

Revue des sciences de l'éducation

Les paradoxes de l'évaluation de l'enseignement universitaire

Huguette Bernard, Nadine Postiaux et Alice Salcin

Volume 26, numéro 3, 2000

URI : id.erudit.org/iderudit/000293ar

DOI : [10.7202/000293ar](https://doi.org/10.7202/000293ar)

[Aller au sommaire du numéro](#)

Éditeur(s)

Revue des sciences de l'éducation

[Découvrir la revue](#)

Citer cet article

Bernard, H., Postiaux, N. & Salcin, A. (2000). Les paradoxes de l'évaluation de l'enseignement universitaire. *Revue des sciences de l'éducation*, 26(3), 625–650. doi:10.7202/000293ar

Tous droits de reproduction, de traduction et d'adaptation réservés © Revue des sciences de l'éducation, 2000

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne. [<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>]



Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. www.erudit.org

Les paradoxes de l'évaluation de l'enseignement universitaire

Huguette Bernard
Professeure

Université de Montréal

Nadine Postiaux et Alice Salcin
Étudiantes

Université libre de Bruxelles

Résumé – Cet article tente d'approcher la complexité des processus d'évaluation de l'enseignement tels qu'ils sont vécus et ressentis dans les universités québécoises. Par une méthodologie à la fois quantitative et qualitative sont ici dégagées les grandes tendances de réponses d'une population de chargés de cours et de professeurs au sujet des buts, des moyens, des sources et des modalités de l'évaluation de l'enseignement. Les sujets s'entendent, sans distinction majeure de statut, sur les moyens et les modalités d'évaluation. En revanche, le profil de l'évaluateur idéal semble difficile à cerner. Les autrices relèvent de nombreux paradoxes dans les réponses des sujets interrogés, reflet probable de la complexité d'une question qui reste sensible dans le monde universitaire.

Introduction

L'évaluation de l'enseignement se fait presque automatiquement dans les universités nord-américaines. Pourtant, il n'est pas rare d'entendre des professeurs, des chargés de cours, des étudiants, des directeurs de département ou des doyens de faculté remettre en question la valeur et l'efficacité de ces évaluations.

À la fin de presque trente ans d'évaluation de l'enseignement dans les universités québécoises, il est pertinent de chercher à connaître la perception des professeurs et des chargés de cours à cet égard et de mettre en évidence les paradoxes reliés à ce type d'évaluation. Quelle information possédons-nous sur l'efficacité d'une telle évaluation provenant principalement des étudiants? Cette dernière favorise-t-elle réellement l'amélioration de l'enseignement, comme on le suppose? Les questionnaires généralement utilisés sont-ils un bon moyen pour obtenir l'information désirée? Les professeurs croient-ils que les étudiants ont la compétence pour les évaluer? Croient-ils que les étudiants devraient avoir accès aux résultats de leur évaluation? Questions tout à fait actuelles, puisque les étudiants revendiquent de plus en plus cet accès à l'information croyant par là que c'est la seule façon de faire avancer ce dossier, qui leur semble stagnant depuis plusieurs années? Les professeurs ont-ils vraiment besoin des ressources qu'on semble met-

tre à leur disposition? Autant de questions auxquelles nous désirons obtenir des réponses afin de dresser un portrait de la situation actuelle et de proposer des mesures pour corriger au besoin certaines pratiques d'évaluation de l'enseignement.

Peu d'études ont tenté de connaître l'avis des professeurs à l'égard de l'évaluation qu'ils subissent très souvent cours après cours, année après année. Par ailleurs, une étude sur la perception des étudiants à cet égard (Thivierge et Bernard, 1996) a été réalisée auprès de 391 sujets de premier cycle universitaire et provenant de sept disciplines différentes. Les questions posées aux étudiants portaient principalement sur leurs perceptions de l'importance et de l'utilité de l'évaluation et sur la volonté institutionnelle de valoriser cette pratique. Ainsi, si les étudiants jugent l'évaluation de l'enseignement importante et croient qu'ils ont un rôle à jouer dans cette opération, ils doutent de son utilité parce qu'ils sont convaincus que les professeurs et l'institution y accordent peu de valeur; ils qualifient d'opération bidon cette évaluation qu'ils font à la fin des cours. Est-il possible de modifier cette perception? Que disent les professeurs et les chargés de cours à ce sujet?

Recension des écrits

Pour réaliser cette étude, nous avons consulté les écrits en fonction des dimensions suivantes, jugées les plus importantes de l'évaluation de l'enseignement, soit les buts poursuivis par l'évaluation, les sources d'information consultées, les moyens d'évaluation utilisés, les idées préconçues et les biais reliés à l'évaluation ainsi que les stratégies d'amélioration de l'enseignement consécutives aux évaluations.

Buts de l'évaluation de l'enseignement

Plusieurs auteurs dont Centra (1993), Murray (1984), Doyle (1983), McKeachie (1979) se sont prononcés sur les buts de l'évaluation de l'enseignement et tous s'entendent sur les deux plus importants: l'évaluation formative, faite à des fins d'amélioration de l'enseignement, et l'évaluation administrative, faite à des fins de promotion, permanence et renouvellement de contrats. Par ailleurs, les deux buts doivent-ils être confondus dans une même évaluation ou doivent-ils être réalisés séparément et différemment? Les auteurs diffèrent d'opinion sur cette question. Pour certains, il faut dissocier les buts puisque chacun demande une démarche totalement différente. Centra (1993) et Bernard (1992) proposent que le but formatif précède le but administratif. D'autres recommandent de concilier les deux buts en faisant en sorte que le but administratif serve aussi à des fins formatives. C'est du moins ce qui ressort de l'étude des politiques et des pratiques d'évaluation de l'enseignement. Il est peu fréquent qu'on dissocie les deux buts, ce qui fait dire

à certains que ce type d'évaluation confondue ne remplit convenablement ni sa mission administrative ni sa mission formative. Un troisième but s'ajoute, celui de permettre aux étudiants de faire leurs choix de cours en fonction des résultats des évaluations. Ce qui suppose de diffuser largement les résultats. Cette pratique existe principalement dans les universités anglophones.

Dans notre étude, il importe de savoir si les professeurs et les chargés de cours perçoivent l'évaluation comme une aide pour identifier les forces et les faiblesses de leur enseignement, s'ils considèrent qu'elle devrait aussi servir à des fins de promotion et de renouvellement de contrat, et s'ils acceptent que les étudiants aient accès aux résultats de leurs évaluations.

Sources d'information

Les études ont principalement porté sur la validité de l'évaluation faite par les étudiants (Abrami et d'Apollonia, 1990; Marsh, 1987). Ainsi, on ne remet plus en cause le rôle et la valeur des évaluations faites par les étudiants. Murray (1984) juge qu'ils sont les seuls témoins constants de la prestation de l'enseignement, ils sont les seuls à pouvoir indiquer si les exposés du professeur sont clairs, si les commentaires sur les travaux les aident à progresser, si les documents utilisés favorisent la compréhension de la matière. Ils sont une source privilégiée d'information, mais ils ne peuvent être l'unique source pour juger de la qualité de l'enseignement des professeurs.

D'autres sources, telles que les pairs, le professeur évalué, les diplômés et les spécialistes en enseignement, peuvent être mises à contribution (Centra, 1993; Brinko, 1991, 1993; Bernard, 1992; French-Lazovik, 1981). Chacune peut participer dans la mesure où les rôles sont bien définis, les buts bien précisés et les moyens différenciés, correctement construits et adaptés à la situation.

