

Mesure et évaluation en éducation

MESURE
et
ÉVALUATION
en Éducation

Validation de la version française de l'Échelle du sentiment de responsabilité des enseignants (*Teacher Responsibility Scale*)

Cynthia Vaudroz et Jean-Louis Berger

Volume 41, numéro 3, 2018

Réception : 23 novembre 2015

Version finale : 9 janvier 2018

Acceptation : 16 janvier 2018

URI : <https://id.erudit.org/iderudit/1065166ar>

DOI : <https://doi.org/10.7202/1065166ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Vaudroz, C. & Berger, J.-L. (2018). Validation de la version française de l'Échelle du sentiment de responsabilité des enseignants (*Teacher Responsibility Scale*). *Mesure et évaluation en éducation*, 41(3), 87–117.
<https://doi.org/10.7202/1065166ar>

Résumé de l'article

L'objectif de cet article est de présenter la validation de la version française de l'Échelle du sentiment de responsabilité des enseignants (*Teacher Responsibility Scale*) élaborée par Lauermaann et Karabenick (2013). Dans une première étude, nous avons soumis à 154 enseignants en formation la version française de l'Échelle du sentiment de responsabilité, une échelle mesurant le sentiment d'efficacité personnelle et une échelle de contrôle de la désirabilité sociale. Dans une seconde étude, nous avons testé l'invariance de l'Échelle du sentiment de responsabilité entre hommes et femmes, entre enseignants novices et chevronnés ainsi que selon le type de support utilisé pour la passation du questionnaire (papier ou en ligne) auprès d'un échantillon constitué de 533 enseignants, en réalisant des analyses factorielles confirmatoires multigroupes. Les résultats indiquent que l'échelle possède des propriétés psychométriques de qualité et peut être utilisée à des fins de recherche et/ou de formation.

Validation de la version française de l'Échelle du sentiment de responsabilité des enseignants (*Teacher Responsibility Scale*)*

Cynthia Vaudroz

*Unité de recherche pour le pilotage des systèmes pédagogiques
Département de la formation, de la jeunesse et de la culture,
État de Vaud (Suisse)*

Jean-Louis Berger

*Institut fédéral des hautes études en formation professionnelle,
Renens (Suisse)*

MOTS CLÉS : sentiment de responsabilité, validité, analyse factorielle confirmatoire multigroupes

L'objectif de cet article est de présenter la validation de la version française de l'Échelle du sentiment de responsabilité des enseignants (Teacher Responsibility Scale) élaborée par Lauermann et Karabenick (2013). Dans une première étude, nous avons soumis à 154 enseignants en formation la version française de l'Échelle du sentiment de responsabilité, une échelle mesurant le sentiment d'efficacité personnelle et une échelle de contrôle de la désirabilité sociale. Dans une seconde étude, nous avons testé l'invariance de l'Échelle du sentiment de responsabilité entre hommes et femmes, entre enseignants novices et chevronnés ainsi que selon le type de support utilisé pour la passation du questionnaire (papier ou en ligne) auprès d'un échantillon constitué de 533 enseignants, en réalisant des analyses factorielles confirmatoires multigroupes. Les résultats indiquent que l'échelle possède des propriétés psychométriques de qualité et peut être utilisée à des fins de recherche et/ou de formation.

KEY WORDS : sense of responsibility, validity, multi-group confirmatory factor analysis

This article aims at presenting the validity of the French version of the Teacher Responsibility Scale, which was developed by Lauermann and Karabenick (2013). In a first study, we administered the French version of the Teacher Responsibility

* Le processus d'évaluation et la décision d'acceptation de cet article ont été sous la responsabilité du précédent rédacteur européen, M. Christophe Dierendonck.

Scale, a scale measuring teachers' self-efficacy, and a scale controlling for social desirability, to 154 trainee teachers. In a second study, we tested the scale invariance between men and women, between new and experienced teachers, and between a paper and pencil version and an online version in a sample of 533 teachers, using multi-group confirmatory factor analyses. Results indicate that the scale has good psychometric properties and could be used for research and education.

PALAVRAS-CHAVE: sentimento de responsabilidade, validade, análise fatorial confirmatória multigrupos

O objetivo deste artigo é apresentar a validação da versão francesa da Escala do sentimento de responsabilidade dos professores (Teacher Responsibility Scale) desenvolvida por Lauermann e Karabenick (2013). Num primeiro estudo, submetemos a 154 professores em formação a versão francesa da Escala do sentimento de responsabilidade, uma escala que mede o sentimento de eficiência pessoal e uma escala que mede o controlo da desejabilidade social. Num segundo estudo, testamos a invariância da Escala do sentimento de responsabilidade entre homens e mulheres, entre professores principiantes e professores experientes, bem como pelo tipo de suporte usado para preenchimento do questionário (em papel ou online) junto de uma amostra de 533 professores, realizando análises fatoriais confirmatórias multigrupos. Os resultados indicam que a escala possui propriedades psicométricas de qualidade e pode ser utilizada para fins de investigação e formação.

Note des auteurs: La correspondance liée à cet article peut être envoyée à [cynthia.vaudroz@vd.ch] et [jean-louis.berger@iffp.swiss].

La rédaction de cet article a été réalisée dans le cadre du projet *L'évolution des conceptions des enseignant-e-s durant leur formation* (subside n°100019_146351), soutenu par le Fonds national suisse de la recherche scientifique.

Introduction

La « professionnalisation des enseignants » : voilà une expression utilisée depuis quelques années par les chercheurs en sciences de l'éducation, par les formateurs d'enseignants et par les responsables des politiques éducatives (Paquay, Altet, Charlier et Perrenoud, 2012). Cette notion recouvre, selon Altet (2010), une multitude de compétences, de capacités et d'aptitudes portant notamment sur les savoirs spécifiques au métier, sur les représentations de normes collectives sur lesquelles se construit l'identité professionnelle, mais aussi sur des aspects éthiques tels qu'« une responsabilité personnelle dans l'exercice de ses compétences » (Altet, 2010, p. 16). Cette responsabilité présente une importance grandissante dans l'exercice du métier étant donné les attentes sociétales vis-à-vis du corps enseignant. En effet, tant les parents d'élèves que les directions d'école accordent de plus en plus d'importance au rôle de l'enseignant dans la réussite des apprenants et le désignent comme étant largement responsable de celle-ci. Le degré élevé d'autonomie inhérent au métier requiert de l'enseignant qu'il assume personnellement un certain nombre de responsabilités qui lui sont attribuées (notamment faire apprendre, engager les apprenants, évaluer les apprentissages), mais aussi des responsabilités morales propres à son rôle telles que celle d'établir et de maintenir des relations de confiance avec les apprenants.

En dépit de son évidente pertinence, la responsabilité de l'enseignant est un sujet moins abordé par les recherches, qu'elles soient francophones ou anglo-saxonnes, que ne le sont les aspects cognitifs de l'enseignant, notamment ses processus de pensée (Fives et Gill, 2015). Toutefois, selon la discussion de la responsabilité d'un point de vue éthique rédigée par Oser (1994), tout acte d'enseignement comprendrait un aspect moral en son cœur et la responsabilité de l'enseignant constituerait une motivation de type moral. Ainsi, il est nécessaire de s'intéresser également aux aspects moraux du métier d'enseignant si l'on cherche à comprendre comment et pourquoi le métier est exercé. Partant du constat selon lequel les systèmes de responsabilité fondés sur la performance des apprenants (*performance-based accountability*) instaurés dans les écoles donnent pour responsabilité

aux enseignants d'améliorer les résultats des apprenants, une ligne de recherche sur ce sujet a été récemment développée aux États-Unis d'Amérique (Lauermann et Karabenick, 2011a, 2013). Bien que les systèmes éducatifs diffèrent sensiblement d'un pays à l'autre, le cœur des activités (gestion de la classe, planification de cours, évaluation des apprentissages) qu'ils assument est largement similaire. Ainsi, la notion de sentiment de responsabilité des enseignants n'est pas seulement pertinente dans le pays d'origine de cette nouvelle ligne de recherche; elle l'est également dans nombre de systèmes éducatifs européens, dont le système suisse dans le cadre duquel s'inscrit la présente étude.

Depuis les années 1980, les chercheurs dans le champ de la psychologie de l'éducation s'intéressent à la façon dont les enseignants conçoivent leurs responsabilités professionnelles (Guskey, 1982; Potvin et Rousseau, 1993; Rose et Medway, 1981). En particulier, ces études s'appuient sur les théories du locus de contrôle et des attributions causales afin de comprendre dans quelle mesure les enseignants pensent être à l'origine des résultats positifs et/ou négatifs des apprenants. Autrement dit, il s'agit de comprendre les croyances des enseignants relatives à leur contrôle sur les résultats et aux raisons pour lesquelles les apprenants réussissent ou échouent. Sur la base de ces travaux et d'études ayant porté sur le concept de responsabilité au-delà de son application à l'enseignement, Lauermann et Karabenick (2011a) ont développé théoriquement la notion de *sentiment de responsabilité des enseignants*.

