

***Élaboration d'une échelle de mesure et
preuves de validité structurelle par analyses
factorielles exploratoires***

***Illustration sur une échelle multidimensionnelle
d'engagement des étudiants***

Géraldine Heilporn, Ph.D.

Atelier-webinaire 21 octobre 2021



Plan du webinar

- Développement et validation d'une échelle de mesure
- Exploration de la structure interne d'une échelle par des analyses factorielles
- Illustration sur une échelle de mesure multidimensionnelle d'engagement des étudiants

Heilporn, G., Lakhal, S., Bélisle, M. et St-Onge, C. (2021). Engagement des étudiants : une échelle de mesure multidimensionnelle appliquée à des modalités de cours hybrides universitaires. *Mesure et évaluation en éducation*, 43(2), 1-34. <https://doi.org/10.7202/1081043ar>

Introduction

- Intention de départ pour le webinaire
- Pourquoi devrait-on apporter des preuves de validité d'une échelle de mesure? Pourquoi devrait-on étudier sa structure interne?

DÉVELOPPEMENT ET VALIDATION D'UNE ÉCHELLE DE MESURE

Principales étapes (DeVellis, 2016)

1. Déterminer clairement l'objet de mesure
2. Générer un bassin initial d'énoncés (items)
3. Sélectionner un format de mesure
4. Faire réviser le bassin initial d'énoncés par des experts
5. Considérer l'inclusion d'énoncés de validation
6. Tester les énoncés auprès un échantillon de participants
7. Évaluer les énoncés
8. Optimiser la longueur de l'échelle

Perspective intégratrice

Perspective intégratrice et unifiée de **validité** de l'interprétation des résultats d'une échelle **pour un usage proposé** (Messick, 1995; *Standards for Educational and Psychological Testing* de l'*American Educational Research Association*, 2014)

- Éléments de preuve de contenu
- Éléments de preuve de processus de réponse
- Éléments de preuve de la structure interne (construit)
- Éléments de preuve de conséquence

1. Déterminer l'objet de mesure

Délimitation claire et précise des frontières du phénomène d'intérêt

- engagement des étudiants dans des modalités de cours hybrides universitaires
- perspective multidimensionnelle d'engagement (Fredricks et al., 2004, 2016, 2019)

Prise en compte du niveau de spécificité ou de généralité désiré

- Contexte multidisciplinaire, enseignement supérieur
- Modalités de cours hybrides (activités synchrones + asynchrones)

Objet de mesure : illustration

Frontières : Engagement \neq motivation (et \neq stratégies d'engagement)

~~Antécédents~~

Objet de mesure

~~Conséquences~~

Point de départ : dimensions comportementale, émotionnelle, cognitive de Fredricks et a. (2004)

Participation aux activités
Suivi des règles, des normes



Réactions affectives
Sentiment d'appartenance

Cognitif

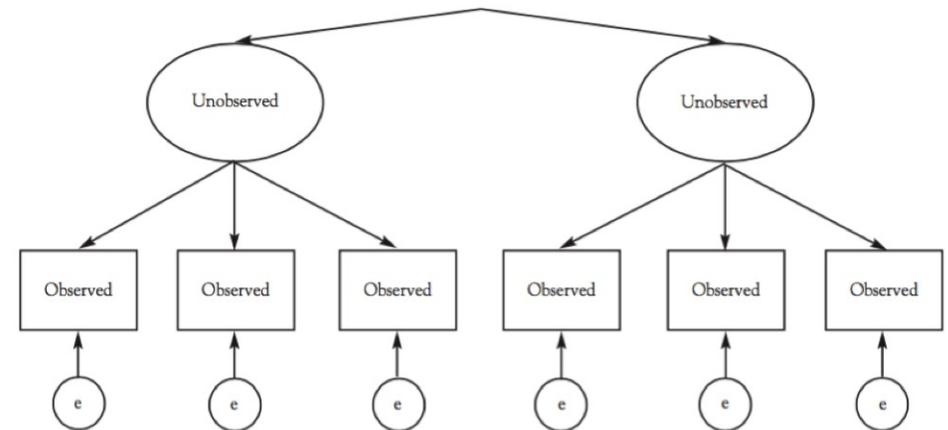
Investissement psychologique
Stratégies d'apprentissage et métacognitives

