

Vers un cadre pour l'évaluation de la compétence d'enseignant

Erik Roelofs

Chercheur principal au Centre d'information et de recherche psychométrique du Cito, Institut national néerlandais de recherches et d'élaboration de tests à Arnhem

Piet Sanders

Directeur du Centre d'information et de recherche psychométrique du Cito, Institut national néerlandais de recherches et d'élaboration de tests à Arnhem

RÉSUMÉ

L'élaboration d'instruments permettant d'évaluer la compétence de l'enseignant nécessite un modèle de performance compétente qui peut guider à la fois la collecte et l'évaluation des preuves dans des situations fonctionnelles. Selon Kane (1992), la validation des déclarations sur la compétence des enseignants est synonyme d'évaluation de l'argumentation interprétative.

Partant des idées que l'on se fait aujourd'hui de l'enseignement et de l'apprentissage, l'article décrit un modèle interprétatif de la performance compétente qui, loin d'être prescriptif par nature, ouvre la voie à diverses formes de performance professionnelle responsable. Les conséquences de la performance professionnelle pour les étudiants, les classes ou l'organisation constituent la base du modèle. Elles permettent de déduire les interventions acceptables et les processus de prise de décision sous-jacents, ainsi que les parties associées d'une base de connaissances professionnelles. Les conséquences de ces perceptions pour le développement de domaines de compétence et la collecte de preuves sont passées en revue.

Mots clés

Competence assessment, teaching personnel, teaching quality, quality of education, performance appraisal, personnel assessment

Évaluation des compétences, personnel didactique, qualité de l'apprentissage, qualité de l'enseignement, évaluation de la performance, évaluation du personnel

1. Introduction

L'évaluation de la compétence des enseignants connaît, au niveau international, un intérêt croissant, né de la demande d'assurance de qualité et d'une reconnaissance accrue de la profession enseignante (Verloop, 1999). Les États-Unis ont derrière eux une longue tradition d'évaluation des enseignants qui se reflète tant dans le volume d'articles et ouvrages publiés sur la recherche que dans les instruments mis au point. Aux États-Unis, l'attention dirigée vers l'évaluation des enseignants a été essentiellement stimulée par le principe de la comptabilité envers le contribuable.

Plusieurs instruments permettant d'évaluer les enseignants à divers stades de leur carrière professionnelle dans le cadre de la sélection, de la certification et du développement professionnel ont ainsi été élaborés (Dwyer, 1998). De nombreuses informations contenues dans le présent article sont extraites de cette documentation fournie. Cela ne signifie pas pour autant que les autres pays ne disposent pas d'une tradition d'évaluation des enseignants. Pelkmans (1998), par exemple, décrit des pratiques d'évaluation des enseignants en Angleterre, au Pays de Galles, en Allemagne, en Australie et aux Pays-Bas. Ces pays sont toutefois moins expérimentés que les États-Unis.

Aux Pays-Bas, comme dans d'autres pays (Pelkmans, 1998), une plus grande attention est accordée à l'évaluation des enseignants compte tenu de l'élargissement du champ libre laissé aux établissements scolaires pour élaborer des politiques, ce qui permet, entre autres, de différencier les fonctions et les salaires (Verloop, 1999; Straetmans et Sanders, 2001). L'accentuation croissante de la formation basée sur les compétences augmente également la demande d'évaluation des compétences d'enseignant. Une loi adoptée aux Pays-Bas stipule, en outre, que les professionnels de l'enseignement doivent satisfaire à des exigences de compétence. La fondation néerlandaise pour la compétence enseignante professionnelle (SBL, 2003) a formulé des exigences jugées cruciales pour les enseignants débutants dans sept domaines de compétence. Ceux-ci sont détaillés sous le point 3.2.1.1.

C'est aux collèges de formation des enseignants et aux organisations scolaires qu'il incombe de veiller à la satisfaction de ces exigences. Enfin, le transfert d'autres professions vers l'enseignement a également été facilité, afin de lutter contre la menace de pénurie d'enseignants qui sévit aux Pays-Bas (Klarus, Schuler et Ter Wee, 2000; Tillema, 2001).

Les développements décrits ci-dessus requièrent une approche cohérente de l'évaluation de la compétence des enseignants. Cet article présente quelques éléments de base d'un cadre d'évaluation de la compétence des enseignants. Nous commençons par élaborer un modèle interprétatif pour l'évaluation de la performance compétente (des enseignants), qui repose sur des conceptions théoriques différentes du bon enseignement. Nous poursuivons en examinant divers instruments pour l'évaluation de la compétence d'enseignant, en utilisant les exigences qui découlent de la notion de validité conceptuelle telle que Messick (1996) l'expose. Enfin, nous abordons quelques points d'étude future.

2. Un modèle interprétatif de la compétence d'enseignant

Il n'existe aucune définition généralement acceptée du concept de compétence. Dernièrement, plusieurs auteurs (par exemple, Bos, 1998; Mulder, 2001; Van Merriënboer, Van der Klink et Jansen, 2002) ont passé la

littérature y afférente en revue et ont élaboré des définitions complètes. Une première distinction importante peut être établie, en anglais, entre *competence* (compétence) et *competency* (compétences, aptitudes). Selon Mulder, *competence* est un concept global qui désigne l'ensemble des aptitudes ou capacités présentées par des personnes ou organisations, tandis que *competency*, plus spécifique, constitue une partie de la compétence. *Competency* (au pluriel *competencies*) est un concept plus étroit et plus atomistique, qui est utilisé pour désigner des aptitudes particulières (cf. également McConnell, 2001). S'inspirant de l'étude de centaines de définitions de la compétence (par exemple, Bunk, 1994; Spencer et Spencer, 1993; Parry, 1996), Mulder (2001) a déduit une définition qui regroupe la pensée des principaux auteurs: «la compétence est la capacité d'une personne ou d'une organisation à atteindre des niveaux particuliers de performance» (p. 76). Citant divers auteurs, il ajoute que les capacités des personnes individuelles se composent de:

- maîtrises d'actions intégrées,
- constituées de groupes de structures de connaissances,
- d'aptitudes cognitives, interactives, émotionnelles et, si nécessaire, psychomotrices,
- ainsi que d'attitudes et de valeurs qui sont nécessaires pour:
- exécuter des tâches,
- résoudre des problèmes,
- et, d'une manière plus générale, la capacité d'évoluer dans :
 - un emploi,
 - une organisation,
 - une fonction,
 - un rôle particuliers.

