
Evaluation des enseignements par les étudiants

Que nous apprennent les évaluations d'enseignants identiques par des étudiants de facultés différentes ?

Marie Bocquillon*, **Antoine Derobertmeasure****, **Frédérique Artus***** & **Marc Demeuse******

** Université de Mons
Doctorante FNRS-FRESH
Institut d'Administration Scolaire
Place du Parc, 18
B - 7000 Mons
marie.bocquillon@umons.ac.be*

*** Université de Mons
Cellule Facultaire de Pédagogie Universitaire
Place du Parc, 18
B - 7000 Mons
antoine.derobertmeasure@umons.ac.be*

**** Université de Mons
Service d'Appui Pédagogique
Rue du Rossignol, 1
"Verlaine" Building (1st floor)
B - 7000 Mons
frederique.artus@umons.ac.be*

***** Université de Mons
Institut d'Administration Scolaire
Place du Parc, 18
B - 7000 Mons
marc.demeuse@umons.ac.be*

RÉSUMÉ. L'évaluation des enseignements par les étudiants est maintenant implantée dans bon nombre d'institutions d'enseignement supérieur. La recherche portant sur les réponses des étudiants est abondante, notamment en ce qui concerne les réponses des étudiants à des questions fermées. L'article suivant a la double particularité de (1) s'intéresser, sur la base d'une analyse thématique, à l'analyse des 699 réponses d'étudiants à une question ouverte, (2) de porter sur le cas particulier de 5 enseignants amenés à enseigner dans deux facultés de la même institution universitaire. Cette démarche de recherche permet d'étudier les caractéristiques génériques et spécifiques des évaluations des enseignements par les étudiants à travers l'étude des variations mais aussi des convergences (entre les EEE) d'un même enseignant dans des contextes (ici, facultés) différents. Les principaux résultats montrent une cohérence interfacultaire en ce qui concerne le degré de satisfaction des étudiants ainsi qu'au niveau du contenu de leur propos ; en effet, ce sont les mêmes aspects de la situation d'enseignement-apprentissage qui sont les plus commentés.

MOTS-CLÉS : Évaluation des enseignements, analyse thématique, commentaires, questions ouvertes

1. Introduction

L'élaboration d'un espace européen de l'enseignement supérieur (Rege Colet, 2010 ; Younès, Rege Colet, Detroz & Sylvestre, 2013) conduit à uniformiser certaines démarches en œuvre dans les institutions de formation. Ainsi, la tendance actuelle en Europe, notamment en contexte francophone, est à la généralisation de la procédure d'évaluation des enseignements par les étudiants (Daele, 2013 ; Romainville, 2009), celle-ci étant pratiquée depuis les années soixante dans les universités anglo-saxonnes (Younès, 2007).

Ce processus ne reçoit pas forcément un accueil positif de la part de tous les enseignants, certains construisant leurs résistances autour du principe de liberté académique et de la non-reconnaissance d'un statut d'évaluateur compétent en la personne de l'étudiant (Endrizzi, 2014, p. 23). La compétence évaluative de l'étudiant ne fait d'ailleurs pas non plus l'objet d'un consensus dans la littérature scientifique consacrée au sujet, comme l'avancent Felton, Koper, Mitchell et Stinson (2008), Freng et Webber (2009) et Worthington (2002). Selon ces auteurs, des variables telles que l'apparence physique de l'enseignant, la difficulté du cours et la sévérité de notation de l'enseignant, le statut de l'enseignant et sa reconnaissance nationale/internationale, son genre... entrent en ligne de compte lorsque l'étudiant est amené à formuler une évaluation de l'enseignement reçu. A ces éléments, nous semble important d'ajouter les propos de Detroz (2008, p.129) lequel indique que « *des méta-analyses mettant en œuvre des modèles statistiques performants relativisent la plupart du temps l'impact des résultats de ces études de laboratoire, soit en mettant en évidence l'inconstance des interactions mises à jour, soit en minimisant leur impact sur le score EEE. Les auteurs de ces études reconnaissent toutefois que, bien que faibles, certaines interactions peuvent influencer les résultats à l'EEE* ».

Dans une optique formative (Berthiaume & Sylvestre, 2012 ; Winer, Di Genova, Vungoc & Talsma, 2012 ; Zimmaro, Gaede, Heikes, Shim & Lewis, 2006), l'un des intérêts de ces évaluations consiste à fournir aux enseignants de l'enseignement supérieur un retour sur leur pratique en vue, *in fine*, d'améliorer la qualité de l'enseignement, la condition sine qua non à l'atteinte de cet objectif étant la mise en place de mesures concrètes notamment en ce qui concerne le soutien pédagogique à offrir aux enseignants (Bernard, 2011 ; Heyde, Ledjouis & Sylvestre, 2008).

Afin de dépasser le caractère parfois désorganisé leur étant attribué, il est nécessaire, pour le chercheur d'analyser de manière systématique (Le Duc, Dillenbourg & Ricci, 2012 ; Lewis, 2001), les commentaires des étudiants. Cet intérêt est d'autant plus grand que plusieurs auteurs (Alhija & Fresko, 2009 ; Zimmaro & al., 2006) ont mis en évidence le manque de recherches sur le contenu des commentaires des étudiants, les recherches étant davantage centrées sur l'analyse des réponses à des questions fermées, offrant un autre type d'informations à l'enseignant afin qu'il détermine les aspects de la situation d'enseignement-apprentissage à améliorer (Winer, Di Genova, Vungoc & Talsma, 2012 ; Zimmaro & al., 2006).

Plusieurs recherches ont déjà été effectuées au départ des données générées par l'analyse du contenu des commentaires rédigés par les étudiants dans le cadre de l'EEE à l'Université de Mons. L'une, (Bocquillon, et al. 2014), s'est intéressée à présenter, de manière descriptive, la manière selon laquelle les étudiants se positionnent, selon leur faculté et en regard de leur année d'études, dans le cadre de l'évaluation des enseignements. Les résultats montrent notamment que les étudiants peu avancés dans le parcours sont identifiés comme plus « sévères » dans leur évaluation que les étudiants de fin de master. Différents profils de répondants, liés à la faculté d'appartenance, ont également été identifiés. L'exploration de ces profils contrastés a fait l'objet d'un article (Bocquillon, Derobertmeasure, Artus, Kozlowski, 2015), qui présente les résultats en ce qui concerne les réponses (N=550) des étudiants de la Faculté de Psychologie et des Sciences de l'Education et de la Faculté Polytechnique. Cette recherche a mis en avant des différences entre les étudiants de ces facultés en termes de focalisation sur certains aspects. A titre d'exemple, dans la catégorie des effets associés aux actions mises en place par les enseignants, les étudiants de sciences humaines ont formulé davantage de commentaires mettant en avant le caractère motivant (ou non) des enseignements proposés tandis que les étudiants de la faculté Polytechnique ont davantage focalisé leurs propos sur les apprentissages et réflexions engendrés (ou non) par les enseignements suivis. Une troisième publication questionnant davantage les compétences évaluatives des étudiants a été publiée (Bocquillon, Derobertmeasure, Artus, Demeuse, 2015). Elle propose la mise en relation des réponses des étudiants aux critères d'efficacité de la pratique enseignante. Les principaux résultats montrent que les aspects majoritairement commentés par les étudiants (discours didactique, gestion de l'apprentissage, activités sur les contenus et choix et organisation des contenus) sont en totale cohérence avec les principales caractéristiques associées aux enseignants efficaces. L'article dont il est ici question ambitionne, à partir de la variable d'appartenance facultaire, identifiée comme intéressante à prendre en compte dans le cadre de ce type d'études (Kember & Leung, 2011), de mettre en relation les commentaires d'étudiants de facultés différentes

ayant cependant eu cours avec le même enseignant, considérant, comme le signalait déjà Marsh en 1987 que les évaluations des enseignements par les étudiants sont en lien avec la qualité pédagogique du cours.

