



MULTI DIMENSIONNALITE DE L'EVALUATION DE L'ENSEIGNEMENT UNIVERSITAIRE DANS LA LITTERATURE ANGLOPHONE

Nathalie Younes

► To cite this version:

Nathalie Younes. MULTI DIMENSIONNALITE DE L'EVALUATION DE L'ENSEIGNEMENT UNIVERSITAIRE DANS LA LITTERATURE ANGLOPHONE. Véronique Bedin. L'évaluation à l'université. Évaluer ou conseiller ?, Presses Universitaires de Rennes, 2009. hal-01121475

HAL Id: hal-01121475

<https://hal.science/hal-01121475>

Submitted on 1 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CHAPITRE

MULTI DIMENSIONNALITE DE L'EVALUATION DE L'ENSEIGNEMENT UNIVERSITAIRE DANS LA LITTERATURE ANGLOPHONE

Nathalie YOUNES

Université Blaise Pascal, France

INTRODUCTION

L'évaluation de l'enseignement universitaire qui est en Europe une préoccupation relativement récente est explorée aux Etats Unis par les chercheurs et pratiquée depuis plus de soixante ans, influençant et préfigurant les évolutions mondiales dans ce domaine. En dépit de cette expérience confirmée, le débat autour de l'évaluation de l'enseignement reste toujours vivace, et même renouvelé par les phénomènes de mondialisation de l'enseignement supérieur devenu un marché concurrentiel, avec toujours plus de pression évaluative venant aussi bien des financeurs et de la société civile que des étudiants. Dans une telle optique compétitive, les avancées technologiques de l'informatique et de l'internet, dans leurs capacités à stocker, traiter, mettre à disposition et diffuser l'information amplifient les démarches fondées sur des indicateurs de résultats quantifiés, utilisés pour les comparaisons et le contrôle. La plupart des pays développent désormais une logique de pilotage de l'éducation qui prend appui sur des critères communs permettant la comparaison et le classement. La vision de l'évaluation-conseil comme mode d'accompagnement des pratiques évaluatives (Bedin & Jorro, 2007), s'inscrit dans une dynamique qui insiste sur l'importance de l'aide à la décision et la veille prospective, pour soutenir les acteurs dans les grandes transformations de l'enseignement supérieur. Nous présentons une lecture des conceptions anglo-saxonnes de l'évaluation de l'enseignement supérieur en interrogeant la multidimensionnalité notamment dans ses composantes réflexives et régulatrices.

1. DE LA QUALITE

Dans l'enseignement supérieur en redéfinition, non seulement les offres de formation sont plus nombreuses et les demandes plus importantes avec la diversification et l'augmentation du public étudiant mais désormais s'impose l'obligation de rendre des comptes sur les moyens et les résultats des formations en place. La mise en place de mécanismes d'accréditation et d'assurance qualité pour la régulation de la formation dans un marché concurrentiel est l'une des pierres angulaires des réformes de l'enseignement supérieur. Sous ce terme fédérateur, différentes définitions de la qualité sont repérables ainsi que différentes formes d'évaluation. L'accréditation et le benchmarking sont deux formes particulièrement significatives de ce modèle basé sur l'explicitation de standards de qualité et leur comparaison.

1.1. Qualité et accréditation comme modèle normatif

Les démarches d'accréditation ont d'abord été conçues dans et pour le contexte américain afin de réguler un système d'enseignement supérieur compétitif et totalement décentralisé (Murray, 2001). Aux Etats-Unis, les agences privées d'accréditation (depuis 1952) et les

associations d'université¹ sont les structures incontournables de l'enseignement supérieur, l'état intervenant notamment par la reconnaissance d'un certain nombre d'agences d'accréditation². Les principales fonctions de l'accréditation sont : faire la vérification externe qu'un programme ou une institution est à la hauteur des standards, établir des critères de certification, aider les étudiants à situer les similitudes et les différences des institutions et des programmes, aider les institutions à déterminer l'acceptabilité du transfert de crédits (Gill, 2006). Reconnaissance officielle externe d'un programme validé selon des standards spécifiés comme les qualifications des équipes, les activités de recherche, l'admission des étudiants, les ressources d'apprentissage, l'accréditation peut aussi être basée sur des standards académiques ou sur l'estimation des compétences professionnelles (Harvey, 2004). Elle repose sur l'autoévaluation, l'évaluation par les pairs/experts et le rapport public ainsi que sur une grande variété de méthodes : analyses de documents, études des indicateurs de performance, entretiens, observations de cours, enquêtes de satisfaction, enquêtes auprès des anciens étudiants, enquêtes auprès des employeurs.

Ces démarches ont été importées en Europe à partir des années 90 dans le double objectif de garantir la qualité mais aussi l'amélioration des formations proposées (Sursock, 2000). Sheele (2004) distingue les démarches qui se développent dans des pays tels que l'Allemagne, l'Irlande, l'Italie, les Pays-Bas, la Norvège, l'Autriche, l'Espagne et la Belgique mettant l'accent sur l'assurance qualité et des formes d'accréditation existant en Europe de l'Est (Russie, Pologne, Hongrie...) développées par l'état pour réguler l'émergence rapide d'établissements d'enseignement supérieur privés en garantissant un minimum de standards de qualité. Les communiqués des ministres de l'enseignement supérieur dans le cadre du processus de Bologne (1999), Prague (2001), Berlin (2003), Bergen (2005), Londres (2007) font appel de façon croissante à une coopération européenne et entre les organisations universitaires (ENQUA, European Network of Quality Assurance EUA, European University Association, ESU et EURASCHE) afin de développer un ensemble de règles et de procédures en matière d'assurance qualité.

Bien que les définitions et les pratiques de l'accréditation soient différentes en Europe et aux Etats-Unis et à l'intérieur de l'Europe, une sorte de convergence s'établit. Les procédures américaines d'accréditation incorporent les principes de l'assurance qualité (Alstete, 2004) alors que les principes de l'assurance qualité développés en Europe de l'Ouest intègrent progressivement les habitudes américaines. Les standards européens et les guides de l'assurance qualité définis par l'ENQA (2007) ainsi que le développement d'agences d'accréditations comme l'ECA (European Consortium for accreditation) manifestent ce mouvement.