Ainsi, le spécialiste en enseignement pourrait s'exprimer spécifiquement sur la qualité de l'enseignement à des fins formatives alors que le pair, spécialiste du domaine d'enseignement du professeur évalué, pourrait intervenir à des fins de promotion ou de renouvellement de contrat en jugeant la pertinence des contenus enseignés et la qualité du matériel proposé aux étudiants. Le professeur aurait un rôle important en décrivant sa situation d'enseignement, expliquant ses approches pédagogiques et mettant en évidence les variables qui ont pu influencer son enseignement. Il pourrait témoigner de son investissement en enseignement (Bernard, 1992).

Les diplômés, quant à eux, sont fréquemment consultés au moment de l'évaluation des programmes d'études, mais rarement en évaluation des enseignements.

D'ailleurs, se souviennent-ils de leurs professeurs si ce n'est de ceux qu'ils jugent exceptionnels ou incompetents?

Ainsi, dans notre étude, on se préoccupera de savoir si les professeurs et les chargés de cours considèrent que les étudiants ont la compétence pour les évaluer, et si les diplômés ou les spécialistes en enseignement ou encore leurs collègues ou eux-mêmes peuvent jouer un rôle dans l'évaluation de l'enseignement.

Moyens d'évaluation

Le questionnaire rempli par les étudiants est le moyen d'évaluation le plus répandu pour évaluer l'enseignement des professeurs (Bernard et Bourque, 1999; CRÉPUQ, 1996; Donald et Saroyan, 1991; Donald 1991). Centra (1993), Abrami (1990), Marsh (1987) et plusieurs autres chercheurs ont mis au point des questionnaires validés d'évaluation de l'enseignement par les étudiants. Ces questionnaires, s'ils sont bien élaborés, peuvent effectivement devenir une source d'information nécessaire, mais pas forcément suffisante pour porter un jugement sur la qualité de la prestation des professeurs et des chargés de cours en classe. Par ailleurs, d'autres moyens, tels un questionnaire d'analyse du matériel d'enseignement, une grille d'observation en classe et le dossier de l'enseignement ou portfolio, peuvent être utilisés afin de compléter l'information fournie par les étudiants (Bernard, 1992; Seldin, 1991, 1993; French-Lazovik, 1981; Millman, 1981).

Des études antérieures ont montré que le dossier de l'enseignement et l'observation en classe sont des moyens d'évaluation utilisés dans des milieux anglophones, mais très peu dans les universités francophones (Bernard et Trahan, 1988). L'utilisation de ces moyens est un sujet très controversé par certains auteurs. Par exemple, Scriven (1981) considère que l'autoévaluation peut être considérée comme *self-serving* et que l'observation en classe est inacceptable lorsqu'elle est utilisée à des fins administratives parce qu'elle est trop ponctuelle, mal outillée, artificielle.

Les professeurs et les chargés de cours jugent-ils le questionnaire étudiant comme un bon moyen d'évaluation? Croient-ils que le matériel d'enseignement devrait être soumis à l'évaluation? Que pensent-ils des commentaires écrits des étudiants?

Les mythes et les biais de l'évaluation

De nombreuses études (Feldman, 1976, 1978, 1979, 1988; Marsh, 1987) montrent que les auteurs veulent connaître les variables susceptibles d'influencer

les résultats de l'évaluation de l'enseignement par les étudiants. On cherche à savoir si l'enseignement à des groupes nombreux provoquent des résultats aux évaluations plus faibles que dans les groupes restreints, si les cours obligatoires sont évalués plus faiblement que les cours au choix, si la personnalité du professeur peut influencer l'évaluation, si les notes accordées aux étudiants influencent leur évaluation, etc. Indépendamment des résultats de ces recherches, il nous importait de connaître la perception des professeurs et des chargés de cours sur ces sujets.

Les stratégies d'amélioration de l'enseignement consécutives aux évaluations

Quelles sont les stratégies mises en place à la suite des évaluations de l'enseignement? Des ressources existent-elles pour aider les professeurs et les chargés de cours qui le désirent? Les collègues ou les directeurs peuvent-ils jouer un rôle dans l'amélioration de l'enseignement? L'institution doit-elle agir en proposant des actions concrètes afin d'améliorer l'enseignement?

Il s'agit là de questions auxquelles certains auteurs ont tenté de répondre (Weimer et Firing Lenze, 1991; Seldin, 1990). Pour sa part, la CRÉPUQ (1996) a déjà fait état des mesures prises par les universités afin d'améliorer l'enseignement, telles l'évaluation, la formation des nouveaux professeurs, la consultation pédagogique, les journées sur l'enseignement, l'octroi de fonds de développement de l'enseignement, l'octroi de prix d'excellence en enseignement. Wright (1995) a réalisé une étude auprès des autorités concernés (vice-recteur, directeur de service à l'enseignement, etc.) des universités nord-américaines, canadiennes, australiennes et britanniques pour cerner leurs perceptions de l'efficacité d'un certain nombre de stratégies d'amélioration de l'enseignement. Il constate notamment que l'évaluation de l'enseignement faite en fin de session arrive en trente-sixième position sur les trente-six mesures proposées au Canada et que la prise en compte de l'enseignement, et non seulement de la recherche, au moment de l'étude des dossiers de promotion et de renouvellement de contrat, est considérée de façon internationale comme la mesure la plus efficace d'amélioration de l'enseignement. Malgré toutes les actions généralement proposées par les établissements, les professeurs et les chargés de cours ont-ils l'impression que leur université se préoccupe de la qualité de l'enseignement? Proposent-ils des moyens plus efficaces que ceux suggérés?

Méthodologie

Pour obtenir des réponses à nos questions, nous avons interrogé, par questionnaire écrit, des professeurs et des chargés de cours provenant de cinq établissements universitaires québécois¹.

Le questionnaire

Une liste de 75 énoncés issus de la recension des écrits et de notre expérience en ce domaine a été élaborée. Ceux-ci sont répartis dans les huit dimensions suivantes de l'évaluation: 1) les buts poursuivis et l'utilité de l'évaluation; 2) les sources d'information consultées; 3) les moyens utilisés; 4) le personnel évalué; 5) l'accès aux résultats; 6) les modalités de l'évaluation; 7) les mythes et les biais reliés à l'évaluation; 8) les suites de l'évaluation et la valorisation de l'enseignement.

Les énoncés ont été soumis à sept juges, des professeurs provenant de disciplines différentes et des spécialistes en enseignement supérieur et en évaluation, qui avaient à se prononcer sur la pertinence et la clarté de chacun des énoncés.

Le questionnaire final compte trois parties. La première présente les 40 énoncés jugés les plus pertinents et les plus clairs dans chacune des 8 dimensions précitées. Les répondants doivent se prononcer sur une échelle d'accord-désaccord en cinq points pour chaque item. La deuxième partie regroupe les informations signalétiques relatives aux professeurs et les chargés de cours (département, faculté, statut, nombre d'années d'enseignement). Enfin, dans une troisième partie, les sujets sont invités à commenter librement les pratiques d'évaluation de l'enseignement et à faire des suggestions d'amélioration.

L'échantillon

Près de 1 200 professeurs (40%) et chargés de cours (60%) provenant de cinq établissements universitaires québécois ont répondu au questionnaire administré à l'automne 1997 et à l'hiver 1998. Nous avons choisi de ne présenter ici que les réponses d'un seul établissement universitaire, dont 393 sujets ont répondu, soit 217 professeurs (55%) et 176 chargés de cours (45%). Plus de 50% d'entre eux ont une expérience d'enseignement de plus de 10 ans, une très faible minorité a une expérience de moins de trois ans.

