

Évaluation de la fidélité interjuges d'une grille d'observation des comportements d'offre active de services en français dans le cadre d'activités d'apprentissage par simulation

Coralie Vincent, Alexandra Maria Bodnaruc, Jacinthe Savard, Cris-Carelle Kengneson, Josée Benoît et Isabelle Giroux

Numéro 21, 2023

URI : <https://id.erudit.org/iderudit/1097641ar>

DOI : <https://doi.org/10.7202/1097641ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Institut canadien de recherche sur les minorités linguistiques / Canadian Institute for Research on Linguistic Minorities

ISSN

1927-8632 (numérique)

[Découvrir la revue](#)

Citer cet article

Vincent, C., Bodnaruc, A. M., Savard, J., Kengneson, C.-C., Benoît, J. & Giroux, I. (2023). Évaluation de la fidélité interjuges d'une grille d'observation des comportements d'offre active de services en français dans le cadre d'activités d'apprentissage par simulation. *Minorités linguistiques et société / Linguistic Minorities and Society*, (21). <https://doi.org/10.7202/1097641ar>

Résumé de l'article

Dans les régions canadiennes où les communautés francophones vivent en situation minoritaire, l'offre active (OA) de services en français consiste à proposer des services en français de façon proactive et dès le premier contact. Dans le contexte des services sociaux et de santé, l'OA contribue à la qualité, à la sécurité et à l'équité des soins destinés à ces communautés. La Grille d'observation des indicateurs de compétence en OA a été créée pour évaluer cette compétence chez les futurs professionnels dans un contexte d'apprentissage par simulation. Le présent projet de recherche visait à poursuivre l'évaluation métrologique de la grille en mesurant sa fidélité interjuges grâce au coefficient de corrélation intraclasse (CCI). Les résultats ont démontré un niveau de fidélité globale élevé (CCI = 0,87), ce qui permet d'encourager l'utilisation de la grille dans les milieux d'apprentissage. Des recherches supplémentaires permettront d'évaluer la fidélité de la grille dans divers contextes et d'assurer l'amélioration continue de la formation des futurs professionnels sur l'OA.

Varia

Évaluation de la fidélité interjuges d'une grille d'observation des comportements d'offre active de services en français dans le cadre d'activités d'apprentissage par simulation

Coralie VINCENT

Université d'Ottawa

cvinc064@uottawa.ca

Alexandra Maria BODNARUC

Université d'Ottawa

abodn049@uottawa.ca

Jacinthe SAVARD

Université d'Ottawa

jsavard@uottawa.ca

Cris-Carelle KENGNESON

Université d'Ottawa

ckeng055@uottawa.ca

Josée BENOÎT

Groupe de recherche sur la formation et les pratiques en santé et service social en contexte francophone minoritaire (GRFoPS)

josee.benoit@uottawa.ca

Isabelle GIROUX

Université d'Ottawa

igiroux@uottawa.ca

Note éditoriale

Toutes les auteures sont des chercheuses, associées ou assistantes de recherche membres du Groupe de recherche sur la formation et les pratiques en santé et en service social en contexte francophone minoritaire (GRFoPS).

Résumé

Dans les régions canadiennes où les communautés francophones vivent en situation minoritaire, l'offre active (OA) de services en français consiste à proposer des services en français de façon proactive et dès le premier contact. Dans le contexte des services sociaux et de santé, l'OA contribue à la qualité, à la sécurité et à l'équité des soins destinés à ces communautés. La Grille d'observation des indicateurs de compétence en OA a été créée pour évaluer cette compétence chez les futurs professionnels dans un contexte d'apprentissage par simulation. Le présent projet de recherche visait à poursuivre l'évaluation métrologique de la grille en mesurant sa fidélité interjuges grâce au coefficient de corrélation intraclasse (CCI). Les résultats ont démontré un niveau de fidélité globale élevé (CCI = 0,87), ce qui permet d'encourager l'utilisation de la grille dans les milieux d'apprentissage. Des recherches supplémentaires permettront d'évaluer la fidélité de la grille dans divers contextes et d'assurer l'amélioration continue de la formation des futurs professionnels sur l'OA.

Mots-clés: outils d'évaluation; offre active de services en français; formation des professionnels de la santé; apprentissage par simulation; fidélité interjuges

Abstract

In Canadian Francophone minority communities, the active offer (AO) of French language services involves proactively offering services in French from the first point of contact. In the context of health and social services, the AO contributes to the quality, safety and equity of care within these communities. The AO Competency Indicator Observation Grid was created to assess AO competency among future professionals in a simulated learning environment. The purpose of this research project was to continue the metrological evaluation of the grid by measuring its inter-rater reliability using the intraclass correlation coefficient (ICC). The results demonstrated a high level of global reliability (ICC = 0.87), thus supporting the use of the grid in learning environments. Further research will evaluate the reliability of the grid in a variety of settings and ensure continued improvement in the training of future professionals on the AO.

Keywords: assessment tools; active offer of French language services; health professional education; simulated learning; inter-rater reliability

Introduction. Minorités francophones et accès aux soins et services de santé

Les minorités de langue officielle, soit les anglophones vivant dans une province ou un territoire majoritairement francophone et les francophones vivant dans une province ou un territoire majoritairement anglophone, font partie intégrante du portrait démographique du Canada. L'anglais est la langue officielle utilisée de façon majoritaire dans 12 des 13 provinces et territoires canadiens, alors que le français est l'unique langue officielle de la province du Québec, où il est utilisé de façon majoritaire (Statistique Canada, 2022a). En 2021, 3,3 % des Canadiens vivant à l'extérieur du Québec ont dit utiliser le français comme première langue officielle parlée, ce qui représente près d'un million de francophones vivant en situation linguistique minoritaire au Canada (Statistique Canada, 2022b).

Bien que le bilinguisme officiel assure l'offre de services en français et en anglais au sein des institutions fédérales, des inégalités linguistiques demeurent présentes au niveau provincial et territorial, où l'une des deux langues officielles prédomine. Puisque les services sociaux et de santé sont de compétence provinciale ou territoriale, l'accès à des services en français n'est garanti que dans certaines provinces et certains territoires ou, parfois, que dans certains établissements désignés de ces provinces et territoires. Dans d'autres endroits, la possibilité de recevoir des services en français peut relever du hasard d'une rencontre fortuite entre une personne francophone et un intervenant pouvant s'exprimer en français.

Il a été démontré que les francophones vivant en situation minoritaire, particulièrement lorsqu'ils ont une maîtrise raisonnable de l'anglais, n'ont pas tendance à demander des services en français si ceux-ci ne sont pas offerts de façon proactive (Charbonneau, 2011). Par exemple, une étude de l'Institut canadien de recherche sur les minorités linguistiques effectuée en Nouvelle-Écosse a trouvé que seulement 13,6 % des francophones sondés affirmaient « qu'il y a de fortes probabilités qu'il[s] demande[nt] un service français s'il[s] n'[ont] pas été invité[s] à le faire, c'est-à-dire en l'absence d'incitatifs pour le faire », alors que près de 70 % des répondants ont affirmé qu'il était fort probable qu'ils continuent en français « si le prestataire du service amorce la conversation en français, par exemple en disant “Hello, bonjour!” » (Deveau *et al.*, 2010, s.p.).

Les personnes appartenant à une minorité de langue officielle sont confrontées à des barrières linguistiques qui peuvent, à différents échelons du système de la santé, limiter leur accès aux services et diminuer la qualité des soins qu'elles reçoivent (Bouchard et Desmeules, 2011, 2013; Bowen, 2001, 2015; de Moissac et Bowen, 2019). Il est indéniable que la nécessité de communiquer dans une langue autre que sa langue maternelle ou sa langue officielle prédominante peut constituer un défi particulier dans un contexte de sollicitation et de prestation de soins de santé (Drolet *et al.*, 2017). Comme le démontrent de nombreuses études, cette situation peut influencer sur la qualité des échanges entre les bénéficiaires et les professionnels de la santé et avoir des répercussions négatives sur le choix de traitement, la compréhension du traitement et l'adhérence au traitement, ainsi que sur la santé et la sécurité des bénéficiaires (Bowen, 2015; de Moissac et Bowen, 2019; Drolet *et al.*, 2014, 2017). De ce fait, l'invitation active des bénéficiaires à se prévaloir de services sociaux et de santé dans la langue officielle de leur choix est d'une importance majeure.