Cette étude est réalisée sur la base de l'analyse thématique de 699 commentaires afin de répondre aux questions suivantes : les étudiants de deux facultés différentes ayant eu cours avec le même enseignant ont-ils la même perception de la prestation de celui-ci ? Leurs commentaires ont-ils la même polarité (« positif », « négatif », « recommandation », « neutre ») et portent-ils sur les mêmes aspects de la situation d'enseignement-apprentissage ? L'analyse des résultats met en évidence des éléments permettant d'étudier des caractéristiques génériques et spécifiques des évaluations des enseignements par les étudiants à travers l'étude des variations mais aussi des convergences (entre les EEE) d'un même enseignant dans des contextes (ici, facultés) différents. Ce texte présente tout d'abord le contexte au sein duquel la démarche d'EEE dont il est question est mise en place. Il présente ensuite brièvement le modèle théorique ayant permis de réaliser l'analyse thématique au cœur de cette recherche. La méthodologie et les principaux résultats sont ensuite présentés, d'un point de vue descriptif (ventilation générale des résultats) et d'un point de vue de leur contenu. L'article se termine par une conclusion ciblant les convergences relevées ainsi que les améliorations potentielles et les limites du dispositif d'évaluation mis en place dans l'institution concernée.

2. Contexte local

À l'Université de Mons, l'évaluation des enseignements par les étudiants a été généralisée en 2010 au niveau de toute l'institution. L'objectif de l'évaluation, est de recueillir un avis global sur la prestation d'un enseignant. Il s'agit d'une évaluation annuelle qui a lieu l'année qui suit la réussite de l'enseignement par l'étudiant. Elle comprend une question fermée : « Comment évalueriez-vous globalement les prestations pédagogiques de cet enseignant ? » à laquelle l'étudiant doit répondre en se positionnant sur une échelle de réponses à 5 valeurs : « excellentes », « bonnes », « satisfaisantes », « insuffisantes », « très insuffisantes ». Si l'étudiant répond « insuffisantes » ou « très insuffisantes », il est obligé de répondre à la question ouverte « Quels sont les points à améliorer ? ». Les étudiants qui ont répondu « excellentes », « bonnes » ou « satisfaisantes » peuvent également rédiger un commentaire (Bocquillon et al., 2014).

L'un des intérêts de la mise en place de cette démarche, consiste, pour l'institution, outre l'aspect strictement administratif, à fournir aux enseignants des informations quant à la qualité perçue de leurs enseignements, mettant ainsi à leur disposition une ressource mobilisable dans un processus formatif (Younès, 2007).

3. Modèle théorique

Afin de pouvoir « lire » et « décoder » ces commentaires en regard d'une grille de lecture et déterminer les objets de la situation d'enseignement-apprentissage sur lesquels les commentaires portent, le modèle de la situation d'enseignement-apprentissage de Derobertmasure et Dehon (2012), adapté à partir des travaux consacrés aux commentaires des étudiants (Alhija & Fresko, 2009 ; Le Duc & al., 2012 ; Lewis, 2001 ; Winer & al., 2012 ; Zimmaro & al., 2006 ; Younès, 2007) (figure 1) a été utilisé. Il a été utilisé dans le cadre de l'analyse thématique à partir de laquelle a été générée une partie significative des résultats présentés dans cet article.

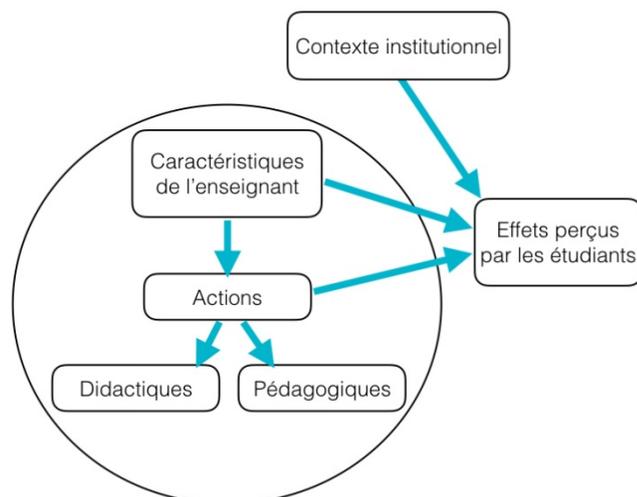


Figure 1. *Modèle de la situation d'enseignement-apprentissage adapté pour l'analyse des commentaires des étudiants (Bocquillon, Derobertmasure, Artus, Kozlowski, 2015).*

Ce modèle propose quatre grandes catégories permettant de coder les réponses des étudiants dans le cadre de l'EEE. Une catégorie reprend l'ensemble des caractéristiques de l'enseignant en tant que « personne » (sympathie, humour...). Les commentaires peuvent également porter sur ce que l'enseignant met en place, c'est-à-dire, ses actions de nature didactique (sélection et organisation des contenus, transposition...) ou pédagogique (gestion des rythmes, du climat...). L'étudiant peut également commenter le contexte (horaire, locaux...) dans lequel l'enseignement prend place. Finalement, dans le cadre de ses commentaires, l'étudiant peut également formuler un propos relatif aux effets (motivation, apprentissage...) qu'il associe aux actions mises en place par l'enseignant.

4. Méthodologie

4.1. Échantillonnage

À partir d'une liste anonymisée des enseignants ayant donné cours dans deux facultés différentes en 2010-2011 et 2011-2012 à l'Université de Mons, nous avons retenu les 5 enseignants pour lesquels un minimum de 20 commentaires par faculté avait été rédigé par les étudiants de premier cycle lors des enquêtes pédagogiques, ce qui nous donne un total de 699 commentaires analysés (sur un total de 15 720 commentaires obtenus pour toutes les facultés de l'Université de Mons lors des enquêtes pédagogiques portant sur les années 2010-2011 et 2011-2012). Le tableau 1 présente la répartition des commentaires entre les enseignants et les facultés. Il est important de préciser que le nombre inscrit dans chaque case représente le nombre total de commentaires des étudiants d'une faculté reçus par un enseignant. Par exemple, l'enseignant 1 a reçu 21 commentaires de la part des étudiants de la faculté 1. Bien que parfois faibles ou présentant de grands écarts en fonction des facultés, ces nombres représentent le résultat de l'application des précautions méthodologiques explicitées ci-avant et ils représentent donc la totalité des commentaires formulés par les étudiants concernés par les prestations en question.

Enseignant 1	
Nombre de commentaires faculté 1	Nombre de commentaires faculté 2
21 (7.7%)	251 (92.3%)
Enseignant 2	
Nombre de commentaires faculté 3	Nombre de commentaires faculté 4
91 (71.1%)	37 (28.9%)
Enseignant 3	
Nombre de commentaires faculté 2	Nombre de commentaires faculté 5
132 (77.6%)	38 (22.4%)
Enseignant 4	
Nombre de commentaires faculté 1	Nombre de commentaires faculté 5
23 (44.2%)	29 (55.8%)
Enseignant 5	
Nombre de commentaires faculté 1	Nombre de commentaires faculté 2
31 (40.3%)	46 (59.7%)

Tableau 1. Répartition des commentaires (N=699) entre les enseignants et les facultés.

