

C2i2e à distance : l'activité de l'étudiant est-elle représentative de son évaluation?

Recherche scientifique avec données empiriques

Résumé

Cet article rend compte d'une dynamique d'innovation consistant à mettre en œuvre un dispositif de certification à distance pour de très grands effectifs. Destinée à des publics ayant pour l'essentiel expérimenté des dispositifs de formation classiques, cette solution est à même de traiter l'ensemble des participants à égalité de traitement, tout en gérant un dispositif pour de très grands effectifs. À partir de travaux sur l'étude des traces d'étudiants à distance, cet article tente de caractériser les éventuels croisements entre implication et performance.

Mots clés

Certification via Internet, C2i2e, Participation des étudiants, Interaction, Formation des enseignants

Abstract

This paper examines a dynamic innovation: a distance certification system that can handle very large numbers of students. Designed for students who are already familiar with traditional educational structures, it treats all participants equitably while managing a large-scale database. Based on the research on student logs in distance education, potential interactions between student involvement and performance are characterized.

Key words

Certification via Internet, student participation, interaction, teacher training



©Auteur(s). Cette œuvre, disponible à http://ritpu.ca/IMG/pdf/RITPU_v09_n01-02_101.pdf, est mise à disposition selon les termes de la licence Creative Commons Attribution - Pas de Modification 2.5 Canada : <http://creativecommons.org/licences/by-nd/2.5/ca/deed.fr>

Introduction

Comment s'assurer que les enseignants utiliseront les technologies de l'information et de la communication pour l'éducation (TICE) dans leurs pratiques futures? Depuis les tentatives des années 50 qui cherchaient à déployer le cinéma et le son comme vecteurs d'enseignement dans les classes (Jaillet, 2008), le musée des initiatives pédagogiques est rempli de toutes les approches visant à éloigner l'enseignant du bureau ou du tableau et à enrichir son attitude magistrale par l'entremise des technologies (Depover, 2009). Cela ne veut pas dire que celles-ci ne sont pas parées d'un ensemble de vertus liées à leurs caractéristiques, mais c'est moins celles-ci qui comptent que le fait de permettre à l'enseignant de changer de gestes professionnels et donc de conceptions pédagogiques ou de l'obliger à le faire. La réponse classique est celle de la formation. Mais après des années de tentatives d'à peu près toutes les formules possibles sans résultats notables (Chaptal, 2003), de nouvelles mesures tentent d'infléchir les pratiques en reprenant l'idée selon laquelle l'évaluation va surdéterminer les pratiques. Après avoir fait intégrer les pratiques des technologies dans les curricula des formations d'enseignants, il s'est dès lors agi de faire certifier les compétences liées à l'usage des technologies afin d'être titularisé dans le métier d'enseignant. Ainsi, le certificat Informatique et Internet niveau 2 enseignant (C2i2e) est requis en même temps que d'autres conditions. Cela pose évidemment des problèmes de tous ordres et ne se passe pas sans difficulté. Pour une part, l'exigence d'une certification supplémentaire vient rajouter à la crise dans la formation des enseignants en France à la suite de la maîtrise des formations. L'Université de Cergy-Pontoise (UCP) a cherché à utiliser cette contrainte nouvelle pour mettre en œuvre une dynamique originale autour de la certification. Celle-ci se déroule intégralement à distance pour près de 1 000 étudiants. Ils sont répartis en groupe de 16 et disposent d'un tutorat qui n'a pas pour mission de former les étudiants, mais de s'assurer qu'ils ont bien saisi les consignes liées à la certification. Le dispositif repose sur l'évaluation par les pairs. Le

fait de réaliser l'ensemble du processus à distance permet de disposer de l'intégralité des traces des comportements des étudiants en regard notamment des interactions qui sont encouragées et des résultats qui sont produits. Peut-on relever des concordances entre la participation des étudiants au dispositif et leur évaluation?

Contexte

La difficulté à intégrer les TICE dans les pratiques éducatives est un constat générique dans la quasi-totalité des pays dont on cherche sans fin les possibles martingales (Albero, 2010). Les plans d'équipement des écoles, les opérations lourdes de distribution d'ordinateurs portables ou de tablettes n'ont fait bouger que légèrement les lignes (Blandin, 2010). À côté des approches matérielles, l'État a mis en place successivement plusieurs initiatives de formation et d'accompagnement des enseignants. Suivant les initiatives de quelques universités (cartable électronique, établissement scolaire virtuel...) (Jaillet, 2004), une nouvelle approche a conduit l'État à déployer le concept d'environnement numérique de travail. Le principe de ces solutions informatiques est assez simple. Partant du constat que toutes les sphères professionnelles du commerce, de l'industrie et de la santé disposaient à présent de logiciels intégrés qui assurent l'ensemble de la gestion des flux d'information, il fallait doter les établissements scolaires de dispositifs de ce genre. Autrement dit, il s'agissait de faire en sorte que les enseignants, les parents, l'administration et les élèves disposent d'un logiciel intégré qui regroupe toutes les informations liées à l'école. On espérait ainsi naïvement qu'à côté des usages liés à l'administration et à la scolarité, des fonctionnalités plus en lien avec la pédagogie amènent progressivement les enseignants à modifier leurs pratiques. Malheureusement, les politiques publiques se sont surtout préoccupées de questions techniques et de compatibilité des flux d'information relevant de la scolarité et de l'administration qui se formalisent par le schéma directeur des espaces numériques de travail (SDET, ministère de l'Éducation nationale,

