

Les qualités psychométriques des évaluations des enseignements par les étudiants algériens en sciences et techniques des activités physiques et sportives selon le genre et la moyenne académique : revue systémique de la littérature.

The psychometric qualities of evaluations of teaching by Algerian students in sciences and techniques of physical and sports activities according to gender and academic average: systemic review of the literature.

Hakim Hariti, Ph.D.,² Jaouad Alem, Ph.D.,¹ Malika Nada Chihaoui, M.Sc.² et Hanane Khaled, M.Sc.².

²Laboratoire Sciences et Pratiques des Activités Physiques Sportives et Artistiques SPAPSA Institut de l'Éducation Physique et Sportive (IEPS), Université d'Alger 3 (Alger), Algérie. hariti.hakim@univ-alger3.dz, malikanadachihaoui@gmail.com, hanane.khaled1995@gmail.com

¹Groupe de Recherche sur l'Évaluation et le Développement des Compétences en Activité Physique et en Santé (GRÉDCAPS), École de kinésiologie et des sciences de la santé. Faculté de la santé, Université Laurentienne, Sudbury (ON), Canada. jalem@laurentian.ca

Reçue le : 05/04/2021

Expertisé le : 01/05/2021

Publié le : 09/07/2021

Résumé

La validité des évaluations anonymes des enseignements par les étudiants (EEE) repose sur l'hypothèse qu'en assistant à des conférences, les étudiants observent la capacité des instructeurs, et qu'ils rapportent honnêtement les évaluations (Cashin, 1999 ; Feldman, 1988 ; Howard et al., 1985 ; Marsh, 2007 ; Marsh et Dunkin, 1992 ; Stark et Freishtat, 2014 ; Wright, 2006 ; Spooren et al., 2013). Par contre, d'autres études ont démontré que les étudiants ne sont pas si impartiaux (Johnson, 2003 ; Braga et al., 2014) et que les (EEE) pouvaient être influencées par le niveau d'étude, le type de questionnaire, le type de passation du questionnaire. Cette revue systémique de la littérature présente deux facteurs de biais potentiels : les effets du genre des étudiants et des professeurs et les effets de la moyenne académique sur les EEE. Même si les résultats sont plutôt contradictoires et mitigés, cette revue permet de libeller des hypothèses de recherche qui mériteraient d'être testées : 1- Les EEE sont influencées par le genre des étudiants et des enseignants,

en faveur du

genre masculin, 2- la corrélation entre les EEE et la moyenne académique est positive et moyenne, 3- les items des questionnaires EEE sont discriminants et interreliés. Il conviendrait aussi de mener d'autres études qui intégreraient des variables contextuelles telles que la culture évaluative des étudiants et des professeurs.

Mots clés : Évaluation de l'enseignement par les étudiants, biais liés au genre des étudiants et des professeurs, biais relié à la moyenne académique, consistance interne et discrimination des items.

Abstract

The validity of anonymous student teaching evaluations (EEE) rests on the assumption that by attending lectures, students observe the ability of instructors, and that they honestly report evaluations (Cashin, 1999; Feldman, 1988; Howard et al., 1985; Marsh, 2007; Marsh and Dunkin, 1992; Stark and Freishtat, 2014; Wright, 2006; Spooen et al., 2013). On the other hand, other studies have shown that students are not so impartial (Johnson, 2003; Braga et al., 2014) and that (EEE) could be influenced by the level of study, the type of questionnaire, the type of administration of the questionnaire. This systemic review of the literature presents two potential bias factors: the effects of the gender of students and professors and the effects of the academic average on EEE. Even if the results are rather contradictory and mixed, this review makes it possible to formulate research hypotheses that would deserve to be tested: 1- EEE are influenced by the gender of students and teachers, in favor of the masculine gender, 2- the correlation between the EEE and the academic average is positive and average, 3- the items of the EEE questionnaires are discriminating and interrelated. Other studies should also be carried out that incorporate contextual variables such as the evaluation culture of students and teachers.

Keyword : Evaluation of teaching by students, gender bias of students and teachers, bias related to academic average, internal consistency and item discrimination.

