



Utilisation de tableaux de bord numériques pour l'évaluation des compétences scolaires : une étude de cas

Matthieu Cisel, Georges-Louis Baron

► To cite this version:

Matthieu Cisel, Georges-Louis Baron. Utilisation de tableaux de bord numériques pour l'évaluation des compétences scolaires : une étude de cas. Questions Vives, Université de Provence Aix-Marseille 1, Département des sciences de l'éducation, 2019. hal-03051588

HAL Id: hal-03051588

<https://hal.archives-ouvertes.fr/hal-03051588>

Submitted on 10 Dec 2020

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.

Utilisation de tableaux de bord numériques pour l'évaluation des compétences scolaires : une étude de cas

Matthieu Cisel et Georges-Louis Baron. Université Paris Descartes, Laboratoire EDA

Résumé : Au cours de la dernière décennie, deux injonctions se sont intensifiées au niveau du système scolaire français : utiliser davantage le numérique en classe d'une part, et développer « l'approche par compétences » d'autre part. Les tableaux de bord associés à des artefacts numériques ont été utilisés dans de nombreux contextes éducatifs, et se situent à la confluence de ces deux injonctions. Nous explorons au prisme de la théorie de l'activité d'Engeström les contradictions qu'induirait l'introduction d'un tel artefact dans les pratiques d'évaluation des compétences scolaires. La contribution est fondée sur une étude de cas, la conception d'une application numérique instrumentant la démarche d'investigation. Des entretiens collectifs fondés sur des maquettes sont menés avec un panel d'enseignants. Soit le coût temporel de l'appropriation du tableau de bord est jugé trop élevé pour permettre une application, soit il y a dévolution à l'artefact d'une partie trop importante du diagnostic, grâce à l'interopérabilité croissante entre outils d'évaluation. Aveugles aux caractéristiques des élèves et aux spécificités des situations didactiques, les applications risquent de produire des diagnostics aberrants.

Mots-clés : tableau de bord, traces d'interaction, approche par compétences, recherche orientée par la conception

Title in English: Using Digital Dashboards in the context of skill-based assessments in primary and secondary education: a case study

Summary: Over the course of the last decade, teachers have been increasingly pressured to use digital artefacts in the classroom, and to shift from grade-based assessment to skill-based assessment. Dashboards have been used in various educational settings; they could help bridge the gap between these two expectations. Based on Engeström's activity theory, we explore the contradictions that could rise if educational dashboards were to be used in an evaluation context. This article is based on a case study, the design of an application. Collective semi-structured interviews based on dashboards' mock-ups were conducted with a panel of teachers from both elementary and middle school. Either the additional time required to deepen the evaluation is too high to enable the use of the dashboard, or the teacher would rather save time by relying heavily on the diagnostic abilities of the application. Given that the artefact is blind to the characteristics of both the didactic situation and of the learners, the dashboard is likely to produce incoherent diagnostics.

Keywords: dashboard, learning analytics, skills' evaluation, design-based research

Introduction

Numérique et évaluation par compétences

Transition vers l'évaluation par compétences : quelles potentialités pour le numérique ?

Le développement de l'évaluation par compétences (EpC) constitue depuis plusieurs années une injonction forte (BOEN, 2007, 2016), tant au niveau du primaire que du secondaire (Guimard, 2010). La transition d'une évaluation par notes à une évaluation par compétences s'accompagne d'un certain nombre de problèmes et de mises en question de la part de praticiens et de chercheurs (Hirtt, 2009 ; Coulet, 2010). Certains auteurs y voient cependant une opportunité d'explorer les potentialités offertes par le numérique (Blais et Gilles, 2011). Une piste explorée, en particulier, est l'étayage des diagnostics réalisés par des traces collectées via des applications numériques. Des Tableaux de Bord (TDB) sont susceptibles de collecter les performances des élèves (Mottus, Graf et Chen, 2015). Cette approche est en cohérence avec le fait que les enseignants sont incités depuis plusieurs années à utiliser au sein de leurs pratiques des applications numériques, notamment dans le cadre du travail collaboratif entre élèves (BOEN, 2016, p. 94). Elle soulève la problématique suivante : Comment les enseignants se représentent-ils l'utilisation de tableaux de bord dans un contexte d'évaluation par compétences ?

Si elle ouvre la voie à des développements prometteurs, l'utilisation de tels TDB est également susceptible d'introduire dans l'activité enseignante diverses formes de contradictions, que nous nous sommes attaché à caractériser dans le contexte d'un projet d'innovation autour du développement d'un environnement informatisé pour l'apprentissage humain (EIAH). Au travers d'entretiens focalisés de groupe menés sur la base de maquettes de TDB avec des enseignants du primaire et du secondaire, nous avons cherché à mettre au jour les représentations des enseignants sur les tensions que pourrait induire une telle évolution de l'EpC. La section suivante a vocation à préciser le contexte dans lequel a émergé ce questionnement.

Une réflexion fondée sur une étude de cas : le CNEC

Nous avons exploré cette question dans le cadre d'un projet de développement du *Cahier Numérique de l'Elève-Chercheur*, ou CNEC (CDC, 2016). Le CNEC est un EIAH visant à favoriser la collaboration d'élèves autour de la rédaction d'écrits scientifiques, selon une démarche de pédagogie par projet promue par le programme Savanturiers (Cisel, Barbier et Baron, 2019). Ce dernier vise à développer des "simili-projets de recherche", principalement à l'école primaire et au collège, et encadrés par des mentors généralement issus du milieu académique (Pirone, 2018 ; Carosin et Demeuze, 2018). Il s'agit d'initier les élèves aux démarches d'investigation (Coquidé, Fortin et Rumelhard, 2009), en leur faisant jouer une part active aux différentes étapes de la démarche, de la formulation de la question de recherche à l'interprétation des résultats. Le programme s'inscrit explicitement dans une logique de renouvellement des approches

pédagogiques dans l'enseignement des sciences, mettant l'accent sur la dimension méthodologique de l'activité scientifique.

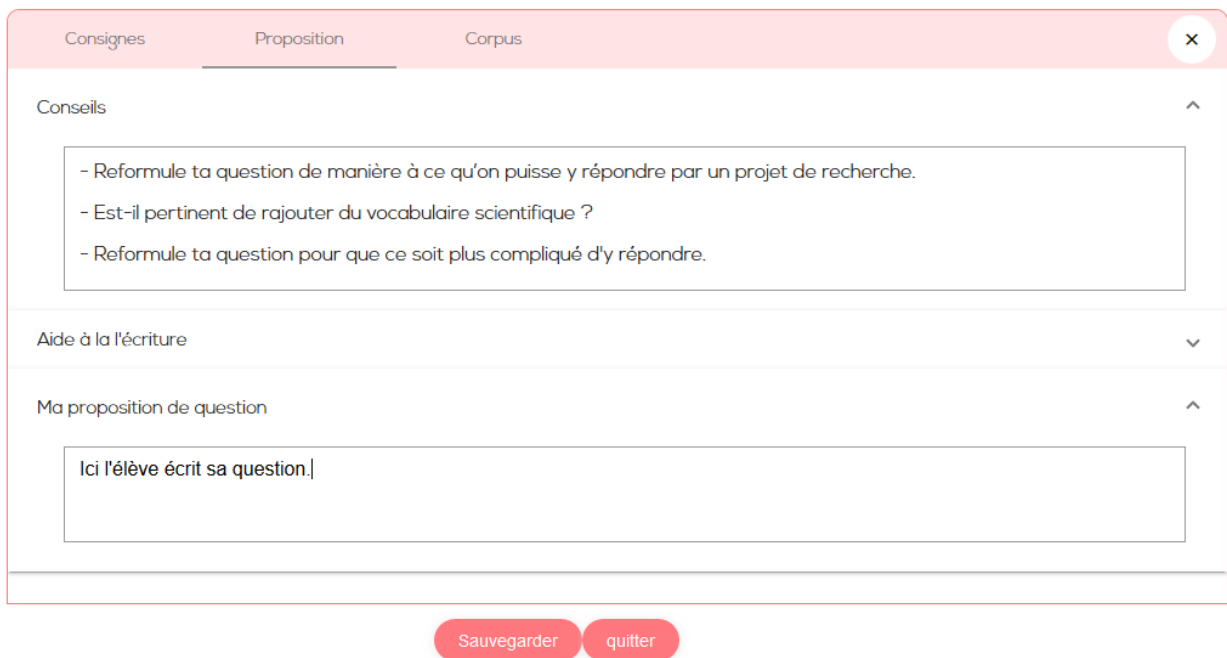
Le CNEC est ainsi susceptible d'instrumenter par des étayages l'acquisition des rudiments des démarches scientifiques - et notamment du raisonnement hypothético-déductif, ainsi qu'à développer la capacité à coopérer au sein d'un groupe pour réaliser des écrits collectifs. Chaque fois qu'un élève réalise une action avec l'artefact, des traces sont collectées. Elles peuvent donner lieu à la mise au point de profils apprenants (Jean-Daubias, 2011) visibles à partir d'un TDB, dont les usages possibles dans un contexte d'évaluation sont au cœur de la réflexion présentée dans cette contribution. L'analyse d'utilisations potentielles en contexte d'évaluation des compétences est née d'une proposition réalisée au sein du consortium.