Il faut savoir, lorsque l'on mesure les dimensions de la compétence, que ces dernières ne s'observent pas directement, mais se manifestent dans la performance dans une situation spécifique (Spencer et Spencer, 1993). La compétence peut, en outre, être développée à un niveau particulier, par exemple, débutant, expérimenté et expert. Mulder (2001) souligne que la compétence peut être présente chez des individus (compétence personnelle) et auprès de systèmes (compétence de système ou d'équipe). Enfin, les aspects de la compétence se transfèrent, dans une certaine mesure, d'une situation à une autre (Thijssen, 1998, 2001).

On peut se poser diverses questions lorsque l'on mesure la compétence de personnes. D'où proviennent les affirmations sur la compétence? Quelles sont les hypothèses et notions théoriques qui sont à la base des mesures de la compétence? Il est important, pour répondre à des questions de ce type, d'utiliser un modèle d'interprétation suffisamment descriptif et explicatif (Shepard, 1993).

Il n'existe pas de cadre scientifique pertinent et généralement accepté pour déterminer ce qui constitue l'enseignement compétent et à partir duquel des déductions peuvent être tirées pour évaluer la compétence d'enseignant (Haertel, 1991). Il existe, en revanche, divers cadres dont

le contenu dépend largement de la vision sous-jacente de la performance professionnelle (Dwyer, 1994, 1998) et des approches théoriques adoptées (cf. Reynolds, 1992).

Les théoriciens qui mettent au point des instruments d'évaluation de l'enseignant évoluent généralement vers une vision partagée de l'enseignement compétent qui est obtenue à la suite d'une interaction entre les théoriciens et les représentants de la profession. La vision résultant de cette interaction peut varier largement: il peut s'agir d'un hybride de tous les types de vision de l'enseignement, mais aussi d'une vision relativement spécifique, par exemple, d'un enseignement «orienté vers le programme», ou d'un enseignement «orienté vers le développement» dans le cadre de l'éducation pendant la prime enfance ou de l'enseignement préscolaire. La première approche est logique et permet de formuler des cadres de performance compétente à appliquer à de grands groupes d'enseignants, par exemple, des exigences de compétences nationales. La seconde approche convient mieux aux organisations qui suivent une mission spécifique.

En dehors de la vision de l'enseignement, l'angle théorique sur la performance professionnelle détermine également la forme que prend un modèle interprétatif. Différents éléments de la compétence de l'enseignant ont été soulignés à travers l'histoire de l'évaluation des enseignants dans la littérature y afférente. L'examen de celle-ci permet de distinguer divers avis sur ce que constitue un bon enseignant et un bon enseignement (Creemers, 1991; Verloop, 1999):

- (a) différencier les traits de personnalité requis pour devenir un enseignant qui réussit (Getzels et Jackson, 1963; Creemers (1991);
- (b) décrire les éléments de connaissance qui impliquent le contenu d'une matière et la façon dont les enseignants raisonnent dans une certaine discipline (Bruner, 1963; Tom et Valli, 1990);
- (c) décrire des formes de comportement enseignant qui contribuent à la performance de l'enseignement (Brophy et Good, 1986; Simon et Boyer, 1974);
- (d) décrire les processus cognitifs et décisionnels des enseignants (Kagan, 1990; Verloop, 1988);
- (e) décrire la connaissance pratique que les enseignants appliquent à des situations spécifiques dans lesquelles ils se trouvent (leur classe ou leur domaine thématique) et la façon dont ils élaborent des théories sur ces situations (Beijaard et Verloop, 1996).

Des techniques d'évaluation spécifiques ont été utilisées pour chacune de ces conceptions du bon enseignement. Afin d'évaluer les traits de personnalité, des questionnaires et des tests psychologiques permettant d'identifier certaines caractéristiques souhaitables ou indésirables ont été utilisés. Dans cette conception de l'enseignement, l'accent a été mis moins sur le bon enseignement lui-même, que sur les caractéristiques de bonne citoyenneté. Une conception toujours dominante est le savoir des bons enseignants. Ce savoir impliquait autrefois la connaissance de faits et d'éléments bien distincts; par la suite, l'accent a été mis sur la structure d'une

discipline (telle que les mathématiques ou la physique) et, plus tard encore, sur la façon dont les professionnels agissent et réfléchissent dans une certaine discipline. Le savoir concerne aussi la connaissance pédagogique, par exemple, sur les méthodes permettant de dispenser l'instruction, d'élaborer les programmes et de regrouper les étudiants. Ces éléments d'érudition sont de plus en plus tirés de la recherche pédagogique (Bellon, Bellon et Blank, 1990). Une méthode fréquemment utilisée pour évaluer le savoir consiste à passer des tests de connaissances normalisés (par exemple, Latham, Gitomer et Ziomek, 1999).

Réagissant à la fois à l'accent mis sur les caractéristiques stables de l'enseignant et à l'accent unilatéral placé sur le savoir, l'enseignement est également perçu comme la représentation d'un comportement effectif. Cette conception tient compte de ce que l'enseignant montre à la classe. Divers instruments d'observation ont été développés afin de se concentrer sur des unités (à petite échelle) de comportement qui sont jugées en rapport avec des résultats d'enseignement fructueux (Stodolsky, 1990). Cette approche accorde toutefois peu d'attention à ce qui se passe dans l'esprit des enseignants: ce qu'ils pensent ou décident et pourquoi ils décident de cette façon. Différents instruments d'évaluation ont été utilisés afin de découvrir la façon de penser de l'enseignant, notamment la réflexion à haute voix sur des protocoles tout en résolvant un problème pédagogique et des entretiens en rappel stimulé. Lors de ces derniers, les enseignants observent leur performance enregistrée sur vidéo et répondent à des questions visant à déterminer ce qu'ils pensaient à un moment bien précis. Le concept de l'enseignement, en tant qu'exposition d'une riche base de connaissances pratiques, implique des méthodes d'évaluation qui se concentrent sur des situations spécifiques auxquelles les enseignants sont confrontés. Les instruments utilisés ne diffèrent pas, pour l'essentiel, de ceux qui sont utilisés afin de révéler les processus de réflexion. L'accent est toutefois mis davantage sur le cadre de travail spécifique (par exemple, le thème spécifique dans une classe déterminée) dans lequel les enseignants exercent leurs activités (Meijer, Verloop et Beijjaard, 1999).

Suite à la récente révolution qui s'est opérée au niveau de la réflexion sur l'apprentissage, qui est reprise sous le terme «nouvel apprentissage», il existe une conception supplémentaire de l'enseignement, à savoir: promouvoir de solides activités d'apprentissage auprès des apprenants. Dans cette conception, le bon enseignement suppose que les enseignants ne doivent pas tant faire preuve d'un répertoire bien défini de «bons» comportements, mais doivent plutôt montrer qu'ils contribuent à l'apprentissage réussi de leurs élèves (Simons, 1999; Vermunt et Verschaffel, 2000).

