

Traduction, adaptation et évaluation psychométrique préliminaire d'une mesure d'engagement et d'une mesure de charge cognitive en contexte d'apprentissage numérique

Translation, adaptation and preliminary psychometric evaluation of a measure of engagement and a measure of cognitive load in a digital learning context

Guillaume FONTAINE^{1,2*}, Sylvie COSSETTE^{1,2}, Marc-André MAHEU-CADOTTE^{1,2,3}, Tanya MAILHOT⁴, Patrick LAVOIE^{1,2,5}, Marie-Pierre GAGNON^{6,7}, Véronique DUBÉ^{1,3} et José CÔTÉ^{1,3}

¹ Faculté des sciences infirmières, Université de Montréal, Montréal, Québec, Canada

² Centre de recherche, Institut de Cardiologie de Montréal, Montréal, Québec, Canada

³ Centre de recherche, Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada

⁴ Bouvé College of Health Sciences, Northeastern University, Boston, Massachusetts, USA

⁵ Centre d'innovation en formation infirmière, Université de Montréal, Montréal, Québec, Canada

⁶ Faculté des sciences infirmières, Université Laval, Québec, Québec, Canada

⁷ Centre de recherche, CHU de Québec, Québec, Québec, Canada

*Correspondance et offprints : Guillaume FONTAINE, Faculté des sciences infirmières, Université de Montréal et Centre de recherche de l'Institut de Cardiologie de Montréal, Bureau S-2490, 5000 rue Bélanger, Montréal, Québec, Canada, H1T 1C8. Courriel : guillaume.fontaine@umontreal.ca

Contribution réelle et respective de chaque auteur : Guillaume Fontaine a contribué à la conception du protocole de recherche, au recueil des données, aux analyses statistiques, à l'interprétation des résultats, ainsi qu'à l'écriture, la révision et l'approbation du manuscrit. Sylvie Cossette a contribué à la conception du protocole de recherche, aux analyses statistiques, à l'interprétation des résultats, ainsi qu'à l'écriture, la révision et l'approbation du manuscrit. Marc-André Maheu-Cadotte, Tanya Mailhot, Patrick Lavoie, Marie-Pierre Gagnon, Véronique Dubé et José Côté ont contribué à l'interprétation des résultats ainsi qu'à l'écriture, la révision et l'approbation du manuscrit.

Résumé (Français, 250 mots)

Contexte : L'apprentissage numérique chez les professionnels de la santé amène des défis comme le manque d'engagement et l'excès de charge cognitive. Toutefois, il n'existe pas de mesures d'engagement et de charge cognitive validées en français. De telles mesures permettraient de considérer ces variables dans le développement de formations numériques et l'évaluation de leurs effets. **Buts :** 1) Traduire et adapter la *User Engagement Scale – Short Form* et le *Cognitive Load Index* en français ; 2) Effectuer une évaluation psychométrique préliminaire des échelles afin d'examiner leurs caractéristiques, les corrélations entre les sous-échelles et la cohérence interne. **Méthode :** La traduction et l'adaptation ont été réalisées selon Sousa et Rojjanasrirat (2011). Pour l'évaluation des propriétés psychométriques, des étudiants en sciences infirmières ont répondu en ligne à l'Échelle d'engagement de l'utilisateur – Forme abrégée (EEU-FA) et à l'Indice de charge cognitive (IDCC). Les données ont été analysées par des statistiques descriptives, des coefficients de corrélation de Pearson, des coefficients omega (ω) de McDonald et des coefficients alpha (α) de Cronbach. **Résultats :** Cinquante-sept participants ont complété l'étude. La majorité des distributions étaient normales. Les corrélations entre les sous-échelles étaient entre 0,03 et 0,64 pour l'EEU-FA, et entre -0,12 à 0,41 pour l'IDCC. Les coefficients ω pour l'EEU-FA (0,77 à 0,89) et α pour l'IDCC (0,70 à 0,96) démontrent une bonne cohérence interne. **Conclusions :** Cette étude a permis de traduire et d'adapter en français des mesures d'engagement et de charge cognitive en contexte numérique. Les échelles font état de bonnes qualités psychométriques préliminaires.

Mots clés : apprentissage numérique ; engagement ; charge cognitive ; échelle de mesure ; évaluation psychométrique.

Résumé (Anglais, 250 mots)

Context: Digital learning in health professionals generates challenges such as the lack of engagement and the excess of cognitive load. However, there are no measures of engagement and cognitive load validated in French. Such measures would allow these variables to be considered in the development of digital training programs and in the evaluation of their effects. **Goals:** 1) Translate and adapt the User Engagement Scale – Short Form and the Cognitive Load Index in French; 2) Conduct a preliminary psychometric evaluation of translated scales to examine the characteristics of items, the correlations between subscales, and internal consistency. **Method:** Translation and adaptation were performed according to the method of Sousa and Rojjanasrirat (2011). For the evaluation of psychometric properties, nursing students responded online to the *Échelle d'engagement de l'utilisateur – Forme abrégée* (EEU-FA) and the *Indice de charge cognitive* (IDCC). The results were analyzed using descriptive statistics, Pearson's correlation coefficients, McDonald's omega (ω) coefficients and Cronbach's (α) alpha coefficients. **Results:** Fifty-seven participants completed the study. The majority of the distributions were normal. The correlations between the subscales ranged from 0.03 to 0.64 for the EEU-FA, and from -0.12 to 0.41 for the IDCC. McDonald's ω coefficients ranged from 0.76 to 0.89 for the EEU-FA, and Cronbach's α coefficients ranged from 0.70 to 0.96 for the IDCC, demonstrating good internal consistency. **Conclusions:** This study translated and adapted into French a measure of engagement and cognitive load in a digital context. Both measurement scales show satisfying preliminary psychometric qualities.

Key words: digital learning; engagement; cognitive load; measurement scale; psychometric evaluation.

Introduction

La formation des professionnels de la santé intègre de façon croissante les technologies de l'information et de la communication [1]. Les environnements numériques d'apprentissage utilisent ces technologies afin d'offrir un accès accru au contenu éducatif et de soutenir l'apprentissage [2]. Les environnements numériques d'apprentissage présentent plusieurs avantages potentiels comparativement à la formation en présentiel, dont une augmentation de l'accessibilité. De plus, leur efficacité apparaît supérieure à celle de la formation en présentiel [1]. Toutefois, face aux progrès des dernières décennies dans le domaine des environnements numériques d'apprentissage pour la formation des professionnels de la santé, certains défis émergent. En effet, au même titre que l'interaction apprenant-enseignant en classe, l'interaction apprenant-ordinateur est modérée par des processus cognitifs complexes qui influencent l'apprentissage [3]. À ce titre, deux processus cognitifs complexes ont été étudiés au fil des dernières décennies afin d'optimiser l'apprentissage : l'engagement et la charge cognitive.