Accréditation individuelle et /ou d'équipe : l'impulsion anglaise

La place des étudiants dans l'accréditation individuelle des enseignants et dans l'importance accordée à la formation continue a été déterminante en Angleterre. En effet, le développement de l'accréditation des enseignants dans ce pays provient de l'union nationale des étudiants (National Commission on Higher Education, 1997). Appuyés par des parents qui étaient de plus en plus sollicités financièrement et par le gouvernement, ils réclamaient le développement de la formation obligatoire des enseignants. L'idée du dispositif SEDA (Staff and Educational Development Association) était de définir le résultat requis d'une formation professionnelle, pour avoir une base commune et que les responsables institutionnels y accordent les ressources nécessaires (Beaty, 2001). Ce système trace les grandes lignes des

¹ Parmi les plus importantes le CHEA regroupe 3000 établissements d'enseignement supérieur (<http://www.chea.org>) et valide 60 programmes d'accréditation

² La liste des agences d'accréditation est disponible sur le site du département américain de l'éducation. http://www.ed.gov/admins/finaid/accred/accreditation_pg8.html

résultats d'apprentissage et un ensemble de valeurs et de principes sous-jacents. Les programmes de développement professionnel des enseignants du supérieur, démontrant qu'ils soutiennent et évaluent les enseignants suivant ce cadre de travail, sont reconnus. Les enseignants qui participent à un tel programme (nécessitant à peu près une année d'études à temps partiel) reçoivent un certificat d'accréditation. Initiés en 1992, il y avait 58 programmes reconnus en 2001, fondés principalement sur l'observation de l'enseignement, l'évaluation du portfolio et sur un accompagnement des enseignants (Beaty, 1997). Il a été en effet considéré que les enseignants ont besoin de soutien pour développer les compétences requises, ce qui a entraîné la recherche de spécialistes du développement universitaire. Les pressions sur les ressources et le « rendre compte » (« accountability ») ont nécessité de développer le travail d'équipe (Beaty, 2001), reconnu par les différents acteurs (étudiants, administrateurs et enseignants) comme facteur d'efficacité (QAA, 2000).

1.2 . Le benchmarking et le ranking

Parmi les méthodes issues du monde de l'entreprise et du management de la qualité, le benchmarking est un processus de comparaison sociale sur des standards identifiés. Cette notion de positionnement relatif, littéralement marquage des frontières et des territoires, vient de l'art militaire et consistait à connaître le plus précisément possible la composition et l'armement des troupes de l'adversaire pour mieux le contrer. Dans le monde contemporain, c'est une notion économique suivant laquelle l'avantage vient tout autant de la position relative que l'on a sur un marché que de la force absolue de l'entreprise. Ce système d'évaluation et d'amélioration de la qualité consiste à identifier une organisation qui réussit quelque chose, à étudier comment elle le fait, et à appliquer ensuite les résultats de l'étude pour optimiser les processus et les résultats de sa propre institution. En fournissant un standard de comparaison, le benchmarking peut être considéré comme un support aidant à s'améliorer, utilisable dans le monde de l'enseignement supérieur (Jackson & Lund, 2000). La littérature montre que cette démarche a d'abord surtout été appliquée aux procédures administratives et dans les activités d'enseignement liées au commerce et à la formation continue (Alstete, 1996). Dans le manuel élaboré par McKinnon, Walker et Davis (2000) à destination des institutions d'enseignement supérieur australiennes, 67 « benchmarks » sont regroupés en 9 catégories : la gouvernance, l'impact, la finance, l'enseignement et l'apprentissage, le soutien des étudiants, la recherche, l'internationalisation, l'équipe. Les dix « benchmarks » concernant l'enseignement sont les plans d'enseignement et d'apprentissage, les processus d'établissement de cours, l'expertise d'enseignement, l'environnement d'enseignement, les comptes-rendus universitaires, l'utilité des cours, l'apprentissage, l'équité, la satisfaction des étudiants, l'insertion professionnelle. Chaque benchmark contient de l'information à la fois sur ce qui le caractérise, sur les sources de données, des exemples de bonne pratique et des profils descriptifs de niveaux de performance. Au Royaume Uni, une liste indicative précise cinq paramètres : habiletés clés, résultats de l'apprentissage, expérience de l'étudiant, systèmes départementaux pour gérer les standards, environnement d'apprentissage et processus éducationnels (Jackson & Lund, 2000), ce qui met en évidence le large spectre retenu pour appréhender la qualité.

Le ranking qui établit un classement des établissements parachève ce dispositif de comparaison hiérarchisée. Mais l'efficacité de ces dispositifs en termes d'amélioration est loin d'être prouvée. Par exemple, le manuel de McKinnon n'est pas perçu comme un outil d'aide à l'amélioration mais plutôt comme un outil de contrôle ; rigide, formel et peu productif (Garlick, Pryor, 2004).

1.3. Une position alternative interprétative et critique de l'évaluation de l'enseignement

Car si la tendance générale dans les différents pays montre que le recours aux indicateurs va augmentant, à la fois comme mesure directe et comme partie des évaluations qualitatives d'experts et qu'il constitue une norme et une obligation, il est à noter qu'une position interprétative et critique de l'évaluation de l'enseignement universitaire s'est également affirmée comme alternative.

A partir du constat du peu de connaissance objective de l'enseignement, Cranton argumente pour l'inscription de l'évaluation dans un paradigme interprétatif et critique de la connaissance communicative et émancipatrice. D'après ce chercheur (2001, p. 13), la connaissance de l'enseignement objective, empirique et basée sur des principes de causalité invariants s'avère bien faible. Tout enseignement dépend non seulement de sa forme propre mais également des individus qui apprennent et travaillent ensemble, ainsi que du contexte dans lequel il prend place. Si les principes communément admis d'un enseignement efficace sont examinés, il apparaît qu'ils ne sont pas invariants. Par exemple, est-il toujours préférable de bien organiser un enseignement ? Certains étudiants peuvent être étouffés par la structure et certains enseignants peuvent être naturellement plus intuitifs et coulants dans leur style. Certains sujets ne conduisent pas non plus à une organisation claire. La pensée critique des étudiants est peut-être favorisée par des tours et détours dans le traitement de la classe... Quantité d'exemples montrent que ce qui fonctionne le mieux dépend des personnes et des contextes. Il n'y a pas une seule façon d'enseigner. On dispose de certaines généralisations fondées sur les recherches empiriques mais pas de vérités invariables. Ceci veut dire qu'en dépit du désir d'avoir une connaissance instrumentale sur l'enseignement, il y en a très peu même si des accords entre les étudiants, les enseignants et les administrateurs sur les caractéristiques d'un bon enseignement ont pu être mis en évidence.