Le sentiment de responsabilité est défini comme un sens d'obligation et d'engagement interne à produire ou prévenir certains résultats (Lauermann et Karabenick, 2011a). Les individus adoptent un comportement non pas parce que celui-ci leur est agréable ou bénéfique, mais parce qu'ils ressentent une obligation de se comporter ainsi. La responsabilité dont il est ici question ne se réfère pas à des obligations venant de l'institution ou d'autres sources externes, mais bel et bien d'une source interne, soit l'enseignant. Ce sentiment pourrait être qualifié de croyance sur le soi enseignant, inscrivant ce sentiment dans la notion des croyances sur soi (*self-beliefs*; Pajares, 1992), au même titre que le sentiment d'efficacité personnelle à enseigner par exemple. Il s'agit d'une croyance subjective, bien qu'elle soit influencée par des connaissances objectives de ce que le système éducatif, la direction de l'établissement scolaire ou encore les lois requièrent formellement d'un enseignant. La définition du senti-

ment de responsabilité en matière de croyances des enseignants signifie également qu'il est probable que celui-ci exerce une influence instrumentale sur la définition des tâches et sur la sélection des instruments cognitifs avec lesquels l'enseignant interprète, planifie et prend des décisions relatives à ces tâches. Ainsi, il jouerait un rôle critique dans la définition des comportements et dans l'organisation de l'information (Pajares, 1992). Le sentiment de responsabilité se différencie des concepts de locus de contrôle, d'attributions causales et de sentiment d'efficacité personnelle, bien que tous ces concepts soient en lien (Lauermann et Karabenick, 2011a). Alors que le locus de contrôle constitue un trait de personnalité général, le sentiment de responsabilité des enseignants est spécifique aux objets de cette profession. Les attributions causales, quant à elles, portent sur les causes perçues d'un événement; en contraste, le sentiment de responsabilité s'intéresse au degré de contrôle personnel perçu sur cet événement. Enfin, si le sentiment d'efficacité personnelle se réfère aux perceptions de contrôle et de compétence à exécuter des actions, il n'implique pas forcément que l'individu se sente responsable du résultat.

Cinq objets principaux de responsabilité ont été distingués théoriquement par Lauermann et Karabenick (2011a, 2013), soit un sentiment de responsabilité pour :

- 1) la motivation des apprenants (intérêt, enthousiasme et valeur du sujet enseigné);
- 2) la réussite des apprenants (apprentissage, performance et progrès scolaire durant l'année);
- 3) la confiance en soi des apprenants (l'enseignant a pour tâche d'augmenter la confiance des élèves en leurs propres performances);
- 4) les relations avec les apprenants (avoir confiance en l'enseignant, aller demander de l'aide à l'enseignant en cas de problème et savoir que l'enseignant se soucie d'eux); et
- 5) la qualité de l'enseignement (l'enseignant rend ses leçons aussi efficaces et engageantes que possible).

Les premières études, fondées sur le cadre théorique développé par ces chercheurs, indiquent que le sentiment de responsabilité varie de manière interindividuelle (d'un enseignant à l'autre) et intra-individuelle (d'un objet de responsabilité à l'autre pour un même enseignant). Ce sentiment revêt

une pertinence particulière en ce qu'il est associé à des variables importantes, tant en matière d'identité professionnelle que de styles d'enseignement.

La pertinence du sentiment de responsabilité des enseignants

Sentiment de responsabilité et identité professionnelle

De manière consistante au travers de deux études, l'une aux États-Unis d'Amérique et l'autre en Allemagne, Lauermann et Karabenick (2011b) ont observé que les motivations au choix de la carrière enseignante expliquaient certains objets du sentiment de responsabilité (mesuré par la *Teacher Responsibility Scale*; Lauermann et Karabenick, 2013). Les deux études ont montré que les motivations dites d'utilité sociale (désir de travailler avec des jeunes, de favoriser leur développement) expliquaient le degré auquel l'enseignant se sentait responsable de la motivation des apprenants, de leur réussite ainsi que de ses relations avec les apprenants. En contraste, les motivations extrinsèques et intrinsèques ne jouaient qu'un rôle mineur dans le sentiment de responsabilité. Notre propre étude mettant en lien les motivations à devenir enseignant, le sentiment de responsabilité et les styles de gestion de classe (Girardet et Berger, 2016) confirme que la valeur d'utilité sociale est associée positivement aux trois mêmes objets de responsabilité. De plus, un choix de devenir enseignant fondé sur la valeur intrinsèque du métier (intérêt pour les activités de l'enseignement) est associé à un plus fort sentiment de responsabilité pour la qualité de l'enseignement.

Le sentiment de responsabilité jouerait également un rôle sur le plan identitaire (Berger et Lê Van, 2019). En effet, par l'utilisation d'une analyse d'échelonnement multidimensionnel¹ de variables traitant de l'identité enseignante, nous avons observé que le sentiment de responsabilité (sans distinction des objets) se rapprochait non seulement de l'utilité sociale et de la valeur intrinsèque, mais aussi de la perception d'expertise pédagogique et de l'importance des aptitudes dans le choix de devenir enseignant.

Sentiment de responsabilité et style d'enseignement

Le sentiment de responsabilité permettrait de comprendre pourquoi les enseignants adoptent certaines pratiques. En d'autres termes, il pourrait s'agir d'un déterminant du comportement des enseignants en classe. En effet, les études de Lauermann et Karabenick (2011b) indiquent que cer-

taines formes du sentiment de responsabilité expliquent les structures de buts, soit les diverses pratiques de classe qui rendent les buts de maîtrise ou de performance² saillants, ainsi que les messages explicitement liés à des buts que les enseignants communiquent à leurs élèves (Kaplan, Middleton, Urdan et Midgley, 2002). Spécifiquement, le niveau de responsabilité pour la réussite des apprenants est associé positivement à l'adoption d'une structure de buts de maîtrise, tandis que la responsabilité pour la motivation des apprenants est liée positivement à une structure de buts de performance (Lauermann et Karabenick, 2011b). Les associations entre les objets de responsabilité et les structures de buts sont d'ampleur moyenne. Cela révèle que les pratiques de gestion de classe des enseignants peuvent être influencées par le sentiment de responsabilité.

Notre propre étude sur la responsabilité et les styles de gestion de classe (soit les tendances à interagir d'une certaine façon avec les élèves pour les engager dans les apprentissages et à réagir aux problèmes de comportement) indique que le sentiment de responsabilité pour la qualité de l'enseignement joue un rôle central dans l'adoption de l'un ou l'autre de ces styles rapportés par les enseignants. En effet, cet objet de responsabilité est lié positivement au soutien à l'autonomie rapporté (comportements de l'enseignant visant à identifier, nourrir et développer les ressources motivationnelles internes aux élèves) et à la structuration rapportée (quantité et clarté des informations fournies par l'enseignant quant à ses attentes et à la façon dont les élèves sont censés atteindre de manière effective les résultats désirés). Au contraire, il est lié négativement au contrôle rapporté (comportements de l'enseignant qu'on pourrait qualifier de coercitifs, visant à imposer aux élèves une façon spécifique de penser, de se percevoir et de se comporter) et au laisser-faire rapporté (messages de l'enseignant confus, contradictoires; consignes et attentes communiquées de manière peu claire; manque d'information sur la façon dont les élèves sont censés atteindre les objectifs fixés) (Girardet et Berger, 2016). Toutes les associations entre objets de responsabilité et styles d'enseignement sont d'ampleur moyenne. Ainsi, les enseignants qui se sentent le plus responsables de la qualité de leur travail rapportent des styles de gestion de classe démontrés par la recherche comme étant favorables pour l'engagement des apprenants.

Les résultats des études, encore peu nombreuses, indiquent globalement que le sentiment de responsabilité est un concept pertinent pour comprendre les aspects psychologiques du fonctionnement des enseignants.

L'échelle Teacher Responsibility Scale de Lauermann et Karabenick

À partir d'une revue de la littérature sur les mesures de responsabilité des enseignants, Lauermann et Karabenick (2013) ont conçu une échelle mesurant le sentiment de responsabilité des enseignants en se concentrant sur différentes facettes importantes de leur travail. Celles-ci étaient au départ au nombre de cinq, soit la responsabilité pour :

- 1) la motivation des apprenants ;
- 2) la réussite des apprenants ;
- 3) la confiance en soi des apprenants ;
- 4) les relations positives avec les apprenants ; et
- 5) la qualité de l'enseignement.