* étudiants-étudiants

2. Générer un bassin initial d'énoncés

Chaque énoncé devrait refléter la variable latente sous-jacente

Relations entre v. latentes (facteurs)
et v. observées (énoncés)



(Schreiber, Nora, Stage, Barlow et King, 2006, p. 324)

Redondances :

- Oui à l'étape initiale
- ~~Formulations trop similaires~~
- Même niveau de spécificité
- Sous-dimensions : attention au nombre d'énoncés

Voir DeVellis (2016), ch. 5

2. Générer un bassin initial d'énoncés

Formuler de « bons » énoncés :

- **Clarté**
- Énoncés courts
- Formulations simples (adaptées aux participants visés!)
- ~~Questions doubles (attention aux « et », « ou », « parce que », « ce qui », etc.)~~
- ~~Doubles négations~~

Échelle d'engagement :

- 69 énoncés initialement (21 comp, 20 émo, 28 cogn)
- Inspiration échelles anglophones et indicateurs pour chaque dimension

3. Sélectionner un format de mesure

... et justifier les choix 😊

Réflexions :

- Quels buts sont poursuivis par l'échelle de mesure?
- Quelles méthodes d'analyse de données seront utilisées?

Échelle de Likert : degré d'accord avec des opinions ou attitudes

< 5 choix de réponses => perte d'informations? (Streiner et al., 2015)

> 7 choix de réponses => surcharge cognitive?

Nombre pair ou impair de choix de réponses (Sturgis et al., 2014)

Échelle d'engagement :

➤ 6 choix de réponses

1) tout à fait en désaccord, 2) en désaccord, 3) plutôt en désaccord,
4) plutôt d'accord, 5) d'accord et 6) tout à fait d'accord

4. Faire réviser le bassin initial d'énoncés par des experts

Évaluation de chaque énoncé :

Échelle de Likert
(1-faible, 2-moyenne, 3-élevée)

- **Pertinence** p/r objet de mesure
 - Fournir les définitions nécessaires
 - Montrer une ouverture à des suggestions / commentaires
- **Clarté** et concision

Échelle d'engagement :

- Énoncés examinés par 7 professeurs universitaires
- Pertinence (resp. clarté) notée : Sur 69 énoncés : 13 reformulés, 30 supprimés → 39 énoncés

**VOUS (équipe de recherche) prenez les décisions finales.
Quelles règles de décision?**

5. Considérer l'inclusion d'énoncés de validation

Biais de désirabilité sociale?

Inclusion d'une échelle de désirabilité sociale

p. ex., Balanced Inventory of Desirable Responding (Paulhus, 1998) ou version abrégée (D'Amours-Raymond, 2011)

Validité de construit

Inclusion d'une échelle mesurant un construit apparenté.

Échelle d'engagement :
➤ Pas d'autre échelle incluse

6. Tester les énoncés auprès un échantillon de participants

Dans certains cas : prétest de la clarté des énoncés auprès d'un petit échantillon de participants.

Administration de l'échelle :

- représentativité de la population ciblée
- Échantillon de grande taille pour AFE (aussi AFC)
- Pour des AFE, présentation des énoncés en ordre aléatoire

EXPLORATION DE LA STRUCTURE INTERNE D'UNE ÉCHELLE PAR DES ANALYSES FACTORIELLES

Taille de l'échantillon de données

Grand échantillon → stabilité des résultats

Pas de consensus :

200 acceptable, 300 bon, 500 très bon, 1000 excellent
(Comrey et Lee, 1992)

≥ 5-10 participants par énoncé (p.ex., Costello et Osborne, 2005; Stevens, 1996)

Taille minimale dépend des communautés obtenues
(McCallum et al., 1999) → jugement post hoc

Échelle d'engagement :

- 39 énoncés → objectif : min 195 participants (5:1)
- Obtenu : 465 participants

Échelle d'engagement : échantillon

- 39 énoncés → objectif : min 195 participants (5:1)
- Obtenu : 465 participants
- Division aléatoire en deux échantillons ($n_1 = 234$, $n_2 = 231$)
 - But : stabilité test-retest