Par ailleurs, le constat du peu d'efficacité pour l'amélioration des cadres de benchmarking tels que celui utilisé en Australie a conduit Garlick et Pryor (2004) à proposer un dispositif basé sur la collaboration des différents acteurs concernés par l'évaluation et mettant en jeu l'implication, le dialogue réflexif et l'apprentissage collectif dans la recherche, la planification et l'analyse d'actions d'amélioration. La visée praxéologique et l'implication des différents types d'acteurs rapproche cette démarche de celle de l'évaluation conseil.

2 – DES FORMES D'EVALUATION COMPLEMENTAIRES

La littérature anglo-saxonne a largement décrit différentes formes d'évaluation de l'enseignement. Les résultats des étudiants, l'évaluation de l'enseignement par les étudiants, l'évaluation par les pairs et les portfolios constituent les quatre modalités principales qui bien que se présentant de façon spécifique sont à articuler pour une évaluation multidimensionnelle (Scriven, 1983b), étant donné qu'elles permettent d'en appréhender différentes facettes et points de vue : l'enseignement et l'apprentissage, l'étudiant et l'enseignant, l'individu et le collectif ainsi que leurs interactions.

2.1. Evaluation de l'enseignement par les résultats des étudiants

L'évaluation de l'enseignement supérieur tend de plus en plus à montrer le bénéfice en « valeur ajoutée » en se centrant sur les résultats des étudiants en termes de compétences acquises. On peut voir ce mouvement au niveau des procédures d'accréditation américaines (Gill, 2006). Mais la question se pose de savoir si les résultats des étudiants peuvent être utilisés comme des indicateurs valides de la qualité de l'enseignement dans les environnements universitaires complexes d'aujourd'hui. En effet, il est difficile de les définir

et de les mesurer. De plus, les résultats des étudiants sont difficiles à corrélérer avec l'enseignement. Au moins trois sources majeures d'influence, autres que l'enseignement en lui-même, interviennent sur les performances des étudiants : institutionnelles, sociales et dispositionnelles. Par exemple, Ramsden et Entwistle (1981) ont étudié comment le climat d'un département ou d'une institution influençait l'apprentissage des étudiants en favorisant ou en défavorisant leur engagement. De larges études empiriques comme celle d'Astin (1978) ont montré comment l'établissement choisi influençait les changements des étudiants et le développement de carrière. Les facteurs dispositionnels affectant l'apprentissage des étudiants incluent l'attitude des étudiants par rapport à l'instruction, les capacités individuelles, l'étape de développement et la volonté à accepter la responsabilité de l'apprentissage. Parmi les facteurs sociaux, l'argent, les expériences et les valeurs sont aussi décisifs. Les travaux sur l'apprentissage situé ont montré qu'on ne peut isoler un facteur et en mesurer l'impact causal sur l'apprentissage individuel parmi ceux qui s'entrelacent d'une manière complexe dans une communauté d'apprentissage (Brown and Duguid, 1996 ; Kirschner and Witson, 1997). Enfin, le temps est aussi une dimension critique de l'apprentissage. Quels résultats sont observables pendant les cours ? En quoi sont-ils différents de ceux observés à la fin de la période d'instruction ? Les résultats les plus significatifs seront peut-être observés plusieurs années plus tard.

Fenwick and Parsons (2000) ont repéré les différentes approches de la mesure des résultats des étudiants en listant les avantages et les inconvénients. Il s'agit d'apprécier comment les utiliser pour évaluer l'enseignement de manière productive. Ces chercheurs catégorisent ces approches en indicateurs « micro », ciblés au niveau d'un enseignement particulier et indicateurs « macro », ciblés sur des cohortes, des institutions, des programmes. Parmi les indicateurs micro, sont distingués les tests et examens construits par les enseignants, les productions d'étudiants, les observations de performances d'étudiants, les observations en cours et les autoévaluations des étudiants. Les critères étant variables, la validité de ces indicateurs est limitée. Les indicateurs « macro » sont ciblés aux niveaux de programmes entiers ou d'institutions. Ils emploient à la fois des méthodes quantitatives et qualitatives : Enquêtes d'attitudes et d'intérêts, tests de réussite standardisés, études longitudinales, enquêtes auprès des anciens étudiants.

Alors que les tests de réussite standardisés se répandent en tant qu'outils de contrôle de la qualité malgré leurs limites avérées, Fenwick et Parsons (2000) argumentent pour la combinaison de différentes mesures aussi bien dans l'évaluation d'un enseignement que dans l'évaluation d'un programme. Fenwick (2001) insiste particulièrement sur l'utilisation de ces résultats dans une perspective de soutien et d'amélioration de l'enseignement en préconisant leur contrôle par les enseignants eux-mêmes.

2.2. Evaluation de l'enseignement par les étudiants (EEE)

L'évaluation de l'enseignement supérieur est née en Amérique du Nord avec l'évaluation de l'enseignement par les étudiants. Doyle (1983) a montré que ce sont les associations étudiantes qui en sont à l'origine : l'utilisation des questionnaires a commencé à se répandre à partir des années 60 quand les étudiants ont réussi à obtenir une plus grande représentation dans les universités. L'origine de l'évaluation est donc contemporaine d'une autre position, plus active et engagée des étudiants par rapport à l'enseignement. Elle s'est beaucoup développée dans les universités américaines, canadiennes, australiennes, anglaises et québécoises³, pour devenir la principale forme d'évaluation de l'enseignement universitaire. Elle est principalement utilisée à des fins administratives d'embauche et de

³ 29 % des collèges américains la pratiquaient en 1973, 68 % en 1983 et 86 % en 1993 (Seldin, 1993). De nos jours 90 à 100% des collèges et des universités américaines sont engagés dans cette pratique (Trout, 2000).

sélection des enseignants, et d'amélioration de l'enseignement mais aussi comme information pour choisir les cours (Wilhelm & Comegys, 2004). Dans la communauté universitaire, elle reste controversée dans ses objectifs comme dans son utilisation.