Cet environnement s'inscrit dans la lignée de nombreux artefacts visant à instrumenter ce que l'on qualifie d'enseignement scientifique fondé sur l'investigation (Edelson, Gordin et Pea, 1999 ; Quintana, Zhang et Krajcik, 2005). Il s'agit d'une application web, développée au sein du consortium eFRAN¹ *les Savanturiers du Numérique* (CDC, 2016). Elle comporte plusieurs modules, dont la *Fiche-Recherche*, qui se situe au centre de la réflexion sur le TDB. Celle-ci permet la production d'écrits structurés autour de cinq sections reflétant les différentes étapes d'une démarche d'investigation (Pedaste *et al.*, 2015) - question de recherche, hypothèse, protocole, données, conclusion. Un élève ou un groupe d'élèves rédige une proposition pour une section donnée, et l'enseignant peut alors rédiger un commentaire, et qualifier le travail réalisé en choisissant parmi trois options : validé, à corriger, ou fermé.

Dans une interface articulée avec la Fiche-Recherche, des étayages fixes (Bonnat *et al.*, 2019), identiques pour tous les élèves mais dépendant des sections concernées (Figure 1), aident à la formulation des propositions. Chaque élève au sein d'un groupe est susceptible de formuler des propositions à partir de son compte, propositions qui sont ensuite mutualisées et articulées dans une « Fiche Recherche » commune à l'ensemble du groupe. La proposition est ensuite envoyée à l'enseignant, qui la valide ou l'invalide, et dans ce dernier cas, la renvoie aux élèves pour modification. L'idée d'utiliser à des fins d'évaluation ce processus de rédaction incrémentale d'écrits scientifiques a émergé rapidement pendant le projet, à la suite de la découverte de travaux sur l'utilisation de tableaux de bord dans le champ des méthodes d'investigation dans l'enseignement supérieur (Specht *et al.*, 2013).

¹Environnements de Formation, de Recherche, et d'Animation Numérique. Lancés dans le cadre du Plan d'investissement d'avenir 2 (PIA2) (CDC, 2016)

Figure 1 : Capture d'écran d'un des modules du CNEC, le Brouillon de recherche, visant à aider les élèves de l'école primaire et du collège à formuler des questions de recherche.



Au cours de la seconde année du projet, l'équipe de conception du consortium a produit un certain nombre de propositions en matière de TDB pour la démarche d'investigation. Ces propositions préfigurent une évolution possible, et possiblement déjà effective, des pratiques d'évaluation. L'analyse des observations de projets Savanturiers suggère que des TDB sont déjà utilisés pour instrumenter des EpC. Dans un collège parisien, nous avons ainsi pu constater qu'une enseignante de mathématiques utilisait les informations de connexion à une application dont elle avait prescrit l'utilisation, Labomep (Pilet *et al.*, 2013), pour effectuer une évaluation par compétences. Par ailleurs, des enseignants au collège attribuent des notes sur la base de TDB fournies par le Projet Voltaire, spécialisé dans l'apprentissage de l'orthographe, en fonction du niveau d'engagement de leurs élèves (Mezin, 2019). Néanmoins, l'utilisation de tels indicateurs est susceptible d'accroître de biaiser le diagnostic du praticien, soit parce que leur calcul est trop complexe à appréhender pour comprendre ce qu'ils reflètent, soit parce qu'ils sont trop simples et ne reflètent dès lors pas une compétence. On peut par exemple voir une contradiction dans le fait que des enseignants établissent des compétences à partir d'une simple connexion d'un élève à une application, ce qui ne constitue pas une performance *stricto sensu*.

Ces anecdotes attirent l'attention sur des évolutions de pratiques d'évaluation encore peu documentées dans la littérature scientifique. Elles invitent au développement de recherches portant en particulier sur les contradictions que peut introduire dans la pratique enseignante une telle utilisation des TDB, et plus généralement des traces collectées par les applications numériques. Cette réflexion nous est apparue d'autant plus importante dans le cas du CNEC que les indicateurs

envisagés étaient complexes – ils se fondaient sur des actions variées des élèves. Le propos peut néanmoins être élargi au-delà de ce seul cas d'étude et se veut plus général.

Une synthèse de travaux sur les TDB pour l'approche par compétences en éducation

L'approche par compétences et son instrumentation numérique

Comme le souligne Guimard (2010), l'approche par compétences s'est développée en France « à l'image de ce qui se fait dans bien d'autres pays (Canada, États-Unis, Suisse, Pays-Bas, Royaume-Uni, Belgique...) et dans une logique que certains n'hésitent pas à soupçonner d'être inspirée par celle qui prévaut dans les entreprises (Hirtt, 2009) ». Mobilisé déjà dans la formation d'adultes depuis plusieurs décennies, le terme « compétence » a concentré un certain nombre de critiques du fait de son caractère protéiforme. Son utilisation dans le milieu scolaire a également suscité des réflexions critiques quant aux malentendus qu'il est susceptible d'engendrer (Coulet, 2010). Malgré le flou notionnel qui l'entoure, le terme a été utilisé tant par les praticiens que dans les textes officiels (BOEN, 2007), et il est entré dans le vocabulaire courant du milieu enseignant, quand bien même les praticiens ne s'accorderaient pas sur une définition consensuelle (Guimard, 2010). Nous avons adopté dans cette contribution une définition en extension du terme *compétence*, c'est-à-dire que nous qualifierons de *compétence* tout élément qu'un enseignant désigne comme étant une compétence qu'il évalue, que cela soit officiellement, en fin de cycle, ou durant l'année, en utilisant les termes officiels ou une terminologie qui lui est propre. Nous parlons également indistinctement d'item, pour désigner un élément à remplir par l'enseignant dans les livrets de compétences.

Des outils spécifiques ont été développés pour répondre aux injonctions officielles de développer l'approche par compétences. Le Livret Scolaire Unique Numérisé (LSUN), qui a remplacé le Livret Personnel de Compétences en 2016, a ainsi vocation à suivre « les acquisitions, progrès et difficultés des élèves du CP jusqu'à la fin du collège » (BOEN, 2016). Désormais, les enseignants utilisent des logiciels de vie scolaire comme Pronotes, qui communiquent dans une large mesure automatiquement avec le LSUN eu égard aux évaluations réalisées. Les réflexions sur les modalités de représentation informatique des compétences, issues notamment de la théorie des champs conceptuels (Coulet, 2011), n'ont pas eu semble-t-il d'influence sur la manière dont ont été représentées les compétences dans les outils utilisés au quotidien par les enseignants pour). Les échelles de valuation utilisées dans ces artefacts sont généralement fondées sur quatre niveaux d'acquisition (BOEN, 2016), de non-acquis à acquis, avec des variations terminologiques autour de ces termes.