Formation à l'offre active de services en français

L'offre active (OA) de services en français est une pratique visant à améliorer l'offre de services aux communautés francophones vivant en situation minoritaire. Dans le contexte des services sociaux et de santé, elle consiste à inviter, de manière proactive et dès le premier contact, les bénéficiaires à s'exprimer dans la langue officielle de leur choix, et à assurer la continuité des services dans la langue officielle choisie (Bouchard *et al.*, 2012). La pratique de l'OA est importante afin d'assurer que les bénéficiaires savent qu'ils ont accès à des services en français tout au long du processus de prestation de soins et de services, mais également que la qualité des services sera la même peu importe la langue officielle choisie. En effet, des données empiriques récentes montrent que les francophones vivant dans une province ou un territoire majoritairement anglophone perçoivent le manque de fournisseurs de services sociaux et de santé parlant français (45 %) et la crainte de devoir attendre plus longtemps pour obtenir des services (26 %) comme étant des obstacles à l'obtention de services en français (Santé Canada, 2020, p. 10).

Bien que la sécurisation culturelle et linguistique fasse partie des compétences essentielles d'entrée à la pratique chez les professionnels de la santé et des services sociaux, les travaux de Bouchard et Vézina (2009) ont montré que les nouveaux diplômés de ces domaines, même lorsqu'ils étaient formés en français, étaient peu outillés pour reconnaître les difficultés qu'engendre la minorisation linguistique, ainsi que pour poser des gestes concrets permettant de faciliter l'accès à des services en français pour les minorités francophones. L'importance de cette compétence a été particulièrement renforcée au cours des dernières années, ce qui a encouragé la mise en place de nouvelles activités d'apprentissage sur l'OA au sein des établissements de formation. Malgré la disponibilité d'un nombre grandissant de ressources d'apprentissage et de promotion de l'OA (Casimiro *et al.*, 2018; Dubouloz *et al.*, 2014; Duchesne *et al.*, 2019; Savard *et al.*, 2015), il n'existait, jusqu'à récemment (Vincent *et al.*, 2021), aucun outil validé permettant d'évaluer les comportements observables de l'OA chez les futurs professionnels dans le cadre d'activités d'apprentissage par simulation, lesquelles permettent l'application pratique et concrète des concepts théoriques de l'OA.

La Grille d'observation des indicateurs de compétence en OA

Afin de répondre au besoin d'un tel outil, les travaux de notre équipe de recherche ont permis la création et la validation de la Grille d'observation des indicateurs de compétence en offre active de services en français (Vincent *et al.*, 2021). Cet outil d'évaluation a été élaboré afin de faciliter l'apprentissage expérientiel par simulation de la compétence en OA chez les étudiants des domaines de la santé et des services sociaux en identifiant les indicateurs de compétence associés aux comportements observables d'OA. Cette grille permet ainsi de guider les éducateurs dans la création de simulations portant sur une diversité de comportements liés à l'OA, ainsi que les étudiants vers l'acquisition et l'application pratique de cette compétence.

Validité et fidélité d'un outil de mesure

Un outil de mesure est dit fidèle et valide lorsqu'il mesure, de manière constante (fidélité) et sans contamination par d'autres caractéristiques, le(s) concept(s) qu'il prétend mesurer (validité) (Moskal et Leydens, 2000). L'évaluation métrologique globale d'une grille d'observation est un

processus qui inclut généralement deux étapes principales, soit l'évaluation de validité de contenu et l'évaluation de la fidélité.

La validité de contenu d'une grille d'observation réfère à la pertinence et à l'exhaustivité des items inclus dans l'outil par rapport aux concepts à évaluer (Lynn, 1986). Cette étape permet d'assurer la cohérence entre ces concepts et les connaissances sous-jacentes, ainsi que la clarté des directives d'utilisation, des énoncés et des choix de réponses ou de cotation (Lynn, 1986). La validité de contenu peut être établie à l'aide de recensions de la littérature, par le biais de groupes d'experts ou par une combinaison de ces deux méthodes.

La fidélité d'un outil réfère à sa capacité d'engendrer des résultats constants (Downing, 2004; Lewis *et al.*, 2008). Lorsqu'un outil prend la forme d'une grille d'évaluation ou d'observation, la fidélité peut être évaluée de différentes façons, notamment en mesurant la fidélité interjuges, définie comme une mesure du niveau de concordance entre les cotes attribuées par au moins deux évaluateurs différents lors de l'administration d'un instrument de mesure donné, chez le(s) même(s) sujet(s), au même moment et dans les mêmes conditions (Koo et Li, 2016). Dans les écrits en français, les termes fiabilité ou fidélité inter-évaluateurs, inter-observateurs ou inter-codeurs sont parfois utilisés en tant que synonymes de la fidélité interjuges. Comme le mentionnent Moskal et Leydens (2000), l'évaluation de la fidélité interjuges permet de répondre à trois questions principales sur la clarté et l'objectivité d'une grille d'observation : 1) est-ce que les différentes catégories de l'échelle de cotation établie dans la grille d'évaluation sont bien définies? 2) est-ce que les différences entre les cotes de réussite sont claires? et 3) est-ce que des juges indépendants donneraient la même cote de réussite pour le même étudiant en se basant sur la grille d'évaluation donnée? L'évaluation de la fidélité interjuges permet donc d'assurer la fiabilité et l'objectivité d'un outil de mesure en vérifiant si la structure et le contenu de la grille et de son échelle de cotation présentent un niveau de clarté adéquat pour engendrer des évaluations constantes et cohérentes de la part de différents juges.

L'importance de mesurer ce paramètre est rehaussée par l'augmentation graduelle de la taille moyenne des cohortes d'étudiants, phénomène qui conduit fréquemment les programmes offrant l'apprentissage par simulation à devoir faire appel à de multiples évaluateurs distincts pour noter la performance des étudiants d'un même groupe (Lucas *et al.*, 2017). Dans un tel contexte, il est primordial d'utiliser un outil qui présente une bonne fidélité interjuges, ce qui permet d'assurer une évaluation équivalente des étudiants, indépendamment de leurs évaluateurs. La recherche démontre que les méthodes d'évaluation sont un facteur majeur affectant la motivation des étudiants dans leur apprentissage, d'où la nécessité d'assurer que les outils d'évaluation utilisés sont justes, valides et fiables (Tsingos *et al.*, 2015).

Évaluation métrologique de la Grille d'observation des indicateurs de compétence en OA

La Grille d'observation des indicateurs de compétence en OA a été créée à partir d'une revue de la littérature sur les outils d'évaluation de format similaire, ainsi que d'une révision des outils actuels portant sur l'apprentissage et l'évaluation de l'OA. Le contenu et le format de la grille ont ensuite été validés par un groupe d'experts interdisciplinaire composé de professionnels de la santé, de chercheurs et d'éducateurs ayant des connaissances et des expériences pertinentes liées à l'utilisation et à l'enseignement de l'OA ainsi qu'à l'évaluation de simulations étudiantes. Ce processus de validation a permis de confirmer l'importance et la clarté de chaque indicateur de compétence inclus, ainsi que le format d'utilisation de l'outil. Le processus détaillé du développement du contenu et du format de la grille ainsi que de l'évaluation de la validité de celle-ci a été décrit dans une publication précédente (Vincent *et al.*, 2021).

Par conséquent, l'objectif de la présente étude était de continuer l'évaluation métrologique de la grille en mesurant sa fidélité interjuges lorsqu'utilisée dans le cadre d'activités d'apprentissage par simulation de niveau universitaire offertes à de futurs professionnels de la santé. Le projet a été évalué et approuvé par le Bureau d'éthique et d'intégrité de la recherche de l'Université d'Ottawa.

1. Méthode

1.1. Contenu et format de la grille

La Grille d'observation des indicateurs de compétence en OA prend la forme d'une grille à échelle descriptive comprenant trois éléments principaux, soit les indicateurs de compétence, les catégories d'indicateurs de compétence et l'échelle de cotation.

1.1.1. Indicateurs de compétence et catégories

La grille d'observation comprend un total de 20 indicateurs de compétence, lesquels correspondent à des comportements observables d'OA pouvant être démontrés par l'étudiant dans un contexte de simulation. Les indicateurs de compétence sont groupés en cinq catégories qui réfèrent à des composantes générales d'une rencontre typique entre bénéficiaire et professionnels de la santé et des services sociaux, ainsi qu'à des habiletés linguistiques de base. Les indicateurs de compétence ainsi que leurs catégories sont présentés dans le tableau 1.