Tableau 1
Caractéristiques sociodémographiques des participants (n=465)

Caractéristiques	Fréquences en %		
Genre	73,1% féminin	26% masculin	0,9% préfèrent ne pas répondre
Âge	53,3% ≤25 ans	19,3% [26-35 ans]	16,8% [36-45 ans] 10,6% [>45 ans]
Cycle	66,5% 1 ^{er} cycle	33,5% 2 ^e -3 ^e cycle	
Modalité	46,2% hybride	32,7% hybride en ligne	21,1% hybride-mixte

Examen initial des données

Analyse d'items : difficulté, discrimination, asymétrie

Voir Laveault et Grégoire (2014)

+ test normalité des données

Matrice des corrélations :

- Corrélations item-échelle ≥ 0.30
- Nombre de corrélations item-item ≥ 0.30

(DeVellis, 2016 ; Pituch et Stevens, 2016)

- Corrélations item-item ≤ 0.80 (colinéarité)

(Bourque et al., 2007 ; Tabachnick et Fidell, 2007)

Échelle d'engagement : examen initial

➤ Fortes asymétries, indices difficulté élevés (> 0.90) et discrimination faibles (< 0.20) ; certains $r(\text{item}, \text{échelle}) < 0.30$ et $\text{gd nb-r}(\text{item}, \text{item}) < 0.30$

« J'ai participé aux séances synchrones (en classe ou virtuelles) du cours »

« J'ai visité la plateforme en ligne du cours chaque semaine »

« J'ai remis les travaux du cours à temps »

« J'ai remis les travaux du cours à temps »

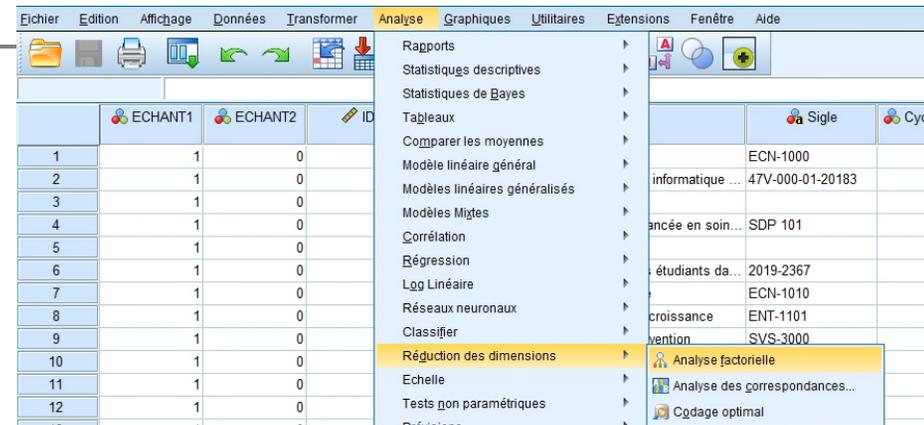
➤ $R(\text{item}, \text{item}) > 0.80$ trop élevé

« Je me suis senti(e) enthousiaste par rapport à ce cours »

« J'ai eu du plaisir dans ce cours » ($r = 0.83$)

« J'avais hâte de participer aux activités de ce cours » ($r = 0.82$)

Analyses factorielles exploratoires



Adéquation avec des analyses factorielles :

- Test de sphéricité de Bartlett significatif ($p < 0.05$)
(corrélations entre les énoncés)
- Indice de Kaiser-Meyer-Olkin (KMO) ≥ 0.7
(ajustement des énoncés aux facteurs sous-jacent)

(Tabachnick & Fidell, 2007; Yong & Pearce, 2013)

Échelle d'engagement :