de la Jeunesse et de la Vie associative, 2011). En conséquence, les sociétés privées incitées à venir prendre part à ce nouveau marché ont rapidement signé la fin des innovations possibles en repositionnant l'essentiel des fonctions accessibles aux enseignants dans les limites de ce qui était attendu et donc loin de la pédagogie. Dans le même temps, l'État français a mis en place une politique de certification des étudiants. Fondé sur le fait que les solutions d'enseignement ne suffisent pas, le principe consistait à développer les démarches de vérification des compétences. Cela a débuté en 2001 par le brevet informatique et Internet (B2i, ministère de l'Éducation nationale, de la Jeunesse et de la Vie associative, 2009), qui s'adresse aux élèves lors de leur scolarité initiale, puis par le certificat Informatique et Internet (c2i.education.fr), qui s'adresse aux étudiants, et le certificat de compétences en langues de l'enseignement supérieur (CLES, certification-cles.fr) en plusieurs niveaux. Ce socle de base, qui considère que des étudiants ayant suivi un cursus d'enseignement supérieur doivent être autonomes et compétents pour ce qui est des outils informatiques qui nous entourent, s'est vu complété avec le C2i niveau 2. Celui-ci se focalise sur les compétences informatiques liées aux futurs domaines professionnels. Ainsi, il existe un C2i niveau 2 pour les juristes, les ingénieurs, etc. Il existe donc un C2i niveau 2 pour les enseignants, le C2i2e. Les étudiants doivent valider 23 des 27 compétences que comporte le C2i2e pour être certifiés et il est à présent obligatoire de le détenir pour être titularisé. Le terme obligatoire est à relativiser. En effet, le certificat pour les langues l'était également, mais devant la pression des syndicats d'enseignants et des organisations d'étudiants, l'État a fait marche arrière en considérant que le fait d'avoir suivi des enseignements en langues étrangères était équivalent à ce certificat. Si ce n'est donc qu'un moratoire peut venir hypothéquer la généralisation de l'obtention de ce C2i2e par l'ensemble des futurs enseignants, une relative frénésie s'est emparée de l'ensemble des universités qui préparent au métier d'enseignant. Jusqu'à présent, cette certification était intégrée dans des cours, avec des observa-

tions lors des stages sur le terrain. Pour l'université considérée, celle de Cergy-Pontoise, l'ensemble du processus se déroule de plus dans un système de contraintes multiples. Tout d'abord, l'Institut universitaire de formation des maîtres (IUFM), qui porte la formation des futurs enseignants, a décidé il y a quelques années de ne plus faire d'enseignements spécifiques aux TICE. Ces enseignements sont donc intégrés aux enseignements disciplinaires qui doivent préparer les étudiants à ces compétences. Par ailleurs, la récente mastérisation des formations des futurs enseignants a conduit les IUFM à nourrir la difficile ambition de les préparer à la fois à un métier, avec des compétences et des gestes professionnels, et à un concours, avec tous les aspects liés au bachotage. Comme la situation n'était pas encore suffisamment simple, le concours n'est pas placé avant la formation ou après la formation, mais en même temps que la formation. Tous ces éléments concourent à créer une situation de stress important pour ces étudiants qui se voient imposer une épreuve de plus dans ce qu'ils considèrent être un calvaire.

Problématique

La situation de base consiste à former plus de 1 000 étudiants par an, sous contraintes de temps et de budget. La solution élaborée se base sur un dispositif à distance à partir de l'expertise menée en enseignement à distance depuis 12 ans par la formation UTICEF, devenue ACREDITE, portée par l'UCP, Genève et Mons, puis étendue à d'autres formations par l'UCP (Jaillet, 2004). Ceci porte le nombre d'étudiants formés à distance par l'UCP à près de 1 250 étudiants. Le C2i2e est une certification et, à l'inverse des pratiques anciennes ou de celles qui ont été mises en œuvre dans d'autres universités, il a été décidé de dissocier la formation et la certification. En effet, précédemment, les enseignants réalisaient un enseignement lié aux technologies et évaluaient les compétences mises en œuvre dans la foulée, éventuellement lors de visites d'étudiants en situation professionnelle. L'essentiel se rapportait à des gestes et à des pratiques d'enseignants en situation magistrale. Comme

