère une visualisation au format HTML. Pour cela, l'objet est comparé à 8 implémentations reconnues de graphes. S'il correspond à l'une d'entre elles, une image est générée selon la représentation détectée. Deux observations importantes sont à l'origine d'AutoGraph :

- le "typage dynamique" de Python permet de passer à une même fonction `autograph` tous types d'implémentations sans la dupliquer en autant de variantes, ce qui est plus contraint avec les langages C/C++ ou Java.
- le nombre de possibilités *raisonnables* d'implémenter d'un graphe, c'est-à-dire qu'on rencontre dans les milieux professionnels, est limité (nous en avons énuméré pour l'instant 8).

La validation automatique de graphes permet aux étudiants de vérifier que leur compréhension d'une implémentation est correcte, et ensuite d'observer l'évolution de la structure à différents moments d'un algorithme. Il est également possible de créer un graphe par manipulation directe dans l'interface et de l'exporter en code, comme sur la plateforme <https://visualgo.net/>.

En revanche, AutoGraph n'est pas capable de détecter les implémentations erronées de graphes, car il y a trop d'erreurs possibles à prendre en compte. Il est alors important que l'équipe enseignante oriente les étudiants au préalable vers des implémentations correctes, ou qu'ils commencent par dessiner dans l'outil avant d'exporter le code correspondant. Ensuite l'outil n'est pas conçu a priori pour des usages avancés de visualisation de données ou de problèmes algorithmiques complexes. Pour ces cas de figure, AutoGraph peut convertir les graphes d'entrée pour le module externe `networkx` plus avancé.

4 Conclusion et perspectives futures

Cet article a présenté notre étude en cours sur l'utilisation de visualisations interactives de graphes, pour améliorer l'apprentissage des algorithmes. Nous avons introduit AutoGraph, un outil versatile conçu pour une utilisation non supervisée par les étudiants. À l'avenir, nous envisageons de faire évoluer cet outil pour appuyer son adoption hors du cadre expérimental, et améliorer la portée écologique des résultats. Ainsi nous considérons des scénarios tels que la génération d'images pour des sujets de TD, le support de graphes de grande taille, ou encore la visualisation d'algorithmes durant les cours.

References

1. Berry, J., Dean, N., Goldberg, M.K., Shamon, G.E., Skiena, S.: LINK: A system for graph computation. *Software: Practice and Experience* **30**(11), 1285–1302 (2000). [https://doi.org/10.1002/1097-024X\(200009\)30:11<1285::AID-SPE340>3.0.CO;2-W](https://doi.org/10.1002/1097-024X(200009)30:11<1285::AID-SPE340>3.0.CO;2-W)
2. Chen, D.Y., Chuang, T.R., Tsai, S.C.: JGAP: A Java-based graph algorithms platform. *Software: Practice and Experience* **31**(7), 615–635 (2001). <https://doi.org/10.1002/spe.379>
3. Eaton, D.A.R.: Toytree: A minimalist tree visualization and manipulation library for Python. *Methods in Ecology and Evolution* **11**(1), 187–191 (2020). <https://doi.org/10.1111/2041-210X.13313>
4. Hundhausen, C.D., Douglas, S.A., Stasko, J.T.: A Meta-Study of Algorithm Visualization Effectiveness. *Journal of Visual Languages & Computing* **13**(3), 259–290 (Jun 2002). <https://doi.org/10.1006/jvlc.2002.0237>
5. Ministère de l'Éducation nationale: Programme de l'enseignement de spécialité de numérique et sciences informatiques de la classe terminale de la voie générale. <https://www.education.gouv.fr/bo/19/Special8/MENE1921247A.htm> (Jul 2019)
6. Schliep, A., Hochstättler, W.: Developing Gato and CATBox with Python: Teaching Graph Algorithms through Visualization and Experimentation. In: Borwein, J., Morales, M.H., Rodrigues, J.F., Polthier, K. (eds.) *Multimedia Tools for Communicating Mathematics*. pp. 291–309. *Mathematics and Visualization*, Springer, Berlin, Heidelberg (2002). https://doi.org/10.1007/978-3-642-56240-2_18

Les *Teaching Analytics* pour l'auto-évaluation par analyse comportementale de l'enseignant sur un LMS

Ibtissem Bennacer, Rémi Venant et Sébastien Iksal

Université du Mans, Avenue Olivier Messiaen, 72085 Le Mans
Ibtissem.Bennacer@univ-lemans.fr Remi.Venant@univ-lemans.fr
Sebastien.Iksal@univ-lemans.fr

Résumé Ces dernières années, l'utilisation des systèmes de gestion de l'apprentissage (LMS) s'est considérablement développée. Cependant, de nombreux enseignants éprouvent de grandes difficultés pour concevoir des apprentissages pertinents pour leurs étudiants. À cet effet, il est nécessaire que l'enseignant puisse s'engager dans une activité d'apprentissage afin d'améliorer son usage du LMS. Dans cette perspective, le présent article propose une étude du comportement des enseignants sur un LMS, pour leur permettre à terme de s'évaluer et se positionner par rapport à leurs pairs. Nous proposons un modèle avec quatre axes pour décrire le comportement de l'enseignant sur un LMS, issues de modèles empiriques et théoriques. À l'aide de ce modèle, nous étudions par la suite les structures comportementales enseignantes pour identifier des profils exploitables pour l'auto-évaluation.

Mots-clés : LMS - *Teaching analytics* - *Learning analytics* - auto-évaluation - clustering.

Abstract. In recent years, the use of learning management systems (LMS) has grown considerably. However, many teachers find it very difficult to design relevant learning for their students. To this end, It is necessary for the teacher to be able to engage in a learning activity in order to improve his use of the LMS. In this context, this article proposes a study of teachers' behavior on an LMS, to allow them to assess and position themselves in relation to their peers. We propose a model with four axes to describe the behavior of the teacher on an LMS, derived from empirical and theoretical models. Using this model, we then study the behavioral structures of teachers in order to identify profiles that can be used for self-assessment.

Keywords: LMS - *Teaching analytics* - *Learning analytics* - self-assessment - clustering

1 Introduction

La tendance à l'utilisation des système de gestion de l'apprentissage (LMS) se répand maintenant rapidement dans tous les domaines de l'éducation. La plupart des universités offrent des LMS comme solution technologique « *one size*

fits all » pour tous les enseignants de toutes les disciplines. Cependant, de nombreux enseignants rencontrent plusieurs difficultés pour intégrer ces plateformes dans leurs pratiques. Leurs principaux problèmes semblent être d'ordre technique ou organisationnel, en raison du manque de soutien et de temps consacré à son apprentissage [2]. D'un autre côté, l'utilisation des LMS permet de capturer de grandes quantités de données quantitatives concernant le comportement des utilisateurs, et ouvre ainsi la voie aux *Teaching analytics* (TA) qui se réfèrent aux méthodes et aux outils pour aider les enseignants à analyser et à améliorer leurs conceptions pédagogiques. Dans cet article, nous souhaitons exploiter le domaine des TA afin d'aider les enseignants à évaluer leurs comportements sur leur LMS institutionnel. Notre objectif est de doter les enseignants d'outils d'*awareness* personnelle et sociale [3], pour les engager dans un processus d'apprentissage du LMS. À cette fin, nous visons définir, en nous appuyant sur les seules traces du LMS, les indicateurs qui permettraient de représenter son usage par les enseignants, et déterminer une typologie d'utilisation de la plateforme par ces derniers.

2 Travaux connexes

Certains chercheurs ont étudié les TA pour comprendre comment les enseignants dispensent leurs cours. Par exemple, afin de soutenir le processus d'investigation de l'enseignant (*Teacher inquiry*), [6] ont défini les TA comme un composant nécessaire, exploité en synergie avec les *Learning Analytics* (LA). Dans cette optique, [5] ont employé les TA pour extraire automatiquement les actions des enseignants. Pour s'affranchir de la dépendance au contexte technologique, [1] ont proposé un référentiel théorique des bonnes pratiques d'e-learning (DISC), tandis que [4] ont construit empiriquement un modèle pour décrire les systèmes d'apprentissage hybrides. Ces différents travaux montrent l'importance de mobiliser des outils d'analyse sur les actions de l'enseignant lui-même, dans des approches variées et à des fins diverses. Cependant, il apparaît que l'exploitation de ces comportements pour l'auto-évaluation n'a pas encore été explorée.

3 Approche méthodologique

Dans l'objectif de définir des indicateurs et un modèle comportemental permettant de représenter l'utilisation du LMS par l'enseignant, nous avons adapté le modèle de [4] à notre contexte de terrain (un LMS utilisé principalement en présentiel par l'ensemble de l'Université). Le modèle construit se fonde sur plusieurs axes liés à l'utilisation de la plateforme, ce qui permet aux enseignants d'évaluer leurs comportements sous plusieurs aspects. Nous avons donc commencé par identifier les indicateurs qui peuvent être calculés à partir de notre LMS et correspondant à chaque axe. À l'issue de cette étape, nous avons effectué une analyse globale de l'activité des utilisateurs avec toutes les variables afin de pouvoir identifier les profils types d'usage de l'ensemble de la plateforme par

les enseignants. Nous avons ensuite conduit une seconde analyse par axe pour déterminer les enseignants qui sont plus ou moins actifs selon chaque axe.

Le modèle comportemental ainsi élaboré se compose de quatre axes. Le premier axe « Participation des étudiants » concerne les activités des étudiants sur l'ensemble des fonctionnalités du LMS et comprend 15 variables (ex : accès aux ressources, forum, chat, etc), cet axe représente la partie LA qui est utilisée en synergie avec les TA exploités dans les trois axes suivants. Le second axe « Accompagnement humain » représente le support à la collaboration et à la communication par l'enseignant et comporte 19 variables (ex. : création de discussion de forum, réponse dans le chat, etc). Le troisième axe « Environnement technopédagogique » se caractérise par l'utilisation de ressources multimédias et d'outils de gestion et comprend 13 variables (ex : utilisation d'images et de vidéos, manipulation du calendrier, etc). Le dernier axe « Degré d'ouverture » est consacré aux outils externes à la plateforme et à la collaboration entre enseignants et comporte 3 variables (ex : utilisation de liens, nombre d'enseignants-éditeurs inscrits au cours).

Pour la méthode d'analyse, nous avons commencé par récupérer les traces d'activités des enseignants et des étudiants à partir de Juin 2016 jusqu'à Juillet 2018 et d'Octobre 2019 jusqu'à Novembre 2020 (une panne informatique sur le LMS de l'université ayant provoquée la perte des données entre les deux périodes). Pour faciliter l'analyse, une première transformation des données a été mise en place pour permettre de stocker celles-ci sous la forme de traces xApi¹ au sein d'un *Learning Record Store* ElasticSearch. Une fois les traces collectées, nous avons retiré les valeurs aberrantes (identifiants nuls, enseignants et cours inconnus), supprimé les cours "fantômes" (i.e. : cours où les activités des étudiants sont inexistantes) et les enseignants fantômes (i.e. : utilisateurs ayant le statut d'enseignant mais n'étant actif sur aucun cours non fantôme). Nous avons aussi gardé que les variables dont la variance était supérieure à 0.9. À la fin, nous avons obtenu un jeu de données totalisant 268 enseignants et 543 cours ayant des valeurs sur chacun de nos axes.

Par la suite, nous avons entamé l'analyse générale en combinant les indicateurs de tous les axes. Dans un premier temps, nous avons effectué une analyse en composantes principales (ACP) pour vérifier et corriger la cohérence du modèle. Dans un second temps, nous avons procédé à une analyse de clustering en testant plusieurs algorithmes (K-Means, DbSCAN, Agglomerative clustering, Gaussian Mixture) pour identifier des groupes d'enseignants en fonction de l'activité globale sur le LMS. En revanche, pour l'analyse par axe, nous avons comparé les mêmes algorithmes de clustering utilisés pour la première analyse afin d'identifier des profils comportementaux par axe.