Afin de distinguer la responsabilité attribuée à l'enseignant de manière externe (par la loi ou par l'établissement) de la responsabilité personnelle, ces auteurs ont opté pour un degré de spécificité moyen et pour l'utilisation du conditionnel (« *je me sentirais personnellement responsable si...* ») avec une valence négative (« *...un de mes étudiants n'était pas intéressé par la matière que j'enseigne* »). À partir de ces considérations, ils ont rédigé 21 items portant sur les 5 facettes citées ci-dessus, que les enseignants devaient évaluer sur une échelle allant de 0 (*pas du tout responsable*) à 100 (*totalement responsable*).

Les questionnaires ont été passés à un échantillon de 315 enseignants en formation dans une université allemande. Une analyse factorielle exploratoire a conduit à l'élimination de 8 items problématiques et de 1 facteur – le sentiment de responsabilité pour la confiance en soi des apprenants. L'échelle finale était alors constituée de 3 items pour chacune des facettes (*responsabilité pour la motivation des apprenants*, *responsabilité pour les relations avec les apprenants* et *responsabilité pour la qualité de l'enseignement*) et de 4 items pour la facette *responsabilité pour la réussite des apprenants*, pour un total de 13 items.

La structure factorielle obtenue dans la première étude a été testée dans une seconde étude, dont l'échantillon était constitué de 412 enseignants aux États-Unis d'Amérique. Les items étaient identiques à ceux retenus grâce aux résultats de l'étude précédente (après traduction de l'allemand à l'anglais par une procédure de retraduction [*forward-backward*]), sauf pour un item : « *Je me sentirais personnellement responsable si un apprenant échouait à mon cours* ». Cet item a été exclu, car il n'est pas possible d'échouer dans certaines classes. De plus, les critères de réussite diffèrent selon les écoles. Les indices d'adéquation de l'analyse factorielle confirmatoire de la structure en quatre facteurs ont montré que le modèle s'adaptait très bien aux données (Lauermann et Karabenick, 2013). Ainsi, la version finale de l'échelle, soit celle qui a été traduite en français, est constituée de quatre dimensions, chacune mesurée par trois items.

La traduction en français

La *Teacher Responsibility Scale* a été traduite selon une procédure de retraduction (Behling et Law, 2000). Quatre personnes ont été impliquées dans le processus de traduction. Dans un premier temps, le second auteur de cet article, familier des travaux des auteurs de l'échelle, a procédé à une traduction de l'anglais au français (*forward*). Ensuite, la version française a été retraduite en anglais (*backward*) par une personne bilingue, experte dans le domaine de l'éducation. Les différences entre les formulations originales et les formulations retraduites du français ont été examinées par Lauermann et Karabenick, et ces différences ont été discutées avec le second auteur de cet article. Seules des différences mineures ont été relevées et résolues. Ce processus a permis de s'assurer de la qualité de la traduction. La validité de cette version française a été examinée par deux études qui font l'objet de cet article.

Procédure de validation de la version française

Deux études ont été conduites afin d'évaluer la validité de construit de la traduction de l'échelle. La validité de construit a été conceptualisée selon les travaux de Messick (1989), que reprennent Bertrand et Blais (2004). Il s'agit, d'une part, de confirmer les dimensions composant l'instrument traduit en français par l'utilisation d'analyses dimensionnelles. Ces analyses sont réalisées non seulement sur deux échantillons d'enseignants, mais aussi sur des sous-échantillons distingués selon les caractéristiques des enseignants (sexe, expérience) et selon le mode de passation de

l'instrument. D'autre part, il s'agit de situer le concept de sentiment de responsabilité personnelle par rapport à d'autres concepts partageant des similitudes, mais théoriquement différents, ce qui est réalisé en examinant la validité divergente par rapport au sentiment d'efficacité personnelle à enseigner³. À cela s'ajoute une question de validité interne de l'échelle: il est pertinent d'examiner dans quelle mesure les scores de l'Échelle du sentiment de responsabilité sont liés à la désirabilité sociale, soit la tendance à «déformer plus ou moins intentionnellement les descriptions de soi dans une direction jugée favorable (Gough, 1952) en exagérant ses qualités et/ou en minimisant ses défauts» (Juhel et Rouxel, 2005, p. 60).

Ainsi, la première étude visait à tester la validité de construit par des analyses factorielles confirmatoires et par des analyses corrélationnelles avec une mesure de la désirabilité sociale et une mesure du sentiment d'efficacité personnelle. La seconde étude visait à évaluer l'invariance factorielle de l'échelle en fonction de plusieurs facteurs: le sexe, le degré d'expérience d'enseignement et le type de support utilisé pour la passation du questionnaire (papier ou en ligne).

Toutes les analyses factorielles ont été réalisées à l'aide du logiciel Mplus version 7, avec la procédure Maximum Likelihood Robust (MLR), afin de tenir compte des distributions anormales. Les corrélations ont quant à elles été calculées avec SPSS Statistics version 23.

Étude 1 : analyse dimensionnelle et validité convergente/divergente

Cette étude poursuivait les objectifs suivants: premièrement, il s'agissait d'évaluer la validité interne de la traduction sur le plan factoriel. Pour ce faire, des analyses factorielles confirmatoires ont été réalisées ainsi qu'une évaluation de la fidélité. Deuxièmement, les corrélations entre les scores de l'Échelle du sentiment de responsabilité et d'autres scores d'échelles ont été calculées: la tendance à la désirabilité sociale (validité interne) et le sentiment d'efficacité personnelle (validité divergente).

Méthodologie

Participants

L'échantillon était constitué de 154 enseignants de la formation professionnelle en formation à l'enseignement en Suisse romande. Parmi eux, 58 étaient des femmes et 94 étaient des hommes (deux données manquantes).

L'âge moyen des participants était de 40 ans et 2 mois (écart-type [ET] = 6 ans et 11 mois; étendue de 25 ans à 57 ans). Leur nombre d'années d'expérience avant la formation variait entre 1 et 29 années, avec une médiane de 3 ans et une moyenne de 4 ans et 7 mois (ET = 4 ans et 6 mois).

Instruments

Sentiment de responsabilité: La version traduite en français de l'échelle *Teacher Responsibility Scale* (Lauermann et Karabenick, 2013) décrite ci-dessus a été utilisée (voir Annexe pour l'échelle complète). Les participants devaient évaluer dans quelle mesure ils se sentaient personnellement responsables d'avoir dû anticiper différentes situations, sur une échelle de 0 (*pas du tout responsable*) à 100 (*totalelement responsable*). Il y avait 3 items par objet de responsabilité (*responsabilité pour la motivation des apprenants*, *responsabilité pour les relations avec les apprenants*, *responsabilité pour la qualité de l'enseignement* et *responsabilité pour la réussite des apprenants*), pour un total de 12 items.

Sentiment d'efficacité personnelle: Une traduction et adaptation de la *Ohio State Teacher Efficacy Scale* (Tschannen-Moran et Woolfolk Hoy, 2001) a été utilisée pour évaluer le sentiment d'efficacité personnelle de l'enseignant. Chaque item commençait par « *Considérant mes récentes réalisations, ressources et opportunités en tant qu'enseignant-e...* ». Les répondants devaient alors évaluer chaque item sur une échelle allant de 1 (*en total désaccord*) à 6 (*tout à fait d'accord*). Le sentiment d'efficacité personnelle était évalué pour les trois objets suivants, chacun par quatre items : 1) gérer la classe, 2) engager les apprenants (traduits par Dumay et Galand, 2012) et 3) planifier les cours (développés pour cette étude; exemple d'item : « *...planifier des activités qui couvrent précisément le temps de cours* »). Pour cette étude, nous avons utilisé le score de sentiment d'efficacité générale, calculé par la moyenne des 12 items (α de Cronbach = 0,88).

Tendance à la désirabilité sociale: Des versions courtes des échelles conçues par Tournois, Mesnil et Kop (2000) ont été utilisées afin d'évaluer deux formes de désirabilité sociale : les tendances à l'autoduperie (6 items; p. ex. « *Je suis toujours optimiste* »; α de Cronbach = 0,67) et à l'hétéroduperie (6 items; p. ex. « *Je dis toujours des choses favorables sur les autres* »; α de Cronbach = 0,71). L'autoduperie ou illusion sur soi représente une façon optimiste de s'autoévaluer qui contribue à préserver et amplifier une image de soi positive (Paulhus, 1984). L'hétéroduperie ou

gestion de l'impression représente une façon de se présenter en donnant aux autres une image de soi positive de manière intentionnelle (Paulhus, 1984)⁴.

