



HAL
open science

Caractéristiques personnelles et explications dans les EIAH : une étude exploratoire dans l'enseignement de l'informatique

Esther Félix, Franck Amadiou, Julien Broisin

► To cite this version:

Esther Félix, Franck Amadiou, Julien Broisin. Caractéristiques personnelles et explications dans les EIAH : une étude exploratoire dans l'enseignement de l'informatique. STICEF (Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation), 2023, 10.23709/sticef.30.1.2 . hal-04652899

HAL Id: hal-04652899

<https://hal.science/hal-04652899>

Submitted on 18 Jul 2024

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.

Caractéristiques personnelles et explications dans les EIAH : une étude exploratoire dans l'enseignement de l'informatique

Personal Characteristics and Explanations in TEL Environments: an Exploratory Study in Computer Science Education

Esther FÉLIX¹; Franck AMADIEU²; Julien BROISIN¹

¹IRIT - Institut de Recherche en Informatique de Toulouse, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France

²CLLE - Cognition, Langues, Langage, Ergonomie, CNRS, Toulouse, France

Résumé. L'intelligence artificielle explicable dans l'éducation est devenue une préoccupation majeure. Dans cet article, nous introduisons des explications dans un système à base de compétences, dans le contexte de l'apprentissage de l'informatique. Notre objectif global est d'accroître la compréhension par les apprenants des niveaux de maîtrise de leurs compétences qui sont calculés automatiquement, et ainsi augmenter leur confiance dans le système. En particulier, nous étudions quelles caractéristiques personnelles des étudiants affectent la consultation et la compréhension des explications, et quels sont les impacts de ces explications sur leur perception et leur comportement au sein du système. Notre étude porte sur 98 étudiants de première année en informatique dans l'enseignement supérieur. Les analyses quantitatives montrent que les étudiants qui ont une faible perception de leur capacité à réussir sont moins susceptibles d'accéder aux explications et de les comprendre, et que les étudiants peu engagés et peu performants sont moins susceptibles de comprendre les explications. Les analyses qualitatives, quant à elles, nous aident à identifier certains besoins des apprenants en termes de contenu, forme et temporalité des explications.

Mots-clés : explicabilité, apprentissage de l'informatique, eiah, intelligence artificielle

Abstract. *Explainable artificial intelligence in education recently became a major concern. In this paper, we introduce explanations in a competency profile, in the context of computer education. Our overall objective is to increase learners' understanding and trust in the system. In particular, we are interested in studying which students' personal characteristics affect the consultation of explanations, and what are the impacts of these explanations on their perception of and their behavior within the system. Our study involves 98 first-year students from a computer science degree. Our quantitative analyses show that students with a low perception of their ability to succeed are less likely to access and understand explanations, and also that low-engaged, low-performing students are less likely to understand explanations. Qualitative analyses, meanwhile, help us to identify certain learner needs in terms of content, form and timing of explanations.*

Keywords: *explainability, computer education, tel, artificial intelligence*

1. INTRODUCTION

La recherche sur l'intelligence artificielle explicable, ou *xAI* (*eXplainable Artificial Intelligence*), vise à rendre les systèmes fondés sur l'intelligence artificielle (IA) plus transparents et compréhensibles par les humains. Jusqu'à présent, ce domaine était principalement destiné aux experts en IA, mais des recherches récentes montrent la nécessité d'adresser des explications à des utilisateurs non experts afin d'améliorer leur compréhension et ainsi renforcer leur confiance dans les systèmes qu'ils utilisent (Kraus *et al.*, 2020).

Les environnements informatiques pour l'apprentissage humain (EIAH) sont aujourd'hui souvent soutenus par des méthodes d'IA pour interagir avec les apprenants ou les enseignants. Pour améliorer la compréhension et la transparence de ces mécanismes intelligents pour les utilisateurs, et ainsi favoriser l'adoption des EIAH munis d'IA, une solution consiste à ajouter des explications (Conati *et al.*, 2018). Pour être efficaces, les explications fournies doivent être adaptées au niveau d'expertise de l'utilisateur et plus généralement aux caractéristiques individuelles qui lui sont propres (Tomsett *et al.*, 2018). Notre objectif à long terme vise donc à personnaliser les explications fournies aux utilisateurs dans les EIAH. Toutefois, pour atteindre ce but, un objectif intermédiaire qui fait l'objet de cet article est d'identifier les caractéristiques individuelles susceptibles d'avoir un impact sur la consultation, la perception et la compréhension des explications, ainsi que les types d'explications susceptibles d'être les plus efficaces (Khosravi *et al.*, 2022). Des caractéristiques ont déjà été identifiées dans des domaines tels que la recommandation musicale (Millecamp *et al.*, 2019), mais à notre connaissance, elles n'ont pas été identifiées dans le domaine de l'éducation (Khosravi *et al.*, 2022). Dans cet article, nous visons donc à répondre aux questions de recherche suivantes :

- **QR1** : Quelles caractéristiques personnelles sont corrélées avec la consultation, la compréhension et la perception de l'utilité des explications ?
- **QR2** : Comment les caractéristiques personnelles des apprenants influencent-elles leur perception du système (QR2.1), et dans quelle mesure la consultation, la compréhension et la perception de l'utilité des explications affectent-elles cette perception du système (QR2.2) ?

Pour répondre à ces questions, nous menons une étude exploratoire en contexte d'apprentissage réel à partir d'explications relatives au profil de compétences des apprenants. Ces explications sont intégrées dans une plateforme d'apprentissage de l'informatique. Nous étudions l'influence des explications selon six caractéristiques personnelles d'apprenants de première année de Bachelor Universitaire de Technologie (BUT) informatique.

La section suivante présente un état de l'art des notions d'IA explicable pour l'éducation et de personnalisation des explications dans le contexte des EIAH. La section 3 décrit l'expérimentation que nous avons menée et la section 4 détaille les analyses des données quantitatives et qualitatives. Les résultats de ces analyses sont présentés dans la section 5 et discutés dans la section 6. Les conclusions de l'étude et les orientations futures de nos travaux terminent cet article.

2. ÉTAT DE L'ART

2.1. INTELLIGENCE ARTIFICIELLE EXPLICABLE DANS LE DOMAINE DE L'ÉDUCATION

L'explicabilité des modèles d'IA est un domaine de recherche qui a pris de l'ampleur ces dernières années en raison de la nécessité de comprendre comment les algorithmes prennent leurs décisions. L'objectif de l'IA explicable est de rendre les systèmes plus transparents et

d'accroître la confiance des utilisateurs dans leur fonctionnement et dans les résultats qu'ils produisent (Kraus *et al.*, 2020).

L'idée de fournir des explications à des utilisateurs non experts n'a commencé à être explorée que récemment dans le monde de la recherche en IA, où de telles explications étaient initialement conçues par et pour des experts en IA cherchant à mieux comprendre leurs algorithmes afin d'en améliorer le fonctionnement. En 2019, Arya *et al.* ont conçu une taxonomie des différentes techniques d'explicabilité en IA pour répondre à la nécessité d'adapter les explications à un public d'utilisateurs plus large que les experts, avec l'exemple d'un algorithme d'attribution de crédit pour une banque devant être expliqué de manière transparente aux clients. Les méthodes d'explicabilité peuvent être classées en différentes approches (Barredo Arrieta *et al.*, 2020 ; Gilpin *et al.*, 2018) selon qu'elles expliquent l'ensemble du modèle lui-même (approches globales) ou un résultat individuel (approches locales), et selon que le modèle s'explique de lui-même ou que ses résultats doivent être expliqués par la suite avec des explications post-hoc (Arya *et al.*, 2019 ; Khosravi *et al.*, 2022). Ribera distingue trois rôles d'utilisateurs qui nécessitent différents types d'explications : les développeurs et chercheurs en IA, les experts du domaine, et les utilisateurs profanes (Ribera et Lapedriza, 2019). Les experts et chercheurs en IA sont susceptibles d'avoir besoin d'explications concernant le modèle lui-même dans le but de le tester et l'améliorer, avec par exemple des modélisations intermédiaires (*proxy models*) qui se comportent de manière similaire au modèle original mais qui sont plus simples à expliquer (Gilpin *et al.*, 2018). Les méthodes visant à décomposer les réseaux de neurones en arbres de décisions (Sato et Tsukimoto, 2001) sont des exemples d'explications de ce type. Pour les autres rôles d'utilisateurs non spécialistes en IA, en particulier les utilisateurs profanes, des explications locales et post-hoc sont préconisées, qui n'expliquent pas le modèle entier mais plutôt comment un résultat particulier a été obtenu (Keane *et al.*, 2021). Ces explications sont à favoriser pour ces types d'utilisateurs car elles sont plus proches du mode de fonctionnement humain (Lipton, 2018) et contiennent une quantité plus restreinte d'informations, ce qui facilite leur compréhension (Wachter *et al.*, 2017). Les explications post-hoc peuvent être divisées en plusieurs familles (Barredo Arrieta *et al.*, 2020) qui comprennent les explications textuelles (Bennetot *et al.*, 2019), les explications visuelles (Krause *et al.*, 2016 ; Lipton, 2018), les explications fondées sur l'importance des variables ou *feature relevance explanations* (Lundberg et Lee, 2017), ou encore les explications par l'exemple, comme les explications contrefactuelles (Wachter *et al.*, 2017).

Il est important de noter que le terme « explicabilité » en IA fait généralement référence aux techniques d'explicabilité qui s'appliquent à des modèles de type « boîte noire » (algorithmes de *machine learning* et *deep learning*). Cependant, selon Rizzo *et al.* (2023), l'explicabilité est plutôt un gradient allant des modèles complètement boîte noire aux modèles complètement transparents, sans seuils précis. Ils soutiennent que la différence entre ces modèles réside principalement dans le niveau de difficulté à concevoir des explications les concernant, et que la clé pour déployer des modèles explicables est d'intégrer les humains dans la boucle de conception de l'explicabilité. Leur notion de difficulté est définie par la complexité interne du modèle, qui est liée à la charge cognitive que l'utilisateur peut supporter. Par conséquent, les techniques d'explicabilité post-hoc utilisées pour les modèles à boîte noire peuvent également servir à expliquer des modèles considérés comme transparents, et le type d'explication requis dépend du niveau d'expertise de l'utilisateur et de ses objectifs. De plus, bien que la notion d'explicabilité en tant qu'*explainable AI* soit récente, des tentatives d'explication des systèmes d'IA basés sur des règles à leurs utilisateurs ont été faites depuis le développement de ce type de système. Par opposition à l'IA numérique, ces systèmes mettent en œuvre une IA symbolique, c'est-à-dire des algorithmes qui visent à re-

produire le raisonnement humain. Pour ces systèmes également, il y a toujours eu une volonté de fournir aux utilisateurs des informations qui les aident à comprendre comment le système fonctionne et pourquoi il produit un résultat donné, afin d'améliorer la confiance dans le système. Toutefois, dans ces cas, le terme « explicabilité » n'est pas explicitement utilisé. Les chercheurs soulignent plutôt la nécessité pour les systèmes d'IA d'avoir la « capacité d'expliquer leur (ligne de) raisonnement » (Clancey, 1983a; Leake & Mcsherry, 2005). Lacave et Diez (2004, p.134) ont proposé une définition de ce qu'une explication devrait être : selon eux, « expliquer consiste à exposer quelque chose de manière **compréhensible** pour le destinataire de l'explication, afin d'améliorer ses connaissances sur l'objet de l'explication, et **satisfaisante**, au sens où l'explication répond aux attentes du destinataire ». Lacave et Diez donnent des recommandations sur ce que les explications doivent inclure dans les systèmes experts heuristiques (par exemple, l'adaptation à l'utilisateur, la compréhensibilité et l'accent mis sur la causalité) et ont fait une revue de ces systèmes capables de fournir ce type d'explications. En particulier, un système d'explications a été intégré à un ITS (*Intelligent Tutoring System*) sous le nom de GUIDON (Clancey, 1983b). Ce tuteur utilisait le dialogue avec les étudiants pour les aider à résoudre leurs problèmes, et était capable de justifier ses règles auprès des étudiants en fonction de leurs connaissances et de leurs intérêts.