la certification n'était pas une dimension déterminante de la formation des futurs enseignants, les exigences n'étaient pas très élevées. La garantie de l'unicité de l'évaluation n'était pas assurée et les critères étaient à géométrie variable et flous. L'habilitation de délivrance demandée au ministère a donc consisté à différencier la partie formation et la partie certification. Comme pour le permis de conduire, celui qui forme n'est pas celui qui évalue. Par conséquent, les enseignants sont normalement en charge de la formation des étudiants dans leurs enseignements qui doivent intégrer les technologies. Une enquête interne en direction des enseignants a été conduite en début d'année, pour tenir compte de leurs pratiques pour la certification. Dix pour cent des enseignants interrogés ont répondu. Puis, le dispositif de certification intégralement à distance s'est mis en œuvre. Les étudiants en situation de certification sont inscrits dans des groupes de travail de 16 personnes. Chacun est encadré par un tuteur dont la charge ne consiste pas à les former aux compétences visées par la certification, mais à assurer leur compréhension des consignes, à repositionner le travail fourni par les uns et les autres, à inciter les étudiants à la collaboration et à rappeler les échéances et l'organisation nécessaire. L'une des clés du dispositif consiste à réaliser le travail d'évaluation par les pairs, selon des modalités issues d'un travail de recherche sanctionné par une thèse (Bouzidi et Jaillet, 2009), c'est-à-dire que les travaux sont évalués par les étudiants eux-mêmes. Chaque étudiant doit corriger trois travaux et le sien, alors que le tuteur évalue lui aussi tous les étudiants de son groupe. La problématique traitée dans cet article réfère à la massification d'un dispositif à distance. Comment traiter et accompagner de très grands effectifs à partir de principes qui relèvent de l'artisanat? Deux aspects sont traités dans le format de cet article : l'un a trait à la caractérisation des pratiques d'étudiants à distance, l'autre aux éventuels liens entre ces pratiques et les résultats d'évaluation. Cela nous conduit à resserrer nos questions de recherche à partir de données quantitatives : Quelles sont les pratiques d'utilisation du dispositif à distance? Quelles sont les interactions à l'œuvre? Sur le plan du positionnement des étudiants, quel

lien faire entre la participation et les résultats de l'évaluation?

Corpus et méthodologie

L'étude porte sur la moitié de l'effectif engagé, soit 450 étudiants qui font partie de la première vague de certification. Près de 426 d'entre eux vont au bout de la première vague. Ils relèvent de huit formations qui mènent à un master d'enseignement. Il s'agit de formations qui préparent aux fonctions de professeur des écoles (instituteur), de professeur des collèges et lycées ou de conseiller principal d'éducation (qui assure le fonctionnement de la vie scolaire).

Méthodologiquement, il s'agit d'observer si l'activité des étudiants dans le dispositif de certification peut être reliée à différentes catégories de variables (Djouad, Mille, Reffay et Benmohamed, 2009). Autrement dit, peut-on mettre en évidence la plus ou moins grande implication de l'étudiant dans le processus et son origine de formation ou bien l'équipe dans laquelle il travaille? Au final, les résultats d'évaluation et ses actions sont-ils également corrélables? Pour ce faire, en référence à de précédents travaux sur l'activité des étudiants sur la base de traces dans les dispositifs de formation à distance (Jaillet, 2009), il s'agit de les traiter pour avoir une représentation de leur implication par le triplet d'activité (Jaillet, 2005). Celui-ci s'élabore avec des traces de nature complémentaire pour avoir une représentation réaliste de ce que font les usagers. Raisonner à partir des données brutes ne peut permettre de construire une représentation efficiente (Soller, Martínez, Jermann et Muehlenbrock, 2005). Considérer uniquement le temps ou le nombre de connexions aboutit à de très grands biais (Santos, Rodríguez, Gaudioso et Boticario, 2003). C'est le cas pour un étudiant qui reste connecté une journée entière à une plateforme sans jamais l'utiliser, ou un autre qui réalise des successions importantes de connexions sans être davantage actif. Le triplet d'activité a donc été forgé pour articuler différentes données de sorte que le construit obtenu