Procédure de passation de l'instrument

Les instruments décrits ci-dessus faisaient partie d'une étude plus vaste, conduite par questionnaires papier auprès des enseignants en formation, durant une période de cours de 45 minutes en 2012. Tous les participants étaient informés de la nature et de l'objectif de la recherche ainsi que rendus attentifs au fait qu'ils pouvaient se retirer de l'étude à tout moment. La participation n'était ni obligatoire ni rémunérée. Tous les participants présents le jour de la passation ont accepté de remplir le questionnaire.

Plan d'analyse

Tout d'abord, nous avons testé, à l'aide d'une analyse factorielle confirmatoire (AFC), un modèle à quatre facteurs, soit les quatre facteurs de responsabilité décrits ci-dessus : *responsabilité pour la motivation des apprenants, pour les relations avec les apprenants, pour la qualité de l'enseignement et pour la réussite des apprenants*. Ce modèle a été mis en concurrence avec deux modèles alternatifs : un modèle à un seul facteur de responsabilité générale ainsi qu'un modèle hiérarchique avec un facteur de second ordre (responsabilité générale) et quatre facteurs de premier ordre. Si les données s'ajustent mieux au modèle à un seul facteur de responsabilité générale, cela signifie que les réponses des participants ne permettent pas de distinguer quatre objets de responsabilité et, ainsi, que le sentiment de responsabilité est un concept unidimensionnel. Si les données s'ajustent mieux au modèle hiérarchique, cela signifie que le sentiment de responsabilité personnelle peut être conçu tant comme multidimensionnel que résumable par un score général. Les indices d'adéquation ont été comparés aux critères pour l'évaluation des indices d'adéquation des modélisations par équation structurelle (*structural equation modeling*; SEM) formulés par Schermelleh-Engel, Moosbrugger et Müller (2003), qui figurent dans le tableau 1. Afin de tester la supériorité d'un modèle sur les autres, nous avons également calculé la différence de khi carré (χ^2) entre les modèles. Nous avons utilisé ces différentes informations afin d'aboutir à un jugement concernant le meilleur modèle.

Tableau 1
*Critères pour l'évaluation des indices d'adéquation des SEM
 selon Schermelleh-Engel et al. (2003)*

Test ou indice	Valeur suggérant une adéquation acceptable du modèle	Valeur suggérant une bonne adéquation du modèle
valeur-p χ^2	$0,01 \leq p \leq 0,05$	$0,05 < p \leq 1,00$
RMSEA	$0,05 < \text{RMSEA} \leq 0,08$	$0 \leq \text{RMSEA} \leq 0,05$
CFI	$0,95 \leq \text{CFI} \leq 0,97$	$0,97 \leq \text{CFI} \leq 1,00$
SRMR	$0,05 < \text{SRMR} \leq 0,10$	$0 \leq \text{SRMR} \leq 0,05$

Note. RMSEA = erreur quadratique moyenne de l'approximation (*root mean square error of approximation*); CFI = indice d'ajustement comparatif (*comparative fit index*); SRMR = valeur moyenne quadratique pondérée (*standardized root mean square residual*).

Résultats

Analyses dimensionnelles

Le tableau 2 présente les résultats des différentes AFC effectuées.

Tableau 2
Adéquation des modèles d'AFC

Modèle	χ^2	ddl	χ^2/ddl	valeur-p	CFI	RMSEA	SRMR
1 facteur	236,525	54	4,38	< 0,001	0,642	0,148	0,105
4 facteurs	60,928	48	1,27	0,100	0,975	0,042	0,056
4 facteurs de premier ordre et 1 facteur général	72,511	50	1,45	0,020	0,956	0,054	0,062

Note. $n = 154$; ddl = degrés de liberté; CFI = indice d'ajustement comparatif (*comparative fit index*); RMSEA = erreur quadratique moyenne de l'approximation (*root mean square error of approximation*); SRMR = valeur moyenne quadratique pondérée (*standardized root mean square residual*).

D'après les indices d'adéquation au modèle, le modèle à quatre facteurs semble être le plus adéquat. Cela est confirmé par les analyses de différences de χ^2 ajustées selon la correction de Satorra-Bentler (Satorra et Bentler, 2001) : le modèle à quatre facteurs s'ajuste mieux que le modèle à un facteur [$\Delta\chi^2(6) = 96,9988$; $p < 0,001$] et le modèle hiérarchique est plus adéquat que le modèle à un facteur [$\Delta\chi^2(2) = 11,213$; $p = 0,004$]. En d'autres termes, les quatre objets du sentiment de responsabilité personnelle des enseignants représentent autant de dimensions distinctes de ce sentiment. Ainsi, il est judicieux de traiter le sentiment de responsabilité comme un concept multidimensionnel.

Le tableau 3 indique, pour chacun des 12 items, le libellé, la moyenne, l'écart-type, la saturation et, pour chacun des 4 facteurs, le coefficient de détermination des scores factoriels (représentant la fidélité de la variable latente⁵) ainsi que l'alpha de Cronbach (représentant la cohérence interne des scores d'échelle).

Le tableau 4 indique le sens et l'amplitude des coefficients de corrélation entre les différents objets de responsabilité. Ces corrélations sont toutes significatives et de taille moyenne. Le lien le plus fort est celui entre la responsabilité concernant la réussite des apprenants et celle concernant leur motivation.

Validité divergente et validité interne

Afin d'évaluer la validité divergente de l'Échelle du sentiment de responsabilité des enseignants, les corrélations avec le score à l'Échelle du sentiment d'efficacité personnelle ont été calculées. Les résultats figurent dans le tableau 5. Les corrélations sont toutes significatives et de taille moyenne. Bien que corrélés significativement, les deux concepts sont empiriquement distincts, et ce, au travers des divers objets de responsabilité.

Tableau 3
Statistiques descriptives des items et résultats de l'AFC en quatre facteurs

Item	Libellé	M	ET	Sat.	DF	α
Responsabilité pour la motivation des apprenants					0,90	0,79
rsm1	<i>l'un de mes apprenants ne s'intéressait pas à la matière que j'enseigne.</i>	47,66	22,54	0,67		
rsm2	<i>l'un de mes apprenants n'aimait pas la matière que j'enseigne.</i>	29,05	22,84	0,79		
rsm3	<i>l'un de mes apprenants n'accordait pas de valeur à la matière que j'enseigne.</i>	36,17	24,02	0,77		
Responsabilité pour les relations avec les apprenants					0,93	0,78
rrs1	<i>l'un de mes apprenants pensait qu'il ne pourrait pas compter sur moi s'il avait besoin d'aide.</i>	75,97	25,26	0,63		
rrs2	<i>l'un de mes apprenants ne pensait pas pouvoir me confier les problèmes qu'il rencontre à l'école ou en dehors de l'école.</i>	58,00	28,85	0,69		
rrs3	<i>l'un de mes apprenants ne croyait pas que je me soucie vraiment de lui.</i>	65,66	26,99	0,89		
Responsabilité pour la qualité de l'enseignement					0,92	0,71
rte1	<i>un cours que j'ai enseigné ne reflétait pas au mieux mes habiletés d'enseignant.</i>	69,78	25,07	0,57		
rte2	<i>un cours que j'ai enseigné n'avait pas favorisé l'apprentissage des apprenants autant que ce que j'aurais pu faire.</i>	73,66	18,57	0,62		
rte3	<i>un cours que j'ai enseigné n'était pas aussi stimulant pour les apprenants que ce que j'aurais pu faire.</i>	73,75	19,86	0,89		
Responsabilité pour la réussite des apprenants					0,93	0,81
rsa1	<i>l'un de mes apprenants ne réalisait pas de grands progrès durant l'année scolaire.</i>	54,07	20,87	0,84		
rsa2	<i>l'un de mes apprenants ne parvenait pas à apprendre la matière demandée au programme.</i>	56,07	19,61	0,63		
rsa3	<i>l'un de mes apprenants ne réussissait pas bien.</i>	49,78	20,78	0,84		

Note. $n = 154$; M = moyenne; ET = écart-type; Sat = saturation factorielle; DF = coefficient de détermination du facteur; α = alpha de Cronbach de l'échelle. Toutes les saturations sont significatives à $p < 0,001$.

Tableau 4
Corrélations d'ordre zéro entre les objets du sentiment de responsabilité

Objet de responsabilité	Relations avec les apprenants	Qualité de l'enseignement	Réussite des apprenants
Motivation des apprenants	0,40**	0,22*	0,47**
Relations avec les apprenants		0,40**	0,48**
Qualité de l'enseignement			0,30**

Note. $n = 154$; * = corrélation significative à $p < 0,05$; ** = corrélation significative à $p < 0,01$ (bilatérale).