L'engagement est le degré de l'investissement de l'utilisateur lorsqu'il interagit avec un environnement numérique [3]. Le concept d'engagement n'est pas exclusif au domaine de l'apprentissage en contexte numérique; il est étudié, entre autres, dans les domaines du marketing, de la e-santé, des réseaux sociaux et des jeux vidéo [3]. Dans tous les cas, on vise à rendre les environnements numériques intéressants, conviviaux, visuellement plaisants et gratifiants afin de soutenir l'engagement de l'utilisateur. L'engagement est un concept bidimensionnel. La première dimension est *expérientielle*, soit l'expérience subjective relative à l'utilisation de l'environnement numérique caractérisée par le niveau d'attention, la convivialité perçue, l'attrait esthétique et la gratification [4]. La deuxième dimension est *comportementale*, soit le degré (p. ex. fréquence, durée) d'utilisation de l'environnement numérique [5]. Les échelles de mesure et les entretiens individuels figurent parmi les stratégies les plus utilisées pour mesurer la dimension expérientielle. Les métriques telles que le nombre de pages visitées et la durée de navigation sont des mesures de la dimension comportementale [6]. En somme, la mesure de l'engagement peut être utile en contexte de formation numérique afin de mieux calibrer la forme, le contenu et le dosage de la formation et en contexte de recherche pour contrôler l'impact de cette variable modératrice sur l'apprentissage [4].

À notre connaissance, il n'existe pas d'échelle de mesure validée en français permettant de mesurer la dimension expérientielle de l'engagement en contexte d'apprentissage numérique. À ce titre, la *User Engagement Scale* est une échelle de mesure auto-rapportée proposée en 2010 par Heather L. O'Brien et Elaine G. Toms et qui a été largement utilisée pour mesurer la dimension expérientielle de l'engagement en contexte numérique [7]. L'échelle a été développée à partir d'entretiens individuels et d'une revue des écrits, puis testée auprès de trois échantillons (N = 427, N = 779, N = 344) d'utilisateurs naviguant sur un environnement numérique de commerce en ligne [4]. La version courte de l'échelle, nommée *User Engagement Scale – Short Form*, a été publiée en 2018 et est particulièrement pertinente en contexte de recherche vu la facilité d'y répondre, la réduction du risque de réponses manquantes et ses très bonnes qualités psychométriques [4].

La *charge cognitive* renvoie au degré auquel la mémoire de travail est sollicitée chez l'apprenant [8]. La mémoire de travail ne peut traiter qu'un nombre limité d'éléments d'information durant une période déterminée. Cela explique en partie pourquoi, lors d'une formation, certains éléments d'information font l'objet de peu d'attention. Il existe trois types de charges cognitives

qui s'additionnent et composent la charge cognitive globale [8] : les charges intrinsèque, extrinsèque et essentielle. La *charge intrinsèque* est déterminée par quatre facteurs : 1) la complexité de la tâche d'apprentissage, 2) le nombre d'éléments d'information présenté à un moment donné, 3) le degré auquel ces éléments interagissent entre eux et 4) le niveau de familiarité de l'apprenant avec les éléments présentés [8, 9]. En effet, la charge intrinsèque augmente en fonction de la complexité de la tâche d'apprentissage pour l'apprenant (p. ex. : une tâche touchant un sujet plus avancé sera plus complexe). De plus, elle augmente en fonction du nombre d'éléments d'information présentés (p. ex., l'apprentissage de trois facteurs de risque cardiovasculaire est plus exigeant que l'apprentissage d'un seul). Enfin, la charge intrinsèque augmente lorsque l'interactivité entre les éléments d'information augmente (p. ex. : l'apprentissage des liens physiopathologiques entre le tabagisme et l'hypertension artérielle est plus ardu que l'apprentissage du fait que le tabagisme est un facteur de risque). Toutefois, le niveau de familiarité de l'apprenant peut pallier le haut nombre d'éléments d'information et l'interactivité entre ceux-ci. La *charge extrinsèque*, quant à elle, augmente à mesure que des éléments superflus s'ajoutent à la tâche d'apprentissage. Ces éléments peuvent être des distractions, des explications peu vulgarisées, des informations superflues ou des instructions difficiles à suivre. De plus, la charge extrinsèque augmente lorsque l'information présentée excède la capacité des canaux de traitement visuel et auditif [9]. Enfin, la *charge essentielle* représente la charge imposée par les ressources cognitives nécessaires à l'apprentissage, comme l'élaboration et l'automatisation des scripts mentaux—des réseaux de connaissances hautement développés et organisés dans la mémoire à long terme [9]. La charge essentielle est régulée par l'apprenant et s'additionne aux charges intrinsèque et extrinsèque. En résumé, si la tâche d'apprentissage est trop complexe (charge intrinsèque) et les explications ou instructions sont difficiles à suivre (charge extrinsèque), il est probable que l'apprenant n'ait plus suffisamment de ressources cognitives dans sa mémoire de travail pour l'élaboration et l'automatisation des scripts mentaux (charge essentielle).

À notre connaissance, il n'existe pas d'échelle de mesure validée en français permettant de mesurer les différents types de charge cognitive en contexte d'apprentissage numérique. À ce titre, le *Cognitive Load Index*, proposé par Leppink en 2013 [10], est l'une des premières échelles de mesure auto-rapportées permettant de mesurer les trois types de charge cognitive en contexte d'apprentissage. Le *Cognitive Load Index* peut être utilisé par les formateurs désirant évaluer une formation offerte en présentiel ou via un environnement numérique d'apprentissage, et pour mesurer la charge cognitive des apprenants pendant ou après une formation en contexte de recherche. La capacité du *Cognitive Load Index* de mesurer à la fois les charges intrinsèque, extrinsèque et essentielle est plus pertinente que les mesures globales de la charge cognitive (p. ex. : Paas [11] et le NASA Task Load Index [12]). En effet, la mesure des différents types de charge cognitive permet de mieux comprendre, au regard de la complexité du contenu éducatif, des directives et des explications, comment une formation engendre différents degrés d'apprentissage. Prenons l'exemple de deux apprenants pour qui la formation engendre une même charge cognitive globale. Chez l'un d'eux, la charge cognitive globale peut résulter du cumul d'une charge intrinsèque plus basse, vu un niveau d'expertise élevé, et d'une charge extrinsèque plus élevée, vu la non-familiarité avec la manière dont les instructions sont formulées. Chez l'autre apprenant, l'inverse peut se produire avec une charge intrinsèque élevée et une charge extrinsèque basse. Par conséquent, la distinction des différents types de charge cognitive offre des avenues pour arrimer la formation aux caractéristiques des apprenants. Bien que certains auteurs se soient penchés sur la mesure des différents types de charge cognitive [13], le *Cognitive Load Index* représente l'échelle

de mesure la plus complète au plan des trois sous-échelles et la plus robuste au plan psychométrique.

En somme, l'engagement et la charge cognitive sont deux concepts importants à évaluer en contexte d'apprentissage numérique. Plus particulièrement, la *User Engagement Scale – Short Form* et le *Cognitive Load Index* pourront être utilisées auprès d'étudiants et de professionnels dans le domaine de la santé, en enseignement et en recherche, afin d'évaluer leur engagement et leur charge cognitive en lien avec l'utilisation d'un environnement numérique d'apprentissage. Plus particulièrement, ces deux échelles ont un potentiel intéressant pour déterminer des cibles actionnables (p. ex. convivialité, attrait esthétique, complexité de la tâche d'apprentissage, instructions et explications) pouvant être utilisées afin d'optimiser les formations dispensées en contexte numérique.