Dans les pays anglo-saxons, un champ vivace de recherches existe quant aux modalités, à la validité et à l'utilité des EEE.

2.2.1 Modalités

Parmi les instruments aujourd'hui largement diffusés dans les pays anglo-saxons, on trouve notamment l'IDEA (Instructional Development and effective Assessment), initié par Hoyt (1973) et révisé par Hoyt, Chen, Pallett et Gross (1999), et le SEEQ (Student Evaluation of Educational Quality, développé et validé par Marsh (1982, 1984, 1987) pour évaluer un enseignement en 9 dimensions de l'enseignement et 35 items. D'autre part pour évaluer un programme dans sa globalité le *questionnaire d'expérience de la formation* (CEQ) a été développé en Australie pour rassembler le feedback des diplômés sur l'expérience de leur cursus dans cinq domaines : la qualité de l'enseignement ; la clarté des buts ; la quantité de travail ; les méthodes d'évaluation ; le développement d'habiletés génériques (Johnson, 1999 ; Ramsden, 1991). Son développement est basé sur une recherche de Ramsden et Entwistle (1981) en Grande Bretagne sur les facteurs qui promeuvent l'apprentissage en profondeur. L'instrument se distingue des évaluations par les étudiants classiques, qui sont centrées sur les enseignements (voire les enseignants) pris individuellement, car il s'intéresse au programme d'étude dans son ensemble. Cette mesure en 24 items a été montrée comme étant stable, fiable et significative quant à l'expérience des étudiants de leurs programmes d'étude (Hand & Trembath, 1998). (Exemples d'items : échelle de bon enseignement : l'équipe de ce programme me motivait à travailler de mon mieux, échelle de quantité de travail : la quantité de travail était trop importante, habiletés génériques : le programme développait ma capacité à résoudre les problèmes). Les données des CEQ sont désormais collectées chaque année auprès des diplômés en Australie et utilisées à des fins de diagnostics par les institutions et unités organisationnelles comme les facultés et les départements. Les résultats sont aussi utilisés par les universités pour publier des guides assistant les étudiants dans le choix de leur programme d'enseignement. Le CEQ est un indicateur de performance utile qui semble avoir permis des améliorations de la qualité de l'enseignement et du service aux étudiants (Johnson, 1999). Il permet aussi des comparaisons dans le temps et entre institutions à l'intérieur d'un champ disciplinaire. Néanmoins, la signification des scores doit toujours être interprétée dans le contexte des buts spécifiques institutionnels et de programmes (Hand & Trembath, 1998).

2.2.2 Validité et utilité des évaluations de l'enseignement par les étudiants

Les EEE sont valides si elles reflètent l'efficacité de l'enseignement. Les chercheurs s'accordent pour considérer qu'il n'y a pas de critère univoque de l'efficacité de l'enseignement et Scriven (1983a) a particulièrement bien analysé les biais des indicateurs les plus couramment utilisés que sont des mesures de l'apprentissage des étudiants, les jugements des anciens étudiants et les observations d'experts. La question de la validité des EEE a donc été traitée par l'étude des corrélations entre celles-ci et d'autres mesures imparfaites et partielles de l'efficacité de l'enseignement. De nombreuses recherches (voir Marsh, 1987 ; Abrami et d'Apollonia, 1990 pour des revues) montrent de manière non équivoque que les enseignants bien évalués par les étudiants sont aussi ceux dont les étudiants apprennent le plus, ceux à propos desquels ils font des commentaires favorables dans les remarques et les entretiens, ceux qui sont bien évalués par des observateurs extérieurs, les anciens étudiants et les collègues. Ces résultats en faveur de la validité des EEE peuvent paraître contre intuitifs

mais ce sont les résultats les plus éprouvés scientifiquement à ce sujet. Les biais affectant les EEE ne doivent cependant pas être minimisés⁴.

Les EEE sont censées améliorer l'enseignement à deux niveaux. Au niveau institutionnel, elles peuvent influencer les décisions prises au sujet des enseignants quant à leur titularisation ou promotion et quant à l'attribution des cours. Au niveau individuel, il s'agit de « feedback » à destination de l'enseignant l'informant de la perception de son enseignement par les étudiants. Les recherches anglo-saxonnes ont principalement étudié les effets au niveau individuel des enseignants et montré l'utilité des feedback provenant des étudiants pour l'amélioration de l'enseignement (Murray, Jelley, & Renaud, 1996) mais aussi la variabilité de cette utilité, dépendante d'un certain nombre de facteurs comme la valeur et la pertinence de ces feedback pour les enseignants (Geis, 1991 ; Yao, Weissinger & Grady, 2003) et le contexte de cette utilisation. Ce sont surtout les enseignants qui en plus de ces évaluations participaient à des groupes avec discussion de stratégies d'amélioration qui amélioreraient à la fois leurs résultats d'évaluation par les étudiants et les résultats des étudiants (Marsh et Roche, 1993).

2.3. Evaluation par les pairs dans un but formatif

Il existe plusieurs formes d'évaluation de l'enseignement par les pairs, basées sur des interactions et du soutien présentées dans les années 90. Ainsi les *programmes pairs-partenaires* décrivent des stratégies qui impliquent que des collègues - en général des paires d'enseignants - travaillent ensemble de manière systématique pendant au moins un semestre dans une relation d'aide qui inclut des visites de cours mutuelles, des discussions avec les étudiants et des rencontres régulières pour discuter de leur enseignement et des façons de l'améliorer (Morrison, 1997). Le *diagnostic sur l'enseignement en petit groupe*, (SGID, Lenze, 1997) est une technique d'analyse de l'enseignement qui requiert 20 ou 30 minutes du temps de cours dirigé par un collègue ou un animateur d'équipe entraîné dans l'utilisation de cette méthode.

Pour Smith (2001), ces activités d'évaluation par les pairs sont très importantes car elles reflètent un aspect de ce que veut dire être plus professionnel dans son enseignement à savoir être activement engagé dans le contrôle de l'impact de son propre travail et dans l'amélioration de son efficacité. Cependant, elles peuvent ne pas entraîner de transformation radicale. Il se réfère au travail d'Argyris et Schön (1974) qui distinguent deux boucles d'apprentissage. La première boucle conduit à utiliser des stratégies d'action plus efficaces pour atteindre ses buts sans questionner ses valeurs de base ; par exemple améliorer ses cours magistraux sans questionner le principe du cours magistral. Une seconde boucle implique le changement des critères ou valeurs de base par lesquels l'efficacité est jugée ; par exemple améliorer l'apprentissage plutôt que le cours, et peut-être ne plus faire de cours dans sa forme traditionnelle.