Les systèmes basés sur des règles sont donc de bons candidats pour fournir des explications sur leur raisonnement (Hayes-Roth, 1985). Ainsi, dans notre étude, **nous nous intéressons en particulier à l'explicabilité au sens « capacité à expliquer le raisonnement » d'un système à des utilisateurs non experts**, plutôt qu'au sens « *explainable AI* » couramment utilisé pour évoquer les dernières techniques permettant de comprendre le fonctionnement des algorithmes de *machine learning* « boîtes noires » les plus modernes. Nous avons choisi de générer les explications à l'aide de méthodes d'explicabilité post-hoc : les explications fournies sont **locales**, car les travaux antérieurs montrent qu'elles sont bien adaptées aux utilisateurs novices (Lipton, 2018); elles s'appuient également sur des **exemples** (*example-based*), et elles exposent aux apprenants pourquoi et comment le système a produit le résultat proposé.

Appliqué à l'éducation, le domaine de l'IA explicable vise ainsi à améliorer la transparence des systèmes d'IA utilisés dans les contextes éducatifs afin de renforcer la confiance des utilisateurs (Khosravi *et al.*, 2022). Il s'agit d'expliquer le fonctionnement interne de ces systèmes et leurs processus décisionnels aux utilisateurs finaux, tels que les étudiants, les enseignants ou les parents. Les recherches proposant des explications sur les systèmes d'IA dans le domaine de l'éducation sont encore peu nombreuses. Par exemple, Conati *et al.* (2021) ont proposé des conseils accompagnés d'explications dans un système de tutorat intelligent (*Intelligent Tutoring System*, ITS) en mettant l'accent sur *pourquoi* une suggestion est utile et *comment* l'IA prend la décision de faire cette suggestion. L'application de l'explicabilité au domaine de l'éducation se distingue par le fait que son objectif est de fournir des explications *actionnables* (Khosravi *et al.*, 2022). La notion d'actionnabilité est définie comme se rapportant à des « données permettant d'établir une procédure corrective, ou une boucle de feedback, pour un ensemble d'actions » (Jørnø et Gynther, 2018). Les explications dans le domaine de l'éducation doivent donc, en plus d'expliquer comment certains résultats sont calculés, être associées à une forme de feedback incitant à s'engager dans une réflexion ou une action (Winne, 2021).

Ainsi, certains travaux en EIAH cherchent à fournir des explications actionnables menant par exemple à des recommandations explicables de ressources d'apprentissage (Abdi *et al.*, 2020; Barria-Pineda *et al.*, 2021; Palominos *et al.*, 2019), à l'utilisation optimale du temps des enseignants pour examiner le travail des étudiants (Darvishi *et al.*, 2020), à l'apport d'un

soutien aux apprenants qui en ont le plus besoin (Khosravi *et al.*, 2021), à des recommandations d'actions pédagogiques à effectuer en fonction de prédictions fournies (Pardo *et al.*, 2016), à l'amélioration de l'auto-régulation des apprenants (Afzaal *et al.*, 2021), ou encore à une meilleure compréhension du feedback donné sur du contenu de cours (Ali *et al.*, 2012). Les notions d'explication actionnable et de feedback sont donc liées et il apparaît important de tenir compte des recherches portant sur les feedback dans l'éducation, en particulier dans les EIAH, afin d'intégrer efficacement des explications actionnables dans ces environnements.

2.2. FEEDBACK ET BESOINS DE PERSONNALISATION DANS LES EIAH

En 1995, Winne et Butler définissent le feedback comme une « information qui permet à un apprenant de confirmer, d'ajouter, d'écraser, d'ajuster ou de restructurer de l'information en mémoire, qu'il s'agisse de connaissances du domaine, de connaissances métacognitives, de croyances sur soi et sur les tâches, ou de stratégies cognitives » (Butler et Winne, 1995, p.275). Hattie et Timperley (2007), dans leur revue de littérature devenue une référence sur la notion de feedback, proposent un modèle du feedback et évaluent son efficacité selon différents aspects. Ils définissent le feedback comme une information fournie par un agent (par exemple un professeur, un parent, un pair) concernant la performance ou la compréhension d'un individu. L'objectif du feedback est de réduire l'écart entre, d'une part, la compréhension et la performance d'un apprenant atteintes à un moment donné (« là où il se trouve ») et, d'autre part, les objectifs fixés à atteindre (« là où il veut aller »). Cet objectif n'est cependant pas évident à atteindre, car le feedback peut être inadapté, rejeté ou mal interprété et donc ne pas avoir l'impact désiré. D'après Hattie et Timperley (2007), le feedback, pour être efficace, doit ainsi répondre aux questions suivantes : *Quels sont mes objectifs ?* (« *feed-up* », il s'agit de comparer la situation actuelle à l'objectif visé) ; *Quel est mon avancement actuel ?* (« *feed-back* », c'est l'analyse de la progression qui a été effectuée) ; *Quelle est la prochaine étape ?* (« *feed-forward* » : la réponse à cette question doit fournir une explication de l'objectif ciblé à partir de l'état d'avancement actuel, en donnant des informations sur ce qui est compris et sur ce qui ne l'est pas). C'est ce qui mène à une adaptation de l'apprentissage, à l'élaboration de nouvelles stratégies, et aide au processus d'auto-régulation. De plus, le feedback peut être différencié selon plusieurs niveaux : le feedback au niveau de la tâche, qui permet d'indiquer si une tâche précise a été bien accomplie ou comprise ; le feedback au niveau du processus, qui donne un retour sur les stratégies de résolution qui ont été mises en place par l'apprenant ; le feedback au niveau de l'auto-régulation, qui est focalisé sur la régulation par l'apprenant de ses stratégies ; et enfin le feedback sur la personne, qui consiste en un retour sur l'apprenant mais pas sur la tâche en elle-même.

Narciss *et al.* ont proposé un cadre conceptuel qui apporte un autre degré de précision en distinguant trois catégories principales de feedback dans les EIAH (Narciss, 2008 ; Narciss *et al.*, 2002) : la connaissance des résultats (KR, *Knowledge of Result*), la connaissance de la réponse correcte (KCR, *Knowledge of Correct Response*) et le feedback élaboré (EF, *Elaborated Feedback*). Keuning *et al.* (2018) ont adapté ce cadre et détaillé différents types de feedback élaboré dans le domaine de l'enseignement de l'informatique : la connaissance des erreurs (KM, *Knowledge of Mistakes*) qui concerne les échecs des tests, les erreurs de compilation ou les problèmes de performance ; la connaissance de la manière de procéder (KH, *Knowledge on How to proceed*) pour les conseils liés aux étapes de traitement des tâches ; et la connaissance de la métacognition (KMC) pour le feedback sur les stratégies métacognitives. La plupart des outils de feedback pour l'enseignement de l'informatique utilisent le feedback de type KM. Nous avons également étudié l'effet du feedback métacognitif sur les changements de stratégies des apprenants dans une plateforme d'apprentissage de l'informa-

tique, mais n'avons pas trouvé d'effet significatif (Félix *et al.*, 2022) ; nous avons néanmoins observé que les programmeurs novices sont à la recherche d'un feedback au niveau de la tâche qui soit directement et plus facilement exploitable pour eux. Dans le présent papier, notre étude propose aux étudiants des explications qui peuvent être considérées comme une forme de feedback (Khosravi *et al.*, 2022). Ce feedback fourni est un feedback élaboré de type « connaissance de la manière de procéder » selon le cadre de Narciss *et al.* (Narciss, 2008 ; Narciss *et al.*, 2002), et au niveau du processus d'après la classification de Hattie et Timperley (2007). Comme conseillé par Hattie et Timperley, notre feedback comporte des informations concernant le niveau de l'apprenant sur les compétences à acquérir (« *Quel est mon avancement actuel ?* », information utile à l'auto-évaluation), donne une explication sur comment ce niveau de compétence a été calculé de manière à les aider à comprendre les attendus du module (« *Quels sont mes objectifs ?* », information utile à l'auto-régulation), et donne un conseil pour améliorer leur manière de travailler (« *Quelle est la prochaine étape ?* », information également utile à l'auto-régulation).

De plus, une récente étude souligne « le besoin de développer des explications exploitables ("actionable") et personnalisées » comme l'un des cinq domaines majeurs pour favoriser l'adoption des systèmes éducatifs munis d'IA (Khosravi *et al.*, 2022, p.19). Plus précisément, les auteurs soulignent la nécessité d'identifier : « les types d'explications dont les différents utilisateurs finaux ont besoin » ; « comment les explications devraient être fournies » ; et « si ces facteurs peuvent dépendre de différences individuelles » telles que les capacités, les traits de personnalités, ou encore les états cognitifs et affectifs. Kouki *et al.* (2019) soulignent également la nécessité de prendre en compte les différences individuelles dans la recherche visant à identifier les types d'explications requises par divers utilisateurs et la manière dont les explications doivent être fournies pour être acceptables. Au-delà du niveau d'expertise et des objectifs, des recherches récentes montrent que des explications adaptées à d'autres caractéristiques personnelles d'un utilisateur (Putnam, 2020), telles que les traits de personnalité (Millecamp *et al.*, 2019), le style de prise de décision (Naveed *et al.*, 2018) et les capacités cognitives (Riefler *et al.*, 2022), peuvent moduler la perception et l'effet des explications. Cependant, les travaux portant sur les caractéristiques des utilisateurs qui influencent la compréhension des explications des systèmes d'IA sont, à notre connaissance, rares. L'effet des caractéristiques personnelles sur la perception d'explications a été étudié dans certains cas particuliers tels que les recommandations musicales (Millecamp *et al.*, 2019), mais très peu à notre connaissance dans les EIAH (Khosravi *et al.*, 2022), même si certains chercheurs de la communauté ont commencé à étudier cette question dans le but de fournir un enseignement personnalisé aux apprenants en fonction de leurs besoins, de leurs états et de leurs capacités (Conati *et al.*, 2021).

Afin d'apporter des premiers éléments de réponse concernant l'importance des caractéristiques des individus sur leurs utilisation, compréhension et perception des explications dans les EIAH, nous proposons d'introduire des explications sur les résultats d'un algorithme utilisé dans notre plateforme d'apprentissage de l'informatique. Le choix de se concentrer sur les résultats est fondé sur les recommandations de Ribera et Lapedriza (2019) pour les utilisateurs non experts en IA. Notre étude explore, dans le contexte d'un cours de programmation destiné à des étudiants en première année de BUT, les questions de recherche **QR1** et **QR2** qui concernent respectivement les caractéristiques personnelles qui peuvent influencer la consultation, la compréhension et l'utilité perçue des explications, et le rôle de ces caractéristiques personnelles et de l'utilisation des explications dans la perception du système.

3. MÉTHODES

3.1. OUTILS ET CONTEXTE EXPÉRIMENTAL

3.1.1. La plateforme Lab4CE

L'expérimentation a été réalisée à l'aide de la plateforme Lab4CE (Broisin *et al.*, 2017), dédiée aux travaux pratiques en informatique. Elle intègre divers outils et visualisations permettant des sessions d'activités pratiques à distance. Lab4CE fournit à chaque apprenant un laboratoire virtuel personnel, et collecte les données d'interaction des apprenants (par exemple, l'exécution d'une commande ou d'un script, la modification du code source d'un programme, ou encore l'accès à une visualisation spécifique). Toutes ces actions sont contextualisées temporellement.

3.1.2. Le système à base de compétences

Lab4CE intègre également une approche fondée sur les compétences : un ensemble de compétences est associé à la discipline étudiée par les apprenants, et le niveau de maîtrise des apprenants pour chaque compétence est calculé automatiquement sur la base de leurs activités sur la plateforme. Les apprenants peuvent aussi sélectionner une ou plusieurs compétences et demander des recommandations de matériel d'apprentissage portant sur ces compétences. L'interface dédiée au **système à base de compétences** est illustrée dans la figure 1. Elle est accessible à partir d'un bouton de la barre de menu de Lab4CE. La partie de gauche montre le **profil de compétences** d'un apprenant, la partie à droite est consacrée au **système de recommandation**.

Le profil de compétences. Ce profil est affiché sous la forme d'une structure arborescente qui représente l'ensemble des compétences associées à la discipline étudiée. Les compétences de la figure 1 sont les compétences impliquées dans un cours de programmation Shell, qui est le sujet de notre expérimentation. Pour ce thème, les enseignants ont défini un ensemble de 55 compétences, dont certaines sont composées de compétences de granularité plus fine. Par exemple, la compétence générale « *Répéter des instructions* » comprend les compétences « *Connaître la syntaxe de l'instruction while* » et « *Connaître la syntaxe de l'instruction for* ».