Nous préparons actuellement une étude expérimentale dans laquelle les versions françaises de ces mesures seront utilisées pour évaluer l'engagement et la charge cognitive chez le personnel infirmier suite à une formation suivie par le biais d'un environnement numérique d'apprentissage [14-17]. C'est dans ce contexte que nous avons élaboré la présente étude, qui visait à :

1. Traduire et adapter la *User Engagement Scale – Short Form* et le *Cognitive Load Index* en français ;
2. Effectuer une évaluation psychométrique préliminaire des deux échelles de mesure traduites et adaptées en français afin d'examiner :
 - a. Les caractéristiques des items ;
 - b. Les corrélations entre les sous-échelles ;
 - c. La cohérence interne des sous-échelles.

Matériel et méthodes

Plan d'étude et cadre méthodologique

Cette étude méthodologique transversale était basée sur la méthodologie de Sousa and Rojjanasrirat [18] pour la traduction, l'adaptation culturelle et la validation des instruments de mesure dans le domaine de la santé.

Description de la *User Engagement Scale – Short Form*

La *User Engagement Scale – Short Form* comprend 12 items répartis au sein de quatre sous-échelles afin de mesurer les attributs de l'engagement en contexte numérique : 1) attention soutenue (*Focused Attention*, 3 items) ; 2) convivialité perçue (*Perceived Usability*, 3 items) ; 3) attrait esthétique (*Aesthetic Appeal*, 3 items) ; et 4) gratification (*Reward*, 3 items) [4]. Les coefficients omega (ω)—un coefficient permettant de mesurer la cohérence interne d'échelles au même titre que le coefficient alpha (α) de Cronbach mais qui permet de contrer certaines limites inhérentes au coefficient α [19]—varient de 0,75 à 0,82 pour la sous-échelle « attention soutenue », de 0,70 à 0,86 pour la sous-échelle « convivialité perçue », de 0,84 à 0,88 pour la sous-échelle « attrait esthétique » et de 0,79 à 0,81 pour la sous-échelle « gratification ». L'échelle de réponse à chacun des items est de type Likert à cinq niveaux (de 1 [fortement en désaccord] à 5 [fortement en accord]). Les items de la sous-échelle « Convivialité perçue » doivent être codés à l'inverse avant de calculer les scores (c.-à-d., 1 → 5, 2 → 4, 3 → 3, 4 → 2, 5 → 1). Les scores de chacune

des quatre sous-échelles peuvent être calculés en additionnant les valeurs des réponses pour les trois items contenus dans chaque sous-échelle et en les divisant par trois. Un score global d'engagement (sur 5) peut être calculé en additionnant tous les items et en divisant par douze [4]. Plus les scores de chaque dimension et le score global de l'échelle sont élevés, plus les utilisateurs sont engagés dans l'interaction avec l'environnement numérique.

Description du *Cognitive Load Index*

Le *Cognitive Load Index* comprend 10 items répartis en trois sous-échelles afin de mesurer les différents types de charge cognitive : 1) charge intrinsèque (*Intrinsic Load*, 3 items) ; 2) charge extrinsèque (*Extrinsic Load*, 3 items) et 3) charge essentielle (*Germane Load*, 4 items) [10]. Le *Cognitive Load Index* a été validé en langue anglaise auprès d'étudiants au baccalauréat en psychologie et dans le domaine de la santé avec une très bonne cohérence interne pour ses trois sous-échelles : les coefficients α de Cronbach étaient de 0,81 à 0,88 pour la sous-échelle « Charge intrinsèque », de 0,80 à 0,85 pour la sous-échelle « Charge extrinsèque » et de 0,89 à 0,93 pour la sous-échelle « Charge essentielle » [10]. L'échelle de réponse à chacun des items est de type Likert à 11 niveaux (de 0 [pas du tout en accord] à 10 [tout à fait en accord]). Les scores des sous-échelles peuvent être calculés en additionnant les valeurs des réponses aux items contenus dans une sous-échelle et en divisant le total par le nombre d'items de la sous-échelle [10]. Plus le score de charge intrinsèque est élevé, plus la tâche d'apprentissage est perçue comme complexe par les apprenants. Plus le score de charge extrinsèque est élevé, plus les instructions et les explications lors de la formation sont perçues comme inefficaces par les apprenants. Enfin, plus le score de charge essentielle est élevé, plus l'apprenant perçoit que l'activité a amélioré sa connaissance et sa compréhension des concepts et des sujets abordés. Aucun score global n'est calculé dans cette échelle de mesure.

Traduction et adaptation des échelles de mesure

La traduction et l'adaptation en français de la *User Engagement Scale – Short Form* et du *Cognitive Load Index* ont été réalisées en quatre étapes préalablement à l'évaluation psychométrique ([Figure 1](#)). La traduction et l'adaptation ont été réalisées par deux doctorants et une postdoctorante bilingues dont la langue maternelle est le français (traducteurs A, B et C), deux traducteurs professionnels d'une firme externe engagée pour l'étude (traducteurs D et E) et une chercheuse senior avec une expertise en développement et en validation d'outils de mesure (SC).

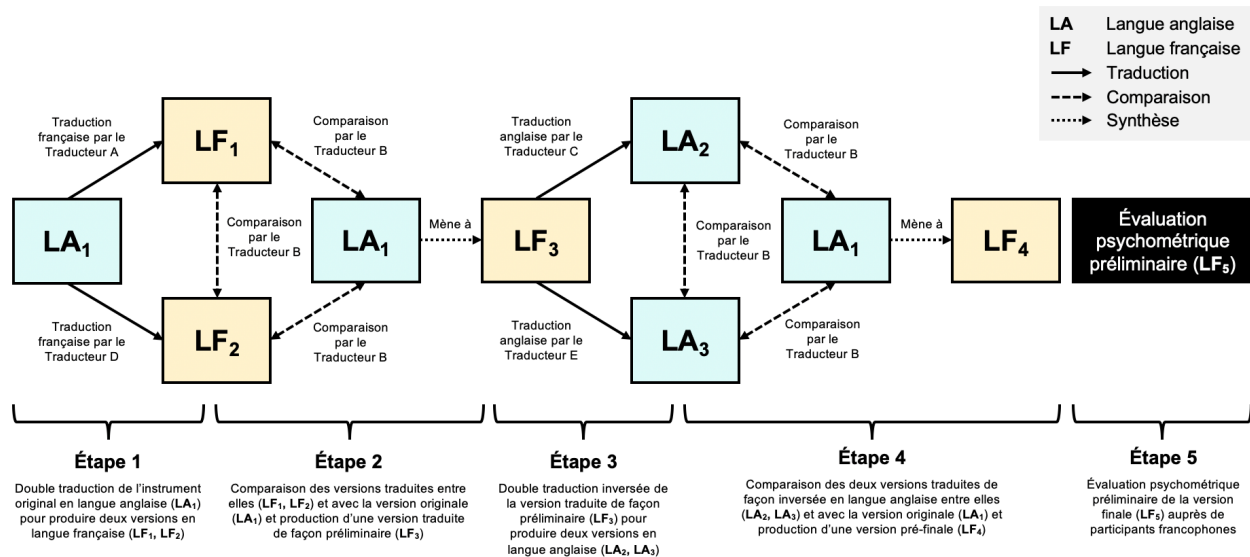


Figure 1. Étapes de la présente étude.
Abréviations : LA = langue anglaise ; LF = langue française.