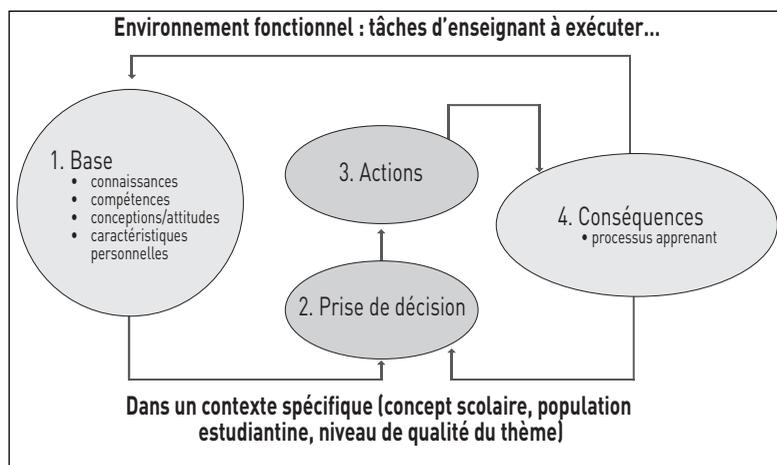
Si toutes les conceptions distinctes de l'enseignement couvrent certains aspects de la compétence d'enseignant, aucune d'elles ne décrit ou n'explique, en revanche, entièrement le sens de l'enseignement compétent. Un examen de la définition des compétences caractéristiques, telle qu'elle est donnée ci-dessus, révèle le besoin de disposer d'un concept détaillé unifié de la compétence d'enseignant qui tient compte de tous les éléments dif-

férents de la compétence d'enseignant, c'est-à-dire les traits de personnalité, le savoir, le comportement, la réflexion, la prise de décision dans une situation spécifique et les activités apprenantes qui en résultent. Roelofs et Sanders (2003) ont développé un modèle de performance compétente pour évaluer la compétence d'enseignant qui rend justice aux aspects de la performance compétente décrits ci-dessus. Le modèle suit la définition générale des compétences décrite par Mulder (2001).

Le point de départ de ce modèle, qui est illustré sous le graphique 1, est que la compétence des enseignants se reflète dans les conséquences des actes des enseignants, les plus importantes étant les activités apprenantes des étudiants. D'autres exemples de conséquences sont le climat (harmonieux ou perturbateur) qui règne dans la classe, le sentiment de bien-être chez les étudiants et les bonnes relations avec les parents et les collègues. Partant des conséquences, les autres éléments du modèle peuvent être établis en progressant à reculons. Premièrement, le volet des «actions» se réfère aux activités professionnelles, par exemple, la dispense d'instructions, la rétrocession d'informations aux étudiants et l'instauration d'un climat de classe coopératif. Deuxièmement, toute activité enseignante se déroule dans un contexte spécifique dans lequel l'enseignant doit prendre de nombreuses décisions à long terme (projection dans l'avenir) ou immédiates dans une situation de classe (cf. Doyle, 1983). À titre d'exemple, les enseignants devront planifier leur instruction et l'adapter en fonction des circonstances (par exemple, différents styles d'apprentissage étudiant et différentes conditions d'organisation). Troisièmement, lorsqu'ils prennent des décisions et exécutent des activités, les enseignants devront partir d'une base de connaissance professionnelle et de certaines caractéristiques personnelles.

L'évaluation de différents domaines de compétence d'enseignant, par exemple, l'instruction et la prise en charge d'une classe, implique que des

Graphique 1: modèle interprétatif de la performance compétente (basé sur Roelofs et Sanders, 2003)



déductions soient faites concernant les enseignants (cf. Kane, 1992). En combinant différents aspects de l'enseignement en un modèle global de performance compétente, les chances de faire des déductions valables sont meilleures qu'en utilisant des modèles réductionnistes qui se concentrent sur des éléments distincts du processus d'enseignement.

3. Évaluation de la compétence d'enseignant

Avant de décrire les implications de notre modèle général de collecte de preuves sur la compétence d'enseignant, nous commençons par quelques observations sur l'importance de la validité conceptuelle en tant que trait d'union pour déterminer la qualité de l'évaluation. Plusieurs exigences de qualité spécifiques peuvent être déduites de ce cadre de validité conceptuelle.

3.1 Critères de la validité conceptuelle des instruments de compétence

Les exigences imposées pour les instruments de compétence varient selon l'objectif de l'évaluation. Compte tenu des conséquences pour le candidat, les exigences imposées pour les instruments «à grand enjeu» (sélection, certification) sont plus strictes que celles revendiquées pour les instruments de développement professionnel (Pelkmans, 1998).

Le cadre le plus extensif pour déterminer la qualité d'instruments a été élaboré par Messick (1996). Ce dernier indique qu'il est nécessaire, pour chaque forme d'évaluation, d'envisager six aspects de la validité conceptuelle:

- (a) le contenu,
- (b) la théorie et les modèles de traitement,
- (c) la structure,
- (d) la généralisabilité,
- (e) les aspects extérieurs,
- (f) les conséquences.

Le «contenu» concerne la pertinence et la représentativité de l'évaluation. La question qui se pose est la suivante: «dans quelles limites peut-on tirer des conclusions de l'évaluation?» L'aspect «théorie et modèles de traitement» concerne la mesure dans laquelle les tâches sélectionnées nécessitent la prise de l'action pertinente de la part d'un candidat et vise à savoir si l'influence de facteurs inadaptés sur le plan conceptuel est minimisée. La «structure» vise à déterminer si les critères de performance reflètent correctement les critères utilisés par les experts, ainsi que la précision et la cohérence de la cotation et de l'évaluation de la performance.

La «généralisabilité» désigne la mesure dans laquelle les évaluations peuvent être généralisées à un univers de tâches et d'environnements, par exemple.

Les aspects extérieurs de la validité concernent l'étendue selon laquel-

le les résultats de mesure convergent avec d'autres mesures et concepts ou, au contraire, divergent. Les «conséquences» – ou validité conséquente – examinent dans quelle mesure l'instrument a des effets secondaires positifs ou négatifs sur le mode d'apprentissage de l'étudiant et le mode d'enseignement de l'enseignant.

