



## Usages et mésusages des analyses factorielles exploratoires : un exemple à partir de la version courte-française du Questionnaire for Teacher Interaction (QTI)

## Uses and Misuses of Exploratory Factor Analyses: An Example from the French-Canadian Short Version of the Questionnaire for Teacher Interaction (QTI)

Ibtissem BEN ALAYA<sup>1</sup>, David DUMAS<sup>2</sup>, Vincent GRENON<sup>3</sup>, Jean-François DESBIENS<sup>4</sup> & Naila BALI<sup>5</sup>

**Article Type:** Research Article

**Application Date:** 23.03.2019

**Accepted Date:** 10.11.2020

**To Cite This Article:** Ben Alaya, I., Dumas, D., Grenon V., Desbiens, J. F. & Bali, N. (2020). Usages et mésusages des analyses factorielles exploratoires : un exemple à partir de la version courte-française du Questionnaire for Teacher Interaction (QTI). *Anadolu University Journal of Education Faculty*, 4(4), 344-364.

**RESUMÉ :** Les démarches de validation factorielle des questionnaires psychométriques traduits en français sont entachées par la présence de plusieurs faiblesses méthodologiques et statistiques. L'analyse factorielle exploratoire (AFE) est considérée parmi les méthodes d'extraction les plus utilisées pour vérifier la validité des construits des questionnaires en sciences humaines et sociales (Osborne et Costello, 2009). Néanmoins, la plupart des études basées sur les AFE sont marquées par des usages non appropriés ou des interprétations litigieuses de ce type d'analyses. Dans cet article, le *Questionnaire on Teacher Interaction* (QTI) a été utilisé comme exemple des questionnaires psychométriques traduits en français. Très récemment, ses démarches de validation factorielle ont fait l'objet d'une critique approfondie (Ben Alaya, Grenon, Desbiens et Bali, 2018). Pour avancer les recherches dans ce sens, nous avons comparé les résultats de deux démarches d'AFE : la première est inspirée des études antérieures ayant validé le QTI alors que la deuxième suit les étapes et les recommandations des spécialistes en AFE. En comparant les deux résultats, nous avons montré à quel point les décisions opérées et l'usage non approprié des AFE peuvent influencer le résultat et la solution factorielle trouvée.

**Mots clés :** Questionnaire on Teacher Interaction, Validité, Factorielle, Analyse factorielle exploratoire.

<sup>1</sup> Laboratoire d'Études sur les parcours scolaires et les influences sociales, LÉPSIS, Université de Québec à Montréal (UQAM). ben\_alaya.ibtissem@uqam.ca , ORCID:0000-0003-2032-5446

<sup>2</sup> Centre de services scolaires du Val-des-Cerfs. David.Dumas@Usherbrooke.ca , ORCID:0000-0002-2059-6177

<sup>3</sup> Université de Sherbrooke, Faculté d'éducation, CRIFPE. Vincent.Grenon@usherbrooke.ca , ORCID:0000-0002-7085-0081

<sup>4</sup> Université de Sherbrooke, Faculté des sciences de l'activité physique, CRIFPE. Jean-Francois.Desbiens@usherbrooke.ca , ORCID:0000-0003-4668-8675

<sup>5</sup> Institut Supérieur du sport et de l'Éducation Physique de Ksar Saïd, Université de la Manouba, Laboratoire Optimisation de la Performance Sportive. naila\_bali@yahoo.fr , ORCID:0000-0001-7751-6273

## 1. INTRODUCTION

Les analyses factorielles (AF), acception francisée des «factor analysis», trouvent leur origine dans les travaux de Spearman effectués au début des années 1900 sur les compétences et les habiletés humaines (Harman, 1976; Yong et Pearce, 2013). Elles désignent l'ensemble des méthodes statistiques ayant au moins trois utilisations principales: 1. la compréhension de la structure d'un ensemble de variables; 2. la construction d'un questionnaire qui permet de mesurer des variables latentes; 3. la réduction d'un ensemble de données en un petit groupe de facteurs faciles à gérer et qui conserve le maximum d'informations originales (Field, 2009). Beavers, Lounsbury, Richards, Huck, Skolits et Esquivel (2013) ont précisé qu'il ne s'agit pas d'une méthode statistique unique, mais plutôt d'un ensemble de démarches statistiques qui partagent une méthodologie et une fonctionnalité particulières. Il existe, une gamme de choix et de décisions auxquels les chercheurs peuvent avoir recours pour améliorer la précision de leurs analyses et la qualité de leurs solutions factorielles (Fabrigar, Wegener, MacCallum et Strahan, 1999).