4.2. L'analyse de contenu et le logiciel NVivo®

Bien que mobilisée dans de nombreux travaux, l'analyse de contenu est une méthode difficile à définir de façon exhaustive et consensuelle. On peut considérer que l'analyse de contenu d'un document consiste à « *rechercher les informations qui s'y trouvent, dégager le sens ou les sens de ce qui est présenté, formuler et classer tout ce que « contient » ce document* » (Mucchielli, 2006, p. 24) en limitant le recours à l'intuition et à la subjectivité (Derobertmeasure & Robertson, 2013).

Parmi plusieurs types d'analyses de contenu, c'est l'analyse de contenu thématique qui a été choisie. Ce type d'analyse se base sur un découpage du texte selon des unités de sens. Le chercheur doit donc associer un code à chaque unité du texte en fonction de son contenu.

Le logiciel NVivo® a été choisi, car il est fréquemment utilisé pour les analyses thématiques et de manière générale dans les recherches en sciences humaines, ce qui semble lui conférer une certaine reconnaissance dans le monde scientifique (Derobertmeasure & Robertson, 2013). De plus, dans le cadre présent, la fonctionnalité permettant de réaliser des matrices a été particulièrement utile.

4.3. Opérationnalisation du modèle théorique en une grille de codage

Lors de l'analyse de contenu, chaque commentaire (N=699) a été découpé en unités de codage (N=1536) (Bardin, 2001, p. 135). Dans le cadre de cette étude, étant donné qu'un même commentaire peut porter sur plusieurs thèmes, les unités de codage ont été découpées sur la base du critère sémantique (Alhija & Fresko, 2009 ; Le Duc & al., 2012 ; Winer & al., 2012).

Chaque unité de codage a d'abord été codée en fonction de son thème. Pour ce faire, le modèle théorique a été opérationnalisé en une grille de codage comprenant 8 catégories principales et 23 sous-catégories (tableau 2). Celle-ci est alimentée par une revue de la littérature reprenant les auteurs suivants : Lewis (2001), Zimmaro & al. (2006), Younès (2007), Alhija & Fresko (2009), Derobertmeasure (2012), Le Duc & al. (2012), Winer & al. (2012).

Catégories	Sous-catégories
Actions didactiques et psychopédagogiques	Choix et organisation des contenus
	Discours didactique
	Support didactique
	Préparation du cours
	Actes de langage
	Activités sur les contenus
	Climat-espace
	Dynamique corporelle
	Gestion de l'apprentissage
	Rythme-temps
	Prise en compte du niveau/ profil de l'étudiant
Effets	Apprentissage/ raisonnement/ réflexion/ compréhension
	Motivation/ intérêt de l'étudiant
	Participation des étudiants
	Impact du cours
Caractéristiques de l'enseignant	
Contexte institutionnel	Horaires, durée
	Environnement physique
	Nombre d'étudiants
	Composition du « groupe-classe »
	Programme des cours
	Matériel disponible
Appréciation générale	Appréciation générale du professeur
	Appréciation générale du cours
Sentiments des étudiants	
Autres	
Illisible	

Tableau 2. *Catégories de codage et sous-catégories.*

Ensuite, chaque unité de codage a été codée en fonction de sa polarité :

- la catégorie « positif » concerne les propos par lesquels les étudiants mettent en évidence les points positifs, les forces de la prestation de l'enseignant, sans proposer de pistes d'amélioration ;
- la catégorie « négatif » reprend les propos par lesquels les étudiants mettent en évidence les points négatifs, les faiblesses de la prestation de l'enseignant, sans proposer de pistes d'amélioration ;
- la catégorie « recommandation » concerne les propositions de pistes d'amélioration, qui sont souvent caractérisées par l'utilisation du conditionnel ;
- la catégorie « neutre » reprend les propos des étudiants qui déclarent ne pas pouvoir juger l'enseignant ou qui le font sur un ton neutre (ni positif, ni négatif).

Comme l'illustre l'exemple suivant, un commentaire peut comporter des unités relevant de plusieurs thèmes et polarités : « *Très disponible et dynamique* [caractéristiques de l'enseignant, positif], *mais support de cours insuffisant* [support didactique, négatif] *et explications parfois un peu brouillonnes* [discours didactique, négatif] *et difficiles à suivre* [apprentissage-compréhension, négatif] ».

4.4. Validation de la procédure de codage

Afin de limiter les biais associés à l'analyse de contenu, telles que la subjectivité du codeur, deux procédures de validation ont été retenues (Mukamurera & al., 2006) : la définition opérationnelle des catégories et des sous-catégories de codage et le contrôle de la fiabilité inter-codeurs (reproductibilité). Cette dernière a été calculée grâce à la formule de Miles et Huberman (2003, p. 126) : nombre d'accords entre les codages / (nombre d'accords + nombre de désaccords entre les codages). La fiabilité inter-codeurs est considérée comme satisfaisante à partir de 80 % d'accord. Dans le cadre de cette recherche, 50 unités ont été codées par un second chercheur qui disposait de la définition opérationnelle des différentes catégories et sous-catégories. Le résultat obtenu était de 94 % d'accord en ce qui concerne la polarité des unités et 78 % d'accord au niveau des thèmes de celles-ci. Après avoir discuté des désaccords, le résultat obtenu est de 100 % pour la polarité et 98 % pour les thèmes.

5. Analyse des résultats

5.1. Analyse descriptive

5.1.1. Valeurs attribuées aux prestations d'enseignement par les étudiants selon leur faculté d'origine

De manière générale, pour l'ensemble de l'université, les évaluations quantitatives des enseignants sont positives : 1% d'évaluations « très insuffisantes », 5% d'évaluations « insuffisantes », 23% d'évaluations « satisfaisantes », 43% d'évaluations « bonnes » et 28% d'évaluations « excellentes ».

Le tableau 3 présente, pour chacun des enseignants de l'échantillon, la répartition des réponses quantitatives (N=2290) en fonction de la faculté et des valeurs « très insuffisantes », « insuffisantes », « satisfaisantes », « bonnes » et « excellentes ».

Comme le montre ce tableau, on constate des convergences générales dans les évaluations quantitatives d'un même enseignant par des étudiants issus de facultés différentes. Ainsi, pour chaque enseignant, la valeur majoritairement attribuée par les étudiants d'une faculté (ex. : « bonnes » pour l'enseignant 1) n'est jamais éloignée de plus d'une valeur (sur l'échelle à 5 valeurs) de celle attribuée par les étudiants de l'autre faculté (ex. : « excellentes » pour l'enseignant 1). Malgré ces convergences générales, on remarque des décalages notamment au niveau des évaluations « excellentes » et « insuffisantes ». Par exemple, l'enseignant 3 a reçu 8,5% de réponses « insuffisantes » pour la faculté 2 contre 16,8% pour la faculté 5. Il est important de rappeler que ces constats sont élaborés sur la base d'un nombre de commentaires parfois très différent entre facultés et que certains résultats peuvent être impactés par ces différences, notamment dans le cas de très petits échantillons.

Cette première information est néanmoins, d'un point de vue de l'institution, plutôt rassurante : les enseignants amenés à intervenir dans différentes facultés obtiennent des appréciations proches, ce qui signifie que leur enseignement est perçu de manière globalement similaire par des publics issus de deux facultés.

