pedocs
Open Access Erziehungswissenschaften

DIPF

Laveault, Dany

Le jugement professionnel: foyer de tensions et de synergies nouvelles en évaluation scolaire

Schweizerische Zeitschrift für Bildungswissenschaften 30 (2008) 3, S. 483-500

urn:nbn:de:0111-opus-42303

in Kooperation mit / in cooperation with:

Schweizerische Zeitschrift für Bildungswissenschaften
Rivista svizzera di scienze dell'educazione
Revue suisse des sciences de l'éducation

http://www.rsse.ch/index.html

Nutzungsbedingungen / conditions of use

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

We grant a non-exclusive, non-transferable, individual and limited right to using this document.
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.
By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

pedocs
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Mitglied der Leibniz-Gemeinschaft
Informationszentrum (IZ) Bildung
Schloßstr. 29, D-60486 Frankfurt am Main
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Le jugement professionnel: foyer de tensions et de synergies nouvelles en évaluation scolaire

Dany Laveault

Le jugement professionnel d'évaluation intervient dans les prises de décision et il se révèle particulièrement utile lorsque les mesures de contrôle externe du risque d'erreur se sont avérées inappropriées ou insuffisantes. Il est appelé à soutenir la qualité de l'évaluation dans les trois grandes catégories d'actes professionnels: les actes logiques, psychologiques et moraux. Particulièrement en évaluation certificative, il fait l'objet de nombreuses tensions entre dispositifs de contrôle externe qui visent à encadrer ce jugement et facteurs de contrôle interne qui favorisent la réflexion critique. Le développement de nouvelles synergies apparaît comme une condition essentielle au dépassement de ces tensions. À ce sujet, il est possible d'envisager que les critères utilisés pour attester de la validité de la recherche, quantitative ou qualitative, puissent également s'appliquer à l'évaluation de la crédibilité du jugement d'évaluation.

L'évaluation scolaire et particulièrement l'évaluation sommative impliquent que l'enseignant prenne une série de décisions: du choix de ce qui sera évalué aux meilleurs moyens pour le faire, jusqu'à la décision finale pour déterminer si les résultats sont conformes aux attentes et satisfont aux exigences d'un programme de formation. Chacune de ces décisions peut être plus ou moins adéquate ce qui fait que l'entreprise même de l'évaluation scolaire demeure une *approximation risquée* (Cardinet, 1990) qui fait appel au jugement professionnel de l'enseignant.

Pour réduire le risque associé à l'évaluation scolaire, plusieurs dispositifs de contrôle peuvent être envisagés. Facione, Facione et Giancarlo (1997) en identifient cinq qu'il est possible de généraliser à l'évaluation scolaire:

1. Automatiser ce qui peut l'être. Le recours à des examens standardisés et à des questions à choix multiples correspond à une telle disposition.
2. Élaborer des directives et des protocoles qui réduisent des processus complexes à une série d'étapes à suivre.
3. Contrôler ou limiter l'accès à la prise de décision à travers différents types d'obligations.

ISSN 1424-3946, Academic Press Fribourg

4. Limiter le pouvoir discrétionnaire des individus (révision des décisions par une autorité centrale, pouvoir de signature, etc.).
5. Former les personnes à réfléchir.

Ce modèle décrit une progression hiérarchique dans le contrôle du risque qui se termine par une boucle d'autorégulation où la capacité de l'enseignant à réfléchir sur sa propre pratique constitue la dernière mesure de contrôle. Lorsqu'il n'est plus possible d'ajouter des contrôles supplémentaires ou lorsque ceux-ci s'avèrent limités, le jugement professionnel constitue le dernier recours en matière d'évaluation scolaire. Il réside dans la capacité et la volonté de l'enseignant à penser le problème en dehors de son cadre habituel.

L'intérêt pour le jugement professionnel n'est pas nouveau, mais les réformes en éducation et particulièrement celles concernant l'évaluation ont fait en sorte qu'il prenne une importance accrue. Ces changements et ces innovations ont donné lieu à une série de tensions, qu'il est possible de regrouper en deux catégories:

1. *Tensions liées au caractère de plus en plus complexe de l'évaluation scolaire*. L'approche par compétences s'est traduite par l'évaluation de la performance des élèves au moyen de tâches complexes. Ces pratiques sont loin de donner lieu à des interprétations aussi *automatiques* que les examens traditionnels. L'enseignant est amené à conjuguer ses pratiques anciennes avec les pratiques nouvelles et à confronter ses propres jugements avec ceux de ses pairs. Ce phénomène de collaboration s'étend à toutes les personnes intéressées par les résultats: l'élève au premier titre, mais aussi ses parents et les personnes responsables de l'efficacité et du bon fonctionnement du système scolaire. Il a pris une importance accrue avec l'introduction d'évaluations externes dont les résultats sont rendus publics.
2. *Tensions entre contrôle externe et contrôle interne de l'évaluation scolaire*. Le contrôle externe du risque dans la prise de décision privilégie une évaluation fondée sur l'adhésion à des référentiels communs ou à des protocoles préétablis. L'exercice du jugement professionnel, par ailleurs, nécessite une certaine marge de manœuvre. Une telle tension est bien résumée par la question suivante du Conseil supérieur de l'éducation du Québec (2001): «Quand et comment un enseignant peut-il exercer son jugement professionnel en matière d'évaluation s'il doit constamment se référer au régime pédagogique et à la politique d'évaluation des apprentissages en la matière?» (p. 7).

Dans ce nouveau contexte, l'exercice du jugement professionnel en évaluation scolaire est appelé à s'étendre, à s'adapter et à évoluer. Le but de cet article est de définir d'abord en quoi consiste le jugement professionnel en évaluation scolaire et quels en sont les différentes catégories et principaux indicateurs. Ensuite, l'article cherchera à documenter les tensions précédentes et identifiera quelles nouvelles synergies sont devenues nécessaires pour que le jugement professionnel puisse s'exercer en limitant ces tensions et éventuellement en les dépassant.

Les catégories de jugement

Le modèle de Facione et al. (1997) décrit des dispositifs de contrôle laissant une marge de manœuvre plus ou moins grande à l'exercice du jugement. On peut se demander s'il y a une opposition entre l'exercice du jugement professionnel et les mesures de contrôle visant à réduire les risques d'erreur dans la prise de décision. Les modèles de Tripp (1993) et de Coles (2002) n'autorisent pas une telle interprétation, car tous deux considèrent que le jugement comporte plusieurs facettes.