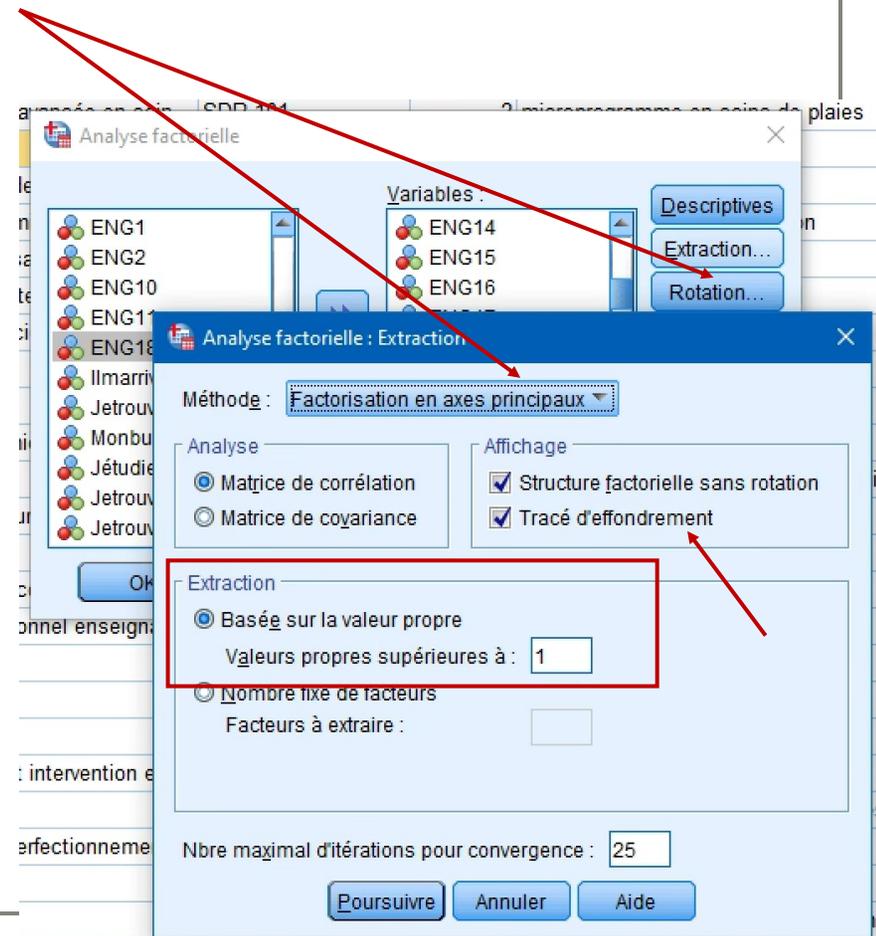
➤ Bartlett $p = 0.000$, KMO = 0.94 excellent

Analyses factorielles exploratoires

Choix de la méthode :

- factorisation en axes principaux
- Rotation oblique (oblimin)
- Si intercorrélations entre les facteurs
- Autrement : rotation orthogonale (ex. Varimax)

(Costello et Osborne, 2005 ; Worthington et Whittaker, 2006)



Nombre de facteurs retenus

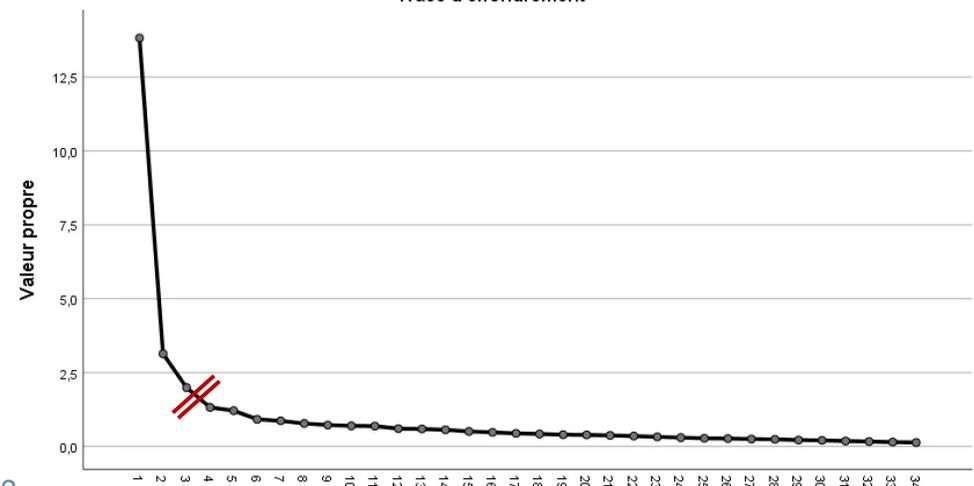
Sur base de plusieurs critères :