3.2 Recueillir des preuves de compétence

L'utilisation de notre modèle global de compétence d'enseignant en tant que cadre interprétatif pour l'évaluation, ainsi que la prise en considération de critères pour la validité conceptuelle, permettent de décrire les conséquences pour la construction d'un domaine thématique et des procédures pour la collecte d'éléments probants. Trois des aspects de Messick méritent que l'on s'y intéresse particulièrement: 1) le contenu, 2) la théorie et les modèles de traitement et 3) la généralisabilité.

3.2.1 Développer un domaine de compétence

Suivant l'aspect «contenu» de Messick, le contenu d'évaluation devrait être pertinent et représentatif de la profession enseignante. Les examens du contenu sont utilisés pour fixer les limites dans lesquelles des déductions concernant la compétence des enseignants sont faites. Diverses procédures complémentaires sont généralement adoptées lors de l'établissement d'un domaine de compétence: analyses empiriques sur le fonctionnement des enseignants, consultation de sommités de l'enseignement, recherche empirique sur des variables contribuant à une meilleure performance enseignante et consultation de comités de professionnels praticiens (cf. Verloop, Beijaard et Van Driel, 1998). Le mélange de perspectives scientifiques et de perspectives pratiques contribue à l'acceptation et à l'utilisabilité pratique des instruments (Beijaard et Verloop, 1996; Duke et Stiggins, 1990; Uhlenbeck, 2002).

La base de toute évaluation doit contenir une vue d'ensemble d'aspects de la compétence, des situations dans lesquelles ces aspects doivent être démontrés et le degré de maîtrise souhaité. Trois questions nécessitent une réponse:

- (a) Quel est le contenu le plus important de la compétence?
- (b) De quelle manière les critères de performance sont-ils définis?
- (c) Comment les niveaux de compétence peuvent-ils être évalués?

3.2.1.1 Sélection du contenu

Afin de démarquer les domaines de compétence, il est indispensable de sélectionner ce qui est caractéristique du fonctionnement professionnel adéquat et ce qui est vital pour le fonctionnement. Partant de notre modèle de compétence, les principales questions sont les suivantes:

- Quelles sont les performances que les enseignants sont censés démontrer et dans quelles situations fonctionnelles?
- À quel degré de difficulté des situations fonctionnelles les enseignants

doivent-ils pouvoir faire face?

- Quels résultats estudiantins («conséquences») peut-on attendre des activités de l'enseignant?
- Quelles actions et processus de prise de décision pourraient mettre les enseignants en mesure de contribuer aux résultats des étudiants?

Diverses descriptions de domaine pour la compétence d'enseignant ont été élaborées, tant aux États-Unis qu'aux Pays-Bas. Danielson et McGreal (2000) ont distingué quatre grands domaines fonctionnels professionnels pertinents: 1) la planification et la préparation, 2) l'instruction, 3) l'environnement de classe et 4) les responsabilités professionnelles. S'agissant de l'environnement de classe, ils déclarent: «[...] ces activités et tâches établissent un environnement de classe confortable et respectueux qui cultive une culture d'apprentissage et crée un lieu sûr pour la prise de risque» (op. cit., p. 31). Concernant l'instruction, ils écrivent: «[...] Les enseignants qui excellent dans le troisième domaine [instruction] créent une atmosphère marquée par l'enthousiasme sur l'importance de l'apprentissage et la signification du contenu» (op. cit., p. 32). Ces déclarations illustrent l'accent qui est mis sur les conséquences des actions plutôt que sur les actions elles-mêmes.

La description donnée par Danielson et McGreal a constitué également la base d'un ensemble d'instruments d'évaluation baptisé «Praxis III» dans le cadre de la fameuse série Praxis, élaborée par le service des tests éducatifs pour mesurer, notamment, la pratique en classe des enseignants dans chacun des quatre domaines mentionnés. Dans l'évaluation Praxis-III, les observations en classe du comportement de l'enseignant et de l'étudiant sont combinées à des entretiens de pré- et post-observation (Dwyer, 1998), le dernier traitant de la façon dont les enseignants prennent les décisions.

Le Conseil national américain pour les normes professionnelles de l'enseignement (*American National Board for Professional Teaching Standards*) réduit la multiplicité des tâches dans l'évaluation des enseignants (expérimentés) aux dimensions de l'expertise enseignante, telles que l'improvisation, le degré de défi, la passion pour l'enseignement et l'apprentissage (Bond et al., 2000).

Aux Pays-Bas, la fondation néerlandaise pour la compétence professionnelle d'enseignement (SBL) a mis au point dernièrement un ensemble de critères initiaux de compétence pour les enseignants de l'enseignement primaire et de l'enseignement secondaire relevant du ministère de l'éducation, en se basant sur sept grands domaines de compétence. Le premier élément de la description des exigences de la SBL est le fonctionnement des classes et des étudiants individuels d'enseignants compétents, ce qui illustre une approche cohérente de l'enseignement. Les domaines suivants ont été développés:

- (a) la compétence interpersonnelle: capacité de créer un climat chaleureux et coopératif et de développer une communication ouverte;
- (b) la compétence pédagogique: capacité de créer un environnement

- apprenant psychologiquement sain pour les étudiants et de contribuer ainsi à leur bien-être;
- (c) la compétence thématique et didactique: la capacité de guider les étudiants vers l'acquisition des bases de la matière scolaire et la manière dont cette connaissance peut être utilisée dans la vie quotidienne et professionnelle;
 - (d) la compétence organisationnelle: capacité de créer un bon climat de concentration sur la tâche dans leur classe;
 - (e) la compétence collégiale: capacité d'orienter son propre travail sur celui des collègues et de contribuer à l'organisation scolaire d'une manière générale;
 - (f) la compétence de coopération avec l'environnement scolaire: capacité de contribuer à la coopération avec les personnes (parents) et organisations dans le contexte scolaire;
 - (g) la compétence de réflexion et développement: la capacité de remettre sa propre compétence en question et de s'adapter aux exigences et développements en évolution de la profession.

Cet ensemble d'exigences constituera la base de nombreux instruments d'évaluation au niveau des collèges de formation des enseignants.