	Ens. 1		Ens. 2		Ens. 3		Ens. 4		Ens. 5		Total
	Fac 1	Fac 2	Fac 3	Fac 4	Fac 2	Fac 5	Fac 1	Fac 5	Fac 1	Fac 2	
Très insuffisantes	0 0%	0 0%	1 0,3%	3 3%	3 1%	6 4,6%	3 2,6%	5 2%	2 1,4%	3 1,3%	26 1,1%
Insuffisantes	4 3,5%	2 0,4%	15 4,5%	10 10%	26 8,5%	22 16,8%	6 5,3%	13 5,2%	3 2,1%	10 4,5%	111 4,8%
Satisfaisantes	9 8%	14 2,4%	72 21,6%	28 28%	130 42,3%	64 48,9%	50 43,8%	97 39%	11 7,8%	51 22,9%	526 23%
Bonnes	58 51,3%	164 28,3%	164 49,3%	46 46%	114 37,1%	31 23,7%	49 43%	108 43,4%	56 39,7%	112 50,2%	902 39,4%
Excellentes	42 37,2%	399 68,9%	81 24,3%	13 13%	34 11,1%	8 6,1%	6 5,3%	26 10,4%	69 48,9%	47 21,1%	725 31,7%
Total de réponses quantitatives	113 100%	579 100%	333 100%	100 100%	307 100%	131 100%	114 100%	249 100%	141 100%	223 100%	2290 100%

Tableau 3. Répartition, pour chaque enseignant, des réponses (N=2290) en fonction de la faculté et des valeurs « très insuffisantes », « insuffisantes », « satisfaisantes », « bonnes » et « excellentes ».

5.1.2. Ventilation des commentaires en fonction de la faculté d'appartenance et de la valeur attribuée aux prestations d'enseignement

Si l'on s'intéresse ensuite à la répartition des commentaires rédigés par les étudiants selon les valeurs attribuées à leur évaluation¹ (N=699) (tableau 4), on constate, en lien logique avec les résultats précédents, une répartition inégale des commentaires au sein des notes « très insuffisantes », « insuffisantes », « satisfaisantes », « bonnes » et « excellentes ». Par exemple, en ce qui concerne l'enseignant 1, il y a davantage de commentaires associés aux notes « insuffisantes » (19%) dans la faculté 1 que dans la faculté 2 (0,4%) où il y a davantage de commentaires associés aux notes « excellentes » (74,9% contre 57,1%).

	Ens. 1		Ens. 2		Ens. 3		Ens. 4		Ens. 5		Total
	Fac 1	Fac 2	Fac 3	Fac 4	Fac 2	Fac 5	Fac 1	Fac 5	Fac 1	Fac 2	
Très insuffisantes	0 0%	0 0%	1 1,1%	3 8,1%	3 2,3%	3 7,9%	2 8,7%	2 6,9%	2 6,5%	2 4,3%	18 2,6%
Insuffisantes	4 19%	1 0,4%	10 11%	8 21,6%	23 17,4%	19 50%	6 26,1%	7 24,1%	3 9,7%	9 19,6%	90 12,9%
Satisfaisantes	1 4,8%	7 2,8%	23 25,3%	13 35,1%	55 41,7%	9 23,7%	9 39,1%	9 31%	2 6,5%	8 17,4%	136 19,5%
Bonnes	4 19,1%	55 21,9%	39 42,9%	10 27%	40 30,3%	6 15,8%	5 21,7%	7 24,1%	11 35,5%	14 30,4%	191 27,3%
Excellentes	12 57,1%	188 74,9%	18 19,7%	3 8,1%	11 8,3%	1 2,6%	1 4,4%	4 13,8%	13 41,9%	13 28,3%	264 37,8%
Total de commentaires	21 100%	251 100%	91 100%	37 100%	132 100%	38 100%	23 100%	29 100%	31 100%	46 100%	699 100%

Tableau 4. Répartition, pour chaque enseignant, des commentaires (N=699) en fonction de la faculté et des valeurs « très insuffisantes », « insuffisantes », « satisfaisantes », « bonnes » et « excellentes »

Pour une part, ce tableau reproduit les grandes tendances avancées dans le tableau précédent. L'analyse du contenu des commentaires permet d'en apprendre davantage sur les différents aspects couverts par ces commentaires (catégories et polarités).

¹ Dans quelques cas, on constate, pour les évaluations « très insuffisantes » et « insuffisantes » que le nombre de commentaires est inférieur au nombre d'évaluations y étant associées alors qu'une stricte égalité aurait dû être constatée étant donné que dans ces situations, les étudiants étaient obligés de formuler un commentaire. Dans les faits, certains étudiants ont contourné la consigne en utilisant certains subterfuges typographiques du type « slash », « points de suspension », « espaces » ... ce qui, d'un point de vue informatique, leur permettait de valider leur évaluation.

5.2. Analyse de contenu des commentaires

5.2.1. Polarité des commentaires

Le tableau 5 présente, pour chaque enseignant, la répartition des unités de codage (à l'exception des unités codées dans « Illisible ») (N=1528) en fonction de la faculté et de la polarité. Comme le montre ce tableau, on constate, pour chaque enseignant, des différences en ce qui concerne la polarité positive et négative des unités en fonction de la faculté considérée. Par exemple, pour l'enseignant 1, les étudiants de la faculté 2 ont rédigé des commentaires comprenant davantage d'unités positives (92,6%) que les étudiants de la faculté 1 (63,6%) dont les commentaires sont assortis de plus d'unités négatives (11,4% contre 3,3% pour la faculté 2) et « recommandations » (18,2% contre 3,1% pour la faculté 2). Concernant la catégorie « recommandations », il est d'ailleurs intéressant de souligner que cette catégorie devrait, théoriquement dans le cas d'une application stricte de la consigne donnée aux étudiants, être celle recueillant le plus d'unités de codage étant donné que la question posée est « Quels sont les points à améliorer ? ». Or, elle ne recueille en moyenne que 10,9% des commentaires et ne dépasse pas les 23,1% (enseignant 2, faculté 3) sur l'ensemble des résultats. On remarque également que les commentaires des étudiants sont généralement orientés, la catégorie « neutre » n'étant associée qu'à très peu de commentaires.

	Ens. 1		Ens. 2		Ens. 3		Ens. 4		Ens. 5		Total
	Fac 1	Fac 2	Fac 3	Fac 4	Fac 2	Fac 5	Fac 1	Fac 5	Fac 1	Fac 2	
Positif	28 63,6%	536 92,6%	80 40,2%	24 30,8%	98 33,1%	5 8,5%	8 22,9%	10 20,8%	46 61,3%	48 41,7%	883 57,8%
Négatif	5 11,4%	19 3,3%	71 35,7%	43 55,1%	159 53,7%	39 66,1%	20 57,1%	35 72,9%	22 29,3%	42 36,5%	455 29,8%
Recommandation	8 18,2%	18 3,1%	46 23,1%	11 14,1%	36 12,2%	13 22%	7 20%	3 6,3%	7 9,3%	18 15,7%	167 10,9%
Neutre	3 6,8%	6 1%	2 1%	0 0%	3 1%	2 3,4%	0 0%	0 0%	0 0%	7 6,1%	23 1,5%
Total	44 100%	579 100%	199 100%	78 100%	296 100%	59 100%	35 100%	48 100%	75 100%	115 100%	1528 100%

Tableau 5. Répartition, pour chaque enseignant, des unités de codage (N=1528) en fonction de la faculté et de la polarité.