Analyse des résultats

Dans un premier temps, nous développons les diverses étapes de la validation de l'instrument dans une perspective exploratoire. Nous nous attardons sur l'analyse factorielle exploratoire, car celle-ci semble offrir un modèle convaincant et sans doute plus adapté à la pensée des sujets. C'est pour cette raison que nous optons pour une présentation des résultats par item (moyenne, écart-type, médiane et pourcentage) au sein de la structure factorielle. Dans un deuxième temps, nous

abordons l'analyse qualitative en mettant en évidence les commentaires et suggestions d'amélioration du processus d'évaluation de l'enseignement proposés par les professeurs et chargés de cours. Nous terminerons par une discussion des grandes tendances des résultats émanant des deux analyses.

L'analyse quantitative

L'analyse factorielle a été effectuée dans une perspective exploratoire afin de cerner les facteurs qui ressortent des réponses des sujets sans *a priori* théorique. Nous avons gardé autant d'items que l'alpha de Cronbach le permettait et avons utilisé l'analyse factorielle Oblimin qui ne suppose aucun lien entre les facteurs.

L'analyse de l'alpha nous a amenées à recoder 8 items. Les items recodés et reformulés sont les items 1, 2, 12, 20, 23, 25, 32 et 38. Par contre, 9 items (13, 28, 29, 33, 34, 36, 37, 39, 40) ont une faible corrélation item-test (inférieures à 0,20) ; ils ont donc été retirés de l'analyse. Une faible corrélation est ici considérée comme résultante de ce que l'énoncé est trop général ou jugé hors sujet par les répondants, donc sans lien avec les facteurs importants.

Après ces opérations, l'alpha de Cronbach est de 0,90 pour les 31 items restants et soumis à l'analyse factorielle. En 20 itérations, l'analyse Oblimin a proposé 7 facteurs. Le coefficient Kaiser-Meyer-Olkin est de 0,89 et la probabilité du test de Bartlett est inférieure à l'alpha.

Nous présentons au tableau 1 les dimensions retenues pour le modèle théorique et les facteurs proposés par l'analyse factorielle. Il est intéressant de constater que l'analyse factorielle a confirmé certaines dimensions (personnel évalué), a regroupé certaines autres (moyen et source étudiante), a mis en évidence certaines dimensions (autres sources) et en aura séparé d'autres (biais externes et internes).

Tableau 1
Dimensions du modèle théorique et facteurs proposés
par l'analyse factorielle

Dimensions du modèle théorique	Facteurs proposés par l'analyse factorielle
Buts et utilités	Moyen et source étudiante
Sources d'information	Personnel évalué
Moyens d'évaluation	Autres sources
Personnel évalué	Accès et usage des résultats
Accès à l'information	Biais externes aux professeurs
Modalités d'évaluation	Suites de l'évaluation
Mythes et biais	Biais internes aux professeurs
Suites de l'évaluation et valorisation de l'enseignement	

Observons, par facteur, les résultats de chacun des items. Outre la moyenne, l'écart-type et la médiane, nous présenterons le pourcentage d'accord (réponses 4 et 5 dans le questionnaire), le pourcentage de désaccord (réponses 1 et 2 dans le questionnaire) ainsi que le point neutre (réponse 3 dans le questionnaire).

De manière générale, les résultats ne montrent pas de différence significative majeure entre les deux populations (chargés de cours et professeurs). Lorsque celle-ci existe, nous mettrons en évidence cette différence.

– Facteur 1 : moyen et source étudiante

Le premier facteur (tableau 2) regroupe la source étudiante et le moyen d'évaluation qui y est lié, soit le questionnaire. Si la distinction *moyen* et *source* apparaissait logique dans le modèle théorique, elle ne semble pas refléter la perception des sujets. Dans la plupart des cas, la seule modalité d'évaluation que les sujets connaissent est le questionnaire rempli par les étudiants.

Tableau 2
Moyen et source étudiante

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% \pm 3
1. L'évaluation faite par les étudiants m'aide à identifier les forces et les faiblesses de mon enseignement.	3,512	1,078	4	56,1	15,2	28,8
12. Les commentaires écrits des étudiants sont une source intéressante de suggestions pour améliorer l'enseignement.	3,603	1,225	4	59,2	18,5	22,2
11. Les questionnaires remplis par les étudiants ne sont pas de bons moyens pour évaluer l'enseignement.	2,948	1,276	3	30,7	37,5	31,8
5. Les évaluations faites par les étudiants sont inutiles.	2,008	1,188	2	13,3	69,7	16,9
4. Les résultats de l'évaluation faite par les étudiants sont trop imprécis pour me permettre d'améliorer mon enseignement.	2,776	1,201	3	27	45,5	27,5
2. Les étudiants ont la compétence nécessaire pour évaluer l'enseignement.	3,109	1,179	3	26,4	38,9	34,7
14. Les commentaires écrits des étudiants à la fin des questionnaires sont un lieu de défolement.	3,286	1,242	3	46,3	27,3	26,3
24. Les étudiants devraient s'identifier lorsqu'ils évaluent l'enseignement.	2,892	1,682	3	41,4	46,8	11,8

On note que 70 % des sujets se prononcent favorablement quant à l'utilité de l'évaluation (item 5) et près de 60 % reconnaissent son rôle formatif (items 1 et 12). À retenir cependant que si les commentaires écrits sont jugés source intéressante de suggestions, 46 % estiment qu'il s'agit également pour les étudiants d'une occasion de défolement (item 14). Sur la question de l'identification des étudiants (item 24), les résultats séparés des professeurs et des chargés de cours montrent une différence significative (test de Mann et Whitney, $\alpha < 0,01$). Les chargés de cours se prononcent à 49 % en faveur de l'identification alors que 55 % des professeurs s'y opposent.

– Facteur 2: personnel évalué

Concernant le personnel évalué (tableau 3), les sujets se prononcent clairement en faveur d'un système d'évaluation universel qui touche chacun de façon égalitaire. Il s'agit de loger tout le monde à la même enseigne et non de pointer les débutants, les « moutons noirs » ou encore uniquement les cours de premier

cycle. Seulement 22% estiment que le professeur doit être libre de faire évaluer ses cours.

Tableau 3
Personnel évalué

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% ± 3
18. On devrait évaluer uniquement les cours donnés par les chargés de cours.	1,270	0,744	1	3	93,1	3,9
17. On devrait évaluer uniquement les cours donnés par les nouveaux professeurs.	1,604	1,052	1	8,7	84,4	6,9
19. Seuls les professeurs dont se plaignent les étudiants devraient être évalués.	1,657	1,129	1	10,8	81,7	5,5
25. Les cours de 2 ^e et 3 ^e cycles devraient aussi être évalués.	4,570	0,852	5	91,1	3,6	5,2
15. Le professeur devrait être libre de faire évaluer ses cours.	2,204	1,447	2	21,7	66,5	11,7

– Facteur 3: autres sources

Ce troisième facteur (tableau 4) regroupe les sources possibles d'évaluation, autres que les étudiants. Cette deuxième évocation de l'évaluateur semble démontrer une certaine sensibilité des sujets sur cette question.

Les résultats montrent que les sujets, même s'ils sont globalement favorables à l'évaluation, semblent éprouver des difficultés à cerner le profil de l'évaluateur idéal: 71% des sujets s'opposent à ce que l'étudiant soit le seul à jouer ce rôle (item 4) et 50% ne semblent pas attribuer de compétence en la matière à leurs collègues (item 6). Le spécialiste en enseignement et l'étudiant diplômé (items 8 et 7) ne remportent pas non plus un franc suffrage. Seul le professeur, lui-même, paraît clairement compétent à 86% (item 10). Il est intéressant de constater que l'enseignant ne semble pas avoir besoin d'un regard extérieur sur sa pratique et que son seul jugement semble lui suffire.

