

Les étudiants peuvent-ils objectivement évaluer la qualité d'un enseignement ?

par [Anne Boring](http://www.anneboring.com), OFCE-PRESAGE-Sciences Po et LEDa-DIAL, www.anneboring.com.

L'auteure présentera ses travaux au [Symposium international – Les biais de Genre dans la gouvernance et l'évaluation de la Recherche](#) organisé par [EGERA, Effective Gender Equality in Research and the Academia](#), qui se déroulera le 23 février 2015 à Sciences Po, dans les locaux du CERI, Paris.

Les universités anglo-saxonnes s'appuient largement sur les évaluations des enseignements par les étudiants pour mesurer la qualité d'un enseignement. Elles font l'hypothèse que les étudiants seraient les mieux placés pour juger de la qualité d'un enseignement dans la mesure où ils observent les enseignants à longueur de cours. Les évaluations ont généralement deux finalités. D'une part, elles sont utilisées comme outil de pilotage pédagogique pour les enseignants eux-mêmes, en leur fournissant des suggestions pour améliorer leur enseignement ; d'autre part, ces évaluations sont aussi souvent utilisées par l'administration pour ses décisions de promotion ou de prolongation de contrats d'enseignements. Les évaluations ont alors un objectif incitatif : elles encourageraient les enseignants à donner le meilleur d'eux-mêmes afin d'être reconduits le semestre suivant ou d'obtenir des promotions.

En France, la pratique des évaluations des enseignements est encore peu répandue, mais de nombreux établissements d'enseignement supérieur envisagent de la mettre en place. Certains établissements privés l'utilisent déjà dans leur politique de recrutement ou de prolongement de contrats de

vacation d'enseignement. Les établissements publics, quant à eux, ne peuvent utiliser les évaluations des enseignements que pour aider l'enseignant à améliorer ses pratiques pédagogiques. En effet, les établissements publics sont dans l'obligation de respecter la directive du ministère de l'Enseignement supérieur et de la recherche, qui précise que les « résultats de l'évaluation » ne peuvent être communiqués « qu'à l'enseignant concerné, et non au responsable pédagogique ou au directeur de la composante »[\[1\]](#). Cette directive confirme une décision du Conseil d'Etat, datant de 1997, qui indique que la procédure d'évaluation des enseignements « tend seulement à permettre aux enseignants d'avoir une meilleure connaissance de la façon dont les éléments pédagogiques de leurs enseignements sont appréciés par les étudiants » et « qu'elle ne comporte ni n'implique aucune incidence sur les prérogatives ou la carrière des enseignants ». Ainsi, seul l'enseignant intéressé peut avoir « connaissance des éléments de cette forme de l'évaluation »[\[2\]](#).

Que l'utilisation finale de cet outil de pilotage soit l'amélioration de la pédagogie ou la gestion des équipes, les universités doivent s'assurer que l'évaluation par les étudiants soit une mesure objective de la qualité d'un enseignement. Pour cela, elles doivent vérifier au moins trois conditions. Il est nécessaire :

- 1) que les étudiants sachent mesurer la qualité d'un enseignement, c'est-à-dire qu'ils soient en mesure d'établir les critères qui définissent la qualité d'un enseignement et de juger l'enseignant selon ces critères ;
- 2) que les étudiants ne soient pas biaisés dans leurs jugements et appréciations ;
- 3) que les enseignants ne puissent pas adopter de comportements stratégiques pour obtenir de bonnes évaluations ; autrement dit que l'objectif d'obtenir de bonnes évaluations

pour un enseignant n'induit pas de comportement pouvant porter préjudice à la qualité de l'enseignement.

Les étudiants savent-ils juger de la qualité d'un enseignement ? (Condition 1)

Quel enseignant n'a pas assisté à une discussion entre collègues où chacun défendait sa propre méthode pédagogique comme étant « la meilleure » ? Ces discussions portent généralement sur le contenu des enseignements, la façon de transmettre ce contenu, ainsi que sur les différentes approches concernant les modalités de contrôle des connaissances des étudiants. Déterminer les critères qui définissent un enseignement de qualité n'est pas chose aisée et les professionnels ne sont pas d'accord entre eux. Pourtant le système des évaluations suppose que les élèves soient au moins en partie en mesure de le faire.