Les enseignants sont encore libres, en France, d'utiliser au quotidien soit exclusivement des notes, soit exclusivement des compétences, soit un mélange des deux approches. L'évaluation peut être associée à un devoir sur table présentant des barèmes explicites, comme être réalisée sans traces écrites, notamment lorsqu'il s'agit de compétences liées à la capacité des élèves à travailler en

groupe. Mis à part à certains jalons répartis tous les trois ans en primaire et au collège – les fins de cycles où l'EpC est obligatoire, les enseignants peuvent choisir d'évaluer les compétences de leur choix, en utilisant les formulations officielles, ou s'écarter des items officiels des programmes et créer des items à évaluer *ad hoc*. Là où des auteurs comme Sayac (2017) se penchent sur les pratiques évaluatives effectives des enseignants, nous nous intéressons, dans le cadre de cette contribution, à la manière dont les pratiques d'EpC pourraient évoluer à travers l'utilisation de TDB, qui font l'objet d'une littérature grandissante, dont les paragraphes qui suivent présentent des éléments de lecture.

Les tableaux de bord en éducation

Nous avons sommairement divisé en deux groupes les travaux sur les TDB en éducation publiés au cours de la dernière décennie, entre ceux qui appartiennent avant tout au champ de l'informatique, et ceux qui relèvent des sciences humaines. Si la synthèse des avancées récentes dans le domaine dépasse le périmètre de cet article, nous évoquerons quelques travaux notables afin de préciser dans quelle lignée de travaux nous nous inscrivons par la présente contribution. Le premier axe, correspondant à une approche informatique, est centré sur les méthodes de collecte et de traitement des traces d'apprentissage laissées par les apprenants. Cet axe a connu des développements significatifs dans la littérature francophone, dont on peut trouver des synthèses dans des recherches doctorales récentes (Dabbebi, 2019).

S'agissant des travaux mobilisant des approches relevant davantage des sciences humaines, nous pouvons distinguer plusieurs axes de recherche dans la littérature scientifique (Schwendimann *et al.*, 2016), ces différents axes étant poursuivis de manière simultanée dans certains travaux (Iandoli, Quinto, De Liddo et Buckingham Shum, 2014). L'un d'entre eux consiste à établir, au moment de la conception, des liens entre théories de l'apprentissage et choix des indicateurs affichés dans le TDB (Sedrakyan *et al.*, 2018). Le deuxième axe d'intérêt relève de l'évaluation des TDB, et s'inscrit plus généralement dans le champ de l'évaluation des EIAH (Jamet, 2006). Il vise à évaluer l'acceptabilité, l'utilité, et/ou l'utilisabilité (Nielsen, 1993 ; Tricot *et al.*, 2003) de TDB développés ou en cours de développement (Ali *et al.*, 2012). Nous y incluons les recherches fondées sur la mise à l'épreuve en classe de TDB fonctionnels (Faber, Luyten et Visscher, 2017), ou sur des maquettes de TDB (Stephens-Martinez, Hearst et Fox, 2014 ; Scheffel *et al.*, 2017).

L'avantage d'une telle approche est de permettre des ajustements en amont de tout développement informatique. Nous pouvons distinguer au moins deux fonctions possibles des TDB : d'une part l'orchestration, ou la régulation, de l'activité de la classe (Verbert *et al.*, 2014), sans diagnostic sur les apprentissages, notamment pour des interventions pour identifier un désengagement (Hu, Lo, et Shih, 2014), d'autre part une logique d'aide au diagnostic humain, qui correspond à la démarche menée au sein du consortium. Cette approche est présente dans Pépidiag (Jean-Daubias, 2002), un EIAH qui permet d'aider les enseignants de mathématiques à diagnostiquer les compétences des élèves en calcul à partir d'indicateurs fondés sur des traces laissées lors de la réalisation d'exercices par les élèves. On la trouve aussi dans les travaux sur Renkan (Carrillo *et al.*, 2018) ; dans cette

dernière recherche, les auteurs décrivent le processus de conception d'un TDB visant à suivre les étapes de construction d'une carte mentale. Réalisant des entretiens basés sur des maquettes, ils mettent au jour certaines des réticences des praticiens quant à l'utilisation de traces laissées par les élèves, notamment en matière d'évaluation de l'engagement de l'élève.

Nous nous inscrivons dans la lignée de ces travaux, mais en nous focalisant ici sur l'évaluation des compétences scolaires, telles que formulées dans les programmes officiels. La question exacte à laquelle nous apportons des éléments de réponse dans cet article est la suivante : Quelles sont les représentations d'enseignants innovants inscrits dans le programme Savanturiers, quant à l'utilisation de TDB à des fins d'évaluation des compétences scolaires au cours de projets fondés sur des démarches d'investigation ? Pour la traiter, nous avons mobilisé la théorie de l'activité d'Engeström (1987), que nous présentons succinctement dans les paragraphes qui suivent.

Présentation de la théorie de l'activité

Présentation de la théorie de l'activité et définition du système observé

La théorie de l'activité, qui a connu une diffusion certaine dans les dernières décennies, représente un cadre interdisciplinaire trouvant ses origines dans les travaux soviétiques en psychologie du développement, notamment Vygotski et Leontiev. Au fil de ses formes successives, elle s'est développée pour devenir ce que l'on nomme parfois la théorie de l'activité de troisième génération (Engeström, 1987, 2000). Mobilisée par ses créateurs dans des contextes variés – dont les transformations des établissements scolaires (Engeström, 2008) – notre terrain d'étude, elle vise à appréhender les actions individuelles au prisme du système d'activité dans lesquelles elles s'inscrivent. Elle est largement utilisée dans le champ de l'éducation, et plus particulièrement dans le champ des technologies éducatives, de la formation à distance (Dir et Simonian, 2015), ou de la conception d'environnements informatiques (Bourguin, 2000 ; Bourguin, Derycke et Tarby, 2005).

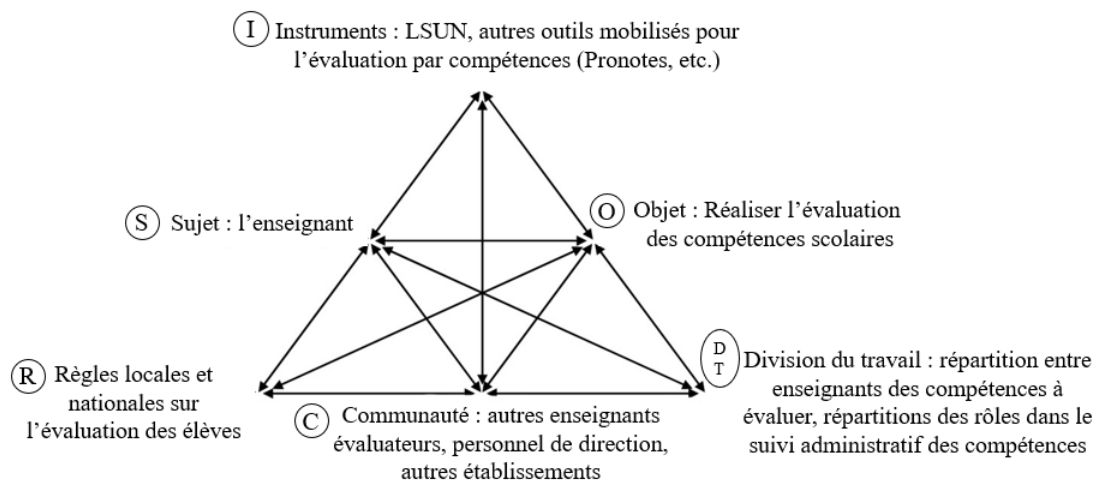
Engeström (1987) représente l'activité par un modèle empirique constitué de six pôles en interaction, qui constituent alors le système d'activité (SA) (Figure 2) : le sujet de l'activité (S), son objet (O), les règles qui la régissent (R), la communauté (C), la division du travail (DT), et les instruments (I), qui constituent la focale de notre réflexion. Dans notre cas, le sujet de l'activité, l'enseignant, réalise une action sur l'objet de l'activité, l'évaluation des compétences des élèves, au moyen d'instruments : devoirs et cahiers des élèves, grilles de notation, et, dans notre cas, des tableaux de bord.