Tabachnick, Fidell et Osterlind (2001) qualifient l'AF de processus cyclique continuellement raffiné lors duquel il est important de comparer les solutions, de justifier les choix et de prendre les bonnes décisions jusqu'à l'atteinte de la solution factorielle la plus satisfaisante. Les chercheurs qui souhaitent effectuer ce type d'analyses sont confrontés à deux problèmes majeurs: 1) la documentation scientifique est pauvre en informations relatives au bon usage des AF (Field, 2013); 2) l'interprétation de l'AF ne nécessite pas seulement une bonne connaissance des statistiques et de la méthodologie; elle doit aussi s'accompagner d'une compréhension approfondie des aspects théoriques et conceptuels de l'instrument utilisé (Beavers et al., 2013).

Certains auteurs ne font pas la différence entre l'AF et l'analyse en composantes principales (ACP) (Osborne et Costello, 2009; Steiger, 1990). Cette confusion peut être expliquée par le fait que l'ACP forme le choix par défaut dans de nombreux logiciels statistiques comme SPSS et SAS (Osborne et Costello, 2009). La différence entre les deux est simple: l'ACP est une méthode d'extraction qui se base sur la variance totale (ne distingue pas la variance commune, unique et d'erreur) (Beavers et al., 2013). L'AF est une technique d'analyse qui élimine la variance unique et l'erreur des calculs pour estimer les facteurs à partir des variances communes (Field, 2013). Elle regroupe plusieurs méthodes d'extractions: 1) l'analyse en composantes principales (ou ACP); 2) la méthode du maximum de vraisemblance; 3) la méthode des moindres carrés non pondérés; 4) la méthode des moindres carrés non généralisés; 5) la méthode alpha maximisation; 6) la méthode de factorisation en projection et 7) la méthode de factorisation en axes principaux assimilée à l'analyse factorielle exploratoire (AFE) (Osborne et Costello, 2009). Selon Beavers et al. (2013), l'AFE figure parmi les méthodes de validation les plus citées en sciences humaines et sociales pour au moins trois raisons: 1) elle constitue la première étape de construction de nouveaux questionnaires (Osborne et Costello, 2009); 2) elle sert à valider la structure latente des questionnaires et 3) elle ne dépend pas de la normalité des variables (Fabrigar et al., 1999). La documentation anglophone récente comprend quelques articles et livres qui présentent les meilleures pratiques de l'AFE (Beavers et al., 2013; Field, 2009; Osborne et Costello, 2009; Yong et Pearce, 2013). Cependant, les guides et les ouvrages scientifiques traitant ce sujet en français demeurent plutôt rares. C'est la raison pour laquelle les processus de validation des questionnaires y compris ceux traduits d'une langue tierce vers le français présentent fréquemment des faiblesses méthodologiques qui ont préoccupé (Vallerand, 1989) et préoccupent encore aujourd'hui les chercheurs francophones. Ces faiblesses peuvent être analysées sur deux plans : 1) la qualité de la traduction et 2) la démarche de validation

proprement dite. C'est sur ce dernier aspect que s'attarde le présent article. Il traite plus spécifiquement de la validation factorielle effectuée à l'aide des AFE.

Nous avons choisi d'appliquer des AFE sur la version courte du « Questionnaire on Teacher Interaction » (QTI) comme exemple des questionnaires traduits en français et critiqués récemment pour des faiblesses constatées dans les démarches de validation factorielle. Malgré la large diffusion de ce questionnaire (utilisé dans plus de 40 pays et traduit dans plus de 20 langues (den Brok, Brekelmans, Levy et Wubbels, 2002), la plupart des démarches de validation factorielle du QTI sont entachées par des lacunes méthodologiques plus au moins importantes (Ben Alaya, Grenon, Desbiens et Bali, 2018).