Tableau 4
Autres sources d'évaluation

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% \pm 3
8. Un spécialiste en enseignement serait plus compétent que les étudiants pour juger l'enseignement.	2,863	1,261	3	32	39,9	28,2
6. Les collègues sont plus compétents que les étudiants pour juger la qualité de l'enseignement.	2,499	1,448	2	18,5	50,1	31,4
9. Les étudiants ne devraient pas être les seuls à juger de la qualité de l'enseignement des professeurs.	3,940	1,172	4	71,4	13,6	15
7. Les diplômés sont de meilleurs juges de la qualité de l'enseignement des professeurs que les étudiants actuels.	3,053	1,231	3	38,5	31,8	29,7
10. Le professeur évalué doit pouvoir se prononcer sur la qualité de son enseignement.	4,332	0,940	5	86	5,9	8,1

– Facteur 4: accès et usage des résultats

Le facteur 4 regroupe les items liés à l'accès aux résultats et à l'usage qu'on leur réserve implicitement (usage formatif) et explicitement (usage administratif) (tableau 5).

Alors qu'ils n'attribuent pas facilement de crédit à la compétence de l'évaluateur, 76% des sujets ne jugent pas que l'enseignement est une fonction trop complexe pour être évaluée (item 16). Par leur réponse à l'item 22, ils contestent à 68% l'accès aux résultats uniquement à des fins administratives. Ils se prononcent aussi en faveur de l'intervention d'un comité départemental à 54% (item 23). Qui plus est, près de 60% des sujets refusent que le professeur soit le seul à avoir accès à ses résultats. Ce pourcentage est encore plus marqué chez les professeurs (67%) que chez les chargés de cours (48%) (test de Mann et Whitney, $\alpha < 0,01$).

Tableau 5
Accès et usage des résultats d'évaluation

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% \pm 3
21. Le professeur évalué devrait être le seul à avoir accès aux résultats de l'évaluation de son enseignement.	2,412	1,365	2	21,7	58,8	19,4
22. Le directeur ne devrait avoir accès aux résultats de l'évaluation de l'enseignement qu'au moment d'une demande de promotion ou de renouvellement.	2,081	1,276	2	16	67,6	16,5
23. Un comité départemental devrait avoir accès aux résultats des évaluations faites par les étudiants.	3,439	1,384	4	53,8	24,2	22,1
20. Les étudiants devraient avoir accès aux résultats de l'évaluation de l'enseignement.	2,742	1,542	3	33,9	46,6	19,6
16. L'enseignement est une fonction trop complexe pour être évaluée.	1,818	1,110	1	11	76,4	12,7
3. Les résultats de l'évaluation faite par les étudiants ne devraient pas être utilisés à des fins de promotion, de permanence et de renouvellement de contrats.	2,894	1,442	3	34,6	42	23,3

L'accès des étudiants aux résultats reste une question sensible qui divise les enseignants (item 20). On relève 47% de désaccord, pudeur marginale dans le paysage nord-américain. Par contre, en ce qui concerne le rôle administratif de l'évaluation (item 3), les professeurs et les chargés de cours ne partagent pas le même avis (test de Mann et Whitney, $\alpha < 0,01$). Nous constatons en effet que si 52% des professeurs n'ont pas d'objection concernant cette question, les chargés de cours au contraire sont moins catégoriques et s'y opposent à 43%. On comprend facilement que les chargés de cours soient moins catégoriques sur cette question, car ils ont beaucoup plus à perdre.

– Facteur 5: biais externes aux professeurs

Ce facteur regroupe au tableau 6 les biais objectifs en dehors du champ de responsabilité du professeur.

Tableau 6
Biais externes aux professeurs

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% \pm 3
26. Les professeurs qui enseignent à des groupes nombreux obtiennent des résultats plus faibles à leur évaluation que ceux qui enseignent à des groupes restreints.	3,173	1,236	3	42,1	26	31,9
27. Les professeurs qui donnent des cours au choix obtiennent de meilleurs résultats à leur évaluation que ceux qui dispensent des cours obligatoires.	3,134	1,236	3	42,2	28,5	29,4

On constate que les sujets ne se prononcent pas clairement sur ces deux questions. Notons que la recherche donne raison aux 42% de sujets interrogés qui pensent que la taille du groupe et le caractère obligatoire du cours peuvent influencer les résultats de l'évaluation. Si les résultats sont peu marqués sur ces questions, on peut avancer l'hypothèse que les professeurs et les chargés de cours sont peu informés des biais susceptibles d'influencer les résultats de l'évaluation de l'enseignement.

– Facteur 6: suites de l'évaluation

Ce facteur regroupe au tableau 7 les deux aspects des suites de l'évaluation, soit les suites administratives (item 38) et les suites formatives (items 32 et 35).

Tableau 7
Aspects relatifs aux suites de l'évaluation

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% \pm 3
38. L'université devrait prendre des mesures disciplinaires à l'égard des professeurs qui éprouvent des difficultés constantes en enseignement.	3,360	1,321	4	53	26,7	20,4
32. Les résultats des évaluations faites par les étudiants sont présentés de façon claire.	3,178	1,272	3	29,3	45,9	24,8
35. Le professeur ne dispose pas de ressources nécessaires lorsque les résultats des évaluations indiquent des aspects à améliorer.	3,256	1,232	3	46,1	27,5	26,4

Un total de 46% des sujets estiment que les conditions ne sont pas remplies pour que l'évaluation puisse jouer pleinement son rôle formatif: manque de clarté des résultats et manque de ressources en cas de difficulté (items 32 et 35). Ce manque de ressources se fait plus cruellement sentir, en toute logique, chez les chargés de cours (55% contre 36% chez les professeurs) (test de Mann et Whitney, $\alpha < 0,01$).

On peut s'étonner que 53% des sujets se prononcent en faveur de mesures disciplinaires en cas de mauvaises évaluations récurrentes. Cette réaction qui paraît excessive est peut-être attribuable à un sentiment grandissant d'injustice dont la responsabilité échoit sans doute à l'établissement plutôt qu'aux enseignants. Une telle réponse dénote peut-être un besoin d'équité ou encore une préoccupation pour la qualité de l'enseignement.

– Facteur 7: biais internes aux choix du professeur

Ce facteur regroupe au tableau 8 les biais liés à des choix du professeur, soit celui de sympathiser avec les étudiants et celui de coter de manière généreuse.

Tableau 8
Biais internes aux choix du professeur

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% \pm 3
31. Les professeurs qui sympathisent avec leurs étudiants obtiennent des résultats plus élevés à leur évaluation que les professeurs réservés.	3,675	1,105	4	63,7	14,3	22
30. Les professeurs qui donnent des notes élevées à leurs étudiants obtiennent de meilleurs résultats à leur évaluation que ceux qui sont sévères.	3,421	1,272	4	51,6	24,7	23,6

Il semble que ces deux biais, non clairement démontrés par la recherche, soient solidement ancrés dans la pensée collective des enseignants. On relève des taux de réponses plus marqués pour les biais internes que pour les biais externes (facteur 5).

– Items supprimés

Les items supprimés sont ceux qui ont obtenu une corrélation item-test inférieure à 0,20 (tableau 9). Bien qu'ils ne soient pas retenus par l'analyse factorielle, ils nous semblent apporter des compléments d'information non négligeables.