Le modèle de Tripp (1993) définit quatre catégories de jugement allant du simple au complexe: jugement pratique, jugement diagnostique, jugement réflexif et jugement critique. Au premier niveau, les décisions se prennent sans réfléchir, sur le champ, de manière automatique. Les deux autres niveaux font intervenir des choix ou encore des stratégies plus complexes nécessitant un minimum de réflexion (jugement diagnostique et jugement critique). Enfin, le dernier niveau définit une réflexion plus poussée, particulièrement dans les situations nouvelles, imprévues ou indéterminées (jugement critique).

Coles (2002) a repris les quatre catégories de Tripp (1993) et leur a associé quatre domaines d'exercice du jugement. Ceux-ci se présentent comme suit:

1. Le jugement *intuitif* a pour domaine les questions urgentes qui exigent une réponse immédiate, presque réflexe. Le professionnel cherche alors à répondre à la question: «Qu'est-ce que je fais immédiatement?».
2. Le jugement *stratégique* a également un caractère d'urgence, mais il porte sur un plus grand nombre d'éventualités. Le problème fait appel à des compétences procédurales et exige de prendre en considération et de choisir parmi plusieurs solutions pratiques. La question qui correspond le mieux à ce type de jugement est: «Qu'est-ce que je pourrais bien faire maintenant?».
3. Le jugement *réflexif* cible les situations de pratiques professionnelles empreintes d'incertitude. Il est personnel et se réalise lorsque le professionnel prend en considération ses propres habiletés, pondère les multiples facettes d'une situation complexe telle qu'il la comprend au moyen de sa pratique établie et choisit d'aller de l'avant avec la solution qu'il a choisie. Ce faisant, il cherche à répondre à la question: «Qu'est-ce qu'il me serait possible de faire maintenant?».
4. Le jugement *délibératif* porte sur les dilemmes impliquant un conflit entre valeurs morales et idéaux. Les enjeux moraux émergent de préoccupations où le bien de l'élève est au centre des préoccupations et nécessite que soient remises en question certaines connaissances ou croyances prises pour certaines. La question que se pose le professionnel est de l'ordre non plus de la compétence, mais de la déontologie: «Qu'est-ce que je devrais faire maintenant?».

Les modèles de Tripp (1993) et de Coles (2002) permettent d'établir un certain nombre de caractéristiques essentielles du jugement exercé par un professionnel. D'abord, il importe de distinguer entre le «jugement d'un professionnel» qui

s'exerce de plusieurs façons et dans des domaines d'application variés et le «jugement professionnel» à proprement parler qui se caractérise par une forme d'autocritique et de réflexivité. Pour Tripp (1993), ces différents jugements s'exercent probablement les uns à la suite des autres. Par exemple, pour réfléchir à une décision d'évaluation et éventuellement la reconsidérer, l'enseignant doit d'abord être en mesure d'utiliser son savoir d'expérience et son savoir théorique pour décrire et comprendre les divers aspects de la situation (jugement diagnostique). Le jugement réflexif fait suite à un jugement diagnostique, de même qu'un jugement critique intervient après un jugement réflexif.

Pour Coles (2002), les jugements intuitifs et stratégiques peuvent difficilement être qualifiés de professionnels puisqu'ils ne mettent pas en jeu la nature discrétionnaire des décisions du professionnel. De plus, de tels jugements sont plus susceptibles de se produire dans des situations où le problème à résoudre est bien défini et lorsqu'il existe un protocole préétabli. Si la présence de règles ou encore le caractère automatique de certaines décisions ne favorisent pas l'exercice du jugement professionnel, c'est principalement parce que la réflexion sur l'utilisation de ces règles est suspendue. Toutefois, rien n'empêche le professionnel de réfléchir sur ces automatismes et de les remettre en question. «We need to question the notion of objectivity and look to improve our deliberative subjectivity» (Coles, 2002, p. 9).

Paramètres du jugement professionnel

Selon Fenstermacher et Richardson (2000), la qualité de l'enseignement repose sur trois grandes catégories d'actes professionnels:

1. les actes *logiques*: démontrer, expliquer, corriger, interpréter;
2. les actes *psychologiques*: motiver, encourager, récompenser, planifier et évaluer;
3. les actes *moraux*: honnêteté, courage, tolérance, compassion, respect, équité.

Le jugement professionnel est appelé à soutenir la qualité de l'évaluation scolaire dans ces trois grandes catégories d'actes professionnels. Au niveau des actes logiques, il repose sur l'emploi rigoureux d'un savoir théorique et d'expérience essentiel à la correction et à l'interprétation des résultats. Il fait aussi partie des actes psychologiques, chaque fois que l'enseignant évalue et utilise les résultats de cette évaluation pour motiver l'élève et supporter son apprentissage. Enfin, il fait partie des actes moraux qui nécessitent des arbitrages parfois difficiles entre les différentes obligations de l'enseignant.

Au niveau des actes logiques, les modèles de Tripp (1993) et de Coles (2002) associent le jugement professionnel aux capacités de réflexion et d'autocritique. Ces capacités sont intimement liées aux processus métacognitifs de l'enseignant. Le déclenchement de processus métacognitifs à la base de cette réflexion peut se faire en réponse à une situation lorsque celle-ci est en grande partie indétermi-

née. La situation rencontrée peut être nouvelle, imprévue, ambigüe ou faire appel à des demandes difficilement compatibles auxquelles ni le savoir d'expérience, ni le savoir théorique de l'enseignant, ni les procédures habituellement prévues ne peuvent apporter facilement de solution. De telles situations requièrent une attention accrue. Par ailleurs, le déclenchement des processus métacognitifs peut être le résultat de dispositions particulières du sujet. Facione et al. (1997) énoncent sept habitudes intellectuelles qui prédisposent au jugement professionnel: la recherche de la vérité, l'ouverture d'esprit; la capacité d'analyse; la rigueur; la confiance en la pensée critique; la curiosité, la maturité du jugement. Sans la présence de telles dispositions intellectuelles, il est fort probable que le jugement professionnel ne se développera, ni ne s'exercera à l'initiative du sujet.