soit le reflet de ce que fait effectivement l'utilisateur. À partir de ces données, on peut avoir une bonne représentation de l'activité d'apprentissage et d'interaction. On peut concevoir plusieurs types de triplets d'activité pour préciser soit le niveau d'implication dans les scénarios d'apprentissage, soit le niveau d'interaction. Le principe d'élaboration du triplet consiste à croiser plusieurs types de traces. Celles-ci permettent d'obtenir des données sur différents registres d'actions de l'utilisateur. Par exemple, on mesure l'assiduité par le nombre de connexions à la plateforme, à l'espace de travail ou à l'espace de clavardage. On mesure l'implication par le temps passé dans ces mêmes endroits et l'on mesure la participation au scénario d'apprentissage par l'addition des traces d'actions, dépôts de documents, consultations, etc., ou encore, on mesure l'interaction en relevant le nombre de messages ou parfois le nombre de caractères par message. Le triplet est avant tout un instrument de suivi. Autrement dit, il permet de positionner l'activité d'un étudiant par rapport aux autres étudiants qui travaillent avec lui dans le même groupe en délivrant ses scores en pourcentage de celui qui a eu l'activité la plus forte. Les données du triplet s'envisagent donc toujours selon trois types de croisements. Par exemple, l'étudiant X qui est caractérisé par un triplet d'« implication » 58, 25, 40 a réalisé 58 % des connexions de celui de son groupe qui s'est connecté le plus, il l'a fait en 25 % du temps de connexion de celui de son groupe qui y a passé le plus de temps et a réalisé 40 % d'actions sur la plateforme de celui de son groupe qui en a réalisé le plus dans le même temps. Il obtient un score général de 123. Autrement dit, le premier croisement est un positionnement de l'individu par rapport à celui qui est le plus actif dans la catégorie. On utilise la référence du plus fort score afin d'éviter l'effet d'écrasement de la moyenne. Le deuxième type de croisement se fait entre les trois catégories. Cela permet de vérifier l'équilibre des scores. Ainsi, dans cet exemple, il ne sert à rien de rester connecté un très long moment, voire de se connecter souvent, si dans le même temps, on ne réalise pas les actions associées à l'engagement supposé.

Le troisième type de croisement se réalise par la comparaison des scores des différents triplets généraux des étudiants entre eux. On additionne les trois valeurs de chaque triplet pour obtenir une valeur représentative du niveau d'ensemble. Les étudiants fortement impliqués auront des scores au-delà de 200, les étudiants faiblement impliqués en dessous de 50. Les écarts sont donc très visibles.

Ainsi, le triplet permet de caractériser le profil de l'étudiant. Le taux de connexion important (58 %), la faible durée de connexion (25 %) et une activité d'un peu moins de la moitié de celui qui agit le plus (40 %) laissent présager un étudiant impliqué, qui va à l'essentiel en rationalisant ses interventions de façon très concentrée. Un triplet 80, 80, 5 rendrait compte d'un étudiant qui passe beaucoup de temps à faire croire qu'il est actif, mais dont le 5 permet d'observer qu'il l'est très peu par rapport à ce qui lui est demandé.

Pour manipuler le triplet d'activité, on utilise le logiciel original fondé sur les travaux de recherche depuis plus de 10 ans en enseignement à distance E-Space, développé par une équipe de recherche-développement (innovation, formation, technologies numériques) (IFTN, www.iftn.fr) conjointe de l'Université de Strasbourg et de l'UCP. Ce logiciel permet de proposer aux usagers une plateforme d'enseignement à distance qui est basée sur des principes d'organisation et de scénarisation pédagogique qui reprennent les travaux réalisés sur l'ergonomie de l'enseignement à distance à partir d'une analogie à un campus réel (Jaillot, 2004). Dans le logiciel E-Space, des fonctions de supervision avancée permettent de visualiser les interactions et les implications pour chaque étudiant et groupe d'étudiants.

Les données dans leur ensemble sont prises en compte pour ce qu'elles sont : connexions, durées, activités, etc., soit en triplet. Elles sont ensuite croisées avec des données catégorielles : genre, formation, modalités de formation et évaluation. Le but de cette investigation consiste à évaluer si les données récoltées permettent d'expliquer les variations selon les différentes caractéristiques catégorielles. Ainsi peut-on repérer un effet groupe de

travail dans l'activité qui est déployée, ce qui pourrait conduire à se demander s'il y a un effet tuteur qui inciterait plus ou moins les étudiants à travailler ensemble, à communiquer. Y a-t-il un effet formation d'origine? Autrement dit, peut-on observer que les étudiants de certaines formations sont plus impliqués, assidus que d'autres? Pour les étudiants qui suivent un parcours à distance, sont-ils plus ou moins impliqués que les autres, eux qui ont déjà la maîtrise des processus à distance? Enfin, il est peu probable que l'on puisse ramener à une équation simple l'implication dans le dispositif et dans les interactions à distance et les notes obtenues à l'évaluation, mais il est intéressant de tenter de repérer les variables qui y seraient sensibles. Autrement dit, est-ce que dans certains groupes le travail en commun, comparativement au travail individuel, se corrèle avec les résultats de l'évaluation?

Résultats

Tout d'abord, on prend en compte les données qui seront utiles pour le triplet d'activité consacré à la participation à la certification des 426 étudiants :

- La durée de connexion dans le séminaire E-Space dans toutes ses fonctionnalités (moyenne : 22 heures; écart-type : 36 heures);
- Le nombre de connexions dans le séminaire E-Space (moyenne : 12,6; écart-type : 8);
- Le nombre d'actions d'un panel de possibilités (le nombre de consultations de documents par l'étudiant (moyenne : 6,5; écart-type : 2,7), le nombre de courriers électroniques envoyés (moyenne : 1,2; écart-type : 1,8), le nombre de messages publiés sur le forum) (moyenne : 11,7; écart-type : 12,4).