Tableau 9
Items ayant obtenu une corrélation item-test inférieure à 0,20

Items	\bar{x}	σ	Médiane	% + 4 et 5	% - 1 et 2	% ± 3
13. Le matériel d'enseignement distribué aux étudiants (plans de cours, notes polycopiées, examens, travaux, etc.) devrait être soumis à l'évaluation par les collègues.	2,987	1,351	3	40,7	37	22,2
28. Plus les étudiants avancent dans leurs études, plus les évaluations qu'ils font des professeurs sont fiables.	3,227	1,130	3	47,1	23,2	29,7
29. Les professeurs qui sont d'excellents communicateurs obtiennent de meilleures évaluations que ceux qui ont des difficultés à communiquer.	4,264	0,792	4	89,2	3,8	7
33. Les délais de remise des résultats des évaluations faites par les étudiants sont trop longs.	3,829	1,298	4	65,3	19,4	15,2
34. Une évaluation faite en cours de session serait plus utile pour améliorer l'enseignement.	3,363	1,365	3	49	16	25
36. Il est facile de consulter un collègue lorsque les résultats de l'évaluation démontrent des difficultés en enseignement.	2,487	1,235	2	20,6	53,2	26,2
37. Les directeurs ont la compétence nécessaire pour aider les professeurs qui éprouvent des difficultés en enseignement.	2,181	1,079	2	12	61,6	26,4
39. L'université devrait entreprendre des actions concrètes afin d'améliorer l'enseignement dans son institution.	4,473	0,787	5	90,9	3,1	6
40. L'université se soucie peu de la qualité de l'enseignement dans son institution.	3,251	1,281	3	44,5	29,6	25,9

Dans la lignée de ce que nous venons de mentionner au dernier facteur, près de 90 % des sujets sont convaincus que les professeurs doués du sens de la communication obtiennent de meilleures évaluations. La question du charisme, même si

elle est controversée dans la documentation scientifique, remporte le plus d'adhésion chez les sujets. La première vertu d'un professeur serait-elle d'être le *Robin Williams* ou la *Liza Minelli* de l'enseignement?

En outre, les sujets revendiquent à 90 % des actions concrètes pour améliorer l'enseignement (item 39). À titre illustratif, citons la question des délais des résultats que les sujets trouvent trop longs (item 33: 65 %) et la volonté de voir l'évaluation jouer son rôle formatif en cours de session (item 34: 50 %).

Par ailleurs, 62 % des personnes interrogées ne jugent pas que leurs directeurs ont la compétence pour les aider, contrairement aux croyances optimistes de l'établissement concernant le rôle du directeur (item 37). Si on observe uniquement les résultats des directeurs de l'échantillon, on constate qu'eux-mêmes doutent à 50 % d'avoir cette compétence et que seuls 26 % d'entre eux pensent la posséder.

Enfin, les sujets interrogés jugent également difficile de consulter un collègue en cas de difficultés (53 %: item 36). Pour eux, ces deux ressources peu coûteuses ne seraient pas la solution idéale. Si les professeurs et les chargés de cours n'arrivent pas à identifier l'évaluateur idéal, ils semblent aussi éprouver de la difficulté à identifier la personne ressource idéale.

En somme, que se dégage-t-il de l'analyse des résultats? D'abord des consensus: 1) l'enseignement n'est pas une fonction trop complexe pour être évaluée; 2) toutes les catégories du personnel enseignant devraient être évaluées sans égard à leur rang et à leur statut; 3) les étudiants ne peuvent être l'unique source d'information consultée, le professeur évalué doit pouvoir s'exprimer sur son enseignement; 4) les excellents communicateurs reçoivent de meilleures évaluations; enfin 5) l'institution devrait entreprendre des actions concrètes afin d'améliorer l'enseignement.

Ensuite, des énoncés acceptés par un peu plus de la moitié des répondants: 1) l'évaluation faite par les étudiants constitue une opération utile qui aide à identifier les points forts et les aspects à améliorer de son enseignement; 2) les commentaires des étudiants sont une source intéressante de suggestions même s'ils peuvent constituer, dans certains cas, une occasion de défoulement; 3) les résultats ne peuvent être accessibles uniquement au professeur, le directeur pourrait les consulter en tout temps et un comité pourrait y avoir accès; 4) les collègues ne sont pas plus compétents que les étudiants pour juger de la qualité de l'enseignement; 5) les professeurs qui sympathisent avec leurs étudiants et ceux qui leur donnent des notes élevées obtiennent de meilleures résultats que les autres; 6) on juge que les directeurs n'ont pas la compétence nécessaire pour aider les professeurs qui éprouvent des difficultés en enseignement et que les collègues ne sont pas davantage accessibles.

Enfin, des perceptions contradictoires: 1) les questionnaires sont-ils de bons moyens d'évaluation; 2) les étudiants sont-ils compétents; 3) devraient-ils s'identifier; 4) les diplômés sont-ils de meilleurs juges; 5) les résultats devraient-ils être utilisés à des fins de promotion?

L'analyse qualitative des commentaires et des suggestions des professeurs et des chargés de cours de cette université apportera peut-être des réponses à ces dernières questions, étant entendu que les répondants commentent principalement les aspects qui leur tiennent à cœur ou qui les dérangent fondamentalement.

L'analyse qualitative

À la fin du questionnaire, les sujets pouvaient exprimer librement leur opinion en répondant à l'invitation suivante: «Commentaires sur les pratiques d'évaluation de l'enseignement et suggestions d'amélioration». Sur les 393 sujets ayant répondu à la partie quantitative, 233 (soit 59 %) ont exprimé des commentaires et suggestions dont 120 professeurs et 113 chargés de cours.

Après une première lecture exploratoire, nous avons noté que les dimensions théoriques citées dans l'introduction pouvaient servir de premier canevas au codage des commentaires. Deux chercheurs ont extrait un total de 620 unités significatives de commentaires dont 319 pour les professeurs et 301 pour les chargés de cours. Ces unités significatives ont été codées selon un réseau de catégories. La validité inter-chercheurs est de 96 %, les 4 % de cas litigieux ont été soumis à la validité de l'expert.

Afin d'être plus conformes aux réponses des sujets, nous avons effectué des modifications mineures aux dimensions retenues. D'abord, nous avons ajouté une dimension appelée «difficultés structurelles». Celle-ci correspond aux paramètres invoqués par les sujets comme étant en dehors de leur champ de responsabilité (état des locaux, programme de cours, disponibilité des outils multimédias, etc.). Ensuite, nous avons subdivisé la dimension «suite de l'évaluation et amélioration de l'enseignement». Il ressort que les sujets, à juste titre, envisagent les stratégies d'amélioration comme étant parfois indépendantes de l'évaluation des enseignements. Enfin, nous avons regroupé les dimensions «moyens» et «sources» car elles nous semblaient bien souvent confondues (par exemple, «l'évaluation par le questionnaire des étudiants est juste et utile»). Notons que la documentation scientifique américaine regroupe généralement ces deux dimensions (Seldin, 1984; Centra, 1993) et que l'analyse factorielle les a également regroupées.

Après avoir classé les commentaires selon le canevas revu des dimensions, nous avons dessiné des sous-réseaux de deux à quatre niveaux, selon les cas, afin d'être le

plus fidèles possible aux commentaires des répondants. Nous présentons les dimensions, catégories et sous-catégories, qui ont un effectif suffisant ou qui ont été jugées pertinentes en fonction du sujet de la recherche.

Analyse des résultats par dimension

Les résultats (tableau 10) montrent, pour chaque dimension, le pourcentage de commentaires (par exemple, 34 % des commentaires portent sur la première dimension : moyens et sources) et le pourcentage de sujets ayant exprimé un commentaire (par exemple, 62 % des sujets ont commenté les moyens et sources d'évaluation). Afin de faciliter la lecture, les pourcentages sont arrondis à l'unité.