Cet article s'inscrit dans la foulée de cette critique. Nous avons comparé deux types de démarches d'AFE. Dans la première, nous avons suivi les mêmes étapes d'AFE que celles utilisées dans les démarches de validation du QTI utilisées antérieurement même si elles ne correspondent pas rigoureusement aux étapes recommandées par les experts. Dans la seconde, nous avons cherché à appliquer aussi systématiquement que possible les critères et les étapes recommandés par les spécialistes des AFE (ex. Osborne et Costello, 2009). En comparant les résultats obtenus par chacune de ces deux démarches, nous souhaitons montrer à quel point les décisions opérées peuvent influencer le résultat et la solution factorielle trouvée. Nous avons choisi de diffuser cet article en français en décrivant les étapes et les exigences des AFE (démarche 2) pour offrir au lectorat francophone un guide en français sur l'usage approprié des AFE.

## 2. METHODE

Cette étude consiste en une analyse secondaire des données collectées par Dumas (2009). Dans cette partie, nous proposons en premier lieu une courte description de l'étude de laquelle sont tirées les données utilisées dans cet article. En second lieu, l'instrument utilisé pour la collecte des données sera présenté. Enfin, nous décrirons les deux démarches d'AFE utilisées pour répondre aux objectifs de cette étude.

### 2.1. Description de l'étude source

La recherche menée par Dumas (2009) avait pour objectif l'étude de la relation enseignant-élève dans les cours d'éducation physique et à la santé (ÉPS) selon les perceptions des élèves. Les données recueillies ont été collectées entre novembre 2004 et décembre 2005 à partir d'un échantillon non probabiliste constitué de 2261 élèves de secondaire 1 à 5 (60,5% de garçons et 39,5% de filles). Ces élèves ont été sélectionnés sur une base entièrement volontaire à partir d'un bassin de neuf écoles secondaires réparties dans quatre régions administratives (Laval, Montréal, Chaudière-Appalaches et Montérégie) du Québec (Canada). Le chercheur intéressé peut trouver une description détaillée de cet échantillon à l'adresse : (<http://savoirs.usherbrooke.ca/handle/11143/868>). Nous avons effectué les analyses à partir des données anonymisées encodées dans la base de données. Les données brutes n'étaient pas accessibles.

### 2.2 Instrument de collecte des données

Les données du présent article ont été collectées à partir du « Questionnaire on Teacher Interaction » (QTI). Cet instrument a vu le jour aux Pays-Bas et il a été adapté au contexte scolaire par (Wubbels, Creton et Hooymayers, 1985). Le QTI est basé sur un modèle théorique bidimensionnel en circomplexe de Leary (1957). Les deux dimensions sont nommées respectivement soutien et contrôle.

Elles sont représentées par deux axes supposés orthogonaux du circomplexe. Ce circomplexe est formé de huit octants qui correspondent aux huit échelles du QTI à savoir : le leadership, la bienveillance, l'indulgence, la permissivité, l'incertitude, l'insatisfaction, la punition et l'exigence (pour plus de détails voir Ben Alaya et al., 2018). On peut distinguer plusieurs versions du QTI selon le nombre d'énoncés par échelle et la langue dans laquelle il a été traduit. La version utilisée dans la recherche de Dumas (2009) et celle à laquelle nous référerons dans cet article renferme 32 énoncés, soit 4 énoncés par échelle. Il s'agit de la version la plus courte de ce questionnaire. Elle a été traduite et adaptée en français par les deux chercheurs Lapointe et Legault (1999) à partir de la version anglophone de 64 énoncés. Ces deux chercheurs prétendent que le QTI présente une validité et une fidélité suffisantes pour être utilisé dans le contexte québécois. Plus spécifiquement, en se basant sur un échantillon de 1764 élèves québécois, Lapointe et Legault ont vérifié aussi bien sa fidélité (la consistance interne) que sa validité factorielle et structurale.

### 3. ANALYSES STATISTIQUES

Cette étude présente les résultats des deux démarches d'AFE dans le but de montrer les conséquences du bon usage et du mésusage des AFE sur la vérification des qualités psychométriques de la version courte du QTI dans le contexte québécois de l'ÉPS au secondaire.

Dans la première démarche, nous avons adopté les étapes de l'AFE utilisées dans les démarches de validation précédentes du QTI (ex. Sivan, Chan et Kwan, 2014; Telli, den Brok et Cakiroglu, 2007; Wei, den Brok et Zhou, 2009) basées sur des normes moins strictes. Dans la deuxième démarche, nous avons adopté les critères et les étapes recommandés par les spécialistes des AFE (ex. Osborne et Costello, 2009). Le logiciel SPSS (version 19.0) a été utilisé pour faire les analyses statistiques. Dans les prochains développements, nous allons décrire en détail chacune de ces deux démarches.