Tableau 1

Indicateurs de compétence et catégories de la Grille d'observation des indicateurs de compétence en OA

Code	Indicateur de compétence
Catégorie A – Introduction	
A-1	Porte une identification en évidence qui mentionne les deux langues (ex., épinglette « Bonjour/Hello »).
A-2	Accueille une nouvelle personne dans les deux langues officielles (français et anglais).
A-3	Demande au bénéficiaire dans quelle langue (français ou anglais) il/elle souhaite poursuivre la conversation.
A-4	Demande au bénéficiaire dans quelle langue (français ou anglais) il/elle souhaite remplir son formulaire d'accueil.
A-5	Poursuit l'entrevue dans la langue officielle choisie par le/la bénéficiaire.
A-6	Vérifie la langue de préférence de chacun des bénéficiaires et accompagnateurs (s'il y a lieu).
A-7	Inscrit la langue de préférence du bénéficiaire dans son dossier.
Catégorie B – Habiletés linguistiques	
B-1	Utilise le vouvoiement de façon appropriée lorsque la communication se fait en français (c.-à-d. vouvoyer tous les adultes, tutoyer les enfants, les adolescents et les adultes qui ont clarifié qu'ils préféreraient être tutoyés).
B-2	Démontre un niveau d'habiletés linguistiques dans la langue de choix du bénéficiaire assez élevé pour comprendre le bénéficiaire et se faire comprendre (ex. fluidité, compréhension, etc.).
Catégorie C – Intervention	
En présence d'un client ayant un vocabulaire, des expressions ou un niveau de langue différents des siens, l'étudiant :	
C-1	S'ajuste au niveau de langue du bénéficiaire pour qu'il/elle se sente à l'aise de communiquer en français (ex. niveau de langage, régionalisme, culture, etc.).
C-2	Valide que le/la bénéficiaire a compris les explications et les recommandations (ex., en lui demandant de les répéter ou de réexpliquer dans ses propres mots).
Catégorie D – Ressources informatives	
D-1	Offre des documents dans les deux langues officielles ou selon le choix du bénéficiaire.
D-2	Vérifie que les outils d'éducation et d'information disponibles en français sont adaptés au contexte local du bénéficiaire (ex. contexte culturel ou réglementaire).
D-3	S'assure que le/la bénéficiaire comprend le niveau de langue utilisé dans les documents écrits.
D-4	Tient compte de la langue préférée de chaque proche ou membre de la famille dans la planification des ressources à offrir.
D-5	Veille à ce qu'il y ait un affichage indiquant les comptoirs d'accueil où il y a la possibilité d'être servi dans les deux langues.
D-6	Veille à ce qu'il y ait des documents informatifs et des outils de divertissement en français (ex., dépliants, revues, journaux, radio, télévision, jeux) dans les aires communes (ex. salle d'attente des bénéficiaires ou autres aires communes).
Catégorie E – Suivi	
E-1	Tient compte de la langue préférée du bénéficiaire lorsque celui-ci est dirigé vers une autre ressource ou un autre intervenant.
E-2	Informe verbalement les autres membres de l'équipe de la langue préférée du bénéficiaire, ou demande à un autre intervenant de le faire.
E-3	Note la langue préférée du bénéficiaire sur le formulaire de référence.

Source : Adapté de Vincent *et al.* (2021)

La catégorie A – Introduction comprend sept indicateurs de compétence visant à évaluer si l'étudiant applique les principes de l'OA dès le premier contact avec le bénéficiaire et, le cas échéant, avec son ou ses accompagnateur(s). La catégorie B – Habiletés linguistiques et la catégorie C – Intervention comprennent deux indicateurs de compétence chacune, visant à évaluer le niveau d'efficacité de la communication entre l'étudiant et le bénéficiaire. Plus précisément, les

indicateurs de compétence de la catégorie B évaluent si les habiletés linguistiques de l'étudiant dans la langue de choix du bénéficiaire sont suffisantes pour permettre une communication efficace, alors que les indicateurs de compétence de la catégorie C évaluent la capacité de l'étudiant à adapter son niveau de langue aux besoins du bénéficiaire tout en vérifiant sa compréhension. Enfin, la catégorie D – Ressources informatives et la catégorie E – Suivi comprennent six et trois indicateurs de compétence, respectivement. Les indicateurs de compétence de la catégorie D visent à évaluer si l'étudiant applique les principes de l'OA lors du choix des ressources informatives et de leur remise auprès du bénéficiaire, en s'assurant notamment qu'elles sont offertes dans les deux langues officielles ou dans la langue officielle préférée du bénéficiaire, adaptées au niveau de langue de ce dernier et pertinentes compte tenu du contexte de milieu de vie du bénéficiaire. Les indicateurs de la catégorie E visent, quant à eux, à évaluer si l'étudiant prend en considération la langue officielle de préférence du bénéficiaire et en informe les autres professionnels lors de la planification d'un suivi ou lors de l'aiguillage des patients vers un autre service. Bien que certains des comportements visés par les indicateurs de la catégorie A et D puissent, dans un contexte de travail, être démontrés par d'autres membres de l'équipe de soin (par ex. gestionnaire, préposé à l'accueil), il est important que tous les futurs professionnels de la santé et des services sociaux développent et comprennent l'ensemble des aspects de la compétence en OA.

1.1.2. Échelle de cotation et guide d'utilisation

Le niveau de performance d'un étudiant pour chacun des indicateurs de compétence de la grille d'observation est noté sur une échelle de Likert à quatre niveaux, générant un score de 1 à 4 points. Le niveau 1) *n'atteint aucunement les attentes* est utilisé lorsque l'étudiant n'a pas démontré le comportement associé à l'indicateur de compétence. Le niveau 2) *atteint partiellement les attentes* réfère à la performance d'un étudiant qui n'a pas entièrement démontré le comportement décrit par l'indicateur ou qui l'a démontré de façon sous-optimale. Le niveau 3) *atteint les attentes* est utilisé lorsque l'étudiant démontre entièrement et adéquatement le comportement décrit par l'indicateur de compétence, alors que le niveau 4) *dépasse les attentes* est choisi lorsque l'étudiant démontre l'ensemble des composantes du comportement décrit par l'indicateur de compétence et les démontre plus d'une fois ou à un niveau supérieur à ce qui est attendu d'un professionnel novice. Puisqu'il est rare qu'une simulation permette individuellement l'évaluation de tous les comportements d'OA visés par les indicateurs de performance inclus dans la grille, le score total de la grille peut varier selon le nombre d'indicateurs inclus dans la simulation utilisée.

Un guide d'utilisation a également été créé pour accompagner la grille. Celui-ci vise à fournir des explications générales sur le fonctionnement de la grille, à guider les formateurs dans l'élaboration de scénarios de simulation en précisant certains éléments à inclure pour permettre aux étudiants de démontrer les comportements d'OA visés par les indicateurs de compétence de la grille et, enfin, à offrir des explications détaillées sur la façon de coter chaque indicateur de compétence selon les niveaux de performance.

1.2. Description des participants et de l'activité d'apprentissage par simulation

1.2.1. Participants

Les participants à cette étude étaient des étudiants de trois cohortes consécutives (2018-2020) du programme de maîtrise ès sciences de la santé en ergothérapie de l'Université d'Ottawa. Dans le cadre d'un cours de première année du programme, tous les étudiants de ces cohortes devaient participer à une activité d'apprentissage par simulation qui a été filmée à des fins éducatives. Parmi ces étudiants, 93 ont consenti à ce que l'enregistrement vidéo de la simulation soit utilisé pour l'évaluation de programme ou la recherche. Les enregistrements vidéo étaient identifiés par

des codes alphanumériques et sécurisés par des mots de passe auxquels seuls les chercheurs et le personnel de recherche autorisés avaient accès.

1.2.2. Activité d'apprentissage par simulation

L'activité d'apprentissage par simulation utilisée pour évaluer la fidélité interjuges de la Grille d'observation des indicateurs de compétence en OA a été filmée à l'aide de caméras liées au logiciel de gestion de simulations audiovisuelles *CAE Learning Space* (2018-2019) ou via l'application *Zoom* (2020). L'utilisation d'enregistrements audiovisuels de simulations d'étudiants présentant différents niveaux de performance est une méthode courante et adéquate pour l'évaluation de la fidélité interjuges d'une grille d'observation (Cross *et al.*, 2001; Williams *et al.*, 2003).