5.2.2. Répartition des unités de sens en fonction des facultés et des thèmes d'analyse

De manière générale, des éléments de convergence sont relevés dans le tableau 6. En effet, malgré quelques nuances en ce qui concerne l'importance attribuée aux actions pédagogiques et didactiques de l'enseignant, c'est toujours la catégorie la plus commentée par les étudiants. De la même manière, les éléments relatifs au contexte institutionnel, aux sentiments et à la catégorie « autres » sont également des points de convergence entre les différentes évaluations puisque ces catégories ne sont que peu commentées. Au sein des évaluations enseignant par enseignant, des convergences sont également identifiées, ainsi, les catégories « effets » et « caractéristiques de l'enseignant » sont toujours, quelle que soit la faculté considérée, abordées dans des proportions proches.

Malgré ces convergences générales, des divergences sont également relevées en ce qui concerne les thèmes généraux abordés en fonction de la faculté considérée (à l'exception du cas de l'enseignant 4). Ainsi, pour les enseignants 1 et 2, on constate que les commentaires des étudiants d'une des deux facultés (la faculté 2 dans le cas de l'enseignant 1 et la faculté 4 dans le cas de l'enseignant 2) sont davantage généraux (respectivement 22,5% et 24,4% des unités codées dans « appréciation générale ») que les commentaires des étudiants de l'autre faculté (la faculté 1 dans le cas de l'enseignant 1 et la faculté 3 dans le cas de l'enseignant 2), qui portent davantage sur les actions (respectivement 45,5% et 56,8% des unités) et les caractéristiques de l'enseignant (respectivement 29,5% et 17,6%). En ce qui concerne l'enseignant 3, on constate que les étudiants de la faculté 2 sont davantage sensibles aux actions de l'enseignant (45,9% contre 33,9% pour les étudiants de la faculté 5) et au contexte institutionnel (8,1% contre 3,4%) que les étudiants de la faculté 5, qui commentent davantage les caractéristiques de l'enseignant (18,6% contre 10,8%) et les effets (30,5% contre 23%). Enfin, pour ce qui est de l'enseignant 5, les étudiants de la faculté 2 commentent davantage le contexte institutionnel (10,4%) que les étudiants de la faculté 1 (1,3%). Ces résultats devraient, idéalement, concourir quelque peu à « rassurer » les enseignants réfractaires à la procédure qui pensent que les étudiants commentent plus leur « personnalité » que leurs compétences : on remarque ici que les caractéristiques de l'enseignant, à l'exception d'un cas sur les 10, ne

ont jamais commentées à raison de plus de 20% des unités de codage. Par ailleurs, très peu d'unités ont été codées dans la catégorie « sentiments » (ex. : « *J'adore ce prof* »), confirmant les résultats d'une autre étude (Bocquillon, Derobertmeasure, Artus & Kozlowski, 2015).

	Ens. 1		Ens. 2		Ens. 3		Ens. 4		Ens. 5		Total
	Fac 1	Fac 2	Fac 3	Fac 4	Fac 2	Fac 5	Fac 1	Fac 5	Fac 1	Fac 2	
Actions didactiques et psycho-pédagogiques	20 45,5%	224 38,7%	113 56,8%	34 43,6%	136 45,9%	20 33,9%	20 57,1%	27 56,3%	32 42,7%	49 42,6%	675 44,2%
Caractéristiques de l'enseignant	13 29,5%	110 19%	35 17,6%	10 12,8%	32 10,8%	11 18,6%	5 14,3%	8 16,7%	14 18,7%	17 14,8%	255 16,7%
Appréciation générale	5 11,4%	130 22,5%	17 8,5%	19 24,4%	33 11,2%	6 10,2%	4 11,4%	6 12,5%	16 21,3%	19 16,5%	255 16,7%
Effets	5 11,4%	92 15,9%	31 15,6%	15 19,2%	68 23%	18 30,5%	5 14,3%	7 14,6%	12 16%	11 9,6%	264 17,3%
Contexte institutionnel	0 0%	4 0,7%	1 0,5%	0 0%	24 8,1%	2 3,4%	0 0%	0 0%	1 1,3%	12 10,4%	44 2,9%
Sentiments	0 0%	13 2,2%	1 0,5%	0 0%	1 0,3%	0 0%	0 0%	0 0%	0 0%	0 0%	15 1%
Autres	1 2,3%	6 1%	1 0,5%	0 0%	2 0,7%	2 3,4%	1 2,9%	0 0%	0 0%	7 6,1%	20 1,3%
Total	44 100%	579 100%	199 100%	78 100%	296 100%	59 100%	35 100%	48 100%	75 100%	115 100%	1528 100%

Tableau 6. Répartition, pour chaque enseignant, des unités (N=1528) en fonction des facultés et des catégories principales.

5.2.3. Répartition des unités de codages de la thématique « actions de l'enseignant »

La comparaison peut être affinée si l'on s'intéresse à la répartition, pour chaque enseignant, des unités codées dans la catégorie « Actions » (N=675) en fonction des facultés et des différentes actions (tableau 7).

Comme le montre le tableau 7, certains éléments de convergence apparaissent : la préparation, les actes de langage et la dynamique corporelle sont quasiment passés sous silence par la plupart des étudiants tandis que le discours didactique recueille le plus d'unités de codage associées aux actions de l'enseignant.

De la même manière, des convergences apparaissent entre les étudiants de deux facultés lorsque des catégories identifiées comme étant de moindre importance, voire minoritaires au niveau de l'institution (Bocquillon et al. 2015), sont mobilisées dans le cadre de commentaires : c'est le cas de la référence aux activités sur les contenus qui apparaissent de manière importante pour l'enseignant 5, pour la faculté 1 et la faculté 2 ; cette logique est également retrouvée dans le cas de la gestion de l'apprentissage chez le même enseignant et dans le cas des références à la gestion du climat et de l'espace pour l'enseignant 4.

L'importance accordée à cette catégorie par les étudiants lorsqu'ils évaluent la prestation de l'enseignant 4 est intéressante dans la mesure où celle-ci a été très peu relevée dans notre échantillon précédent. En effet, elle occupait 1,4% des unités codées dans la catégorie actions (N=589) (Bocquillon, Derobertmeasure, Artus, Kozlowski, 2015). L'échantillon précédent était constitué de 550 commentaires sélectionnés au hasard dans deux facultés de l'Université de Mons en 2010-2011. Ces résultats indiquent que certaines catégories qui étaient « marginales » dans notre étude précédente peuvent être plus saillantes lorsque l'on analyse les commentaires adressés à des enseignants en particulier plutôt qu'un échantillon de commentaires portant sur l'ensemble des enseignants.

C'est également le cas de la catégorie « Actes de langage », qui représentait 3,2% des unités codées dans la catégorie « Actions » dans notre étude précédente. Dans la présente étude, cette catégorie est assez bien commentée par les étudiants de la faculté 2 lorsqu'ils évaluent l'enseignant 3 (11,8%). Une logique identique peut être associée à la catégorie « Prise en compte du niveau et du profil des étudiants », qui a été peu relevée dans les commentaires composant notre échantillon précédent (2% des unités codées), et qui représente un cinquième des unités codées dans « Actions » pour les étudiants qui ont évalué l'enseignant 5 (18,8% pour les étudiants de la faculté 1 et 20,4% pour les étudiants de la faculté 2). Par contre, les catégories « Préparation » et

« Dynamique corporelle », qui étaient marginales dans l'étude précédente (respectivement 0,2% et 0% des unités codées dans « Actions ») le restent dans la présente étude.