Au niveau des actes psychologiques, le système de croyances et de valeurs d'un enseignant est tout aussi important que ses compétences en matière d'actes logiques. Laveault (2005) observe que le domaine des croyances permet d'intégrer tout l'aspect plus ou moins conscient du jugement de l'enseignant. Par exemple, plusieurs recherches démontrent que des attentes et des standards élevés imprègnent l'environnement des écoles efficaces (Groupe d'études sur les écoles efficaces 2001, juin). Dans de telles écoles, les enseignants ne se laissent pas facilement abattre et croient en leur capacité d'intervenir favorablement auprès de l'élève (Bandura, 1997). Le système des croyances des enseignants joue un rôle crucial dans la perception qu'ils se font d'eux-mêmes et de la situation. Le jugement devient professionnel lorsqu'il y a réflexion et remise en question de ces croyances, ce que Coles (2002) a décrit comme la «reconstruction critique de la pratique».

De la même façon que le système des croyances, le rapport à la discipline et au savoir peuvent constituer une source de biais. Dunn, Parry et Morgan (2002) font ressortir que les enseignants de différentes disciplines ne privilégient pas le même type de capacités ni ne formulent les niveaux de rendement ou d'exigence de la même manière. Par exemple, en physique, les enseignants tendent à choisir des tâches qui donnent lieu à des réponses correctes ou incorrectes et laissant peu de place à de l'interprétation. En histoire, les enseignants ont tendance à avoir recours à des tâches qui nécessitent des niveaux élevés d'inférence comme par exemple dans les projets, les dissertations et les rapports de recherche. Parce que qu'ils s'inscrivent dans des usages propres à chaque discipline, les jugements d'évaluation décrits par Dunn, Parry et Morgan (2002) appartiennent aux catégories de jugement intuitif et stratégique. Ils ne peuvent être qualifiés de «professionnels» lorsqu'ils relèvent d'habitudes ou de pratiques préétablies qui ne font pas l'objet de remise en question et de réflexion approfondie.

Au niveau des actes moraux, le jugement professionnel d'évaluation exige que l'enseignant arbitre différentes obligations dans l'intérêt de l'élève. Hadji (1997) identifie plusieurs de ces obligations, complétées par Laveault (2005). Comment conjuguer devoir d'honnêteté et devoir de bienveillance, le droit à la vérité de la personne évaluée et le droit à préserver son estime de soi dans des situations

d'échec? Comment être loyal envers tous les élèves et bienveillant de façon équitable, lorsque l'on porte davantage attention et que l'on vient en aide plus souvent à l'élève qui en a le plus besoin? Ces conflits d'obligations entraînent des dilemmes moraux dont la solution repose sur un ensemble de compétences, mais surtout sur une conscience déontologique, une réflexion éthique et une ouverture à la remise en question personnelle de ses croyances. À la disposition intellectuelle de Facione et al. (1997), il serait possible d'ajouter une disposition éthique qui la complète, car comme le dit Hadji (1997): «On entre dans l'éthique dès que l'on commence à considérer que tout ne va pas de soi et qu'il y a lieu de s'interroger sur le sens et la valeur de ce que l'on fait» (p. 25).

Tensions reliées au caractère complexe de l'évaluation sommative

Parmi l'ensemble des situations d'évaluation scolaire, l'évaluation sommative à des fins de certification est sans doute celle qui fait intervenir le plus souvent les niveaux réflexifs et critiques du jugement professionnel de l'enseignant. Il y a pour cela plusieurs raisons:

- Les enjeux de l'évaluation, qu'elle soit interne ou externe, sont élevés et intéressent plusieurs personnes à des titres divers: élèves, parents, enseignants, institution scolaire.
- Les buts poursuivis sont multiples et pas toujours compatibles. L'évaluation sommative est utilisée pour faire le bilan des apprentissages des élèves (orientation, certification, promotion), communiquer avec les parents ou entre collègues, mais aussi – à tort ou à raison – pour évaluer le rendement des enseignants et des institutions.
- L'évaluation certificative place l'enseignant dans un conflit de loyautés (Cardinet, 1986): loyauté envers l'élève à laquelle les pratiques d'évaluation formative l'ont habitué; mais aussi loyauté envers la société dans l'exercice de sa fonction de certification des apprentissages. L'enseignant cumule ainsi deux fonctions: celle de *coach* et celle de *juge* (Harlen, 2005).

Face à cette problématique particulière, deux attitudes semblent prévaloir. La première consiste à lever le voile sur le jugement de l'enseignant, à faire pénétrer l'élève dans la «boîte noire» de l'évaluation en favorisant son implication dans le processus (Speck, 1998). L'autre consiste à réduire les tensions entre les fonctions certificative et formative de l'évaluation scolaire et à développer une synergie entre les deux (Harlen, 2005).

Selon Speck (1998), l'exercice du jugement professionnel devrait être plus transparent et impliquer l'élève davantage. Il considère que la notation peut avoir une fonction formative lorsqu'elle permet à l'élève de mieux comprendre les exigences du travail scolaire et comment il peut les satisfaire. Selon Speck (1998), le

rôle de l'enseignant ne consiste pas seulement à rendre des jugements professionnels mais aussi à aider les élèves à comprendre autant que possible comment ces jugements sont rendus. C'est une chose d'informer les élèves des barèmes de notation ou encore des conditions de réalisation d'une tâche, c'en est une autre de lui faire prendre conscience des exigences de qualité qui font que son travail est bien fait. L'implication accrue de l'élève devrait faire en sorte qu'il se pose les bonnes questions et parvienne ainsi à développer son aptitude à l'autoévaluation.

Une plus grande participation de l'élève ne répond pas uniquement à des exigences de transparence et de justice car il ne serait pas juste de pénaliser un élève pour ne pas avoir atteint des niveaux de rendement qui ne lui ont jamais été expliqués. Elle joue également un rôle important dans la construction même du savoir de l'élève. Ce *savoir évaluer* (*assessment as learning*) transmis de l'enseignant à l'élève, constitue un ajout important aux fonctions certificative (*assessment of learning*) et formative (*assessment for learning*) de l'évaluation scolaire (Stiggins, Arter, Chappuis & Chappuis, 2004). Il ne vise pas à rendre l'élève apte à se noter, mais plutôt à exercer un jugement autonome et un meilleur contrôle sur ses apprentissages. Il existe plusieurs méthodes pour y parvenir dont les différentes formes d'autoévaluation (Laveault, 2004), le dossier progressif d'apprentissage, le recours à des copies types et les conférences avec l'élève (Stiggins et al., 2004). Plusieurs recherches ont démontré l'efficacité de ces méthodes et l'importance de mieux faire comprendre à l'élève les critères d'évaluation de son rendement en l'impliquant davantage (Church, 1997; Laveault & Miles, 2008; Orsmond, Merry & Reiling, 2002).