La somme des trois valeurs détermine le triplet d'implication. Les trois valeurs relatives se compensent. Ainsi, un étudiant qui tente de donner une image de forte implication en renouvelant sa connexion sans rien faire d'autre n'aura qu'une forte valeur sur trois.

Puis, on s'intéresse au triplet d'activité consacré à l'interaction synchrone dans la certification, toujours des 426 étudiants :

- La durée de connexion au salon de discussion de la plateforme E-Space dans toutes ses fonctionnalités (moyenne : 39 minutes; écart-type : 31 minutes);
- Le nombre de connexions au salon de discussion de la plateforme E-Space (moyenne : 9,1; écart-type : 15,5);
- Le nombre de messages produits au salon de discussion (moyenne : 93,5; écart-type : 137).

Comme pour le triplet d'implication, le triplet d'interaction synchrone s'obtient par la somme des trois valeurs.

Puis, les deux triplets s'additionnent pour obtenir la valeur globale de participation.

Du point de vue statistique, on étudie si les traces d'activités peuvent se différencier selon le genre ou le groupe d'étudiants et la formation d'origine.

La première question statistique porte sur un éventuel effet groupe, puis un effet sexe et enfin un effet formation d'origine. Ainsi, on va différencier l'étude tout d'abord du point de vue de l'implication dans le séminaire (tableau 1), c'est-à-dire que l'on se focalise sur les temps de connexion, les durées de connexion au séminaire et le nombre d'activités. Le nombre d'activités pris en compte comprend les messages sur le forum et la messagerie écrits et lus, les documents déposés et lus, et le nombre de consultations du cours. Ces trois items, temps de connexion, durée, nombre d'activités, composent le triplet d'implication dans le séminaire. Puis, on s'intéresse aux données qui vont constituer le triplet d'interaction (tableau 2), c'est-à-dire aux données liées à la participation au salon de discussion synchrone (*chat*). On s'intéresse alors à la durée de connexion dans le clavardage, au nombre de connexions et au nombre de messages produits. On s'intéresse également à l'évaluation seule, croisée avec les trois mêmes facteurs étudiés (tableau 3). Pour chacun de ces facteurs, le test d'homogénéité de la variance de Levene renvoie une preuve d'hé-

térogénéité. En conséquence, plutôt qu'un test ANOVA, il est nécessaire d'effectuer un test non paramétrique de Kruskal-Wallis.

Tableau 1. Résultat Khi2 Kruskal-Wallis pour les facteurs qui composent le triplet d'implication (ensemble des actions dans le séminaire) selon le groupe, la formation d'origine et le sexe

Facteurs	Groupe (1 à 29) ddl : 28	Formation (1 à 8) ddl : 7	Sexe (M et F) ddl : 1
Nombre de connexions au séminaire	89,86; $p < 0,01$	35,5; $p < 0,01$	7,95; $p < 0,01$
Durée de connexion au séminaire	57,6; $p < 0,01$	NS : 5,83; $p < 0,55$	8,53; $p < 0,01$
Nombre d'actions d'implication	72,44; $p < 0,01$	20,34; $p < 0,01$	7,43; $p < 0,01$
Triplet d'implication	42,34; $p < 0,01$	25,49; $p < 0,01$	14,45; $p < 0,03$

Tableau 2. Résultat Khi2 Kruskal-Wallis pour les facteurs qui composent le triplet d'interaction (clavardage) selon le groupe, la formation d'origine et le sexe

Facteurs	Groupe (1 à 29) ddl : 28	Formation (1 à 8) ddl : 7	Sexe (M et F) Ddl : 1
Nombre de connexions au clavardage	60,7; $p < 0,01$	NS : 9,98; $p < 0,1$	NS : 0; $p < 0,96$
Durée de connexion au clavardage	77,78; $p < 0,01$	18,5; $p < 0,01$	4,91; $p < 0,02$
Nombre de messages du clavardage	56; $p < 0,01$	18,37; $p < 0,01$	6,57; $p < 0,01$
Triplet d'interaction	77,33; $p < 0,01$	21,35; $p < 0,01$	4,41; $p < 0,03$

Tableau 3. Résultat Khi2 Kruskal-Wallis pour l'évaluation selon le groupe, la formation d'origine et le sexe

	Groupe (1 à 29) ddl : 28	Formation (1 à 8) ddl : 7	Sexe (M et F) Ddl : 1
Évaluation	152,80; $p < 0,01$	19,75; $p < 0,01$	2,34; $p < 0,12$ (NS)

Le test ne permet pas de fixer la nature de la relation entre les différents facteurs, mais celle-ci existe. C'est particulièrement clair pour les 29 groupes de formation. Pour étudier l'éventuelle relation, nous reprenons les deux triplets que nous avons forgés, le triplet d'implication (% du nombre de connexions, % de la durée de connexion, % des actions de l'étudiant) et le triplet d'interaction (% du nombre de

connexions, % de la durée de connexion, % des messages de l'étudiant) pour élaborer le triplet d'activité générale qui est la somme des deux. À partir de celui-ci, nous allons vérifier s'il existe une relation possible entre l'évaluation et ce triplet par une régression linéaire que nous calculerons pour chacun des facteurs : sexe, formation d'origine et groupe de certification.