Tableau 5
Corrélations d'ordre zéro entre les objets du sentiment de responsabilité, le sentiment d'efficacité personnelle et les tendances à la désirabilité sociale

Objet de responsabilité	Sentiment d'efficacité personnelle	Désirabilité sociale	
		Autoduperie	Hétéroduperie
Motivation des apprenants	0,21**	0,06	0,13
Relations avec les apprenants	0,25**	-0,07	0,24**
Qualité de l'enseignement	0,32**	0,09	0,08
Réussite des apprenants	0,16*	-0,16*	0,11

Note. $n = 154$; * = corrélation significative à $p < 0,05$; ** = corrélation significative à $p < 0,01$ (bilatérale).

Concernant la validité interne, les corrélations des scores de l'Échelle du sentiment de responsabilité des enseignants avec les deux dimensions de la désirabilité sociale révèlent que la responsabilité pour les relations avec les apprenants est liée, dans une mesure modeste, à la tendance à l'hétéroduperie. La responsabilité pour la réussite des apprenants est quant à elle liée à la tendance à l'autoduperie.

Discussion de l'étude 1

Cette étude révèle que :

- la version française de l'Échelle du sentiment de responsabilité présente une validité factorielle adéquate en ce qu'elle reflète les résultats de la version originale ;

- b) la fidélité des scores produits par les quatre sous-échelles, que ce soit sous la forme des coefficients de détermination des facteurs ou de la cohérence interne, est adéquate;
- c) la validité interne, comprise comme le degré d'association entre les scores aux sous-échelles et les deux facettes de la tendance à la désirabilité sociale, est satisfaisante.

Les associations significatives entre sentiment de responsabilité et désirabilité sociale peuvent être expliquées de la façon suivante. L'hétéroduperie vise à entretenir des relations sociales agréables, ce qui revient à agir de manière conforme aux normes sociales et aux attentes vis-à-vis de son rôle. Ainsi, les enseignants qui se sentent responsables d'entretenir des relations de qualité avec les apprenants sont également ceux qui disent agir de manière conforme à leur rôle. La corrélation entre la responsabilité pour la réussite des apprenants et la tendance à l'autoduperie, bien que faible, indique que les enseignants qui tendent à s'évaluer de manière optimiste disent assumer une moins forte responsabilité envers l'échec des apprenants. En somme, cela pourrait être interprété comme un mécanisme de défense consistant à attribuer de manière externe les échecs des apprenants afin qu'ils ne détériorent pas l'image de soi de l'enseignant. Finalement, les associations entre le sentiment de responsabilité et le sentiment d'efficacité personnelle reflètent les résultats des auteurs de l'échelle, à savoir que les enseignants qui se sentent le plus largement responsables pensent être les plus compétents (Lauermann et Karabenick, 2013).

Étude 2: analyse d'invariance de l'échelle

L'objectif de cette seconde étude était de tester l'hypothèse selon laquelle la structure factorielle de l'Échelle du sentiment de responsabilité des enseignants est similaire selon le sexe (hommes ou femmes), selon les années d'expérience d'enseignement (novices ou experts) et selon le type de support utilisé pour la passation du questionnaire (papier ou en ligne). Il est important de tester l'invariance d'une mesure, car un même instrument devrait mesurer le même concept chez des individus différents, même si ceux-ci diffèrent selon d'autres variables (Millsap, 2011). Dans le cas présent, il est souhaitable que deux individus ayant théoriquement un même niveau de sentiment de responsabilité obtiennent des scores similaires

aux différentes échelles de celui-ci, qu'ils aient passé le questionnaire en format papier ou en ligne, qu'ils soient un homme ou une femme et qu'ils soient experts ou novices dans l'enseignement.

Méthodologie

Participants

L'échantillon était constitué de 533 enseignants exerçant dans le contexte de la formation professionnelle en Suisse romande. Au moment de la passation, 116 d'entre eux étaient en formation en vue d'obtenir un diplôme d'enseignant en écoles professionnelles; ils ont rempli les questionnaires en format papier. Les 417 autres participants les ont remplis en ligne.

L'âge moyen des participants était de 45 ans et 11 mois ($ET = 9$ ans et 5 mois; étendue de 24 ans à 65 ans). Leur nombre d'années d'expérience variait entre 0 et 45, avec une moyenne de 12 ans et 11 mois ($ET = 10$ ans et 9 mois). Parmi eux, 37,7% sont des femmes et 61,5% sont des hommes; 0,8% n'a pas indiqué son sexe.

Instrument

Sentiment de responsabilité: L'Échelle du sentiment de responsabilité utilisée dans cette seconde étude est la même que dans la première étude décrite ci-dessus, si ce n'est que les enseignants devaient évaluer dans quelle mesure ils se sentaient personnellement responsables de *n'avoir pas su anticiper* au lieu d'*avoir dû anticiper*. Ce changement de formulation s'est avéré nécessaire, car la première formulation était syntaxiquement incorrecte. Lors de la passation en format en ligne, le répondant devait positionner un curseur sur une échelle allant de 0 à 100 pour évaluer chaque item.

Procédure de passation de l'instrument

Le questionnaire en format papier a été distribué lors d'un cours donné dans un institut de formation à l'enseignement en septembre 2013 à l'ensemble des enseignants en formation en début de première année, incluant des enseignants suivant une validation des acquis de l'expérience. La passation n'était pas obligatoire, mais tous les enseignants présents ont accepté d'y participer. Le questionnaire était composé de différentes échelles, dont l'Échelle du sentiment de responsabilité des enseignants.

Le questionnaire en ligne, créé avec le logiciel d'enquête *LimeSurvey*, a été envoyé par courriel à toutes les directions des établissements de formation professionnelle de Suisse romande (N = 91), avec la demande de le transmettre aux enseignants exerçant dans les établissements qu'ils dirigeaient. Au total, 49 directions (53,85 %) ont confirmé avoir transmis le questionnaire à leurs enseignants. La participation était facultative et anonyme. Le questionnaire était composé de différentes échelles, dont l'Échelle du sentiment de responsabilité des enseignants.

Plan d'analyse

Premièrement, nous avons réalisé une AFC sur l'ensemble de notre échantillon afin de vérifier que sa structure factorielle était la même que dans la précédente étude.

Deuxièmement, nous avons calculé des modèles de base pour chaque paire de groupes, dont l'équivalence factorielle sera ensuite testée.

Le tableau 6 récapitule le nombre de participants dans chaque groupe selon les variables sexe, années d'expérience (divisées en deux groupes : novices, soit moins de 10 ans d'expérience, et experts, soit 10 ans d'expérience et plus) et type de support (papier ou en ligne). Le total de 533 participants varie pour le sexe et pour les années d'expérience à cause des données manquantes.

Tableau 6
Récapitulatif du nombre de participants dans chaque paire de comparaisons

Variable	Modalité 1	Modalité 2	Total
Sexe	201 (38 %) femmes	328 (62 %) hommes	529
Années d'expérience	251 (49 %) novices	259 (51 %) experts	510
Type de support	116 (22 %) papier	417 (78 %) en ligne	533

Concernant le nombre de participants dans chaque groupe, nous remarquons qu'il y a presque autant d'enseignants novices que d'experts. Par contre, notre échantillon comporte plus d'hommes que de femmes : cela est fréquemment le cas parmi les enseignants de la formation professionnelle en Suisse. Le nombre de participants ayant passé le questionnaire en format papier est beaucoup plus faible que le nombre de participants ayant passé le questionnaire en ligne.

Les différents groupes ne sont pas indépendants. Si la proportion d'hommes et de femmes est pratiquement similaire dans chaque groupe, il n'en va pas de même pour la proportion de novices et d'experts ayant passé le questionnaire en format papier ou en ligne : il y a 87,9% de novices parmi les participants en format papier et seulement 37,8% de novices parmi les participants en format en ligne. Cela est dû au fait que les participants en format papier sont des enseignants en formation, bien que certains aient déjà plusieurs années d'expérience d'enseignement. Nous sommes conscients de ce biais et serons donc prudents dans l'interprétation des résultats.

Troisièmement, nous avons effectué une série d'AFC multigroupes sur chaque paire de groupes, selon les variables sexe, années d'expérience et type de support. Selon les recommandations de Byrne (2012)⁶, nous avons appliqué successivement, pour chaque paire :

- 1) un modèle d'invariance configurale dans lequel les facteurs sont mesurés par les mêmes indicateurs dans les deux groupes afin de vérifier qu'il y a le même nombre de facteurs ainsi que le même nombre d'items par facteur entre les deux groupes. Aucune contrainte d'égalité n'est imposée ;
- 2) un modèle d'invariance métrique afin de vérifier que les saturations ne diffèrent pas entre les deux groupes ;
- 3) un modèle d'invariance structurelle afin de vérifier que les variances et covariances des facteurs ne diffèrent pas entre les deux groupes ;
- 4) un modèle d'invariance scalaire afin de vérifier que les seuils ne diffèrent pas entre les deux groupes.