3.2.1.2 Critères de performance

À présent que les domaines de compétence ont été décrits, une grande question se pose: comment formuler les critères par rapport auxquels mesurer la performance d'enseignant? Conformément à notre modèle, une approche complète de la formulation de critères s'avère souhaitable, pour éviter de trop dépendre d'activités enseignantes isolées, d'aspects de connaissance distincts ou des résultats obtenus par les étudiants. Ces éléments de performance compétente devraient, de préférence, être combinés dans le cadre d'éléments verbaux de description de critères. Des exemples de critères unilatéraux, parfois tautologiques, sont faciles à trouver: «l'enseignant indique clairement», «choisit correctement la matière». Suivant notre modèle de compétence (cf. graphique 1), les critères de performance commencent par des activités apprenantes souhaitables et l'enregistrement de résultats chez l'étudiant, éléments à partir desquels des actions enseignantes et des décisions acceptables peuvent être tirés. L'acceptabilité des décisions d'un enseignant est liée à la qualité de la base de connaissance professionnelle par rapport à des situations d'enseignement spécifiques. Frederiksen et al., (1998) parlent de «critères fonctionnels». Roelofs et Van den Berg (2005) présentent un exemple de tels critères, extrait de la compétence instructionnelle d'enseignants de l'école maternelle ciblant l'acquisition conceptuelle de jeunes enfants: «grâce à des activités instructionnelles (questions, explications, tâches de représentation et discussions), l'enseignant réussit dans sa mission dès lors que les enfants exécutent des activités qui contribuent à leur faire mieux connaître un ensemble sélectionné de concepts (par exemple, l'automne)».

3.2.1.3 Niveaux de performance

Lorsque des critères peuvent être considérés comme des déclarations de performance compétente dans le cadre de situations fonctionnelles, les normes examinent la qualité des actions et leurs résultats. Pour les objectifs que nous poursuivons, une discussion sur le développement des normes de performance dépasse les limites du présent article. Nous préférons souligner l'importance de disposer d'un modèle interprétatif capable de décrire et d'expliquer les différences dans les niveaux de performance. Un modèle qui tient compte des différences entre les novices et les experts au sein d'une profession peut accroître la validité conceptuelle de l'évaluation d'enseignant. Les travaux de Berliner (2001) dans le domaine du développement de l'expertise sont particulièrement significatifs. Berliner résume les différences entre experts et novices. Les experts:

- (a) excellent dans leur propre domaine spécialiste et dans des contextes spécifiques;
- (b) automatisent les actions récurrentes;
- (c) sont plus opportunistes et plus souples;
- (d) sont plus sensibles aux exigences fonctionnelles et aux situations lorsqu'ils résolvent des problèmes;
- (e) représentent des problèmes d'une manière qualitativement différente (plus riche);
- (f) reconnaissent des modèles de situation de travail d'une manière plus rapide et plus précise;
- (g) respectent des modèles plus significatifs dans leur domaine d'expérience;
- (h) se préparent plus longuement pour résoudre des problèmes et utilisent des sources d'information plus riches et plus personnelles.

Les caractéristiques des enseignants experts, telles que Berliner les a relevées, ont été utilisées récemment dans un modèle visant à examiner la validité conceptuelle du système de certification NBTPS pour enseignants expérimentés (Bond et al., 2000). Les résultats ont montré que, dans la majorité des dimensions, les enseignants certifiés étaient plus performants que les enseignants non agréés.

3.2.2 Sources d'éléments probants de compétence

L'accent, lors de l'élaboration d'instruments, est mis sur l'obtention des meilleures preuves possibles de compétence chez un candidat. Si l'on tient compte des exigences de validité qui sont associées à la représentation de la matière, à la théorie sous-jacente et aux modèles de traitement et des exigences de généralisabilité, la cote «nec plus ultra» fait référence à la représentativité des tâches et des situations fonctionnelles et au degré auquel les processus escomptés et effets de la performance compétente sont dûment représentés dans l'évaluation.

Le premier choix pour la collecte de preuves de compétence porte sur la nature de la preuve. Les formes de preuve fondamentales sont: documents de préparation du cours, observation du cours (directe ou enregis-

trée, accent sur les actions de l'enseignant ou les activités des étudiants), journaux de classe des enseignants (concentration sur les actions), entretiens de réflexion (concentration sur les processus de prise de décision), comptes rendus de réflexion (concentration sur les processus de prise de décision), tests des étudiants (concentration sur les résultats), épreuve écrite passée par l'enseignant (concentration sur la base de connaissances ou sur les processus de prise de décision) et épreuve multimédia de l'enseignant (concentration sur la base de connaissances ou le processus de prise de décision).

Suivant notre modèle de compétence, toutes les preuves de compétence devraient être enregistrées et interprétées dans des situations d'enseignement spécifiques. Les instruments de compétence diffèrent fortement à cet égard.

Les observations de cours vont, par exemple, des évaluations hors contexte, basées sur des visites, aux cours à proprement parler («l'enseignant explique clairement») en passant par des rapports narratifs sur des épisodes de cours, ou des séquences vidéo non filtrées. Cela s'applique aussi au recueil de documents didactiques qui peuvent se présenter, par exemple, sous la forme d'éléments provenant d'une série complète de cours, ou d'extraits de ce que l'enseignant considère comme étant ses meilleurs travaux. Le contenu de la documentation peut également varier fortement. Il peut se rapporter aux plans de cours de l'enseignant, à des (exemples de) répercussions de l'information de l'enseignant vers l'étudiant, ou à des (exemples d') informations que l'étudiant retire d'une telle répercussion de l'information. Il convient en tout cas de déterminer qui se trouve dans la meilleure position pour fournir les éléments probants nécessaires pour justifier la compétence: l'enseignant lui-même, ses collègues, les étudiants, les parents, les dirigeants, des experts externes, ou d'autres personnes? Selon notre modèle de compétence, tout participant à l'évaluation devrait être en mesure de fournir des preuves de compétence représentatives et convaincantes en rapport avec les conséquences, les actions des enseignants et les processus décisionnels. Peterson (2002) décrit, pour différentes sources de données, les avantages et les inconvénients de l'implication de chacun de ces participants.

Le deuxième choix dans la collecte de preuves de compétence inclut la sélection d'un ensemble de tâches et de situations fonctionnelles pouvant être considéré comme représentatif du domaine de compétence en cours d'étude, tant sur le plan quantitatif que sur le plan qualitatif. Les questions suivantes appellent une réponse:

- (a) la situation choisie, ou l'ensemble des situations retenues, sont-elles représentatives d'actions se produisant dans le contexte professionnel?
- (b) la ou les tâches sélectionnées et la ou les situations fonctionnelles retenues sont-elles pertinentes ou critiques pour la compétence à prouver?
- (c) quel est le degré de difficulté ou de complexité de la tâche ou de la situation fonctionnelle?

- (d) le candidat a-t-il la possibilité de fournir les preuves de compétence nécessaires?
- (e) l'ensemble sélectionné de tâches ou de situations fonctionnelles couvre-t-il toutes les tâches et situations fonctionnelles?
- (f) les déclarations sur les situations mesurées peuvent-elles, par extrapolation, être étendues à la situation de travail?