Tableau 4. Résultat de régression linéaire score triplet d'activité générale et évaluation selon la formation d'origine

Formation d'origine	Corrélation de Pearson	Valeur de p	ANOVA de la régression	Équation de la régression
1	0,231	0,01	14,02 (1,248) p < 0,01	$Y = x(0,021) + 38,52$
4	0,632	0,01	7,29 (1,11) p < 0,02	$Y = x(0,164) + 13,28$
7	0,34	0,02	3,9 (1,30) p < 0,05	$Y = x(0,022) + 40,23$
2, 3, 5, 6, 8	NS			

Tableau 5. Résultat de régression linéaire score triplet d'activité générale et évaluation selon le sexe

Sexe	Corrélation de Pearson	Valeur de p	ANOVA de la régression	Équation de la régression
Masculin	0,317	0,01	7,81 (1,70) p < 0,01	$Y = x(0,032) + 36,06$
Féminin	0,218	0,01	17,55 (1,355) p < 0,01	$Y = x(0,019) + 39,45$

Tableau 6. Résultat de régression linéaire score triplet d'activité générale et évaluation selon le groupe de travail

Groupes de travail	Corrélation de Pearson	Valeur de p	ANOVA de la régression	Équation de la régression
1, 2, 8, 9, 10, 13, 16, 17, 18, 19, 20, 22, 23, 24, 26, 27, 28	NS			
3	0,54	0,02	4,91 (1,12) p < 0,04	$Y = x(0,048) + 37,38$
4	0,71	0,01	11,21 (1,11) p < 0,01	$Y = x(0,044) + 31,28$
5	0,63	0,01	8,55 (1,13) p < 0,01	$Y = x(0,068) + 28,07$
6	0,451	0,01	3,32 (1,13) p < 0,09	$Y = x(0,059) + 31,98$
7	0,547	0,01	5,98 (1,14) p < 0,02	$Y = x(0,04) + 29,52$
11	0,557	0,01	6,29 (1,14) p < 0,02	$Y = x(0,056) + 36,65$
12	0,469	0,04	3,39 (1,12) p < 0,09	$Y = x(0,052) + 35,26$
14	0,403	0,06	NS	
15	0,472	0,03	4,02 (1,14) p < 0,06	$Y = x(0,038) + 28,15$
21	0,506	0,03	3,78 (1,11) p < 0,07	$Y = x(0,053) + 27,61$
25	0,551	0,01	8,11 (1,14) p < 0,02	$Y = x(0,022) + 44,83$
29	0,606	0,01	6,94 (1,12) p < 0,02	$Y = x(0,057) + 31,69$

Interprétations

Tout d'abord, les dimensions descriptives permettent de comprendre l'hétérogénéité de la distribution. Entre ceux qui font le minimum et ceux qui surinvestissent, le panel d'implication est très vaste. Il n'y a donc rien d'étonnant à constater que l'ANOVA n'est pas utilisable pour tenter de déterminer si le croisement des éléments qui constituent les triplets et les facteurs sexe, formation d'origine et groupe de travail permet de relever d'éventuelles différences. Le test non paramétrique de Kruskal-Wallis révèle que l'on peut rejeter l'hypothèse nulle pour ce qui concerne l'activité des étudiants liée au séminaire (tableau 1) et à l'interaction clavardage (tableau 2), sauf pour ce qui concerne le sexe et le nombre de connexions. Même exception liée au sexe et à l'évaluation, on ne peut rejeter l'hypothèse nulle, tandis qu'on le peut pour ce qui concerne les groupes de travail et les formations d'origine (tableau 3). Il y a un effet sexe, formation d'origine et groupe de travail. Ces facteurs et l'activité ou l'évaluation ont donc une relation à caractériser plus finement.

Une approche par régression linéaire nous en dit un peu plus long. En utilisant le triplet d'activité générale – qui est la somme du triplet d'implication et de celui d'interaction – et l'évaluation, on constate que trois formations d'origine qui représentent plus de 60 % de l'effectif sont caractérisables par une régression linéaire (tableau 4). Pour la formation la plus importante et ses 250 étudiants, le coefficient n'est pas très élevé (0,231), mais il est significatif, et la régression l'est, tout comme l'équation de la régression. Concrètement, cela signifie que la connaissance du score au triplet d'activité générale permet pour la majorité des étudiants de déterminer leur plage de score à l'évaluation. Idéalement, un étudiant qui obtient un score de 202 au triplet d'activité générale obtient par l'équation une valeur de 46,3 alors que l'évaluation qui est indépendante de l'ensemble des variables étudiées le place à 47,5. Même chose pour le croisement sexe et notation (tableau 5) : les coefficients ne sont pas très élevés mais ils sont pertinents. Enfin, pour 11 groupes sur

29 (tableau 6), il y a une corrélation relativement forte, autour de 0,5, et des régressions significatives avec cependant des différences de coefficients d'équation de régressions extrêmement diverses.