Chaque modèle a été comparé au modèle précédent en utilisant la différence de χ^2 ajusté. Dans le cas où l'adéquation des modèles estimés n'était pas satisfaisante, des modifications ont été effectuées, indiquant une invariance partielle.

Résultats

Modèle factoriel global et modèles factoriels de base des groupes examinés

L'AFC sur le deuxième échantillon au complet montre que les données s'ajustent de manière satisfaisante à une solution factorielle identique à celle choisie à la suite de la première étude (maximum likelihood : $\chi^2_{(48)} = 80,598$ et $p = 0,002$; CFI = 0,974; RMSEA = 0,046; SRMR = 0,036). Le tableau 7 indique les indices d'adéquation des modèles de base pour chaque groupe étudié.

Tableau 7
Adéquation des modèles de base par groupe

Groupe	N	ddl	χ^2	valeur-p	CFI	RMSEA	SRMR
Femmes	201	48	84,40	0,001	0,959	0,061	0,047
Hommes	328	48	68,89	0,027	0,984	0,036	0,037
Novices	251	48	72,03	0,014	0,976	0,045	0,045
Experts	259	48	74,46	0,009	0,975	0,046	0,038
Papier	116	48	48,68	0,446	0,998	0,011	0,048
En ligne	417	48	102,83	0,000	0,967	0,052	0,039

Note. ddl = degrés de liberté; CFI = indice d'ajustement comparatif (*comparative fit index*); RMSEA = erreur quadratique moyenne de l'approximation (*root mean square error of approximation*); SRMR = valeur moyenne quadratique pondérée (*standardized root mean square residual*).

Les indices d'adéquation sont satisfaisants (les indices de CFI étant supérieurs à 0,95, et les indices RMSEA et SRMR étant inférieurs à 0,05) et aucune modification importante n'est suggérée par le logiciel Mplus. Autrement dit, il n'existe pas de saturations croisées ni de covariances élevées et théoriquement justifiables entre les items. Nous pouvons donc réaliser les AFC multigroupes.

AFC multigroupes

La suite du texte présente les trois comparaisons multigroupes selon le sexe, les années d'expérience et le type de support. Pour chaque comparaison, l'invariance a été testée de manière séquentielle sur les plans configural, métrique, structurel et scalaire.

Invariance de la mesure entre les hommes et les femmes : Le tableau 8 montre les indices d'adéquation des différents modèles d'AFC multigroupes pour la variable « sexe ».

Tableau 8
Adéquation des modèles d'AFC multigroupes par sexe

Modèle	N ^b re param.	ddl	χ^2	$\Delta \chi^2$	Δ ddl	Δ valeur-p	CFI	RMSEA	SRMR
M ₁ : invariance configurale	84	96	153,28	—	—	—	0,974	0,047	0,041
M ₂ : invariance métrique	76	104	159,00	5,72	8	0,681	0,975	0,045	0,044
M ₃ : invariance structurelle	66	114	169,34	10,34	10	0,495	0,975	0,043	0,053
M ₄ : invariance scalaire	54	126	189,70	18,50	12	0,101	0,971	0,044	0,056

Note. $n = 529$; N^bre param. = nombre de paramètres; ddl = degrés de liberté; CFI = indice d'ajustement comparatif (*comparative fit index*); RMSEA = erreur quadratique moyenne de l'approximation (*root mean square error of approximation*); SRMR = valeur moyenne quadratique pondérée (*standardized root mean square residual*). Les différences de khi carré sont calculées entre le modèle n-1 et le modèle n. Les différences ont été ajustées selon la correction proposée par Satorra et Bentler (2001).

Étant donné que a) les différences de χ^2 ne sont pas significatives entre le modèle le plus parcimonieux et le précédent, et que b) les indices d'adéquation des modèles sont satisfaisants, nous pouvons conclure que l'invariance est totale sur les plans configural (les quatre facteurs sont mesurés par les mêmes indicateurs dans les deux groupes), métrique (les saturations sont invariantes dans les deux groupes), structurel (les variances et covariances des facteurs sont similaires dans les deux groupes) et scalaire (les seuils sont invariants dans les deux groupes).

Invariance de la mesure entre les enseignants novices et les enseignants experts : Le tableau 9 montre les indices d'adéquation des différents modèles d'AFC multigroupes pour la variable « expérience d'enseignement ».

Tableau 9
*Adéquation des modèles d'AFC multigroupes
 par expérience dans l'enseignement*

Modèle	N ^{bre} param.	ddl	χ^2	$\Delta \chi^2$	Δ ddl	Δ valeur-p	CFI	RMSEA	SRMR
M ₁ : invariance configurale	84	96	146,33	—	—	—	0,975	0,045	0,042
M ₂ : invariance métrique	76	104	150,74	4,28	8	0,831	0,977	0,042	0,045
M ₃ : invariance structurelle	66	114	155,18	4,11	10	0,942	0,980	0,038	0,053
M ₄ : invariance scalaire	54	126	174,66	19,21	12	0,083	0,976	0,039	0,061

Note. $n = 510$; N^{bre} param. = nombre de paramètres; ddl = degrés de liberté; CFI = indice d'ajustement comparatif (*comparative fit index*); RMSEA = erreur quadratique moyenne de l'approximation (*root mean square error of approximation*); SRMR = valeur moyenne quadratique pondérée (*standardized root mean square residual*). Les différences de khi carré sont calculées entre le modèle $n-1$ et le modèle n . Les différences ont été ajustées selon la correction proposée par Satorra et Bentler (2001).

Tout comme pour la comparaison selon le sexe, les tests d'invariance selon les années d'expérience nous indiquent que l'invariance est totale sur les plans configural, métrique, structurel et scalaire. En effet, les différences de χ^2 ne sont pas significatives entre le modèle le plus parcimonieux et le précédent, et les indices d'adéquation des modèles sont satisfaisants.

Invariance de la mesure entre les questionnaires en format papier et en ligne : Le tableau 10 montre les indices d'adéquation des différents modèles d'AFC multigroupes pour la variable « type de support ».

Dans la même logique que les tests précédents, l'invariance est totale sur les plans configural, métrique et structurel. Par contre, pour l'invariance scalaire, nous ne pouvons statuer que sur une invariance partielle des seuils. En effet, étant donné que le modèle M_{4a} s'est avéré significativement moins adéquat que le modèle M₃, les indices de modification donnés par le logiciel ont été consultés. Sur cette base, les seuils de trois items et de deux facteurs latents ont été libérés de manière séquentielle, en allant de l'item dont la modification améliorerait le modèle de façon la plus conséquente au facteur dont la modification améliorerait le modèle de façon la moins conséquente. Une fois que les seuils de ces trois items et de ces deux facteurs latents ne sont plus contraints à être égaux dans les différents

groupes (modèle M_{4f}), la différence entre ce dernier modèle et le modèle M₃ n'est plus significative et les indices d'adéquation au modèle sont satisfaisants.

Tableau 10
Adéquation des modèles d'AFC multigroupes par type de support

Modèle	N ^{bre} param.	ddl	χ^2	$\Delta \chi^2$	Δ ddl	Δ valeur-p	CFI	RMSEA	SRMR
M ₁ : invariance configurale	84	96	148,54	—	—	—	0,974	0,045	0,041
M ₂ : invariance métrique	76	104	155,69	6,78	8	0,561	0,975	0,043	0,044
M ₃ : invariance structurelle	66	114	163,11	6,55	10	0,767	0,976	0,040	0,049
M _{4a} : invariance scalaire	54	126	210,22	50,13	12	0,000	0,959	0,050	0,075
M _{4b} : invariance scalaire sauf rrs3	55	125	203,63	43,86	11	0,000	0,962	0,049	0,076
M _{4c} : invariance scalaire sauf rrs3, rte3	56	124	196,88	36,52	10	0,000	0,965	0,047	0,076
M _{4d} : invariance scalaire sauf rrs3, rte3, rte2	57	123	196,00	32,89	9	0,000	0,965	0,047	0,075
M _{4e} : invariance scalaire sauf rrs3, rte3, rte2 rsa	58	122	189,46	29,15	8	0,000	0,967	0,046	0,067
M _{4f} : invariance scalaire sauf rrs3, rte3, rte2 rsa, rsm	59	121	172,43	9,97	7	0,191	0,975	0,040	0,055

Note. $n = 533$; N^{bre} param. = nombre de paramètres; ddl = degrés de liberté; CFI = indice d'ajustement comparatif (*comparative fit index*); RMSEA = erreur quadratique moyenne de l'approximation (*root mean square error of approximation*); SRMR = valeur moyenne quadratique pondérée (*standardized root mean square residual*). Les différences de khi carré sont calculées entre le modèle n-1 et le modèle n. Les différences ont été ajustées selon la correction proposée par Satorra et Bentler (2001).