4 Conclusion et perspectives

Dans cet article, nous avons étudié le comportement des enseignants sur un LMS institutionnel dans l'objectif de leur fournir un outil d'aide à l'engagement dans

¹ <https://xapi.com/developer-overview/>

un processus d'apprentissage. Sur la base des travaux de recherche théorique et empirique sur le comportement de l'enseignant, nous avons proposé un modèle d'utilisation à travers quatre axes qui décrivent l'usage de la plateforme par les enseignants. De ce modèle, nous avons effectué une analyse générale qui nous a permis d'identifier sept profils d'enseignants (un groupe d'enseignants peu actifs, 3 groupes d'enseignants peu actifs avec des étudiants actifs, 2 groupes d'enseignants peu actifs avec une utilisation particulière et un groupe d'enseignants très actifs). La deuxième analyse par axe nous a montré que la plupart des enseignants utilisent généralement le LMS de la même manière faible ou moyenne. Toutefois, les résultats des algorithmes de clustering semblent converger vers une détection des valeurs particulières. Ceci permet de distinguer les enseignants novices de ceux plus experts selon les différents axes.

En revanche, notre étude présente plusieurs limites. Nous avons intégré toutes les traces des enseignants sur le LMS pour analyser leur comportement, mais il est important de noter que de nombreux enseignants utilisent d'autres outils pour gérer leur enseignement auxquels nous n'avons pas accès. De plus, notre étude ne prend pas en compte les activités d'enseignement en présentiel. D'autre part, nous considérons l'ensemble des enseignants de la même manière. Si cela a l'avantage de dégager des tendances indépendantes du contexte, la prise en compte de ce dernier pourrait offrir des profils plus fins, notamment en distinguant les domaines enseignés et les niveaux universitaires visés.

Dans nos travaux futurs, nous continuons à explorer le comportement des enseignants en améliorant les axes, les indicateurs et les modèles de clustering utilisés dans cette étude. De plus, nous développons l'outil de support centré sur les enseignants, pour les doter d'*awareness* personnelle et sociale, leur permettre de se situer selon les différents axes de notre modèle, et ainsi les engager dans un processus d'apprentissage autonome ou par les pairs.

Références

1. Coomey, M., Stephenson, J. : Online learning : it is all about dialogue, involvement, support and control-according to the research. *Teaching and learning online : Pedagogies for new technologies* pp. 37–52 (2001)
2. Fung, H., Yuen, A. : Factors affecting students' and teachers' use of lms-towards a holistic framework. In : *International Conference on Hybrid Learning*. pp. 306–316. Springer (2012)
3. Gutwin, C., Stark, G., Greenberg, S. : *Support for workspace awareness in educational groupware* (1995)
4. Peraya, D., Charlier, B., Deschryver, N. : Apprendre en présence et à distance : une définition des dispositifs hybrides. *Distances et savoirs* 4, 469–496 (2006)
5. Prieto, L.P., Sharma, K., Dillenbourg, P., Jesús, M. : Teaching analytics : towards automatic extraction of orchestration graphs using wearable sensors. In : *Proceedings of the sixth international conference on learning analytics & knowledge*. pp. 148–157 (2016)
6. Sampson, D. : Teaching and learning analytics to support teacher inquiry. In : *2017 IEEE Global Engineering Education Conference (EDUCON)*. pp. 1881–1882. IEEE (2017)

Rapport au savoir en contexte muséal : le cas du jeu *Geome*

Catherine Bonnat¹, Gil Oliveira¹, Simon Morard¹, Elsa Paukovics¹, Eric Sanchez¹

¹ Université de Genève, TECFA, 1211 Genève 4, Suisse
Catherine.bonnat@unige.ch; Gil.oliveira@unige.ch;
Simon.morard@unige.ch; Elsa.paukovics@unige.ch;
Eric.sanchez@unige.ch

Résumé. *Geome* a été conçu pour amener des élèves de l'enseignement secondaire à s'interroger sur la manière dont ils appréhendent les savoirs scientifiques et leur production lors de visites scolaires. Nous avons conduit une analyse *a priori* du jeu selon une approche par composants qui sont mis en relation avec les dimensions de l'épistémologie personnelle. Cette analyse permet de comprendre comment ces dimensions peuvent être mobilisées lorsque les élèves jouent.

Mots-clés : Jeu, musée, épistémologie personnelle, analyse *a priori*.

Abstract. *Geome* aims to bring secondary school students to question the way they understand scientific knowledge and its production. We conducted *a priori* analysis of the game using an approach by its components. The game components identified were linked to personal epistemology dimensions. This analysis allows us to understand how these dimensions can be mobilized when they play.

Keywords: Game, museum, personal epistemology, *a priori* analysis.

1 Introduction

Le jeu *Geome* vise à faire vivre à des élèves de l'enseignement secondaire (12 – 15 ans), dans le cadre de visites scolaires en contexte muséal, une expérience susceptible de les conduire à s'interroger sur leur rapport au savoir ou épistémologie personnelle. Le jeu aborde la question de l'Anthropocène à travers l'étude des relations homme-nature. Ce thème concerne des problèmes complexes ainsi que des controverses scientifiques tels que la perturbation anthropique du climat, de la biodiversité et des cycles géochimiques.

La recherche que nous menons vise à comprendre comment l'usage du jeu, en contexte de visite scolaire au musée permet, à des apprenants, de s'engager dans la résolution de ces problèmes et de vivre une expérience susceptible de les conduire à réviser leur rapport au savoir. Cet article propose une analyse *a priori* du jeu et de son scénario d'usage, relativement à un modèle théorique d'épistémologie personnelle [1].

2 Cadre théorique

Les problèmes ouverts et complexes [2] sont caractérisés par le fait qu'ils ne peuvent pas être résolus avec un haut degré de certitude. La capacité d'un individu à les traiter est liée en partie à sa culture scientifique et à son rapport au savoir. Il est généralement admis que si les élèves comprennent la source et les limites des savoirs scientifiques, ils seront mieux équipés pour prendre des décisions éclairées sur des problèmes personnels et sociétaux ayant un fondement scientifique [3].

Nous problématisons la question des apprentissages sous l'angle de l'épistémologie personnelle décrite par Hofer et Pintrich [1] comme un système de croyances. Le système comporte 4 dimensions, articulées autour de 2 axes : la nature de la connaissance (croyances d'un individu à propos de ce qu'est le savoir) et la nature de l'acte de connaître (croyances sur la manière dont l'individu en vient à connaître). Ces dimensions (certitude, simplicité, source et justification) sont liées et peuvent se développer selon un continuum de représentations allant de peu à plus sophistiqué (tableau 1). L'objectif de ce papier est de présenter l'analyse *a priori* [5] du jeu *Geome*, c'est à dire à identifier les dimensions de l'épistémologie personnelle que le jeu est susceptible de mobiliser pendant les différentes phases du scénario de jeu.

3 Méthode

L'analyse *a priori* du jeu repose sur l'identification de composants de jeu (comp.) susceptibles de mobiliser une ou plusieurs dimensions de l'épistémologie. Il s'agit de caractériser les éléments constitutifs du jeu *Geome*. Pour cela, nous nous sommes appuyés sur les travaux de Ke [5], qui décrivent l'architecture du *gameplay* comme un ensemble de composants ordonnés, s'articulant autour des mécaniques de jeu et de la narration. Relativement à notre contexte de jeu au musée, nous en retenons 7 : (1) les actions qui peuvent être entreprises par le joueur, (2) les attitudes attendues, (3) les conditions à remplir pour gagner, (4) la présence de personnages qui contribuent à la structure narrative du jeu, de même que (5) l'univers et l'espace de jeu, (6) les tâches et les missions (ex. une enquête), et (7) l'avatar incarné dans le jeu. L'analyse de ces composants est effectuée à l'aide d'une grille d'indicateurs (Tableau 1), empruntée aux travaux de Kra-mar [6] et adaptée à notre situation.

Tableau 1. Tableau des indicateurs retenus pour les 4 dimensions.

Dimensions	Indicateurs
Certitude	Une information n'a pas à être remise en question
	Une information est utilisée avec précaution/on ne peut jamais être certain
Simplicité	Une information est pensée comme un fait isolé et non contextualisé
	Mise en lien d'une information avec d'autres en tenant compte du contexte
Source	Une information valable vient d'une autorité externe
	L'information est activement construite par l'individu et par les interactions entre individus

Justification	L'information est justifiée par un argument d'observation ou d'autorité, ou par une impression personnelle
	L'information est justifiée par la prise en compte de divers points de vue et du contexte pour une évaluation critique de l'expertise

La troisième étape consiste à identifier les composants de jeu susceptibles de mobiliser ces dimensions et à proposer des conjectures sur les liens entre expérience du joueur et développement épistémique.

4 Discussion des résultats et conclusion

Geome est un jeu d'équipe disponible sur tablette tactile qui permet des interactions avec les collections du Musée de la nature à Sion (CH). Le scénario pédagogique (1h30) comprend deux jeux distincts, et une phase de débriefing après chaque jeu. La conception et l'articulation des deux jeux reposent sur une métaphore qui traduit un changement du rapport de l'Homme à la nature, d'abord chasseur, puis enquêteur. Les joueurs incarnent un professionnel de la nature qui vit et travaille dans un chalet isolé dans une vallée. Bloqué par l'hiver, il sera d'abord amené à exploiter les ressources de son environnement. Les joueurs chassent, capturent ou protègent les animaux rencontrés dans les différentes salles du musée. Leurs actions impactent une barre de points collective, ("arbre de vie") correspondant à l'état de santé du milieu naturel. Cette dernière diminue si bien que les joueurs sont amenés à perdre collectivement.

Dans le second jeu, le joueur est amené, en qualité d'expert, à conduire des enquêtes relatives au milieu naturel. Les joueurs partent alors en quête d'indices qui peuvent être obtenus en scannant un élément de la muséographie. Ils sont amenés à se positionner sur la nature des informations auxquelles ils sont confrontés. La résolution de l'enquête passe par l'identification des liens entre les éléments avec lesquels ils interagissent. Chaque lien établi permet de gagner collectivement des points.

L'analyse *a priori* du jeu permet de formuler des conjectures sur l'expérience du jeu, telle qu'elle est vécue par les élèves du point de vue de son épistémologie personnelle. La dimension simplicité est concernée lorsque les joueurs interagissent (comp. 1) avec la muséographie (comp. 5). En effet, dans la première partie du jeu, lorsqu'ils scannent un objet, ils obtiennent des informations qui leur sont présentées comme des faits isolés et non contextualisés. Cependant, les conséquences globales de leurs actions individuelles (comp. 3) peuvent les amener à prendre conscience des liens systémiques entre ces mêmes objets. Dans la deuxième partie du jeu, ces mêmes informations sont contextualisées selon l'enquête qui est menée. Sa résolution nécessite une mise en relation des informations, en particulier lorsque les joueurs produisent une carte systémique (comp. 6). Ainsi, le jeu propose deux expériences distinctes du point de vue de la simplicité des savoirs pour développer la conscience de la complexité d'une information et de son traitement. Les fiches qui sont données aux élèves pour la résolution des enquêtes peuvent les conforter dans une opinion préalable. Néanmoins, l'information est présentée de manière évolutive au cours du jeu ce qui peut les conduire à remettre en cause des certitudes. En effets, les enquêtes (comp. 6) prennent la forme de problèmes

ouverts et il est attendu qu'ils soient amenés à ne pas considérer les informations comme des faits avérés mais qu'ils comprennent qu'elles peuvent être remises en question.

Ces sources d'information sont multiples dans la seconde partie du jeu. Les joueurs interagissent avec des personnages (berger, biologiste...) qui apportent ces informations (comp. 1 et 4), consultent les fiches descriptives (comp. 6) et la muséographie (comp. 5). Ces informations proviennent donc d'autorités externes dont la légitimité peut difficilement être remise en question. C'est particulièrement vrai pour le musée qui possède *a priori* une autorité scientifique. Néanmoins, le joueur a la possibilité de se forger sa propre opinion qui peut varier au cours de l'enquête. Nous faisons donc l'hypothèse que les joueurs pourront progressivement s'identifier comme de producteurs de solutions en se questionnant sur la source des informations dont ils disposent.

Les fiches descriptives et les indices récoltés doivent faire l'objet d'une évaluation critique au sein des équipes (comp. 1) lors du jeu 2 (comp. 6), car la réussite de la mission dépend d'une bonne appréciation de la qualité de l'information disponible (comp. 3), ainsi que d'une certaine résistance aux discours d'autorité (comp. 4). Il est donc attendu que le jeu mobilise l'épistémologie personnelle des élèves sur la dimension justification et qu'ils parviennent à remettre en cause certaines informations non soutenues par des arguments solides.