Pour Harlen (2005), il n'y a pas lieu de continuer à traiter séparément évaluation formative et évaluation sommative, car les deux peuvent être utilisées pour supporter l'apprentissage de l'élève. Il existe une relation de réciprocité entre les deux fonctions que l'on aurait avantage à mieux exploiter. D'une part, l'évaluation formative, hautement critériée et contextualisée, possède un caractère volatile et peu fiable lorsque l'on suit de près les hésitations qui se produisent dans l'apprentissage continu de l'élève. D'autre part, le bilan des apprentissages offert par l'évaluation sommative, quoique plus stable, requiert une information riche comme celle recueillie dans le cadre de l'évaluation formative pour interpréter les résultats de l'élève et leur donner tout leur sens.

Si le but ultime de l'évaluation scolaire est d'appuyer l'élève dans ses apprentissages, alors le jugement professionnel doit pouvoir s'exercer en tenant compte de toutes les informations à la disposition de l'enseignant. Ce qui est préconisé par Harlen ne consiste pas à intégrer les évaluations formatives de l'élève dans la notation de l'élève, mais dans l'interprétation et la réflexion qui en découle. D'ailleurs, lorsque l'enseignant doit trancher dans des cas limites, l'étude de Lafortune et Allal (2008) a démontré que ceux-ci font intervenir de nouveaux critères ou des informations supplémentaires pour mieux informer leur processus de prise de décision. De telles pratiques sont acceptables lorsqu'elles se font

dans l'intérêt de l'élève et sont conformes aux standards de l'évaluation scolaire émis conjointement par l'AERA/APA/NCME (2002):

- Standard 11.20: «toute information collatérale qui peut mener à une interprétation différente de la performance de l'examiné à un test devrait être prise en considération» (p. 117).
- Standard 13.7: «Tout renseignement pertinent devrait être pris en compte s'il est susceptible d'améliorer la validité globale d'une décision» (p. 146).

La prise en compte de renseignements obtenus en marge de l'évaluation certificative soulève de nombreuses questions. Le risque d'erreur dans la prise de décision n'est pas aboli. Il est simplement déplacé des données (sont-elles fidèles, valides?) à l'interprétation des données et à la prise de décision. Or, si le risque d'erreur en évaluation formative a peu d'impact à long terme, il a des conséquences beaucoup plus graves lorsque l'évaluation sert à prendre une décision concernant l'obtention de crédits en vue de l'obtention d'un diplôme.

En mettant l'accent sur l'acquisition de compétences, les réformes scolaires ont contribué à poser en des termes nouveaux toute la question du jugement professionnel pour l'évaluation certificative. L'élaboration de situations d'apprentissage et de tâches complexes associe étroitement l'évaluation et l'acquisition de compétences en continu de sorte que le jugement est appelé à s'exercer dans de nouvelles conditions où les démarcations entre évaluation et apprentissage, entre évaluation formative et évaluation certificative, tendent à s'estomper. Dans de telles situations complexes, l'interprétation de l'enseignant doit pouvoir s'appuyer sur un référentiel commun qui lui permette de situer l'élève sur une échelle de progression.

Plusieurs tentatives de développer de tels référentiels ont eu lieu récemment et se poursuivent, qu'il s'agisse des échelles de niveaux de rendement (Ministère de l'éducation de l'Ontario [MÉO], 2000), des échelles de niveaux de compétence (Ministère de l'éducation du Québec [MÉQ], 2002) ou de socles de compétences (Inspection générale de l'éducation nationale [IGEN], 2007; Rey, Carette, Defrance & Kahn, 2003). Il ne s'agit pas d'évaluer l'élève au moyen d'un cumul d'indicateurs disparates, mais de situer l'élève sur un «parcours de développement» (Tardif, 2006) commun à tous les élèves. Voici par exemple comment le Ministère de l'éducation des loisirs et du sport du Québec (MÉLS, 2006) conçoit l'articulation du jugement de l'enseignant avec l'utilisation de telles échelles:

> *Les échelles fournissent des références qui guident le jugement global que l'enseignant doit porter sur les compétences des élèves à la fin du cycle. La responsabilité première de ce jugement revient à l'enseignant qui a suivi et évalué les apprentissages de l'élève, mais il serait souhaitable qu'il y ait concertation pour que les enseignants d'une même discipline se donnent une représentation claire et commune des niveaux de chaque échelle (on pourrait notamment illustrer les niveaux à l'aide de quelques exemples recueillis en classe afin de mettre en évidence ce qui les caractérise).* (p. 76)

Les recherches sur l'évaluation scolaire ont démontré l'importance des échanges maître-élèves afin de rendre les élèves aptes à interpréter et à comprendre leurs résultats et ainsi à mieux soutenir leur progression au moyen de cibles réalistes de rendement. Il en va de même pour l'enseignant. C'est à travers différentes formes d'interactions sociales et d'échanges professionnels entre collègues qu'il peut le mieux entreprendre une démarche réflexive et critique, similaire à celle de l'autoévaluation de l'élève, afin de vérifier son interprétation des résultats et valider ses décisions. De telles interactions correspondent à la notion d' *évaluation négociée* définie par Cardinet (1990/1987):

> *L'évaluation a nécessairement, dans ce cas, des référentiels multiples, parce qu'elle prend comme cadre la réunion des référentiels individuels des acteurs. La réalité est définie alors comme la somme des faits objectifs* […] *et des représentations subjectives* […]. *La meilleure façon d'explorer cette réalité, c'est de confronter tous les points de vue possibles et pour cela de susciter des interactions entre partenaires.* (p. 150)

Les référentiels communs et les échanges professionnels sont autant d'outils permettant à l'enseignant d'ajouter une boucle de régulation supplémentaire, tout comme le jugement de l'enseignant constitue une boucle de régulation qui doit permettre à l'élève de valider sa propre évaluation et de se fixer des cibles à atteindre dans son parcours de développement. Ces régulations du jugement professionnel, à la fois internes et externes, entre enseignants et entre enseignants et élèves, permettent de s'assurer non seulement que la meilleure décision est prise, mais aussi qu'elle est bien comprise par tous les intéressés. Ce n'est que dans ces conditions que le jugement professionnel devient jugement délibératif et que l'évaluation négociée permet le développement de représentations partagées et d'un référentiel commun.