Ce tableau permet également de s'interroger sur une éventuelle centration des étudiants sur certaines catégories en fonction de leur appartenance facultaire. En effet, les étudiants des facultés 1, 2 et 5 ont évalué au moins deux des enseignants ciblés par cette recherche. Ainsi, les étudiants de la faculté 1 ont évalué les prestations des enseignants 1, 4 et 5 ; ceux de la faculté 2 ont évalué les prestations des enseignants 1, 3 et 5 et ceux de la faculté 5 ont évalué les prestations des enseignants 3 et 4. Prendre en considération les catégories auxquelles ils ont attribué de l'importance permet de s'interroger sur une éventuelle focalisation de ces étudiants sur certaines dimensions de l'acte d'enseignement quelle que soit la prestation évaluée. A la lecture de certains résultats concernant notamment le « discours didactique » lorsqu'il est évalué par les étudiants de la faculté 1 (respectivement associé à 35%, 10% et 9,4% des unités relatives aux enseignants 1, 4 et 5) ou le « choix et organisation de contenus » pour la faculté 2 (respectivement associé à 4,9%, 25,7% et 10,2% des unités relatives aux enseignants 1, 3 et 5), il apparaît que les étudiants d'une même faculté ne concentrent pas, *a priori*, leurs commentaires sur certains aspects au détriment d'autres. Ce constat indique une capacité de discernement des étudiants dans le cadre des évaluations des enseignements, l'évaluation étant davantage liée aux diverses situations d'enseignement rencontrées qu'à une focalisation systématique sur certains aspects indépendamment de la situation observée.

	Ens. 1		Ens. 2		Ens. 3		Ens. 4		Ens. 5		Total
	Fac 1	Fac 2	Fac 3	Fac 4	Fac 2	Fac 5	Fac 1	Fac 5	Fac 1	Fac 2	
Choix et organisation des contenus	0 0%	11 4,9%	26 23%	2 5,9%	35 25,7%	1 5%	2 10%	2 7,4%	4 12,5%	5 10,2%	88 13%
Discours didactique	7 35%	67 29,9%	32 28,3%	16 47,1%	33 24,3%	7 35%	2 10%	7 25,9%	3 9,4%	6 12,2%	180 26,7%
Préparation	0 0%	1 0,4%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	1 0,1%
Support didactique	4 20%	21 9,4%	7 6,2%	3 8,8%	7 5,2%	1 5%	0 0%	2 7,4%	1 3,1%	3 6,1%	49 7,3%
Actes de langage	0 0%	2 0,9%	0 0%	0 0%	16 11,8%	4 20%	1 5%	5 18,5%	0 0%	0 0%	28 4,2%
Activités sur les contenus	5 25%	30 13,4%	22 19,5%	6 17,6%	8 5,9%	1 5%	0 0%	0 0%	10 31,3%	11 22,5%	93 13,8%
Climat-espace	0 0%	13 5,8%	1 0,9%	0 0%	15 11%	1 5%	15 75%	8 29,6%	0 0%	1 2%	54 8%
Dynamique corporelle	0 0%	0 0%	0 0%	0 0%	1 0,7%	0 0%	0 0%	1 3,7%	0 0%	0 0%	2 0,3%
Gestion de l'apprentissage	1 5%	47 21%	8 7,1%	4 11,8%	17 12,5%	5 25%	0 0%	1 3,7%	8 25%	11 22,5%	102 15,1%
Prise en compte du niveau-profil	0 0%	7 3,1%	9 8%	2 5,9%	0 0%	0 0%	0 0%	0 0%	6 18,8%	10 20,4%	34 5%
Rythme-temps	3 15%	25 11,2%	8 7,1%	1 2,9%	4 2,9%	0 0%	0 0%	1 3,7%	0 0%	2 4,1%	44 6,5%
Total	20 100%	224 100%	113 100%	34 100%	136 100%	20 100%	20 100%	27 100%	32 100%	49 100%	675 100%

Tableau 7. Répartition, pour chaque enseignant, des unités codées dans la catégorie « Actions » (N=675) en fonction des facultés et des différentes actions didactiques et psychopédagogiques.

6. Conclusion

Cette recherche a été réalisée à partir de l'ensemble des commentaires (N=699) issus de l'évaluation des enseignements de 5 enseignants de l'Université de Mons ayant donné cours dans deux facultés différentes en 2010-2011 et 2011-2012. À partir de ces données, différents résultats, du plus macro au plus micro ont été présentés dans cet article. Il s'agit (1) de la ventilation générale des évaluations (d'une prestation jugée « excellente » à « très insuffisante »), (2) de la ventilation des commentaires en fonction du niveau d'évaluation octroyé, (3) de la vue d'ensemble de la polarité assignée à chacune des 1528 unités de sens (contenues dans les 699 commentaires), (4) de l'analyse thématique des 1528 unités de sens en fonction du modèle de la situation d'enseignement apprentissage (voir figure 1), (5) de la focalisation de l'analyse thématique sur les sous-catégories composant la catégorie « actions de l'enseignant ». Certains effectifs à la base de cette étude sont

faibles, mais représentent l'ensemble des commentaires obtenus. Il s'agissait donc du seul échantillon disponible pour étudier le contenu des commentaires relatifs à une même prestation, mais formulés par des étudiants inscrits dans des facultés différentes.

L'analyse de ces commentaires permet d'avancer vers une meilleure connaissance des démarches évaluatives des étudiants (Pan & al., 2009). En effet, de nombreuses convergences sont observées dans le cadre des évaluations, quelle que soit la faculté d'appartenance des étudiants, lorsque l'on plonge réellement dans le contenu de ceux-ci : ils évaluent beaucoup plus les actions de l'enseignant que les caractéristiques intrinsèques de celui-ci, ils sont capables de justifier leur appréciation positive comme leur appréciation négative (et évitent les commentaires de polarité « neutre ») et ils ne semblent pas être influencés par un quelconque effet « facultaire » en ce qui concerne les objets sur lesquels portent leur évaluation. De manière plus globale, une convergence est également constatée en ce qui concerne la satisfaction générale par rapport aux enseignements, même si des décalages apparaissent entre les facultés : un écart apparaît au niveau des appréciations générales, notamment en ce qui concerne les catégories d'évaluations les plus extrêmes, à savoir « excellentes » et « insatisfaisantes » ce qui peut suggérer des attentes différentes des étudiants en fonction de leur faculté d'appartenance. Cette tendance est également mise en avant par Pan et ses collègues (2009) qui insistent sur l'importance de l'analyse des commentaires pour comprendre la signification des appréciations en fonction des différents groupes d'étudiants concernés. Le relevé de ces différences de notation entre les facultés n'est pas neuf et Felton et ses collègues (2008), à partir des contributions de Feldman (1978), Cohen (1981), Cashin (1990), Cerrito (2000) et Obenchain, Abernathy & West (2001) en proposent une synthèse explicitant les raisons pour lesquelles des différences d'appréciation sont presque systématiquement relevées entre les facultés selon leur composante « sciences humaines », « sciences appliquées » ou « sciences exactes ».