Ainsi, les dimensions sont dépendantes d'une combinaison de différents composants du jeu. Néanmoins, le jeu ne mobilise pas suffisamment le fait que les élèves incarnent un personnage (comp. 7). Par ailleurs, les deux temps du jeu offrent au joueur deux expériences bien distinctes. Une première expérience (jeu 1) les conduit à agir en utilisant des informations isolées sans que leur épistémologie personnelle ne soit sollicitée au-delà de la dimension simplicité, alors que le jeu 2 est de nature à les amener à s'interroger sur les quatre dimensions de leur épistémologie personnelle. Ces conjectures ont été mises à l'épreuve lors d'expérimentations du jeu avec des classes de cycle d'orientation. Les résultats participeront à l'évolution du jeu et viendront enrichir les modèles théoriques en vue d'une seconde itération.

Références

1. Hofer, B., Pintrich, P.: The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88-140 (1997).
2. Jonassen, D. Toward a design theory of problem solving. *Educational Technology Research and Development*, 48(4), 63-85 (2000).
3. Lederman, N., Lederman, J., Antink, A.: Nature of science and scientific inquiry as contexts for the learning of science and achievement of scientific literacy. *International Journal of Education in Mathematics, Science and Technology*, 1(3), 138-147 (2013).
4. Chevallard, Y.: Emploi et analyse du contrat didactique, Sur l'analyse didactique, deux études sur les notions de contrat et de situation, IREM de Marseille (1988).
5. Ke, F.: Designing and integrating purposeful learning in game play: a systematic review. *Education Tech Research Dev* 64, 219-244 (2016).
6. Kramar, N. : Apport d'un jeu sérieux pour l'instauration d'un nouveau rapport au savoir du point de vue épistémologique : cas du jeu Clim@ction. Mémoire de master, Lyon (2012).

Etude de cas : schèmes et invariants opératoires chez les professionnels de la formation par et pour le numérique – Projet Interreg - Teach Transition

Elfi Casanave¹, Pierre-André Caron²

¹ Université de Lille, ULR 4354, CIREL, F-59000 Lille, France
elfi.casanave@univ-lille.fr

² Université de Lille, ULR 4354, CIREL, F-59000 Lille, France
pierre-andre.caron@univ-lille.fr

Résumé. Notre étude propose une analyse d'entretiens de formateurs de type ethnographique aux profils variés afin de mettre en évidence les principes décisionnels sous-jacents à l'activité de formation par et pour le numérique (conception et mise en œuvre). Cette première analyse a pour but d'initier et d'éclairer la réflexion sur le design et l'ingénierie collaborative d'une formation universitaire transfrontalière au numérique pour les enseignants (projet Interreg - Teach Transition).

Mots-clés : formation, numérique, enseignants, distanciel, schèmes.

Abstract. Our study exposes an analysis of ethnographic interviews involving a varied pool of professional development practitioners. Our goal is to highlight the underlying decisional process taking place when designing and implementing digital professional training. This tentative analysis aims at sparking and feeding the reflective practice of collaborators currently engineering a new a transborder, university-led digital training-course for teachers (Interreg - Teach Transition project).

Keywords: Professional development, digital, teachers, distance, schemes.

1 Introduction

La crise de la COVID qui sévit dans le monde depuis le printemps 2020 a fait basculer sans transition le monde de l'éducation vers un usage intensif du numérique en formation. D'une intégration limitée du numérique dans les pratiques [1], écoles, universités et formations sont passées du jour au lendemain à un mode distanciel puis hybride forcé dans lequel le numérique a pris de facto une place centrale. Parallèlement à cet épiphénomène, le projet Interreg - Teach Transition, qui associe divers partenaires éducatifs Lillois et Wallons, a pour ambition de créer une formation universitaire transfrontalière ayant pour projet de former les enseignants au numérique pour leur permettre d'adapter

leurs pratiques à ces modes distanciels ou hybrides qui pourraient s'installer durablement dans le paysage éducatif.

Notre étude menée au printemps 2020 interroge un panel de professionnels de la formation par le biais d'entretiens de type ethnographique dans lesquels les répondants sont invités, par le récit, à verbaliser les principes sous-jacents à l'élaboration et à la mise en œuvre de leurs formations par et pour le numérique. Cette étude, qui constitue une première étape au sein des six phases composant notre recherche, est complétée par un questionnaire à destination des participants potentiels à cette formation afin d'en élaborer un premier profil, qui ne sera pas abordé dans notre communication.

Notre recherche a pour objectif macro-systémique d'accompagner l'élaboration d'une nouvelle formation puis d'analyser itérativement son déroulement selon les principes de la recherche collaborative orientée par la conception [2]. Nous souhaitons, à partir des résultats découlant d'une première démarche inductive, élaborer un cadre de design potentiellement utilisable pour l'ingénierie d'autres formations similaires, et évaluer la pertinence de ce cadre du point de vue des formateurs et des formés afin de le faire éventuellement évoluer au sein de notre démarche itérative.

Notre analyse de ces entretiens propose un mode de traitement mixte original et complémentaire en deux phases : la première analyse s'appuie sur une méthode venue du design [3], consistant à faire émerger collectivement et librement les points saillants des récits collectés; une seconde analyse est de type compréhensive et s'appuie méthodologiquement sur la mise en évidence de schèmes et d'invariants opératoires chez les professionnels de la formation par et pour le numérique [4].

2 Méthode de recueil des entretiens des formateurs/concepteurs

Ces entretiens semi-directifs nous ont permis de recueillir les propos de 17 formateurs et/ou concepteurs expérimentés côté français et belge aux profils divers : spécialisés dans le numérique, issus du milieu universitaire, primaire ou secondaire, enseignants ou formateurs à distance, en présentiel ou en mode hybride. Les formateurs sélectionnés ont été invités à partager avec nous la façon dont ils conçoivent en général leurs formations (numériques ou non), et à détailler un exemple spécifique de formation accompagné d'éléments concrets de leur choix (documents, tableaux, photos, schémas, dessins, etc.). Notre objectif est alors d'enregistrer un récit le plus ininterrompu et personnel possible nous permettant de recueillir les représentations de ces formateurs, de la conception d'une formation à sa mise en œuvre et de favoriser dans leur récit les digressions éventuelles, sources d'association qui pourraient nourrir notre démarche inductive. En fin d'entretien, nous avons inscrit dans notre protocole de recueil la possibilité de demander des clarifications sur l'ordonnement des activités, le suivi, l'évaluation, les compétences, les savoirs, les difficultés et les réussites, si cela n'a pas été abordé dans le récit. Les entretiens enregistrés, d'une durée moyenne comprise entre 45 minutes et une heure, ont été ensuite résumés sous format texte afin de faciliter le partage et l'analyse par les différents collaborateurs du projet. Ces résumés sont accompagnés des documents que le formateur interrogé a souhaité associer pour éclairer son propos.

3 Analyse du corpus

La première analyse du corpus a été effectuée par l'ensemble des collaborateurs du projet dans une approche design qui vise à favoriser, à partir des données recueillies, une phase d'idéation [5]: les collaborateurs ont lu chacun de leur côté l'ensemble des résumés des entretiens postés au préalable sur un tableau blanc virtuel et interactif, puis ont partagé en synchrone via un outil de visioconférence leur vision des points saillants de chaque entretien, consignés sur le tableau blanc sous forme de post-its électroniques. L'objectif de cette approche est de ne pas brider la créativité des échanges afin de favoriser l'émergence de lectures singulières et collectives originales des entretiens.

Dans un deuxième temps, notre équipe a souhaité s'inscrire dans une démarche compréhensive des propos recueillis afin de mieux cerner l'activité des formateurs par ou pour le numérique et en déduire une première approche des compétences qu'ils mobilisent dans leur quotidien professionnel. Pour cela nous avons entrepris d'analyser le corpus d'entretiens à la lumière des modèles théoriques de Vergnaud [4] et d'Engeström [6]. Nous avons choisi d'effectuer une analyse thématique de chaque entretien, puis une analyse thématique transversale afin de faire émerger dans une démarche itérative des catégories essentielles d'activités décrites par les formateurs.

Une meilleure compréhension des schémas des formateurs-concepteurs nous paraît à même de favoriser la construction d'une formation universitaire à visée professionnalisante, d'une part en contribuant à leur identification et à leur enseignement, d'autre part en contribuant au développement même de ces schémas chez les apprenants. Cette démarche s'inscrit dans une exploration que nous menons dans le cadre notre projet sur la nature des TICE : entre savoir disciplinaire et pratiques professionnelles.

4 Premiers résultats

La plupart des activités décrites par les formateurs dans les entretiens, tant au niveau de la conception que de la mise en œuvre, peuvent être classées selon deux types de démarches. Les activités semblent relever tantôt d'une démarche anticipative ou régulative, tantôt d'une démarche adaptative au groupe ou à l'individu [4].

En outre, ces actions semblent toujours influencées par un certain nombre de facteurs au niveau macro, méso, et micro [9], que le formateur internalise dans sa démarche. Ces influences sont parfois verbalisées clairement par le formateur mais semblent à d'autres moments intériorisées sans verbalisation. Nous avons fait une place particulière dans notre analyse, par emprunt à la théorie de l'activité d'Engeström [6], aux activités collectives et collaboratives dans la forme comme dans la participation, qui dénotent une vision de l'individu comme apprenant social, au sein d'une communauté. Nous avons synthétisé notre analyse de l'activité de formateur dans la figure 1 (Fig.1).

Lien vers le schéma synthétique de l'activité des formateurs et des différents niveaux d'influence (Fig.1) : <https://teachtransition.eu/content/uploads/2021/04/fig-1-schema-synthetique-de-lactivite-de-formateur-et-des-differents-niveaux-dinfluence.jpg>

5 Mise en perspective

Notre approche mixte dans le projet Interreg - Teach Transition intègre des techniques issues du design et de l'analyse de l'activité aux processus plus classiques d'ingénierie pédagogique, afin d'articuler l'art du faire pratique et l'art du faire scientifique comme le suggère Paquelin [7]. Les premiers résultats émanant de notre analyse mixte de ces entretiens de formateurs corroborent les récentes recherches indiquant la tension constante ressentie par le formateur-concepteur, oscillant entre rigueur prescriptive et nécessité de s'adapter aux apprenants [8].

L'approche brièvement décrite ci-dessus, grâce aux emprunts pluridisciplinaires et à son fonctionnement itératif, nous permet, dans le cadre de ce projet, une triple mise en abîme : elle interroge notre action de chercheurs-concepteurs-formateurs et donc la démarche de design que nous proposons. Elle questionne les compétences que notre formation souhaite contribuer à développer chez les apprenants, et elle influence les contenus à enseigner. Notre démarche s'inscrit et trouve son écho dans le débat lancé en mars 2020 dans la revue *Distances et Médiations des Savoirs* [9] sur la forme et le rôle de l'ingénierie pédagogique.

Références

1. Brunel, S.: De la didactique des usages numériques. Éditions Universitaires Européennes (2014).
2. Caron, P.-A.: Ingénierie dispositive et enseignement à distance au temps de la COVID 19. *Distances et médiations des savoirs*, 30, (2020).
3. Engeström, Y. et al.: *Perspectives on Activity Theory*. Cambridge University Press (1999).
4. Neumann, M.: Ideation reference process model for the early phase of innovation. Université de Grenoble (2013).
5. Nova, N. et al.: De l'ethnographie au design, du terrain à la création : tactiques de traduction. *Sciences du Design*. Presses Universitaires de France. 1–1, 86–93 (2015).
6. Paquelin, D.: Repères pour une ingénierie interactionniste situationnelle. *Distances médiations et savoirs*, 32, (2020).
7. Peraya, D., Peltier, C.: Ingénierie pédagogique : vingt fois sur le métier remettons notre ouvrage. *Distances médiations et savoirs*, 29, (2020).
8. 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. *Éducation et didactique*. Presses Universitaires de Rennes. 9–2, 73–94 (2015).
9. Tarrit, C.R.: Étude d'un Dispositif Pédagogique Instrumenté Complexe et de ses propriétés. Une approche épistémique et méthodologique, le cas d'une université brésilienne. *Univ. Lille Cirel-Trigone* (2015).
10. Vergnaud, G.: La théorie des champs conceptuels. *Recherches en didactique des mathématiques*. 10, 2–3, (1990).