Extrait de code 1 : Algorithme de calcul du niveau de maîtrise

```

1 listeCompetencesGranulariteFine, listeCompetencesGenerales
2 profilCompetencesEtudiant, tracesEtudiant
3 for each competence in listeCompetencesGranulariteFine do
4     maitrise <- 0
5     for each regex in competence do
6         resultatsCorrespondances <-
7             TrouverCorrespondances(regex, tracesEtudiant)
8         evalRegex <- EvaluerRegex(regex, resultatsCorrespondances)
9         maitrise <- Maximum(0, Minimum(1, maitrise + evalRegex))
10    profilCompetencesEtudiant[competence].maitrise <- maitrise
11
12 for each competence in listeCompetencesGenerales do
13     maitrise <- MoyenneMaitrisesSousCompetences(competenceGenerale)
14     profilCompetencesEtudiant[competenceGenerale].maitrise <- maitrise
15
16 Fonction EvaluerRegex(regex, resultatsCorrespondances) :
17     nbExecutionsCorrectes <- 0
18     nbExecutionsErreurs <- 0
19     ppos <- regex.poidsPositif
20     pneg <- regex.poidsNegatif

```



```

21     nbMaxCorrectes <- regex.nbMaxCorrectes
22     nbMaxErreurs <- regex.nbMaxErreurs
23     for each correspondance in resultatsCorrespondances do
24         if VerifierExecutionCorrecte(correspondance) == "correct" then
25             nbExecutionsCorrectes <- nbExecutionsCorrectes + 1
26         else
27             nbExecutionsErreurs <- nbExecutionsErreurs + 1
28     return ppos x Minimum(nbMaxCorrectes, nbExecutionsCorrectes) -
29           pneg x Minimum(nbMaxErreurs, nbExecutionsErreurs)
30
31 Fonction TrouverCorrespondances(regex, tracesEtudiant):
32     # Renvoie la liste des correspondances trouvees dans les commandes
33     # et les scripts de letudiant pour la regex donnee en entree.
34
35 Fonction MoyennesMaitrisesSousCompetences(competence):
36     # Renvoie la moyenne des maitrises des sous-competences comprises
37     # dans la competence generale donnee en entree.
38
39 Fonction VerifierExecutionCorrecte(correspondance):
40     # Verifie si la correspondance (commande ou script)
41     # donnee en entree produit une erreur a l'execution ou non.

```

Pour chaque apprenant, le niveau de maîtrise de chacune des compétences est calculé de manière automatique par le système. Ce niveau de maîtrise est compris entre 0 et 1, et résulte d'un algorithme qui s'appuie sur des règles d'analyse des commandes et des scripts exécutés sur la plateforme Lab4CE par l'apprenant. Chaque compétence est associée à une ou plusieurs expressions régulières déterminées par des experts, et l'algorithme vérifie si les commandes ou les scripts écrits par les étudiants correspondent à ces expressions régulières. Plus précisément, chaque expression régulière à deux couples de valeurs (poids et nombres maximum d'occurrences) : le premier couple contient un **poids positif** p_{pos} , un **poids négatif** p_{neg} , et le deuxième comprend un **nombre maximal d'occurrences correctes** $nbMaxCorrectes$ et un **nombre maximal d'occurrences incorrectes** $nbMaxErreurs$. Pour obtenir le niveau de maîtrise d'une compétence d'un apprenant, l'algorithme cherche dans les traces de cet apprenant les correspondances avec les différentes expressions régulières associées à la compétence, et pour chaque correspondance trouvée, le niveau de maîtrise augmente de p_{pos} , dans une limite de $nbMaxErreurs$ augmentations possibles pour une même expression régulière (cf. la fonction `EvaluerRegex` dans Extrait de code 1). De plus, la valeur du taux de maîtrise est plafonnée à 1 et ne peut pas descendre en-dessous de 0.

Par exemple, la compétence « *Connaître les variables spéciales* » est associée à 7 expressions régulières correspondant aux variables spéciales `$#`, `$-`, `$?`, `$$`, `$!`, `$0` et `$*` (cf. Extrait de code 2 qui détaille cet exemple de structure `Compétence`), avec pour chacune un poids positif de 0,15 et un nombre maximal d'occurrences correctes de 1. Ainsi, si un étudiant utilise l'une de ces expressions dans un script ou une commande qui s'exécute correctement, le niveau de maîtrise de la compétence augmente de 0,15. Cependant, même si une variable spéciale est trouvée plusieurs fois, le niveau de maîtrise n'augmente que pour la première correspondance détectée. Les étudiants sont donc obligés d'utiliser au moins une fois chaque variable spéciale pour valider complètement la compétence.

Extrait de code 2 : Expressions régulières, nombres maximaux d'occurrences et poids associés à la compétence « *Connaître les variables spéciales* »

```

1     {
2         "competencyName": "Connaitre_les_variables_speciales",
3         "regexList": [
4             {

```

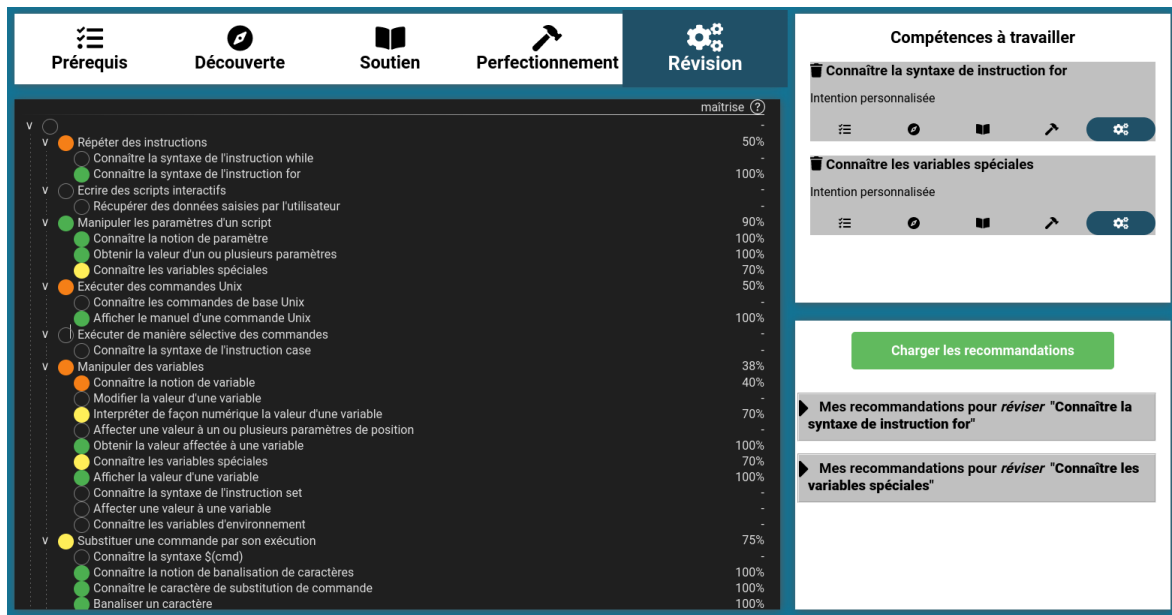


Figure 1 : Le profil de compétences et le système de recommandations d'un apprenant

```

5      "regex": ".*\\$#.*",
6      "poids": {0.15, 0.08},
7      "nombreMaxOccurrences": {1,1}
8    },
9    {
10     "regex": ".*\\$\\-.*",
11     "poids": {0.15, 0.08},
12     "nombreMaxOccurrences": {1,1}
13   },
14   {
15     "regex": ".*\\$\\?.*",
16     "poids": {0.15, 0.08},
17     "nombreMaxOccurrences": {1,1}
18   },
19   {
20     "regex": ".*\\$\\$.*",
21     "poids": {0.15, 0.08},
22     "nombreMaxOccurrences": {1,1}
23   },
24   {
25     "regex": ".*\\$!.*",
26     "poids": {0.15, 0.08},
27     "nombreMaxOccurrences": {1,1}
28   },
29   {
30     "regex": ".*\\$0.*",
31     "poids": {0.15, 0.08},
32     "nombreMaxOccurrences": {1,1}
33   },
34   {
35     "regex": ".*\\$\\*.*",

```

```

36         "poids": {0.15, 0.08},
37         "nombresMaxOccurrences": {1,1}
38     }
39 ]
40 }

```

Un autre exemple est la compétence « *Connaître la syntaxe de l'instruction for* », qui ne comprend qu'une seule expression régulière (`. *for.*in.*do.*done.*`) d'un poids de 0,5 et d'un nombre maximal d'occurrences de 2. Pour valider complètement cette compétence, un apprenant doit donc utiliser au moins deux fois la syntaxe d'une boucle `for` dans des scripts s'exécutant correctement.

Si des correspondances d'expressions régulières sont trouvées dans des scripts ou des commandes qui ne se sont pas exécutés correctement, le niveau de maîtrise de la compétence associée **baisse** de p_{neg} , dans une limite de $nbMaxErreurs$ baisses de niveau. Si p_{neg} n'est pas spécifié, il vaut par défaut la moitié de p_{pos} , et si $nbMaxErreurs$ n'est pas spécifié, il est par défaut égal à $nbMaxCorrectes$.

Le niveau de maîtrise est ensuite converti en pourcentage et affiché dans la colonne **maîtrise** (voir figure 1). Le cercle de couleur associé à chaque compétence reflète la valeur du niveau de maîtrise de l'apprenant pour cette compétence. La couleur verte correspond à une maîtrise supérieure à 75 %, la couleur jaune à une maîtrise comprise entre 50 et 75 %, la couleur orange à une maîtrise comprise entre 25 et 50 %, la couleur rouge à une inférieure à 25 %, et la couleur noire correspond à une compétence qui n'a pas encore été travaillée ou pour laquelle aucune correspondance avec une expression régulière n'a été trouvée.

Le système de recommandation. Le système de recommandation correspond à la partie droite de la figure 1. Chaque compétence du profil d'un étudiant peut être glissée et déposée dans la partie *Compétences à travailler*, et une intention de travail peut être spécifiée : Prérequis, Découverte, Soutien, Révision ou Perfectionnement. L'intention Prérequis consiste à travailler sur des compétences qui sont des prérequis des compétences sélectionnées. L'intention Découverte permet aux étudiants de travailler sur les compétences pour la première fois. L'intention Soutien propose des contenus pour travailler des compétences déjà étudiées mais non maîtrisées. L'intention Révision suggère des ressources pédagogiques qui aident à réviser pour un examen. L'intention Perfectionnement consiste à travailler sur des compétences qui sont des complexifications des compétences ciblées. Une fois que l'étudiant a sélectionné les compétences à travailler, il peut demander des recommandations en cliquant sur le bouton *Charger les recommandations*. Les recommandations apparaissent alors dans un menu déroulant (voir figure 2) et correspondent à des exercices ou des rappels de cours adaptés au profil de compétences de l'apprenant, aux compétences à travailler et aux intentions de travail. L'algorithme de recommandation est décrit dans (Sablayrolles *et al.*, 2021).