Discussion

Y a-t-il un effet genre, formation d'origine, groupe de travail, aussi bien dans l'évaluation que dans l'activité repérable et objectivable par les triplets d'activité notamment? La réponse est oui, partiellement. Le déroulé de l'ensemble des observations montre qu'il y a une différence liée au sexe. Si l'on extrait les rangs moyens utilisés pour le test de Kruskal-Wallis, on observe que les filles sont plus participantes avec un rang moyen de 221, contre 174 pour les garçons. Cependant, le différentiel d'évaluation n'est pas pertinent pour le sexe. Pour les groupes de travail, il est impossible de déterminer par ces seules données les raisons de l'effet, mais il existe. Est-il dû aux tuteurs ou à la composition du groupe selon les formations d'origine? En effet, les groupes sont constitués en favorisant la plus grande hétérogénéité possible des formations d'origine afin d'éliminer le travail en groupe en présence au profit de l'interactivité à distance garantissant l'égalité de traitement, mais la présence de trois groupes dominants influe forcément sur la distribution. La structuration des triplets d'activité est certainement à parfaire plus finement, peut-être en intégrant une dimension temporelle au début, au milieu et à la fin pour intégrer la question du rythme, de la régularité. Évidemment, les interventions des tuteurs ne sont pas comptabilisées ici, ce qui peut constituer un biais important et mériterait d'être mis en référence puisqu'à l'évidence, il y a des différences selon les groupes. Les résultats obtenus en observant uniquement les données quantitatives sans lien avec les contenus pour établir le rapport avec l'évaluation sont à la limite presque provocateurs. À l'extrême, on pourrait en venir à élaborer encore plus finement un algorithme qui chercherait à optimiser les équations de régressions, cette fois multifacteur, pour se passer de l'évaluation. On constate dans cette étude l'existence de liens mais à l'évidence, ils ne sont pas strictement mathématiques.

Des facteurs à améliorer

Ce qu'il est possible de montrer, c'est que les distributions ne sont pas complètement éclatées. Il est impossible de définir facilement une heuristique qui expliquerait immédiatement le tout. Sans être totalement insatisfaisante, l'étude ne permet pas de modéliser les dépendances. Il faut noter également qu'un grand nombre de facteurs viennent rendre l'éventuelle détermination très difficile. Par exemple, une petite minorité d'étudiants, malgré de nombreuses mises en garde, se sont mis en situation de fraude. De l'étudiant qui fournit la même copie qu'un de ses collègues, en changeant toutefois le nom, à celui qui a recomposé son travail à partir de trois ou quatre copies, une petite minorité d'étudiants, qui deviendront enseignants, ont considéré que cela n'était finalement pas grave de tricher et que, surtout, cela ne se verrait pas. L'ensemble des copies est évidemment passé par des logiciels de détection de similitudes, et il n'est donc pas difficile de les repérer. Par conséquent, ces étudiants, par un comportement non orthodoxe, perturbent les données à partir desquelles on cherche les dépendances, notamment par le fait qu'ils n'utilisent pas la plateforme pour se mettre d'accord pour frauder, mais se côtoient dans des salles de travail ou utilisent d'autres modalités de communication. Il faut signaler également que cette innovation se déroule dans un climat de nette opposition de la part des étudiants. À leur décharge, dans le contexte de la mastérisation, soulignons que dès lors que les étudiants ont eu connaissance des modalités de certification et du calendrier, ils ont considéré comme une attaque personnelle le sort qui leur était réservé. Ils ont donc très rapidement contre-attaqué par une pétition en ligne qui, malgré les nombreuses pluri-signatures, a atteint rapidement les 600 signataires. Autant dire que c'est dans un climat de franche opposition que le dispositif s'est mis en place. Un bon nombre d'échanges, de connexions, de discussions ont eu lieu, non pas pour travailler sur l'objet de la certification, mais pour contester, remettre en cause l'encadrement du dispositif. Par conséquent, les données de ce type d'échanges ou d'actions sont potentiellement divergentes par rapport à l'objectif de certification.