Tensions liées au niveau de contrôle du jugement professionnel

Si les référentiels communs constituent une contrainte à l'exercice du jugement professionnel de l'enseignant, ils ne sont pas nécessairement préjudiciables dans la mesure où ils sont de nature à appuyer les jugements pratiques et diagnostiques du professionnel. De plus, s'ils influencent l'exercice du jugement professionnel, ils ne l'empêchent pas car ils ne font qu'ajouter un niveau de contrôle et de complexité supplémentaire dans la prise de décision. Ils représentent une autre forme d'interaction entre l'enseignant et sa collectivité professionnelle. Voici à titre d'exemple, comment le programme-cadre québécois de l'évaluation au niveau primaire (MEQ, 2002) justifie l'utilisation d'échelles de compétences communes:

> *La formulation générale des échelles appelle les personnes qui interviennent auprès de l'élève tout au long du cycle à poser, au moment du bilan de fin de cycle, un*

> *jugement professionnel reposant sur des observations pertinentes.* […] *les enseignants auront la même représentation de ce qu'on doit attendre d'un élève au niveau 3 en lecture. Ainsi, le bilan des apprentissages établi à partir des échelles permettra aux enseignants qui reçoivent les élèves de poursuivre le travail amorcé au cycle précédent. Les échelles contribuent donc à la continuité entre les cycles et à un meilleur ajustement des actions pédagogiques aux besoins des élèves* (p. 34).

Il existe une complémentarité nécessaire entre contrôle interne et contrôle externe du jugement professionnel. Un exemple tiré de la politique d'évaluation de l'Ontario (MÉO, 2000) servira à illustrer la complémentarité entre contrôle interne et autonomie du jugement professionnel d'une part, et contrôle externe sous forme de directives d'autre part. À titre d'exemple, prenons les cas hypothétiques de deux élèves pour lesquels cinq résultats numériques en pourcentages ont été calculés et à partir desquels il faut trouver un résultat final:

Élève A: 74, 70, 80, 75, 81 Moyenne = 76
Élève B: 62, 82, 94, 87, 55 Moyenne = 76

L'application automatique de la moyenne nous porterait à attribuer le même score global à ces deux élèves. Dans le cas de l'élève A, la moyenne est représentative de l'ensemble des résultats. Dans le cas de l'élève B, ce n'est pas le cas: les résultats de cet élève varient de façon extrême et aucun ne s'approche à plus de six points de la moyenne. Dans un tel cas, l'enseignant est appelé à se demander si la moyenne de l'élève B est un aussi bon indicateur de rendement que la moyenne de l'élève A. Une réflexion plus poussée le conduirait à rechercher l'explication de telles variations et à en estimer les conséquences avant de prendre une décision définitive.

Voici un exemple de même nature inspiré d'une évaluation de type qualitatif au moyen de l'échelle descriptive de niveaux de rendement de l'Ontario, qui comporte quatre niveaux, le niveau trois représentant la norme provinciale ou niveau «satisfaisant» (MÉO, 2000).

- Élève C: 2 3 3 4 3 3
- Élève D: 1 2 2 3 3 4
- Élève E: 4 3 3 2 2 1

Le curriculum de l'Ontario de la 9e à la 12e année (2000) formule la directive suivante: «[…] la note devrait faire état du niveau de rendement le plus fréquent pour la durée du cours, bien qu'il faille accorder une attention particulière aux manifestations de rendement les plus récentes» (p. 19). Dans le cas de l'élève C, poser un jugement global sur son niveau de rendement ne présente pas de difficulté particulière, car le niveau trois est à la fois le plus représentatif et le plus récent. Par contre, la décision résultant de l'application de ces directives est beaucoup moins évidente dans le cas des élèves D et E. Dans les deux cas, les

performances les plus récentes ne sont pas représentatives de l'ensemble. De plus, les résultats sont disparates, comme dans l'exemple sur les moyennes. Tant l'élève D que l'élève E ont obtenu des résultats inférieurs à la norme provinciale dans la moitié des cas. Cependant, l'élève D démontre une progression continue de ses résultats et termine l'année avec une excellente marque. Par contre, l'élève E termine l'année avec un rendement insuffisant, même si sa performance en début d'année était tout à fait satisfaisante.

Dans les cas des élèves A et C, un jugement pratique ou intuitif ne présente pas de problèmes, comme dans la majorité des cas où les résultats sont relativement homogènes. Par contre, les cas des élèves B, D et E nécessitent un jugement plus poussé. Seul un jugement diagnostique pourrait approfondir les raisons des écarts entre les résultats et un jugement réflexif et critique pourrait mieux argumenter la prise de décision. Plusieurs facteurs doivent être pris en considération:

- Les derniers résultats peuvent-ils être expliqués par des difficultés passagères de l'élève, les limites de l'instrument de mesure ou des outils d'observation?
- Est-ce que la réussite aux dernières épreuves porte sur des apprentissages nécessaires à la continuité de la progression de l'élève?
- Quelles conséquences aura la décision sur l'apprentissage et les chances de réussite future de l'élève?

Ces trois cas démontrent que méthodes quantitative et qualitative donnent lieu à des problématiques similaires dans le choix d'un score représentatif et que dans un cas comme dans l'autre, il n'est pas possible de faire l'économie du jugement professionnel. Ils indiquent que l'objectivité dans la prise de décision est tout à fait relative, même encadrée par des directives strictes ou encore avec des résultats quantitatifs. Dans les trois cas précédents, la décision peut difficilement être prise sans recourir à des renseignements complémentaires.

Enjeux du jugement professionnel: le développement de nouvelles synergies

Les travaux récents en matière d'évaluation et de jugement professionnel font ressortir le caractère formateur de l'évaluation, tant formative que certificative et ce, tant pour l'élève que pour l'enseignant. En multipliant les sources de *feedback*, en faisant appel au jugement réflexif et critique tant de l'enseignant que de l'élève, l'évaluation contribue au développement professionnel et à une meilleure communication entre l'enseignant et l'élève grâce à l'établissement d'un dialogue constructif. Nous assistons alors à la création d'un *cercle vertueux* (Laveault, 2005): le jugement professionnel contribue à améliorer les pratiques d'évaluation et ces pratiques permettent en retour une meilleure réflexion sur les résultats, sur le suivi de l'élève et le support à lui apporter.