L'objet de l'activité qui nous intéresse est l'évaluation par les enseignants des compétences développées par les élèves, dans les projets Savanturiers, et ce aussi bien du point de vue de l'acquisition du raisonnement scientifique, que de celui des compétences collaboratives. Le sujet est l'enseignant en position d'évaluateur, qui s'inscrit dans une communauté, que nous définissons ici comme les autres enseignants qui contribuent à l'évaluation des compétences des élèves. La

situation est sensiblement différente en primaire et dans le secondaire, mais il existe un invariant : pour un cycle donné, l'évaluation des compétences n'est généralement pas effectuée par un enseignant isolé. Au collège, des enseignants de plusieurs disciplines peuvent ainsi évaluer la même compétence au cours d'une année donnée, en particulier sur un item aussi transversal que la capacité à collaborer. Au niveau du cycle primaire, c'est généralement sur plusieurs années que la compétence est évaluée, mais le principe de division du travail entre enseignants reste le même.

S'agissant du pôle Division du travail, se pose alors la question de la répartition des tâches entre enseignants, et avec le personnel administratif. Les règles correspondent aux législations et règles locales en vigueur concernant l'évaluation des compétences. Celle-ci est obligatoire en fin de cycle, soit tous les trois ans, notamment à des fins d'orientation au cours de l'année de troisième. Dernière règle, l'évaluation doit être légitimée, et être aussi peu arbitraire que possible.

Figure 2 : Schématisation fondée sur le modèle d'Engeström (1987) du système d'activité représentant l'évaluation des compétences telle que pratiquée aujourd'hui, sans TDB



Enfin, s'agissant du pôle instrument, celui-ci est composé de l'ensemble des artefacts médiateurs mobilisés lors de la phase d'évaluation : le logiciel dans lequel l'enseignant entre les évaluations réalisées (Pronotes², Charlemagne, etc.), les grilles de notation éventuellement utilisées, les supports portant d'éventuelles productions des élèves, et, dans le cas du présent article, un éventuel tableau de bord. Nous représentons le modèle de l'activité ainsi défini dans la Figure 2. Dans la prochaine section, nous revenons sur les contradictions associées.

² <https://www.index-education.com/fr/logiciel-gestion-vie-scolaire.php>

Théorie de l'activité et analyse des contradictions

Un système d'activité évolue, notamment du fait d'innovations technologiques, ou de changements réglementaires, et comporte des contradictions porteuses de transformations. La recherche de dépassement des tensions associées notamment à ces évolutions constitue l'un des mécanismes à l'origine de l'évolution de l'activité. Engeström distingue plusieurs niveaux de contradictions. Les contradictions primaires correspondent aux tensions au sein d'un pôle donné. Par exemple, s'agissant du pôle instrument, différents artefacts peuvent entrer en tension lorsqu'ils sont mobilisés de manière concomitante pour évaluer les compétences des élèves en matière d'acquisition de la démarche scientifique, notamment si l'enseignant utilise les cahiers individuels des élèves et un cahier rassemblant les productions de l'ensemble des membres d'un groupe. Les contradictions secondaires découlent des tensions entre deux pôles, tandis que les ternaires correspondent aux tensions entre anciens et nouveaux éléments d'un pôle donné. Les quaternaires correspondent aux tensions entre différents systèmes d'activité.

Nous analyserons aussi bien les tensions qui traversent le système d'activité avant l'introduction du TDB que les contradictions ternaires qui apparaissent potentiellement, entre ancien et nouveau système d'évaluation. La mise en évidence de telles contradictions implique généralement une observation directe du système d'activité, et/ou des entretiens avec les sujets de l'activité. Nous décrivons dans la section qui suit, consacrée à la méthodologie, l'approche qui a été suivie dans le cadre de cette recherche.

Méthodologie

Les enseignants partenaires du consortium

Le consortium qui développe le CNEC, *Les Savanturiers du Numérique*, rassemble le programme Savanturiers, porteur du projet, un industriel qui a développé l'application, des enseignants partenaires intervenant via leurs académies, et une équipe de chercheurs (Cisel, Barbier et Baron, 2019). Le programme Savanturiers, né en 2013 sous l'impulsion d'une ancienne professeure des écoles, fait partie des acteurs qui bénéficient de cette dynamique. Selon une logique de conception participative commune dans le champ des EIAH (Frascara, 2003 ; Jean-Daubias, 2004), des enseignants partenaires du programme et issus de l'enseignement primaire et secondaire ont été intégrés au consortium. Nous utiliserons des pseudonymes pour les désigner. Leur expérience de la pédagogie par projet était variable, certains appartenant aux Savanturiers depuis plusieurs années, d'autres se lançant pour la première fois dans la démarche (Cisel et Baron, 2018). Tous avaient une pratique de l'évaluation par compétences au cours de l'année, rendue obligatoire par les textes officiels. La majorité des enseignants partenaires se saisissaient de l'opportunité que représentaient ces projets pour évaluer des compétences scolaires.

Les enseignants ont accordé des entretiens et ouvert leurs classes aux chercheurs. On peut distinguer deux périodes ; la première recouvre l'essentiel de l'année scolaire 2016-2017, elle correspond à des observations de situations de classe sans le CNEC (Cisel et Baron, 2018). Nous avons alors recensé les situations didactiques mises en place lors de l'activité de classe, et collecté les instruments mobilisés par ces enseignants pour réaliser l'évaluation des compétences scolaires. La seconde période, qui débute en juin 2017, comporte des tests utilisateurs en classe et des discussions de maquette au cours d'entretiens collectifs. Des versions successives ont été testées en conditions écologiques jusqu'aux derniers mois du consortium (Cisel, Barbier et Baron, 2019). Nous nous sommes positionnés, dans le cadre de cette recherche, dans le paradigme de la Recherche Orientée par la Conception (Sanchez et Monod-Ansaldi, 2015), ou *Design-Based Research* dans la littérature anglo-saxonne (Amiel et Reeves, 2008), nous attachant sur les éléments du processus de conception les plus susceptibles de produire des résultats nouveaux du point de vue de la recherche.

Mise au point des maquettes du Tableau de Bord

En parallèle des premiers tests utilisateurs en classe, le chantier de la conception du Tableau de Bord a été lancé au cours de la seconde année du consortium. Les propositions de différents membres du consortium ont été confrontées, tant en termes d'indicateurs qu'en termes d'utilisations possibles de ces indicateurs. La question de l'évaluation des compétences revenant de manière récurrente dans les discussions au sein du consortium, nous avons alors décidé de nous intéresser aux représentations des enseignants sur ce sujet.

Les indicateurs présentés dans le Tableau 1 sont fondés sur les traces d'interaction laissées par les élèves lors de la navigation avec la plateforme.

« Compétence « évaluée par l'enseignant	Indicateur fourni par le Tableau de Bord du CNEC
Proposer une ou des hypothèses pour répondre à une question ou à un problème	Nombre d'hypothèses proposées et validées par l'enseignant
	Nombre d'itérations nécessaires avant validation de l'hypothèse par l'enseignant
Proposer des expériences simples pour tester une hypothèse	Nombre d'itérations nécessaires avant validation du premier protocole par l'enseignant
	Nombre d'itérations du protocole faisant suite à une collecte de données

Tableau 1 : Exemples d'indicateurs fondés sur des traces collectées par le CNEC et associées à des compétences du socle commun

L'association de ces indicateurs avec des items évalués dans le cadre d'une EpC est une proposition du programme Savanturiers. Ainsi, l'item « Proposer une ou des hypothèses simples pour répondre à une question ou à un problème » – inscrit aux cycles 3 et 4 – est relié dans le TDB au nombre d'hypothèses proposées via la Fiche-Recherche et d'itérations réalisées sur l'hypothèse avant que celle-ci ne soit définitivement validée par l'enseignant.

Compte tenu des débats qu'ont suscités ces approches au sein du consortium, il a été jugé préférable de travailler avec des maquettes avant de réaliser des développements informatiques. Celles-ci ont été présentées aux enseignants dans le cadre d'entretiens collectifs, qui constituent la base empirique de cette contribution, et dont les modalités d'organisation sont présentées dans la section qui suit. Les propositions de tableaux de bord ainsi réalisées ont été soumises au panel d'enseignants partenaires au cours de *focus groups* (Krueger, 2014)³, dont nous précisons la méthodologie dans la section suivante.