***DigiTrainer*, un outil pour former ceux qui forment**

Pierre-François Coen¹, Edmée Runtz-Christan²

¹ Haute école pédagogique de Fribourg, UR-EVIDENS, CH-1700 Fribourg, Suisse
pierre-francois.coen@edufr.ch

² Université de Fribourg, CERF, CH-1700 Fribourg Suisse
edmee.runtz-christan@unifr.ch

Résumé. Les institutions de formation des enseignant·e·s forment leurs étudiant·e·s dans une logique d'alternance entre institution et terrain professionnel. Dans ce cadre, les personnes en formation sont accompagnées selon différentes modalités. *DigiTrainer*, la plateforme présentée dans ce poster, permet de travailler sur cet accompagnement en partant du principe que les feedback des étudiant·e·s sont des moyens efficaces pour permettre aux formateur·rice·s d'améliorer leur fonctionnement. Via des questionnaires classiques, elle permet de déterminer la posture et les modalités d'accompagnement, mais aussi de traiter les entretiens de formation en analysant les tours de parole, en identifiant les locuteur·trice·s et la proportion de leur discours respectif. Un outil de transcription automatique permet de faire des traitements lexicométriques : quantification de certains termes, identification de mots-clés ou locutions remarquables.

Mots-clés : formation des enseignant·e·s, accompagnement, analyse automatique du discours, plateforme numérique.

Abstract. Teacher training institutions train their students on a vocational principle which involves trainees alternating between the training institution and a practical on-the-job work placement. In this context, students are supported in various ways. *DigiTrainer*, the platform presented in this poster, allows us to work on this support based on the principle that feedback from students in training is an effective means of enabling teacher trainers to improve their performance. Using standard questionnaires, the platform allows the posture and the methods of support to be determined. The tool also makes it possible to analyse the training interviews by analysing speech turns, identifying the speakers and the proportion of their respective speeches. An automatic transcription tool enables lexicometric processing: quantification of certain terms, identification of key words or remarkable phrases.

Keywords: teacher training, coaching, automatic speech analysis, digital platform.

1 Introduction

Dans le contexte de la formation des enseignant·e·s, l'alternance entre des cours dispensés dans l'institution de formation et des stages sur le terrain professionnel est quasiment la norme. Ce dispositif de formation peut prendre différentes formes (alternance juxtapositive, associative ou intégrative) [1] qui visent à articuler, avec plus ou moins d'intensité, les apports théoriques dispensés dans l'institution avec les expériences professionnelles vécues sur le terrain. Pour soutenir cette alternance, les institutions ont mis en place un accompagnement des personnes en formation. Selon les contextes, ce suivi est réalisé par des enseignant·e·s formateur·trice·s titulaires (qui accueillent des stagiaires dans leur classe), par des formateur·trice·s de l'institution (qui suivent le développement des étudiant·e·s) et parfois, par des groupes d'enseignant·e·s constitués en communautés de pratiques (qui assurent un encadrement partagé).

2 Accompagner les enseignant·e·s en formation

La manière d'accompagner un stagiaire influence de façon importante son développement professionnel [2]. Pour assurer un encadrement de qualité, les formateur·trice·s de terrain adoptent des postures (définies comme une manière d'agir dans une situation d'accompagnement susceptible d'évoluer en fonction d'une formation, d'une situation ou d'un projet [3]) et des modalités variées et ajustées aux besoins du stagiaire [4]. L'accompagnement se centre sur le développement professionnel alliant une attitude réflexive sur les pratiques et une autoévaluation. Néanmoins, le style d'accompagnement mobilisé par le·la formateur·trice de terrain dépend de divers facteurs circonstanciels [5] tels que le statut de l'accompagnant·e, l'objet de l'accompagnement, la fonction de l'accompagnement, et de forces et faiblesses révélées par la situation d'accompagnement. L'intersubjectivité immanente à la relation existante entre les deux acteur·trice·s va les contraindre à se remettre en question, à effectuer un travail sur eux, à progresser ensemble. En outre, le·la formateur·trice de terrain doit connaître et comprendre les attentes des apprenant·e·s pour les former au mieux [6]. Il doit être capable de changer de posture en fonction des situations [7], pour finalement s'effacer : le paradoxe de tout accompagnement étant de ne plus exister [8].

3 *DigiTrainer*, un outil pour former à l'accompagnement

Le développement des compétences d'accompagnement est au cœur du projet *DigiTrainer*. Il part du principe qu'une réflexion sur son activité conduit à une prise de conscience par les formateur·trice·s de terrain de leur posture d'accompagnement : se rendre compte de comment on accompagne permet d'améliorer son accompagnement. Cette prise de conscience s'appuie sur la collecte et l'exploitation de données produites autant par le formateur lui-même (réflexion) que par la personne accompagnée (feedback). C'est là que les ressources numériques apparaissent comme précieuses parce qu'elles facilitent la saisie de données probantes qui permettent l'accès à la conscience

des aspects opaques de l'activité [9]. Selon Mailles-Viard [10], certains usages du numérique contribuent à développer l'autoévaluation et l'autorégulation. C'est une des perspectives que développe l'analytique des apprentissages (*learning analytics*) qui permet de consolider la prise d'informations, d'organiser, de synthétiser et de modéliser les données recueillies pour les rendre intelligibles aux acteurs et développer leur réflexivité [11, 12]. Contrairement à beaucoup de démarches qui saisissent un grand nombre de données sans permettre aux utilisateurs de les contrôler, le dispositif proposé ici implique fortement les acteurs [13]. La fonction essentielle de *DigiTrainer* - qui se présente sous la forme d'une plateforme web - est de permettre à la personne en formation de contribuer à la formation de la personne qui la forme.

4 Les données prises en compte dans *DigiTrainer*

La plateforme *DigiTrainer* prend en compte deux types de données différentes : 1) des données (provoquées) issues de questionnaires sur les postures et les modalités d'accompagnement ; 2) des données (invoquées) issues d'enregistrements des entretiens d'accompagnement.

4.1 Les données issues des questionnaires

Durant son stage, les acteur·trice·s sont invité·e·s à compléter à différentes reprises, des questionnaires permettant d'identifier les postures et les modalités d'accompagnement. Issus des travaux sur l'accompagnement des enseignant·e·s en formation [14], ces questionnaires permettent de déterminer une posture d'accompagnement (imposeur, organisateur, coconstructeur, facilitateur ou émancipateur) et des modalités regroupées autour de trois axes (direction, relation, réflexion). La plateforme enregistre ces informations et les restitue via un tableau de bord qui les présente sous forme graphique. En intégrant plusieurs sets de données, elle permet de comparer les données de l'étudiant·e avec celles de l'accompagnant·e et de voir l'évolution des réponses tout au long de la formation et de tenter de comprendre les écarts de perception.

4.2 Les données issues des entretiens

Dans le cadre du mentorat, les acteur·trice·s font régulièrement des entretiens pour débriefer une leçon ou pour faire le bilan d'une journée. *DigiTrainer* récupère les enregistrements de ces entretiens et propose différents traitements conversationnels. En analysant les tours de parole, il identifie les locuteurs et la proportion de leur discours respectif. Il donne accès à des transcriptions complètes à partir desquelles différents traitements lexicométriques sont possibles : quantification de certains termes, identification de mots ou locutions remarquables ... Ces transcriptions peuvent être utilisées lors de formations au mentorat.

5 Usages et perspectives

La plateforme agit comme un puissant révélateur des postures et des modalités d'accompagnement et permet aux formateur·trice·s de terrain de disposer de données à la fois quantitatives et qualitatives possibles à articuler entre elles par exemple en repérant des locutions verbales ou des mots associés à certaines postures. Le développement de l'outil est accompagné d'un projet de recherche collaborative auquel sont associés des chercheur·e·s et des formateur·trice·s. Cette étude vise à analyser comment les acteurs s'approprient l'outil, mais aussi de voir comment il impacte leurs pratiques d'accompagnement. Par exemple, il est envisagé de voir comment certains termes (feedback) utilisés caractérisent certaines postures ou modalités d'accompagnement.

References

1. Vanhulle, S., Merhan, F., Ronveaux, C.: Du principe d'alternance aux alternances en formation des adultes et des enseignants : un état de la question. In : Merhan, F., Ronveaux, C., Vanhulle, S. (eds.) *Alternance en formation*, pp. 5-45. DeBoeck, Bruxelles (2007).
2. Crasborn, F., Hennissen, P., Brouwer, N., Korthage, F., Berger, T.: Mapping mentor teachers' roles in mentoring dialogues. *Educational Research Review* 3(2), 168-186 (2008).
3. Runtz-Christan, E.: Postures d'accompagnement. In Runtz-Christan, E., Coen, P.-F.(eds) *Collection de concepts-clés de la formation des enseignant·e·s en Suisse romande et au Tessin*, 208-211. Lausanne, LEP (2021)
4. Paul, M.: *La démarche d'accompagnement : repères méthodologiques et ressources théoriques*. De Boeck, Burxelles (2016).
5. Cosnefroy, L., Hoffmann, C., Douady, J.: *L'accompagnement méthodologique*. *Recherche en formation* 77, 29-44 (2014).
6. Runtz-Christan, E.: *Pour comprendre les attentes des stagiaires-enseignants à l'égard de leurs enseignants-formateurs*. CRIFPE 2017.
7. Crasborn, F., Hennissen, P.: Training mentor teachers for effective supervision : The impact of the SMART programme. In: Arnold, A., Gröschner, A., Hascher, T.(eds.) *Schulpraktika in der Lehrbildung: Theoretische Grundlagen, Konzeption, Prozesse und Effekte*, 377-397. Münster, Waxmann (2014).
8. Biémar, S.: Accompagner un groupe d'enseignants dans une école : une grille de compétence. In : Charlier, E., Biémar, S. (eds.) *Accompagner un agir professionnel*, 19-33. De Boeck, Bruxelles (2012).
9. Coen, P.-F.: Les technologies, des aides précieuses pour développer la réflexivité des apprenant. *Formation et pratiques d'enseignement en questions* 3, 149-160 (2006).
10. Mailles-Viard Metz, S.: L'aide du numérique aux activités d'auto-évaluation. *Revue internationale de pédagogie de l'enseignement supérieur* 31(3) (2015).
11. Ferguson, R.: Learning analytics : drivers, developments and challenges. *International Journal of Technology Enhanced Learning* 4(5-6), 304-317 (2012).
12. Ifenthaler, D.: Are Higher Education Institutions Prepared for Learning Analytics ? *TechTrends*, 61(4), 366-371 (2017).
13. Kruse, A., Ponsajapan, R.: *Student-Centered Learning Analytics*. CNDLS, Thought papers (2012).
14. Colognesi, S., Van Nieuwenhoven, C., Runtz-Christan, E., Bélair, L.: Un modèle de postures et d'interventions comme ensemble dynamique pour accompagner les pratiques en situation professionnelle. *Phronesis*, 5-21 (2019).

Méthodologie d'analyse des contributions rédactionnelles lors d'une co-écriture en ligne

Edith Durand¹, Paul Lotin¹, Damien Chabanal¹, Simon Ensor¹, Loïse Jeannin², Anne-Laure Foucher¹, Lidia Lebas¹, Véronique Quanquin¹, Ciara R. Wigham² et Christine Blanchard¹

¹ Université Clermont Auvergne, LRL, F-63000 CLERMONT-FERRAND, FRANCE
edith.durand@uca.fr, paul.lotin@uca.fr, damien.chabanal@uca.fr,
simon.ensor@uca.fr, a-laure.foucher@uca.fr,
lidia.fraczak@uca.fr, veronique.quanquin@uca.fr,
christine.blanchard@uca.fr

² Université Clermont Auvergne, ACTé, F-63000 CLERMONT-FERRAND, FRANCE
loise.jeannin@uca.fr, ciara.wigham@uca.fr

Résumé. Dans cette contribution, nous présentons une méthodologie de recueil et d'analyse de données afin d'évaluer de quelle manière des étudiants de niveau universitaire collaborent à l'aide d'un support d'écriture collaborative numérique en vue de produire un texte de vulgarisation scientifique. L'objectif est d'étudier les liens entre les styles de collaboration lors du processus d'écriture et la qualité du texte final. Cette méthodologie doit également permettre de mesurer l'influence des commentaires des co-auteurs et des enseignants sur la qualité du texte.