Introduction

Detroz (2008) définit l'évaluation de l'enseignement par les étudiants (EEE) « comme le fait de recueillir l'avis des étudiants sur la qualité des enseignements qu'ils ont suivis afin de permettre un jugement menant à des régulations » (p.118). La généralisation de l'EEE s'est faite de manière spectaculaire vers la fin du XX^e siècle ; elle est privilégiée par diverses institutions d'enseignement supérieur à travers le monde afin d'améliorer la qualité de l'enseignement en se basant sur l'expérience d'apprentissage des étudiants et l'évaluation par ces derniers des enseignements reçus. L'étude longitudinale de Seldin (1993) portant sur 600 collègues américains montre que 29 % de ceux-ci utilisaient l'EEE en 1973, qu'ils étaient 68 % à la pratiquer en 1983 et que ce chiffre atteignait les 86 % en 1993.

Certains chercheurs experts, considèrent même que l'EEE peut devenir une des stratégies de soutien au processus de développement professionnel des enseignants universitaires et ainsi agir en complémentarité avec d'autres stratégies et techniques, comme par exemple le croisement de ces informations avec d'autres telles que les résultats aux examens, les observations de l'enseignant lui-même ou encore l'observation par d'autres professionnels de l'enseignement (Bernard, Postiaux et Salcin , 2000 ; Coggi et Maccario, 2009; Smith, 2008). Les EEE sont utilisées pour des fins académiques et/ou pour des décisions liées à la carrière professorale (Donovan, Mader et Shinsky, 2006 ; Fontaine, 2009, Younès, 2009) et sont donc, par conséquent, une source de controverses dans les milieux universitaires (Dommeyer, Baum, Hanna et Chapman, 2004; Romainville et Coggi, 2009; Harvey et Hebert, 2012).

La validité des évaluations anonymes des étudiants repose sur l'hypothèse qu'en assistant à des conférences, les étudiants observent la capacité des instructeurs, et qu'ils rapportent honnêtement leurs EEE. Dans ce sens, Dunn, Hooks et Kohlbeck (2016) affirment que les évaluations des étudiants sont des mesures valides de la qualité de l'enseignement. Alem, , Hamrouni, Bouguerra et Hariti, (2015) ont analysé trois qualités psychométriques (la fiabilité, le degré de discrimination et la validité de construit) des huit items EEE utilisés par une université ontarienne et l'ont administrés à des étudiants algériens et tunisiens. Il apparaît que la fiabilité et le degré de discrimination des items étaient tout à fait acceptables, par contre la validité de construit restait à démontrer.

Boring, Ottoboni et Stark (2016) affirment pour leur part que les enseignants moins efficaces reçoivent une EEE plus élevée que leurs homologues féminins plus efficaces. Ils suggèrent de ne pas prendre en compte les résultats de l'EEE dans la promotion et le recrutement des enseignants à titre d'exemple.

Sur le plan empirique, Marsh (1983, 1984; Marsh et Bailey, 1993) soutiennent que les EEE sont multidimensionnelles, fiables et stables et relativement valides face à une variété d'indicateurs d'enseignement de qualité; les EEE ne sont pas affectées par des biais hypothétiques, elles sont utiles aussi

Hakim Hariti
Jaouad Alem
Malika Nada Chihouai
Hanane Khaled

Les qualités psychométriques des évaluations des enseignements par les étudiants algériens en sciences et techniques des activités physiques et sportives selon le genre et la moyenne académique : revue systématique de la littérature

bien aux professeurs

à titre de rétroaction à propos de leur enseignement; qu'aux étudiants en tant qu'outil de sélection des cours, qu'aux administrateurs pour fin de décisions administratives liés à la progression de la carrière, qu'aux chercheurs.