Les tâches et situations fonctionnelles diffèrent dans leur degré d'authenticité: situations d'évaluation réelles, réelles simplifiées, simulées et symbolisées. L'évaluation en situation réelle implique que le candidat exécute des tâches qui se produisent dans la réalité de tous les jours, sans intervention dans cette situation. Parmi les exemples de situation réelle, citons le fait de donner cours à sa propre classe, ou de surveiller les étudiants de son propre groupe lors de l'exécution de tâches indépendantes. Toutes les fonctions enseignantes peuvent se produire et doivent être exécutées sur le champ. La réussite ou l'échec de l'exécution des tâches enseignantes a des conséquences directes pour l'étudiant.

Dans une situation fonctionnelle réelle simplifiée, le candidat exécute une tâche réelle, qui est cependant moins complexe que dans la réalité, comme, par exemple, donner une mini-leçon à un petit groupe d'étudiants. On ne retrouve pas dans la situation de travail simulée les conséquences directes pour l'étudiant et la possibilité de «mettre un terme» à la situation de travail. L'authenticité des preuves de compétence dans les situations fonctionnelles réelles simplifiées ou simulées est réduite par comparaison avec les situations de travail réelles. L'avantage, est toutefois, que ces situations permettent de présenter des tâches pertinentes qui ne se produisent pas fréquemment dans des conditions de travail. La compétence peut aussi être évaluée dans des situations de cours symboliques, c'est-à-dire dans des situations qui ne se produisent pas en réalité et où la pression temporelle et l'urgence immédiate font défaut. L'accent est mis sur la collecte d'éléments probants concernant la prise de décisions dans diverses situations fonctionnelles. Bien que le niveau d'authenticité soit faible, les situations de cours décrites permettent une bonne couverture des tâches et des situations.

Les concepteurs d'instruments d'évaluation ne tiennent généralement pas compte des délibérations sur la nature et l'ampleur des preuves requises et se lancent immédiatement dans la conception des instruments. L'importance des diverses sources de preuve peut, dès lors, devenir incertaine. Ce problème se pose lors de l'évaluation de portefeuilles (non structurés). La nature de la collecte d'éléments probants peut varier considérablement dans le cas des portefeuilles. Ces derniers peuvent contenir des preuves directes sous la forme de reproductions de cours, de réalisations d'étudiants et de rapports de réflexion, mais aussi de produits qui sont, en soi, le résultat d'évaluations, tels que les résultats d'épreuves écrites, des lettres de recommandation et des évaluations par des homologues. Un portefeuille peut également être très utile pour réunir diverses preuves de compétence. La capacité d'évaluation reste toutefois fortement dépendante de

la structure du portefeuille et de l'admissibilité, l'observabilité et la scorableté des preuves enregistrées (Heller, Sheingold et Myford, 1998).

4. Discussion

Cet article présente un cadre détaillé d'évaluation de la compétence d'enseignant qui peut constituer la base d'évaluations de qualité. La présente section commence par tirer quelques conclusions, puis discute des avantages éventuels du modèle lorsqu'il est utilisé pour créer des arguments interprétatifs. Elle conclut en examinant des applications récentes du modèle dans l'évaluation vidéo de portefeuille.

La première partie de l'article est consacrée au développement du modèle. En partant d'études de la littérature y afférente, elle a conclu qu'il n'existe aucun cadre scientifique pertinent et généralisé déterminant de ce qui constitue l'enseignement compétent. Plusieurs éléments différents de la compétence d'enseignant ont été mis en exergue dans toute l'histoire de l'évaluation des enseignants, en l'occurrence: les traits de personnalité qui font d'une personne un bon enseignant, des éléments de savoir essentiels impliquant le contenu de la matière, le mode de pensée de l'enseignant dans une discipline, les formes de comportement professoral qui contribuent à la performance de l'enseignant, les connaissances pratiques et les théories subjectives des enseignants qui déterminent les actions des enseignants dans des situations d'enseignement spécifiques et, enfin, l'enseignement en tant qu'élément pour stimuler des activités d'apprentissage puissantes entre apprenants.

Tous les éléments distincts de l'enseignement couvrent certains aspects de la compétence d'enseignant, mais aucun d'entre eux ne décrit pleinement ou n'explique réellement le sens réel de l'enseignement compétent. Un concept détaillé et unifié de la compétence d'enseignant a, dès lors, été introduit afin de tenir compte de tous les éléments différents intervenant.

En résumé, le modèle avance que la compétence d'enseignant se reflète dans les conséquences des actions des enseignants, dont les plus importantes sont les activités apprenantes des étudiants. En partant des conséquences, les autres composants du modèle ont été répertoriés à reculons. Le composant «actions» fait référence aux activités professionnelles qui stimulent l'apprentissage par les étudiants, ou à d'autres conséquences. Le composant «prise de décision» implique qu'un enseignant doit prendre de nombreuses décisions à long terme ou immédiates dans une situation de classe, telles que, par exemple, déterminer si certaines actions doivent ou non être initiées. Il a, en outre, été souligné que la prise de décision, les actions et les conséquences se déroulent dans un contexte spécifique dans lequel les enseignants exécutent leurs tâches professionnelles. Lors de la prise de décision et de l'exécution d'activités, les enseignants doivent enfin partir d'une ba-

se de connaissances professionnelle et de certaines caractéristiques personnelles.

La seconde partie de l'article décrit la manière dont le modèle peut aider à développer des domaines d'évaluation et des critères de performance et à recueillir des preuves de compétence. Elle examine également la manière dont on peut satisfaire trois des critères de validité conceptuelle avancés par Messick: le contenu, les modèles de théorie et de traitement et la généralisabilité.

Suivant le modèle de la performance compétente, les critères devraient démarrer par les activités et les résultats apprenants qui apparaissent souhaitables chez les étudiants et, partant de là, déduire les actions et décisions que l'enseignant devrait prendre. Les domaines de compétence sont de plus en plus affirmés de cette manière, sans favoriser une quelconque ligne d'action, mais en décrivant de grandes catégories d'activités et en développant des activités apprenantes souhaitables. Afin de distinguer des niveaux de performance souhaitables, des comparaisons systématiques peuvent également être établies entre les novices et les experts d'une même profession.

Le développement d'instruments doit se concentrer sur l'obtention de la meilleure preuve de compétence possible des candidats. Si l'on tient compte des exigences de validité associées à la représentation du contenu, à la théorie et aux modèles et de traitement sous-jacents et à la généralisabilité, la cote «*nec plus ultra*» fait référence à la représentativité des tâches et des situations fonctionnelles et au degré auquel les processus et effets escomptés de la performance compétente sont représentés de manière adéquate dans l'évaluation.