La démarche de traduction et d'adaptation a été réalisée de façon indépendante, mais simultanée, pour les deux échelles de mesure. Chaque échelle a fait l'objet d'une double traduction vers le français (étape 1). Les deux versions de chaque échelle ont été comparées entre elles afin de produire une version traduite de façon préliminaire en langue française (étape 2). Ces versions traduites de façon préliminaire en langue française ont été retraduites vers l'anglais par deux traducteurs indépendants (étape 3). Les versions retraduites en langue anglaise ont été comparées entre elles et à leur version originale afin de produire une version pré-finale (étape 4). Toute ambiguïté ou divergence a été discutée et résolue par le biais d'un consensus. Du point de vue de l'adaptation, les versions pré-finales de chaque échelle de mesure ont été revues pour vérifier si la formulation de chaque item représentait bien l'idée originale de l'item. Des ajustements ont été apportés pour produire une version finale qui fut soumise à une évaluation psychométrique préliminaire.

Évaluation psychométrique préliminaire

L'objectif de ce volet de l'étude était d'examiner les distributions des scores, les corrélations entre les sous-échelles et la cohérence interne des versions traduites de la *User Engagement Scale – Short Form* et du *Cognitive Load Index*. Nous soulignons le caractère *préliminaire* de cette évaluation psychométrique des deux échelles, qui ne visait qu'à obtenir des résultats à titre indicatif.

Échantillon et milieu

Afin d'effectuer une évaluation psychométrique préliminaire des versions en langue française des deux échelles de mesure, nous avons ciblé des étudiants en sciences infirmières d'une université francophone québécoise. Pour déterminer la taille d'échantillon, nous nous sommes basés sur les lignes de conduite en matière d'examen psychométrique d'échelles de mesure. Il est recommandé d'avoir autour de 5 participants par item lors d'une évaluation psychométrique préliminaire [18]. Une taille d'échantillon de 50 participants a ainsi été retenue.

Critères d'inclusion et d'exclusion

Nous avons inclus les participants s'ils : 1) étaient inscrits à temps plein ou à temps partiel à un programme de 1^{er}, 2^{ième} ou 3^{ième} cycle en sciences infirmières dans l'université québécoise retenue par l'étude ; 2) étaient en mesure de lire et comprendre le français ; et 3) avaient accès à un appareil électronique connecté à Internet. Il n'y avait aucun critère d'exclusion.

Procédure

Les auteurs des versions originales des deux échelles de mesure (O'Brien, Leppink) ont été contactés en avril 2019 afin d'obtenir la permission de procéder à la traduction et à l'adaptation des échelles de mesure. Les deux auteurs ont donné leur autorisation.

Nous avons recruté les participants par courriel en juin 2019 via les différentes listes de diffusion de l'établissement après approbation éthique et institutionnelle. L'étude s'est déroulée entièrement en ligne. Dans le courriel d'invitation, les participants potentiels ont été invités à lire le formulaire de consentement sur la plateforme *SurveyMonkey* (www.surveymonkey.com) et à cocher si oui ou non ils donnaient leur consentement à participer. La participation à l'étude était volontaire et anonyme ; aucune donnée sociodémographique n'a été collectée. Les participants ont ensuite navigué pendant 10 à 15 minutes sur un site Internet public abordant des sujets relatifs à la santé cardiovasculaire. Nous avons choisi ce site Internet puisqu'il aborde les mêmes sujets qui sont abordés dans une formation numérique que notre équipe déploiera et évaluera sous peu. Par ailleurs, les éléments multimédia (p. ex. graphiques, images, texte, vidéos) employés sur ce site Internet sont variés et correspondent à ceux d'une formation en ligne. Une fois la période de navigation terminée, les participants ont été invités à retourner sur la plateforme *SurveyMonkey* pour compléter les versions traduites en langue française des deux échelles de mesure. La participation à l'étude se terminait après cette étape.

Analyses statistiques

Les analyses statistiques ont été réalisées avec la taille d'échantillon minimale recommandée dans la littérature pour rapporter des valeurs de psychométrie, qui devront être revalidées avec un échantillon plus grand pour effectuer des inférences de plus haut niveau. Par conséquent, les résultats obtenus ne sont que préliminaires.

Toutes les analyses statistiques, hormis le calcul des coefficients omega (ω) de McDonald, ont été réalisées avec le logiciel IBM SPSS Statistics (version 25, IBM Corporation). Nous avons examiné l'ensemble des données afin de repérer les données aberrantes (*outliers*) en utilisant les critères statistiques *Quartile 1 (Q1) - (1,5 x écart interquartile)* pour déterminer le seuil inférieur des valeurs normales et *Quartile 3 (Q3) + (1,5 x écart interquartile)* [20, 21]. Il est suggéré de retirer les données aberrantes dans les cas où celles-ci sont potentiellement attribuables à une erreur de lecture de l'échelle de mesure ou que des données antérieures soutiennent la problématique observée dans le présent échantillon [22, 23].

Nous avons ensuite examiné les scores de chaque échelle de mesure par le biais de statistiques descriptives (c.-à-d., moyenne et écart-type, médiane et écart interquartile) afin d'évaluer si les distributions des scores pour chaque item étaient acceptables. L'utilisation concomitante de la moyenne et de la médiane permet de mieux caractériser lorsque la distribution des scores s'éloigne de la normalité. De plus, afin de décrire la normalité de la distribution des

données collectées, nous avons examiné les coefficients d'asymétrie (*skewness*) et d'aplatissement (*kurtosis*) [22].

Afin d'évaluer les relations entre les sous-échelles propres à chaque échelle de mesure, nous avons examiné les coefficients de corrélation de Pearson. Les corrélations entre les sous-échelles mesurant des concepts reliés au plan théorique devraient être modérées à grandes. À l'inverse, les corrélations entre les sous-échelles mesurant des concepts distincts devraient être faibles [23]. Un seuil de signification alpha de 0,05 a été fixé afin de déterminer les corrélations statistiquement significatives. Nous avons suivi le barème suivant pour l'interprétation des coefficients de corrélation : négligeable ($|r| < 0,1$), faible (0,1–0,29), modérée (0,3–0,49) et grande ($\geq 0,5$) [24].

Enfin, pour décrire la cohérence interne des sous-échelles de chaque échelle de mesure, nous avons examiné les coefficients omega (ω) de McDonald (pour l'EEU-FA) ou les coefficients alpha (α) de Cronbach (pour l>IDCC). Nous avons sélectionné les mêmes coefficients que les auteurs dans les études originales de chaque échelle de mesure afin de pouvoir comparer nos résultats aux résultats originaux. Les coefficients ω permettent d'évaluer la cohérence interne des sous-échelles au même titre que les coefficients α , mais l'utilisation de coefficients ω présente moins de risques de surestimation ou de sous-estimation de la fiabilité. Nous avons calculé les coefficients ω tel que recommandé par O'Brien dans l'étude originale ayant validé la *User Engagement Scale – Short Form* [4]. Plus particulièrement, nous avons utilisé le paquet logiciel MBESS du logiciel R (version 3.6.2, *The R Foundation for Statistical Computing*) et suivi les lignes directrices de Dunn [25]. Nous avons suivi le barème suivant pour l'interprétation des coefficients α et ω : un coefficient entre 0,70 et 0,79 représente une cohérence interne acceptable ; un coefficient entre 0,80 et 0,89 représente une bonne cohérence interne ; un coefficient supérieur ou égal à 0,90 représente une excellente cohérence interne [22, 26].