Hutchings (1996) décrit de manière précise comment le feedback par un travail collaboratif entre pairs peut aider à identifier et à améliorer les méthodes pour créer une érudition sur l'enseignement et l'apprentissage : cercles d'enseignants, observations mutuelles en classe, mentoring et coaching, équipes d'enseignement et aussi collaborations inter campus. Le portfolio est un moyen d'améliorer son enseignement et de le rendre public. Les journaux sont aussi de bons moyens de diffuser une pratique et de la soumettre à critiques. Par

⁴ Comme par exemple, les effets de halo (Aleamoni, 1999) et de récence (Dickey & Pearson, 2005).

exemple, *The Journal of Scholarship of Teaching and Learning* est un journal diffusé sur le Web depuis 2001 et dont les articles sont revus par les pairs⁵.

2.4. Portfolios

On peut considérer le développement des portfolios d'enseignement comme en grande partie réactif, avec un essai de réappropriation du contrôle du processus d'évaluation par les enseignants et en élargissant les bases de l'appréciation de l'enseignement (Knapper, 1995). Cet élargissement tend à assurer que différents contextes sont pris en compte dans l'évaluation de l'enseignement et de l'apprentissage. L'utilisation des portfolios dans l'enseignement supérieur qui date des années 80 est un concept originaire du Canada progressivement répandu dans le monde entier. Shore a le premier exprimé l'idée que c'est aux enseignants universitaires de construire eux-mêmes la démonstration de leur efficacité en tant qu'enseignants, « a portfolio of evidence » (Shore, 1975, p. 8). Le comité publia ensuite son guide pour le dossier d'enseignement (Shore et al., 1986) qu'ils définissent comme un résumé des principales réalisations d'un professeur. Ils défendent une approche pour présenter un bon enseignement qui anticipe la notion de « classroom research » de Cross (1986) et le concept de « savoir sur l'enseignement » (scholarship of teaching) (Boyer, 1990). Ils préconisent alors un dossier de 3 pages, bien qu'aujourd'hui les dossiers soient plus longs. Ils recommandent de clarifier et documenter les responsabilités d'enseignement, de sélectionner les critères de l'enseignement efficace, de les prouver, de les résumer et de rassembler si besoin un matériel exemplaire. Suit une liste de 49 catégories (ou types de preuves) qui peuvent être incluses dans un portfolio et regroupées en chapitres : « les produits d'un enseignement » (par exemple, travaux d'étudiants) ; « matériel personnel » (devoirs, plan de cours, innovations...) et « informations externes » (étudiants, collègues, anciens élèves...). Le portfolio est donc envisagé non comme une nouvelle méthode d'évaluation mais plutôt comme un système de collecte, combinaison, et organisation de l'information à partir d'une large série de sources incluant les approches traditionnelles de l'évaluation de l'enseignement comme les visites de pairs et les évaluations des étudiants. Il était prévu que le portfolio soit utilisé pour les principales décisions de carrière comme la titularisation et la promotion mais aussi dans une perspective formative.

Le concept de portfolio a également été appliqué au niveau local (du département ou de l'UFR). Il a dans ce cas été conçu à l'origine comme une approche pour allouer des ressources aux départements et éviter l'utilisation de mesures simplistes et unitaires de la qualité de l'enseignement (Mullins & Cannon, 1991). Ses composants étaient : une information contextuelle sur le département, des preuves des résultats d'apprentissage à un haut niveau de généralisation, l'évaluation par les étudiants de leur expérience de cours (une idée reprise plus tard dans le CEQ), les preuves des innovations, et la preuve des contributions de la faculté au développement de la qualité d'enseignement.

Les portfolios ont aussi été utilisés pour évaluer l'enseignement au niveau national. Par exemple, en 1993, les universités australiennes ont participé à un processus national d'audit de qualité initié par le gouvernement fédéral en soumettant des portfolios de leur travail. Limités à 20 pages, les portfolios devaient inclure de l'information sur les caractéristiques des étudiants et des enseignants, les dispositions de feedback sur l'enseignement, les liens avec l'industrie, la formation continue, les innovations pédagogiques et le programme d'évaluation internationale. Ces portfolios, complétés par des visites de pairs sur site, ont été utilisés comme preuve pour porter des jugements sur le mérite institutionnel, pour l'allocation de fonds pour soutenir d'autres développements de la qualité (Knapper & Wright, 2001).

⁵ [http :www.iupui.edu/~josotl](http://www.iupui.edu/~josotl)

En dépit de leur acceptation grandissante, les portfolios ne sont pas sans poser des problèmes notamment leur manque de sélectivité et de cohérence, ce qui surcharge la tâche des comités chargés de les évaluer. Certains préconisent d'utiliser une structure admise incluant les buts, la préparation, les méthodes appropriées, les résultats significatifs et une critique réflexive (Glassick, Huber & Maeroff, 1997). Centra (1993) est un des quelques chercheurs qui ont rassemblé des données sur la fiabilité et la validité des portfolios. Il a comparé les décisions de titularisation et de promotion basées sur les portfolios à celles qui étaient prises sans ce support et a trouvé que les jugements étaient raisonnablement fiables (accord inter-juges) quand les juges travaillaient avec des critères spécifiés. Il a conclu qu'utiliser les portfolios pour prendre des décisions sommatives quant à l'enseignement pouvait fournir une image de la performance plus complète et que l'évaluation des portfolios nécessitait de discuter des standards et des critères d'évaluation. Il recommande que les portfolios n'incluent pas seulement ce que les personnes disent mais aussi ce qu'elles font.

Ceci reflète un résultat commun dans de nombreux domaines, à savoir que la fiabilité inter-évaluateurs augmente quand des critères spécifiques sont fournis et étaye la position suivant laquelle, pour juger de l'efficacité de l'enseignement, il faut se mettre d'accord sur ce qu'est l'efficacité de l'enseignement.