#### 3.1. Première démarche d'AFE

Tel que mentionné plus haut, cette démarche a été inspirée d'études antérieures ayant validé le QTI (ex. Telli, et al., 2007; Wei et al., 2009). Dans tous les cas, les chercheurs semblent assumer que les huit échelles du QTI sont valides. En effet, tous amorcent leur démarche pour vérifier les qualités psychométriques du QTI par l'examen de la consistance interne des huit échelles par le biais de l'alpha de Cronbach. La plupart du temps, nous remarquons que la fidélité interne est confirmée malgré la présence de faibles valeurs d'alpha. Par exemple, Ben Alaya et al. (2018) ont constaté que Wei et al. (2009) ont confirmé la fidélité du QTI bien que la valeur alpha soit de  $< ,42$  pour la permissivité et  $< ,6$  pour respectivement l'exigence et l'incertitude. Nous avons donc débuté cette démarche animée par une supposition forte à savoir que les huit échelles du QTI reposent bien sur des regroupement cohérents et suffisants d'énoncés. S'en est suivi l'examen de la cohérence interne des échelles en acceptant volontairement des valeurs pour l'alpha de Cronbach inférieures au seuil fréquemment recommandé de  $,7$  (Hair, 2006).

Après avoir vérifié la consistance interne, les chercheurs effectuent leurs analyses factorielles à partir des échelles et non des énoncés pour prouver la structure bidimensionnelle du QTI (ex. : Telli et al., 2007; Wei et al., 2009). Conformément à ce qui a été rapporté par ces chercheurs, nous avons procédé par une AFE des huit échelles du QTI après avoir vérifié leur consistance interne respective.

Conformément aux démarches adoptées ultérieurement dans la validation du QTI, nous n'avons pas vérifié les prémisses des AFE qui touchent principalement : les communalités, la distribution normale des données dans chacune des échelles, la matrice de corrélation et la matrice anti-image.

Les AFE nécessitent l'extraction d'un nombre bien déterminé de facteurs sur SPSS. Nous avons remarqué que la majorité de chercheurs qui ont entrepris de valider le QTI assument l'existence de deux facteurs ou construits (ex. Coll, Taylor et Fisher, 2002 ; Fisher, Fraser et Rickards, 1997 ; Goh et Fraser, 1996 ; Henderson, Fisher et Fraser, 2000) : le soutien et le contrôle. Ce choix est expliqué par le fait que le modèle théorique sur lequel est basé le QTI suppose l'existence de ces deux dimensions indépendantes. C'est la raison pour laquelle dans ces études, l'extraction de deux facteurs n'est pas justifiée par un ou plusieurs critères d'extraction (den Brok et al., 2003). En conséquence, nous avons décrété, dans cette première démarche, l'extraction de deux facteurs sans appuyer cette décision sur quelque analyse que ce soit. Il s'agit d'une solution forcée à deux facteurs. Cette étape a produit une solution factorielle bifactorielle. Nous avons remarqué que dans la plupart des démarches de validation examinées, c'est la solution avec la rotation *Varimax* qui est retenue probablement parce que le modèle théorique du QTI suppose l'indépendance de soutien et du contrôle (Ben Alaya et al., 2018). Ceci justifie notre recours à la rotation *Varimax*.

Pour montrer que la solution factorielle reflète bel et bien la structure théorique du QTI, les chercheurs ayant validé ce questionnaire examinent la corrélation entre les variables qui supportent chacun de deux facteurs (ex. Legault et Lapointe, 1999). La création de ces deux variables nécessite le recours à des équations mathématiques pour calculer les « construits » associés à chaque facteur. Nous avons utilisé l'équation la plus simple, celle utilisée par Lapointe et Legault : « Soutien = (Bienveillance + Indulgence) - (Punition + Insatisfaction) ; contrôle = (Leadership + Exigence) - (Permissivité + Incertitude) » (p.7).

Étant toujours conforme aux démarches précédentes de validation du QTI, nous avons considéré que les deux variables seraient indépendantes à deux conditions : 1) s'il y a absence de corrélation significative ou 2) s'il y a une corrélation significative avec un coefficient de détermination marginal (ex.  $\leq 1\%$ ). Le choix du coefficient de corrélation de Spearman ou de Pearson est tributaire de la normalité de la distribution des données pour les deux variables.