Figure 2 : Menu déroulant des recommandations associées à la compétence « *Connaître les variables spéciales* »

3.1.3. Explications relatives au calcul du niveau de maîtrise des compétences.

Contenu. Les explications que nous avons conçues visent à expliquer aux apprenants comment leur niveau de maîtrise, pour chaque compétence de leur profil, est calculé. Nous avons choisi de fournir des explications locales, car ce type d'explication a été mis en œuvre avec succès dans des travaux antérieurs dans des contextes similaires à notre expérimentation (Conati *et al.*, 2021). Ces explications s'appuient sur les productions des apprenants. Nous avons distingué trois cas : un niveau de maîtrise supérieur à 0,75 ; inférieur à 0,75 et supérieur à 0 ; égal à 0. Dans les cas 1 et 2, l'explication est personnalisée et composée de quatre parties (voir figure 3a) :

1. Le libellé de la compétence.
2. Un message textuel associé au niveau de maîtrise de la compétence (dans le cas 1, le message est « Bravo, vous avez une bonne maîtrise de cette compétence car vous avez utilisé dans votre code suffisamment d'instructions correctes correspondant à cette compétence. » ; dans le cas 2, le message est « Bravo, vous commencez à maîtriser cette compétence car vous avez utilisé de manière correcte certaines commandes ou instructions dans votre code. »). Ce message contient un compliment envers l'accomplissement d'une tâche : dans un feedback, c'est ce type de compliment qui permet d'augmenter l'auto-efficacité, au contraire des compliments dirigés vers la personne (Hattie et Timperley, 2007).
3. Un extrait de code issu des productions de l'étudiant qui a contribué à élever le niveau de maîtrise.
4. Un message encourageant l'apprenant à continuer à s'entraîner pour mieux maîtriser la compétence. Lorsque le niveau de maîtrise est à 0 (cas 3), l'explication contient 2 parties (voir figure 3b) :
 - Le libellé de la compétence.
 - Un message textuel expliquant que le niveau de maîtrise de la compétence augmentera lorsque les commandes liées à cette compétence seront exécutées. Aucun exemple de commande attendue n'est donné.

En complément de ces explications locales, les apprenants avaient également accès à une explication globale du fonctionnement de l'algorithme. L'objectif de cette explication était de décrire, dans les grandes lignes et avec des phrases simples, le fonctionnement de l'algorithme décrit dans l'Extrait de code 1. Cette explication globale est montrée dans la figure 4.

Forme. Les explications pour chaque niveau de maîtrise peuvent être obtenues en cliquant sur le niveau de maîtrise (voir le profil de compétences dans la figure 1), ce qui fait apparaître une fenêtre pop-up contenant les éléments textuels décrits ci-dessus. Le format textuel a été choisi car il semble plus adapté pour expliquer du code, mais également car certains résultats montrent que des explications textuelles sont plus convaincantes que des explications visuelles (Kouki *et al.*, 2019) et que les explications visuelles, même lorsqu'elles sont préférées par les utilisateurs, ne sont pas nécessairement les plus efficaces (Tsai et Brusilovsky, 2019). L'explication globale est quant à elle accessible en survolant le symbole « ? » visible à côté de la colonne « maîtrise » dans la figure 1.

Temporalité. Les explications sont fournies à la demande et non imposées à des moments précis durant la session de travail. En effet, dans de précédents travaux, nous avons fait le choix d'imposer du feedback avec des explications aux étudiants à deux reprises durant chaque session, mais ce feedback avait été peu lu et pris en compte par les étudiants (Félix *et al.*, 2022). Les étudiants avaient exprimé une préférence pour un fonctionnement à la demande de manière à ne pas attirer leur attention sur d'autres éléments que le travail à réaliser.

Explication de la compétence « Connaître la syntaxe de l'instruction for » ×

Bravo, vous avez une **bonne maîtrise** de cette compétence car vous avez utilisé dans votre code suffisamment d'instructions correctes correspondant à cette compétence.

Voici un **extrait de votre code** qui a contribué à faire **monter votre taux de maîtrise** pour cette compétence :

```
#!/bin/bash for i in 1 2 3 4 5 do echo "Hello World $i" done
```

Continuez à **pratiquer régulièrement** pour consolider vos acquis !

(a) Taux de maîtrise de 1 (cas 1)

Explication de la compétence « Connaître la syntaxe de l'instruction for » ×

Le taux de maîtrise de cette compétence augmentera lorsque vous taperez des commandes ou des scripts liés à cette compétence.

(b) Taux de maîtrise de 0 (cas 3)

Figure 3 : Exemples d'explications fournies aux apprenants.

Taux de maîtrise (mis à jour une fois par jour)

Le taux de maîtrise est calculé à partir d'un **algorithme** qui récupère le code et les commandes que vous tapez sur Lab4CE.

L'algorithme cherche ensuite des **motifs** dans votre code et fait **monter votre niveau de maîtrise des compétences** en conséquence.

Par exemple, la maîtrise de la compétence "Connaître la syntaxe de l'instruction seq" augmentera si l'algorithme trouve des instructions **seq** valides dans vos commandes ou votre code.

Figure 4 : Explication globale du système

3.1.4. Contexte expérimental.

L'expérimentation s'est déroulée dans le cadre d'un cours de programmation Shell et comprenait une session pratique de 90 minutes par semaine. Ces sessions consistaient en des exercices à réaliser sur la plateforme Lab4CE, laquelle était disponible 24 heures sur 24 pendant toute la durée de l'expérimentation. Cette dernière a duré 4 semaines avec 98 étudiants inscrits en première année de BUT d'informatique. Avant le début de l'expérimentation, les étudiants ont participé à une session au cours de laquelle ils ont été guidés par leurs enseignants sur la manière d'utiliser les fonctionnalités de la plateforme, y compris

l'interface du système à base de compétences. Afin de répondre à nos questions de recherche, deux groupes de 49 étudiants ont été constitués de manière aléatoire. Les étudiants des deux groupes (contrôle et expérimental) ont effectué les séances de travaux pratiques durant les 4 semaines de l'expérimentation et avaient accès au profil de compétences, mais seuls les étudiants du groupe expérimental avaient accès aux explications du profil. Les étudiants du groupe contrôle avaient accès au profil, mais sans les explications. Pour tous les étudiants, le profil était mis à jour tous les soirs à minuit (phase 2 dans la figure 5).

Avant le début de l'expérimentation (phase 1 dans la figure 5), tous les apprenants ont passé un test pré-expérimental qui évaluait leurs connaissances théoriques préalables sur le thème du cours, et ils ont été invités à remplir un questionnaire pré-expérimental Q1 fondé sur le MSLQ, *Motivated Strategies for Learning Questionnaire* (Pintrich et Others, 1991). Après l'expérimentation (phase 3 dans la figure 5), ils ont passé un examen terminal noté et ont également été invités à remplir un questionnaire post-expérimentation Q2 concernant leurs perceptions et compréhension du système. Les deux questionnaires sont disponibles en ligne (<https://osf.io/tc29a>) et leurs contenus sont détaillés dans la section 3.2.2. Enfin, nous avons organisé un focus group après l'expérimentation avec 6 étudiants du groupe contrôle et 2 étudiants du groupe expérimental afin de recueillir leurs opinions et points de vue sur les explications. Le focus group avait deux objectifs principaux : premièrement, identifier les raisons pour lesquelles ils ont consulté et utilisé le système à base de compétences et les explications associées s'ils y avaient accès; deuxièmement, recueillir leurs opinions et besoins sur le contenu et la forme que devraient avoir les explications.

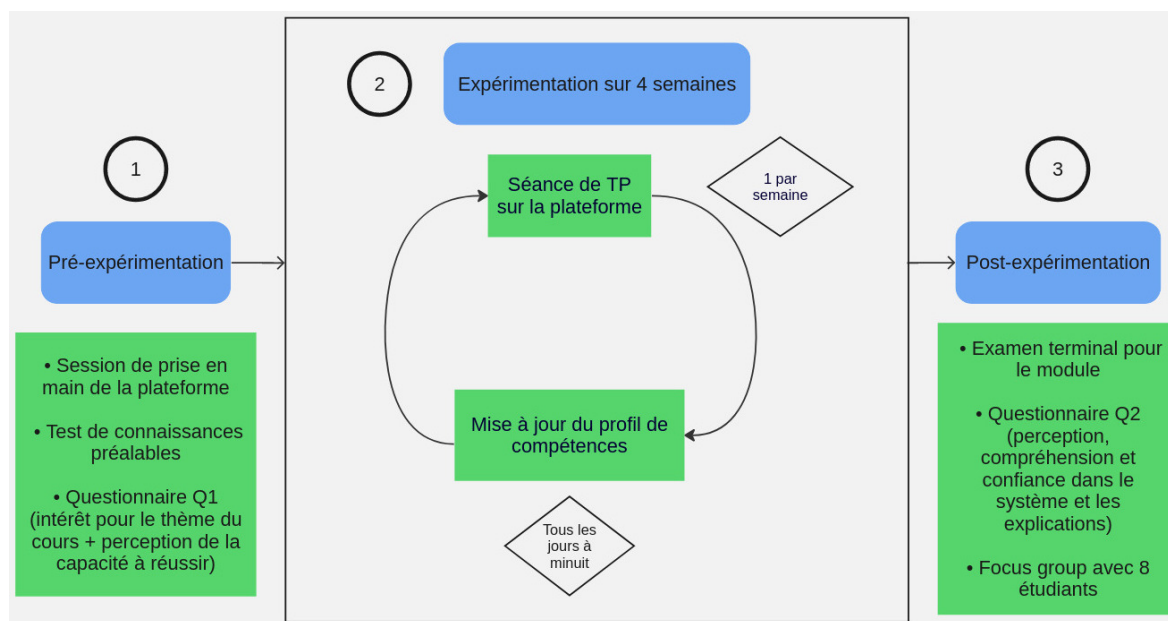


Figure 5 : Déroulement de l'expérimentation pour tous les étudiants (groupe contrôle et groupe expérimental)

3.2. DONNÉES COLLECTÉES

3.2.1. Données quantitatives.

À partir des données d'interaction collectées par Lab4CE, nous avons extrait, pour chaque étudiant, des données relatives à l'exécution de commandes et de scripts. Ces données reflètent la principale activité d'apprentissage des apprenants et ont été considérées comme une

mesure de leur **engagement**. En effet, même si Fredricks *et al.* (2004) ont identifié au moins trois types d'engagement (comportemental, émotionnel et cognitif), nous nous référons dans cet article à l'engagement comportemental sur la tâche; ainsi, par engagement, nous nous référons à l'implication des étudiants dans des activités pratiques sur la plateforme. Nous avons également calculé, pour chaque apprenant, le nombre de commandes et de scripts qui ont généré des erreurs au moment de l'exécution, par rapport à toutes les exécutions. Nous avons considéré cet indicateur comme une variable reflétant la **performance sur la tâche d'apprentissage**, où un pourcentage élevé d'erreurs reflète une faible performance sur la tâche d'apprentissage. Nous avons également recueilli le nombre d'explications consultées par les étudiants du groupe expérimental. Cependant, pour des raisons techniques, seules les consultations des explications locales ont pu être relevées, les accès à l'explication globale n'ayant pas pu être comptabilisés. Enfin, nous avons utilisé les notes obtenues par les étudiants lors du test préalable à l'expérimentation comme leurs **connaissances préalables**, et leurs notes à l'examen terminal comme leur **performance académique**.

3.2.2. Données qualitatives.

Le questionnaire Q1 comprenait des questions de Likert à 7 niveaux fondées sur le questionnaire MSLQ de Pintrich *et al.* (Pintrich *et Others*, 1991). Cinq items mesuraient l'**intérêt des apprenants pour le thème du cours** et huit items mesuraient leur **perception de leur capacité à réussir dans le module**.

Le questionnaire Q2 mesurait, à l'aide de 15 questions de Likert à 7 niveaux, la **perception de l'utilité** (Davis, 1989), la **compréhension** et la **confiance** des étudiants dans le système à base de compétences. Les étudiants du groupe expérimental ont été invités à répondre à 5 questions de Likert à 7 niveaux sur leur **perception de l'utilité et leur compréhension des explications**. Ces étudiants ont également été invités à répondre à 5 questions ouvertes sur le contenu et la forme des explications et sur leurs éventuels besoins supplémentaires. Les étudiants du groupe contrôle, en plus des 15 questions sur le système à base de compétences, ont été invités à répondre à 2 questions pour savoir s'ils souhaitaient ou non des explications sur leur profil de compétences, et quels types d'explications ils souhaitaient.

À partir des données qualitatives et quantitatives, nous avons donc représenté les apprenants avec **6 caractéristiques personnelles** : intérêt pour le thème du cours, perception de leur capacité à réussir, connaissances préalables, engagement, performance sur la tâche d'apprentissage et performance académique. Nous distinguons deux catégories parmi ces 6 caractéristiques : les caractéristiques *initiales* (intérêt pour le thème du cours, perception de leur capacité à réussir, connaissances préalables) et les caractéristiques *comportementales* (engagement, performance sur la tâche d'apprentissage et performance académique).