Tableau 10
Pourcentages de commentaires et de sujets ayant exprimé un commentaire au sein de chaque dimension

Dimensions	Chargés de cours		Professeurs		Total	
	Sur commentaires N = 301	Sur sujets N = 113	Sur commentaires N = 319	Sur sujets N = 120	Sur commentaires N = 620	Sur sujets N = 233
	%		%		%	
1. Moyens et sources	33	65	34	59	34	62
2. Modalité d'évaluation	20	39	19	35	20	37
3. Buts et utilité	11	27	13	30	12	28
4. Mythes et biais	11	25	13	22	12	23
5. Suites de l'évaluation	10	23	7	17	8	29
6. Difficultés structurelles	9	23	6	16	7	19
7. Amélioration, suggestion hors évaluation	4	8	4	11	4	11
8. Accès aux résultats	2	-	2	-	2	-
9. Personnel évalué	1	-	1	-	1	-

En général, comme pour l'analyse quantitative, les deux populations, professeurs et chargés de cours, se distinguent peu dans leurs réponses. De plus, si on observe les pourcentages totaux, on constate que les sujets sont davantage préoccupés par ce qui est directement lié au moyen choisi et à ses modalités d'évaluation (dimensions 1 et 2). Les buts et utilités de l'évaluation ainsi que les suites (dimensions 3 et 5) sont cités dans une même proportion de sujets (28 % et 29 %). Vient ensuite la dimension 4, mythes et biais, commentée par 23 % des sujets. Les difficultés structurelles (dimension 6) sont mentionnées par 19 % des sujets et, enfin, 11 % des sujets commentent la dimension « amélioration ». Voyons maintenant plus en détail ce que les sujets ont relevé.

– Première dimension : moyens et sources

Soixante-deux pourcent (62 %) des sujets ont commenté cette dimension (tableau 11). Une fois encore, le moyen est souvent confondu avec la source. Trois grandes sous-dimensions ont été commentées, à savoir le questionnaire, les commentaires écrits des étudiants et les autres moyens d'évaluation. Par ailleurs, les sujets font de nombreuses suggestions. Mentionnons que 50 % des commentaires au sein de cette dimension portent sur le questionnaire des étudiants de manière générale.

Tableau 11
Pourcentages calculés sur le nombre de commentaires
pour la catégorie mentionnée en tête de colonne

Moyens et sources N=208	Chargés de cours	Professeurs	Total
	%	%	%
1. Questionnaire par les étudiants			50
a. mise en cause de la compétence de l'étudiant	2	8	10
b. mise en cause de la qualité du questionnaire	8	16	25
2. Commentaires écrits des étudiants sont			16
a. un espace de dévouement, de règlement de compte	6	4	10
b. de grande utilité	3	3	6
3. Autres moyens			13
a. Entretien oral collectif au sein de la classe	3	4	7
b. Questionnaire informel à but formatif	1	1	2
c. Grille d'observation	1	1	2
d. Évaluation du matériel d'enseignement	1	1	2
4. Suggestions			18
a. Diversifier les moyens	3	2	5

Concernant la compétence de l'étudiant à porter un jugement sur la qualité de l'enseignement (10 % des commentaires), les sujets expriment des réserves, les professeurs davantage que les chargés de cours. Plusieurs d'entre eux ont jugé nécessaire de différencier la compétence pédagogique, qu'ils accordent plus volontiers à l'étudiant, de la compétence scientifique. Certains ont également émis l'idée que l'étudiant ne pouvait donner qu'un avis de consommateur en terme de « j'aime et je n'aime pas » et non une opinion construite et approfondie.

Au sujet de la qualité du questionnaire, 25 % des commentaires évoquent des critiques ou des suggestions (surtout les professeurs) et se positionnent en faveur d'une révision du questionnaire actuellement utilisé dans leur département. Il est difficile de dégager des tendances générales parmi ces commentaires. Ceux-ci appa-

raissent souvent contradictoires: trop long, trop court, trop détaillé, pas assez précis. L'hétérogénéité des commentaires peut s'expliquer de différentes manières: d'une part, il est probable que les sujets ont tout simplement des avis différents; d'autre part, étant de secteurs différents, les questionnaires dont ils parlent reflètent des réalités différentes. Cependant, on peut constater que l'ensemble des sujets souhaitent un questionnaire moins standardisé et le plus adapté possible à la spécificité du cours (cours magistral, cours pratique, séminaire, stage). Ceci laisse supposer une prédominance du but formatif de l'évaluation.

Seize pour cent (16%) des observations abordent la question des commentaires écrits des étudiants. La majorité exprime le fait que les commentaires sont une occasion de dévouement ou de règlement de compte et une plus faible proportion que les commentaires sont de grande utilité.

Plusieurs suggestions ont été émises. On demande que les moyens d'évaluation soient plus diversifiés au lieu de se limiter au seul questionnaire des étudiants; on suggère, à titre d'exemples, des moyens comme l'entretien oral, le questionnaire informel, l'évaluation du matériel ou encore la grille d'observation, entre autres.

– Deuxième dimension: modalités d'évaluation

En général, les sujets ont largement commenté cette dimension (37% des sujets) (tableau 12). Dans l'ensemble, ils désirent un meilleur encadrement et une plus grande réflexion concernant les modalités d'application de l'évaluation par les étudiants. Ils soulèvent différents aspects du problème.

Tableau 12
Pourcentages de commentaires calculés sur le nombre
de commentaires pour la catégorie citée

Modalité d'évaluation N=121	Chargés de cours	Professeurs	Total
	%	%	%
1. Identification de l'étudiant	15	11	26
2. Moment d'évaluation			20
3. Délais plus rapides	10	8	18
4. Suggestions :			27
Il faut former les étudiants à l'évaluation	3	3	6
Il faut former les enseignants à l'évaluation	1	3	4
Il faut tenir compte du nombre de cours fréquentés	2	1	3
Seuls les volontaires devraient répondre	0	2	2

Un nombre important de commentaires (26%) portent sur le souhait de voir les étudiants s'identifier lors de l'évaluation par questionnaire, ce qui confirme les résultats de l'analyse quantitative. Certains d'entre eux estiment que ceux-ci auraient une attitude plus responsable et des commentaires plus réfléchis s'ils assumaient la paternité de leurs propos; voici la citation d'un commentaire à ce propos: «De niveau universitaire, l'étudiant qui porte critique sur l'enseignement d'un professeur devrait être en mesure d'assumer son dit».

Les répondants estiment aussi que les délais de remise de résultats sont trop longs (18% des commentaires) et ne permettent pas au processus de jouer pleinement son rôle formatif: «Nous devrions avoir accès aux évaluations plus tôt pour nous aider à améliorer notre enseignement pour la prochaine session», résultat confirmé par l'analyse quantitative.

Notons aussi que 14% des commentaires se prononcent en faveur d'une évaluation à la mi-session, ce qui manifeste un intérêt pour l'aspect formatif de l'évaluation.

Il est intéressant de relever pour cette dimension que 27% des commentaires expriment des suggestions concernant les modalités d'évaluation: former les professeurs et les étudiants à l'évaluation, encadrer correctement le processus, tenir compte de la présence au cours, évaluer plusieurs fois un cours dans la même session.

– Autres dimensions commentées

Généralement, les sujets s'expriment positivement au sujet de l'utilité de l'évaluation. En ce qui concerne les buts, les sujets adhèrent surtout au but formatif, les chargés de cours davantage que les professeurs. À ce sujet, un répondant dit ceci: «Ces évaluations me semblent un outil parmi d'autres pour favoriser l'amélioration de l'enseignement».