Par contre, d'autres recherches affirment que les résultats de l'EEE ont tendance à refléter en partie les caractéristiques des étudiants, des cours et des enseignants et qui ne sont pas forcément liés à la qualité de l'enseignement proprement dit (Berk 2005; Isely et Singh 2005; MacNell et al. 2015; McPherson et al. 2009). Par exemple, le biais de la note attendue par les étudiants pourrait offrir aux enseignants des incitations à l'inflation des notes, quels que soient les résultats d'apprentissage (Ewing 2012; Langbein 2008; McPherson 2006). Cependant les corrélations entre les performances académiques et les EEE sont significatives mais plutôt faibles ($R=0,36$) selon Younès (2009).

Analysé sous l'angle du genre, il existe l'hypothèse du biais des préjugés sexistes qui peut impacter les résultats de l'EEE. Une étude de Boring (2017) a porté sur la discrimination sexuelle que subissent les enseignantes en sciences sociales, l'auteur conclut que les étudiants font preuve de discrimination dans les évaluations des professeurs, les enseignantes obtiennent des scores EEE inférieurs, malgré que les enseignantes étaient des instructrices aussi efficaces que leurs homologues masculins.

Toutefois, Marsh (2007) et Spooren, Brockx et Mortelmans (2013) concluent que les études trouvent généralement de faibles influences parfois même non significatives pour un ensemble particulier de caractéristiques observées des enseignants et des étudiants pour expliquer la variation de l'ensemble. En effet, Basow (1995), Centra et Gaubatz (2000) ont constaté qu'il n'y avait pas de différence dans les scores de l'EEE attribuée par les étudiants des deux genres pour leurs enseignants, seulement il a été constaté que les étudiantes avaient tendance à attribuer une note globale plus élevée aux enseignantes.

Dans ce contexte et en absence d'une démarche et d'outils de l'EEE en Algérie, nous nous demandons si l'EEE administré à une population d'étudiant algérien en EPS, est affectée par les biais liés au genre et à la moyenne académique. Par ailleurs, nous nous assurerons également des qualités métriques des EEE en termes de discrimination et de consistance interne des items.

Cadre conceptuel.

La validité des EEE

La validité des évaluations anonymes de la qualité de l'enseignement par les étudiants repose sur l'hypothèse que, en assistant à des conférences, les étudiants observent la capacité des enseignants, et qu'ils rapportent honnêtement leur EEE. Si ce point de vue est plausible à certains égards, néanmoins de nombreuses études telles que celle de Johnson (2003) remettent en doute la validité de l'EEE comme mesure de l'efficacité de l'enseignement. Il existerait également comme le souligne Braga et al. (2014), des raisons de remettre en question l'EEE du fait que les étudiants ne sont pas des évaluateurs impartiaux de la performance des enseignants.

Néanmoins, des auteurs tels que Marsh et Dunkin (1992) considèrent l'EEE comme l'une des meilleures façons d'évaluer l'enseignement offert dans le cadre d'un cours par une ressource professorale et qu'elle est peu affectée par des biais potentiels. D'ailleurs, beaucoup de recherches plaident en faveur de la validité de l'EEE (Cashin, 1999 ; Feldman, 1988 ; Howard et al., 1985 ; Marsh, 2007 ; Marsh et Dunkin, 1992 ; Stark et Freishtat, 2014 ; Wright, 2006). Spooren et al., (2013).

Feistauer & Richter (2017) examinaient la fiabilité inter-évaluateurs de l'EEE à l'aide de modèles multiniveaux croisés. Un ensemble de données de 4224 évaluations (données issues de questionnaire) des cours de psychologie organisés sur trois ans à l'Université de Kassel, en Allemagne ont été analysées. Les évaluations ont été fournies par 480 étudiants (73% de femmes). Les conférences et séminaires ont été dispensés par 53 enseignants différents (dont 30 femmes). L'étude était basée sur un questionnaire EEE standardisé largement utilisé en Allemagne, la version pour les conférences contenait 31 items, la version pour les séminaires 34 items. Des analyses ont été effectuées avec des modèles à effets mixtes linéaires recoupés (Baayen et al., 2008) qui ont permis de séparer les composantes principales de la variance des enseignants, des cours et des étudiants. L'interaction des étudiants et des enseignants était la source de variance la plus forte (26% dans les séminaires et 32% dans les conférences). Cette étude confirme le doute quant à l'égard de la validité des EEE en tant qu'indicateurs de la qualité de l'enseignement et suggère que les notes d'évaluation agrégées soient utilisées avec prudence.