Enfin, nous avons transcrit les enregistrements du **focus group** et saisi dans Nvivo les verbatims qui en résultaient. Les verbatims ont été analysés par deux membres de l'équipe de recherche afin d'identifier les thèmes principaux qui émergeaient, à l'aide d'une approche de codage multithématique (Ayache & Dumez, 2011).

4. ANALYSE DE DONNÉES

L'analyse des données a été réalisée à l'aide des outils jamovi et Jupyter Notebook. Nous avons utilisé l'alpha de Cronbach pour évaluer la cohérence interne de l'intérêt pour le thème du cours et de la perception de la capacité à réussir, qui ont été mesurés à l'aide d'items de Likert. Nous avons obtenu une cohérence supérieure à 0,7 et donc satisfaisante pour les items mesurant l'intérêt ($\alpha = 0,757$) ainsi que pour les items mesurant la perception de la capacité

à réussir ($\alpha = 0,935$). Ceci nous conduit à exprimer l'intérêt pour le thème du module et la perception de la capacité à réussir comme deux variables constituées des sommes des items Likert correspondants.

Pour répondre à QR1 sur la corrélation entre les caractéristiques personnelles des apprenants et la consultation, la compréhension et l'utilité perçue des explications, nous avons d'abord classé les apprenants en différents groupes en fonction de leurs caractéristiques personnelles. Pour chaque caractéristique personnelle, nous avons créé trois catégories : élevée, moyenne et faible. Par exemple, les étudiants qui ont obtenu une note dans le premier tiers de la classe à l'examen final ont été classés dans la catégorie « performance académique élevée », les étudiants qui ont obtenu une note dans le dernier tiers de la classe ont été classés dans la catégorie « performance académique faible » et les autres, dans la catégorie « performance académique moyenne ». Le même processus a été appliqué pour diviser les apprenants en trois groupes pour toutes les autres caractéristiques personnelles. Ensuite, pour les étudiants du groupe expérimental, une analyse de corrélation de Pearson a été menée entre toutes les caractéristiques personnelles et leur perception, leur compréhension et leur utilisation des explications afin de déterminer l'existence de relations linéaires significatives entre les paires de variables de ces deux ensembles. Les corrélations ont été calculées d'abord pour tous les étudiants qui ont reçu des explications, puis, parmi ces étudiants, pour les sous-groupes qui appartenaient aux catégories haute et basse de chaque caractéristique personnelle. Les conditions requises pour l'analyse de Pearson (variables normalement distribuées, linéarité, absence de valeurs aberrantes) ont toujours été vérifiées.

Pour répondre à QR2.1 concernant l'influence des caractéristiques personnelles des apprenants sur leur perception du système à base de compétences, nous avons effectué une analyse de corrélations entre les caractéristiques personnelles des apprenants du groupe expérimental et leur utilité perçue, compréhension et confiance dans le système à base de compétences. Les corrélations ont également été calculées pour différentes catégories d'étudiants de la même manière que l'analyse effectuée pour la question QR1.

Pour répondre à QR2.2 qui porte sur la mesure dans laquelle la consultation, la compréhension et l'utilité perçue des explications affectent la perception du système à base de compétences, nous avons effectué deux analyses. Nous avons d'abord effectué une analyse de corrélation pour les étudiants qui ont consulté au moins une explication entre, d'une part, la compréhension et l'utilité perçue des explications, et d'autre part, la compréhension, l'utilité perçue et la confiance dans le système à base de compétences. L'objectif de cette analyse était d'explorer une relation potentielle entre les variables liées aux explications et la perception du système. Ensuite, nous avons effectué des tests t à deux échantillons entre d'une part le groupe des étudiants ayant reçu des explications et le groupe contrôle, afin de voir si leur perception et leur compréhension du système à base de compétences était différente.

5. RÉSULTATS

5.1. CORRÉLATIONS ENTRE LES CARACTÉRISTIQUES PERSONNELLES DES APPRENANTS ET LA PERCEPTION, LA COMPRÉHENSION ET LA CONSULTATION DES EXPLICATIONS (QR1)

Le tableau 1 expose les moyennes et écarts-types des six caractéristiques personnelles considérées.

Pour différentes catégories d'étudiants du groupe expérimental ayant reçu des explications, le tableau 2 montre les corrélations entre les caractéristiques personnelles des apprenants et la consultation, la compréhension et l'utilité perçue des explications. Ce tableau ne

Tableau 1 : Moyenne, médiane et écart-type pour les six variables correspondant aux caractéristiques personnelles

Variables	Moyenne	Médiane	Écart-type
Nombre d'actions effectuées sur la plateforme (engagement)	4,24	3,47	4,30
Connaissances préalables (note au questionnaire de début du cours)	11,4	11,3	3,26
Performance sur la tâche d'apprentissage (% d'erreurs)	7,13	3,73	126
Performance académique (note à l'examen)	7,70	8,0	4,17
Intérêt pour le thème du cours (Likert 7 items)	5,26	5,20	0,818
Perception de la capacité à réussir (Likert 7 items)	4,9	5,0	0,954

présente que les résultats significatifs des différentes analyses, c'est-à-dire ceux pour lesquels on obtient une valeur p inférieure à 0,05. Nous utilisons le standard de Cohen pour évaluer l'ampleur de l'effet (petit effet lorsque $r > 0,10$, effet moyen lorsque $r > 0,30$, et grand effet lorsque $r > 0,50$), car il permet de mesurer la taille de l'effet par rapport à l'hypothèse nulle.

Tableau 2 : Corrélations significatives entre les caractéristiques personnelles des apprenants et les variables liées aux explications, pour différentes catégories d'apprenants du groupe expérimental.

Catégories d'étudiants	Caractéristiques personnelles	Variables liées aux explications	r de Pearson	p
Tous les étudiants	Perception de la capacité à réussir	Consultation des explications	0,366	0,018
Haute performance sur la tâche	Perception de la capacité à réussir	Consultation des explications	0,568	0,043
Haute performance sur la tâche	Perception de la capacité à réussir	Compréhension des explications	0,758	0,048
Faible engagement	Performance sur la tâche d'apprentissage	Compréhension des explications	0,754	0,031

L'analyse des corrélations effectuée avec l'ensemble des étudiants du groupe expérimental montre une corrélation positive entre la perception de la capacité à réussir et la consultation des explications ($r = 0,366$, $p = 0,018$). Cette corrélation est encore plus forte chez les apprenants caractérisés par une performance élevée sur la tâche ($r = 0,568$, $p = 0,043$). Pour les étudiants de ce groupe, il existe également une forte corrélation entre leur perception de leur capacité à réussir et leur compréhension des explications ($r = 0,758$, $p = 0,048$).

Lorsque l'on considère les étudiants caractérisés par un faible engagement dans la tâche

d'apprentissage, nous trouvons une corrélation positive importante entre la performance sur la tâche d'apprentissage et la compréhension des explications ($r = 0,754, p = 0,031$).

Les résultats ne montrent pas de corrélations significatives entre les autres variables. En particulier, l'utilité perçue des explications n'est corrélée à aucune caractéristique personnelle.

5.2. CORRÉLATIONS ENTRE LES CARACTÉRISTIQUES PERSONNELLES DES APPRENANTS ET LA PERCEPTION DU SYSTÈME À BASE DE COMPÉTENCES (QR2.1)

Le tableau 3 montre les corrélations entre, d'une part, les caractéristiques personnelles des apprenants du groupe expérimental ayant reçu des explications et, d'autre part, la compréhension, l'utilité perçue et la confiance dans le système à base de compétences pour lesquelles on obtient une valeur p inférieure à 0,05.

Tableau 3 : Corrélations significatives entre les caractéristiques personnelles des apprenants et les variables liées au système à base de compétences, pour différentes catégories d'apprenants du groupe expérimental.

Catégories d'étudiants	Caractéristiques personnelles	Variabes liées aux système	r de Pearson	p
Tous les étudiants	Performance sur la tâche d'apprentissage	Compréhension du système	0,351	0,042
Hautes connaissances préalables	Intérêt dans le thème du cours	Utilité perçue du système	0,583	0,029
Haut engagement	Perception de la capacité à réussir	Utilité perçue du système	0,807	0,003
Haut engagement	Performance académique	Compréhension du système	0,627	0,039
Faible engagement	Performance académique	Utilité perçue du système	-0,576	0,039
Faible performance sur la tâche	Engagement	Utilité perçue du système	0,646	0,032

Les résultats montrent une corrélation positive moyenne entre la performance sur la tâche et la compréhension du système à base de compétences pour tous les étudiants du groupe expérimental ($r = 0,351, p = 0,042$). En ne retenant que les étudiants avec des connaissances préalables élevées, les résultats montrent une forte corrélation positive entre l'intérêt des apprenants pour le thème du cours et leur perception de l'utilité du système ($r = 0,583, p = 0,029$). Pour les étudiants très engagés, la perception de leur capacité à réussir est positivement et fortement corrélée à l'utilité perçue du système ($r = 0,807, p = 0,003$). De même, la performance académique de ces étudiants est positivement et fortement corrélée à leur compréhension du système ($r = 0,627, p = 0,039$). En revanche, pour les étudiants faiblement engagés, les résultats académiques sont négativement et fortement corrélés à la perception de l'utilité du système ($r = -0,576, p = 0,039$). Enfin, pour les étudiants peu performants sur la tâche d'apprentissage, l'engagement est positivement et fortement corrélé à la perception de l'utilité du système ($r = 0,646, p = 0,032$).

Les résultats ne montrent pas d'autres corrélations significatives. En particulier, aucune corrélation impliquant la confiance dans le système n'a été trouvée.

5.3. CORRÉLATIONS ENTRE LA PERCEPTION DES EXPLICATIONS PAR LES APPRENANTS ET LEUR PERCEPTION DU SYSTÈME (QR2.2)

En étudiant les apprenants du groupe expérimental qui ont consulté au moins une explication, les résultats montrent des corrélations significatives très fortes entre toutes les variables que nous avons considérées (voir tableau 4). En effet, l'utilité perçue des explications est positivement et fortement corrélée avec la compréhension du système ($r = 0,866$, $p < 0,001$) et la confiance dans le système ($r = 0,860$, $p < 0,001$). De même, la compréhension des explications est positivement et fortement corrélée à la confiance dans le système ($r = 0,786$, $p < 0,001$) et à l'utilité perçue du système ($r = 0,599$, $p = 0,018$).

Tableau 4 : Corrélations significatives entre les caractéristiques personnelles des apprenants et les variables liées au système à base de compétences, pour différentes catégories d'apprenants du groupe expérimental.

Variables liées aux explications	Variables liées au système à base de compétences	<i>r</i> de Pearson	<i>p</i>
Utilité perçue des explications	Compréhension du système	0,866	< 0,001
Utilité perçue des explications	Confiance dans le système	0,860	< 0,001
Compréhension des explications	Utilité perçue du système	0,599	< 0,001
Compréhension des explications	Confiance dans le système	0,786	< 0,001

Les tests *t* n'ont pas montré de différence statistiquement significative entre les groupes étudiés. Cependant, la figure 6 montre une tendance à une plus grande compréhension et confiance envers le système dans le groupe des apprenants qui ont consulté des explications (groupe expérimental) par rapport au groupe de ceux qui n'en ont pas consulté (groupe contrôle).