Conclusion

Malgré les difficultés inhérentes à la mise en place de tout dispositif, l'ensemble a fonctionné avec des résultats positifs en matière d'évaluation pour les étudiants. Pour ce qui concerne la problématique de caractérisation de l'activité à partir de traces, on peut constater qu'elle est pertinente et que les triplets d'activité permettent pour partie de caractériser une typologie d'étudiants en fonction de leurs actions et de leurs résultats. Cependant, on voit bien que cela ne permet pas d'expliquer l'intégralité de la variable, pour les raisons de biais exposées à propos de la contestation et vraisemblablement parce qu'il faut également avoir accès à des données plus qualitatives. On peut imaginer que les niveaux des étudiants dans les matières classiques ont une importance quantitative forte. Comment expliquer autrement que des étudiants avec des triplets généraux très forts ont en même temps des résultats moyens? La qualité et la nature des interactions sont certainement importantes. On devrait également observer les contenus des échanges et le temps passé à éclaircir ce qui leur pose problème ou bien la posture relative des groupes. Ainsi, l'étude partielle des interactions entre étudiants montre qu'une grande majorité de futurs enseignants adopte de réelles postures d'enseignants en alliant volonté et compétence pour accompagner les autres étudiants, en posant des questions, en répondant et en cherchant à comprendre les difficultés de leurs collègues. À la différence de cette majorité, d'autres étudiants, minoritaires certes mais repérables cependant, ont beaucoup plus de difficultés à dépasser le stade de la récrimination et s'enferment plutôt dans une posture de contestation et de dénonciation. Il y a des effets d'entraînement et, selon l'état d'esprit du groupe dans lequel l'étudiant est affecté, (positif, attentiste, négatif) les scores risquent forts d'être différents. Sur le fond, le choix de tenter une innovation massive est un risque. Les étudiants, les futurs enseignants et certains de leurs enseignants n'y sont pas particulièrement favorables. Il n'est certes pas politiquement correct de l'évoquer ainsi, mais est-ce un signe du potentiel d'un système éducatif ?

Références

- Albero, B. (2010). Une approche sociotechnique des environnements de formation. *Éducation et didactique*, 4(1), 7-24. Manuscrit [récupéré](#) de l'archive HAL, section *Sciences de l'homme et de la société* : <http://halshs.archives-ouvertes.fr>
- Blandin, B. (2010). Dispositifs techniques pour l'éducation, la formation et l'apprentissage – Point de vue des industriels sur les évolutions. *Distances et savoirs*, 2(8), 223-234.
- Bouzidi, L. et Jaillet, A. (2009). Can online peer assessment be trusted? *Journal of Educational Technology & Society*, 12(4), 257-268. [Récupéré](#) du site de la revue : <http://www.ifets.info>
- Chaptal, A. (2003). Réflexions sur les technologies éducatives et les évolutions des usages : le dilemme constructiviste. *Distances et savoirs*, 1(1), 121-147. [Récupéré](#) du site de la revue : <http://www.cairn.info/revue-distances-et-savoirs.htm>
- Depover, C. (2009). *La recherche en technologie éducative : un guide pour découvrir un domaine en émergence*. Paris, France : Les Éditions des archives contemporaines et Agence universitaire de la francophonie.
- Djouad, T., Mille, A., Reffay, C. et Benmohamed, M. (2009). Ingénierie des indicateurs d'activités à partir de traces modélisées pour un environnement informatique d'apprentissage humain. *STICEF*, 16. [Récupéré](#) du site de la revue : <http://sticef.univ-lemans.fr>
- Jaillet, A. (2004). *L'école à l'ère numérique*. Paris, France : L'Harmattan.
- Jaillet, A. (2005). Peut-on repérer les effets de l'apprentissage collaboratif à distance? *Distances et savoirs*, 3(1), 49-66. [Récupéré](#) du site de la revue : <http://www.cairn.info/revue-distances-et-savoirs.htm>
- Jaillet, A. (2008). Les segments de films pédagogiques, instruments de supplémentation. *Éducation et formation*, e289, 33-43. [Récupéré](#) du site de la revue : <http://ute3.umh.ac.be/revues>
- Jaillet, A. (2009). Traces et histoires de traces. Dans F. Larose et A. Jaillet (dir.), *Traces numériques en enseignement et formation : analyses et usages* (p. 15-36). Paris, France : L'Harmattan.
- Ministère de l'Éducation nationale, de la Jeunesse et de la Vie associative (2009, mise à jour décembre). *Le brevet informatique et internet (B2i) école-collège-lycée et CFA*. [Récupéré](#) le 1^{er} mai 2012 du portail Éduscol : <http://eduscol.education.fr>
- Ministère de l'Éducation nationale, de la Jeunesse et de la Vie associative (2011, mise à jour septembre). *Préconisations techniques. Le schéma directeur des espaces numériques de travail (SDET)*. [Récupéré](#) le 1^{er} mai 2012 du portail Éduscol : <http://eduscol.education.fr>
- Santos, O.-C., Rodríguez, A., Gaudio, E. et Boticario, J.-G. (2003). *Helping the tutor to manage a collaborative task in a web-based learning environment*. Dans R. A. Calvo et M. Grandbastien (dir.), *AIED03 online supplementary proceedings, vol. IV: Workshop towards intelligent learning management systems*. [Récupéré](#) du site de la conférence AIED 2003 : <http://rp-www.cs.usyd.edu.au/~aied>
- Soller, A., Martinez, A., Jermann, P. et Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education*, 15, 261-290. Manuscrit [récupéré](#) de l'archive HAL, section *Sciences de l'homme et de la société* : <http://halshs.archives-ouvertes.fr>