Résultats

Traduction et adaptation culturelle de la *User Engagement Scale – Short Form* en langue française

La version en langue française de la *User Engagement Scale – Short Form* a été nommée Échelle d'engagement de l'utilisateur – Forme abrégée (EEU-FA). La traduction de l'échelle de mesure a été effectuée de l'anglais au français, puis du français à l'anglais afin de repérer des incohérences. Peu de changements ont été apportés aux libellés des items de la version traduite en langue française lors du retour à l'anglais. Les items de la version pré-finale l'EEU-FA ont été examinés afin de déterminer s'ils étaient culturellement cohérents en contexte québécois et canadien. Aucun changement de fut apporté à ce titre. Par contre, deux items de la version en langue française (1, 9) posaient problème au plan du sens. La traduction littérale de l'item 1 « Je me suis perdu dans cette expérience » a été remplacée par l'item 2 de la version longue de la *User Engagement Scale*, qui évalue aussi le concept « attention soutenue ». Cet item fut traduit ainsi : « Cette expérience m'a tellement absorbé que j'en ai perdu la notion du temps ». De façon similaire, la traduction littérale de l'item 9 « Cette Application X a éveillé mes sens » a été remplacée par l'item 20 de la version longue de la *User Engagement Scale*, qui évalue aussi le concept « attrait esthétique ». Cet item fut traduit ainsi : « La présentation visuelle de l'Application X était plaisante à regarder ». La version mise à l'essai de l'EEU-FA est présentée au Tableau I.

Traduction et adaptation culturelle du *Cognitive Load Index* en langue française

La version en langue française du *Cognitive Load Index* a été nommée par l'Indice de charge cognitive (IDCC). Les items de la version pré-finale l'IDCC ont été examinés afin de déterminer s'ils étaient culturellement cohérents en contexte québécois et canadien. Aucun changement de fut apporté à ce titre. Toutefois, l'échelle de mesure a été développée pour une utilisation dans le contexte d'un cours portant sur les notions en statistiques. Par conséquent, comme prévu par les auteurs de l'échelle originale en langue anglaise, il fut nécessaire d'apporter des changements à trois items de l'IDCC qui sont spécifiques au contenu de la formation (2, 8 et 9). Par exemple, l'item 2 « L'activité traitait de [formules statistiques] qui me sont apparues très complexes » a été modifié pour « L'activité traitait [d'informations sur la santé cardiovasculaire] qui me sont apparues très complexes ». La version mise à l'essai de l'IDCC est présentée au Tableau II.

Évaluation psychométrique préliminaire

Cinquante-sept participants ont donné leur consentement et participé à l'évaluation psychométrique préliminaire de l'EEU-FA et de l'IDCC en juin 2019. Tous les participants ont répondu à l'ensemble des questions pour les deux échelles de mesures.

L'examen des données pour chacun des items de l'EEU-FA a révélé une seule donnée potentiellement aberrante sur un total de 684 observations (N items de l'échelle * N participants), ce qui fut jugé négligeable. L'examen des données aberrantes pour chacun des items de l'IDCC a révélé huit données potentiellement aberrantes sur un total de 570 observations, dont cinq provenaient des réponses de cinq participants à l'item 4 de l'IDCC : « Les directives ou les explications fournies durant l'activité étaient très difficiles à saisir ». Ces participants ont répondu « Tout à fait en accord » (9, 10) à cet item alors qu'ils ont répondu à l'opposé pour les deux autres items de cette même sous-échelle mesurant le même concept. Ceci résulte en des valeurs aberrantes qui peuvent compromettre les résultats d'analyse. Après discussion, les données de ces cinq participants ont été retirées des analyses pour l'IDCC (N=52 pour ces analyses).

Analyse des items et sous-échelles de l'Échelle d'engagement de l'utilisateur – Forme abrégée

Les caractéristiques des items de l'EEU-FA sont présentées au Tableau I. Les items pour lesquels les indices de *skewness* et de *kurtosis* s'éloignent le plus de 0 sont les items 4 et 5 de la sous-échelle de la convivialité perçue et l'item 0 de la sous-échelle de l'attrait esthétique, ce qui indique que la distribution des réponses des participants pour ces items s'éloigne de la normalité.

Tableau I. Caractéristiques des items et des sous-échelles de l'Échelle d'engagement de l'utilisateur – Forme abrégée (N=57).^{a, b}

Sous-échelles	Items ^a	Caractéristiques des items				Caractéristiques des sous-échelles	
		Moyenne ± ÉT ^c	Médiane (ÉI) ^c	Skewness	Kurtosis	Moyenne ± ÉT ^c	Médiane (ÉI) ^c
Attention soutenue	1. Cette expérience m'a tellement absorbé que j'en ai perdu la notion du temps.	2,77 ± 1,09	3,00 (1,00)	0,22	-0,40	3,04 ± 0,94	3,00 (1,17)
	2. Le temps a filé lorsque j'ai utilisé l' <u>Application X</u> .	3,18 ± 1,07	3,00 (1,00)	-0,54	-0,16		
	3. J'ai été absorbé par cette expérience.	3,16 ± 0,94	3,00 (1,00)	-0,46	0,40		
Convivialité perçue	4. J'ai éprouvé de la frustration lors de l'utilisation de l' <u>Application X</u> .	4,44 ± 0,89 ^e	5,00 (1,00)	-1,48	1,20	4,52 ± 0,62	5,00 (1,00)
	5. J'ai trouvé que l' <u>Application X</u> était difficile à utiliser.	4,51 ± 0,69 ^e	5,00 (1,00)	-1,42	2,12		
	6. L'utilisation de l' <u>Application X</u> était pénible.	4,61 ± 0,53 ^e	5,00 (1,00)	-0,86	-0,47		
Attrait esthétique	7. L' <u>Application X</u> était attrayante.	4,23 ± 0,73	4,00 (1,00)	-0,95	1,39	4,24 ± 0,62	4,00 (0,67)
	8. L' <u>Application X</u> était visuellement agréable.	4,25 ± 0,66	4,00 (1,00)	-0,70	1,14		
	9. La présentation visuelle de l' <u>Application X</u> était plaisante à regarder.	4,25 ± 0,79	4,00 (1,00)	-1,62	4,71		
Gratification	10. L'utilisation de l' <u>Application X</u> en valait la peine.	4,16 ± 0,75	4,00 (1,00)	-0,27	-1,16	3,99 ± 0,61	4,00 (0,83)
	11. Mon expérience s'est avérée gratifiante.	3,77 ± 0,78	4,00 (1,00)	-0,28	-0,17		
	12. J'ai été intéressé par cette expérience.	4,04 ± 0,60	4,00 (0,00)	-0,53	1,97		

a. Abréviations : ÉT = écart-type ; ÉI = écart interquartile ; Z = Z-Score.

b. Dans la présente étude, le terme « Application X » était remplacé par le nom du site Internet visité par les participants.

c. Étendue : 1–5. Échelle de réponse : 1) fortement en désaccord; 2) en désaccord; 3) ni en accord, ni en désaccord; 4) en accord; 5) fortement en accord.

d. Les scores de ces trois items ont été codés à l'inverse.