Dans la région du Maghreb, Alem et al. (2015), ont voulu vérifier les conclusions de Marsh (2007) relatives aux qualités psychométriques de l'EEE. Le questionnaire utilisé était celui de l'Université Laurentienne (8 items), il avait été administré à des étudiants algériens (N1= 83) et tunisiens (N2= 56) en formation en éducation physique et sportive au cours du mois de septembre 2012. L'analyse factorielle exploratoire avait permis de confirmer seulement deux dimensions théoriques du questionnaire : le climat et les objectifs d'apprentissage. La fiabilité du questionnaire telle que vérifiée par le calcul des coefficients alpha de Crombach était supérieure à 0,83. La méthode 'Éducational Testing Service' avait permis d'estimer les degrés de discrimination des items : les huit items étaient bien tous discriminants. Cette étude suggère que les EEE sont effectivement multidimensionnelles, fiables et discriminantes, par contre leur validité reste encore à démontrer, dans la mesure où la solution factorielle n'indiquait pas 7 dimensions clairement distinctes.

L'impact du genre des étudiants sur les EEE

L'influence des préjugés sexistes sur l'évaluation des étudiants a fait l'objet de plusieurs études, ces recherches ont analysés les évaluations des enseignants en fonction du genre de l'étudiant et de la note académique (Centra et Gaubatz, 2000 ; Huston, 2006; Miller et Chamberlin, 2000).

Basow (1995), Centra et Gaubatz (2000) ont constaté qu'il n'y avait pas de différence dans les scores de l'EEE attribuée par les étudiants. Par contre, Feldman, (1993) et Lueck et al., (1993) ont constaté que les étudiantes avaient tendance à attribuer une note globale plus élevée aux enseignantes tandis que les étudiants (les garçons) attribuaient de meilleures notes aux enseignants (les

Hakim Hariti
Jaouad Alem
Malika Nada Chihaoui
Hanane Khaled

Les qualités psychométriques des évaluations des enseignements par les étudiants algériens en sciences et techniques des activités physiques et sportives selon le genre et la moyenne académique : revue systématique de la littérature

hommes) qui, par

ailleurs sont mieux notés sur les thèmes de connaissances et d'enthousiasme. Pour certains auteurs ces résultats pouvaient être attribués à la perception des étudiants-es- des « rôles » du genre de l'enseignant. Signalons que Bodhe et Jankar (2015) ont noté que les étudiants-es- ne considèrent pas le genre de l'enseignant comme un biais de discrimination. Jesus Silva et al., (2017), affirment qu'il n'y a aucune influence du genre de l'enseignant sur la clarté du cours.

Huebner et Magel (2015) ont comparés les scores moyens de l'EEE dans des classes enseignées par des enseignantes et dans lesquelles la note moyenne des étudiants était supérieure à la note moyenne de leurs homologues féminins. Ils ont constaté que les variances moyennes des réponses des classes pour les deux genres de l'enseignant étaient significatives ou légèrement différente sur 13 des 16 items en faveur des enseignants. 10 items parmi 16 étaient en faveur des enseignants (p compris entre .014 à .047), à l'exception de l'item portant sur l'équité des procédures de notation ou les écarts moyens étaient plus élevés pour les enseignantes (39% de la variation des réponses moyennes de la classe).

Punyanunt-Carter et Carter (2015), affirment que les différences selon le genre de l'étudiant pour évaluer les enseignements en fonction du genre de l'enseignant ne sont pas significatives. Par exemple, pour l'enseignante, 82% des étudiants et 67% des étudiantes étaient d'accord ou fortement d'accord pour dire que l'enseignante était efficace. 69% des étudiants et étudiantes étaient d'accord ou fortement d'accord pour dire que l'enseignant était efficace).