Présentation de maquettes au cours de focus groups

Dix-huit réunions collectives impliquant les enseignants partenaires ont été organisées au cours de la période 2016-2019, lors de ce que l'on nomma les Comités de Conception (Cocon). Ces Cocons

³ Si le terme entretiens focalisés de groupe est courant dans la littérature francophone, le terme anglophone *focus group* est plus usité, et sera celui que nous utiliserons dans cette contribution.

ont eu pour fonction de déterminer les grandes orientations technologiques du projet, puis de recueillir les représentations des enseignants sur les propositions réalisées par les différents membres du consortium. Le rôle des chercheurs était double dans ce type de réunion : noter les propos des enseignants pour en réaliser une médiation auprès de l'industriel, et collecter des données pour produire des résultats de recherche.

Six Cocons ont porté spécifiquement sur la question des tableaux de bord, et ont donné lieu à des focus groups d'une heure chacun environ, chacun réunissant entre quatre et huit enseignants issus des cycles primaire et secondaire (Cisel, Barbier et Baron, 2019). L'intérêt de l'approche par focus group est de permettre la confrontation entre points de vue de praticiens aux perspectives distinctes (Krueger, 2014). Leur animation était à la charge du chercheur, et fondée sur un canevas d'entretien, ce dernier incluant des questions comme : Dans quelle mesure utiliseriez-vous un tel Tableau de bord pour réaliser des évaluations par compétences ? Quels sont selon vous les obstacles à son utilisation dans un tel contexte ? Les enseignants partenaires étaient familiers avec l'artefact du fait des tests en classe, mais nous préférons, pour clarifier les propositions effectuées, présenter en amont des modules concernés, qui étaient visualisés au tableau. Les propositions réalisées par la consultante étaient ensuite projetées et discutées avec l'ensemble des enseignants présents. Ce n'est qu'après s'être assuré que l'ensemble des indicateurs était bien compris que le focus group débutait. Les questions étaient formulées de sorte à mettre au jour la diversité des positions des enseignants vis-à-vis des indicateurs proposés. Les verbatims ont été enregistrés puis retranscrits dans leur intégralité.

Résultats

La présentation des résultats qui suit est organisée en trois segments. En premier lieu, nous reviendrons sur le coût temporel de l'EpC, et en particulier sur le gain de temps que les enseignants recherchent à travers de nouveaux instruments. Des enseignants voient le TDB comme une opportunité de dépasser des contradictions qui traversent le système d'activité dans sa forme actuelle, liées aux injonctions administratives à mettre en œuvre une EpC avec des instruments non nécessairement adaptés. Mais ce gain de temps présente néanmoins un risque pour certains, s'il conduit à dévoluer aux applications l'essentiel du processus d'évaluation.

Nous reviendrons ensuite sur les réticences qu'ils expriment quant aux indicateurs quantitatifs fondés exclusivement sur les actions des élèves, qui peuvent conduire à ignorer les spécificités des situations didactiques au sein desquelles l'application a été utilisée. Ils soulignent notamment le risque que les élèves, une fois qu'ils ont appréhendé le fonctionnement de l'application, adaptent leur comportement pour biaiser l'évaluation en leur faveur. Dans un dernier temps, nous mettrons au jour la distinction que les enseignants établissent entre compétences eu égard à la pertinence d'une instrumentation numérique. Pour présenter les contradictions, nous aurons recours à la modélisation en triangle présentée dans le cadre théorique, représentant les contradictions par des flèches de couleur reliant des pôles de différents systèmes d'activité (Figure 3).

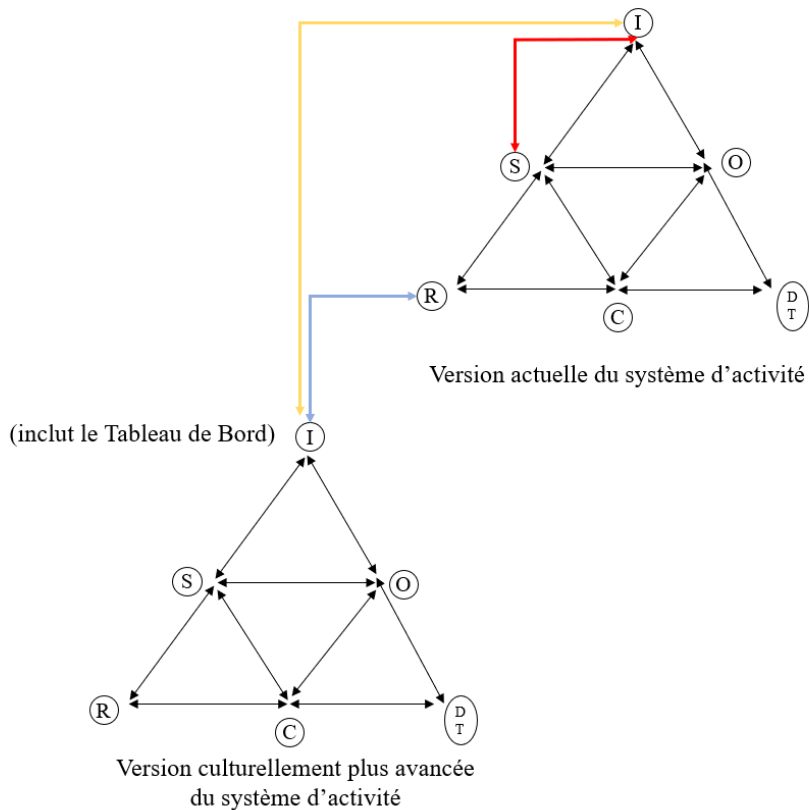
Le TDB : un gain de temps potentiel, mais à quel prix ?

L'évolution des cadres réglementaires a imposé depuis quelques années aux enseignants de réaliser des évaluations de compétences à chaque fin de cycle, auxquelles s'ajoutent les évaluations réalisées de manière facultative. Cette obligation est vécue par les enquêtés comme une charge supplémentaire, qui ne devrait pas leur incomber, et qui vient s'ajouter aux injonctions régulières à utiliser le numérique en classe (BOEN, 2016). Ainsi, le système d'activité dans sa forme actuelle porte pour les enseignants une tension, que nous avons identifiée comme une contradiction secondaire entre le pôle Sujet, et le pôle Instruments (Flèche Rouge, Figure 3). La contradiction tient dans le fait que le rôle qu'on impose à l'enseignant via les outils comme le LSUN est en tension avec celui de pédagogue – la question de l'identité professionnelle relève du pôle sujet. Ils doivent jouer un rôle d'administration alors que l'EpC aurait davantage pour eux vocation à faire évoluer leurs pratiques pédagogiques.

Plusieurs des enseignants enquêtés souhaiteraient que le TDB permette de dépasser partiellement cette contradiction secondaire en les déchargeant de cette tâche ressentie comme plus administrative que pédagogique. Eloïse, enseignante de primaire, déclare ainsi : « *Ce que j'aimerais c'est que l'application remplisse automatiquement le LSU. Après, je repasserais derrière bien sûr* ». Valérie, enseignante de physique-chimie au collège, tient un discours analogue. Dans la mesure où l'outil qu'elle utilise pour effectuer l'évaluation par compétences, Pronotes, communique directement avec le LSU, elle souhaiterait que le TDB puisse en faire autant : « *Est-ce que je vais l'utiliser si le CNEC il dialogue pas avec le LSU, pas sûr. J'ai trop pris l'habitude en fait de me connecter direct à Pronotes, et de pas avoir besoin de te connecter au LSU* ». Cela suppose une interopérabilité entre les applications et le LSU. Si elle faciliterait le travail de l'enseignant, le risque que pose l'interopérabilité entre outils, est souligné par Matéo, enseignant de SVT en collège :

Si ton logiciel par exemple, [...] s'il va dialoguer direct avec le LSU, là c'est dangereux. Parce que du coup tu te soumetts au fait que certaines personnes vont se dire, mes évaluations vont être transmises au LSU, et il ne va même pas vérifier dans le LSU au fait.