Analyse des items et sous-échelles de l'Indice de charge cognitive

Les caractéristiques des items de l'IDCC sont présentées au Tableau II. Les items pour lesquels les indices de *skewness* et de *kurtosis* s'éloignent le plus de 0 sont les items 4 et 6 de la sous-échelle de la charge intrinsèque.

Tableau II. Caractéristiques des items et des sous-échelles de l'Indice de charge cognitive (N=52).^a

Sous-échelles	Items ^b	Caractéristiques des items				Caractéristiques des sous-échelles	
		Moyenne ± ÉT ^c	Médiane (ÉI) ^c	Skewness	Kurtosis	Moyenne ± ÉT ^c	Médiane (ÉI) ^c
Charge intrinsèque	1. Le ou les sujets abordés durant cette activité étaient très complexes.	2,13 ± 2,21	2,00 (3,00)	1,27	0,90	1,43 ± 1,42	1,00 (2,00)
	2. L'activité traitait <i>d'informations sur la santé cardiovasculaire</i> qui me sont apparues très complexes.	1,35 ± 1,37	1,00 (2,00)	1,24	1,82		
	3. L'activité traitait de définitions et de concepts qui me sont apparus très complexes.	1,12 ± 1,34	1,00 (2,00)	1,11	0,45		
Charge extrinsèque	4. Les directives ou les explications fournies durant l'activité étaient très difficiles à saisir.	1,04 ± 1,52	0,00 (2,00)	1,53	1,88	1,37 ± 1,48	1,00 (2,00)
	5. Les directives ou les explications étaient inefficaces pour mon apprentissage.	1,54 ± 1,95	1,00 (2,75)	1,25	0,59		
	6. Les directives ou les explications incluait beaucoup de termes difficiles à comprendre.	0,77 ± 1,22	0,00 (1,00)	2,10	5,66		
Charge essentielle	7. L'activité a vraiment amélioré ma compréhension du ou des sujets abordés.	5,56 ± 2,49	6,00 (4,50)	-0,10	-0,85	5,48 ± 2,57	5,75 (4,50)
	8. L'activité a vraiment amélioré ma connaissance et ma compréhension de <i>la prévention cardiovasculaire</i> .	5,62 ± 2,84	5,50 (5,00)	-0,03	-1,05		
	9. L'activité a vraiment amélioré ma compréhension des <i>informations sur la santé cardiovasculaire</i> abordées.	5,54 ± 2,78	5,00 (5,00)	0,12	-1,08		
	10. L'activité a vraiment amélioré ma compréhension des concepts et des définitions.	5,12 ± 2,86	5,50 (4,00)	-0,11	-0,97		

a. Abréviations : ÉT = écart-type ; ÉI = écart interquartile.

b. Les termes en italique sont spécifiques au contenu de l'activité dans la présente étude; ils devraient être remplacés dans le cadre d'un autre contexte de formation.

c. Étendue : 0-10. Échelle de réponse : 0) pas du tout d'accord; 10) tout à fait d'accord.

Cohérence interne et corrélation des sous-échelles de l'Échelle d'engagement de l'utilisateur – Forme abrégée et de l'Indice de charge cognitive

Les coefficients omega (ω) de McDonald et les coefficients alpha (α) de Cronbach sont présentés dans la diagonale du Tableau III pour apprécier, respectivement, la cohérence interne de l'EEU-FA et de l'IDCC. Pour l'EEU-FA, trois des quatre sous-échelles (attention soutenue, convivialité perçue et gratification) démontrent une bonne cohérence interne avec des coefficients ω se situant entre 0,80 et 0,90. Une sous-échelle (attrait esthétique) obtient un coefficient ω de 0,77, démontrant un niveau acceptable de cohérence interne. Pour l'IDCC, les coefficients α observés atteignaient tous le niveau d'acceptabilité. Ils étaient de 0,83 pour la charge intrinsèque, de 0,70 pour la charge extrinsèque, et de 0,96 pour la charge essentielle.

Tableau III. Cohérence interne (sur la diagonale) et corrélation des sous-échelles (sous la diagonale) de l'Échelle d'engagement de l'utilisateur– Forme abrégée (EEU-FA) et de l'Indice de charge cognitive (IDCC).

		EEU-FA				IDCC		
		AS ^a	CP ^a	AE ^a	Gr ^a	CI ^a	CEx ^a	CEs ^a
EEU-FA	AS	0,89 ^b						
	CP	0,03	0,89 ^b					
	AE	0,04	0,42**	0,77 ^b				
	Gr	0,64**	0,20	0,45**	0,83 ^b			
IDCC	CI	0,04	-0,41**	-0,31*	0,07	0,83 ^c		
	CEx	0,06	-0,43**	-0,32*	-0,06	0,41**	0,70 ^c	
	CEs	0,31*	0,24	0,28*	0,59**	0,23	-0,12	0,96 ^c

** P < 0,01 ; * P < 0,05.

- Abréviations : AS = attention soutenue ; CP = convivialité perçue ; AE = attrait esthétique ; Gr = gratification ; CI = charge intrinsèque ; CEx = charge extrinsèque ; CEs = charge essentielle.
- N = 57. Les coefficients omega (ω) de McDonald sont présentés pour l'appréciation de la cohérence interne des sous-échelles de l'EEU-FA. L'intervalle de confiance à 95 % de ω : AS = [0,76; 0,94] ; CP = [0,81; 0,94] ; AE = [0,51; 0,89] ; Gr = [0,74; 0,87].
- N = 52. Les coefficients alpha (α) de Cronbach sont présentés pour l'appréciation de la cohérence interne des sous-échelles de l'IDCC.

Les coefficients de corrélation entre les sous-échelles de, respectivement, l'EEU-FA et de l'IDCC, sont présentés au Tableau III (en italique sous la diagonale). Parmi les corrélations au sein de l'EEU-FA, la plus élevée est entre l'attention soutenue et la gratification ($r = 0,64$), alors que l'attention soutenue est très peu reliée à la convivialité perçue ($r = 0,03$) et à l'attrait esthétique ($r = 0,04$). Pour l'IDCC, la corrélation la plus élevée et la seule significative au plan statistique est entre la charge intrinsèque et de la charge extrinsèque ($r = 0,41$). Parmi les corrélations entre les sous-échelles de l'EEU-FA et de l'IDCC (en gras dans le tableau), la sous-échelle de la charge essentielle est fortement corrélée avec la sous-échelle de la gratification ($r = 0,59$).

Discussion

Cette étude a permis de traduire et d'adapter en français deux échelles de mesure visant à évaluer l'engagement et la charge cognitive en contexte d'apprentissage numérique. De plus, nous avons mené une évaluation psychométrique préliminaire de ces deux échelles de mesure, l'EEU-

FA et l'IDCC, auprès de 57 étudiants de niveau universitaire en sciences infirmières. Les résultats indiquent que les échelles traduites ont des niveaux de cohérence interne satisfaisants pour une utilisation dans le cadre d'études. Une évaluation psychométrique auprès d'un plus grand échantillon serait toutefois requise afin de poursuivre l'examen des propriétés psychométriques des échelles de mesure.