3. MODELES ET EFFETS

Derrière les pratiques et les politiques de l'évaluation, se profilent des modèles et des visées plus ou moins explicités. La littérature anglophone met en évidence leur pluralité et leur rapport au contexte, notamment à travers la recherche sur l'évaluation de programme (Younès, 2006). Un parallèle peut être établi entre ce cadre de réflexion et l'évaluation de l'enseignement universitaire.

3.1. Pluralité

Les chercheurs ne privilégient pas tous le même type de connaissance : ainsi, la tradition objectiviste représentée par Campbell et Scriven dans l'évaluation de programme valorise la connaissance causale. Dans ce cas, l'évaluation est externe et une attention particulière est portée au contrôle des biais. A l'opposé, un paradigme initié par Stake (1975, 1980) puis développé par Guba et Lincoln (1981), Patton (1980), Cranton (2001), qui s'enracine dans une épistémologie constructiviste, interprétative et herméneutique, favorise la connaissance utile, pluraliste, la découverte. Cronbach et Rossi cherchent pour leur part à intégrer recherche d'objectivité et pragmatisme, construction de connaissance et utilisation de la connaissance. Rossi définit comme compréhensives des évaluations qui réalisent trois types d'activités : les analyses reliées à la conceptualisation et au modèle d'interventions, le contrôle de l'implantation du programme et l'évaluation de l'utilité du programme à savoir ses effets, ses impacts et le rapport coût/efficacité (Rossi & Freeman, 1985, p. 38). Pour ces auteurs, la recherche en évaluation est plus que l'application de méthodes. « C'est aussi une activité politique et managériale, un input dans la mosaïque complexe dans laquelle les décisions politiques émergent pour la planification, la modélisation, l'implantation et la perpétuation des programmes pour améliorer la condition humaine » (Rossi & Freeman, 1985, p. 27). Tous les théoriciens de l'évaluation de programme étudiés s'accordent sur le fait que l'évaluation ne peut se résumer à une seule perspective. Pour Scriven par exemple, « l'évaluation est une multiplicité de multiples » (1983b, p.257) aussi bien au niveau des critères de mérite que des perspectives, c'est à dire qu'il est nécessaire de disposer de plusieurs perspectives avant que la synthèse ne soit faite. Les différentes perspectives sont aussi bien celles des enseignants concernés, celles des étudiants, des pairs voire celles d'autres acteurs. Car il n'y a pas de modèle unique de l'enseignement. Dans une perspective

académique, les critères sont principalement ceux de la discipline tels que définis par les pairs. Dans le cas d'un modèle pédagogique, les critères sont principalement envisagés du point de vue des processus d'apprentissages. Une approche économique détermine la valeur de l'enseignement du point de vue de son efficacité dans le champ socio-économique en privilégiant les critères de l'insertion des étudiants, de l'efficience et de l'attractivité.

3.2 Vers une évaluation régulation

La question du processus d'évaluation entraîne également une interrogation quant à son impact. Les effets peuvent être positifs, négatifs, nuls ou ambivalents, directs, indirects ou incertains. Ils sont généralement loin d'être clairement identifiables, étant donné que l'évaluation s'inscrit dans un contexte universitaire complexe, nécessitant de prendre en compte les différentes logiques qui s'affrontent. Pour que les effets de l'évaluation correspondent aux effets recherchés, à savoir qu'ils aient un impact positif, plutôt qu'un impact négatif, pervers ou nul, les auteurs soulignent qu'il est important que l'évaluation fasse l'objet d'une appropriation par les acteurs concernés et de décisions internes au système (Rossi & Freeman, 1985, p. 357 ; Gueissaz et al., 1998, p. 179). Ce qui détermine la régulation qui peut s'ensuivre.

L'évaluation compréhensive conceptualisée par Rossi constitue un modèle élargi de l'évaluation englobant aussi bien la conception que l'implantation et l'analyse des effets. Ce type d'évaluation prend en compte le contexte spécifique dans lequel elle est effectuée. L'intérêt d'un modèle pragmatique, utile, contextualisé et négocié de l'évaluation comme régulation, que ce soit dans l'évaluation de programme en général ou plus particulièrement dans l'évaluation de l'enseignement supérieur a été largement souligné dans la littérature. A une problématique statique de l'évaluation comme saisie objective, est substituée une problématique contextualisée, formative, dynamique et régulée. La validité d'un système d'évaluation dépend aussi de sa capacité à entraîner une transformation des pratiques, qui donne lieu à des changements utiles à tous les niveaux du système de formation et des acteurs que sont les enseignants, les étudiants, les administratifs, les décideurs. Cette forme de régulation de l'action implique une interactivité et une médiation entre savoirs et action. C'est parce que l'évaluation est trop peu intégrée dans une régulation qu'elle est trop peu facteur de progrès. Trois paramètres sont ainsi déterminants de la régulation produite : la cohérence entre les objectifs et l'utilisation des résultats de l'évaluation, l'évaluation comme outil de contrôle et l'évaluation comme outil d'amélioration de la situation pédagogique étant généralement perçues comme contradictoires ; la participation des différentes catégories d'acteurs à l'ensemble du processus et la capacité du système à s'ajuster à un contexte spécifique et à évoluer.

CONCLUSION

La recherche anglo-saxonne a largement balisé le champ de l'évaluation depuis les années 60 jusqu'à nos jours, nous permettant un travail de reprise critique. Il apparaît que les questions méthodologiques sont indissociables de celles des modèles et plus largement d'une pensée portant sur le sens même de l'enseignement. Les travaux pionniers de l'évaluation de programme montrent à quel point il importe de relier les différents facteurs en jeu dans l'évaluation. Le primat d'une seule approche quantitative et standardisée s'avère insuffisant pour appréhender et accompagner les dynamiques formatives. En effet, l'évaluation de l'enseignement universitaire requiert un dispositif évolutif situé articulant des données qualitatives et quantitatives avec leur mise en débat au regard de pratiques et de perspectives plurielles. La multidimensionnalité de l'évaluation conduit ainsi à une forme de déplacement.

A l'état des lieux contrôle versus formatif, s'ajoute ce qui correspond à une autre potentialité de l'évaluation, sa dimension de conseil. Dans cette acception, l'accompagnement est central pour mieux connaître l'évaluation dans ses utilités comme dans ses limites, favoriser les liens entre les phénomènes, les situations et les acteurs et accompagner des dynamiques d'évolution et de changement.