- valeurs propres > 1.0
- tracé d'effondrement
- caractère interprétable de la matrice des facteurs obtenue

Variance totale expliquée

Facteur	Valeurs propres initiales				Sommes extraites du carré des chargements		
	Total	% de la variance	% cumulé	Total	% de la variance	% cumulé	
1	13,820	40,647	40,647	13,419	39,466	39,466	
2	3,137	9,225	49,872	2,719	7,996	47,462	
3	1,992	5,858	55,730	1,569	4,615	52,077	
4	1,322	3,888	59,618	,838	2,466	54,543	
5	1,212	3,565	63,183	,757	2,227	56,769	
6	,920	2,705	65,888				
7	,866	2,546	68,434				
8	,777	2,286	70,720				

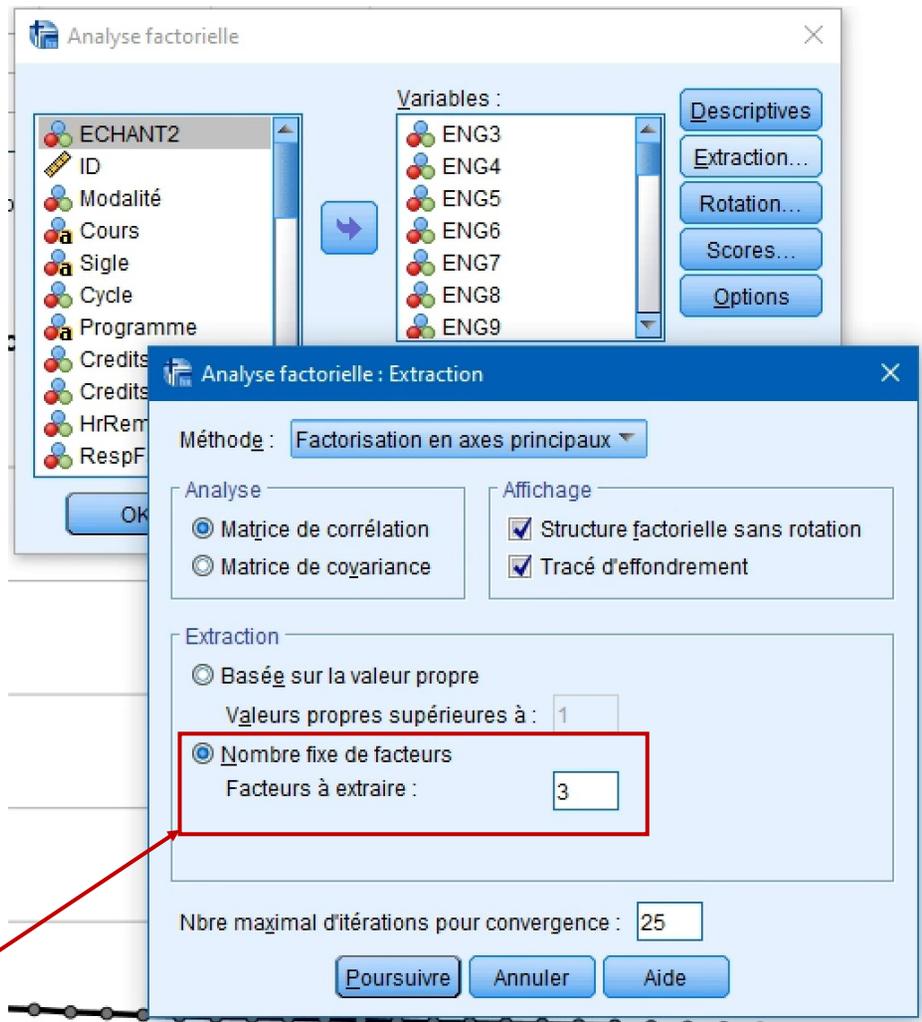
Tracé d'effondrement



Échelle d'engagement :

- AFE initiale, 5 facteurs de v.p. > 1
- Pas d'interprétation claire des facteurs obtenus
- Tracé d'effondrement suggérait plutôt 3 facteurs