### 3.2. Deuxième démarche d'AFE

Contrairement à la première démarche, nous avons essayé dans celle-ci de suivre un protocole bien établi, cohérent et exhaustif pour vérifier les qualités psychométriques de notre instrument de mesure (QTI). Ce protocole est inspiré des recommandations de spécialistes des AF (Beavers et al., 2013; Bourque et al., 2006; Ferguson et Cox, 1993; Field, 2013; Osborne et Costello, 2009). Selon ces derniers, il faut partir des énoncés pour parvenir à établir des échelles (facteurs) et vérifier leur conformité avec les attentes du cadre théorique qui sous-tend le questionnaire. Toutefois, pour pouvoir entamer des AFE, il faut avant tout s'assurer que les données utilisées s'y prêtent bien, c'est-à-dire qu'il faut en vérifier les prémisses. Il faut par la suite justifier le nombre de facteurs à extraire, le type de rotation et la méthode d'extraction avant d'amorcer les itérations menant à la solution factorielle finale. La comparaison entre cette dernière et la structure factorielle théorique du questionnaire permet de confirmer ou d'infirmer la validité factorielle du questionnaire. Après ces étapes seulement, il devient pertinent d'examiner la cohérence interne entre les énoncés de chaque facteur ou échelle (ex. par l'alpha de Cronbach) dans le but d'étudier la fidélité du questionnaire.

Nous avons vérifié, en premier lieu, si les six prémisses des AF sont respectées : 1) une taille d'échantillon d'au moins 300 sujets (Field, 2013); 2) des données de type continu ou ordinal (Yong et Pearce, 2013); 3) des communalités après extraction  $> ,4$  pour les 32 énoncés du QTI. Chaque item qui ne répondait pas à cette dernière prémisse a été retiré des analyses. 4) Les AF requièrent habituellement une distribution normale des données (Yong et Pearce). Cette quatrième prémisse peut être transgressée si l'objectif n'est pas de généraliser les résultats sur la population (Field, 2013). Fabrigar et al. (1999) et Osborne et Costello (2009) suggèrent le recours sur SPSS à des AF de type « *Factorisation en axes principaux* » (analyse factorielle exploratoire (AFE)) si la prémisse de la normalité est violée. 5) Il est aussi important de vérifier si la matrice de corrélation est appropriée à des AF. Il s'agit de s'assurer en premier lieu que cette matrice ne contienne pas trop d'énoncés présentant des corrélations faibles ( $< ,2$ ) ou élevées ( $> ,8$ ). Ceci diminue les risques de multicolinéarité et augmente les chances de produire des facteurs latents communs exprimant la variance totale (Ferguson et Cox, 1993). Deuxièmement, la valeur du déterminant de la matrice doit être supérieure ou égale au seuil minimum de  $10^{-5}$  pour qu'une solution factorielle soit mathématiquement possible (Field, 2013). Par la suite, le test de Bartlett doit être significatif pour pouvoir poursuivre les AF. Un test non significatif indique que la matrice est une matrice identité et nécessite l'arrêt immédiat des AF (Pett, Lackey et Sullivan, 2003). Il est important aussi d'examiner l'ajustement des énoncés aux facteurs trouvés. L'ajustement est insuffisant si l'indice de Kaiser-Meyer-Olkin (KMO) est inférieur à  $,6$ . Dans cette éventualité, l'AF est immédiatement interrompue (Bourque et al., 2006). 6) La sixième prémisse concerne la matrice de corrélation anti-image. Dans cette dernière, il est important de supprimer les éléments de la diagonale qui sont  $< ,5$  afin d'accroître les chances de produire des facteurs distincts et fiables (Yong et Pearce, 2013).

Après avoir vérifié toutes ces prémisses, il faut décider du nombre de facteurs à extraire. Contrairement aux démarches précédentes de validation du QTI qui choisissent toujours deux facteurs en se basant sur le modèle théorique-deux facteurs, les spécialistes en méthodologie proposent quelques méthodes d'extraction afin de justifier le nombre de facteurs à extraire (Beavers et al., 2013) : 1. la méthode de Kaiser (1960); 2. la méthode de Cattell (1966), 3. La méthode de Horn (1965) (O'Connor, 2000). Cette dernière figure parmi les méthodes les plus rigoureuses. Néanmoins, il est recommandé de comparer les résultats procurés par plusieurs méthodes (Field, 2013). C'est la raison pour laquelle nous avons comparé les résultats obtenus à l'aide des trois méthodes d'extraction pour décider du nombre de facteurs à extraire.