Mots-clés : Apprentissage collaboratif médiatisé par ordinateur, environnement numérique de travail, pédagogie universitaire, visualisation, Etherpad.

Abstract. This study presents a methodological approach which enables the investigation of how higher education students co-create articles aimed at popularising science for non-academic readers using an online collaborative writing application. In future research, we will identify the modes of collaboration employed by the students during the writing process and explore how their different ways of working affect the quality of their final texts. We will also study the influence of co-author and teacher comments on final text quality.

Keywords: Computer supported collaborative learning, virtual working environment, higher education pedagogy, visualisation, Etherpad.

1 Introduction

L'approche par compétences (APC) est aujourd'hui au centre des réflexions sur la formation, notamment dans le supérieur [1]. La présente étude s'intéresse plus particulièrement à la façon dont les étudiants collaborent lorsqu'une activité d'écriture est mise en œuvre avec des pairs. Cette activité, dont plusieurs études montrent les bénéfices [2–6], vise en particulier le développement d'une littératie numérique. Cette recherche fait partie du projet Acol@d¹, qui étudie les données d'un cours (116 étudiants répartis en 29 groupes de 3 à 5 étudiants) proposant une situation d'apprentissage mixte (synchrone et asynchrone) hybride, soutenue par un outil d'écriture dynamique (Etherpad) et l'utilisation de visio-conférence. L'une des visées du projet concerne une meilleure compréhension de ce processus en vue de proposer un guidage pour participer à un projet d'écriture collaborative, répartir les tâches et co-écrire.

Nous nous intéressons ici à la manière dont les étudiants co-écrivent à l'aide du support Etherpad, en dégageant des styles de collaboration. Le style de collaboration en co-écriture, initialement défini par Lowry et al. [7], est reflété par la progression dans la co-écriture du texte. Cette progression est désormais accessible et analysable par les éditeurs de texte actuels qui permettent une visualisation historique dynamique de la co-construction de textes [8], ce qui est innovant. De récentes études ont montré une variabilité des styles de travail parmi des groupes de co-auteurs et que les textes dans lesquels les co-auteurs participaient de façon équilibrée étaient jugés de meilleure qualité [3,4]. De plus, les commentaires réalisés sur Etherpad par les apprenants et les enseignants peuvent avoir un effet sur la performance en co-écriture [9].

Suivant la méthodologie innovante de visualisation dynamique de co-écriture et selon les effets potentiels des commentaires dans la performance d'écriture, nous proposons d'étudier deux questions de recherche : 1. quelle est l'influence du style de collaboration et de la balance de participation entre les co-auteurs sur la qualité du texte final ? 2. quelle est l'influence du type de commentaires (co-auteurs et enseignants) sur la production du texte ? Nous présentons ici la méthodologie qui nous permettra d'apporter dans nos travaux futurs des réponses à ces questions.

2 Terrain d'étude

Le projet Acol@d a pour terrain l'enseignement « Plateformes de communication à distance », dispensé au semestre 2 en 3ème année de Licence Sciences du Langage Département de Linguistique, UFR Lettres, Culture et Sciences Humaines, Université Clermont Auvergne). Dans ce cours hybride (séances synchrones en présentiel et à distance), les étudiants co-écrivent en petits groupes des articles de vulgarisation scientifique, à partir de leurs connaissances acquises dans leur formation de spécialité. Pour

¹ Apprentissage Collaboratif à Distance, Laboratoire de Recherche sur le Langage, UFR Lettres, Culture et Sciences Humaines, Université Clermont Auvergne, France, Programme Learn'in Auvergne, CAP20/25, Agence Nationale de la Recherche, Programme Investissements d'Avenir (16-IDEX-0001 CAP 20-25)

ce faire, ils suivent un processus rédactionnel [10] (planification, mise en texte, révision). Lors de la mise en texte, les étudiants utilisent des pads (page d'écriture générée par Framapad© supportant Etherpad) : les contributions de chaque participant sont identifiées car surlignées par une couleur. Le processus d'écriture-même est identifiable, grâce à la fonction historique dynamique : chaque saisie au clavier est mémorisée.

3 Méthodologie d'analyse des données

Pour répondre à notre première question de recherche, un utilitaire a été développé par notre équipe pour extraire à chaque temps de révision enregistré dans le pad, le nombre et le type de révision (ajout/suppression) ainsi que le nombre de caractères écrits par chaque co-auteur en suivant les métriques proposées dans Olson et al. [3] et Wang et al. [8]. Les contributions et leur temporalité peuvent ainsi être relevées en tenant compte de tout le processus d'écriture. Nous utilisons le nombre de caractères écrits par auteur comme mesures de participation au texte (procédure d'Olson et al.).

La balance de participation est calculée pour chaque co-auteur en regard du nombre de caractères inscrits dans le texte. Elle est mesurée par la variance de la contribution de participation en nombre de caractères à chaque fin de session d'écriture pour chaque co-auteur ; la balance de participation de l'équipe est de 0. La normalisation de participation de chaque co-auteur par session permet d'évaluer la progression de contribution de chaque co-auteur dans chaque session puis en moyenne sur la totalité des sessions.

Pour mesurer le style de collaboration, la nomenclature d'Olson et al. est appliquée en observant le codebook publié par ces auteurs en lien avec la visualisation des données extraites des pads (e.g. le style From scratch partant de zéro et élaborant progressivement le texte versus le style Outline où un membre définit le plan qui sera complété collectivement).

Pour mesurer la qualité de l'article final, les textes sont distribués de façon aléatoire entre membres de l'équipe de recherche. Après une première cotation, qui s'appuie sur la grille élaborée par l'équipe, les codeurs évaluent leurs différences, clarifient les rubriques au besoin, et évaluent les travaux une seconde fois. Un score final inter-juge (Krippendorff's alpha score) est calculé pour s'assurer de la fiabilité de l'évaluation.

Enfin, l'analyse interactionnelle nous permettra d'identifier les variables pouvant expliquer le mieux la qualité finale du texte produit. Pour ce faire, une analyse de régression multiniveaux est appliquée en suivant une méthodologie similaire à Temperman et al.[2] ayant mené ces analyses sur un nombre similaire de groupes. Cette technique d'analyse permet de mettre en relation une variable prédite, ici la qualité du texte final, avec un ensemble de variables prédictives, ici la balance de participation, le nombre et type de commentaires.

Pour répondre à la deuxième question de recherche, nous procédons suivant la méthodologie suivante : les données de clavardage de chaque pad, incluant les commentaires des co-auteurs et des enseignants, sont exportées en Word pour l'identification et la visualisation des moments de commentaires des enseignants.

Une analyse qualitative du type de commentaires - polarité (positif/négatif), organisation (assignement des rôles, structure du texte) - sera menée à l'aide du logiciel d'analyse Nvivo©. Des analyses statistiques (corrélation) permettront de mesurer les effets des types de commentaires sur la qualité du texte.

4 Conclusion et perspectives

La méthodologie présentée ici permet une analyse du processus de co-écriture dans toutes les étapes d'une production écrite. Nos analyses à venir permettront de dégager des recommandations pratiques pour la conduite de projets d'écriture collaborative.

Références

1. Chauvigné C. L'approche par compétences : un nouveau paradigme pour la pédagogie universitaire ? *Revue française pédagogie* 2010.
2. Temperman G, Walgraeve S, Lièvre B de, Boumazguida K. Développer des compétences de conceptualisation et d'analyse avec un forum de discussion et un etherpad. *Sci Technol l'Information la Commun pour l'Éducation la Form* 2017. <https://doi.org/10.3406/stice.2017.1729>.
3. Olson JS, Wang D, Olson GM, Zhang J. How people write together now: Beginning the investigation with advanced undergraduates in a project course. *ACM Trans Comput Interact* 2017. <https://doi.org/10.1145/3038919>.
4. Sundgren M, Jaldemark J. Visualizing online collaborative writing strategies in higher education group assignments. *Int J Inf Learn Technol* 2020. <https://doi.org/10.1108/IJILT-02-2020-0018>.
5. Blanchard Rodrigues C. Bénéfices des pads à la co-écriture numérique 2020. <https://doi.org/10.4000/ticetsociete.1314>.
6. Blanchard Rodrigues C, Andriamasinoro M. Écriture numérique à distance : interventions de co-écriture dans un environnement dynamique. s. d.
7. Lowry PB, Curtis A, Lowry MR. Building a taxonomy and nomenclature of collaborative writing to improve interdisciplinary research and practice. *J Bus Commun* 2004. <https://doi.org/10.1177/0021943603259363>.
8. Wang D, Olson JS, Zhang J, Nguyen T, Olson GM. DocuViz: Visualizing collaborative writing. *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2015- April, New York, NY, USA: Association for Computing Machinery; 2015, p. 1865-74. <https://doi.org/10.1145/2702123.2702517>.
9. Abdullah MY, Hussin S, Alnaseeri M. The Effect of Peers' and Teacher's E-Feedback on Writing Anxiety Level Through CMC Applications Supyan Hussin. *Artic Int J Emerg Technol Learn* 2018. <https://doi.org/10.3991/ijet.v13i11.8448>.
10. Flower L, communication JH-C composition and, 1981 undefined. A cognitive process theory of writing. JSTOR s. d.

Articuler présence et distance à la suite du confinement : genèse des pratiques hybrides ?

Gaëlle Lefer Sauvage¹, Cendrine Mercier²

¹ Centre Universitaire de Formations et de Recherches de Mayotte
² Université de Nantes, site Le Mans

Résumé. À la suite du premier confinement (mars 2020), la reprise en classe a amené une enseignante à la rédaction d'un cahier journal ayant fait l'objet d'un dialogue avec une chercheuse, qui met en valeur une multitude de ressources et de dispositifs distanciés et multimodaux [2]. Cet exemple exemplaire de pratiques pédagogiques interroge la genèse des pratiques hybrides et ses généralisations dans le premier degré. À partir des connaissances issues des travaux effectués à l'université et dans le supérieur (notamment à la suite du projet HySup), et de la réflexion plus générale sur le numérique à l'école [16], des enjeux de la genèse des pratiques enseignantes hybrides dans le premier degré sont questionnés.

Mots-clés : hybridation, distanciel, COVID-19, pratiques enseignantes.

Abstract. Following the first confinement (March 2020), a teacher writes a diary notebook after the return to class. A dialogue with a researcher is created, which provides information on the situation in the classroom : a multitude of distant and multimodal resources and systems are then discussed with the researcher [2]. This 'exemplary example' of pedagogical practices has led to a reflection on the genesis of hybrid practices and its generalizations in the first degree. Based on the knowledge resulting from work carried out at university and higher education (particularly following the HySup project), and from more general reflection on digital technology at schools [16] the issues involved in the genesis of the blended learning of teaching practices at primary level are questioned.

Keywords: blended learning, hybrid approach, online learning, COVID-19.

1 Introduction

La pandémie de coronavirus (COVID-19) a mis le monde au défi et particulièrement le monde de l'éducation en modifiant les pratiques enseignantes et les organisations associées dans une nouvelle définition de la forme scolaire, des processus d'enseignement-apprentissage et des liens avec les familles [13, 14]. De nombreux rapports de recherches nationaux [1, 9, 15, 16] et de synthèses [13, 14], sur les pratiques enseignantes

avant et pendant le confinement ont vu le jour. Tous mettent en lumière un bouleversement probable dans les pratiques, avec un enseignement à distance d'urgence [14] imposé, qui amène à bricoler, innover, adapter, ou approfondir, leurs pratiques instrumentées. Les recherches mettent en valeur une utilisation quotidienne des outils numériques dans le cadre d'un usage scolaire professionnel, tels l'ENT (Environnement Numérique de Travail) ou les messageries [13], bien que ces données soient à relativiser selon que les enseignants travaillent en premier ou second degré, mais aussi selon les contextes d'enseignement équipés en infrastructures numériques et nouvelles technologies [1]. Des innovations par rapport aux usages antérieurs, tels que les classes virtuelles ou encore les sites internet, ont même vu le jour. Ne pouvant parler de travail en distanciel, mais plutôt à distance [2], la question est de comprendre comment le contexte de pandémie a encouragé ou favorisé des pratiques instrumentées hybrides. On part du principe que les pratiques hybrides existent dans le premier degré (1er degré), avec une entité propre par rapport à l'université : elles sont probablement multiformes, micro-situées et font intervenir des outils nouveaux par rapport au quotidien [4]. Une étude de cas d'une enseignante chevronnée (traces laissées dans un journal de bord hebdomadaire et un entretien individuel post confinement), nous permet de questionner le concept d'hybridation (conçu au départ comme une articulation présence-distance à travers un support technopédagogique) et de proposer des pistes scientifiques sur la genèse des pratiques hybrides.