Dans la conception anglo-saxonne la qualité est principalement une procédure de contrôle. Elle est fortement basée sur l'explicitation de critères finis ainsi que sur l'explicitation du respect de la procédure et des résultats. Il est patent que cette démarche qui entraîne une définition des exigences, la vérification qu'elles sont atteintes et les actions correctives conduit à élargir cette conception vers ce qui va faciliter des démarches d'accompagnement. L'accompagnement que peut apporter le conseil en évaluation est tout à fait déterminant pour expliciter les objectifs, les critères et les indicateurs retenus et pour accompagner toute la dynamique institutionnelle que cela suppose en termes de délibération, d'ajustement et de changement. Mais il ne peut y avoir une liste finie de critères, l'accompagnement est à concevoir comme une mise en mouvement. Il est fondamental de maintenir le système ouvert.

CHAPITRE

Abrami, P.C. & d'Apollonia, S. (1990). The Dimensionality of ratings and their use in personnel decisions. In M. Theall & J. Franklin (Eds.): *Students ratings of instruction : Issues for improving practice. New Directions for Teaching and Learning*, 43, 97-111.

Aleamoni, L.M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*. 13(2), 153-166.

Alstete, J. W. (1996). *Benchmarking in Higher Education: Adapting Best Practices to Improve Quality*. Washington DC : The George Washington University Graduate School of Education and Human Development.

Alstete, J.W. (2004). *Accreditation matters. Achieving academic recognition and renewal*. ASHE ERIC Higher Education Report, n°4, vol. 30. Jossey-Bass.

Argyris, C. & Schön, D.A. (1974). *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass.

Astin, A.W. (1978). *Four Critical Years: Effects of College on Beliefs, Attitudes and Knowledge*. San Francisco: Jossey-Bass.

Beaty, L. (2001). Teaching Evaluation and Accreditation. In C. Knapper and P. Cranton (Eds.), *Fresh Approaches to the Evaluation of Teaching*, New Directions for Teaching and learning, n° 88, p. 75-86.

Beaty, L. (1997). The Professional Development of Teachers in Higher Education : Structures, Methods, and Responsibilities. *Innovations in education and Training International*, 35, 99-107.

Bedin, V. & Jorro, A. (2007). L'évaluation-conseil en éducation et formation. *Les dossiers des sciences de l'éducation*, 18.

Boyer, E.L. (1990). *Scholarship Reconsidered : Priorities of the Professoriate*. Princeton. N.J. : Carnegie Foundation for the advancement of Teaching.

Brown, J. S. & Duguid, P. (1996). Universities in the Digital Age. *Change, the Magazine of Higher Learning*, 28, 4, 11-19.

- Centra, J. A. (1993). *Reflexive faculty evaluation. Enhancing teaching and determining faculty effectiveness*. San Francisco : Jossey-Bass.
- Cranton, P. (2001). Interpretative and Critical Evaluation, *New Directions for Teaching and Learning*, 88, 87- 97, San Francisco : Jossey-Bass.
- Cross, K.P. (1986). A Proposal to Improve Teaching or What Taking Teaching Seriously Should Mean. *American Association for Higher Education Bulletin*, 39, 1, 9-14.
- Dickey, D. & Pearson C. (2005). Recency Effect in College Student Course Evaluations. *Practical Assessment Research & Evaluation*, 10(6). Disponible en ligne : <http://pareonline.net/getvn.asp?v=10&n=6>
- Doyle, K.O. Jr. (1983). *Evaluating teaching*. Toronto : Lexington Books, D.C. Heath and Company.
- Fenwick, T.J. & Parsons, J. (2000). *The Art of Evaluation: A Handbook for Educators and Trainers*. Toronto: Thompson Educational Publishing.
- Garlick, S. & Pryor, G. (2004). *Benchmarking the University : Learning about improvement*. Report for the department of education, science and training, Commonwealth of Australia.
- Geis, G. L. (1991). The moment of truth: Feeding back information about teaching. *New Directions for Teaching and Learning*, 48, 7-19.
- Gill, W.E. (2006). Outcomes Assessment in the Accreditation Process. Paper presented at the *Annual Meeting of the American educational research association*. San Francisco, CA.
- Glassick, C.E., Huber, M.T. & Maeroff, G. I. (1997). *Scholarship Assessed : Evaluation of the Professoriate*. San Francisco : Jossey-Bass.
- Guba, E. G. & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass.
- Gueissaz, A., Häyrynen-Alesto, M., Fischer-Bluhm K., & Snell K. (1998). Les acteurs de l'évaluation, la décision d'évaluer. Dans P. Dubois (Ed.), *EVALUE, (Évaluation et auto-évaluation des Universités en Europe)*. Commission Européenne (DG XII, TSER).Rapport final.
- Hand, T. & Trembath, K. (1998). *Enhancing and Customising the Analysis of the Course Experience Questionnaire*. Canberra : Department of Employment, Training and Youth Affairs.
- Harvey, L. (2004). The Power of Accreditation: Views of Academics. in P. Di Nauta ; P-L Omar , A. Schade & J. P. Scheele (eds.) *Accreditation Models in Higher Education. Experiences and Perspectives*, ENQA Workshop Reports 3, Helsinki. pp 5-18.
- Hoyt, D. P. (1973). Measurement of instructional effectiveness. *Research in Higher Education*, 1, 367- 378.
- Hoyt, D.P., Chen, Y., Pallett, W.H., & Gross, A.B. (1999). *Revising the IDEA System for Obtaining Student Ratings of Instructors and Courses*. Technical Report 11: The IDEA Center, Manhattan: Kansas State University.
- Hutchings, P. (Ed.). (1996). *Making Teaching Community Property: A Menu for Peer Collaboration and Peer Review*. Washington : American Association for Higher Education.
- Jackson, N. & Lund, H. (2000). *Benchmarking in Higher Education*. Buckingham, England : SHRE and Open University Press. McKinnon, K.R., Walker, S. H. & Davis, D. (2000).

Benchmarking, a manual for Australian Universities. Department of Education, Training and Youth Affairs (DETYA). Disponible sur : <http://www.dest.gov.au/archive/highered/otherpub/bench.pdf>.

Johnson, T. (1999). *Course Experience Questionnaire, 1998*. Parkville, Victoria : Graduate Carrers Concil of Australia.