La dimension «mythes et biais» est commentée de façon très diversifiée par 23% des sujets. Le premier biais qui attire leur attention est l'influence des notes. Dans les commentaires, les professeurs aussi bien que les chargés de cours sont convaincus que les étudiants bien cotés font des évaluations favorables de leurs cours et que les étudiants moins bien cotés font des évaluations négatives des enseignements qui leur sont dispensés. C'est pourquoi certains d'entre eux souhaiteraient voir les étudiants s'identifier: ainsi, ils pourraient trier les commentaires qu'ils jugent acceptables, à savoir ceux du peloton de tête de la classe; paradoxalement, ils se priveraient des commentaires des étudiants en difficulté. Par exemple, un sujet affirme que «les étudiants devraient s'identifier afin que l'enseignant puisse pondérer la valeur des commentaires».

En général, les sujets sont critiques à l'égard des suites de l'évaluation, estimant que bien souvent les résultats restent lettre morte. Les sujets, surtout les chargés de cours, mettent l'accent sur la nécessité d'envisager une prise en charge de l'enseignant consécutive aux évaluations. Par exemple, un sujet déclare que «le problème, ce n'est pas le processus d'évaluation, c'est le suivi et le manque d'encadrement et de support aux professeurs».

Enfin, quelques sujets font des propositions concernant l'amélioration de l'enseignement indépendamment de tout processus d'évaluation, soit sous forme d'encadrement pédagogique, soit sous forme de groupe de discussion. Quelques suggestions sont intéressantes : favoriser les contacts interuniversitaires, élaborer des projets motivateurs, développer une culture de formation continue, etc.

Bref, les commentaires et suggestions rejoignent les principales préoccupations des répondants et les aspects contradictoires des résultats quantitatifs, soit 1) la mise en cause de la qualité du questionnaire d'évaluation de l'enseignement; 2) la mise en cause de la compétence des étudiants comme évaluateurs principaux; 3) leurs commentaires à la fois utiles et dévastateurs; 4) les modalités d'évaluation qu'on devrait revoir; ainsi que 5) l'absence de suivi de l'évaluation et de son caractère formatif.

Discussion et conclusion

Contrairement à ce qui a été véhiculé pendant des années, l'enseignement, au même titre que la recherche, n'est pas une fonction trop complexe pour être évaluée. On a longtemps considéré que la recherche présentait des caractères objectifs plus facilement abordables, mais les sujets interrogés disent clairement que l'enseignement peut également être évalué. Ils reconnaissent en tout premier lieu l'utilité formative de l'évaluation qui les aide à identifier leurs forces et faiblesses, ils vont même jusqu'à dire que les commentaires écrits des étudiants peuvent être une source de suggestions intéressantes même si, pour certains sujets, ils deviennent une occasion de règlement de compte. Un premier paradoxe ressort : bien qu'ils reconnaissent son utilité, ils doutent de la compétence de l'étudiant.

Si les sujets privilégient l'aspect formatif, quelques-uns n'en nient pas le rôle administratif. D'un côté, on le sait, ils souhaitent une reconnaissance de la fonction d'enseignement dans la promotion des carrières et, d'un autre côté, ils n'acceptent pas massivement cette réalité. Est-ce parce qu'ils ne veulent pas que les étudiants soient la seule source d'information? Est-ce parce qu'ils ne considèrent pas que l'évaluation est équitable et qu'ils voudraient bien que tous les acteurs (chargés

de cours, nouveaux professeurs et professeurs permanents) soient évalués de la même manière?

Concernant le profil de l'évaluateur idéal, les choses sont encore moins claires. Oui, la question intéresse les sujets interrogés; on peut même dire qu'elle les préoccupe. À qui reconnaissent-ils le droit et la compétence de les juger? À en croire les résultats, ils ne reconnaissent cette compétence qu'à eux-mêmes. L'enseignement reste donc une question qui se règle entre les étudiants et le professeur. Il semble difficile d'imaginer qu'un autre acteur, si compétent soit-il, puisse s'en mêler. On demande à la fois de diversifier les sources d'information et d'un même souffle, on n'accepte pas un regard extérieur.

Les sujets disent également qu'ils ne sont pas satisfaits du processus tel qu'il existe et qu'ils manquent de ressources pour donner suite aux évaluations. Ce qui incombe certainement à l'institution, c'est de revoir le processus: adaptation du questionnaire à la situation particulière d'enseignement, présentation claire des résultats, délais moins longs et meilleur encadrement de chacune des étapes (notamment celle de l'administration des questionnaires en classe). Mais lorsque le sujet parle d'évaluation à la mi-session, sans doute doit-il en prendre la responsabilité. Ce type d'évaluation informelle à but strictement formatif est de son ressort et il ne doit pas laisser à d'autres le soin de l'organiser. Lui seul peut décider du moment, du moyen, de l'utilisation des résultats et de la pertinence des commentaires qui remettent ou non son travail en cause.

Concernant le but formatif de l'évaluation, ne faut-il pas redéfinir ce terme et envisager de distinguer deux sous-dimensions? La première est celle dont nous venons de parler, que nous appelons le but micro-formatif: le professeur explore ses résultats de l'évaluation et reconstruit son cours du plus petit détail à l'option la plus centrale, en fonction de ce qu'il estime devoir faire. Pour cela, il ira, au besoin, chercher de l'aide pour peu que l'institution lui en propose, mais il refuse de consulter son directeur dont il ne reconnaît pas la compétence nécessaire en ce domaine. La seconde dimension est celle que nous appelons le but macro-formatif, qu'il ne faudrait pas forcément confondre avec le but administratif. Il s'agit de toutes les mesures qui sont sous la responsabilité collective d'une discipline ou d'un programme et qui ont un impact direct sur la formation des étudiants. C'est dans ce sens, croyons-nous, que les professeurs acceptent de donner à un comité départemental un libre accès à leurs résultats.

Un résultat surprenant et relativement important mérite sans doute quelques commentaires: plus de la moitié des sujets se prononcent en faveur de mesures disciplinaires en cas de difficultés récurrentes en enseignement. Ainsi, le but administratif qui se voulait initialement une reconnaissance et une valorisation de

l'enseignement devient aussi une acceptation de la sanction. En ce sens, les professeurs rejoignent les étudiants, lorsque ces derniers recommandent la publication des résultats et osent même afficher sur le WEB les évaluations tant quantitatives que qualitatives des professeurs. On peut penser que cette prise de position de la part des professeurs et des étudiants tente de mettre en évidence une seule et même réalité. La situation est grave et comme les sujets l'ont par ailleurs exprimés, ils sont en demande d'équité et d'efficacité.

Peu à peu, le profil des sujets interrogés apparaît avec toutes ses contradictions: le professeur, le chargé de cours est méfiant mais demandeur, il est sans illusion, mais plein d'idées; il craint le regard de l'autre, mais se soucie de parfaire son enseignement, il craint les représailles, mais accepte que certaines situations soient inacceptables.

Ces nombreux paradoxes sont sans doute le reflet d'une réalité complexe. Devant cette réalité, il faudra que l'institution puisse restaurer la confiance et le sentiment d'équité. Il lui faut se mobiliser afin de trouver les actions concrètes qui pourront améliorer l'enseignement. L'évaluation de l'enseignement, correctement menée, pourrait être une de ces actions concrètes, une parmi d'autres, car l'évaluation n'est pas, de façon magique, synonyme d'amélioration de l'enseignement.