5.4. FOCUS GROUP

L'analyse de la retranscription du focus group a été effectuée par deux membres de l'équipe de recherche. Le codage multithématique (Ayache & Dumez, 2011) a fait émerger les thèmes présentés dans la figure 7. En ce qui concerne le thème « utilisation du système », les étudiants signalent majoritairement une faible utilisation du système à base de compétences et des explications qui y sont associées. L'outil de programmation a quant à lui été utilisé lors des travaux pratiques, mais peu lors du travail personnel des étudiants. Les obstacles à l'utilisation du système se recoupent grandement avec le thème « dysfonctionnements du système », avec une évocation, par la moitié des étudiants du focus group, de problèmes de lenteur et de connexion à la plateforme ayant conduit à limiter leur usage de celle-ci ainsi que leur consultation des explications associées. Pour la compréhension du système, parmi les deux étudiants caractérisés par une faible performance ou un faible engagement, l'une a fait des remarques témoignant de conceptions erronées concernant le fonctionnement du système de mise à jour automatique des compétences. Parmi les six autres étudiants caractérisés par

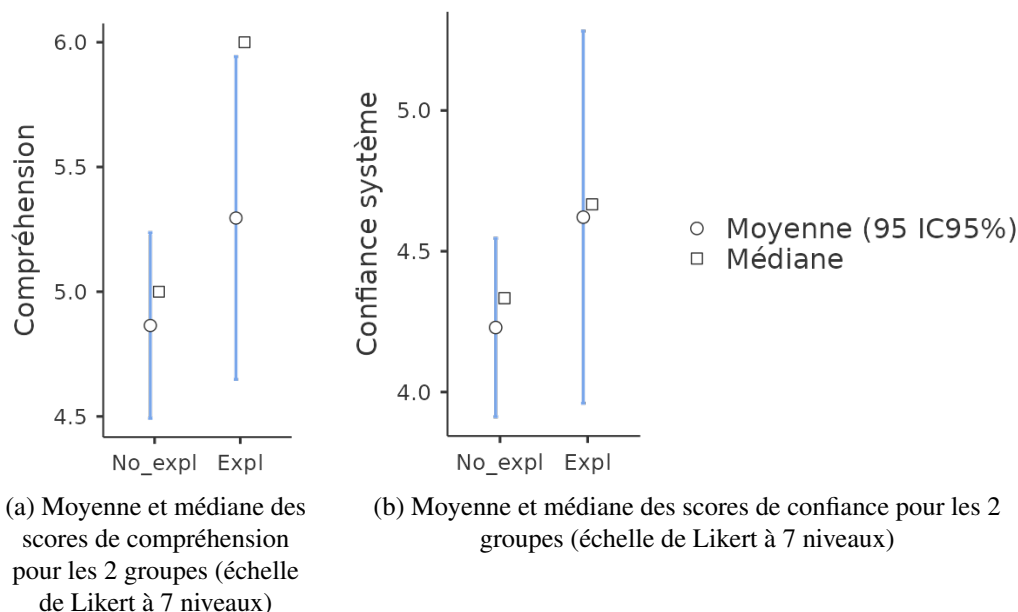


Figure 6 : Comparaison des scores de compréhension et de confiance dans le système à base de compétences des apprenants ayant consulté des explications (Expl) par rapport à ceux qui n'en ont pas consulté (No_expl)

une haute performance et un fort engagement, deux ont exprimé une bonne compréhension du système et une bonne compréhension des explications associées, mais peu d'intérêt pour ces dernières, en évoquant un manque d'utilité. Les quatre autres ont également exprimé une bonne compréhension du système et des explications, et montraient un intérêt pour ces dernières, avec une volonté forte de comprendre le fonctionnement du système. Concernant le thème « idées d'amélioration et besoins exprimés », nous avons également observé des recoupements avec le thème « dysfonctionnements du système », avec des demandes d'amélioration de la vitesse du système et d'une meilleure compatibilité avec certains navigateurs. D'autres demandes concernaient directement le design des explications, avec deux étudiants exprimant un besoin de recevoir des exemples plus précis de dysfonctionnements dans leur code et de problèmes les ayant empêchés de monter leurs compétences. Ces étudiants souhaiteraient que ces exemples soient accompagnés de conseils de corrections à appliquer.

6. DISCUSSION ET LIMITATIONS

6.1. QR1 - CARACTÉRISTIQUES PERSONNELLES EN CORRÉLATION AVEC LA CONSULTATION, LA COMPRÉHENSION ET LA PERCEPTION DES EXPLICATIONS

Nos résultats indiquent une corrélation positive moyenne entre la perception des étudiants de leur capacité à réussir et la consultation des explications. Cette relation est particulièrement forte chez les étudiants qui font preuve d'une grande performance sur la tâche d'apprentissage. De plus, chez ces étudiants, la perception de leur capacité à réussir est également positivement corrélée à leur compréhension des explications. Ces résultats sont cohérents avec une étude qui a montré la contribution des croyances d'efficacité personnelle dans l'apprentissage autorégulé (Zuffianò *et al.*, 2013). Ces résultats tendent donc à montrer que **les étudiants qui ont une faible perception de leur capacité à réussir sont moins susceptibles d'accéder aux explications et de les comprendre**, et donc d'en tirer profit.

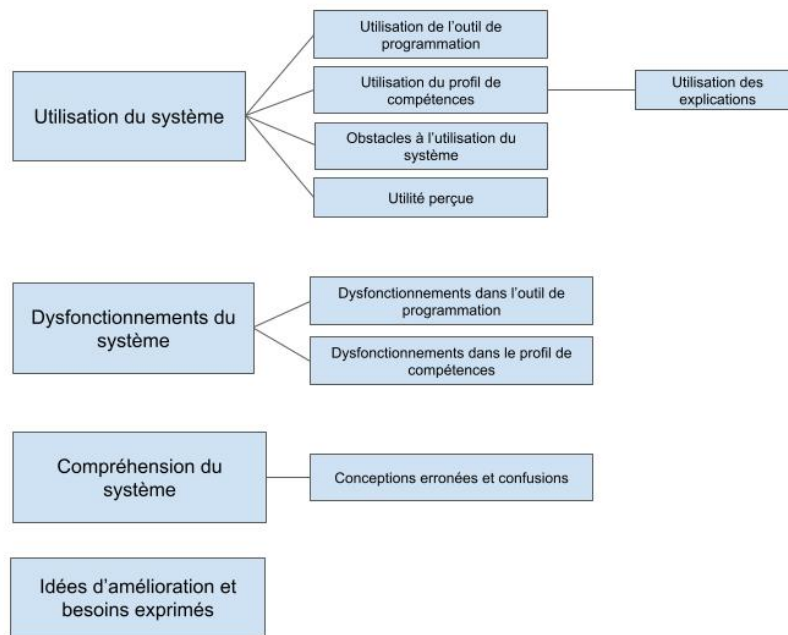


Figure 7 : Thèmes et sous-thèmes identifiés dans le focus group.

Cette tendance apparaît également lorsque l'on considère uniquement le groupe d'étudiants peu engagés qui ont reçu des explications : la corrélation forte et positive observée entre la performance sur la tâche d'apprentissage et la compréhension des explications suggère que **les étudiants peu motivés et peu performants sont moins susceptibles de comprendre les explications**. Inversement, les étudiants peu engagés mais très performants dans la tâche comprennent mieux les explications. Une fois encore, il semble que les étudiants qui comprennent le mieux les explications sont ceux qui en ont le moins besoin : les étudiants dont l'engagement sur la tâche d'apprentissage est faible, mais dont les performances sont élevées, sont des apprenants compétents qui réussissent sans avoir besoin de s'appuyer sur les explications. Ainsi, les résultats tendent à montrer que nos explications locales ne sont pas suffisantes pour la compréhension des explications par les étudiants caractérisés par un faible engagement et une faible performance sur la tâche.

En analysant les réponses des étudiants ayant une faible performance sur la tâche aux questions ouvertes du questionnaire post-expérimentation concernant les explications, différents types de demandes ont émergé. Tout d'abord, en ce qui concerne la *forme des explications*, certains étudiants demandent des « visualisations pour mieux comprendre » sur « une page dédiée à cela ». Nos résultats quantitatifs et qualitatifs tendent donc à montrer que **le format textuel ne serait pas toujours le plus adapté pour délivrer des explications**. Il reste à déterminer si tous les étudiants en difficulté bénéficieraient effectivement plus d'explications non-textuelles (par exemple avec des visualisations) que d'explications purement textuelles. Il est également possible que des explications textuelles conviennent plus ou moins bien aux apprenants en fonction d'autres caractéristiques telles que leur aisance à la lecture (*reading proficiency*) (Conati *et al.*, 2021), ou qu'à l'inverse des explications visuelles soient plus adaptées uniquement si leur aisance à comprendre des visualisations (*visualisation literacy*) est élevée (Millecamp *et al.*, 2019). Cela reste à évaluer plus précisément dans notre contexte avec des étudiants dans l'enseignement supérieur, tout en prenant en compte le fait que certaines explications peuvent plus ou moins bien se prêter à une adaptation sous forme visuelle.

Ensuite, concernant le *contenu des explications*, seuls des étudiants ayant une faible perception de leur capacité à réussir ou un faible engagement dans la tâche d'apprentissage ont exprimé dans le questionnaire post-expérimentation le besoin de plus de détails et d'exemples dans les explications : « les explications ne sont pas assez explicites » ; « [j'aimerais] plus d'exemples ». Plus précisément, ces étudiants demandent des détails sur les « erreurs qui ont empêché le niveau de maîtrise de monter ». À la lumière de ces retours, il semble intéressant de fournir aux apprenants des explications contrefactuelles ou de **combiner des explications locales avec des explications contrefactuelles**. Une explication contrefactuelle d'un résultat d'algorithme donne une indication sur ce qu'il aurait fallu faire différemment pour que l'algorithme prenne une autre décision (Wachter *et al.*, 2017). Cette idée est soutenue par plusieurs auteurs qui font l'hypothèse que les explications contrefactuelles pourraient être fournies à des utilisateurs non experts (Bove *et al.*, 2022 ; Khosravi *et al.*, 2022 ; Ribera et Lapedriza, 2019). Dans notre contexte, il s'agirait de fournir aux apprenants une explication sur les commandes qui ont généré des erreurs et qui n'ont pas contribué à élever les niveaux de maîtrise, ou de donner des exemples de commandes correctes qui contribuent à augmenter ces niveaux de maîtrise pour les aider lorsqu'ils sont bloqués. D'autre part, si l'on examine les retours obtenus dans le focus group et le questionnaire post-expérimentation par les étudiants très performants dans la tâche, deux catégories distinctes peuvent être identifiées : les étudiants qui déclarent avoir compris les explications et les avoir trouvées utiles, et ceux qui ont compris les explications mais ne les ont pas trouvées utiles. Le focus group a révélé que la première catégorie est composée d'étudiants curieux, désireux de comprendre le fonctionnement interne du système. En revanche, la seconde catégorie se compose d'étudiants qui pensent avoir déjà une compréhension approfondie de ce qui doit être fait et de la manière dont cela doit être accompli, et qui n'ont donc pas besoin d'explications. Les informations à notre disposition ne permettent pas d'identifier quelles caractéristiques de ces étudiants performants les rendent plus ou moins réceptifs aux explications. Une piste à explorer consiste à évaluer si le « besoin de cognition » (*Need for Cognition*) est une caractéristique qui permet de prédire l'intérêt ou non de ces étudiants envers les explications, comme le suggèrent certains travaux (Conati *et al.*, 2021 ; Millecamp *et al.*, 2019).

Enfin, concernant la *temporalité des explications*, les apprenants caractérisés par une faible perception de leur capacité à réussir ou un faible engagement dans la tâche d'apprentissage, expriment le besoin d'une indication plus forte du lien entre les activités réalisées dans le cadre de la tâche d'apprentissage et leur effet sur le niveau des compétences associées. Ils souhaiteraient par exemple obtenir **des notifications leur indiquant, en temps réel et en fonction des activités qu'ils réalisent sur la plateforme d'apprentissage, que leur niveau de maîtrise d'une ou de plusieurs compétences augmente**. Le fait de recevoir des notifications en temps réel est perçu comme plus motivant par les apprenants, ce qui est cohérent avec des éléments de la littérature sur le feedback dans des environnements de programmation, qui suggèrent que du feedback en temps réel impacte positivement l'engagement des étudiants (Marwan *et al.*, 2020). Dans notre contexte, au-delà des effets bénéfiques potentiels sur l'engagement, le fait de fournir des notifications en temps réel pourrait renforcer l'explicabilité de notre système en améliorant sa compréhension par les apprenants : les explications textuelles viendraient en complément d'une première représentation mentale formulée par les apprenants eux-mêmes en temps réel, à travers le lien entre les activités qu'ils réalisent et l'amélioration/la régression des niveaux de compétences en cours d'acquisition. Cependant, il est délicat de savoir à quel moment exact délivrer une explication de manière à ce que l'apprenant comprenne son contexte, mais sans le gêner dans son apprentissage. L'approche en temps réel a déjà été adoptée pour des explications dans un tuteur intelligent permettant à des apprenants d'explorer un algorithme de satisfaction de

contraintes (Putnam et Conati, 2019), et a conduit les participants de l'étude à exprimer le besoin d'explications supplémentaires sur *pourquoi* un feedback leur était délivré à ces moments précis. Des recherches complémentaires doivent donc être menées pour identifier plus précisément la forme et la temporalité les plus adaptées pour fournir des explications aux apprenants, en particulier aux apprenants en difficulté.