Selon les étudiants, quels sont les critères importants pour déterminer de la qualité d'un enseignement ? La littérature suggère qu'un critère essentiel du point de vue des étudiants est l'extraversion et le dynamisme de l'enseignant, c'est-à-dire sa capacité à capter l'attention (e.g. Radmacher et Martin, 2001). Plusieurs travaux de recherche tendent à confirmer que les étudiants semblent juger en premier lieu la façon dont un enseignement est dispensé, plutôt que la qualité pédagogique ou le contenu même de l'enseignement.

L'effet « Dr. Fox » (Naftulin, Ware et Donnelly, 1973) fait par exemple référence à des enseignants sympathiques qui peuvent obtenir de bonnes évaluations en donnant l'impression d'être compétents, sans pour autant enseigner un contenu pertinent ou de bonne qualité. Dans cette expérience bien connue aux Etats-Unis, des chercheurs ont embauché un acteur pour donner un cours sur un sujet fictif. Le cours comportait de nombreux néologismes et contre-sens et l'idée des trois chercheurs qui ont embauché l'acteur était de déterminer si les personnes assistant à ce cours étaient en mesure de les

détecter, sans être aveuglées par l'aplomb, l'assurance et l'autorité académique affichée de l'enseignant (il était en effet présenté avec un faux cv : une panoplie complète de faux diplômes prestigieux et de faux articles de recherche). A la fin de l'enseignement, les personnes ayant assisté au cours du Dr. Myron Fox ont évalué positivement son enseignement. L'expérience montre d'une part que la perception que les étudiants ont de l'autorité académique d'un enseignant compte et, d'autre part, que les étudiants ne sont pas toujours capables de juger du contenu d'un enseignement.

Selon Carrell et West (2010) aussi, la perception que les étudiants ont de la qualité d'un enseignement n'est pas forcément corrélée avec la qualité réelle de cet enseignement, lorsque celle-ci est mesurée par la réussite à long terme. Ces auteurs montrent que les évaluations sont positivement corrélées avec la réussite à court terme des étudiants, mais peu corrélées avec la réussite à plus long terme. Leurs résultats laissent penser que les enseignants dont les techniques pédagogiques favorisent le bachotage peuvent être mieux évalués que des enseignants favorisant des techniques pédagogiques plus exigeantes et difficiles mais incitant davantage à l'apprentissage d'un savoir de longue durée. En effet, les élèves sont souvent d'abord préoccupés par leur réussite à l'examen final, plutôt que par l'utilité future des connaissances acquises au cours du semestre. Or, une université devrait créer des incitations pour que les enseignants utilisent des méthodes pédagogiques permettant l'apprentissage de long terme, méthodes qui ne semblent pas toujours être récompensées par les étudiants dans leurs évaluations.

Les jugements des étudiants sur la qualité d'un enseignant sont-ils non-biaisés ? (Condition 2)

L'évaluation des compétences peut être sujette à des biais de la part des évaluateurs. La littérature en psychologie sociale, notamment, suggère qu'il est plus difficile pour une

personne issue d'une minorité d'être perçue comme étant compétente (quand bien même elle le serait), alors qu'il est plus difficile pour une personne issue de la majorité d'être perçue comme étant incompétente (quand bien même elle le serait). Les effets de stéréotypes et de doubles standards d'évaluation s'appliquent dès qu'il s'agit de déterminer de la compétence individuelle (e.g. Basow, Phelan et Capotosto, 2006 ; Foschi, 2000). Leur impact peut avoir des conséquences particulièrement négatives pour certaines minorités, notamment pour les femmes enseignantes à l'université qui restent minoritaires.

Une étude[3] sur les évaluations d'étudiants de première année d'un établissement français d'enseignement supérieur montre que les étudiants appliquent bien des stéréotypes de genre dans la manière dont ils évaluent leurs enseignants. Les résultats de l'analyse économétrique montrent que les étudiants garçons tendent à mieux évaluer les enseignants hommes que femmes. Les enseignants hommes bénéficient en moyenne d'un biais favorable de la part d'étudiants garçons sur la quasi-totalité des dimensions de l'enseignement, en particulier la qualité de l'animation, la capacité à être en lien avec l'actualité et la participation au développement intellectuel de l'étudiant. Les filles ont aussi tendance à évaluer les hommes plus favorablement sur ces critères, mais accordent des évaluations plus favorables aux femmes sur d'autres dimensions de l'enseignement, notamment la préparation et l'organisation des séances, l'utilité des supports pédagogiques, la clarté des critères d'évaluation et la pertinence des commentaires de correction. Les biais des réponses des étudiants garçons et filles en faveur des hommes sur les critères liés à l'animation du cours notamment génèrent des scores de satisfaction globale plus élevés pour les enseignants masculins. Or, d'autres mesures de la qualité des enseignements (telle que la réussite aux examens) tendent à montrer que les enseignements dispensés par des femmes sont d'aussi bonne qualité que ceux dispensés par des hommes. De

plus, certaines tâches d'enseignement davantage valorisées chez les enseignantes (uniquement par les étudiantes) tendent à être chronophages. Les enseignantes se retrouvent ainsi avec moins de temps pour d'autres activités professionnelles, telles que les activités de recherche par exemple.