Rappelons que le modèle théorique du QTI sous-entend l'indépendance des dimensions de soutien et de contrôle. Ceci justifie l'utilisation de la rotation *Varimax* par les chercheurs qui ont validé ce questionnaire. Néanmoins, il ne faut pas oublier que l'AFE forme une phase exploratoire qui a pour but de confirmer ou infirmer la validité de construit. Par ailleurs, il n'est pas acceptable de prendre des décisions non justifiées et de supposer l'indépendance de deux facteurs sans la démontrer. La comparaison des résultats procurés par les rotations orthogonale (*Varimax*) oblique (*Oblimin*) semble être la meilleure façon de dépasser les critiques associées à chacune et pour vérifier la conformité entre la structure réelle et la structure théoriquement attendue du QTI (Beavers et al., 2013). Stafford et Bodson (2006) proposent dans ce cas d'interpréter la matrice de transformation factorielle (*Varimax*) et la matrice de corrélation factorielle (*Oblique*) pour décider du type de rotation le plus appropriée. Si on observe dans ces matrices que les corrélations entre les facteurs sont proches de zéro, il faut conserver la rotation orthogonale qui supporte l'indépendance des facteurs. Ce type de rotation est rejeté aussitôt que des corrélations faibles à moyennes apparaissent entre les facteurs (Field, 2013) même si le modèle théorique suppose le contraire.

La solution factorielle finale produite par des AFE doit respecter trois critères recommandés par les experts (ex. Osborne et Costello, 2009). 1. Les facteurs conservés ne doivent pas contenir d'énoncés complexes (Beavers et al., 2013). Entre autres, les énoncés ayant un écart de chargement sur deux facteurs ou plus  $< ,2$  doivent être supprimés. 2) Les énoncés de chaque facteur doivent avoir des communalités  $> ,4$ . 3) Seulement les facteurs ayant trois énoncés et plus sont retenus.

La consistance interne de chacun des facteurs conservés doit être testée à l'aide de l'alpha de Cronbach (Field, 2016). Contrairement à la première démarche, nous avons fixé un seuil plus rigoureux pour les valeurs d'alpha ( $> ,7$ ) pour au moins deux raisons. Le seuil de l'alpha de Cronbach varie positivement en fonction du nombre d'énoncés par échelle (Peterson, 1994). En plus, pour une étude réalisée en 2017, on ne peut plus considérer les seuils de 1967 (Nunnally) alors que des travaux plus récents ont recommandé des seuils plus rigoureux (Nunnally, 1978). Pour neutraliser l'effet des facteurs secondaires sur la valeur de l'alpha, Field (2009) recommande le recours à deux autres indices : l'alpha de Cronbach en cas de suppression d'item et la corrélation complète d'énoncés corrigés. Pour cette dernière, les énoncés ayant une corrélation  $< ,3$  doivent être supprimés. Cette étape nous permet de vérifier si les facteurs trouvés ont une consistance interne suffisante pour affirmer la fidélité de l'instrument, mais aussi elle nous autorise à créer sur SPSS de nouvelles variables associées à chacun des facteurs trouvés. Un examen de la corrélation entre les construits créés à partir des équations citées dans la première démarche est important pour vérifier s'ils sont effectivement indépendants comme le suppose la théorie.

## 4. DISCUSSION et RESULTATS

Les résultats de toutes les analyses statistiques sont présentés en deux parties. Nous présentons les résultats de la première démarche d'AFE selon les étapes décrites ci-haut. Nous exposons par la suite les résultats obtenus avec la deuxième démarche que nous pouvons qualifier d'une démarche organisée systématiquement et basée sur des critères et seuils rigoureux. Nous discutons à la fin les résultats trouvés.