Caractère itératif des AF exploratoires



Suggestions :

- 1) **Avancer à petits pas,**
- 2) **toujours prendre en compte le sens conceptuel**
- 3) **Préserver une logique dans les décisions**

Examen des communautés

Condition nécessaire :

Communautés extraites ≥ 0.32

(Tabachnick et Fidell, 2007)

↳ part de la variance de cette variable
qui est expliquée par les facteurs
communs

Échelle d'engagement :

X « J'ai réalisé les activités proposées dans les séances synchrones (en classe ou virtuelles) du cours »

X « J'ai réalisé les activités proposées en ligne dans ce cours »

X « J'ai exploré certains sujets du cours plus en profondeur que ce qui était demandé »

Analyse factorielle

Qualités de représentation

	Initiales	Extraction
ENG3 J'ai réalisé les lectures suggérées dans ce cours	,469	,403
ENG4 J'ai été attentif dans ce cours	,665	,596
ENG5 J'ai révisé mes notes de cours pour m'assurer de maîtriser le contenu	,509	,409
ENG6 J'ai consacré du temps à ce cours	,564	,493
ENG7 J'ai travaillé chaque semaine pour ce cours	,486	,445
ENG8 J'ai réalisé les activités proposées dans les séances synchrones (en classe ou virtuelles) du cours	,430	,251
ENG9 J'ai réalisé les activités proposées en ligne dans ce cours	,486	,354
ENG12 J'ai échangé avec d'autres étudiants à propos du cours	,623	,589
ENG13 J'ai participé activement à des discussions dans ce cours	,512	,431
ENG14 J'ai trouvé des façons de rendre le cours pertinent pour moi	,641	,608
ENG15 Je me suis senti intrigué par les activités de ce cours	,610	,547
ENG16 J'ai été intéressé par ce cours	,808	,750

Examen des saturations

Conditions nécessaires :

- **Saturations ≥ 0.32 sur UN facteur**

(Tabachnick et Fidell, 2007 ; Worthington et Whittaker, 2006)

→ À éliminer :

- Énoncés qui ne saturent sur aucun facteur
- Énoncés qui saturent sur plusieurs facteurs

! Processus ITÉRATIF !

Matrice de forme^a

	Facteur		
	1	2	3
ENG3 J'ai réalisé les lectures suggérées dans ce cours	,006	-,015	,585
ENG4 J'ai été attentif dans ce cours	,264	,118	,538
ENG5 J'ai révisé mes notes de cours pour m'assurer de maîtriser le contenu	-,055	,181	,620
ENG6 J'ai consacré du temps à ce cours	-,044	-,041	,749
ENG7 J'ai travaillé chaque semaine pour ce cours	,004	-,040	,641
ENG12 J'ai échangé avec d'autres étudiants à propos du cours	-,263	,845	,080
ENG13 J'ai participé activement à des discussions dans ce cours	,232	,459	,105
ENG32 J'ai développé de meilleures habilités de jugement dans ce cours	,666	,105	,070
ENG33 Je désirais vraiment maîtriser le contenu de ce cours	,266	-,027	,621
ENG34 j'ai discuté du cours avec d'autres étudiants en dehors des temps d'activités en classe et en ligne	,002	,633	,056
ENG35 J'ai évalué les forces et les faiblesses de mes propres idées sur un sujet du cours	,313	,151	,294
ENG36 J'ai réfléchi sur ma façon d'apprendre dans ce cours	,246	,042	,393
ENG38 J'ai donné le meilleur de moi-même dans ce cours	,097	-,052	,599
ENG39 Je me suis senti complètement absorbé dans les activités du cours	,725	-,062	,221

Recherche d'une structure simple

Structure finale = structure simple :

- Idéalement, saturations ≥ 0.40 sur UN facteur
- Sens conceptuel / interprétable des facteurs

Pituch et Stevens, 2016; Tabachnick et Fidell, 2007 ;

Worthington et Whittaker, 2006

Échelle d'engagement :

- Structure finale : 20 énoncés
- Pas d'optimisation de longueur nécessaire
- Stabilité test-retest (cf. 2^e échantillon)