Figure 3 : Une schématisation fondée sur le modèle d'Engeström (1987) des contradictions induites par l'introduction du Tableau de Bord dans une pratique d'évaluation par compétences



En somme, si le TDB ne présente qu'un jeu d'indicateurs que l'enseignant doit interpréter avant de réaliser manuellement ses évaluations par compétences, c'est une charge supplémentaire qui nuira sensiblement à l'acceptabilité de la proposition. Mais si une économie de temps est permise, le risque est de laisser aux applications une responsabilité jugée trop conséquente dans le diagnostic des compétences. Dans ce dernier cas, nous interprétons cette tension comme une contradiction ternaire (Flèche bleue, Figure 3) entre le pôle Instruments du SA dans sa forme plus évoluée, et le pôle Règles du SA dans sa forme actuelle. Il y a contradiction dans la mesure où l'évaluation des compétences scolaires, au regard notamment du poids qu'elle peut avoir dans l'orientation des élèves, doit être le moins arbitraire possible ; or l'application introduit une part d'arbitraire dans la mesure où elle est aveugle aux spécificités des situations didactiques. Plus généralement, nous allons voir que les enseignants considèrent que des indices subtils que ne peuvent détecter les EIAH sont plus utiles que les TDB, ce qui va amener certains à avoir une position différenciée selon le type de compétences dont l'évaluation pourrait être instrumentée.

Un instrument dont l'acceptabilité dépend de l'item évalué

L'intérêt des praticiens pour une instrumentation via le TDB varie selon les items à évaluer. Ainsi, Mateo, enseignant de SVT en troisième, voit l'intérêt pour l'évaluation d'items liés à la maîtrise

de la démarche scientifique. Il déclare ainsi : « *Que le logiciel donne une information sur un positionnement d'hypothèse, avec des critères auxquels nous on a accès, c'est intéressant* ». En revanche, des items comme la capacité à travailler en groupe sont exclus par certains enseignants du périmètre des compétences dont l'évaluation est utilement instrumentée par le numérique. C'est ce qu'exprime Anne-Laure, enseignante au primaire :

J'évalue de manière systématique la capacité à collaborer [...] Par contre je ne pense pas que ça soit possible de le faire avec un outil numérique. Je sais en les regardant comment ils collaborent, et c'est une évaluation que je fais sur plusieurs mois. Je n'ai pas besoin d'un tableau de bord pour ça.

Ce témoignage suggère l'existence d'une variation dans l'intensité des contradictions qu'engendrerait l'introduction d'un tableau de bord. Nous pouvons interpréter cette tension comme une contradiction ternaire entre le pôle Instruments du nouveau système d'activité, et le pôle Instruments de l'ancien système (Flèche jaune, Figure 3). Plus précisément, il y a contradiction dans la mesure où l'on imposerait à l'enseignant d'utiliser des outils pour des tâches qui selon lui n'imposent pas d'effectuer un suivi précis des performances des élèves. Le TDB entrerait en contradiction avec les outils qu'ils mobilisent déjà pour établir leur diagnostic – par exemple la participation différentielle à des rendus collectifs pour évaluer la capacité à collaborer.

Dans la dernière partie de la section *Résultats*, nous allons voir que les réticences des enseignants sont d'autant plus marquées que les indicateurs du TDB se basent exclusivement sur des actions réalisées par les élèves sur la plateforme.

De la réticence à se baser sur des indicateurs fondés exclusivement sur des actions d'élèves

La proposition un moment faite de créer un score agrégé pour une compétence donnée, à partir d'indicateurs quantitatifs résultant des actions des élèves sur la plate-forme, a engendré deux types de critiques de la part des enseignants enquêtés. Mateo estime que l'élève, s'il comprend le principe de fonctionnement de l'évaluation, sera enclin à adapter son comportement en conséquence : « *Soumettre 28 idées, si l'élève a compris qu'il allait être évalué là-dessus, il va te soumettre 28 fois n'importe quoi* ».

Il évoque ainsi implicitement la stratégie dite de « *gaming the system* », détourner le système en français (Baker *et al.*, 2008). Avec des indicateurs simples, comme la connexion à l'application, ou le nombre d'actions réalisées, il existe un risque significatif que les élèves s'approprient les stratégies permettant de manipuler à peu de frais leur évaluation. Quand bien même l'élève n'aurait pas l'intention de détourner le système à son profit, une approche de la conception d'indicateurs qui serait fondée uniquement sur les actions de l'élève serait aveugle aux spécificités de la situation didactique, et aux caractéristiques des élèves considérés. Le problème est magnifié par la distance temporelle qui sépare parfois la situation observée de l'évaluation, lorsque celle-ci n'est pas fondée

sur un devoir écrit. Mateo ajoute ainsi, en réaction à la suggestion d'incorporer des éléments liés au nombre de connexions à l'application pour retravailler une production :

Un élève il peut se connecter une fois et faire quelque chose de parfait. Parce que lui il a pas nécessairement une attitude avec les TICE qui est bonne, il préfère interagir avec son propre cahier ou sa feuille. En fait il va faire toute sa démarche sur sa feuille, il va se connecter une fois pour faire son travail final. Mais tu vas pas pouvoir le mettre expert, parce que sur un des critères que le logiciel il te file, ça a été, il y a qu'une seule itération de connexion.

Il suggère que les indicateurs présents dans le TDB doivent se fonder sur des décisions d'enseignants, et non des actions des élèves. L'enseignant propose ainsi de penser le TDB comme un prolongement d'artefacts déjà mobilisés, comme des grilles de notation, et de renoncer à se fonder exclusivement sur les actions des élèves. Une telle approche revient selon nous à produire une redondance entre le TDB et des outils déjà mobilisés par les praticiens, et ce faisant à créer potentiellement des contradictions ternaires entre les pôles instruments du nouveau et de l'ancien système d'activité (Flèche jaune, Figure 3). Cette réticence à se fonder sur des actions d'élèves et la préférence pour la construction d'indicateurs fondés sur ses propres actions fait écho à certains résultats de l'enquête de Carrillo *et al.* (2018). Dans la discussion, nous confronterons les résultats issus de ces différents travaux, après avoir discuté des limites du dispositif de recherche.

Discussion

Limites du dispositif de recherche

Nous avons mené une recherche exploratoire sur un échantillon réduit d'enseignants réunis au sein d'un projet d'innovation, ce qui limite la portée de nos résultats ; nous ne pouvons ni prétendre à la représentativité de notre échantillon et nous n'avons pas collecté de manière exhaustive les représentations des enseignants. Une autre limite de l'enquête présentée ici est la nature du support présenté pour collecter des données : les maquettes, qui imposent aux enseignants de se projeter dans des situations imaginaires.

Nous avons pu constater par exemple qu'un élément important, identifié au cours des observations de classe, n'est pas apparu dans les focus groups : le problème de l'identité de l'utilisateur qui laisse les traces sur lesquelles sont bâtis les indicateurs présents dans les TDB. Les appareils, tablettes et ordinateurs, circulent d'un élève à l'autre même si les consignes l'interdisent, phénomène auquel l'application est aveugle. De même, la pertinence des indicateurs proposés est particulièrement sensible au respect des consignes données par les enseignants. Des mises en situation en classe permettraient de compléter utilement notre travail, en offrant aux praticiens l'opportunité de prendre conscience de problèmes qui leur échapperaient dans le contexte d'une réflexion fondée exclusivement sur des maquettes. Elles permettraient de jeter un regard nouveau

sur ce qui nous semble constituer le point commun entre les différents résultats que nous avons obtenus : la crainte d'une forme de déqualification.

Interpréter les résultats au prisme de la crainte d'une déqualification

Dans sa forme actuelle, où l'EpC est davantage vécue par les enquêtés comme une activité administrative que pédagogique, ce qui fait écho aux critiques formulées dans la littérature (Hirtt, 2009 ; Coulet, 2010). Si les enseignants voient dans le TDB l'opportunité de résoudre certaines contradictions qui traversent le système d'activité actuel, les propos rapportés reflètent selon nous la crainte d'être placé dans un rôle d'exécutant. En effet, nous avons vu en prime abord qu'il existe un risque que l'enseignant, une fois l'évaluation transmise automatiquement au LSUN par exemple, ne revienne plus sur les suggestions de l'EIAH. Un tel fonctionnement pose pour le praticien la question du risque du désengagement du travail de médiation, dans le sens où même si l'opportunité d'éditer et de corriger ces évaluations existe, il est probable qu'elle ne soit pas saisie faute de temps ou d'intérêt.