L'étude de Boring (2017) a porté sur la discrimination existe que subissent les enseignantes en sciences sociales avec une base de données qui contenait un total de 20 197 observations (11 522 évaluations par les étudiantes et 8675 évaluations par les étudiants). Les résultats des analyses de régression montrent que les enseignants obtiennent des scores moyens de satisfaction significativement plus élevés de 0.156 point lorsque les étudiants les évaluent par rapport aux étudiantes, alors que pour les étudiantes qui évaluent les enseignants les scores de satisfaction étaient de 0,192 points de plus pour les enseignantes. L'auteur conclut que les étudiants font preuve de discrimination dans les évaluations des enseignants-es-.

Malgré que les recherches confirment que les scores moyens de l'EEE sont fiables et stables, elles sont plus fonction des caractéristiques et du comportement des enseignants plutôt que du contenu et de la qualité des cours en soi (Pounder, 2007; Marsh, 1987). Ces données de recherche confirment la notion selon laquelle les évaluations en général tendent à refléter les facteurs contextuels et souvent le genre plutôt que de traiter exclusivement du sujet (Greenwald et al., 1998).

Jesus Silva et al. (2017) ont analysé l'évaluation des performances d'enseignement d'un ensemble de 68 enseignants titulaire. L'activité d'enseignement est évaluée par les étudiants sur 9 items en utilisant une échelle de Likert à 7 niveaux. Les biais étudiés étaient le genre, l'expérience, les années d'expérience dans l'établissement, la catégorie professionnelle, le diplôme

universitaire, lieu de travail et le nombre de cours enseignés (en nombre d'heures).. Du point de vue des étudiants, et dans une perspective de genre, il n'y a aucune influence du genre de l'enseignant sur la clarté dans le cours et ce pour les deux semestres d'études (valeur *de p* respectivement égale à 0,248 et 0,343).

L'influence de la moyenne académique obtenue par les étudiants sur les scores des EEE.

Les notes attendues figurent parmi les biais importants qui peuvent influencer les EEE (Centra 2003; Beran et Violato, 2005). Certaines recherches postulent que les enseignants des deux genres ont tendance à « gonfler » les notes de leurs étudiants-es, car ils pensent que la note attribuée (et attendue) a un effet positif sur l'EEE indépendamment des acquis (Ewing 2012 ; Langbein 2008 ; McPherson 2006). Dans ce cas on se situe dans le champ de l'hypothèse inflationniste des notes des cours : les enseignants qui donnent les plus hautes notes sont « récompensés » par une opinion favorable des étudiants concernant l'enseignement qu'ils prodiguent (Greenwald et Gillmore, 1997). La méta-analyse de Cohen (1981) portant sur 41 études rapporte une corrélation plutôt faible de 0,43 entre la note attendue et l'EEE. Par contre les 24 études recensées par Aléamoni (1999) ne montrent pas de lien entre la note et le score à l'EEE, 37 montrent un lien positif et deux montrent un lien négatif.

Fischer et Hänze (2019), en voulant faire avancer la discussion sur la validité des évaluations des étudiants et leurs rapports avec l'EEE ont testé sept variables de biais potentiels : trois caractéristiques de l'enseignant (première impression, enthousiasme, humour) et quatre caractéristiques des étudiants (intérêt antérieur, notes attendues, expérience d'étude, présence en classe). Les résultats ont montré que les trois caractéristiques des enseignants étaient réellement liées au rapport enseignant-étudiant, et expliquaient même la variance de la variable notée par les étudiants lors du contrôle de la note et les mesures d'observations au rapport. Les notes des étudiants étaient en corrélation significative avec les mesures d'observation respectives: $r = 0,67$, $p < 0,001$ pour la participation des étudiants et $r = 0,48$, $p < 0,001$ pour le rapport.

Les deux recherches de Centra et al., (2003) ; Carrel et West (2010) ont montré que les étudiants à long terme confondent les notes et les attentes de notes avec la valeur d'un cours et que l'EEE ne sont pas associées aux étudiants brillants. Cependant Goos et Salomons (2016), contredisent ses résultats et certifient que la note obtenu dans le cours par les étudiants n'influence pas les scores de l'EEE. Au total, pour un échantillon de 28 240 , de 3329 cours dispensés par 1781 enseignants, le score moyen d'une EEE est de 13,0 avec un écart-type de 1,92 et le taux de réussite moyen est de 86%. La note moyenne obtenue par un étudiant est de 11,8 avec un écart-type de 3,22. Comme on le remarque, la note moyenne de l'étudiant est inférieure à celle du cours. Ces chercheurs avaient conclu que la note académique n'influaient pas significativement l'EEE.