Les sous-échelles de l'EEU-FA ont démontrées des niveaux de cohérence interne satisfaisants (c.-à-d. au-dessus de 0,70). Dans la présente étude, les coefficients ω obtenus pour les sous-échelles variaient entre 0,77 et 0,89, tandis qu'ils variaient entre 0,70 et 0,88 dans les deux études originales [4]. La version originale de l'EEU-FA ayant été publiée en 2018, à notre connaissance, il n'y a pas encore d'étude publiée examinant les propriétés psychométriques de cette échelle dans d'autres contextes. Les résultats de la présente étude soulignent le potentiel de l'EEU-FA quant à la mesure de l'engagement en contexte d'apprentissage numérique. En effet, l'EEU-FA, validée initialement pour l'évaluation de l'expérience utilisateur d'individus utilisant des plateformes de e-commerce [4, 7], démontre donc un potentiel intéressant pour mesurer l'engagement en contexte d'apprentissage numérique chez des professionnels de la santé.

Deux des trois sous-échelles de l'IDCC ont démontré une bonne ou une excellente cohérence interne. Un niveau de cohérence interne plus faible, mais acceptable (0,70) a été observé pour la sous-échelle de charge extrinsèque, qui fait référence à la perception des instructions et des explications sur le site web visité par les participants. Nos résultats font écho à une étude de 2017 où la cohérence interne des sous-échelles du *Cognitive Load Index* était de 0,72 pour la charge intrinsèque, de 0,64 pour la charge extrinsèque et de 0,85 pour la charge essentielle [23]. Des résultats similaires ont été observés dans une autre étude [27]. De plus, dans un des quatre tests de l'étude originale de validation du *Cognitive Load Index*, la sous-échelle de charge extrinsèque avait obtenu un coefficient α de Cronbach de 0,63 [10]. Tel que le suggèrent Cook, Castillo [23] et comme nous l'avons observé, il est possible que la formulation de l'item #4 de la sous-échelle de charge extrinsèque (c.-à-d. « Les directives ou les explications fournies durant l'activité étaient *très* difficiles à saisir ») pose un problème d'interprétation puisqu'il inclue un qualificatif extrême (c.-à-d. « très »). Toutefois, d'autres items ayant des qualificatifs extrêmes (p. ex. très, vraiment) comme les items 1, 2 et 3 n'ont pas eu ce problème. En ce sens, à ce stade, il ne nous est pas possible de proposer d'explications additionnelles quant aux données aberrantes obtenues pour l'item 4. Il serait donc indiqué de porter une attention particulière à cet item lors d'études futures.

Certains points méritent d'être soulevés au niveau de la corrélation entre les sous-échelles de l'EEU-FA et de l'IDCC. Pour ce qui est de l'EEU-FA, les corrélations modérées observées certaines des sous-échelles sont attendues (p. ex., attrait esthétique et convivialité perçue [$r = 0,42$], attrait esthétique et gratification [$r = 0,45$]), puisque celles-ci mesurent différents attributs du même concept et sont donc reliées au plan théorique [26]. Toutefois, les corrélations entre les sous-échelles de l'attrait esthétique et de l'attention soutenue ($r = 0,04$) et de la convivialité perçue et de l'attention soutenue ($r = 0,03$), ce qui pourrait laisser présager que l'attention des utilisateurs est davantage soutenue par le contenu que la forme de l'environnement numérique. Pour ce qui est de l'IDCC, notons que sur le plan théorique, les concepts de charge intrinsèque et de charge extrinsèque sont distincts. Il est postulé que ces deux types de charge cognitive sont indépendants et s'additionnent à la charge essentielle pour constituer la charge cognitive globale dans la mémoire de travail de l'apprenant [9, 13, 28, 29]. En ce sens, nous pourrions nous attendre à ce que la corrélation entre les sous-échelles de la charge intrinsèque et de la charge extrinsèque soit nulle si

elles représentent des concepts distincts. Pourtant, nous avons observé une corrélation modérée de 0,41 ($p < 0,01$) entre ces deux sous-échelles. Ceci a aussi été observé dans deux études ayant évalué la version originale du *Cognitive Load Index* [27]. Ces résultats pourraient indiquer que la formulation des items de ces deux sous-échelles n'est pas optimale afin de distinguer les deux types de charge cognitive, c'est-à-dire de distinguer la complexité de la tâche d'apprentissage en elle-même (charge intrinsèque) de la complexité des instructions et des explications (charge extrinsèque) [8, 27, 28]. Cette corrélation modérée entre les sous-échelles de la charge intrinsèque et de la charge extrinsèque pourrait aussi indiquer que ces deux types de charge cognitive sont reliées et s'influencent en contexte d'apprentissage, remettant en cause leur conceptualisation en tant que concepts indépendants. Dans le cadre de l'étude, nous avons aussi observé une faible corrélation négative entre les sous-échelles de la charge extrinsèque et de la charge essentielle. Cette observation est attendue. Une charge extrinsèque plus basse libèrera la capacité cognitive de l'apprenant, lui permettant de s'engager davantage dans l'apprentissage et augmentant ainsi la charge essentielle [23].

Il est également intéressant de noter la forte corrélation ($r = 0,59$; $P < 0,01$) entre les sous-échelles de la gratification de l'EEU-FA et de la charge essentielle de l'IDCC. Ce résultat préliminaire laisse présager une forte association entre *la sensation de plaisir et de gratification* tirée de l'utilisation de l'environnement numérique d'apprentissage et *la perception d'avoir appris* chez les étudiants. En ce sens, il serait particulièrement pertinent de valider cette association au sein d'un échantillon plus large d'étudiants et de professionnels dans le domaine de la santé.

Cette étude présente certaines limites. Premièrement, nous avons utilisé un échantillon de convenance et nous n'avons pas collecté de données sociodémographiques nous permettant de décrire l'échantillon, outre le fait qu'ils étaient tous étudiants dans un programme en sciences infirmières d'une université québécoise. Par conséquent, une évaluation psychométrique de l'EEU-FA et de l'IDCC devrait être réalisée auprès de populations diversifiées. Deuxièmement, l'étude ne comprenait pas de phase formelle d'examen de la validité de contenu des deux échelles de mesure. Cette décision a été prise puisqu'un corpus important de littérature en langue anglaise soutient la validité de contenu de la *User Engagement Scale – Short Form* et du *Cognitive Load Index*. De plus, deux membres de l'équipe de traduction ont examiné les versions traduites de façon préliminaire des échelles de mesure et ont jugé que ces dernières faisaient état d'une bonne *validité apparente*. Troisièmement, nous n'avons pas réalisé d'analyse factorielle exploratoire vu la petite taille de l'échantillon. La réalisation d'une analyse factorielle avec un plus grand échantillon de professionnels de la santé serait pertinente afin d'évaluer la robustesse des dimensions de l'EEU-FA et de l'IDCC auprès de cette population en contexte d'apprentissage numérique. À ce titre, il serait opportun d'explorer la contribution potentielle des travaux récents explorant la pertinence de modèles alternatifs en matière d'analyse factorielle [30]. Quatrièmement, pour certains items de l'EEU-FA et de l'IDCC, nous avons observé des écarts de la normalité au niveau du *skewness* et du *kurtosis* dans la distribution des scores. Toutefois, cela est attribuable à la taille d'échantillon plus petite dans le cadre d'une évaluation psychométrique préliminaire [31]. Nous avons présenté à la fois les moyennes et écarts-types ainsi que les médianes et écarts interquartiles pour mieux caractériser les distributions des scores. Enfin, afin d'examiner la cohérence interne des sous-échelles de l'IDCC, nous avons retiré des données aberrantes pour cinq participants en utilisant un critère statistique reconnu afin de réaliser cette démarche [20]. Cette démarche était indiquée dans la présente étude puisque ces données aberrantes étaient attribuables à une erreur de lecture de l'échelle de mesure et que des données antérieures soutiennent la problématique observée.