Discussion de l'étude 2

Avec un échantillon différent et plus grand d'enseignants, nous avons répliqué la structure factorielle obtenue lors de l'étude 1. Cela confirme que la version française de l'Échelle du sentiment de responsabilité est constituée de quatre facteurs, chacun mesuré par trois items.

Les AFC multigroupes ont montré que la mesure est invariante selon le sexe et les années d'expérience. Ainsi, dans chaque groupe (hommes, femmes, enseignants novices et enseignants experts), les quatre facteurs sont mesurés par les mêmes items. De plus, les saturations, covariances et variances des facteurs ainsi que les seuils sont similaires. Les différences en matière de moyennes des facteurs qu'on pourrait observer entre ces groupes seraient ainsi dues à un sentiment de responsabilité effectivement différent, et non à un fonctionnement différentiel de la mesure.

En ce qui concerne les AFC multigroupes testant l'invariance entre les enseignants qui ont rempli le questionnaire papier et ceux qui ont rempli le questionnaire en ligne, nous concluons à des invariances configurale, métrique et structurelle. Par contre, les seuils des facteurs diffèrent significativement entre les deux groupes pour les items rrs3 («*l'un de mes apprenants ne croyait pas que je me soucie vraiment de lui*»), rte3 («*un cours que j'ai enseigné n'était pas aussi stimulant pour les apprenants que ce que j'aurais pu faire*»), rte2 («*un cours que j'ai enseigné n'avait pas favorisé l'apprentissage des apprenants autant que ce que j'aurais pu faire*») et pour les facteurs *responsabilité pour la réussite des apprenants* et *responsabilité pour la motivation des apprenants*. Cela peut être expliqué par le fait que, lors de la passation papier, les enseignants devaient inscrire un chiffre entre 0 et 100, tandis que, lors de la passation en ligne, les enseignants devaient positionner un curseur entre 0 et 100. Par défaut, le curseur était positionné à 0. Des analyses de contrôle ont été effectuées en amont pour éviter de considérer un 0 alors qu'il s'agit en réalité d'une donnée manquante⁷. Cela dit, il semblerait que les enseignants avaient plus fortement tendance à laisser le curseur sur 0 dans la version en ligne, alors qu'il était plus rare que les enseignants dans la version papier inscrivent «0». De plus, dans la version papier, les enseignants ont presque toujours arrondi à la dizaine près, ce qui n'a pas été le cas pour la version en ligne. Afin d'éviter ce genre de biais de mesure, il ne faudrait pas utiliser un curseur dans la version en ligne, ou trouver un équivalent au curseur dans la version papier, par exemple se positionner sur un segment allant de 0 à 100.

Malgré la disproportion dans les tailles des groupes, les comparaisons se sont révélées valides : les indices d'adéquation obtenus ont été satisfaisants et tous les modèles ont été estimés sans difficulté par le logiciel.

Conclusion

L'objectif de ces études était de valider la version française de l'Échelle du sentiment de responsabilité de Lauermann et Karabenick (2013). Pour ce faire, nous avons réalisé deux études s'appuyant sur des AFC ainsi que sur des mesures de validité externe et interne.

La première étude avait pour but de vérifier la structure factorielle de l'instrument. Un échantillon d'enseignants en formation a répondu à un questionnaire comprenant les items de l'Échelle du sentiment de responsabilité, des items évaluant leur sentiment d'efficacité et des items mesurant leur tendance à la désirabilité sociale. Les analyses factorielles confirmatoires ont montré que la structure en quatre facteurs était la meilleure, tout comme dans la version en anglais. De plus, les coefficients de corrélation significatifs, mais faibles, entre le sentiment d'efficacité générale et les différentes sous-échelles du sentiment de responsabilité montrent qu'il existe un lien entre ces deux concepts, mais qu'ils sont différents. Finalement, nous pouvons conclure que l'Échelle du sentiment de responsabilité n'est pas fortement affectée par la désirabilité sociale.

La seconde étude avait pour but de répliquer la structure factorielle trouvée dans la première étude et de vérifier l'invariance de la mesure au sein de différentes populations d'enseignants. Des analyses factorielles confirmatoires multigroupes ont ainsi été réalisées, en distinguant deux groupes d'enseignants à trois reprises : des enseignants experts et novices ; des enseignants de sexe masculin et de sexe féminin ; et des enseignants qui ont répondu au questionnaire en version papier et en ligne. Nos résultats ont permis d'établir une invariance complète de la mesure lorsque les solutions factorielles ont été comparées selon le sexe et les années d'expérience. Par contre, nous n'avons pu conclure qu'à une invariance partielle dans le cas du type de support du questionnaire.

De manière globale, les deux études soutiennent les qualités psychométriques de la version française de l'Échelle du sentiment de responsabilité. Celle-ci peut ainsi être utilisée en l'état, en restant toutefois prudent lors du choix de support du questionnaire (en ligne ou papier).

Limites

Nous sommes conscients que les résultats des deux études doivent être interprétés en tenant compte de plusieurs limites. Tout d'abord, nous avons réalisé les deux études avec des enseignants exerçant dans un domaine spécifique, soit celui de la formation professionnelle. Ils enseignent à un public d'adolescents et de jeunes adultes, dont certains ont pu éprouver par le passé, ou éprouvent encore aujourd'hui, des difficultés scolaires. Il s'agirait de vérifier la structure factorielle de l'instrument auprès d'autres populations d'enseignants, par exemple de la scolarité obligatoire.

Deuxièmement, l'échelle mesure quatre objets de responsabilité. Il se pourrait que d'autres objets de responsabilité puissent être endossés par les enseignants, selon le contexte éducatif, le contexte culturel ou l'âge des apprenants. Typiquement, les enseignants semblent également se sentir responsables de l'éducation des apprenants lorsque ceux-ci sont encore à l'école obligatoire, comme le montrent les travaux de Périer (2013) sur les nouveaux rôles endossés par les enseignants.

Troisièmement, chaque objet de responsabilité est évalué par seulement trois items, ce qui limite leur envergure. Qui plus est, les formulations de certains items sont en partie redondantes.

Finalement, la qualité des AFC souffre du fait que les groupes ne sont pas complètement indépendants : il y a beaucoup d'enseignants experts parmi les enseignants qui ont passé le questionnaire en ligne, par rapport à ceux qui ont passé le questionnaire papier. Toutefois, le fait d'avoir comparé également les enseignants experts avec les novices nous permet tout de même de conclure, avec prudence, que l'invariance partielle observée ne semble être due qu'au support du questionnaire, et non à l'expérience d'enseignement.

L'Échelle du sentiment de responsabilité validée en français ouvre de nouvelles perspectives de recherche. Elle constituera un outil précieux pour mieux comprendre la psychologie des enseignants et nous aidera à répondre à de nombreuses questions. Pourquoi les enseignants adoptent-ils certaines pratiques d'enseignement plutôt que d'autres ? Quels sont les impacts du sentiment de responsabilité sur les apprenants en matière d'engagement et de réussite ? Est-ce que le sentiment de responsabilité évolue (p. ex., au cours d'une formation à l'enseignement) ?