Marks (2000) ; McPherson et Jewell (2007) ont mis en évidence l'impact de la clémence de la note sur l'EEE: si les étudiants reçoivent des notes inférieures pour une quantité d'apprentissage donnée, ils évaluent le cours moins favorablement. De même la littérature sur l'EEE souligne que les étudiants

Hakim Hariti
Jaouad Alem
Malika Nada Chihaoui
Hanane Khaled

Les qualités psychométriques des évaluations des enseignements par les étudiants algériens en sciences et techniques des activités physiques et sportives selon le genre et la moyenne académique : revue systématique de la littérature

brillants (sur la base

de la note moyenne ou cumulative) sont plus susceptibles d'attribuer des scores très élevés de l'EEE (Adams et Umbach 2012 ; Kherfi, 2011, Spooren et Van Loon 2012). Ces recherches sont en accord avec l'hypothèse de Marsh (1983) : les étudiants qui apprennent plus, sont ceux qui obtiennent les notes les plus élevées et sont les plus satisfaits. Autre biais : l'hypothèse inflationniste : les enseignants qui donnent les plus hautes notes sont « récompensés » par une opinion favorable des étudiants concernant l'enseignement qu'ils prodiguent (Greenwald et Gillmore, 1997).

Cependant, certaines recherches comme celle de Culver et Tech (2010), ont voulu examiner si la qualité de l'engagement des étudiants modère la relation entre les notes attendues et l'évaluation finale de l'enseignant. Bien que les principaux effets des notes attendues [$F(3, 3061.28) = 1020.43, p < .01$] et de la qualité de l'engagement [$F(2, 4882.72) = 2441.36, p < .01$] des étudiants étaient les deux statistiquement significatifs, l'effet d'interaction était lui aussi statistiquement significatif [$F(6, 336.70) = 56.12, p < .01$], indiquant que la relation entre la note finale attribuée par l'enseignant et la note attendue par l'étudiant était modérée par la qualité de l'engagement de l'étudiant.

Conclusion et pistes de recherche futures :

Même si les résultats de cette revue systématique de la littérature sont plutôt contradictoires et mitigés, cette revue permet de libeller des hypothèses de recherche qui mériteraient d'être testées : 1- Les EEE sont influencées par le genre des étudiants et des enseignants, en faveur du genre masculin, 2- la corrélation entre les EEE et la moyenne académique est positive et moyenne, 3- les items des questionnaires EEE sont discriminants et interreliés.

Il conviendrait aussi de mener d'autres études qui intégreraient des variables contextuelles telles que la culture évaluative des étudiants et des professeurs.

Références bibliographiques.

- Adams, M. J., et Umbach, P. D. (2012). Non response and online student evaluations of teaching: Understanding the influence of salience, fatigue, and academic environments. *Research in Higher Education*, 53(5), 576-591.
- Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924-1998. *Journal of Personnel Evaluation in Education*, 13(2), 153-166.
- Alem, J., Hamrouni, S., Bouguerra, I., et Hariti, H. (2015). The student evaluations of teaching by maghreb students in physical education. *International Journal of Innovation and Research in Educational Sciences*, 2 (2), 2349-5219.
- Baayen, R.H., Davidson, D.J., et Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.
- Basow, S.A. (1995). Student Evaluations of College Professors: When Gender Matters. *Journal of Educational Psychology*, 87, 656-665.
- Beran, T.N., et Violato, c. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment & Evaluation in Higher Education*, 30(6), 593-601.