Le scénario de simulation utilisé dans le cadre de l'activité d'apprentissage visait à permettre la démonstration de compétences cliniques de base, ainsi que de la compétence en OA. Les étudiants prenant part à la simulation avaient précédemment participé à des cours théoriques traitant de l'OA et savaient que cette compétence serait évaluée lors de la simulation. Le scénario était conçu pour évaluer un étudiant à la fois et faisait appel à un acteur professionnel qui jouait le rôle du proche aidant d'un bénéficiaire; dans le cas de l'année 2020, l'évaluateur présent lors de la simulation a brièvement joué le rôle d'un travailleur social. En effet, il existe des différences mineures entre le scénario utilisé en 2020 et le scénario initial, utilisé en 2018 et en 2019. L'évaluation de la compétence en OA à l'aide des vidéos de simulation des étudiants qui ont réalisé l'activité d'apprentissage en 2018 et en 2019 a été effectuée durant l'été 2020 et a permis à l'équipe de recherche d'observer que certains indicateurs de compétence de la grille étaient difficiles ou impossibles à évaluer en raison du scénario utilisé (par ex. le scénario n'impliquait pas d'autres professionnels de la santé ou des services sociaux) ou de problèmes logistiques (par ex. il était difficile de voir la langue des ressources et documents utilisés par l'étudiant à l'écran). Ainsi, quelques modifications ont été apportées au scénario et au déroulement de la simulation au cours de l'année 2020 pour permettre l'évaluation des indicateurs A-1, E-1 et E-2 qui n'avaient pas été évalués les années précédentes, ainsi que pour faciliter l'évaluation de l'indicateur D-1. La rétroaction des juges à la suite des évaluations de 2018-2019 a également mené à l'ajout de la catégorie B – Habiletés linguistiques, laquelle n'était pas dans la grille initiale. Outre les différences sur le plan du scénario, l'activité d'apprentissage par simulation s'est déroulée en présentiel en 2018 et 2019, alors qu'elle s'est déroulée en mode virtuel en 2020 en raison de la pandémie de COVID-19. À cause de ces facteurs, la fidélité de certains indicateurs de compétence (n = 7) a été évaluée à partir des vidéos des trois années, alors que celle d'autres indicateurs (n = 5) a seulement été évaluée à partir des vidéos de 2020.

1.3. Description des évaluateurs et du processus d'évaluation

1.3.1. Recrutement des évaluateurs

Un total de cinq évaluateurs (juges) a été recruté pour évaluer, à l'aide de la Grille d'observation des indicateurs de compétence en OA, le niveau de performance des étudiants à la suite de l'activité d'apprentissage par simulation effectuée et filmée en 2018, en 2019 et en 2020. Chaque enregistrement audiovisuel a été indépendamment visionné et évalué par trois juges. Les enregistrements des simulations effectuées en 2018 et en 2019 ont été visionnés et évalués par les trois mêmes juges durant l'été 2020, alors que les enregistrements des simulations effectuées en 2020 ont été évalués par deux nouveaux juges en 2021 ainsi qu'un des juges présents en 2020.

Ces évaluateurs ont été recrutés au sein du personnel et des étudiants diplômés de la Faculté de sciences de la santé de l'Université d'Ottawa. Ils avaient tous de l'expérience dans un ou plusieurs domaines des sciences de la santé, ainsi qu'une maîtrise adéquate des concepts de base de l'OA. Quatre des juges avaient également de l'expérience dans l'évaluation de simulations étudiantes.

Afin de limiter le risque de partialité, les évaluateurs ne connaissaient pas les étudiants présentés dans les simulations à évaluer. Effectivement, il est démontré que, lorsque les évaluateurs connaissent certains des étudiants, ils ont tendance à donner de meilleurs scores à ceux auxquels ils ont enseigné qu'à ceux qu'ils ne connaissent pas, ce qui peut nuire à la qualité de l'évaluation de la fidélité (Lewis *et al.*, 2008).

1.3.2. Formation des évaluateurs

La formation des évaluateurs sur les modalités d'utilisation de l'outil d'évaluation est l'un des éléments principaux qui permettent d'assurer une fidélité interjuges adéquate (Adamson *et al.*, 2012). De nombreuses études portant sur la fidélité interjuges de grilles d'évaluation utilisées dans le cadre de simulations notent que la fidélité est meilleure lorsque les juges ont reçu une formation, une présentation ou un guide d'accompagnement expliquant l'utilisation appropriée de la grille (Adamson *et al.*, 2012; Castorr *et al.*, 1990; Dunbar *et al.*, 2006;). La fidélité semble également être meilleure lorsque les évaluateurs ont la possibilité de mettre à l'essai l'utilisation de l'outil et de poser leurs questions avant de procéder aux évaluations formelles (Keshmiri *et al.*, 2016; Lucas *et al.*, 2017; Reid *et al.*, 2012).

Dans le cadre du projet présenté, les cinq juges ont tout d'abord assisté à une courte présentation sur la grille et ont reçu le guide d'utilisation décrit plus haut. Les juges ont également été informés des indicateurs de compétence de la grille qui s'appliquaient à la simulation évaluée, en fonction des différentes cohortes. Il leur a ensuite été possible de s'exercer en évaluant trois vidéos de simulations, puis de se regrouper avec les chercheurs afin de poser leurs questions et de clarifier les niveaux de cotation des indicateurs de compétence. Ces éléments de formation sont recommandés pour tout éducateur ou programme souhaitant utiliser la Grille d'observation des indicateurs de compétence en OA dans le cadre de l'évaluation de simulations, car c'est une étape permettant de minimiser le risque de partialité et de favoriser une compréhension uniforme des niveaux de cotation parmi les différents évaluateurs.

1.3.3. Processus d'évaluation

Les cinq juges ont évalué les vidéos de simulations à l'aide de la Grille d'observation des indicateurs de compétence en OA de façon indépendante. Ils n'avaient pas accès aux évaluations effectuées par les autres juges lors du processus d'évaluation afin de limiter le risque d'influence entre les différents évaluateurs (Williams *et al.*, 2003).

Avant le début des évaluations, les juges ont été informés qu'ils devaient évaluer les simulations à l'aide des indicateurs A-2, A-3, A-5, A-6, C-2, D1 et D-4 pour les simulations des années 2018-2019, puis de ces mêmes indicateurs avec l'ajout des indicateurs A-1, B1, B2, E-1 et E-2 en 2020. Huit indicateurs (A-4, A-7, C-1, D-2, D-3, D-5, D-6, E-3) n'ont pas été analysés dans la présente étude parce qu'il était impossible de les évaluer à l'aide de la simulation utilisée, en raison soit du scénario, soit de l'utilisation de simulations en mode virtuel et de fichiers audiovisuels.

Lors des évaluations, les juges pouvaient également utiliser la cote alternative « ne s'applique pas » lorsqu'un défaut technique dans les vidéos ou un problème dans le déroulement de la simulation rendait un indicateur impossible à évaluer pour un étudiant spécifique.

1.4. Analyses statistiques

Les analyses statistiques ont été effectuées à l'aide du logiciel *IBM SPSS Statistics* (v27.0). La fidélité interjuges de la grille d'observation a été évaluée à l'aide du coefficient de corrélation intraclasse (CCI). Le CCI permet de mesurer le niveau d'accord entre les cotes données par les différents juges et peut être utilisé avec deux évaluateurs ou plus, pour des données continues (par intervalles ou par ratios) et ordinales (Hallgreen, 2012). Il est possible d'utiliser le CCI pour évaluer la fidélité des

indicateurs de compétence individuels ainsi que pour évaluer la fidélité d'une catégorie d'indicateurs ou de la grille entière. Comme le recommande Hallgreen (2012), les paramètres utilisés pour calculer le CCI dans SPSS étaient : 1) *accord absolu* puisque les cotes données par les juges à l'aide de la grille étaient seulement distribuées sur quatre niveaux de réussite différents, 2) *deux facteurs (aléatoires)*, car les mêmes juges ont évalué les simulations de plusieurs étudiants, ainsi que 3) *mesure unique*, car le résultat d'évaluation accordé à l'étudiant lors d'une simulation en contexte d'apprentissage provient généralement d'un seul évaluateur, et non de la moyenne des évaluations de plusieurs évaluateurs différents.

Le calcul du CCI des indicateurs, des catégories d'indicateurs et du score total de la grille incluait, pour chaque indicateur, les données des étudiants ayant reçu une cote valide de la part de trois juges. Toutefois, l'analyse de la fidélité interjuges de l'indicateur B-1 incluait seulement les évaluations de deux juges puisque le troisième juge a attribué la même cote à tous les étudiants, ce qui a mené à l'exclusion de ses évaluations par SPSS en raison d'un manque de variabilité. Les cotes données par ce troisième juge ont tout de même été incluses dans le calcul du CCI pour les scores totaux de la catégorie B et de la grille.