En conclusion, en dépit de ces limites, la présente étude fait état de plusieurs forces. Le processus de traduction des deux échelles de mesure fut réalisé d'après les plus récents standards [18]. Les traducteurs impliqués étaient experts en contenu ou en traduction. Cette démarche a permis d'obtenir des traductions complémentaires sur le plan scientifique et linguistique afin de préserver la validité de contenu des versions originales des échelles de mesure. Les résultats de cette étude mettent en évidence le potentiel de l'EEU-FA et de l'IDCC pour mesurer l'engagement et la charge cognitive en contexte d'apprentissage numérique chez les étudiants et les professionnels de la santé. Ces échelles ont un potentiel intéressant pour déterminer des cibles actionnables (p. ex. convivialité, attrait esthétique, complexité de la tâche d'apprentissage) pour optimiser les formations dispensées en contexte numérique. Les recherches futures pourront permettre d'examiner en détail les qualités psychométriques de l'EEU-FA et de l'IDCC (p. ex., la fidélité test-retest) et d'entreprendre une analyse factorielle de ces échelles de mesure.

Approbation éthique

L'étude a été approuvée par le Comité d'éthique de la recherche en sciences et en santé (CERSES) de l'Université de Montréal au mois de mai 2019 (CERSES-19-042-D).

Déclaration d'intérêts

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les résultats publiés dans cet article. Les auteurs n'ont aucun lien avec le site Internet qui fut proposé dans le cadre de l'étude.

Financement et bourses

GF a été soutenu financièrement par la bourse d'études supérieures du Canada Vanier (Instituts de recherche en santé du Canada) et par des bourses du Fonds de recherche du Québec — Santé, de la Fondation des infirmières et infirmiers du Canada, de la Fondation de l'Institut de Cardiologie de Montréal, du Ministère de l'Enseignement supérieur du Québec et de la Faculté des sciences infirmières de l'Université de Montréal.

Remerciements

Les auteurs souhaitent remercier les auteurs Heather L. O'Brien, auteure de la version originale de la *User Engagement Scale – Short Form*, et Jimmie Leppink, auteur de la version originale du *Cognitive Load Index*, pour nous avoir autorisés à traduire et à adapter les échelles de mesure.

Références

1. Sinclair PM, Kable A, Levett-Jones T, Booth, D. The effectiveness of Internet-based e-learning on clinician behaviour and patient outcomes: a systematic review. *Inf J Nurs Stud* 2016;57:70-81.
2. Clark R, Mayer R. E-Learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning. 4 ed. Hoboken, NJ : John Wiley & Sons, 2016.

3. O'Brien HL. Theoretical perspectives on user engagement. In : O'Brien HL, Cairns P (eds). Why engagement matters: cross-disciplinary perspectives and innovations on user engagement with digital media. Basel, Switzerland : Springer, 2016:1-26.
4. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *Int J Hum-Comput St* 2018;112:28-39.
5. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2016;7:254-267.
6. Doherty K, Doherty G. Engagement in HCI: conception, theory and measurement. *ACM Comput Surv* 2018;51:1-39.
7. O'Brien HL, Toms EG. The development and evaluation of a survey to measure user engagement. *J Assoc Inf Sci Tech* 2010;61:50-69.
8. Sweller J. Element interactivity and intrinsic, extraneous and germane cognitive load. *Educ Psychol Rev* 2010;22:123-138.
9. Young JQ, van Merriënboer J, Durning S, ten Cate O. Cognitive Load Theory: Implications for medical education. *Med Teach* 2014;36:371-384.
10. Leppink J, Paas F, van der Vleuten CP, van Gog T, van Merriënboer J. Development of an instrument for measuring different types of cognitive load. *Behav Res Methods* 2013;45:1058-1072.
11. Paas F. Training strategies for attaining transfer of problem solving skills in statistics: A cognitive load approach. *J Educ Psychol* 1992;84:429-434.
12. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv Psychol* 1988;52:139-183.
13. DeLeeuw KE, Mayer RE. A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *J Educ Psychol* 2008;100:223-234.
14. Fontaine G. Évaluation de la plateforme de formation en ligne MOTIV@CŒUR sur les interventions motivationnelles brèves auprès d'infirmières en soins aigus cardiovasculaires. Master's Thesis. Montréal, Canada : Université de Montréal, 2016. Available from : <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/16290>
15. Fontaine G, Cossette S, Heppell S, Boyer L, Mailhot T, Simard MJ et al. Evaluation of a web-based e-learning platform for brief motivational interviewing by cardiovascular care nurses: a pilot study. *J Med Internet Res* 2016;8:e224.
16. Fontaine G, Cossette S, Maheu-Cadotte MA, Mailhot T, Heppell S, Roussy C et al. Behavior change counseling training programs for nurses and nursing students: A systematic descriptive review. *Nurse Educ Today* 2019;82:37-50.
17. Fontaine G, Cossette S, Maheu-Cadotte MA, Mailhot T, Deschênes MF, Mathieu-Dupuis G et al. Efficacy of adaptive e-learning for health professionals and students: a systematic review and meta-analysis. *BMJ Open* 2019;9:e025252.
18. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract* 2011;17:268-274.
19. Béland S, Cousineau D, Loye N. Utiliser le coefficient omega de McDonald à la place de l'alpha de Cronbach. *McGill J Educ* 2017;52:791-804.

20. University of Florida Health. Outliers [Internet]. 2019; Available from: <https://bolt.mph.ufl.edu/6050-6052/unit-1/one-quantitative-variable-introduction/understanding-outliers/#extreme>
21. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. *Int J Endocrinol Metab* 2012;10:486-489.
22. Field A. *Discovering statistics using SPSS: introducing statistical method*. 3 ed. Thousand Oaks, CA: Sage Publications, 2009.
23. Cook DA, Castillo RM, Gas B, Artino Jr AR. Measuring achievement goal motivation, mindsets and cognitive load: validation of three instruments' scores. *Med Educ* 2017;51:1061-1074.
24. Cohen J. *Statistical power analysis for the behavioral sciences*. 2 ed. Hillsdale, New Jersey: Lawrence Erlbaum, 1988.
25. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br J Psychol* 2014;105:399-412.
26. DeVellis RF. *Scale development: theory and applications*. 4 ed. Los Angeles, CA: Sage Publications, 2017.
27. Leppink J, Paas F, Van Gog T, van Der Vleuten CP, Van Merriënboer JJ. Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learn Instr* 2014;30:32-42.
28. Sweller J. Cognitive load during problem solving: effects on learning. *Cogn Sci* 1988;12:257-288.
29. Paas F, Tuovinen JE, Tabbers H, Van Gerven PWM. Cognitive load measurement as a means to advance cognitive load theory. *Educ Psychol* 2003;38:63-71.
30. Sellbom M, Tellegen A. Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychol Assess* 2019; 31:1428-1441.
31. Nunnally JC, Bernstein IH. *Psychometric theory*. 3 ed. New York, NY: McGraw-Hill, 1994.