- Bernard, H., Postiaux, N., et Salcin, A. (2000). Les paradoxes de l'évaluation de l'enseignement universitaire. *Revue des sciences de l'éducation*, 26(3), 625-650.
- Berk, R. A. (2005). Survey of 12 strategies to measure teaching effectiveness. *International Journal of Teaching and Learning in Higher Education*, 17, 48-62.
- Bodhe, C.D., et Jankar, D.S. (2015). Teaching effectiveness: how do students evaluate their teacher? *International J. of healthcare and Biomedical Research*, 63 (02), 155-159.
- Boring, A., Ottoboni, K., et Stark P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *Science Open Research*. DOI: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1.
- Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145, 27-41
- Braga, M., Paccagnella, M., et Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, 41, 71-88.
- Carrell, S.E., Page, M.E., West, J.E., (2010). Sex and science: how professor gender perpetuates the gender gap. *Q. J. Econ.*, 125 (3), 1101-1144.
- Carrell, S.E., et West, J.E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *J Polit Econ*, 118(3), 409-432. ISSN 0022-3808. doi: 10.1086/653808.
- Cashin, W.E. (1999). Student ratings of teaching : Uses and misuses. In P. Seldin (éd.), *Changing practices in evaluating teaching. A practical guide to improved faculty performance and Promotion/Tenure decisions* (pp. 25-44). Bolton, MA : Anker Publishing.
- Centra, J. A. et Gaubatz, N. B. (2000). Is there gender bias in student evaluation of teaching. *The Journal of Higher Education*, 70 (1), 17-33.
- Centra, J.A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education* 44(5), 495-518.
- Coggi, C., et Maccario, D. (2009). Contribution des étudiants à l'évaluation de la didactique universitaire : une réduction à de plus justes proportions, in M. Romainville et C. Coggi (dir.), *L'évaluation de l'enseignement par les étudiants (pp. 167-189)*. Bruxelles: De Boeck.
- Cohen, P.A. (1981). Student Ratings of Instruction and Student Achievement: A Meta-analysis of Multisection Validity Studies. *Review of Educational Research*, 51(3), 281-309.
- Culver, S. et Tech, V. (2010). Course Grades, Quality of Student Engagement, and Students' Evaluation of Instructor. *International Journal of Teaching and Learning in Higher Education*, 22(3), 331-336.
- Detroz, P. (2008). L'Évaluation des enseignements par les étudiants : état de la recherche et perspectives. *Revue française de pédagogie*, 165, 117-135.
- Dommeyer, C. J., Baum, P., Hanna, R. W., et Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and on-line surveys: their effects on response rates and evaluations. *Assessment & Evaluation in Higher Education*, 29(5), 611-623.
- Donovan, J., Mader, C. E., et Shinsky, J. (2006). Constructive student feedback: On-line vs. traditional course evaluations. *Journal of Interactive Online Learning*, 5(3), 283-296
- Dunn, K.A., Hooks, K.L., et Kohlbeck, M.J. (2016) Preparing Future Accounting Faculty Members to Teach. *Issues in Accounting Education*, 31 (2), 155-170.
- Ewing, A. M. (2012). Estimating the impact of relative expected grade on student evaluations of teachers. *Economics of Education Review*, 31(1), 141-154.
- Feistauer, D., et Richter, T. (2017). How reliable are students' evaluations of teaching quality? A variance components approach. *Assessment & Evaluation in Higher Education*, 42 (8), 1263-1279.