La centration de cet article sur cinq cas particuliers d'enseignants (et non plus sur une perspective macroscopique où les résultats sont étudiés au niveau de l'institution ou au niveau des facultés) souligne également l'intérêt de révéler des résultats auxquels l'approche globale étaient restée aveugle. Il s'agit de la mise en évidence de certaines thématiques identifiées comme marginales dans le cas de l'approche générale des commentaires alors qu'elles recouvrent parfois, au niveau individuel d'un enseignant, une importance non négligeable, voire majeure. Il en est par exemple le cas pour la « gestion du climat et de l'espace », peu (voire pas) présente selon une approche globale (Bocquillon, Derobertmasure, Artus, Kozlowski, 2015), mais représentant, pour l'enseignant 4, 75 % du contenu des commentaires évaluatifs rédigés par les étudiants de la faculté 1. Ce résultat plaide pour l'articulation des deux types de démarches : une démarche globale permettant de disposer d'une ligne de base de la ventilation des commentaires en regard des différents thèmes d'analyse et une démarche plus clinique, permettant de positionner chaque enseignant par rapport à ces dimensions et à la répartition « moyenne » des commentaires. Cette identification de cas particuliers(s), comme celui de l'enseignant 4 en ce qui concerne la gestion du climat et de l'espace, constitue également pour l'institution, disposant d'un service d'appui pédagogique ayant notamment pour mission de former les enseignants, un indicateur intéressant à prendre en compte dans l'élaboration du catalogue des formations proposées à ses enseignants. Elle permet par ailleurs à l'enseignant concerné, d'entamer une régulation de ses actions ciblée grâce à une analyse réflexive outillée.

Reste également à souligner que la procédure systématique, telle qu'elle est mise en place au sein de l'institution, si elle a le mérite d'exister, est tout à fait perfectible. L'amélioration du dispositif d'EEE pourrait ainsi être réalisée en amont, comme le suggère Svinicki (2001), en formant les étudiants à la rédaction de feedbacks pertinents à destination de leurs enseignants. Cette potentielle amélioration des commentaires pourrait notamment se traduire par une plus grande proportion de commentaires de l'ordre de la recommandation aux enseignants. Les résultats fournis aux enseignants sont, pour assurer l'anonymat des répondants (dans le cadre de cours en petits effectifs par exemple), rendus à l'enseignant de manière globale en fonction du niveau des étudiants (bachelier ou master). La communication de ces feedbacks pourrait être améliorée, comme le suggère Lewis (2001), en connectant la ventilation des appréciations globales sur l'échelle à cinq niveaux aux commentaires y étant associés.

La question posée aux étudiants induit également une diversité des commentaires relativement grande à traiter par les enseignants lorsqu'ils lisent les résultats de leurs enquêtes ; cette question se limitant à demander aux étudiants les améliorations qu'ils proposeraient à leur enseignant. Une focalisation de certaines questions sur des aspects spécifiques de l'enseignement-apprentissage permettrait certainement de réduire cette diversité et d'augmenter, ainsi, la lisibilité des réponses.

Parallèlement à cette procédure automatisée et générale de collecte des appréciations des étudiants, les enseignants (ou les autorités), peuvent solliciter une évaluation de type approfondie. Depuis 2010, seule une

faculté a procédé à une évaluation approfondie de ses cursus et quelques quatre-vingt enseignants en ont bénéficié individuellement. Le caractère récent de la procédure approfondie n'a sans doute pas encore permis de faire entrer cette démarche dans la culture de l'institution (du moins dans les facultés où elle n'était pas pratiquée avant d'être imposée par l'institution).

Il convient sans doute d'encore préciser que si l'anonymat des répondants est totalement garanti dans les procédures au sein de l'Université de Mons, le résultat des évaluations, peut, dans le cadre de renouvellements de mandats ou de promotions être consulté par une commission d'évaluation, ce qui place la question de l'évaluation des enseignements dans une perspective particulière, bien différente de celle dans laquelle, hypothétiquement, pourrait se placer l'enseignant s'il la réalisait, de manière plus personnelle, en recourant à des outils tels que le suivant développé en Suisse (<http://eval-cours.poloSoft.ch/>). En effet, à l'aide de ce type d'outils, l'enseignant peut, au moment où il le désire, avec le groupe qu'il désire, réaliser de manière proactive, une évaluation de ses cours sur la base d'un nombre variable de questions préétablies (et il s'agit là de la seule contrainte de la plateforme prise en exemple) et ce, sans aucun contrôle de l'institution.

7. Limitations et perspectives

Il est également important de préciser les deux éléments importants liés à cette recherche. Le premier concerne le fait que, durant la période où s'est déroulée cette procédure, seuls les étudiants ayant réussi l'enseignement concerné sont sollicités pour réaliser cette évaluation (ce filtre vient d'être modifié avec la révision du décret définissant l'enseignement supérieur en Belgique francophone (Fédération Wallonie-Bruxelles, 2013), le concept d'année d'études étant supprimé). Le second élément est lié au déséquilibre existant entre le nombre de commentaires obtenus par un même enseignant dans deux facultés, soit parce que l'une d'elle est de taille beaucoup plus importante, soit parce que les étudiants de l'une des deux facultés sont plus prolixes. Il est important de rappeler que les cinq enseignants retenus dans le cadre de cette étude sont les seuls ayant enseigné un même cours dans deux facultés différentes qui ont obtenu au moins 20 commentaires d'étudiants de premier cycle dans chacune des deux facultés en question. Les auteurs ne souhaitent pas révéler les facultés dont il est ici question afin de préserver l'anonymat des enseignants ciblés. Dans une optique où plus d'enseignants pourraient intégrer cet échantillon, et permettre, de cette manière, que l'on ne puisse les reconnaître, l'évaluation des enseignements par les étudiants pourrait être mise en lien avec la nature du cours (cours théorique versus cours pratique), de la taille du groupe auquel l'enseignant s'adresse (grand effectif versus petit effectif) et de la caractérisation de la faculté (sciences humaines, sciences exactes ou sciences appliquées). Outre l'intérêt scientifique de ce type de données, cela permettrait également, au sein de l'institution, de donner davantage de sens aux évaluations des enseignements par les étudiants : sont-ils, par exemple, systématiquement « plus satisfaits » lorsque l'enseignement se dispense en petits groupes ou ce type de configuration n'impacte-t-il aucunement leur degré de satisfaction ? Sont-ils globalement plus satisfaits d'un cours à option car davantage en lien avec leurs aspirations professionnelles ou, au contraire, pour les mêmes raisons, ont-ils davantage d'attentes élevées entraînant par là une exigence plus grande et une évaluation plus « sévère » ?

Références Bibliographiques

- Alhija, F.N., & Fresko, B. (2009). Student evaluation of instruction: what can be learned from students' written comments ? *Studies in Educational Evaluation*, 35, 37-44.
- Bardin, L. (2001). *L'analyse de contenu*. Paris : Presses Universitaires de France (10^{ème} édition).
- Bernard, H. (2011). *Comment évaluer, améliorer, valoriser l'enseignement supérieur ?* Bruxelles : De Boeck.
- Berthiaume, D., & Sylvestre, E. (2012). Guide d'interprétation des résultats de l'évaluation d'un enseignement. *Centre de soutien à l'enseignement (cse), Université de Lausanne*. [En ligne]. Page consultée le 11 octobre 2013. http://pedagogieuniversitaire.files.wordpress.com/2013/03/unil-cse_guide_interpretation_resultats_evaluation_v20121003.pdf
- Bocquillon M., Derobertmeasure A., Kozlowski D., Artus F., & Demeuse, M. (2014). Ce que disent les étudiants au sujet des enseignements : analyse textuelle des commentaires. Mons : Actes du colloque AIPU 2014.
- Bocquillon M., Derobertmeasure A., Artus F., & Demeuse M. (2015). Students' comments about university teaching: which links with effectiveness models? @tic. revista d'innovació educativa, 15, 1-11, DOI: 10.7203/atic.15.3931