Lors de l'analyse des CCI du score total et du score des catégories pour lesquelles il y avait une taille d'échantillon variable pour chaque indicateur, les scores totaux des étudiants ont été transformés sur une échelle de 100 points (score total de l'étudiant/score total possible selon le nombre d'indicateurs ayant reçu trois cotes valides pour cet étudiant X 100). Cette transformation visait à corriger l'augmentation de la variabilité des scores et de l'accord interjuges lorsque les étudiants n'avaient pas été évalués selon le même nombre d'indicateurs pour une catégorie donnée ou pour le score total de la grille. À des fins de comparaison, le CCI a également été calculé en utilisant uniquement les données des étudiants ayant reçu des cotes valides de la part des trois juges pour tous les indicateurs de la catégorie ou de la grille.

Tel qu'il est recommandé dans les lignes directrices sur l'utilisation du CCI publiées par Koo et Li (2016), les résultats ont été interprétés à l'aide de l'échelle suivante : un CCI $\geq 0,90$ démontre un niveau de fidélité excellent, un CCI $\geq 0,75$ représente une fidélité élevée, un CCI $\geq 0,50$ démontre une fidélité modérée et un CCI $< 0,50$ dénote un niveau de fidélité faible.

La taille d'échantillon minimale nécessaire pour étudier la fidélité d'une grille d'évaluation à l'aide du CCI a été calculée selon la méthode de Walter *et al.* (1998). Cette méthode permet de choisir une taille d'échantillon appropriée pour la mesure du CCI à l'aide de la valeur de fidélité minimale qui serait considérée comme acceptable pour l'outil, du niveau de signification statistique et de puissance désiré, du nombre d'évaluateurs par sujet (dans ce cas-ci, par simulation/étudiant), du taux d'abandon attendu des participants au projet, ainsi que du niveau de fidélité attendu de la grille. Dans le cadre de ce projet, les paramètres utilisés étaient : 1) valeur de fidélité minimale acceptable de 0,70; 2) niveau de signification statistique fixé à $\alpha = 0,05$ et puissance $(1-\beta)$ fixée à 80 %; 3) nombre d'évaluateurs/répétitions par sujet de $n = 3$; 4) taux d'abandon de 0 % et 5) niveau de fidélité anticipé pour le total de la grille entre 0,80 et 0,85. Ces paramètres généraient une taille d'échantillon minimale allant de 37 participants pour une fidélité attendue plus élevée (CCI = 0,85), jusqu'à 101 participants pour un niveau de fidélité attendu plus faible (CCI = 0,80).

2. Résultats

Les résultats détaillés des analyses de fidélité interjuges sont présentés dans le tableau 2. La fidélité du score total de la grille était élevée, soit un CCI de 0,87 (intervalle de confiance à 95 % [IC 95 %] : 0,82 - 0,90) en utilisant les scores de l'échantillon complet ajustés sur 100 points ($n = 93$), et de 0,88 (IC 95 % : 0,78 - 0,94) en utilisant seulement le score des participants ayant reçu des cotes valides de trois juges pour tous les indicateurs évalués ($n = 24$). En ce qui concerne chaque catégorie

d'indicateurs, les valeurs des CCI indiquaient un niveau de fidélité élevé pour les catégories A – Introduction (CCI : 0,87, IC 95 % : 0,78 - 0,92) et B – Habilités linguistiques (CCI : 0,75, IC 95 % : 0,58 - 0,87), ainsi qu'un niveau modéré pour la catégorie E – Suivi (CCI : 0,53, IC 95 % : 0,29 - 0,73). Le CCI de la catégorie D – Ressources informatives était de 0,81 (IC 95 % : 0,70 - 0,87) selon les scores ajustés de l'échantillon complet, et de 0,71 (IC 95 % : 0,52 - 0,84) selon les scores de l'échantillon réduit incluant tous les indicateurs, ce qui démontre une fidélité de niveau modéré à élevé.

Tableau 2

Résultats de fidélité interjuges de la Grille d'observation des indicateurs de compétence en OA

Indicateur de compétence	N	CCI ^a [IC 95 %]	Niveau de fidélité ^b
Catégorie A – Introduction			
A-1	24	0,40 [0,09 - 0,66]	Faible
A-2	80	0,86 [0,80 - 0,90]	Élevé
A-3	85	0,82 [0,75 - 0,87]	Élevé
A-5	86	0,58 [0,29 - 0,75]	Modéré
A-6	71	0,81 [0,73 - 0,87]	Élevé
Score total – Ajusté sur 100 points ^c	92	0,87 [0,78 - 0,92]	Élevé
Score total – Données complètes seulement ^d	24	0,87 [0,72 - 0,94]	Élevé
Catégorie B – Habiletés linguistiques			
B-1 [*]	26	0,72 [0,47 - 0,86]	Modéré
B-2	26	0,69 [0,50 - 0,83]	Modéré
Score total	26	0,75 [0,58 - 0,87]	Élevé
Catégorie C – Intervention			
C-2	93	0,11 [0,00 - 0,23]	Faible
Catégorie D – Ressources informatives			
D-1	36	0,74 [0,59 - 0,85]	Modéré
D-4	86	0,71 [0,56 - 0,81]	Modéré
Score total – Ajusté sur 100 points ^c	91	0,81 [0,70 - 0,87]	Élevé
Score total – Données complètes seulement ^d	31	0,71 [0,52 - 0,84]	Modéré
Catégorie E – Suivi			
E-1	26	0,35 [-0,01 - 0,64]	Faible
E-2	26	0,60 [0,38 - 0,78]	Modéré
Score total	26	0,53 [0,29 - 0,73]	Modéré
Score total – Ajusté sur 100 points ^c	93	0,87 [0,82 - 0,90]	Élevé
Score total – Données complètes seulement ^d	24	0,88 [0,78 - 0,94]	Élevé

^a Coefficient de corrélation intraclasse (CCI), accord absolu, deux facteurs aléatoires, mesures uniques, intervalle de confiance [IC] à 95 %.

^b Faible : CCI < 0,50; modéré : CCI ≥ 0,50; élevé : CCI ≥ 0,75; excellent : CCI ≥ 0,90.

^c CCI calculé avec les scores ajustés sur une échelle de 100 points pour tous les participants ayant reçu des cotes valides de la part des trois juges pour au moins un indicateur d'une section donnée.

^d CCI calculé seulement avec les données des participants ayant obtenu des cotes valides de la part de 3 juges pour tous les indicateurs évalués dans une section donnée.

*Un juge a été exclu de l'analyse statistique de cet indicateur par SPSS, car il n'y avait aucune variabilité dans les cotes qu'il a accordées. Les évaluations de ce juge pour cet indicateur ont néanmoins été incluses dans le calcul du CCI du score total de la catégorie B et de la grille.

Les valeurs de CCI des indicateurs de compétence individuels indiquent quant à elles que trois indicateurs de la grille (A-2, A-3 et A-6) ont obtenu un niveau de fidélité élevé (CCI entre 0,81 et 0,86), que six indicateurs (A-5, B-1, B-2, D-1, D-4 et E-2) ont obtenu un niveau de fidélité modéré (CCI entre 0,58 et 0,74), et que trois indicateurs (A-1, C-2 et E-1) ont obtenu un niveau de fidélité faible (CCI entre 0,11 et 0,40).

3. Discussion

Les projets antérieurs de notre équipe de recherche ont mené à la création et à la validation de contenu de la Grille d'observation des indicateurs de compétence en OA. L'analyse de la fidélité interjuges présentée dans cet article visait quant à elle à évaluer si les éléments de la grille étaient suffisamment clairs et objectifs pour permettre à plusieurs juges d'évaluer de manière identique ou similaire la performance d'un même étudiant.

Les résultats de l'analyse de fidélité interjuges, basés sur les indicateurs de compétence évalués, indiquent que la fidélité pour le score total de la grille est de niveau élevé, avec un CCI de 0,87. Il peut donc être conclu que la grille offre, globalement, un niveau de fidélité adéquat et permet une évaluation fiable et équitable de la compétence en OA dans un contexte de simulations.

Une analyse plus approfondie des résultats permet cependant de noter des différences entre les niveaux de fidélité de certaines catégories et de certains indicateurs. Tout d'abord, les catégories A – Introduction, B – Habiletés linguistiques et D – Ressources informatives se démarquent par leurs niveaux de fidélité élevés. Ce résultat peut, en partie, être attribué au fait que les indicateurs de ces catégories réfèrent à des comportements précis et facilement observables, ce qui permet une meilleure objectivité de la part des évaluateurs en limitant l'effet potentiel de la perception individuelle.