2 L'hybridation dans le premier degré : une genèse propre et distincte de celle de l'enseignement supérieur ?

Une étude de cas d'une enseignante du 1er degré [2] a mis en valeur la diversité des pratiques pédagogiques et des micro-innovations quotidiennes. A la suite d'un retour en classe après le premier confinement (mars-juillet 2020), la gestion du protocole amène l'équipe pédagogique à repenser l'accueil et le flux des enfants dans l'école (15 élèves maximum) et hors école pour les « élèves empêchés » [15]. L'enseignante expérimentée dans la profession bénéficie d'une infrastructure technologique non négligeable qu'elle utilise très régulièrement en classe et hors classe, avant, pendant et après le confinement : classe virtuelle, livret scolaire unique (qu'elle nomme « livret scolaire numérique »)¹, MP3, dictaphones, tablettes, pad² de classe, vidéo-projecteur, imprimante, messagerie électronique avec les parents, plateforme avec un ENT. Cette étude de cas semble mettre en évidence des formes d'improvisation non conscientisées d'enseignement hybride-flexible (HyFlex), appelées également « comodales » [13], qui permettent de proposer aux élèves des sessions de cours en face à face et en ligne de manière synchrone (et des ressources en mode asynchrone ou synchrone).

L'enseignante du 1er degré innove en mettant en place des classes virtuelles pour le groupe à distance en situation synchrone, en plus d'un enseignement en présentiel. Elle

¹ Le LSU suppose au minimum, l'accès à internet au niveau de l'école, l'accès au portail ARENA pour les directeurs/trices d'école et une clé OTP, mais aussi, pour l'enseignant, un outil support à l'application LSU (tablette ou ordinateur)

² Un texte collaboratif créé à partir d'un éditeur de texte collaboratif en ligne

propose des ressources déposées en asynchrone sur une plateforme ou envoyées par mail, et détourne des outils à des fins pédagogique et administrative (usage de la plateforme ENT entre collègues pour répondre à un dispositif de décloisonnement par exemple). Mais est-ce un idéal vers lequel tendre pour l'ensemble du premier degré ? Notre étude de cas interroge une évolution du paradigme actuel [10], en tenant compte de la spécificité des contextes et la diversité des profils enseignants³. Il s'agit alors de comprendre ce qui participe à des pratiques instrumentées hybrides dans le 1er degré. La lourdeur du dispositif HyFlex amène cette enseignante à évoquer un aspect « *chronophage* ». Elle estime que, pour elle-même, le multimodal est coûteux sur le plan cognitif, et pour les élèves, dangereux sur le plan attentionnel. Les contraintes techniques entraînent des situations d'apprentissages empêchées. Les modalités pédagogiques sont également peu utilisables et acceptables au quotidien (1 élève se connecte uniquement alors qu'ils sont susceptibles d'être 4 en distanciel). Quand bien même cette enseignante est portée par des valeurs inclusives, les limites du dispositif, et inadéquation aux spécificités des jeunes enfant, sont questionnées [16]. Or, un des principes fondateurs du dispositif HyFlex est l'accessibilité [13]. La question se pose de comprendre la manière dont se développent les pratiques hybrides ou HyFlex quotidiennes des enseignants du 1er degré, en décontextualisation par la pédagogie universitaire et recontextualisant au 1er degré.

3 L'hybridation dans le premier degré : une genèse propre et distincte de celle de l'enseignement supérieur ?

L'enjeu de la réflexion sur l'hybridation des pratiques au premier degré est essentiel et largement politique, pour savoir vers quelle société tendre : à vouloir imaginer une scolarisation des enfants en distanciel (tel c'est le cas pour l'enseignement supérieur [3, 5, 6, 7, 15]), c'est omettre les limites de ces pratiques dans l'enseignement supérieur et la complexité des systèmes scolaires dans lesquelles elles s'inscrivent [8, 11]. Cette approche hybride ordinaire dans le premier degré passe par le développement de l'autonomie des élèves, qui demande des ajustements pédagogiques importants [3, 4]. L'engagement des élèves dans leur parcours scolaire est un point qui semble agir sur leur attention, au point de décrocher à la moindre contrainte technique (cela explique probablement le décrochage numérique des élèves durant les classes virtuelles). Plus globalement, l'injonction à l'hybridation prend le risque d'aller vers un affrontement voire un refus des pratiques hybrides naissantes, microscopiques, bricolées et non conscientisées. La réaction de repli et de fermeture [8], mais aussi d'individualisation de l'offre de formation, et d'une école à deux vitesses [17], dont l'étude de cas est un paroxysme, font partie des risques déjà visibles de l'hybridation et probablement du distanciel. Nous questionnons alors la genèse des pratiques enseignantes hybrides à travers cette prise de conscience de la multiplicité des hybridations pédagogiques quotidiennes et ordinaires (supports, ressources, pratiques pédagogiques, rôle de l'enseignant, compétences des élèves).

³ Lutte contre les classes virtuelles : <https://paris-luttes.info/appele-a-nos-collegues-du-premier-13653>

References

1. Genevois, S., Lefer Sauvage, G., Wallian, N.: Questionnaire d'enquête auprès des enseignants "Confinement et continuité pédagogique", Rapport de recherche (2020). Hal-02934483.
2. Mercier, C., Lefer Sauvage, G.: L'école à l'heure de la COVID-19 : des constats de pratiques vers une nouvelle forme scolaire ? Formation et profession (2021, à paraître).
3. Mercier, C.: Formation distanciée et bien-être des étudiants, RIPTU, 17(3), 103-116 (2020a).
4. Mercier, C.: Accompagner les élèves avec autisme dans des espaces co-éducatifs avec un outil numérique de planification. Traduction et Langues, 19(1), 101-116 (2020b).
5. Burton, R., Borrat, S., Charlier, B. et al.: Vers une typologie des dispositifs hybrides de formation en enseignement supérieur. Distances et savoirs, 1(9), 69-96 (2011).
6. Caron, A., Gauthier, M.-C., Provencher, J., Robillard, R.: L'enseignement hybride ou à distance au primaire. Clothilde Bariteau, Québec (2020), [En ligne] <http://iplusinteractif.s3.amazonaws.com/django-assets/media/pedagogie/PAD-Primaire.pdf>, consulté le 21 décembre 2020.
7. Charlier, B., Deschryver, N., Peraya, D. : Apprendre en présence et à distance : une définition des dispositifs hybrides. Distances et savoirs, 4(4), 469-496 (2006).
8. Dumas, P.: Une problématique pluriculturelle de la notion d'hybridation. Distances et médiations des savoirs, 30 (2020), [En ligne] <http://journals.openedition.org/dms/5183>, consulté le 17 juillet 2020.
9. DEPP.: Note d'information n°20-26. [En ligne] <https://www.education.gouv.fr/crise-sanitaire-de-2020-et-continuite-pedagogique-les-eleves-ont-appris-de-maniere-satisfaisante-305214>, consulté le 12 janvier 2021.
10. Felder, J.: Comment prendre en compte l'apprenant autrement que formellement dans la formation ? Distances et médiations des savoirs, 32 (2020), [En ligne] <https://journals.openedition.org/dms/5912#tocto1n3>, consulté le 2 janvier 2021.
11. Fluckiger, C.: La forme universitaire comme analyseur des « effets » de la technologie : perspective critique, Distances et médiations des savoirs, 22 (2020) [En ligne] <http://journals.openedition.org/dms/2329>, consulté le 07 janvier.
12. Fowers, B. J., Mollica, C. O., & Procacci, E. N.: Constitutive and instrumental goal orientations and their relations with eudaimonic and hedonic well-being. The Journal of Positive Psychology, 5(2), 139-153 (2010). <https://doi.org/10.1080/17439761003630045>
13. IFE: Les effets du confinement sur l'activité des enseignants du primaires et secondaire. Rapport de recherche (2020) [En ligne] <http://ife.ens-lyon.fr/ife/recherche/groupe-de-travail/documentation-confinement-et-enseignement/synthese-enseignants>, consulté le 10 décembre 2020.
14. IGEN.: Les usages pédagogiques du numérique au service de la réussite des élèves, n°130. [En ligne] <https://www.education.gouv.fr/les-usages-pedagogiques-du-numerique-en-situation-pandemique-durant-la-period-de-mars-juin-2020-308421>, consulté le 26 janvier 2021.
15. Paquelin, D.: Présence, distance : Vers de nouvelles configurations organisationnelles ? Distances et médiations des savoirs. Distance and Mediation of Knowledge, 2(7), 7 (2014).
16. Tricot, A.: Se former : quelles nouvelles complémentarités entre présence et distance ? Communication orale pour Canopé [En ligne] https://www.youtube.com/watch?v=N0fr1pclbpw&feature=emb_logo (2020), consulté en novembre 2020.
17. Ria, L., Rayou, P.: La forme scolaire en confinement. Enseignants et parents à l'épreuve de l'enseignement à distance. Dossier Pandémie. Formation et Profession 28(4), 1-11 (2020). doi : 10.18162/fp.2020.675

Modèle de méthode pour accompagner le processus de co-conception des jeux épistémiques numériques

Estelle Prior¹

¹ Université de Genève, TECFA, Genève, Suisse
estelle.prior@unige.ch

Résumé. Cet article présente une problématique pour accompagner les phases d'initiation et de prise de décision. Elle prend la forme d'un modèle de méthode pour la conception des jeux épistémiques numériques. Nous nous appuyons sur le concept d'objet frontière pour problématiser la question de la collaboration entre les acteurs. Nous utilisons THEDRE pour proposer une première formalisation de la phase d'initiation. Nous détaillons les expérimentations futures permettant de produire le modèle et d'analyser les prises de décision.

Mots-clés : co-conception, jeu épistémique numérique, initiation d'un projet, prise de décision, objet frontière.

Abstract. This paper presents a problematic to accompany the initiation and decision-making phases. It takes the form of a method model for the design of digital epistemic games. We rely on the concept of boundary object to problematize the question of collaboration between actors. We use THEDRE to propose a first formalization of the initiation phase. We detail the future experimentations to produce the model and to analyze the decision-making.

Keywords: co-design, digital epistemic game, project initiation, decision-making, boundary object.

1 Introduction

Les jeux épistémiques numériques (JEN) suscitent un intérêt croissant [1, 2]. Le terme épistémique “renvoie à la question de la connaissance chez l'apprenant” [3]. L'utilisation de jeux en contexte éducatif présente de nombreux avantages pour des objectifs d'apprentissage très diversifiés [3]. De nombreux travaux de recherche abordent les difficultés liées à leur conception. En effet, les outils adaptés aux différents experts comme les enseignants [4] sont encore rares. Les équipes de conception peuvent aussi travailler en contexte de recherche ce qui rajoute une difficulté supplémentaire car les questions de recherche, l'analyse de données et la production de connaissances sont alors réalisées de manière collaborative entre enseignants et chercheurs, comme proposé par la Recherche orientée par la Conception (RoC) [5, 6]. Les concepteurs peuvent

aussi s'exprimer de manière différente selon leur profession [7], ce qui peut conduire à des incompréhensions. La collaboration peut donc être problématisée comme un partage des connaissances entre ces acteurs en s'appuyant sur le concept d'objet frontière. En outre, comme dans tout processus, des décisions sont prises. Ainsi, notre recherche doctorale aborde les questions suivantes : Comment accompagner les équipes dans l'initiation des projets de co-conception de JEN en contexte de RoC ? Comment faciliter les prises de décisions dans ce processus ? Quels objets frontières peuvent être mobilisés ? Ces questions sont abordées dans le cadre du projet co.LAB qui vise la conception d'une plateforme dédiée à la conception et l'évaluation de JEN pour des équipes pluridisciplinaires travaillant en contexte de recherche. La plateforme comprend des étapes du processus, qui sont affiliées chacune à des "cartes". Ces dernières offrent un lieu de travail collaboratif et des ressources comme des guides ou des références.