Ensuite, nous avons retrouvé la crainte que de trop vastes prérogatives ne soient attribuées à l'EIAH, qui viendrait instrumenter l'évaluation de compétences qui constituent leur pré carré. Enfin, dans la mesure où l'application est aveugle aux spécificités de la situation didactique, l'utilisation des d'indicateurs fondés sur les traces d'interaction laissées par l'activité de l'élève rend perplexe les praticiens. Une remarque analogue avait été rapportée dans les travaux de Carrillo *et al.* (2018), lorsqu'on présenta aux praticiens des indicateurs décrivant quantitativement les caractéristiques de la carte mentale : « *Ici il n'y a que de la quantité et l'enseignant s'intéresse qu'à la qualité* ». Implicitement, ils déclarent être les seuls à même de juger de la qualité, ce qui fait écho au fait que dans le cas du CNEC, ils aimeraient fonder les indicateurs sur leurs actions, davantage que celle des élèves.

Si le risque existe que ceux des élèves qui ont compris le système en détournent le fonctionnement, cela constituerait selon nous une opportunité manquée que de penser le TDB comme un prolongement d'outils d'évaluation déjà existants, en minimisant l'importance des traces d'activité laissées par des élèves. Des outils comme Pronotes utilisés quotidiennement par de nombreux enseignants vont déjà dans cette direction et permettent d'archiver des évaluations d'enseignants, et la valeur ajoutée d'une proposition de TDB qui irait en ce sens serait discutable. Néanmoins, une utilisation pertinente des traces des élèves présente un coût, qui est au cœur même de la tension qu'induirait l'utilisation de TDB dans des pratiques d'EpC : pour que le TDB dispose d'une valeur ajoutée en termes d'aide au diagnostic humain, l'enseignant doit investir davantage de temps qu'il ne le fait actuellement dans l'étape d'évaluation, ce qui est en contradiction avec le désir exprimé de gagner du temps.

En effet, le TDB ne joue son rôle d'aide au diagnostic humain qu'à la condition que le praticien s'investisse dans l'interprétation des données, ce qui suppose une compréhension du processus de collecte et de traitement des traces, et une remémoration des caractéristiques des situations

didactiques au cours desquelles les traces ont été collectées. Ce dernier point est d'autant plus problématique que l'évaluation des compétences, lorsqu'elle ne se fonde pas sur des traces écrites, intervient généralement plusieurs semaines voire plusieurs mois après les séances correspondantes.

Il est possible dès lors que ce soit dans des situations de classe où les effectifs sont particulièrement réduits qu'une telle utilisation de TDB puisse être possible. Cela exclut alors les utilisations en collège, où les enseignants ont parfois des centaines d'élèves. Cela pourrait expliquer le fait que dans la littérature, les seuls exemples de TDB que nous ayons trouvés quant à l'aide au diagnostic humain portent sur des prototypes et des expérimentations à petite échelle – avec Pépidiag par exemple (Jean-Daubias, 2002). Toutes les utilisations à grande échelle (Faber, Luyten et Visscher, 2017) relèvent de la régulation de l'activité de la classe (Michel, Lavoué et Petriac, 2012), ou de l'autorégulation par les élèves eux-mêmes (Ji, Michel, Lavoué et Georges, 2013 ; Park et Jo, 2015), probablement car l'utilisation du TDB dans des classes avec des effectifs chargés est trop compliquée à mettre en œuvre si l'on se place dans une perspective d'EpC.

Conclusion

En guise de conclusion, nous proposons d'ouvrir sur les perspectives du développement d'outils visant à automatiser les diagnostics de compétences, réduisant la part de diagnostic humain. Il existe ainsi des « tuteurs intelligents » permettant d'évaluer automatiquement les capacités des élèves à mener des démarches d'investigation (Gobert *et al.*, 2013), et qui vont un pas plus loin que des EIAH, comme Pépidiag (Jean-Daubias, 2002) ne se réclamant que de l'aide au diagnostic humain. Si de telles approches venaient à se démocratiser, on peut spéculer sur les probables réticences qu'entraînerait l'introduction d'une technologie dont la fonction est de se substituer au diagnostic humain. Ce problème n'est pas sans rappeler les débats sur les mutations du métier d'enseignant, qu'ont entraîné dans les années 1960 les expérimentations relatives aux machines à enseigner qui fournissent aux élèves des chemins d'enseignement préprogrammés (Bruillard, 1997). Ces débats reviennent de manière cyclique chaque fois que l'apparition d'une technologie paraît menacer la marge de manœuvre des praticiens. Comprendre plus finement les enjeux de l'amélioration des techniques de diagnostic automatique est nécessaire si l'on veut éviter une polarisation du débat qui risquerait sans doute de se produire entre promoteurs et détracteurs de ce type d'approche.

Références

Ali, L., Hatala, M., Gašević, D. et Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers et Education*, 58(1), 470–489.
<https://doi.org/10.1016/j.compedu.2011.08.030>

Amiel, T. et Reeves, T. C. (2008). Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda. *Journal of Educational Technology et Society*, 11(4), 29–40

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A. et Koedinger, K. (2008). Why students engage in “gaming the system” behavior in interactive learning environments. *Journal of Interactive Learning Research*, 19(2), 185-224.

Barbier, C. (2019). *Vers l'appropriation de nouveaux instruments par des enseignants : Le cas de la démarche d'Éducation par la Recherche et du Carnet Numérique de l'Élève-Chercheur*. Thèse de Master non publiée. Université Paris-Descartes, laboratoire EDA.

Blais, J.-G. et Gilles, J.-I. (éds) (2011). *Évaluation des apprentissages et technologies de l'information et de la communication : Le futur est à notre porte*. Québec : Presses de l'université Laval.

Bonnat, C., Marzin-Janvier, P., Girault, I. et d'Ham, C. (2019). Modélisation didactique pour la conception d'étayages dans un EIAH : exemple d'une activité de conception expérimentale en biologie. *STICEF*, 25.

Bruillard, E. (1997). *Les machines à enseigner*. Paris : Hermès.

Bulletin Officiel de l'Éducation Nationale (2007). *Livret Personnel de Compétences*, décret n° 2007-860 du 14-5-2007 JO du 15-5-2007. Bulletin officiel n° 22.

Bulletin Officiel de l'Éducation Nationale (2016). *Évaluation des acquis scolaires des élèves et livret scolaire, à l'école et au collège*, décret n° 2015-1929 du 31-12-2015 JO du 3-1-2016. Bulletin officiel n° 3.

Bourguin, G., Derycke, A. et Tarby, J. C. (2005). Systèmes Interactifs en Co-évolution Réflexions sur les apports de la Théorie de l'Activité au support des Pratiques Collectives Distribuées. *Revue d'Interaction Homme-Machine*, 6(1).