Néanmoins, trois indicateurs (A-1, C-2 et E-1) ont obtenu des niveaux de fidélité faibles en comparaison des autres indicateurs de la grille. À la suite de l'analyse des résultats par l'équipe de recherche, des discussions ont été organisées avec les cinq juges afin de déterminer les causes possibles de ces niveaux de fidélité limités.

L'indicateur A-1 fait partie de la catégorie Introduction et évalue si l'étudiant « porte une identification en évidence qui mentionne les deux langues (ex., épinglette "Bonjour/Hello") ». Il n'avait pas été possible d'évaluer cet indicateur lors des simulations de 2018-2019 parce que les étudiants n'avaient pas été informés qu'ils devaient porter une telle identification avant les simulations. Bien que l'indicateur décrive un comportement simple et objectif, l'évaluation de ce comportement à partir des vidéos de simulations de l'année 2020 a démontré un faible niveau de fidélité interjuges. L'analyse des données et les discussions avec les juges ont permis de dégager des pistes d'explications. Puisque les simulations de l'année 2020 se déroulaient en virtuel, les niveaux de luminosité et de clarté des caméras des ordinateurs personnels des étudiants variaient considérablement, ce qui a pu limiter la capacité des juges à lire correctement les épinglettes portées par les étudiants. Dans certains cas, il est également possible que le cadrage ou la position de l'étudiant ait empêché l'apparition adéquate des épinglettes à l'écran. Ces problèmes techniques ont mené à une confusion de la part des évaluateurs sur les cotes de réussite appropriées pour chaque situation. Afin de remédier à ce problème, il est proposé d'informer les étudiants avant la simulation qu'ils ont la responsabilité d'assurer que leur épinglette est visible et lisible durant la simulation. Des indications plus précises ont également été ajoutées au guide d'utilisation de la grille pour informer les évaluateurs que, lors de simulations virtuelles, une épinglette non visible ou difficilement lisible à l'écran occasionne une cote de 1 (ne porte pas d'épinglette) ou 2 (porte une épinglette, mais de façon plus ou moins lisible), respectivement. La cote 2 peut également être

attribuée dans un contexte de simulation en personne si, par exemple, l'épinglette est portée au niveau de la ceinture ou cachée par les cheveux de l'étudiant.

Le second indicateur ayant obtenu un faible niveau de fidélité fait partie de la catégorie C – Intervention et mesure si l'étudiant « valide que le/la bénéficiaire a compris les explications et les recommandations » (C-2). Cet indicateur fait référence à un comportement observable d'OA moins reconnu et plus subjectif que les autres indicateurs de la grille. En effet, la validation de la compréhension du bénéficiaire est une composante essentielle d'une communication efficace entre tous les prestataires de soins et leurs bénéficiaires, peu importe la langue de préférence de ces derniers. Cependant, ce comportement est particulièrement important afin d'assurer la qualité et la sécurité des soins offerts aux communautés francophones vivant en situation minoritaire puisque les défis liés à la langue peuvent compliquer la compréhension du bénéficiaire des informations communiquées par le prestataire de soins (Drolet *et al.*, 2014, 2017; Savard *et al.*, 2015). Lors des discussions avec les juges à la suite de l'analyse des résultats, des différences marquées ont été constatées quant aux éléments de comportement attendus par les différents juges pour chaque niveau de réussite de cet indicateur. Par exemple, certains juges considéraient des questions courtes telles que « C'est bon? » ou « Avez-vous des questions ? » comme une vérification adéquate de la compréhension du bénéficiaire, alors que d'autres juges cherchaient à ce que l'étudiant demande au bénéficiaire de répéter les explications ou de les reformuler dans ses propres mots. Ainsi, des changements ont été apportés au guide d'utilisation de la grille afin d'augmenter l'objectivité de cet indicateur, en expliquant aux évaluateurs quels éléments de comportement précis, démontrés à quelle fréquence et dans quel contexte, sont nécessaires pour obtenir chaque cote de réussite. Des parenthèses contenant des directives additionnelles (« ex., en lui demandant de les répéter ou réexpliquer dans ses propres mots ») ont également été ajoutées à l'indicateur de compétence dans la grille afin de fournir aux étudiants un exemple précis d'une méthode leur permettant de bien démontrer ce comportement et d'obtenir une cote de réussite plus élevée.

Enfin, le troisième indicateur ayant démontré un faible niveau de fidélité fait partie de la catégorie E – Suivi et il mesure si l'étudiant « tient compte de la langue préférée du bénéficiaire lorsqu'il/elle est dirigé vers une autre ressource ou un autre intervenant » (E-1). Le scénario de simulation utilisé en 2018-2019 incluait seulement l'étudiant, qui jouait le rôle du prestataire de soins, ainsi qu'un acteur, qui jouait le rôle du proche d'un bénéficiaire, et ne permettait donc pas l'évaluation de cet indicateur. En 2020, le rôle d'un travailleur social a été ajouté à la simulation afin de faciliter l'évaluation de cet indicateur additionnel. Il était prévu que ce rôle, impliquant une brève interaction à la fin du scénario, serait joué par l'évaluateur déjà présent lors du déroulement normal de la simulation. Les étudiants étaient informés qu'ils devaient interagir avec ce deuxième prestataire à la fin de la rencontre dans le but de faciliter la poursuite des soins multidisciplinaires offerts au bénéficiaire. L'indicateur de compétence visait ainsi à vérifier que l'étudiant s'assure d'indiquer clairement au travailleur social la langue officielle préférée du bénéficiaire lors du transfert des soins, et ce, sans que le bénéficiaire ait à le demander. Lors des simulations en ligne de l'année 2020, l'évaluateur jouant ce rôle allumait sa caméra quelques minutes avant la fin de la simulation afin de signaler « l'arrivée du travailleur social » et prendre part à la suite du scénario. Les juges ont toutefois rapporté qu'il semblait y avoir de la confusion chez les étudiants lorsque l'évaluateur allumait sa caméra. Même s'ils avaient été informés au préalable de cette façon de faire, certains étudiants ont pensé que cette action indiquait la fin de la simulation, et non l'arrivée du travailleur social. Plusieurs ne pensaient pas que c'était leur rôle de transmettre de l'information au travailleur social et s'attendaient plutôt à ce que ce dernier entreprenne la conversation avec le bénéficiaire sans qu'ils interviennent. Cette situation a compliqué l'évaluation de l'indicateur par les différents juges et a permis de souligner l'importance que les étudiants, les acteurs et les évaluateurs soient bien informés de leurs rôles respectifs et des étapes de la simulation et de l'évaluation des comportements d'OA afin d'assurer une évaluation optimale des indicateurs de compétence. Il serait également utile de ré-évaluer cet indicateur à l'aide d'une

simulation conçue spécialement pour inclure une équipe interdisciplinaire comprenant plusieurs prestataires de soins et de services qui travaillent conjointement pour aider le bénéficiaire.

Il existe de multiples facteurs susceptibles d'influer sur les résultats de l'évaluation de la fidélité interjuges d'une grille d'observation. Dans le cadre de cette étude, le recrutement de juges compétents qui ne connaissaient pas les étudiants à évaluer, la formation adéquate des juges avant le début des évaluations ainsi que la réalisation d'évaluations indépendantes par les différents juges sont des facteurs importants ayant permis de minimiser le risque de biais de recherche et d'assurer une approche rigoureuse et systématique au processus d'évaluation (Adamson *et al.*, 2012; Lewis *et al.*, 2008; Williams *et al.*, 2003).

Néanmoins, les résultats présentés doivent être interprétés dans le contexte de certaines limites. Le coefficient de corrélation intraclasse, utilisé ici pour mesurer la fidélité interjuges, tend à produire de meilleurs niveaux de fidélité lorsque l'étendue des scores possibles est large et qu'il y a une bonne variabilité à travers les scores des différents participants (Koo et Li, 2016). L'utilisation d'une échelle de Likert à quatre niveaux de réussite, bien qu'elle permette une cotation simple et rapide pour les évaluateurs, limite l'étendue des scores possibles et la variabilité des scores lors de l'évaluation de la fidélité des indicateurs individuels. Cette lacune pourrait partiellement expliquer pourquoi les résultats de fidélité des indicateurs individuels sont généralement plus faibles que ceux des scores totaux de la grille et de ses catégories. De surcroît, les étudiants ayant participé à la simulation étaient au même niveau d'apprentissage au sein de leur programme et possédaient ainsi un niveau similaire de connaissances sur l'OA. Cette situation a pu occasionner une variabilité plus faible des niveaux de réussite démontrés par les différents participants et affecter négativement le CCI. Afin de prévenir ce problème, certaines études sur la fidélité interjuges de grilles d'observation induisent une variabilité accrue en évaluant un scénario de simulation réalisé par des étudiants, puis un réalisé par des experts (Reid *et al.*, 2012). Cette méthode assure un intervalle plus grand entre les scores de réussite des différents participants et permet de vérifier que la grille identifie adéquatement les niveaux de compétence élevés et faibles. Cela impliquerait, par exemple, de faire une évaluation de fidélité à partir d'une même simulation effectuée par un échantillon qui inclut des étudiants novices dans l'utilisation de l'OA ainsi que des étudiants diplômés ou des professionnels de la santé utilisant l'OA dans leur travail quotidien.