Lors de la sélection des sources d'éléments probants, cela devrait dépendre du degré auquel le processus complet de performance compétente est représenté, à savoir les décisions de l'enseignant, les actions de l'enseignant, ou les actions de l'étudiant. Enfin, il convient de sélectionner un ensemble de tâches et de situations fonctionnelles qui peut être considéré comme représentatif du domaine de compétence en cours d'étude sur le plan tant quantitatif que qualitatif.

Selon Kane (1992), le modèle peut être utilisé pour créer des arguments interprétatifs afin d'étayer des appréciations sur la compétence d'enseignant. Tout comme Kane, nous pensons qu'il est peu probable que la compétence puisse un jour être prouvée. Tout au plus parviendrons-nous au mieux à adopter un argument interprétatif plausible sur la compétence d'enseignant. Si les évaluateurs sont capables d'interpréter les résultats de l'évaluation par rapport aux processus postulés de notre modèle, l'argument interprétatif est alors soutenu. À titre d'exemple, les évaluateurs qui jugent la qualité de l'instruction peuvent interpréter les résultats de l'étudiant sur la façon dont les enseignants prennent leurs décisions lorsqu'ils donnent une instruction, ou la façon dont ils agissent et les conséquences qui en découlent pour les étudiants dans un environnement de classe spécifique.

Un avantage du modèle est que les conceptions variables ou changeantes sur l'enseignement n'affectent pas sa structure. Différentes visions de l'enseignement, telle que l'instruction programmée par opposition à l'apprentissage par la découverte, se refléteront d'une manière ou d'une autre dans un certain type d'activités apprenantes souhaitables, un répertoire d'actions adéquates et des processus de décision connexes de la part de l'enseignant. Quelle que soit la vision de l'enseignement, les concepteurs d'évaluations peuvent baser leurs décisions de collecte de données plus sciemment sur les processus qu'ils aimeraient éclairer dans leurs évaluations.

Le modèle présenté peut constituer la base d'un développement professionnel plus poussé de l'enseignant, étant donné qu'il décrit les processus auxquels celui-ci prend part. Ces processus peuvent être modifiés et adaptés lorsque l'enseignant reçoit un retour d'informations et s'engage dans des activités de réflexion. Il peut donc aider à améliorer la qualité de la formation et de l'apprentissage des enseignants.

Le fait de partir d'un modèle complet de compétence d'enseignant soulève la question de savoir dans quelle mesure les divers aspects de la performance compétente doivent tous être couverts lors d'un même exercice d'évaluation. Autrement dit, comment peut-on combiner plusieurs tâches et sources de preuve différentes en une seule évaluation? Nous aimerions conclure en citant les résultats d'une étude récente menée par l'université de Leyde et l'Institut national néerlandais de recherches et d'élaboration de tests (CITO), où notre modèle a été utilisé comme base de collecte de preuves sur les compétences. Utilisant des portefeuilles vidéo, diverses preuves de performance compétente associées au même éventail de situations d'enseignement ont été recueillies de manière cohérente. Les actions des enseignants (sur vidéo), les processus décisionnels sous-jacents (par entretien), les conséquences observées au niveau des étudiants (sur vidéo) et le contexte pédagogique (situation) dans lequel les actions enseignantes ont été prouvées (à l'aide de documents de cours) ont été enregistrés. Un système de cotation, permettant à des évaluateurs expérimentés de se faire une idée générale de la compétence instructionnelle, a été instauré. Les premiers résultats d'une étude pilote (Roelofs et Van den Berg, 2004, 2005) montrent que les portefeuilles sont des ensembles de preuves cohérents et évaluables. Les évaluateurs ne se servent toutefois pas de l'ensemble des preuves pour parvenir à leur jugement. Les évaluateurs disposaient de preuves sur la façon dont les enseignants prenaient leurs décisions et sur le contexte fonctionnel, mais ils ne les ont pas utilisées pour leur appréciation. Au moment de notifier leur appréciation aux enseignants, ils ont cependant discuté de ces sources de preuve, afin de leur fournir un retour d'informations plus complet. Cela illustre la façon dont le modèle général de compétence a été utilisé pour fournir un argument interprétatif.

Bibliographie

- Beijaard, D.; Verloop, N. Évaluer la connaissance pratique des enseignants. *Studies in Educational Evaluation*, 1996, vol. 22, n° 3, p. 275-286.
- Bellon, J.J.; Bellon, E.C.; Blank, M.A. *Teaching from a research knowledge base*. New York: MacMillan, 1992.
- Berliner, D.C. En apprendre sur les enseignants experts et tirer des enseignements de ceux-ci. *International Journal of Educational Research*, 2001, vol. 35, p. 463-482.
- Bond, L. et al. *The certification system of the National Board for Professional Teaching Standards: a construct and consequential validity study*. Greensboro, NC: University of North Carolina, Center for Educational Research and Evaluation, 2000.
- Brophy, J.; Good, T.L. Le comportement de l'enseignant et les réalisations de l'étudiant. In Wittrock, M.C. (dir.) *Handbook of research on teaching*, 3^e dir. New York: MacMillan, 1986, p. 328-37.
- Bruner, J.S. *The process of education*. Cambridge: Harvard University Press, 1963.
- Creemers, H.P.M. *Effectieve instructie: een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs, 1992.
- Danielson, Ch.; McGreal, T.L. *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association pour la supervision et le développement de programme, 2000.
- Doyle, W. Ouvrage académique. *Review of Educational Research*, 1983, vol. 53, p. 159-199.
- Duke, D.L.; Stiggins, R.J. Au delà de la compétence minimale: évaluation du développement professionnel. In Millmann J.; Darling-Hammond L. (dir.) *The new handbook of teacher evaluation. Assessing elementary and secondary teachers*. Newbury Park, CA: Sage, 1990, p. 116-132.
- Dwyer, C.A. Critères pour les évaluations d'enseignants sur la base de la performance: validité, normes et problèmes. *Journal of Personnel Evaluation in Education*, 1994, vol. 12, n° 2, p. 135-150.
- Dwyer, C.A. Psychométrie de Praxis III: évaluations de la performance d'une classe. *Journal of Personnel Evaluation in Education*, 1998, vol. 12, n° 2, p. 163-187.
- Frederiksen, J.R. et al. Évaluation par portefeuille vidéo de l'enseignement. *Educational Assessment*, 1998, vol. 5, n° 4, p. 225-298.
- Getzels J.W.; Jackson, P.W. La personnalité et les caractéristiques de l'enseignant. In Gage, N.L. (dir.) *Handbook of research on education*. Chicago: Rand McNally, 1963.
- Haertel, E.H. Nouvelles formes d'évaluation de l'enseignant. In Grant, G. (dir.) *Review of research in education*. Washington, DC: American Educational Research Association, 1991, p. 3-29.