Bourguin, G. (2000). *Un support informatique à l'activité coopérative fondé sur la Théorie de l'Activité: le projet DARE*. Manuscrit de thèse non publié, Université Lille 1. Récupéré à l'adresse <http://www-lisic.univ-littoral.fr/publis/1394700486.pdf> (consulté le 1^{er} décembre 2019)

Caisse des Dépôts et des Consignations (2016). Les Savanturiers du Numérique. Récupéré à l'adresse http://www.caissedesdepots.fr/sites/default/files/medias/projet_13_dp_cdc.pdf (consulté le 1^{er} décembre 2019)

- Carosin, E. et Demeuse, M. (2018). *Les Savanturiers - Rapport d'évaluation final* (p. 110). Récupéré à l'adresse https://les-savanturiers.cri-paris.org/wp-content/uploads/2018/11/recherche_rapport-vf.pdf (consulté le 1^{er} décembre 2019)
- Carrillo, R., Renaud, C., Prié, Y. et Lavoué, É. (2018). Suivre l'engagement des apprenants dans l'activité de construction de cartes mentales. *STICEF*, 25 (1).
- Cisel, M. et Baron, G.-L. (2018, juin). Conception d'un EIAH à destination du programme Savanturiers : difficultés engendrées par une approche inductiviste de la spécification des besoins. Dans Boulc'h, L, Voulgre, E (Eds.), *Actes de la 3^{ème} conférence Ecole et Technologies de l'Information et de la Communication (ETIC3)*, Paris, France.
- Cisel, M., Barbier, C. et Baron G.-L. (2019). *Rapport scientifique de synthèse de la recherche Savanturiers du Numérique*. Manuscrit non publié. Université Paris-Descartes, laboratoire EDA.
- Coquidé, M. Fortin, C. et Rumelhard, G. (2009). L'investigation : fondements et démarches, intérêts et limites. *ASTER*. 49, 51-78.
- Coulet, J.-C. (2010). La « référentialisation » des compétences à l'école, conceptions et mises en œuvre. Des attendus aux malentendus. *Recherche et formation*, 64, 47-62.
- Coulet, J. C. (2011). La notion de compétence : un modèle pour décrire, évaluer et développer les compétences. *Le travail humain*, 74(1), 1-30.
- Dabbebi, I. (2019, à paraître). *Conception de tableaux de bords dynamiques, adaptatifs et contextuels*. Manuscrit de thèse non publié. Université du Mans. <http://www.theses.fr/s140053>
- Dir, M. et Simonian, S. (2015). Analyse de l'activité tutorale dans un organisme privé de formation à distance selon le modèle d'Engeström : tensions et écarts au prescrit. *Distances et médiations des savoirs*, 11.
- Edelson, D. C., Gordin, D. N. et Pea, R. D. (1999). Addressing the Challenges of Inquiry-Based Learning Through Technology and Curriculum Design. *Journal of the Learning Sciences*, 8(3-4), 391-450.
- Engeström, Y. (1987). *Learning by expanding*. Helsinki: Prienta-Konsultit Oy.
- Engeström, Y. (2000). Activity Theory as a framework for analyzing and redesigning work. *Ergonomics*, 43(7), 960-974.
- Engeström, Y. (2008). Weaving the texture of school change. *Journal of Educational Change*, 9(4), 379.

Faber, J. M., Luyten, H. et Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers et Education*, 106, 83–96.

Gobert, J. D., Pedro, M. S., Raziuddin, J. et Baker, R. S. (2013). From Log Files to Assessment Metrics: Measuring Students' Science Inquiry Skills Using Educational Data Mining. *Journal of the Learning Sciences*, 22(4), 521–563.

Mezin, G. (2019). *Le Projet voltaire dans le secondaire. Une étude exploratoire des modalités de médiation par les enseignants*. Mémoire de Diplôme Université non publié. Université Paris-Descartes. <https://halshs.archives-ouvertes.fr/halshs-02383589>

Guimard, P. (2010). *L'évaluation des compétences scolaires*. Rennes : Presses Universitaires de Rennes.

Hirtt, N. (2009). L'approche par compétences : une mystification pédagogique. *L'école démocratique*, 9, p. 1-34.

Hu, Y.-H., Lo, C.-L. et Shih, S.-P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, 36, 469-478.

Iandoli, L., Quinto, I., De Liddo, A. et Buckingham Shum, S. (2014). Socially augmented argumentation tools: Rationale, design and evaluation of a debate dashboard. *International Journal of Human-Computer Studies*, 72(3), 298–319.

Jamet, E. (2006). Une présentation des principales méthodes d'évaluation des EIAH en psychologie cognitive. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, 13.

Jean-Daubias, S. (2002). Un système d'assistance au diagnostic de compétences en algèbre élémentaire. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, 9(1), 171-199.

Jean-Daubias, S. (2004). *De l'intégration de chercheurs, d'experts, d'enseignants et d'apprenants à la conception d'EIAH*. Rapport de Recherche. Université de Technologie de Compiègne, 290-97. Récupéré à l'adresse <https://edutice.archives-ouvertes.fr/edutice-00000708/document> (consulté le 1^{er} décembre 2019).

Jean-Daubias, S. (2011). *Ingénierie des profils d'apprenants*. Université Claude Bernard - Lyon I, Habilitation à diriger des recherches. Manuscrit non publié.

Ji, M., Michel, C., Lavoué, E. et George, S. (2013, Juillet). An architecture to combine activity traces and reporting traces to support self-regulation processes. Dans *2013 IEEE 13th International Conference on Advanced Learning Technologies* (pp. 87-91). IEEE. Pékin, Chine.

Krueger, R. A. (2014). *Focus groups: A practical guide for applied research*. Thousand Oaks : Sage publications.

Michel, C., Lavoué, E. et Pietrac, L. (2012, September). A dashboard to regulate project-based learning. In *European Conference on Technology Enhanced Learning* (pp. 250-263). Springer, Berlin, Heidelberg.

Mottus, A., Graf, S. et Chen, N.-S. (2015). Use of dashboards and visualization techniques to support teacher decision making. Dans *Ubiquitous Learning Environments and Technologies* (pp. 181-199). Berlin : Springer Berlin Heidelberg

Nielsen, J. (1993). *Usability Engineering*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Park, Y. et Jo, I.-H. (2015). Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science*, 21, 110-133.

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T. et Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14, 47-61.

Pirone, F. (2018) « Les Savanturiers de la sociologie ». De la recherche en « éducation par la recherche ». *Diversité*, 192.

Pilet, J., Delozanne, E., Chenevotot, F., Grugeon, B., Prévit, D. et El-Kechai, N. (2013). *Les outils Pépite sur LaboMeP: Identifier des besoins d'apprentissage des élèves pour réguler l'enseignement*. Rapport de Recherche, Université Paris-Diderot.

Quintana, C., Zhang, M. et Krajcik, J. (2005). A Framework for Supporting Metacognitive Aspects of Online Inquiry Through Software-Based Scaffolding. *Educational Psychologist*, 40(4), 235–244.

Sanchez, É. et Monod-Ansaldi, R. (2015). Recherche collaborative orientée par la conception. *Education et didactique*, 9(2), 73–94.

Sayac, N. (2017). Étude des pratiques évaluatives en mathématiques de 25 professeurs des écoles français : Une approche didactique à partir de l'analyse des tâches données en évaluation. *Mesure et évaluation en éducation*, 40(2), 1-31.

Scheffel, M., Drachsler, H., Toisoul, C., Ternier, S. et Specht, M. (2017). The Proof of the Pudding: Examining Validity and Reliability of the Evaluation Framework for Learning Analytics. Dans *Data Driven Approaches in Digital Education* (pp. 194–208). Springer, Cham.

Schwendimann, B. A., Rodriguez Triana, M. J., Prieto Santos, L. P., Shirvani Boroujeni, M., Holzer, A. C. et Gillet, D. (2016, avril). Understanding learning at a glance: An overview of learning dashboard studies. Dans *Proceedings of the sixth international conference on learning analytics et knowledge* (pp. 532-533). Edimbourg, Ecosse.

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S. et Kirschner, P. A. (2018, sous presse). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2018.05.004>

Specht, M., Bedek, M., Duval, E., Held, P., Okada, A. et Stefanov, K. (2013, juin). WESPOT: Inquiry based learning meets learning analytics. In *Proceedings of the The Third International Conference on e-Learning* (pp. 15-20). Belgrade.

Stephens-Martinez, K., Hearst, M. A. et Fox, A. (2014). Monitoring MOOCs: Which Information Sources Do Instructors Value? In *Proceedings of the First ACM Conference on Learning @ Scale Conference* (pp. 79–88). New York, NY, USA: ACM. <https://doi.org/10.1145/2556325.2566246>

Tricot, A., Plécat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G. et Morcillo, A. (2003, juin). Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH (pp. 391–402). Dans *Actes de la conférence Environnements Informatiques pour l'Apprentissage Humain 2003*. Strasbourg, France.

Verbert, K., Govaerts, S., Duval, E., Santos, J., Van Assche, F. et Parra, G. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*, 18(6), 1499-1514.