Par ailleurs, la présente évaluation de fidélité incluait seulement 12 des 20 indicateurs de compétence de la grille. Cinq indicateurs (A-4, A-7, D-2, D-3 et E-3) n'ont pas pu être évalués dans le contexte de simulations en ligne parce que, pour ce faire, les évaluateurs devaient vérifier si les étudiants indiquaient la langue officielle préférée du bénéficiaire dans des documents écrits (par ex. sur le formulaire d'accueil ou de référence), et si la langue utilisée dans les ressources informatives remises au bénéficiaire correspond à la préférence de ce dernier. Par ailleurs, l'évaluation des indicateurs D-5 et D-6 requiert un scénario qui amène l'étudiant à veiller à ce que l'affichage et les ressources offertes dans son milieu de travail respectent les principes d'OA. Enfin, l'évaluation de l'indicateur C-1 nécessite quant à elle un scénario incluant un acteur qui présente un niveau de langue ou des expressions et régionalismes particuliers, permettant ainsi à l'étudiant d'ajuster son niveau de langue à celui du bénéficiaire.

Les travaux de notre équipe de recherche ont permis de mettre en lumière que les indicateurs de compétence en OA sont nombreux et variés. Ils portent sur des comportements observables d'OA susceptibles d'être adoptés à toutes les étapes de la prise en charge des bénéficiaires, ainsi qu'à travers la promotion de l'OA dans le milieu de travail du professionnel de la santé ou des services sociaux. Puisqu'une majorité des simulations portant sur la compétence en OA comprennent également une ou plusieurs composantes cliniques liées au domaine spécifique du futur professionnel, il peut être complexe d'inclure l'entièreté des indicateurs de la Grille d'observation des indicateurs de compétence en OA dans une seule simulation. De ce fait, les programmes de formation cherchant à enseigner et à évaluer l'OA par le biais de simulations sont invités à utiliser

la grille pour élaborer plusieurs scénarios de simulations complémentaires permettant aux étudiants de développer leur compétence en offre active de façon progressive et à différents niveaux de difficulté.

3.1. Perspectives d'avenir

Notre équipe de recherche, en collaboration avec divers programmes au sein des facultés des sciences de la santé, de médecine et de travail social à l'Université d'Ottawa, travaille actuellement à la création et à la mise en place d'une gamme de scénarios de simulations interprofessionnelles portant sur la compétence en OA. Ces activités d'apprentissage ont pour but d'améliorer la formation des futurs professionnels en leur permettant de pratiquer graduellement tous les comportements observables d'OA dans des contextes variés et interdisciplinaires. La mise en oeuvre de ces nouveaux scénarios de simulations facilitera également la suite de l'évaluation métrologique de la Grille d'observation des indicateurs de compétence en OA en permettant l'analyse de l'entièreté de ses indicateurs de compétence dans des programmes de formation variés et à différents niveaux d'apprentissage. De plus, il sera possible de vérifier si les modifications apportées à la grille et à son guide d'accompagnement à la suite de l'analyse des résultats de fidélité ont permis d'améliorer la fidélité interjuges des indicateurs ayant obtenu une valeur de CCI plus faible dans la présente étude. D'autres initiatives visent présentement à adapter cet outil afin de faciliter son usage dans les milieux professionnels, notamment dans un contexte de stages ou de formation continue des professionnels de la santé et des services sociaux. De plus, une adaptation de la grille est prévue pour définir les comportements d'OA qui pourraient être manifestés par des étudiants ou des intervenants unilingues anglophones. Cette grille adaptée pourra être utilisée pour la formation et l'évaluation de la compétence de ces derniers. Enfin, de futurs projets pourraient s'inspirer de la grille pour la création d'outils éducatifs sur l'OA de services en français visant d'autres domaines de services gouvernementaux.

Conclusion

L'OA est une compétence complexe qui gagne à être enseignée par le biais d'une combinaison d'activités d'apprentissage théoriques et pratiques pour sensibiliser davantage les futurs professionnels à l'importance des obstacles auxquels sont confrontées les communautés francophones en situation minoritaire. Les simulations offrent aux étudiants la chance de mettre en pratique leurs connaissances au moyen de divers scénarios qui reflètent des situations réelles vécues par les membres des communautés linguistiques en situation minoritaire dans le milieu de la santé et des services sociaux. Afin d'optimiser et d'encadrer l'apprentissage par simulation, il est nécessaire d'utiliser des outils d'évaluation validés permettant d'établir des attentes claires et d'évaluer les étudiants de façon équitable.

La Grille d'observation des indicateurs de compétence en OA a pour but de guider la création d'activités d'apprentissage expérientielles portant sur les différents comportements observables d'OA, en plus de faciliter une évaluation détaillée et structurée des indicateurs de compétence en OA. Ce projet visait à approfondir l'évaluation métrologique de la grille en analysant sa fidélité interjuges à l'aide du CCI. Les résultats démontrent que la grille, dans son ensemble, présente un niveau de fidélité élevé avec un CCI de 0,87.

Des études supplémentaires sont nécessaires pour évaluer la grille d'observation dans différents contextes et pour évaluer l'entièreté des indicateurs de compétence. Néanmoins, les résultats de fidélité, combinés à la validation de contenu effectuée précédemment, permettent d'encourager l'utilisation de cet outil pour l'évaluation formative et sommative de la compétence en OA dans le cadre de simulations.

La formation des futurs professionnels sur la sécurisation linguistique et la compétence en OA est primordiale afin d'assurer l'amélioration continue de l'accès à des soins et à des services en français de qualité pour les communautés francophones minoritaires. L'enseignement et la valorisation de l'OA dans les programmes de formation appuient le développement d'une génération de professionnels qui sont conscients des difficultés particulières vécues par les minorités de langue officielle dans le milieu de la santé et des services sociaux, compétents dans la pratique de l'OA, ainsi que prêts à soutenir et à encourager les initiatives liées à une implantation élargie de l'OA dans leurs futurs milieux de travail.

Financement

Ce projet a été financé par le Consortium national de formation en santé (CNFS) – Volet Université d'Ottawa. A. Bodnaruc a bénéficié de la bourse doctorale en nutrition et en santé mentale (2019-2022) de l'École des sciences de la nutrition. C. Vincent a bénéficié d'une bourse estivale (2020) d'initiation à la recherche du CNFS – Volet Université d'Ottawa, une bourse de formation de maîtrise du Fonds de recherche du Québec – Santé (FRQS), ainsi que la Bourse des études supérieures de l'Ontario (BÉSO).

Bibliographie

- Adamson, K. A., Gubrud, P., Sideras, S. et Lasater, K. (2012). Assessing the reliability, validity, and use of the Lasater Clinical Judgment Rubric: Three approaches. *Journal of Nursing Education*, 51(2), 66-73. (<https://doi.org/10.3928/01484834-20111130-03>)
- Bouchard, L., Beaulieu, M. et Desmeules M. (2012). L'offre active de services de santé en français en Ontario : une mesure d'équité. *Reflète*, 18(2), 38-65. (<https://doi.org/10.7202/1013173ar>)
- Bouchard, L. et Desmeules, M. (2011). *Minorités de langue officielle du Canada : égales devant la santé?* Presses de l'Université du Québec.
- Bouchard, L., et Desmeules, M. (2013). Les minorités linguistiques du Canada et la santé. *Healthcare policy = Politiques de santé*, 9, 38-47. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4750144/#> (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4750144/>)
- Bouchard, P. et Vézina, S. (2009). *L'outillage des étudiants et des nouveaux professionnels : un levier essentiel à l'amélioration des services de santé en français* [rapport de recherche]. Consortium national de formation en santé.
- Bowen, S. (2001). *Barrières linguistiques dans l'accès aux soins de santé*. Santé Canada. (https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/hcs-sss/alt_formats/hpb-dgps/pdf/pubs/2001-lang-acces/2001-lang-acces-fra.pdf)
- Bowen, S. (2015). *Impact des barrières linguistiques sur la sécurité des patients et la qualité des soins*. Société Santé en français. (<https://reseauantene.ca/wp-content/uploads/2018/05/impacts-barrieres-linguistiques.pdf>)
- Casimiro, L., Savard, J., Sauvé-Schenk, K. et Atchessi, N. (2018). Questionnaire d'aiguillage pour dépister les besoins de formation quant à l'offre active de services en français. *Reflète*, 24(2), 182-211. (<https://doi.org/10.7202/1053868ar>)
- Castorr, A. H., Thompson, K. O., Ryan, J. W., Phillips, C. Y., Prescott, P. A. et Soeken, K. L. (1990). The process of rater training for observational instruments: Implications for interrater reliability. *Research in Nursing and Health*, 13(5), 311-318. (<https://doi.org/10.1002/nur.4770130507>)