NOTE

1. Nous remercions toutes les personnes qui nous ont permis de réaliser cette étude. D'abord, les juges qui, par leurs critiques, ont facilité le choix de nos énoncés. Les directions de facultés, départements et écoles ainsi que les syndicats de chargés de cours qui ont pris en charge la photocopie et la distribution de notre questionnaire et, finalement, les professeurs et chargés de cours qui ont accepté d'y répondre. Sans leur collaboration, cette étude n'aurait pas été possible.

Abstract – This article attempts to investigate the complexity of the teaching evaluation process as experienced in Quebec universities. The authors use both a quantitative and a qualitative methodology to examine the major types of responses given by a population of lecturers and professors as related to objectives, means, sources, and modalities of evaluation of teaching. The findings suggest that the subjects, regardless of status, are in agreement as to the means and modalities of evaluation. However, the profile of the ideal evaluator is difficult to describe. Generally, the authors note a paradox in the responses of subjects interviewed. This is probably the reflection of the complexity of a question that is rather delicate in the university milieu.

Resumen – Este artículo trata de abordar la complejidad de los procesos de evaluación de la enseñanza, tales como ellos son vividos y sentidos en las universidades quebequenses. Utilizando una metodología a la vez cuantitativa y cualitativa, el presente estudio pone en relieve las grandes tendencias de las respuestas dadas por una muestra de profesores asistentes y de profesores regulares respecto de los objetivos, medios, fuentes y modalidades de la evaluación de la enseñanza. Se destaca en el estudio que los sujetos concuerdan, sin distin-

ción mayor en lo que hace a su estatus institucional, sobre los medios y las modalidades de la evaluación. Sin embargo, el rol de la evaluación ideal parece difícil de definir. En forma general, los autores ponen en evidencia muchas paradojas en las respuestas de los encuestados, reflejo quizás de la complejidad de un sujeto delicado dentro del mundo universitario.

Zusammenfassung – Dieser Artikel bemüht sich um eine Bestandsaufnahme der sehr komplexen Beurteilungsmechanismen für Lehrveranstaltungen, wie sie an den verschiedenen Universitäten Québecs vorgefunden werden. Die sowohl quantitativ als auch qualitativ orientierte Methodik hat die gegenwärtigen Antwort-tendenzen freigelegt, die sich aus der Befragung der Kursbeauftragten und Professoren über Ziele, Mittel, Ressourcen und Modalitäten im Lehrbereich ergeben. Daraus geht hervor, dass beide Gruppen ohne nennenswerte Unterschiede mit den Mitteln und Modalitäten der Beurteilungen einverstanden sind. Was dagegen das Profil eines idealen Beurteilers angeht, so scheint der Konsens schwieriger zu sein. Ganz allgemein haben die Autorinnen in den Antworten der Befragten zahlreiche Paradoxe aufgedeckt, was als sicheres Zeichen dafür gewertet werden kann, dass die Beurteilungsproblematik für die akademische Welt ein heikles Thema darstellt.

RÉFÉRENCES

- Abrami, P.C. et d'Apollonia, S. (1990). The Dimensionality of ratings and their use in personnel decisions. Student ratings of instruction: Issues for improving practice. *New Directions for Teaching and Learning*, 43, 97-111.
- Bernard, H. (1992). *Processus d'évaluation de l'enseignement: Théorie et pratique*. Laval: Éditions Études vivantes.
- Bernard, H. et Bourque, S. (1999). Portrait des politiques et des pratiques d'évaluation, d'amélioration et de valorisation de l'enseignement des universités québécoises. *Res Academica*, 15(1-2), 33-60.
- Bernard, H. et Normand, S. (À paraître). Réaction des professeurs en fonction de leur discipline à l'évaluation de l'enseignement universitaire *Mesure et évaluation en éducation*.
- Bernard, H. et Trahan, M. (1988). Analyse des politiques institutionnelles d'évaluation de l'enseignement des universités canadiennes. *La revue canadienne d'enseignement supérieur*, XVIII(3), 51-67.
- Brinko, K.T. (1991). The interactions of teaching improvement. Effective practices for improving teaching. *New Directions for Teaching and Learning*, 48, 39-49.
- Brinko, K.T. (1993). The practice of giving feedback to improve teaching. What is effective? *Journal of Higher Education*, 64(5), 574-593.
- Cashin, W.E. (1990). Students do rate different academic fields differently. Student ratings of instruction: Issues for improving practice. *New Directions for Teaching and Learning*, 43, 113-121.
- Centra, J.A. (1993). *Reflective faculty evaluation. Enhancing teaching and determining faculty effectiveness*. San Francisco (CA): Jossey-Bass.
- CRÉPUQ (1996). *Rapport sur les politiques et pratiques des établissements universitaires du Québec en matière d'évaluation et de valorisation de l'enseignement*. Montréal: CRÉPUQ.
- Donald, J.G. (1991). The Commission of inquiry on Canadian university education: The quality and evaluation of teaching. *Revista Iglu*, 1, 157-173.
- Donald, J. et Saroyan, A. (1991). *Assessing the quality of teaching in Canadian university*. Commission of inquiry on Canadian university education, Research Report, n° 3.
- Doyle, K.O. Jr. (1983). *Evaluating teaching*. Toronto: Lexington Books.

- Feldman, K.A. (1988). Effective college teaching from the students' and faculty's view: Matched and mismatched priorities? *Research in Higher Education*, 24(4), 291-344.
- Feldman, K.A. (1978). Course characteristics and college students' ratings of their teachers and courses: What we know and what we don't. *Research in Higher Education*, 9(3), 199-242.
- Feldman, K.A. (1976). Grades and college students evaluations of their courses and teachers. *Research in Higher Education*, 4(1), 69-111.
- Feldman, K.A. (1979). The Significance of circumstances for college students ratings of their teachers and courses. *Research in Higher Education*, 10(2), 149-172.
- French-Lazovik, G. (1981). Peer review: Documentary evidence in the evaluation of teaching. *Handbook of Teacher Evaluation* (p. 73-89). Beverly Hills (CA): Sage.
- Marsh, H.W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 255-388.
- McKeachie, W. J. (1979). Students ratings of faculty: A reprise. *Academe*, 65(6), 384-397.
- Millman, J. (1981). *Handbook of teacher evaluation*. Beverly Hills (CA): Sage.
- Murray, H.G. (1984). The Impact of formative and summative evaluation of teaching in North American Universities. *Assessment and Evaluation in Higher Education*, 9(2), 117-132.
- Scriven, M. (1981). Summative teacher evaluation. In *Handbook of Teacher Evaluation* (p. 244-271). Beverly Hills (CA): Sage.
- Seldin, P. (1993). *Successful use of teachings portfolios*. Bolton (MA): Anker Publishing Co.
- Seldin, P. (1991). *The teaching portfolio. A practical guide to improved performance and promotion/tenure decisions*. Bolton (MA): Anker Publishing Co.
- Seldin, P. (1990). *How administrators can improve teaching. Moving from talk to action in higher education*. San-Francisco (CA): Jossey-Bass.
- Seldin, P. (1984) *Changing faculty evaluation practices*. San Francisco (CA): Jossey-Bass.
- Thivierge, A. et Bernard, H. (1996). Croyances des étudiants à l'égard de l'évaluation de l'enseignement. *Mesure et évaluation en éducation*, 18(3), 59-84.
- Weimer, M. et Firing Lenze, L. (1991). Instructional interventions: A review of the literature on efforts to improve instruction. *Higher Education: Handbook of Theory and Research* (vol. 11). Texte reproduit en 1994, dans K.A. Feldman et M.B. Paulsen (dir.), Teaching and learning in the college classroom (p. 653-682). Needham Heights (MA): Ginn Press.
- Wright, A. (1995) *Teaching improvement practices: Successful strategies for higher education*. Bolton (MA): Anker Publishing Co.