L'évaluation scolaire ne pourra devenir plus juste, tant au sens de justesse que de justice, que si certaines tensions ne sont pas dépassées pour donner lieu à de nouvelles synergies. Voici quelques synergies que l'évaluation scolaire doit encore développer:

1. *Synergie entre évaluation de l'enseignant et celle de l'élève*. Comment l'enseignant peut-il partager son savoir d'évaluateur pour mieux communiquer les exigences et les attentes des programmes de formation? Comment en retour mettre à profit le savoir d'évaluateur de l'élève pour négocier des cibles de réussite, développer un plan d'enseignement individualisé et soutenir ses apprentissages? La capacité de l'enseignant à aider l'élève sera d'autant facilitée que l'élève sera en mesure de formuler clairement ses difficultés et l'aide dont il a besoin.
2. *Synergie entre l'évaluation de l'enseignant et celle de ses pairs*. La communication entre enseignants est primordiale pour que s'instaure une interprétation commune des attentes et des exigences. L'évaluation scolaire est à la fois une responsabilité individuelle et collective. La prise de décision ne devrait pas dépendre de l'interprétation qu'un seul enseignant se fait des résultats de l'élève et de sa progression par rapport aux objectifs de formation, particulièrement pour les cas difficiles aux enjeux élevés. Il est dans l'intérêt, tant des élèves que de la profession, de faire en sorte que le jugement professionnel puisse s'appuyer sur un certain nombre de références communes qui balisent les marges d'interprétation.
3. *Synergie entre l'objectif et le subjectif*. Cardinet (1990) exprime bien l'ambivalence des enseignants entre «mesure objective et quantitative» et «évaluation globale, subjective et qualitative»: «L'évaluation scolaire est prise en tenaille entre le désir de précision, qui lui ôte toute signification sociale, et un souci d'ouverture sur la vie, qui lui ôte toute valeur métrique» (p. 6). La mesure ne peut remplacer l'exercice du jugement professionnel du personnel enseignant et le jugement fait appel à une interprétation des faits, que ceux-ci soient mesurables ou non. D'ailleurs, selon Weiss (1994), «les résistances au changement dans le domaine de l'évaluation […] semblent bien résulter de la difficulté de la plupart à se départir de l'idée que l'évaluation juste et équitable ne peut que résulter de mesures quantitatives […]. Pourtant des recherches ont contesté la validité de ces notes et ont montré au contraire la validité des appréciations qualitatives des enseignants» (p. 67).
4. *Synergie entre le qualitatif et le quantitatif*. Comment les aspects qualitatifs et quantitatifs de l'évaluation peuvent-ils se compléter? Une première possibilité consiste à dépasser l'opposition entre formatif et certificatif qui attribue le monopole de l'information qualitative et riche en contexte au formatif et celui des données quantitatives à une démarche sommative liée à la certification des acquis. Or, l'évaluation de performances complexes, qu'elle se fasse dans un but formatif ou certificatif, requiert que soit pris en compte les deux types d'information. Dans l'apprentissage précoce de la lecture, il est bien

établi qu'une vitesse de décodage minimale est nécessaire au développement de la compréhension. Ne prendre en considération que les données quantitatives portant sur la vitesse de décodage sans prendre en compte celles, autant qualitatives que quantitatives, portant sur la compréhension du texte, handicaperait les prises de décision, tant au niveau formatif qu'au niveau certificatif. L'inverse serait également vrai.

5. *Synergie entre évaluation interne et externe*. Comment l'évaluation externe peut-elle informer l'enseignant sur la validité de ses exigences et des résultats de ses élèves et comment, en retour, les résultats de l'évaluation interne peuvent-ils compléter les données de l'évaluation externe pour remédier aux lacunes dans les apprentissages des élèves? Que faire, enfin, lorsque les évaluations ne concordent pas et comment expliquer les différences pour éventuellement les dépasser?

L'actualisation de ces nouvelles synergies est de plus en plus nécessaire pour que le jugement professionnel d'évaluation joue véritablement son rôle de soutien à l'élève, en étant plus juste, plus transparent et plus rigoureux. Pour accroître la qualité du jugement professionnel, plusieurs approches ont été envisagées au niveau de la formation et du perfectionnement des enseignants. La méthode des incidents critiques développée par Tripp (1993) permet aux enseignants d'approfondir leur réflexion critique. Le travail en *équipe-cycle* (Lafortune, 2004) ou en communauté d'apprentissage (Fullan, 2001) permet d'unir les compétences de plusieurs enseignants dans la prise de décision.

Mais comment un enseignant ou un groupe d'enseignants peut-il attester de la qualité des décisions ou à tout le moins estimer sa marge d'interprétation ou son degré de certitude? Le jugement professionnel serait-il tautologique? Cardinet (1990) affirme que la capacité à établir les marges d'erreur de leurs estimations fait la force des modèles quantitatifs. Cette capacité d'autocritique est de plus en plus accessible par l'intermédiaire des modèles qualitatifs même si ceux-ci ne possèdent pas la même précision. Ce qui importe, dans les deux modèles, c'est que les deux se prêtent à une sorte de contrôle: «Sans être jamais sûrs de dire vrai, nous pouvons contrôler notre probabilité de nous tromper» (Cardinet, 1990, p. 10).

Alors que de nombreux développements se sont produits tant du côté des méthodes de recherche qualitative que du côté des méthodes mixtes (Johnson, Onwuegbuzie & Turner 2007; Teddlie & Tashakkori, 2006), il est paradoxal de constater à quel point ces méthodologies n'ont pas fait l'objet de transferts dans le domaine de l'évaluation scolaire. La rigueur scientifique n'est certainement pas le monopole de la recherche quantitative. Il existe des protocoles de recherche bien établis, des méthodes d'échantillonnage et de contrôle des variables qui sont propres à la recherche qualitative. Or, le jugement du chercheur est primordial en recherche qualitative et si une grande part de subjectivité est admise, l'invocation du jugement du chercheur, soit-il un chercheur de terrain, ne suffit pas à

garantir la crédibilité de la recherche. Pourquoi en serait-il de même pour le jugement d'évaluation?

Dans les faits, plusieurs pratiques d'évaluation scolaire ont recours à des méthodes de recherche qualitative, sans cependant les nommer comme telles. En voici quelques exemples:

1. Lorsque l'on affirme que le jugement de l'enseignant doit se faire à partir d'une variété de méthodes d'évaluation, nous mettons de l'avant une forme de triangulation des méthodes. Lorsque l'on affirme que le jugement de l'enseignant doit se faire en consultation, nous mettons de l'avant une forme de triangulation de sources et ainsi de suite (Huberman & Miles, 1991).
2. Les méthodes d'échantillonnage intentionnel (Fraenkel & Wallen, 2006) sont couramment utilisées en évaluation scolaire au lieu des méthodes d'échantillonnage représentatif davantage associées aux examens. Par exemple, on demandera à l'élève de produire un *échantillon typique* de ses travaux pour une évaluation à des fins de bilan de son portfolio, ou alors un *échantillon hétérogène (maximum variation)* dans le cadre d'une évaluation formative pour illustrer ce qu'il peut faire de mieux et ce qu'il éprouve le plus de difficulté à accomplir.