Réception : 23 novembre 2015

Version finale : 9 janvier 2018

Acceptation : 16 janvier 2018

NOTES

1. Cette méthode d'analyse vise à représenter géométriquement les données (matrices de corrélations) dans un espace euclidien. Elle permet de résumer les informations en les représentant sous une forme immédiatement perceptible (Tournois et Dicks, 1993).
2. Les buts de maîtrise incluent les buts de l'élève visant à l'accroissement de sa compréhension et à l'amélioration de sa performance; une structure de buts de maîtrise offre un climat centré sur l'apprentissage, sur le développement des compétences et sur la maîtrise des savoirs. Les buts de performance incluent les buts de l'élève visant à l'établissement de la supériorité de sa compétence (par rapport aux autres), à l'obtention de jugements favorables ou à l'évitement de jugements négatifs; une structure de buts de performance met l'accent sur les notes et sur la comparaison entre élèves des résultats scolaires, et rend apparentes les différences d'habiletés entre les élèves (Galand, Philippot et Frenay, 2006).
3. Nous avons choisi le sentiment d'efficacité personnelle à enseigner comme critère de validité divergente afin de démontrer que, si le sentiment de responsabilité personnelle envers certains résultats est statistiquement associé à la confiance en ses capacités à obtenir certains résultats ou à les prévenir, il peut tout de même en être empiriquement distingué.
4. Les échelles originales de Tournois et ses collègues (2000) étaient constituées d'un total de 35 items. Nous avons utilisé seulement 12 items, sélectionnés à la suite d'une analyse factorielle exploratoire réalisée sur un échantillon de 631 adolescents et jeunes adultes en formation professionnelle en Suisse (Büchel, Berger et Kipfer, 2011).
5. La détermination du facteur indique la qualité de la mesure du facteur. La valeur de ce coefficient peut se situer entre 0 et 1. Un coefficient de détermination $\geq 0,80$ suggère une haute qualité de la mesure du facteur.
6. La procédure et les termes français sont inspirés de l'article de Rouxel, Briec, Juhel et Le Maner-Idrissi (2013).
7. Pour ce faire, nous avons utilisé la distance de Mahalanobis selon les suggestions de Tabachnick et Fidell (2007). Les participants dont la valeur de Mahalanobis excédait une valeur critique et pour qui figurait un score de 0 à l'un des items ont été éliminés de l'échantillon. Les réponses de 30 participants ont ainsi été éliminées.

RÉFÉRENCES

- Altet, M. (2010). Deux décennies de formation des enseignants dans les IUFM : un processus de formation professionnalisante en héritage. *Recherches en éducation*, 8, 12-23.
- Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. (Sage University paper series on quantitative applications in the social sciences, 07-131). Thousand Oaks, CA: SAGE.
- Berger, J.-L., & Lê Van, K. (2019). Teacher professional identity as multidimensional: Mapping its components and examining their associations with general pedagogical beliefs. *Educational Studies*, 45(2), 163-181. doi: 10.1080/03055698.2018.1446324
- Bertrand, R. et Blais, J.-G. (2004). *Modèles de mesure : l'apport de la théorie des réponses aux items*. Sainte-Foy : Presses de l'Université du Québec.
- Büchel, F. P., Berger, J.-L. et Kipfer, N. (2011). *Questions sur l'apprentissage : un instrument pour le diagnostic pédagogique aux niveaux secondaires I et II*. Lausanne. Lausanne, Suisse : Loisir et Pédagogie.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge.
- Dumay, X., & Galand, B. (2012). The multilevel impact of transformational leadership on teacher commitment: Cognitive and motivational pathways. *British Educational Research Journal*, 38(5), 703-729. doi: 10.1080/01411926.2011.577889
- Fives, H., & Gill, M. G. (Eds.). (2015). *International handbook of research on teachers' beliefs*. New York: Routledge. doi: 10.4324/9780203108437
- Galand, B., Philippot, P. et Frenay, M. (2006). Structure de buts, relations enseignants-élèves et adaptation scolaire des élèves : une analyse multi-niveaux. *Revue française de pédagogie : recherches en éducation*, 155, 57-72. doi: 10.4000/rfp.225
- Girardet, C. et Berger, J.-L. (2016). Motivation, sentiment de responsabilité et styles de gestion de classe. *Formation et pratiques d'enseignement en question*, 21, 287-309.
- Gough, H. G. (1952). *The adjective check list*. Palo Alto, CA: Consulting Psychologists Press.
- Guskey, T. R. (1982). Differences in teachers' perceptions of personal control of positive versus negative student learning outcomes. *Contemporary Educational Psychology*, 7(1), 70-80. doi: 10.1016/0361-476X(82)90009-1
- Juhel, J. et Rouxel, G. (2005). Effets du contexte d'évaluation sur les dimensions de la désirabilité sociale. *Psychologie du travail et des organisations*, 11, 59-68. doi: 10.1016/j.pto.2005.02.006
- Kaplan, A., Middleton, M. J., Urdan, T., & Midgley, C. (2002). Achievement goals and goal structures. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 21-53). Mahwah, NJ: Erlbaum.
- Lauermann, F., & Karabenick, S. (2011a). Taking teacher responsibility into account(ability): Explicating its multiple components and theoretical status. *Educational Psychologist*, 46(2), 122-140. doi: 10.1080/00461520.2011.558818
- Lauermann, F., & Karabenick, S. A. (2011b, April). *Motivation to become a teacher, teacher responsibility, and implications for mastery and performance approaches to instruction*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

- Lauermann, F., & Karabenick, S. A. (2013). The meaning and measure of teachers' sense of responsibility for educational outcomes. *Teaching and Teacher Education*, 30(1), 13-26. doi: 10.1016/j.tate.2012.10.001
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: MacMillan. doi: 10.1002/j.2330-8516.1987.tb00244.x
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Routledge. doi: 10.4324/9780203821961
- Oser, F. K. (1994). Moral perspective on teaching. *Review of Research in Education*, 20(1), 57-127.
- Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332. doi: 10.3102/00346543062003307
- Paquay, L., Altet, M., Charlier, E. et Perrenoud, P. (dir.). (2012). *Former des enseignants professionnels: quelles stratégies? Quelles compétences?* (4^e éd.). Bruxelles : De Boeck (coll. Perspectives en éducation et formation).
- Paulhus, D. L. (1984). Two-component models of socially-desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609. doi: 10.1037//0022-3514.46.3.598
- Périer, P. (2013). Crise et/ou recomposition identitaire des professeurs débutants du secondaire? *Recherche et formation*, 74, 17-28. doi: 10.4000/rechercheformation.2117
- Potvin, P. et Rousseau, R. (1993). Attitudes des enseignants envers les élèves en difficulté scolaire. *Revue canadienne de l'éducation*, 18(2), 132-149. doi: 10.2307/1495186
- Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcomes. *Journal of Educational Research*, 74(3), 185-190. doi: 10.1080/00220671.1981.10885308
- Rouxel, G., Briec, J., Juhel, J. et Le Maner-Idrissi, G. (2013). Analyse de la structure factorielle du tempérament EAS (version parents) auprès de deux échantillons d'enfants âgés de deux à cinq ans et de six à neuf ans. *Revue européenne de psychologie appliquée*, 63(3), 173-183. doi: 10.1016/j.erap.2013.02.002
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507-514. doi: 10.1007/bf02296192
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(8), 23-74.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.
- Tournois, J. et Dickes, P. (1993). *Pratique de l'échelonnement multidimensionnel: de l'observation à l'interprétation*. Bruxelles: De Boeck.
- Tournois, J., Mesnil, F. et Kop, J.-L. (2000). Autoduperie et hétéroduperie: un instrument de mesure de la désirabilité sociale. *Revue européenne de psychologie appliquée*, 50(1), 219-232.
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783-805. doi: 10.1016/S0742-051X(01)00036-1

Annexe

L'Échelle du sentiment de responsabilité des enseignants (traduction de la Teacher Responsibility Scale de Lauermann et Karabenick)

Imaginez que vous viviez les situations suivantes dans votre propre classe. Dans quelle mesure vous sentiriez-vous PERSONNELLEMENT responsable de n'avoir pas su anticiper chacune des situations? Veuillez indiquer pour chaque énoncé un nombre entre 0 et 100.

Je me sentirais PERSONNELLEMENT responsable si...

	0 = Pas du tout responsable à 100 = Totalemt responsable	0-100
a.	...l'un de mes apprenants ne s'intéressait pas à la matière que j'enseigne.	
b.	...l'un de mes apprenants pensait qu'il ne pourrait pas compter sur moi s'il avait besoin d'aide.	
c.	...un cours que j'ai enseigné ne reflétait pas au mieux mes habiletés d'enseignant-e.	
d.	...l'un de mes apprenants n'aimait pas la matière que j'enseigne.	
e.	...l'un de mes apprenants ne réalisait pas de grands progrès durant l'année scolaire.	
f.	...l'un de mes apprenants n'accordait pas de valeur à la matière que j'enseigne.	
g.	...un cours que j'ai enseigné n'avait pas favorisé l'apprentissage des apprenants autant que ce que j'aurais pu faire.	
h.	...l'un de mes apprenants ne pensait pas pouvoir me confier les problèmes qu'il rencontre à l'école ou en dehors de l'école.	
i.	...l'un de mes apprenants ne parvenait pas à apprendre la matière demandée au programme.	
j.	...l'un de mes apprenants ne réussissait pas bien.	
k.	...l'un de mes apprenants ne croyait pas que je me soucie vraiment de lui.	
l.	...un cours que j'ai enseigné n'était pas aussi stimulant pour les apprenants que ce que j'aurais pu faire.	