- Bocquillon, M., Derobertmeasure A., Artus F., & Kozlowski D. (2015). Evaluation des enseignements par les étudiants : que nous disent les commentaires écrits des étudiants ? *Evaluer. Journal international de recherche en éducation et formation*, 1 (1), 93-117.
- Daele, A. (2013). Analyser les commentaires des étudiant-e-s à propos de ses enseignements. *Pédagogie universitaire – Enseigner et apprendre en enseignement supérieur. Ressources pour le conseil et la formation pédagogique dans l'enseignement supérieur*. [En ligne]. Page consultée le 11 octobre 2013. <http://pedagogieuniversitaire.wordpress.com/2013/03/25/analyser-les-commentaires-des-etudiant-e-s-a-propos-de-ses-enseignements/>
- Derobertmeasure, A., & Dehon, A. (2012). Développement de la réflexivité et décodage de l'action : questions de méthode. *Phronesis*, 1 (2), 24-44.
- Derobertmeasure, A., & Robertson, J.E. (2013). Data analysis in the context of teacher training: code sequence analysis using QDA MINER. *Quality and quantity*, 48 (4), 2255-2276.
- Derobertmeasure, A. (2012). *La formation initiale des enseignants et le développement de la réflexivité ? Objectivation du concept et analyse des productions orales et écrites des futurs enseignants*. Université de Mons : thèse doctorale. [En ligne]. Page consultée le 19 juin 2017. <http://tel.archives-ouvertes.fr/tel-00726944>
- Detroz, P. (2008). L'Évaluation des enseignements par les étudiants : état de la recherche et perspectives. *Revue française de pédagogie*, 165, 117-135.
- Endrizzi, L. (2014). La qualité de l'enseignement : un engagement des établissements, avec les étudiants ? *Dossiers de Veille de l'IFE*, 93. [En ligne]. Page consultée le 12 décembre 2014. <http://ife.ens-lyon.fr/vst/DA-Veille/93-juin-2014.pdf>
- Fédération Wallonie-Bruxelles (2013). Décret définissant le paysage de l'enseignement supérieur et l'organisation académique des études. [En ligne]. Page consultée le 9 novembre 2016. http://www.gallilex.cfwb.be/document/pdf/39681_020.pdf
- Felton, J., Koper, P.T., Mitchell, J., & Stinson, M. (2008). Attractiveness, Easiness, and Other Issues: Student Evaluations of Professors on RateMyProfessors.com. *Assessment & Evaluation in Higher Education*, 33 (1), 45-61.
- Freng, S., & Webber, D. (2009). Turning Up the Heat on Online Teaching Evaluations: Does "Hotness" Matter? *Teaching of Psychology*, 36, 189-193.
- Heyde, M., Lediouris, L., & Sylvestre, E. (2008). Une approche qualité pour l'évaluation de l'enseignement par les étudiants dans une université de masse. Montpellier : Actes du 25^e congrès de l'Association internationale de Pédagogie universitaire, *Le défi de la qualité dans l'enseignement supérieur : vers un changement de paradigme*. [En ligne]. Page consultée le 15 octobre 2013. http://icap.univ-lyon1.fr/servlet/com.univ.collaboratif.util.LectureFichier?ID_FICHER=1224247076033&ID_FICHE=114848
- Kember, D., & Leung, D. (2011). Disciplinary Differences in Student Ratings of Teaching Quality. *Research in Higher Education*, 52 (3), 278-299.
- Le Duc, I., Dillenbourg, P., & Ricci, J.-L. (2012). L'analyse des commentaires des étudiants dans les évaluations de l'enseignement montre qu'ils sont plus sensibles au fond qu'à la forme. Trois-Rivières (Québec) : Actes du 27^e congrès de l'Association internationale de Pédagogie universitaire, *Quelle université pour demain ?* [En ligne]. Page consultée le 1^{er} février 2014. https://oraprdnt.uqtr.quebec.ca/pls/public/docs/GSC2220/F1828181615_Programme_complet_et_actes_Affiches_et_communications_individuelles_Session_1_6_Version_finale.pdf
- Lewis, K.G. (2001). Making Sense of Student Written Comments. *New directions for teaching and learning, Techniques and Strategies for Interpreting Student Evaluations*, 87, 25-32.
- Marsh, H.-W. (1987). Students' evaluation of university teaching: Research findings, methodological issues and directions for future research. *International Journal of Educational Research*, 11 (3), 253-288.
- Miles, M.B., & Huberman, A.M. (2003). *Analyse des données qualitatives*. Bruxelles : De Boeck (2^{ème} édition).
- Mucchielli, R. (2006). *L'analyse des documents et des communications*. Issy-Les-Moulineaux : ESF (9^{ème} édition).
- Mukamurera, J., Lacourse, F., & Couturier, Y. (2006). Des avancées en analyse qualitative : pour une transparence et une systématisation des pratiques. *Recherches qualitatives*, 26 (1), 110-138.
- Pan D., Tan, G.S.H., Ragupathi, K., Booluck, K., Roop, R., & Ip, Y.K. (2009). Profiling Teacher/Teaching Using Descriptors Derived from Qualitative Feedback: Formative and Summative Applications. *Research in Higher Education*, 50 (1), 73-100.
- Rege Colet, N. (2010). L'évaluation à l'épreuve d'un dispositif d'académisation de la formation des enseignants. Genève : Actes du Congrès AREF. [En ligne]. Page consultée le 12 avril 2014. <https://plone.unige.ch/aref2010/symposiums-longs/coordonateurs-en-b/les-points-aveugles-dans-l2019evaluation-des-dispositifs-d2019education-et-de-formation/Levaluation%20a%20lepreuve.pdf>

- Romainville, M. (2009). Une expérience d'élaboration collective de critères de qualité. In M. Romainville et C. Coggi (Eds.), *L'évaluation de l'enseignement par les étudiants. Approches critiques et pratiques innovantes* (pp. 145-163). Bruxelles : De Boeck.
- Svinicki, M.D. (2001). Encouraging your Students to give feedback. In *Techniques and Strategies for Interpreting Student Evaluations. New Directions for Teaching and Learning*, 87, 17-24.
- Winer, L., Di Genova, L., Vungoc, P.A., & Talsma, S. (2012). Interpreting end-of-course evaluation results. Montreal: Teaching and Learning Services, McGill University.
- Worthington, A. (2002). The Impact of Student Perceptions and Characteristics on Teaching Evaluations: A Case Study in Finance Education. *Assessment and Evaluation in Higher Education*, 27 (1), 49-64.
- Younès, N. (2007). A quelles conditions l'évaluation formative de l'enseignement par les étudiants est-elle possible en France ? *Revue Française de Pédagogie*, 161, 25-40.
- Younès, N., Rege Colet, N., Detroz, P., & Sylvestre, E. (2013). La dynamique paradoxale de l'évaluation de l'enseignement par les étudiants. In M. Romainville, R. Goasdoué & M. Vantourout (Eds.), *Évaluation et enseignement supérieur* (pp. 109-126). Bruxelles : De Boeck.
- Zimmaro, D.M., Gaede, C.S., Heikes, E.J., Shim, M.P., & Lewis K.G. (2006). A study of students' written course evaluation comments at a public university. Austin: University of Texas.