- Hakim Hariti
 Jaouad Alem
 Malika Nada Chihaoui
 Hanane Khaled
- Les qualités psychométriques des évaluations des enseignements par les étudiants algériens en sciences et techniques des activités physiques et sportives selon le genre et la moyenne académique : revue systématique de la littérature
- Feldman, K. A. (1993). College Students' Views of Male and Female College Teachers: Part II-Evidence from Students' Evaluations of their Classroom Teachers. *Research in Higher Education*, 34, 151-211.
- Feldman, K.A. (1988). Effective college teaching from the students' and faculty's view: matched or mismatched priorities ? *Research in Higher Education*, (28), 291-344.
- Fischer, E., et Hänze, M. (2019). Bias hypotheses under scrutiny: investigating the validity of student assessment of university teaching by means of external observer ratings. *Assessment & Evaluation in Higher Education*, 44 (5), 772-786
- Flanagan, J.C. (1939). Considérations générales dans la sélection des items de test et une méthode courte pour estimer le coefficient produit-moment à partir de données aux extrémités de la distribution. *Journal of Educational Psychology* , 30 (9), 674.
- Fontaine, S. (2009). Des expériences actuelles d'évaluation des enseignements vers des démarches adaptées aux 2e et 3e cycles. In M. romainville & C. Coggi (Eds.), *L'évaluation de l'enseignement par les étudiants : Approches critiques et pratiques innovantes* (pp. 123-144). Bruxelles : De Boeck.
- Garner, R. L. (2006). Humor in Pedagogy. How Ha-ha Can Lead to Aha! *College Teaching*, 54 (1), 177-180.
- Goos, M., et Salomons, A. (2016). Measuring teaching quality in higher education: assessing selection bias in course evaluations. *Research in Higher Education*, 58, 341-364.
- Greenwald, A. G., et Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52(11), 1209-1217
- Greenwald, A.G., McGhee, D.E., et Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, 74 (6), 1464-1480.
- Harvey, L. (2016). Les évaluations de l'enseignement par les étudiants : vers une démarche abrégée. *Mesure et évaluation en éducation*, 39, (2), 1-37.
- Harvey, L., et Hébert, M.H. (2012). Évaluation de la qualité de l'enseignement par les étudiantes et étudiants : qualités psychométriques et comparaison des conditions de passation. *Mesure et évaluation en éducation*, 35(3), 31-60.
- Howard, G.S., Conway, C.G., & Maxwell, S.E. (1985). Construct validity of measures of college teaching effectiveness. *Journal of Educational Psychology*, (77), 187-196.
- Huebner, L., et Magel, R.C. (2015). A Gendered Study of Student Ratings of Instruction. *Open Journal of Statistics*, 5, 552-567. <http://dx.doi.org/10.4236/ojs.2015.56058>.
- Huston, T. A. (2006). Race and gender bias in higher education: Could faculty course evaluations impede further progress toward parity? *Seattle Journal for Social Justice*, 4(2), 591-611.
- Isely, P., et Singh, H. (2005). Do higher grades lead to favorable student evaluations? *The Journal of Economic Education*, 36(1), 29-42.
- Jesus Silva, N., Costa Lobo, C., Santos Pereira, C., et Seabra Durão, N. (2017). Evaluation of teaching performance in higher education: student's perspectives and management indicators. *Conference: International Technology, Education and Development Conference*. DOI: 10.21125/inted.2017.1076
- Johnson, V. E. (2003). *Grade inflation: A crisis in college education*. New York, NY: Springer-Verlag.
- Kherfi, S. (2011). Whose opinion is it anyway? Determinants of participation in student evaluation of teaching. *The Journal of Economic Education*, 42(1), 19-30.

- Langbein, L. (2008). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review*, 27(4), 417-428.
- Lueck, T.L., Endres, K.L., & Caplan, R. E. (1993). The Interaction Effects of Gender on Teaching Evaluations. *Journalism Education*, 48, 46-54.
- MacNell, L., Driscoll, A., et Hunt, A. (2015). Whats in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, 40(4), 291-303.
- Marks, R. B. (2000). Determinants of student evaluations of global measures of instructor and course value. *Journal of Marketing Education*, 22(2), 108-119.
- Marsh, H. W. (2007). Students' evaluations of university teaching: A multidimensional perspective. In R. P. Perry et J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). New York, NY: Springer.
- Marsh, H. W., et Bailey, H. (1993). Multidimensional students' evaluations of teaching effectiveness: A profile analysis. *Journal of Higher Education*, 64(1), 1-18.
- Marsh, H.W., et Dunkin, M. (1992). Students' evaluations of university teaching : a multidimensional perspective. In J. C. Smart (éd.), *Higher Education: Handbook on Theory and Research* (pp. 143-234). Edison : Agathon Press.
- Marsh, H.W. (1987). Student evaluations of university teaching: research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11 (3), 255-388.
- Marsh, H. W. (1984). Students'evaluation of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5), 707-764.