Si la réflexion est au cœur de la démarche autocritique du jugement professionnel, elle s'exercera avec d'autant plus de crédibilité qu'elle répond à un ensemble de critères de rigueur scientifique. La crédibilité de la recherche qualitative comporte quatre dimensions (Fraenkel & Wallen, 2006; Karsenti, 2000) qu'il est possible de transférer à la crédibilité du jugement professionnel en substituant simplement «enseignant» là où il est question de chercheur:

1. *Fidélité interne*: congruence entre le sens dégagé par l'enseignant et le sens compris et véhiculé par les examinés.
2. *Fidélité externe*: généralisabilité des observations et des interprétations permettant à plusieurs enseignants de porter le même jugement.
3. *Validité interne*: concordance entre enseignants sur la nature des observations.
4. *Validité externe*: degré d'assurance que les mêmes résultats seront retrouvés dans des conditions similaires.

Les indicateurs précédents n'ont pas la précision de mesures quantitatives et n'ont pas toujours la possibilité de s'exprimer numériquement. Parfois, une appréciation quantitative est possible, comme dans le cas de la validité externe qui pourrait être évaluée au moyen d'une technique d'enquête Delphi (Rowe & Wright, 2001). Cette technique permettrait de consulter des enseignants considérés comme des experts et, par une méthode d'établissement de consensus, les conduire à exprimer leur degré d'assurance sur la possibilité que les résultats soient reproductibles dans des conditions similaires. De telles méthodes faisant appel au jugement d'experts en évaluation scolaire ne sont pas nouvelles: que l'on

pense aux méthodes Angoff, Nedelsky ou Ebel pour l'établissement de seuils de réussite (Laveault & Grégoire, 2002) ou encore à la méthode plus récente du *signet* (Buckendahl, Smith, Impara & Plake, 2002). Ces méthodes introduisent rigueur et contrôle dans le processus du jugement d'évaluation scolaire. Leur développement doit se poursuivre et s'étendre au plus grand nombre de sphères possibles de l'évaluation. Enfin, la conceptualisation récente des méthodes mixtes de recherche présente un grand potentiel d'application à l'évaluation scolaire car elles permettent d'entrevoir la fin de l'opposition stérile entre méthodes quantitatives et qualitatives.

La caractéristique la plus souvent invoquée du jugement professionnel est la nature autocritique de la réflexion. Elle implique un *jugement sur le jugement*. Or, pour que ce jugement soit davantage qu'une simple disposition intellectuelle ou éthique, mais qu'il s'exerce en tant que capacité réelle d'autocontrôle, il doit pouvoir se faire avec toute la rigueur scientifique possible. Les méthodes quantitative et qualitative sont des leviers importants non seulement d'analyse des résultats des élèves, mais aussi d'étude et d'analyse des jugements d'évaluation. Ces jugements sur les jugements permettent d'ajouter un autre niveau de contrôle au modèle de Facione et al. (1997). En effet, le contrôle interne et l'autonomie du jugement professionnel risquent de bien peu s'actualiser s'ils ne s'accompagnent de véritables *méthodologies* d'autocontrôle.

Références

AERA/APA/NCME. (2002). *Standards for educational and psychological testing*. Washington: Auteur.

Bandura, A. (1997). *Self-efficacy. The exercise of control*. N.Y.: W. H. Freeman and Co.

Buckendahl, C. W., Smith, R. W., Impara, J. C. & Plake, B. S. (2002). A comparison of Angoff and Bookmark standard setting methods. *Journal of Educational Measurement, 39* (3), 253-263.

Cardinet, J. (1986). *Évaluation scolaire et pratique*. Bruxelles: De Boeck-Wesmael.

Cardinet, J. (1990/1987). *Évaluation externe, interne, ou négociée?* [Repris dans Institut romand de documentation pédagogique (1990). Hommage à Jean Cardinet. Cousset: Delval]. Neuchâtel: IRDP, 1987, 17 p.

Cardinet, J. (1990). *Remettre le quantitatif à sa place en évaluation scolaire*. Neuchâtel: Institut romand de documentation pédagogique (IRDP).

Church, A. H. (1997). Do you see what I see? An exploration of congruence in ratings from multiple perspectives. *Journal of Applied Social Psychology, 27* (11), 983-1020.

Coles, C. (2002). Developing professional judgment. *The Journal of Continuing Education in the Health Professions, 22* , 3-10.

Conseil supérieur de l'éducation. (mars 2001). *Projet de politique d'évaluation des apprentissages. Commentaires*. Québec: Auteur.

Dunn, L., Parry, S. & Morgan, C. (2002). *Seeking quality in criterion referenced assessment.* Paper presented at the Learning Communities and Assessment Cultures Conference organized by the EARLI Special Interest Group on Assessment and Evaluation, University of Northumbria, Newcastle.

Facione, P. A., Facione, N. C. & Giancarlo, C. A. F. (1997). *Professional judgment and the disposition toward critical thinking*. Millbrae, CA: The California Academic Press.

Fenstermacher, G. D. & Richardson, V. (2000). *On making determinations of quality in teaching.* National Academy of Sciences: Board on International Comparative Studies in Education.

Fraenkel, J. R. & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th ed.). Montréal: McGraw-Hill.

Fullan, M. (2001). *The new meaning of educational change*. New York, NY: Teachers College Press.

Groupe d'études sur les écoles efficaces. (2001, juin). *Bien faire, mieux faire!* Rapport soumis au Ministre de l'Éducation de l'Ontario, Toronto.

Hadji, C. (1997). Pour une éthique de l'agir évaluationnel. *Mesure et évaluation en éducation, 20* (2), 7-26.

Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal, 16* (2), 207-223.

Huberman, M. & Miles, M. B. (1991). *Analyse des données qualitatives*. Bruxelles: De Boeck Université.

Inspection générale de l'éducation nationale (IGEN). (2007). *Les livrets de compétences: nouveaux outils pour l'évaluation des acquis* (Rapport n° 2007-048). Paris: IGEN.

Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1* (2), 112-133.