Kirschner, D. & Witson, J. (1997). *Situated Cognition: Social, Semiotic, and Psychological Perspectives*, Mahway, N.J.: Erlbaum.

Knapper, C.K. (1995). The Origins of Teaching Portfolios. *Journal on Excellence in College Teaching*, 6, 1, 45-56.

Knapper, C. & Wright, W. A. (2001). Using Portfolios to Document Good Teaching. In C. Knapper & P. Cranton (Eds.), *Fresh Approaches to the Evaluation of Teaching*, New Directions for Teaching and learning, n° 88, pp. 19-29.

Lenze, L. F. (1997). Small group instructional diagnosis (SGID). In K. T. Brinko & R. j. Menges (Eds.). *Practically speaking: A sourcebook for instructional consultants in higher education*, (pp.143–146), Stillwater, OK: New Forums Press.

Marsh, H.W. (1982). SEEQ: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology*, 52, 77-95.

Marsh, H.W. (1984). Students' Evaluations of University Teaching : Dimensionality, Reliability, Validity, Potential Biases, and Utility. *Journal of Educational Psychology*, 76, 5, 707-754.

Marsh, H.-W. (1987). Students' evaluation of university teaching : Research findings, methodological issues and directions for future research. *International Journal of Educational Research*, 11, 3, 253-288.

Marsh, H.-W. & Roche, L. (1993). The use of students evaluation and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, 30, 1, 217-251.

McKinnon, K.R., Walker, S. H. & Davis, D. (2000). *Benchmarking, a manual for Australian Universities*. Department of Education, Training and Youth Affairs (DETYA). Disponible sur : <http://www.dest.gov.au/archive/highered/otherpub/bench.pdf>.

Morrison, D.E. (1997). Overview of Instructional Consultation in North America. In K. Brinko and R. Menges, (Eds.), *Practically Speaking: A Sourcebook for Instructional Consultants in Higher Education*, Stillwater, Oklahoma : New Forums Press.

Mullins, G.P. & Cannon, R.A. (1991). A Teaching Portfolio for Departments. *Research and Development in Higher Education*, 13, 378-383.

Murray, F.B. (2001) From consensus standards to evidence for claims: assessment and accreditation in the case of teacher education' in Ratcliff, J.L, Lubinescu, E.S & Gaffney, M.A (Eds.), 2001, *New Directions for Higher Education, How Accreditation Influences Assessment*, n°113, April 2001, New York, Jossey-Bass.

Murray, H. G., Jelley, R. B., & Renaud, R. D. (1996, April). Longitudinal Trends in Student Instructional Ratings: Does Evaluation of Teaching Lead to Improvement of Teaching? Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

Patton, M.Q. (1980). *Qualitative evaluation methods*. Beverly Hills, Ca : Sage.

QAA Quality Assessment Review Reports, (2000). Disponible sur HTTP : <http://www.qaa.ac.uk>.

Ramsden, P. & Entwistle, N.J. (1981). Effects of Academics Departments on Students' Approaches to Studying. *British Journal of Educational Psychology*, 51, 368-383.

Ramsden, P. A (1991). Performance Indicator of Teaching Quality in Higher Education : The Course Experience Questionnaire. *Studies in Higher Education*, 16, 2, 129-150.

Ramsden, P. & Entwistle, N.J. (1981). Effects of Academics Departments on Students' Approaches to Studying. *British Journal of Educational Psychology*, 51, 368-383.

Rossi, P. H. & Freeman, H.E. (1985). *Evaluation : A Systematic approach* (3ème ed.). Beverly-Hills CA : Sage.

Scheele, J.P. (2004). Licence to Kill: About Accreditation Issues and James Bond. in P. Di Nauta ; P-L Omar , A. Schade & J. P. Scheele (eds.) *Accreditation Models in Higher Education. Experiences and Perspectives*, ENQA Workshop Reports 3, Helsinki. pp 19-25.

Scriven, M. (1983a). Summative Teacher Evaluation. In J. Milman (ed.), *Handbook of Teacher Evaluation*. Thousand Oaks, Calif. Sage.

Scriven, M. (1983b). Evaluation Ideologies. In G.F. Madaus, M. Scriven, D.L. Stufflebeam (Eds.), *Evaluation Models : Viewpoints on educational and human services evaluation* (pp. 229-260). Boston : Kluwer-Nijhoff.

Shore, B.M. (1975). Moving beyond the Course Evaluation Questionnaire in Evaluating University Teaching. *CAUT Bulletin*, 23, 4, 7-10.

Shore, B.M. and Associates (1986). *The Teaching Dossier : A guide to its Preparation and use*. Ottawa : Canadian Association of University Teachers.

Smith, R. (2001). Formative Evaluation and the Scholarship of Teaching and Learning. In C. Knapper and P. Cranton (Eds.), *Fresh Approaches to the Evaluation of Teaching*, New Directions for Teaching and learning, n° 88, pp 51- 62.

Stake, R. E. (1975). To evaluate an art program. In R.E. Stake (Ed.), *Evaluating the arts in education : A responsive approach* , Columbus, OH: Merrill, pp. 13-31.

Stake, R. E. (1980). Program evaluation, particularly responsive evaluation. In W. Dockrell & D. Hamilton (Eds.) *Rethinking educational research*, London: Hodder and Stoughton, pp. 72-87.

Sursock, A. (2000). Towards accreditation schemes for higher education in Europe? paper at CRE workshop, 8–9 Nov. 2000, Vienna.

Trout, Paul (2000). Flunking the Test: The Dismal Record of Student Evaluations. *Academe*, 86(4), 58-61.

Wilhelm, W. B. & Comegys C. (2004). Course selection decisions by students on campuses with and without published teaching evaluations. *Practical Assessment, Research & Evaluation*, 9(16). Disponible sur : <http://PAREonline.net/getvn.asp?v=9&n=16>

Yao, Weissinger, Grady (2003). Faculty use of student evaluation feedback. *Practical Assessment, Research & Evaluation*, 8(21).

Disponible sur : <http://PAREonline.net/getvn.asp?v=8&n=21>.

Younès, N. (2006). *L'effet de l'évaluation de l'enseignement supérieur par les étudiants*. Thèse de doctorat. Université Pierre Mendès France, Grenoble.