- Heller, J.I.; Sheingold, K.; Myford, C.M. Raisonnement sur les preuves en portefeuilles: fondements cognitifs pour une évaluation valable et fiable. *Educational Assessment*, 1998, vol. 5, n° 1, p. 5-40.
- Kagan, D.M. Modes d'évaluation de la cognition de l'enseignant: déductions concernant le principe de Goldilocks. *Review of Educational Research*, 1990, vol. 60, n° 3, p. 419-469.
- Kane, M.T. Une approche de la validité basée sur l'argument. *Psychological Bulletin*. 1992, vol. 112, n° 3, p. 527-535.
- Klarus, R.; Schuler, Y.; Ter Wee, E. *Bewezen geschikt: een assessment-procedure voor aspirant leerkrachten (zij-instromers) in het PO en VO*. Wageningen: Stoas Onderzoek, 2000.
- Latham, A.S.; Gitomer, D.H.; Ziomek, R. Ce que les tests nous disent sur les nouveaux enseignants. *Educational Leadership*, 1999, vol. 56, n° 8, p. 23-26.
- McConnell, E.A. La compétence par opposition à la compétence caractéristique. *Nursing Management*. 2001, vol. 32, n° 5, p. 14.
- Meijer, P.C.; Verloop, N.; Beijaard, D. Explorer la connaissance pratique des enseignants en langue sur l'enseignement de la compréhension à la lecture. *Teaching and Teacher Education*, 1999, vol. 15, n° 1, p. 59-84.
- Merriënboer, J.J.G. van; Klink, M.R. van der; Hendriks, M. *Competenties: van complicaties tot compromis*. La Haye: Onderwijsraad (Conseil de l'enseignement des Pays-Bas), 2002.
- Messick, S. Validité des évaluations de la performance. In Phillips, G.W. (dir.) *Technical issues in large-scale performance assessments*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 1996.
- Mulder, M. *Competentieontwikkeling in organisaties: perspectieven en praktijk*. La Haye: Elsevier bedrijfsinformatie, 2001.
- Parry, S.B. Quête de compétences. *Training*, juillet 1996, p. 48-56.
- Pelkmans, A. *Evaluatie van leraren en schoolleiders*. Wageningen: Leerstoelgroep Agrarische Onderwijskunde, Landbouwuniversiteit Wageningen, 1998.
- Peterson, K.D. *Teacher evaluation: a comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin Press, 2002.
- Reynolds, M. Qu'est-ce qu'un enseignement débutant compétent? Revue de la littérature. *Review of Educational Research*, 1992, vol. 62, n° 1, p. 1-35.
- Roelofs, E.C.; Sanders, P. Beoordeling van docentcompetenties. In Mulder, M. et al. (dir.) *Bevoegd, maar ook bekwaam?* Groningen: Wolters-Noordhoff, 2003, p. 277-299.
- Roelofs, E.C.; Van den Berg, E. *Images of teacher competence. Design characteristics of a multimedia design portfolio*. Document présenté lors de l'assemblée annuelle de l'American Educational Research Association, Montréal, 2005.
- Roelofs, E.C.; Van den Berg, E. *Videodossiers: leerkrachtbekwaamheid*

- in beeld*. Document présenté lors du séminaire sur la recherche éducative en 2004 (onderwijsresearchdagen) à Utrecht, 2004.
- Shepard, L.A. Évaluation de la validité de test. In Darling-Hammond, L. (dir.) *Review of research in education*. Washington, DC: American Educational Research Association, 1993, p. 405-450.
- Simon, A.; Boyer, E.G. *Mirrors for behaviour: an anthology of classroom observation instruments*. Wyncote, PA: Communication Material Center, 1974.
- Simons, P.R. Competentiegerichte leeromgevingen in organisaties en hoger beroepsonderwijs. In Schlusmans, K. et al. (dir.) *Competentiegerichte leeromgevingen*. Utrecht: Lemma, 1999.
- Spencer, L.M.; Spencer, S.M. *Compétence au travail: modèles de performance supérieure*. New York: Wiley, 1993.
- Stichting Beroepskwaliteit leraren en ander personeel. *Bekwaamheidseisen leraren*. La Haye: Stichting Beroepskwaliteit Leraren en ander onderwijspersoneel, 2003.
- Stodolsky, S.S. Observation de classe. In Darling-Hammond, J.; Millmann, L. (dir.) *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park, CA: Sage, 1990, p. 175-190.
- Straetmans, G.J.J.M.; Sanders, P.F. *Beoordelen van competenties van docenten*. Utrecht: Programmamanagement Educatief Partnerschap, 2001.
- Tijssen, J.G.L. Hindernissen voor competentie management. *Opleiding & Ontwikkeling*, 1998, vol. 11, n° 10, p. 17-23.
- Tijssen, J.G.L. Personele flexibiliteit in strategisch perspectief: de betekenis van employability voor organisaties en individu. In Poel, R.F.; Kessels, J.W.M. (dir.) *Human resource development: Organiseren van het leren*. Alphen aan den Rijn: Samsom, 2001, p. 103-116.
- Tillema, H. *Assessment van competenties: van beoordelen naar ontwikkelen*. Alphen aan den Rijn: Kluwer, 2001.
- Tom, A.R.; Valli, L. Connaissance professionnelle pour enseignants. In Houston R.W. (dir.) *Handbook of research on teacher education*. New York: MacMillan, 1990, p. 372-392.
- Uhlenbeck, A. *The development of an assessment procedure for beginning teachers of English as a foreign language*. Leiden: ICLON Graduate School of Education, 2002.
- Verloop, N. *De leraar: reviewstudie uitgevoerd in opdracht van PROO*. La Haye: NWO/PROO, 1999.
- Verloop, N. Examen des connaissances d'enseignant. *Journal of Curriculum Studies*, 1988, vol. 20, p. 81-86.
- Verloop, N.; Beijaard, D.; Van Driel, J.H. Beoordeling van docenten. *Pedagogische Studiën*, 1998, vol. 75, n° 6, p. 171-187.
- Vermunt, J.; Verschaffel, L. Enseignement orienté sur le processus. In Simons, P.R.J.; van der Linden, J.; Duffy, T. *New learning*. Dordrecht: Kluwer, 2000, p. 209-225.