Échelle d'engagement : structure finale

**Dimension
émotionnelle-cognitive
(44.02% variance expliquée)**

Intérêt, enthousiasme,
Investissement cognitif,
absorption

« J'avais hâte de participer
aux activités de ce cours »
« Je me suis senti
complètement absorbé
dans les activités du
cours »

**Dimension
sociale
(9,79 %)**

Interactions, liens de
confiance avec les
autres étudiants

« J'ai échangé avec
d'autres étudiants à
propos du cours »
« Je me suis senti lié au
groupe d'étudiants de
ce cours »

**Dimension
comportementale
(6,67 %)**

Attention, temps et
efforts

« J'ai consacré du
temps à ce cours »
« J'ai donné le meilleur
de moi-même dans ce
cours »

En cohérence avec des études récentes sur
l'engagement des étudiants en enseignement
supérieur (Redmond et al., 2018; Zhoc et al., 2019)

Éléments de preuve de conséquence

Conséquences de l'usage de l'EMEECH :

- rapide à remplir pour les étudiants
- fournit des informations détaillées aux formateurs sur différentes dimensions de l'engagement
- s'abstenir d'évaluer l'engagement des étudiants sur la base du seul score global
 - tenir compte des caractéristiques personnelles des étudiants

RÉFÉRENCES

Références

- American Educational Research Association (AERA). (2014). *Standards for Educational and Psychological Testing*. AERA.
- Bourque, J., Poulin, N. et Cleaver, A. F. (2007). Évaluation de l'utilisation et de la présentation des résultats d'analyses factorielles et d'analyses en composantes principales en éducation. *Revue des sciences de l'éducation*, 32(2), 325-344. <https://doi.org/10.7202/014411ar>
- Clark, L. A. et Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319. <https://doi.org/10.1037/1040-3590.7.3.309>
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (4e éd.). SAGE Publications. **Voir en particulier le chapitre 5 (Guidelines in scale development).**
- Fredricks, J. A., Blumenfeld, P. C. et Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>
- Fredricks, J. A., Filsecker, M. et Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction*, 43, 1-4. <https://doi.org/10.1016/j.learninstruc.2016.02.002>
- Fredricks, J. A., Reschly, A. L. et Christenson, S. L. (2019). Conclusion: Status of student engagement interventions. Dans J. A. Fredricks, A. L. Reschly et S. L. Christenson (dir.), *Handbook of student engagement interventions* (p. 375-389). Elsevier.
- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological measurement*, 66(3), 393-416.
- Laveault, D. et Grégoire, J. (2014). *Introduction aux théories de tests en psychologie et en sciences de l'éducation*. De Boeck Supérieur.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, 4(1), 84.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. <https://psycnet.apa.org/doi/10.1037/0003-066X.50.9.741>
- Pituch, K. A. et Stevens, J. (2016). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS* (6th ed.). Routledge/Taylor & Francis Group.
- Redmond, P., Abawi, L. A., Brown, A., Henderson, R. et Heffernan, A. (2018). An online engagement framework for higher education. *Online Learning*, 22(1), 183-204. <http://dx.doi.org/10.24059/olj.v22i1.1175>
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Streiner, D. L., Norman, G. R. et Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford University Press.
- Sturgis, P., Roberts, C. et Smith, P. (2014). Middle alternatives revisited: How the neither/ nor response acts as a way of saying "I don't know"? *Sociological Methods & Research*, 43(1), 15-38. <https://doi.org/10.1177/0049124112452527>
- Tabachnick, B. G. et Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson.
- Worthington, R. L. et Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838. <https://doi.org/10.1177/0011000006288127>
- Yong, A. G. et Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79-94. <https://doi.org/10.20982/tqmp.09.2.p079>
- Zhoc, K. C., Webster, B. J., King, R. B., Li, J. C. et Chung, T. S. (2019). Higher Education Student Engagement Scale (HESES): Development and psychometric evidence. *Research in Higher Education*, 60(2), 219-244. <https://doi.org/10.1007/s11162-018-9510-6>

Merci de votre attention!

Geraldine.Heilporn@usherbrooke.ca

Twitter : @Gheilporn