2 Etat de l'art

Marfisi [8] et Vermeulen [1] détaillent les étapes préalables au processus de conception : la formulation du besoin client et le détail des objectifs pédagogiques [8], et l'explicitation des objectifs généraux [1]. Ils précisent le besoin d'identifier et décrire le rôle des acteurs impliqués en amont du processus pour faciliter la collaboration. Au contraire de Marfisi, la phase de conception débute pour certains par la formulation d'objectifs pédagogiques [2, 9] et pour d'autres, par la spécification des connaissances initiales nécessaires aux apprenants [1]. De nombreux acteurs sont intégrés au processus (e.g. enseignants, chercheurs, game designer [9]). Si des processus existent avec des phases, des activités et des acteurs, la phase d'initiation n'est pas détaillée.

Quant à la prise de décision, elle apparaît comme un processus complexe à inclure dans la modélisation des activités de conception collaborative. Elle "doit intégrer une approche où rôles, reconnaissances, motivations et inquiétudes ont leur place à côté de processus plus cognitifs" [10]. Rowe et Luecke [11] précisent qu'elle permet d'améliorer l'efficacité et la qualité dans les activités. Pour Allain [12], elle "correspond au fait d'effectuer un choix entre plusieurs modalités d'actions possibles face à un problème." Elle repose sur : 1) l'objet à définir, 2) les informations utiles, 3) les hypothèses de décision qui intègrent "des connaissances et/ou des expériences antérieures".

Nous pensons que les connaissances évoquées par Allain [12] peuvent renvoyer à l'idée d'objet frontière. Ce concept provient des travaux de Star et Griesemer [13]. Cet objet permet un partage de connaissances entre communautés d'acteurs. Ce partage est plus ou moins complexe : syntaxique (simple transfert), sémantique (traduction) et pragmatique (transformation) [14].

3 Contribution

Notre recherche vise à "produire un instrument qui comporte de la connaissance scientifique et un outil. Pour élaborer ces outils, les utilisateurs et leur contexte sont intégrés dans le processus de recherche" [15]. De manière itérative, la connaissance scientifique créée est réifiée en outil manipulable par l'humain. Puis, la connaissance est améliorée

par les données expérimentales produites [15]. Dans notre cas, la connaissance scientifique créée correspond à la modélisation de l'objet frontière et de ses propriétés, et l'outil manipulable, à la maquette de la plateforme et des lieux de travail collaboratifs. Notre travail doctoral porte donc sur la définition des objets frontières dans le cadre de la co-conception des JEN et le suivi de leur évolution dans ce processus, pour améliorer le modèle initial [13, 14]. Le langage THEDRE [15] sera utilisé pour la modélisation.

La Figure 1 présente une modélisation de la phase "Initier le projet" effectuée avec ce langage [15]. Il propose des acteurs, des guides et un ordonnancement des tâches (e.g. T4 s'effectue après T1 et T2, et avec T3 et T5). Les guides proposés dans le modèle proviennent des membres du projet co.LAB et de la méthode THEDRE. De nouveaux guides pourront être ajoutés selon les besoins. Ils permettront d'accompagner les concepteurs dans la réalisation des tâches, comme celles avec prise de décision et d'aboutir au livrable requis. Dans la phase "Initier le projet", les tâches proposées sont effectuées par tous les acteurs (voir Fig. 1). Elles intègrent aussi des éléments identifiés dans l'état de l'art, dont un spécifique à la recherche, et des éléments issus d'observations menées lors du démarrage de la réingénierie du jeu *Programming Game* [9]. Via son guide, la tâche "Lister les objectifs" (T4) invite les concepteurs à répondre à des questions liées aux objectifs généraux du projet (Q1), aux objectifs pédagogiques dans le JEN (Q2), ...

Nous souhaitons identifier les connaissances partagées au cours de cette phase "Initier le projet" et les prises de décision. Nous les formalisons en ajoutant deux symboles (voir Fig. 1), placés sur les tâches selon les observations effectuées. Celui de l'objet frontière est décliné selon ses trois stades : syntaxique (bordeaux), sémantique (gris) et pragmatique (turquoise). Nous faisons l'hypothèse que cette spécification facilitera le travail de co-conception des JEN en contexte de RoC et notamment, lors du lancement du projet, dans le partage des connaissances, mais aussi pour les prises de décisions.

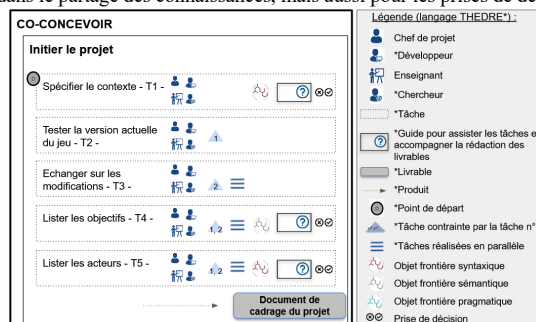


Fig. 1. Modélisation de la phase d'initiation de co-conception d'un JEN selon THEDRE.

4 Conclusion

Ces contributions seront testées lors du processus de co-conception de JEN (e.g. réingénierie du jeu *Programming Game* [9]). Les données seront collectées à travers l'enregistrement vidéo des réunions, le contenu des lieux de travail collaboratifs, la collecte

des traces d'utilisation de la plateforme et par des entretiens et des focus groupes. Ces données seront traitées sous forme d'analyse thématique, appuyée sur une grille d'analyse intégrant les stades de l'objet frontière [14]. Enfin, le modèle de co-conception sera amélioré au travers des expérimentations, puis pourra être déployé sur une plateforme numérique (projet co.LAB). Cette dernière sera manipulée et testée par les utilisateurs.

Nous remercions N. Mandran et E. Sanchez qui encadrent ce travail doctoral ainsi que M. Vermeulen pour son aide. Le projet co.LAB est financé par le Fonds national suisse de la recherche scientifique (PNR 77)

Références

1. Vermeulen, M.: Une approche meta-design des learning games pour développer leur usage, (2018).
2. Marne, B.: Modèles et outils pour la conception de jeux sérieux : une approche meta-design, (2014).
3. Sanchez, E., Romero, M.: Apprendre en jouant. (2020).
4. Oubahssi, L., Piau-Toffolon, C., Loup, G., Sanchez, É.: From Design to Management of Digital Epistemic Games. *IJSG*. 7, 23–46 (2020).
5. 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. *Éducation et didactique*. 9, 73–94 (2015).
6. Monod-Ansaldi, R., Vincent, C., Aldon, G.: Objets frontières et brokering dans les négociations en recherche orientée par la conception. *Education et didactique*. 13, 61–84 (2019).
7. Marne, B., Wisdom, J., Huynh-Kim-Bang, B., Labat, J.-M.: The Six Facets of Serious Game Design: A Methodology Enhanced by Our Design Pattern Library. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., and Hernández-Leo, D. (eds.) *21st Century Learning for 21st Century Skills*. pp. 208–221. Springer, Berlin, Heidelberg (2012).
8. Marfisi-Schottman, I.: Méthodologie, modèles et outils pour la conception de Learning Games, (2012).
9. Plumettaz-Sieber, M., Hulaas, J., Sanchez, E., Jaccard, D.: Co-design of a serious game for computing education. Presented at the The Gamification & Serious Games Symposium 2019 July (2019).
10. Cahour, B.: Décalages socio-cognitifs en réunions de conception participative. *Le travail humain*. Vol. 65, 315–337 (2002).
11. Rowe, A., Luecke, R.: *L'essentiel pour bien décider*. Éditions Prisma (2017).
12. Allain, P.: La prise de décision : aspects théoriques, neuro-anatomie et évaluation. *Revue de neuropsychologie*. Volume 5, 69–81 (2013).
13. Star, S.L., Griesemer, J.R.: Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Soc Stud Sci*. 19, 387–420 (1989).
14. Carlile, P.R.: Transferring, Translating, and Transforming: An Integrative Framework for Managing Knowledge Across Boundaries. *Organization Science*. 15, 555–568 (2004).
15. Mandran, N.: *Traceable Human Experiment Design Research: Theoretical Model and Practical Guide*. John Wiley & Sons, Inc., Hoboken, NJ, USA (2018).

***Geome* et son Compagnon Digital : un dispositif permettant la mise en place de visites muséales ludiques**

Eric Sanchez¹, Catherine Bonnat¹, Gil Oliveira¹, Matthieu Vallat¹, Kevin Manixab²,
Didier Crausaz², Nicolas Kramar³, Sylvia Mueller⁴, Omar Abou Khaled², Elena Mugellini²

¹ Université de Genève, TECFA, 1211 Genève 4, Suisse
eric.sanchez@unige.ch; catherine.bonnat@unige.ch;
gil.oliveira@unige.ch; matthieu.vallat@unifr.ch;

² Université de Fribourg, HumanTech (HELA), 1700 Fribourg, Suisse
kevin.manixab@edu.hefr.ch; didier.crausaz@edu.hefr.ch;
omar.aboukhaled@hefr.ch elena.mugellini@hefr.ch

³ Musée de la Nature, 1950 Sion, Suisse
nicolas.kramar@admin.vs.ch

⁴ Haute Ecole Pédagogique du Valais, 1890 Saint Maurice, Suisse
sylvia.mueller@hepvs.ch

Résumé. Pour cette démonstration nous proposons le jeu *Geome*, qui traite du rapport de l'Homme à la nature, et qui se joue dans un environnement mixte (numérique et tangible), au Musée de la Nature dans le Valais Suisse dans un contexte de visite scolaire (12 – 15 ans). Afin d'aider le maître du jeu à orchestrer la partie, nous présentons également *DigitComp*, un compagnon digital dédié au paramétrage et au suivi de l'activité des joueurs à partir de traces multimodales.

Mots-clés : Jeu, musée, maître du jeu, assistant, traces multimodales.

Abstract. For this demonstration we propose *Geome*, a game dealing with the relationships between humans and nature, and which is played in a mixed environment (digital and tangible), at the Nature Museum in the Swiss Valais, within a school visit context (12 - 15 years old). In order to help the game master to orchestrate, we also present *DigitComp*, a digital companion dedicated to the configuration and monitoring of the players' activity using multimodal learning analytics.

Keywords: Game, museum, game master, assistant, multimodal.

1 Introduction

Les travaux que nous menons portent sur la conception de jeux et l'analyse de leur usage en contexte muséal. Le jeu est alors considéré comme une expérience qui permet,

lorsqu'elle est intégrée dans un scénario pédagogique, d'aborder des objectifs d'apprentissage spécifiques. Nos travaux concernent plus particulièrement deux volets. Un premier volet porte sur la conception de l'expérience ludique du joueur (*play design*) [1]. Ce volet est abordé dans le cadre du projet PLAY dont les objectifs sont de comprendre comment l'expérience ludique peut amener des élèves de l'enseignement secondaire à s'interroger sur leur rapport à la nature d'une part, et au savoir d'autre part. Le second volet concerne l'orchestration de telles situations par le maître du jeu. Il s'agit de comprendre quelles sont les informations dont il doit disposer pour orchestrer le jeu, comment collecter ces informations dans un environnement mixte qui combine des objets tangibles et numériques. Il s'agit également de lui fournir les outils nécessaires pour cette orchestration. Ce second volet est abordé par le projet *DigitComp*.

Les deux projets sont menés par une équipe pluridisciplinaire qui comprend des chercheurs en éducation, des chercheurs en informatique et développeurs, un graphiste ainsi que des praticiens enseignants et médiateurs de musée. Le travail qui est conduit relève de la recherche orientée par la conception [2] et nous présentons ici les résultats des deux développements réalisés (le jeu *Geome* et son compagnon). Ces deux dispositifs ont fait l'objet d'expérimentations en contexte écologique au printemps 2021.