Ces résultats nous conduisent à formuler plusieurs **recommandations pour la conception d'explications** dans un système pour l'apprentissage qui s'appuie sur des algorithmes d'IA. Ces recommandations sont à appliquer différemment en fonction du système utilisé et des données que ce système est en capacité de récolter. En effet, il est possible de mesurer des caractéristiques initiales avec des questionnaires avant que les étudiants ne commencent à utiliser le système, et ainsi personnaliser les explications en fonction de leurs réponses. Cependant, pour les caractéristiques comportementales, la personnalisation des explications en fonction de ces caractéristiques ne peut se faire qu'au cours de l'utilisation du système, ce qui nécessite la possibilité de mettre à jour les explications régulièrement.

1. Pour des étudiants **en difficulté** (faible performance sur la tâche) ou les étudiants avec une **faible perception de leur capacité à réussir**, il vaut mieux **privilégier des explications textuelles très courtes ou des explications sous une forme autre que textuelle**.
2. Pour des étudiants **performants** (haute performance sur la tâche) ou qui **prennent le temps de lire les explications simples**, des explications **textuelles plus détaillées** peuvent être envisagées.
3. Pour tous les étudiants, il est préférable de **fournir des explications contrefactuelles ou locales et de les accompagner d'un feedback correctif**, de manière à ce que les explications soutiennent la compréhension des étudiants des attentes du système, et donc des prochaines actions pertinentes à effectuer pour leur apprentissage.

Pour conclure, notre étude n'a pas mis en avant de corrélations liées à l'intérêt dans le thème du cours ou aux connaissances préalables. Même si cela suggère que ces deux caractéristiques ne sont pas les plus importantes à prendre en compte lors de la conception d'explications, il faut rester prudent. Par exemple, certains travaux concernant des recommandations intégrant des explications (*explainable recommendations*) ont montré qu'il était important de prendre en compte les connaissances préalables (Barria-Pineda *et al.*, 2021). Dans cette étude, les étudiants avec de faibles connaissances préalables faisaient preuve d'un niveau de confiance plus élevé dans les recommandations expliquées et étaient plus disposés à prendre en compte ces recommandations.

6.2. QR2 - VARIABLES CORRÉLÉES À LA CONSULTATION ET À LA PERCEPTION DU SYSTÈME À BASE DE COMPÉTENCES

Dans l'ensemble, nous constatons le schéma suivant : les étudiants les plus performants, qui s'intéressent au cours, qui sont engagés et qui sont performants à l'examen terminal, sont également ceux qui comprennent le mieux le système et son objectif. Inversement, les apprenants peu performants comprennent moins bien le système et n'en perçoivent pas l'utilité. C'est pourtant à ces étudiants que le système devrait apporter le plus grand bénéfice, en leur permettant de prendre conscience de leurs faiblesses, de pratiquer et d'améliorer leurs niveaux de compétences. Cependant, il semble que les explications que nous avons fournies n'ont pas aidé ces étudiants à mieux comprendre, à percevoir l'utilité et à faire confiance au système par rapport aux étudiants plus performants et plus confiants. Nos résultats montrent que, pour les étudiants caractérisés par une faible performance sur la tâche, l'engagement

est corrélé à la perception de l'utilité du système à base de compétences. Ainsi, parmi les étudiants qui ont des difficultés à accomplir les tâches requises par le cours, ceux qui font le plus d'efforts pour être actifs sur la plateforme d'apprentissage sont également ceux qui perçoivent une plus grande utilité du système à base de compétences. À l'inverse, les étudiants en difficulté et peu engagés perçoivent une moins grande utilité au système à base de compétences.

Néanmoins, les résultats qui indiquent de fortes corrélations entre les perceptions de l'utilité et la compréhension des explications ainsi que la compréhension, l'utilité perçue et la confiance dans le système à base de compétences sont très encourageants. Ces corrélations pourraient s'expliquer en partie par le fait que les dimensions d'utilité et de compréhension mesurées ont pu apparaître similaires du point de vue des étudiants ayant répondu aux questionnaires. Cependant, bien qu'il faille faire preuve de prudence dans l'analyse de ces corrélations, elles suggèrent que si nous pouvons améliorer la compréhension et l'utilité perçue des explications, cela pourrait avoir un effet sur la compréhension, l'utilité perçue et la confiance dans le système, offrant ainsi une voie pour aider les étudiants qui éprouvent des difficultés avec ce système. Les comparaisons (tests t) entre les groupes avec et sans explications, bien que non significatives, montrent également une tendance encourageante à cet égard, suggérant que les explications peuvent effectivement contribuer à accroître la confiance, la compréhension et l'utilité perçue du système expliqué. Cette tendance est alignée avec de récentes études qui montrent que les explications améliorent la confiance des utilisateurs non-experts envers le système dans le contexte d'environnements pour l'apprentissage (Conati *et al.*, 2021 ; Khosravi *et al.*, 2022).

De plus, une étudiante du focus group, faisant partie du groupe contrôle et catégorisée comme peu engagée et peu performante, a fourni des retours intéressants. En effet, elle a déclaré ne pas avoir compris que les niveaux de maîtrise des compétences étaient automatiquement mis à jour lorsqu'elle travaillait sur la plateforme et a donc pensé que lorsque les niveaux augmentaient, il s'agissait d'un « bug », ce qui l'a amenée à cesser d'utiliser le système. Cela montre que le manque de compréhension du système pousse certains étudiants peu performants à s'en détourner au lieu d'en tirer profit, et nous encourage à continuer à concevoir des explications du système adaptées afin de les aider à exploiter pleinement les outils à leur disposition.

6.3. LIMITATIONS

Une première limite de notre étude concerne l'utilisation de la plateforme d'apprentissage Lab4CE et du système à base de compétences, qui aurait pu être amplifiée. En effet, si tous les étudiants ont bien utilisé Lab4CE régulièrement pendant les sessions pratiques encadrées en classe, certains étudiants ont expliqué qu'ils préféreraient utiliser des machines virtuelles sur leur ordinateur personnel plutôt que d'utiliser Lab4CE lors du travail personnel à la maison. Dans le focus group, nous avons constaté un chevauchement important entre les obstacles à l'utilisation du système à base de compétences et les dysfonctionnements de Lab4CE. Selon les discussions avec les étudiants, certaines caractéristiques de la plateforme telles que sa lenteur occasionnelle ou certaines occurrences de déconnexion, ont découragé les étudiants de l'utiliser plus régulièrement, ce qui a contribué à limiter leur accès au système à base de compétences et aux explications. De plus, les extraits de code donnés dans les explications pourraient être formatés de manière plus facile à lire pour les étudiants, notamment lors de l'affichage de programmes un peu plus longs que de simples commandes.

Une autre limite concerne l'algorithme utilisé pour calculer les niveaux de maîtrise des compétences des étudiants : les règles sur lesquelles il repose ont été définies avec l'aide d'un enseignant de la matière étudiée, mais l'algorithme n'a pas été évalué ni validé. Il est possible

qu'il fournisse des résultats parfois imprécis, et il est difficile d'évaluer l'impact que pourrait avoir l'amélioration de l'algorithme sur la perception du système à base de compétences et des explications.

Par ailleurs, une autre limite que nous avons identifiée porte sur les explications globales. Des problèmes techniques nous ont en effet empêché de prendre en compte dans notre étude le nombre de consultations de ces explications. Nous n'avons donc pas pu étudier d'éventuels effets de la consultation de cette explication globale comparativement aux explications locales. Cependant, nous pouvons émettre l'hypothèse que cette explication a été très peu consultée et n'a donc pas ou peu influencé les étudiants, car durant le focus group, aucun étudiant ayant reçu des explications ne se souvenait de cette explication globale. Des études supplémentaires sont toutefois nécessaires pour étudier plus en détail ce sujet et ainsi compléter des travaux qui ont déjà comparé la perception des utilisateurs de différents types d'explications, et qui n'ont pas validé l'hypothèse selon laquelle les novices en IA préfèrent les explications locales aux explications globales (Aechtner *et al.*, 2022).

Enfin, nous avons identifié deux limites dans notre protocole expérimental. La première porte sur la constitution du focus group. La participation des étudiants reposant sur du volontariat, nous n'avons pas pu obtenir une part égale de profils d'étudiants. Notamment, les étudiants caractérisés par un faible engagement ou de faibles performances académiques étaient insuffisamment représentés (2 étudiants uniquement sur les 8 participants). Cela a diminué le nombre de retours de la part de ce type d'étudiants, et donc limité les pistes à explorer pour répondre à leurs attentes en termes d'explications. L'autre limite concerne l'identification des connaissances préalables des étudiants à partir du questionnaire pré-expérimentation. Même si la consigne de ce questionnaire mentionnait explicitement que son objectif n'était pas d'évaluer les étudiants, un certain nombre d'entre eux a cherché les réponses aux questions posées à partir de différentes ressources disponibles en ligne afin de ne pas avoir une « mauvaise note ». Ce comportement ne nous a donc pas permis d'évaluer l'influence des connaissances préalables sur l'accès et la compréhension des explications.

7. CONCLUSION ET TRAVAUX FUTURS

Le travail présenté dans cet article est l'une des premières contributions à l'introduction de l'explicabilité dans les EIAH pour l'enseignement de l'informatique, notamment concernant les différentes caractéristiques individuelles susceptibles d'influencer la consultation, la compréhension et la perception des explications. Les analyses statistiques ont révélé des corrélations souvent fortes, voire très fortes, entre certaines caractéristiques des apprenants et leur utilisation et perception des explications ainsi que leur perception et compréhension du système. Les explications fournies ont principalement été consultées et comprises par les étudiants ayant une haute perception de leur capacité à réussir, qui sont performants, engagés et intéressés par le fonctionnement interne du système utilisé. Pour ce type d'apprenants, les explications représentent une source d'information qui peut contribuer à une amélioration supplémentaire de leurs performances. À l'inverse, les analyses tendent à montrer que les étudiants peu performants ou peu engagés dans la tâche d'apprentissage ont des difficultés à accéder aux explications et à les comprendre. Cela renforce la nécessité de leur fournir des explications personnalisées en prenant en compte les besoins spécifiques des apprenants en ce qui concerne le contenu, la forme et le moment des explications. Ces besoins comprennent la combinaison d'explications locales et contrefactuelles, l'utilisation de formats à la fois textuels et visuels, ainsi que la délivrance d'explications en temps réel de manière à renforcer la compréhension du lien entre les actions réalisées par les apprenants et le comportement du système.

En outre, nos résultats soulignent la nécessité d'examiner d'autres caractéristiques personnelles susceptibles d'affecter la perception et la compréhension des explications, afin de fournir des explications adaptées aux apprenants en fonction de ces caractéristiques.