Karsenti, T. & Savoie-Zajc, L. (2000). *Introduction à la recherche en éducation*. Sherbrooke: Éditions du CRP.

Lafortune, L. (2004). *Travailler en équipe-cycle. Entre collègues d'une école*. Québec: Presses de l'Université du Québec.

Lafortune, L. & Allal, L. (Éd). (2008). *Jugement professionnel en évaluation. Pratiques enseignantes à Québec et à Genève.* Québec: Presses de l'Université du Québec.

Laveault, D. & Grégoire, J. (2002). *Introduction aux théories des tests en éducation et en psychologie* (2e éd.). Bruxelles: DeBoeck-Université.

Laveault, D. (2004). Interactions entre formation et évaluation: de la régulation entre acteurs et leurs rôles à la régulation de l'apprentissage. *Mesure et évaluation en éducation, 27* (1), 51-67.

Laveault, D. (2005). Le jugement professionnel de l'enseignant: quel en est impact sur l'acte d'évaluer? *Mesure et évaluation en éducation, 28* (2), 93-114.

Laveault, D. & Miles, C. (2008). Utilité des échelles descriptives et différences individuelles dans l'autoévaluation de l'écrit. *Mesure et évaluation en éducation, 31* (1), 1-29.

Ministère de l'éducation de l'Ontario. (2000). *Le curriculum de l'Ontario de la 9e à la 12e année. Planification des programmes et évaluation*. Toronto: Auteur.

Ministère de l'éducation des loisirs et des sports. (2006). *L'évaluation des apprentissages au secondaire. Cadre de référence* (version préliminaire). Québec: Gouvernement du Québec.

Ministère de l'éducation du Québec. (2002). *L'évaluation des apprentissages au préscolaire et au primaire. Cadre de référence.* Québec: Auteur.

Ministère de l'Éducation du Québec. (2003). *Politique d'évaluation des apprentissages*. Québec: Auteur.

Orsmond, P., Merry, S. & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 27* (4), 309-323.

Rey, B., Carette, V., Defrance, A. & Kahn, S. (2003). *Les compétences à l'école: apprentissage et évaluation*. Bruxelles: DeBoeck.

Rowe, G. & Wright, G. (2001). Expert opinions in forecasting. Role of the Delphi technique. In J. Scott Armstrong (Éd.), *Principles of forecasting: A handbook of researchers and practitioners* (pp. 125-144). Boston: Kluwer Academic Publishers.

Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. St-Laurent: ERPI.

Speck, B. W. (1998). Unveiling some of the mystery of professional judgment in classroom assessment. *New Directions forTeaching andLearning, 74,* 17-31.
Stiggins, R. J., Arter, J. A., Chappuis, J. & Chappuis, S. (2004). *Classroom assessment for student learning. Doing it right – using it well.* Portland, OR: Assessement Training Institute.
Tardif, J. (2006). *L'évaluation des compétences. Documenter le parcours de développement.* Montréal: Chenelière Éducation.
Teddlie, C. & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools, 13* (1), 12-28.
Tripp, D. (1993). *Critical incident in teaching.* London: Routledge.
Weiss, J. (1994). Évaluer autrement. *Mesure et Évaluation en Éducation, 17* (1), 63-74.

Mots clés: jugement professionnel; méthodologies de l'évaluation scolaire; évaluation certificative; prise de décision

Professionelle Schülerbeurteilung: Spannungsherd und neue Synergien in der schulischen Beurteilung

Zusammenfassung

Das professionelle Urteil in der Schule spielt eine Rolle in der Entscheidungsfindung und erweist sich als besonders nützlich, wenn die Maßnahmen der externen Kontrolle zur Verringerung des Fehlerrisikos nicht angemessen waren oder nicht ausreichen. Es ist angesagt, die Qualität der Beurteilung in den drei grossen Kategorien professioneller Handlungen zu unterstützen: die logischen, psychologischen und moralischen Handlungen. Besonders in der zertifizierenden Beurteilung, ist die Schülerbeurteilung Gegenstand zahlreicher Spannungen zwischen der externen Kontrolle, den Abnehmenden, mit diesem Urteil umgehen zu müssen und den Faktoren der internen Kontrolle, die die kritische Reflexion begünstigen. Die Entwicklung neuer Synergien erscheint als eine wesentliche Voraussetzung für die Überwindung dieser Spannungen. In diesem Zusammenhang ist es denkbar, dass Gütekriterien der qualitativen und quantitativen Forschung auch auf die Schülerbeurteilung übertragen und angewendet werden können.

Schlagworte: Lehrerurteil, Schulleistungsbeurteilung, Methoden der schulischen Beurteilung, zertifizierende Beurteilung, Entscheidungsfindung

Il giudizio professionale: luogo di tensioni e di sinergie nell'ambito della valutazione scolastica

Riassunto

Il giudizio professionale di valutazione interviene nel processo decisionale e si rivela particolarmente utile quando le misure di controllo esterno del rischio di er-

rore si sono rivelate inappropriate o insufficienti. Esso è chiamato a sostenere la qualità della valutazione nelle tre grandi categorie di atti professionali: gli atti logici, psicologici e morali. In particolare per quel che riguarda la valutazione certificativa, il giudizio professionale è spesso oggetto di tensioni tra i dispositivi di controllo esterno che cercano di inquadrarlo e altri fattori di controllo interno che favoriscono la riflessione critica. Lo sviluppo di nuove sinergie appare come una condizione essenziale al superamento di tensioni del genere. A questo proposito diventa possibile accettare che i criteri utilizzati per attestare la validità della ricerca, quantitativa o qualitativa, possano anche applicarsi alla valutazione della credibilità del giudizio di valutazione.

Parole chiave: iudizio professionale, metodologie di valutazione scolastica, valutazione certificativa, presa di decisione

Professional judgment: focus of new tensions and synergies in classroom assessment

Abstract

Professional judgment on assessment occurs when decisions must be made. It is especially useful when external controlling measures with regards to wrong decisions have been either inappropriate or lacking. It is intended to uphold assessment quality in three categories of professional action: logical, psychological and moral. Mostly when evaluation is used for certification, professional judgment is subject to numerous tensions between external control procedures which tend to structure this judgment and factors of internal control which make the best use of the professional's critical reflexivity. The development of new synergies appears as an essential condition to overcome such tensions. One such synergy would consist in applying research valid criteria, whether they are associated with quantitative or qualitative data, to the assessment of evaluation judgment credibility.

Key words: professional judgment, methodologies of classroom assessment, certification, decision making