- Charbonneau, F. (2011). Dans la langue officielle de son choix : la loi canadienne sur les langues officielles et la notion de « choix » en matière de services publics. *Lien social et Politiques*, 66, 39-63. (<https://doi.org/10.7202/1008872ar>)
- Cross, V., Hicks, C. et Barwell, F. (2001). Exploring the gap between evidence and judgement: Using video vignettes for practice-based assessment of physiotherapy undergraduates. *Assessment and Evaluation in Higher Education*, 26(3), 189-212. (<https://doi.org/10.1080/02602930120052369>)
- de Moissac, D. et Bowen, S. (2019). Impact of language barriers on quality of care and patient safety for official language minority Francophones in Canada. *Journal of Patient Experience*, 6(1), 24-32. (<https://doi.org/10.1177/2374373518769008>)
- Deveau K., Landry, K. et Allard, R. (2010). *Utilisation des services gouvernementaux en français : une étude auprès des Acadiens et francophones de la Nouvelle-Écosse sur les facteurs associés à l'utilisation des services gouvernementaux en français* [rapport de recherche]. Institut canadien de recherche sur les minorités linguistiques. (https://icrml.ca/images/stories/documents/fr/sommaire_deveau_services_gouv.pdf)
- Downing S. M. (2004). Reliability: On the reproducibility of assessment data. *Medical Education*, 38(9), 1006-1012. (<https://doi.org/10.1111/j.1365-2929.2004.01932.x>)
- Drolet, M., Bouchard, P., Savard, J. et Laforge, M.-J. (2017). Problématique générale : enjeux de l'accessibilité et de l'offre active de services sociaux et de santé au sein de la francophonie canadienne en situation minoritaire. Dans M. Drolet, P. Bouchard et J. Savard (dir.), *Accessibilité et offre active : santé et services sociaux en contexte linguistique minoritaire* (p. 13-32). Presses de l'Université d'Ottawa. (<https://doi.org/10.2307/j.ctv5vdc0.6>)
- Drolet, M., Savard, J., Benoît, J., Arcand, I., Savard, S., Lagacé, J., Lauzon, S. et Dubouloz, C.-J. (2014). Health services for linguistic minorities in a bilingual setting: Challenges for bilingual professionals. *Qualitative Health Research*, 24(3), 295-305. (<https://doi.org/10.1177/1049732314523503>)
- Dubouloz, C.-J., Benoît, J., Guitard, P., Brosseau, L., Savard, J., Kubina, L.-A. et Drolet, M. (2014). Proposition de lignes directrices pour la formation à l'offre active des futurs professionnelles et professionnels en santé et en service social oeuvrant en situation francophone minoritaire. *Reflets*, 20(2), 123-151. (<https://doi.org/10.7202/1027588ar>)
- Duchesne, C., Sauvé-Schenk, K., Benoît, J. et Dubouloz, C.-J. (2019). Élaboration, implantation et évaluation d'une formation à l'enseignement de l'offre active pour les professeurs et formateurs d'université. *Minorités linguistiques et société = Linguistic Minorities and Society*, 11, 29-51. (<https://doi.org/10.7202/1065211ar>)
- Dunbar, N. E., Brooks, C. F. et Kubicka-Miller, T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 31(2), 115-128. (<https://doi.org/10.1007/s10755-006-9012-x>)
- Hallgren K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34. (<https://doi.org/10.20982/tqmp.08.1.p023>)
- Keshmiri, F., Ponzer, S., Sohrabpour, A., Farahmand, S., Shahi, F., Bagheri-Hariri, S., Soltani-Arabshahi, K., Shirazi, M. et Masiello, I. (2016). Contextualization and validation of the interprofessional collaborator assessment rubric (ICAR) through simulation: Pilot investigation. *Medical Journal of the Islamic Republic of Iran*, 30, 1-8. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5038989/>)
- Koo, T. K. et Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. (<https://doi.org/10.1016/j.jcm.2016.02.012>)
- Lewis, L. K., Stiller, K. et Hardy, F. (2008). A clinical assessment tool used for physiotherapy students: is it reliable? *Physiotherapy Theory and Practice*, 24(2), 121-134. (<https://doi.org/10.1080/09593980701508894>)

- Lucas, C., Bosnic-Anticevich, S., Schneider, C. R., Bartimote-Aufflick, K., McEntee, M. et Smith, L. (2017). Inter-rater reliability of a reflective rubric to assess pharmacy students' reflective thinking. *Currents in Pharmacy Teaching and Learning*, 9(6), 989-995. (<https://doi.org/10.1016/j.cptl.2017.07.025>)
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382-385. (<https://doi.org/10.1097/00006199-198611000-00017>)
- Moskal, B. M. et Leydens, J. A. (2000) Scoring rubric development: Validity and reliability. *Practical Assessment, Research, and Evaluation*, 7, 1-6. (<https://doi.org/10.7275/q7rm-gg74>)
- Reid, J., Stone, K., Brown, J., Caglar, D., Kobayashi, A., Lewis-Newby, M., Partridge, R., Seidel, K. et Quan, L. (2012). The Simulation Team Assessment Tool (STAT): Development, reliability and validation. *Resuscitation*, 83(7), 879-886. (<https://doi.org/10.1016/j.resuscitation.2011.12.012>)
- Santé Canada. (2020, 31 mars). *Perceptions des communautés de langue officielle en situation minoritaire : accès aux services de santé dans la langue officielle de son choix* [rapport final]. (<https://publications.gc.ca/site/fra/9.886122/publication.html>)
- Savard, J., Casimiro, L., Bouchard, P., Benoît, J., Drolet, M. et Dubouloz, C.-J. (2015). Conception d'outils de mesure de l'offre active de services sociaux et de santé en français en contexte minoritaire. *Minorités linguistiques et société = Linguistic Minorities and Society*, 6, 131-156. (<https://doi.org/10.7202/1033193ar>)
- Statistique Canada. (2022a, 17 août). *Première langue officielle parlée selon la langue parlée le plus souvent à la maison, les autres langues parlées régulièrement à la maison et la connaissance des langues officielles : Canada, provinces et territoires, divisions de recensement et subdivisions de recensement* (tableau). (<https://doi.org/10.25318/9810019101-fra>)
- Statistique Canada. (2022b, 17 août). Tableau 5 : Première langue officielle parlée, provinces et territoires, 2016 et 2021. *Le Quotidien*. (<https://www150.statcan.gc.ca/n1/daily-quotidien/220817/t005a-fra.htm>)
- Tsingos, C., Bosnic-Anticevich, S. et Smith, L. (2015). Learning styles and approaches: Can reflective strategies encourage deep learning? *Currents in Pharmacy Teaching and Learning*, 7(4), 492-504. (<https://doi.org/10.1016/j.cptl.2015.04.006>)
- Vincent, C., Bodnaruc, A. M., Kengneson, C.-C., Savard, J. et Giroux, I. (2021). Création d'une grille d'observation pour simulations sur l'offre active de services de santé en français. *Reflets*, 27(2), 150-164. (<https://doi.org/10.7202/1093102ar>)
- Walter, S. D., Eliasziw, M. et Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine*, 17(1), 101-110. ([https://doi.org/10.1002/\(sici\)1097-0258\(19980115\)17:1<101::aid-sim727>3.0.co;2-e](https://doi.org/10.1002/(sici)1097-0258(19980115)17:1<101::aid-sim727>3.0.co;2-e))
- Williams, R. G., Klamen, D. A. et McGaghie, W. C. (2003). Cognitive, social and environmental sources of bias in clinical performance ratings. *Teaching and Learning in Medicine*, 15(4), 270-292. (https://doi.org/10.1207/S15328015TLM1504_11)