RÉFÉRENCES

- Abdi, S., Khosravi, H., Sadiq, S., et Gasevic, D. (2020). Complementing educational recommender systems with open learner models. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 360-365. <https://doi.org/10.1145/3375462.3375520>
- Aechtner, J., Cabrera, L., Katwal, D., Onghena, P., Valenzuela, D. P., et Wilbik, A. (2022). Comparing User Perception of Explanations Developed with XAI Methods. *2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. <https://doi.org/10.1109/FUZZ-IEEE55066.2022.9882743>
- Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., Li, X., et Weegar, R. (2021). Generation of Automatic Data-Driven Feedback to Students Using Explainable Machine Learning. Dans I. Roll, D. McNamara, S. Sosnovsky, R. Luckin et V. Dimitrova (dir.), *Proceedings of the International Conference on Artificial Intelligence in Education* (p. 37-42). Springer. https://doi.org/10.1007/978-3-030-78270-2_6
- Ali, L., Hatala, M., Gašević, D., et Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education*, 58(1), 470-489. <https://doi.org/10.1016/j.compedu.2011.08.030>
- Arya, V., Bellamy, R. K. E., Chen, P.-Y., Dhurandhar, A., Hind, M., Hoffman, S. C., Houde, S., Liao, Q. V., Luss, R., Mojsilović, A., Mourad, S., Pedemonte, P., Raghavendra, R., Richards, J., Sattigeri, P., Shanmugam, K., Singh, M., Varshney, K. R., Wei, D., et Zhang, Y. (2019). One Explanation Does Not Fit All : A Toolkit and Taxonomy of AI Explainability Techniques. *arXiv*, 1909(03012). <https://doi.org/10.48550/arXiv.1909.03012>
- Ayache, M., et Dumez, H. (2011). Le codage dans la recherche qualitative une nouvelle perspective ? *Le Libellio d'AEGIS*, 7, 33-46. <https://hal.science/hal-00657490>
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., et Herrera, F. (2020). Explainable Artificial Intelligence (XAI) : Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, (58), 82-115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Barria-Pineda, J., Akhuseyinoglu, K., Želem-Čelap, S., Brusilovsky, P., Milicevic, A. K., et Ivanovic, M. (2021). Explainable Recommendations in a Personalized Programming Practice System. Dans I. Roll, D. McNamara, S. Sosnovsky, R. Luckin et V. Dimitrova (dir.), *Artificial Intelligence in Education* (p. 64-76). Springer International Publishing. https://doi.org/10.1007/978-3-030-78292-4_6
- Bennetot, A., Laurent, J.-L., Chatila, R., et Díaz-Rodríguez, N. (2019). Towards Explainable Neural-Symbolic Visual Reasoning. *arXiv.org*. <https://arxiv.org/abs/1909.09065v2>
- Bove, C., Aigrain, J., Lesot, M.-J., Tijus, C., et Detyniecki, M. (2022). Contextualization and Exploration of Local Feature Importance Explanations to Improve Understanding and Satisfaction of Non-Expert Users. *Proceedings of the 27th International Conference on Intelligent User Interfaces*, 807-819. <https://doi.org/10.1145/3490099.3511139>

- Broisin, J., Venant, R., et Vidal, P. (2017). Lab4CE : a Remote Laboratory for Computer Education. *International Journal of Artificial Intelligence in Education*, 27(1), 154-180. <https://doi.org/10.1007/s40593-015-0079-3>
- Butler, D., et Winne, P. (1995). Feedback and Self-Regulated Learning : A Theoretical Synthesis. *Review of Educational Research - REV EDUC RES*, 65, 245-281. <https://doi.org/10.2307/1170684>
- Clancey, W. J. (1983a). The epistemology of a rule-based expert system —a framework for explanation. *Artificial Intelligence*, 20(3), 215-251. [https://doi.org/10.1016/0004-3702\(83\)90008-5](https://doi.org/10.1016/0004-3702(83)90008-5)
- Clancey, W. J. (1983b). GUIDON. *Journal of Computer-Based Instruction*, 10(1-2), 8-15.
- Conati, C., Porayska-Pomsta, K., et Mavrikis, M. (2018). AI in Education needs interpretable machine learning : Lessons from Open Learner Modelling. *arXiv*. <https://doi.org/10.48550/arXiv.1807.00154>
- Conati, C., Barral, O., Putnam, V., et Rieger, L. (2021). Toward personalized XAI : A case study in intelligent tutoring systems. *Artificial Intelligence*, 298, 103503. <https://doi.org/10.1016/j.artint.2021.103503>
- Darvishi, A., Khosravi, H., et Sadiq, S. (2020). Utilising Learnersourcing to Inform Design Loop Adaptivity. Dans C. Alario-Hoyos, M. J. Rodríguez-Triana, M. Scheffel, I. Arnedillo-Sánchez et S. M. Dennerlein (dir.), *Addressing Global Challenges and Quality Education* (p. 332-346). Springer International Publishing. https://doi.org/10.1007/978-3-030-57717-9_24
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Félix, E., Amadiou, F., Venant, R., et Broisin, J. (2022). Process and Self-regulation Explainable Feedback for Novice Programmers Appears Ineffectual. Dans I. Hilliger, P. J. Muñoz-Merino, T. De Laet, A. Ortega-Arranz et T. Farrell (dir.), *Educating for a New Future : Making Sense of Technology-Enhanced Learning Adoption* (p. 514-520). Springer International Publishing. https://doi.org/10.1007/978-3-031-16290-9_44
- Fredricks, J. A., Blumenfeld, P. C., et Paris, A. H. (2004). School Engagement : Potential of the Concept, State of the Evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>
- Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., et Kagal, L. (2018). Explaining Explanations : An Overview of Interpretability of Machine Learning. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*. <https://doi.org/10.1109/DSAA.2018.00018>
- Hattie, J., et Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Hayes-Roth, B. (1985). A blackboard architecture for control. *Artificial Intelligence*, 26(3), 251-321. [https://doi.org/10.1016/0004-3702\(85\)90063-3](https://doi.org/10.1016/0004-3702(85)90063-3)
- Jørnø, R. L., et Gynther, K. (2018). What Constitutes an ‘Actionable Insight’ in Learning Analytics? *Journal of Learning Analytics*, 5(3), 198-221. <https://doi.org/10.18608/jla.2018.53.13>
- Keane, M. T., Kenny, E. M., Delaney, E., et Smyth, B. (2021). If Only We Had Better Counterfactual Explanations : Five Key Deficits to Rectify in the Evaluation of Counterfactual XAI Techniques. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 4466-4474. <https://doi.org/10.24963/ijcai.2021/609>

- Keuning, H., Jeurig, J., et Heeren, B. (2018). A Systematic Literature Review of Automated Feedback Generation for Programming Exercises. *ACM Transactions on Computing Education*, 19(1), 3 :1-3 :43. <https://doi.org/10.1145/3231711>
- Khosravi, H., Shabaninejad, S., Bakharia, A., Sadiq, S., Indulska, M., et Gašević, D. (2021). Intelligent Learning Analytics Dashboards : Automated Drill-Down Recommendations to Support Teacher Data Exploration. *Journal of Learning Analytics*, 8(3), 133-154. <https://doi.org/10.18608/jla.2021.7279>
- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., et Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education : Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Kouki, P., Schaffer, J., Pujara, J., O'Donovan, J., et Getoor, L. (2019). Personalized explanations for hybrid recommender systems. *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 379-390. <https://doi.org/10.1145/3301275.3302306>
- Kraus, S., Azaria, A., Fiosina, J., Greve, M., Hazon, N., Kolbe, L., Lembcke, T.-B., Muller, J. P., Schleibaum, S., et Vollrath, M. (2020). AI for explaining decisions in multi-agent environments. *Proceedings of the AAAI Conference on Artificial Intelligence*, 13534-13538.
- Krause, J., Perer, A., et Ng, K. (2016). Interacting with Predictions : Visual Inspection of Black-box Machine Learning Models. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5686-5697. <https://doi.org/10.1145/2858036.2858529>
- Lacave, C., et Diez, F. J. (2004). A review of explanation methods for heuristic expert systems. *The Knowledge Engineering Review*, 19(2), 133-146. <https://doi.org/10.1017/S0269888904000190>
- Leake, D., et Mcsherry, D. (2005). Introduction to the Special Issue on Explanation in Case-Based Reasoning. *Artificial Intelligence Review*, 24(2), 103-108. <https://doi.org/10.1007/s10462-005-4606-8>
- Lipton, Z. C. (2018). The Mythos of Model Interpretability : In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31-57. <https://doi.org/10.1145/3236386.3241340>
- Lundberg, S. M., et Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768-4777.
- Marwan, S., Gao, G., Fisk, S., Price, T. W., et Barnes, T. (2020). Adaptive Immediate Feedback Can Improve Novice Programming Engagement and Intention to Persist in Computer Science. *Proceedings of the 2020 ACM Conference on International Computing Education Research*, 194-203. <https://doi.org/10.1145/3372782.3406264>
- Millecamp, M., Htun, N. N., Conati, C., et Verbert, K. (2019). To explain or not to explain : the effects of personal characteristics when explaining music recommendations. *Proceedings of the 24th International Conference on Intelligent User Interfaces*. <https://doi.org/10.1145/3301275.3302313>
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. *Handbook of research on educational communications and technology*, 3, 125-144.
- Narciss, S., Huth, K., et Narciss, D. (2002). How to design informative tutoring feedback for multi-media learning. Dans H. Niegemann, R. Brünken et D. Leutner (dir.), *Instructional design for multimedia learning*.

- Naveed, S., Donkers, T., et Ziegler, J. (2018). Argumentation-Based Explanations in Recommender Systems : Conceptual Framework and Empirical Results. *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, 293-298. <https://doi.org/10.1145/3213586.3225240>
- Palominos, E., Levett-Jones, T., Power, T., et Martinez-Maldonado, R. (2019). Healthcare students' perceptions and experiences of making errors in simulation : An integrative review. *Nurse Education Today*, 77, 32-39. <https://doi.org/10.1016/j.nedt.2019.02.013>
- Pardo, A., Mirriahi, N., Martinez-Maldonado, R., Jovanovic, J., Dawson, S., et Gašević, D. (2016). Generating actionable predictive models of academic performance. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 474-478. <https://doi.org/10.1145/2883851.2883870>
- Pintrich, P. R., et Others, A. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)* (rapp. tech.).
- Putnam, V. (2020). *Toward XAI for Intelligent Tutoring Systems : a case study* [thèse de doct., University of British Columbia]. <https://doi.org/10.14288/1.0389817>
- Putnam, V., et Conati, C. (2019). Exploring the Need for Explainable Artificial Intelligence (XAI) in Intelligent Tutoring Systems (ITS). Dans C. Trattner, D. Parra et N. Riche (dir.), *Joint Proceedings of the ACM IUI 2019 Workshops co-located with the 24th ACM Conference on Intelligent User Interfaces (ACM IUI 2019), Los Angeles, USA, March 20, 2019*. <https://ceur-ws.org/Vol-2327/IUI19WS-ExSS2019-19.pdf>
- Ribera, M., et Lapedriza, À. (2019). Can we do better explanations? A proposal of user-centered explainable AI. *IUI Workshops*. <https://api.semanticscholar.org/CorpusID:84832474>
- Riefle, L., Hemmer, P., Benz, C., Vössing, M., et Pries, J. (2022). On the Influence of Cognitive Styles on Users' Understanding of Explanations.
- Rizzo, M., Veneri, A., Albarelli, A., Lucchese, C., Nobile, M., et Conati, C. (2023). A Theoretical Framework for AI Models Explainability with Application in Biomedicine. *2023 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 1-9. <https://doi.org/10.1109/CIBCB56990.2023.10264877>
- Sablayrolles, L., Guin, N., Lefevre, M., et Broisin, J. (2021). Conception et évaluation d'un processus de personnalisation fondé sur des référentiels de compétences. *10e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain*, 226-237. <https://hal.science/hal-03292896>
- Sato, M., et Tsukimoto, H. (2001). Rule extraction from neural networks via decision tree induction. *IJCNN'01. International Joint Conference on Neural Networks. Proceedings*, 3, 1870-1875 vol.3. <https://doi.org/10.1109/IJCNN.2001.938448>
- Tomsett, R. J., Braines, D., Harborne, D., Preece, A., et Chakraborty, S. (2018). Interpretable to Whom? A Role-based Model for Analyzing Interpretable Machine Learning Systems. *arXiv*. <https://doi.org/10.48550/arXiv.1806.07552>
- Tsai, C.-H., et Brusilovsky, P. (2019). Evaluating Visual Explanations for Similarity-Based Recommendations : User Perception and Performance. *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, 22-30. <https://doi.org/10.1145/3320435.3320465>
- Wachter, S., Mittelstadt, B., et Russell, C. (2017). Counterfactual Explanations Without Opening the Black Box : Automated Decisions and the GDPR. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3063289>

- Winne, P. H. (2021). Open Learner Models Working in Symbiosis With Self-Regulating Learners : A Research Agenda. *International Journal of Artificial Intelligence in Education*, 31(3), 446-459. <https://doi.org/10.1007/s40593-020-00212-4>
- Zuffianò, A., Alessandri, G., Gerbino, M., Luengo Kanacri, B. P., Di Giunta, L., Milioni, M., et Caprara, G. V. (2013). Academic achievement : The unique contribution of self-efficacy beliefs in self-regulated learning beyond intelligence, personality traits, and self-esteem. *Learning and Individual Differences*, 23, 158-162. <https://doi.org/10.1016/j.lindif.2012.07.010>