2 *Geome*, un jeu d'apprentissage en contexte de visite scolaire au Musée

Geome (Fig. 1) a été développé dans le cadre d'un projet financé par le Fonds national de la recherche scientifique (Suisse) par un consortium qui comprend l'Université de Fribourg, le Musée de la Nature à Sion et la Haute école pédagogique du Valais. En jouant à *Geome* dans le cadre de visites scolaires dans un musée, les élèves de l'enseignement secondaire sont conduits à s'interroger sur leur rapport à la nature. Les objectifs pédagogiques concernent la compréhension de l'anthropocène, une période géologique où l'homme ne se considère plus comme faisant partie de cette nature mais la surexploite avec des conséquences sur la biodiversité et des changements climatiques. Un autre volet du jeu concerne le rapport au savoir des élèves concernés. Il s'agit de les amener à s'interroger sur la nature des savoirs et leur mode de production.



Fig. 1. *Geome*, un jeu qui questionne le rapport de l'Homme à la nature

Le développement du jeu s'appuie sur un modèle qui considère le jeu en tant que situation intégrée dans un scénario pédagogique plutôt que comme un artefact [3]. Il s'appuie également sur un modèle qui permet de penser l'intégration des contenus d'apprentissage dans le jeu sous l'angle de la métaphorisation du domaine qui fait l'objet d'un apprentissage [4].

Les joueurs sont amenés, en équipe, et au sein du musée, à incarner un professionnel de la nature qui vit dans une vallée alpine. Dans la première partie du jeu [5], le personnage qu'ils incarnent va se retrouver bloqué en raison des conditions climatiques hivernales. Il est ainsi amené à exploiter son environnement pour survivre. Pour les joueurs, cela consiste à capturer, tuer ou domestiquer les animaux exposés dans le musée. Ce faisant, ils perdent collectivement en épuisant les ressources de « l'arbre de vie ». Dans la seconde partie du jeu, libéré des contraintes de l'hiver, le personnage peut reprendre son travail. Il est appelé, en tant qu'expert, pour résoudre des enquêtes relatives au milieu naturel, basées sur une *fake news*, rumeur ou polémique. La muséographie apporte des indices qui permettent d'élaborer des réponses (Fig. 1). Les équipes gagnent en collaborant au cours de l'enquête. Le joueur adopte ainsi successivement deux postures : celle du prédateur qui exploite son milieu naturel, puis celle de l'enquêteur qui met à jour les relations interspécifiques complexes au sein de l'écosystème. Au cours du débriefing qui succède à la phase de jeu, les élèves sont amenés à se questionner sur le rapport de l'homme à la nature, mais aussi à analyser leur rapport aux médias et à l'information.

3 *DigitComp*, un compagnon digital pour l'orchestration du jeu

DigitComp est un compagnon digital dédié à l'orchestration d'un jeu en réalité mixte (tangible et numérique). Intégré au jeu *Geome*, il a été développé dans le cadre d'un Fonds National Suisse de 12 mois, en collaboration avec les informaticiens de la Haute École d'Ingénierie et d'Architecture de Fribourg.

Il comprend une interface maître du jeu qui permet de configurer la partie et de modifier les variables du jeu définies par défaut. Le maître du jeu peut par exemple, définir/modifier les objectifs de la partie, choisir sa durée, attribuer un type d'énigme par groupe selon leur niveau, configurer des feedbacks automatiques. De plus, au cours de la partie le maître du jeu dispose d'un module de chat de manière à accompagner l'ensemble des joueurs ou groupes de joueurs, en leur proposant, par exemple, un indice pour les aider dans la résolution de leur enquête. Les joueurs, qui évoluent en équipe, sont avertis du message par l'intermédiaire d'une montre connectée, et prennent connaissance du contenu du message sur la tablette.

Afin d'accompagner au mieux les joueurs, le maître du jeu dispose d'une interface de suivi de l'activité en synchrone et, à l'issue de la partie, d'un tableau de bord sur lequel la phase de débriefing du jeu peut s'appuyer [6]. Cependant, la question du type d'information dont le maître du jeu doit disposer se pose, notamment dans un contexte de jeu en réalité mixte. *DigitComp* propose pour cela une fonction qui consiste à observer l'activité des joueurs. Cette fonction repose sur une récolte de traces multimodales [7] stockées au format JSON sur une base de données créée spécifiquement pour le jeu. Ces traces relèvent à la fois des variables du jeu (par exemple l'évolution de l'état de l'arbre de vie), mais aussi les actions réalisées par les joueurs sur la tablette, en interaction avec les éléments du musée (ex. scan d'animaux, Fig.1). Des données physiologiques (ex. fréquence cardiaque) sont également recueillies par l'intermédiaire de la montre connectée, et stockées sur la même base de données. Ces traces multimodales

sont à la fois à disposition du chercheur (données brutes), mais aussi pour certaines d'entre elles à disposition du maître du jeu sur une interface dédiée au suivi et au bilan de l'activité [8]. Le maître du jeu peut accéder en mode synchrone aux traces d'activité d'un seul joueur, d'un groupe de joueurs ou de l'ensemble des groupes. Ces traces sont le résultat d'un traitement statistique élémentaire (ex. nature et nombre d'actions réalisées de type scan). Enfin, les joueurs disposent d'un tableau de bord qui traduit le bilan de la partie (pour leur groupe).

4 Conclusion

Le scénario pédagogique incluant le jeu *Geome* et son compagnon digital ont fait l'objet d'un test pilote au musée en décembre 2020 (enseignants, chercheurs, médiateurs du musée et informaticiens), avant les expérimentations *in situ* avec des élèves du cycle d'orientation des écoles du Valais (CH) en février 2021. Les analyses en cours, portent notamment sur la validité des choix de conception, mais aussi, selon les orientations scientifiques du projet de recherche, sur la posture épistémologique des élèves et leur rapport au savoir. Le processus itératif que nous utilisons, permettra la révision des prototypes et modèles théoriques pour chacune des itérations.

References

1. Genvo, S. : Du game design au play design : éthos et médiation ludique, in Colloque Entre le jeu et le joueur : écarts et médiations. Liège (2018).
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 », *Éducation et didactique*, 9(2), 73-94 (2015).
3. Sanchez, E., Emin-Martinez, V., Mandran, N. : Jeu-game, jeu-play, vers une modélisation du jeu. Une étude empirique à partir des traces numériques d'interaction du jeu Tamago-cours, *STICEF*, 22, 9-45 (2015).
4. Bonnat, C., Sanchez, E., Paukovic, E., Kramar, N.: Didactic transposition and learning game design. Proposal of a model integrating ludicization, and test in a school visit context in a museum. In J. K. Almqvist, K (Ed.), *Didactics in a Changing World. European Perspectives on Teaching, Learning and the Curriculum: EERA Book Series* (sous-presse).
5. Bonnat, C., Oliveira, G., Sanchez, E. : *Geome*, un jeu pour comprendre l'Anthropocène lors de visites scolaires au musée. *Revista de la Asociacion para la Ensenanza de las Ciencias de la Terra*, 28(1), 89 – 98 (2020).
6. Plumettaz-Sieber, M., Bonnat, C., Sanchez, E.: Debriefing and Knowledge Processing. An Empirical Study about Game-Based learning for Computer Education. In A. Liapis, G. Yannakakis, M. Gentile, N. Ninaus (Eds.), *GALA 2019* (Vol. 11899, pp. 3241). Athens: Springer International Publishing (2019).
7. Blikstein, P.: Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 102 (2013).
8. Prieto, L-P, Sharma, K., Kidzinski, L., Dillenbourg, P.: Orchestration Load Indicators and Patterns: In-the-Wild Studies Using Mobile Eye-Tracking. *TLT*, 11(2), 216-229 (2018).

Index des auteurs

- Abou Khaled Omar, 281, 300, 405
Allègre Olivier, 33
Alvarez Lionel, 274, 318
Andriamiseza Rialy, 46
Angelini Leonardo, 281
Arseneau Rosianne, 288
- Badier Anaëlle, 294
Bailly Gérard, 70, 118
Bara Florence, 348
Baron Pierre, 372
Benabbou Azzeddine, 202
Bennacer Ibtissem, 377
Bensebaa Taher, 202
Blanchard Christine, 393
Bonnat Catherine, 300, 381, 405
Bouchet François, 250, 262, 312, 354, 360
Broisin Julien, 46, 166, 226
- Caron Pierre-André, 385
Carron Thibault, 250, 262, 360
Casanave Elfi, 385
Chabanal Damien, 393
Cleuziou Guillaume, 58
Coen Pierre-François, 389
Crausaz Didier, 405
Crespin Benoit, 154
- Daniel Morgane, 250, 262
De Vito Woods Mariateresa, 281
Diard Julien, 118
Durand Edith, 393
d'Ham Cédric, 190
- El Mawas Nour, 214, 306
Ensor Simon, 393
Exbrayat Matthieu, 58
- Favre Alexandre, 281
Fleury Anthony, 94
Flouvat Frédéric, 58
Foucher Anne-Laure, 393
- Gajewski Sebastian, 306
Genevois Sylvain, 336
- Geoffre Thierry, 274, 288, 318
Girault Isabelle, 190
Godde Erika, 70
Group Fluence, 118
Grugeon-Allys Brigitte, 82
Guin Nathalie, 166, 226
- Harrak Fatima, 312
Heutte Jean, 214, 306
Hoffmann Christian, 190
Humeau Jérémie, 94
- Iksal Sebastien, 377
Iván Ramírez Luelmo Sergio, 214
- Jaouadi Mariem, 366
Jeannin Loïse, 393
Jolivet Sébastien, 82
Jouffrais Christophe, 348
- Kaminski Gwenaël, 348
Karoui Aous, 274, 318
Kerneis Jacques, 336
Kramar Nicolas, 405
- Laure Piat-Marchand Anne, 70
Lebas Lidia, 393
Lebis Alexis, 94, 324, 330
Lefer Sauvage Gaëlle, 106, 336, 397
Lefevre Marie, 166, 226, 294
Lefort Mathieu, 294
Lenne Dominique, 178
Lesnes-Cuisiniez Elann, 82
Line Bosse Marie, 70
Loiseau Mathieu, 118
Lotin Paul, 393
Lourdeaux Domitile, 178, 202
Luengo Vanda, 33, 82, 354
- Magnin François, 281
Mandin Sonia, 118
Mandran Nadine, 130, 190, 324, 330, 366
Manixab Kevin, 405
Mansouri Khalifa, 238
Marne Bertrand, 342

Marzin Patricia, 190
Mascle Carolane, 348
Mercier Cendrine, 397
Meyer Svetlana, 118
Michel Christine, 142, 166
Morais Canellas Camila, 354
Morard Simon, 381
Mueller Sylvia, 405
Mugellini Elena, 281, 300, 405
Muratet Mathieu, 82, 342

Ngu Leubou Richard, 154

Oliveira Gil, 381, 405
Oliver-Quelennec Katia, 360

Parmentier Jean-François, 46
Paukovics Elsa, 381
Payre-Ficout Coralie, 118
Pelissero-Witoslawski Luca, 178
Pierrot Laëtitia, 142, 166
Pinçon Claire, 360
Planche Maelle, 190
Poirier Franck, 238
Prior Estelle, 130, 330, 366, 401

Quanquin Véronique, 393

R. Wigham Ciara, 393
Raffaillac Thibault, 372
Rahmouni Nesrine, 202
Ramalho Mario, 274, 318
Reiter Benjamin, 82
Robert Julien, 58
Roche Laurent, 372
Runtz-Christan Edmée, 389

Sablayrolles Louis, 226
Safsouf Yassine, 238
Sanchez Eric, 381
Sanchez Eric, 300, 366, 405
Sehaba Karim, 342
Sergent Thomas, 250, 262
Silvestre Franck, 46

Thion Romuald, 58
Trestini Marc, 154

Valdois Sylviane, 118
Vallat Matthieu, 300, 405
Venant Remi, 377
Venant Rémi, 166
Vermeulen Mathieu, 94, 130, 324, 330

Wajemane Claire, 190
Wallian Nathalie, 336

Yessad Amel, 33, 82

Zaher Ahmed, 118