Les enseignants adoptent-ils des comportements stratégiques au détriment de la qualité de l'enseignement ? (Condition 3)

Enfin, plusieurs études montrent que des enseignants peuvent adopter des comportements stratégiques pour améliorer leurs évaluations. En effet, avec l'introduction des évaluations, les enseignants se trouvent confrontés au problème de l'agent multitâche (Holmstrom et Milgrom, 1991 ; Neal, 2013) : ils doivent bien enseigner, tout en obtenant de bonnes évaluations, objectifs qui ne sont pas nécessairement compatibles comme le montrent Carrell et West (2010). Les deux comportements stratégiques étudiés dans la littérature sont la capacité de démagogie d'un enseignant (*cf.* l'effet Dr. Fox), d'une part, et la notation généreuse des travaux des étudiants, d'autre part. Bien qu'il n'existe pas de consensus quant au lien causal entre bonnes notes données par des enseignants et bonnes évaluations données par les étudiants, il est en revanche démontré que les deux sont corrélées (*e.g.* Isely et Singh, 2005).

Conclusion

Les évaluations par les étudiants ne semblent pas satisfaire les trois conditions de mesure objective de la qualité d'un enseignement. On peut d'ailleurs se poser la question de savoir si la nature même de l'activité d'enseignement peut être mesurée objectivement. Faut-il pour autant ne pas mettre en place des systèmes d'évaluation des enseignements par les étudiants ? Les évaluations peuvent être utiles, mais elles doivent être interprétées avec précaution, en étant prises pour ce qu'elles sont plus probablement : une mesure du plaisir que les étudiants ont à aller en cours plutôt qu'une

mesure unique et objective de la qualité globale d'un enseignement. Le plaisir qu'un étudiant ressent à aller en cours n'est qu'un ingrédient parmi d'autres d'un enseignement de qualité. Il faut par ailleurs veiller à prendre en compte et à corriger les biais que les étudiants peuvent exprimer dans ces évaluations, en pondérant les critères d'évaluation de manière à ne pas décourager ou pénaliser injustement certaines catégories d'enseignants, notamment les femmes qui obtiendraient de moins bonnes évaluations uniquement du fait de stéréotypes de genre.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 612413.

Bibliographie

Basow, S. A., Phelan, J. E., & Capotosto, L. (2006). Gender patterns in college students' choices of their best and worst professors. *Psychology of Women Quarterly*, 30(1), 25-35.

Carrell, S. E., & West, J. E. (2010). Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors. *Journal of Political Economy*, 118(3), 409-432.

Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology*, 21-42.

Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 24-52.

Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations?. *The Journal of Economic Education*, 36(1), 29-42.

Naftulin, D. H., Ware Jr, J. E., & Donnelly, F. A. (1973). The Doctor Fox lecture: A paradigm of educational seduction. *Academic Medicine*, 48(7), 630-635.

Neal, D. (2013). The consequences of using one assessment system to pursue two objectives. *The Journal of Economic Education*, 44(4), 339-352.

Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *The Journal of Psychology*, 135(3), 259-268.



PRESAGE

Programme de Recherche et d'Enseignement
des SAvoirs sur Le GENre

The project EGERA has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 612413.

[1] http://www.sauvonsluniversite.com/IMG/pdf/annexe_2.pdf

[2]

http://www.legifrance.gouv.fr/affichJuriAdmin.do;jsessionid=82259E9F48498C78A1DE5D39FA492F2B.tpdjo02v_3?oldAction=rechExpJuriAdmin&idTexte=CETATEXT000007971674&fastReqId=2009366756&fastPos=1.

[3] Etude effectuée dans le cadre du programme de recherche européen EGERA : Boring, A. (2015). Gender Biases in Student Evaluations of Teachers (document de travail OFCE en préparation).