### 4.1. Vérification des qualités psychométriques du QTI selon la première démarche

#### 4.1.1. La consistance interne

Tel que décrit précédemment, nous avons commencé notre démarche par la vérification de la consistance interne des huit échelles du QTI. Le tableau (1) montre que les valeurs pour les coefficients alpha de Cronbach de chacune des échelles se situent entre ,66 et ,81. Rappelons que des recherches sur le QTI ont toléré de faibles valeurs d'alpha ( $,5$  ou  $,6$  (Ben Alaya et al., 2018)). En nous basant sur ces antécédents, nous considérons que les huit échelles ont des consistances internes suffisantes et par ailleurs une fidélité suffisante pour passer à l'étape suivante, soit l'AFE.

**Table 1:** Consistance interne des huit échelles du QTI

Échelles	Énoncés	Alpha de Cronbach
Leadership	9,17,1, 25	,74
Bienveillance	3, 11, 19, 27	,81
Indulgence	13,21,23, 29	,78
Permissivité	5,7,15, 31	,66
Incertitude	8,16,24,32	,72
Insatisfaction	6, 14, 22, 28	,72
Punition	4,12,20,30	,81
Exigence	2,10,18,26	,71

#### 4.1.2. La validité factorielle

L'AFE des huit échelles du QTI avec la rotation Varimax nous a permis de produire une solution à deux facteurs. Ces derniers expriment 70 % de la variance totale. Le tableau (2) décrit la répartition des échelles selon chaque facteur.

**Table 2:** Solution factorielle après la rotation Varimax.

Échelles	Facteur 1	Facteur 2
Leadership	<b>,81</b>	-,33
Bienveillance	<b>,80</b>	-,28
Indulgence	<b>,75</b>	-,26
Permissivité	<b>,74</b>	-,42
Incertitude	-,38	<b>,78</b>
Insatisfaction	-,38	<b>,76</b>
Punition	-,24	<b>,50</b>
Exigence	-,11	<b>,43</b>

Pour vérifier si les deux facteurs reflètent bien le soutien et le contrôle, nous avons créé deux variables sur SPSS à l'aide des équations de Lapointe et Legault (1999). Les tests de Kolmogorov-Smirnov pour la première variable ( $.13$  ;  $p < .001$ ) et la deuxième variable créée ( $.15$  ;  $p < .001$ ) sont significatifs. Ceci indique que les deux variables ne suivent pas une distribution normale, justifiant par ailleurs notre recours au coefficient de Spearman pour tester la corrélation entre elles. Nous avons observé une corrélation significative ( $r = -.074$  ;  $p < .001$ ) entre les deux variables. Cependant, nous remarquons que le coefficient de détermination ( $r^2 = .005$ ) est très faible. La proportion de la variance expliquée est de ,5 % de la variance totale. Ceci nous permet de justifier l'indépendance des deux variables trouvées. Les deux facteurs trouvés se comportent conformément à la prédiction théorique du QTI.

### 4.1.3. *Qualités psychométriques suffisantes du QTI*

Selon un protocole fondé sur des critères minimaux tirés des études antérieures traitant de la validation du QTI, nous avons établi que les huit échelles de ce questionnaire ont une consistance interne suffisante, ce qui nous permet de confirmer la fidélité du QTI. La validité factorielle est aussi confirmée puisque l'AFE a donné lieu à une structure bidimensionnelle conforme au modèle théorique en termes de nombre de dimensions (deux variables : soutien et contrôle).

## 4.2. *Vérification des qualités psychométriques du QTI selon la première démarche*

Contrairement à la démarche précédente, cette démarche a été organisée systématiquement selon les recommandations des experts des AFE (ex. Beavers et al., 2013; Field, 2013). Tous les choix et les prémisses sont justifiés. Comme expliqué ci-haut, cette démarche débute par la vérification des prémisses des AF.

### 4.2.1. *Prémisses des analyses factorielles*

Avant d'entamer les AF, nous avons vérifié que les six prémisses des AF sont respectées. 1) La taille d'échantillon est suffisante car  $N = 2261 > N = 300$ . 2) Des communalités initiales supérieures à ,4 suggèrent que les chances d'obtenir une solution factorielle stable sont bonnes. 3) Les variables sont de type ordinal et elles 4) ne suivent pas une distribution normale. Ceci justifie notre recours à une méthode d'extraction bien déterminée : l'analyse factorielle exploratoire de type *factorisation en axes principaux* (AFE). 5) En ce qui concerne la matrice de corrélation, nous avons dû supprimer trois énoncés (18, 20 et 22) pour obtenir un déterminant correspondant au seuil minimal de  $10^{-5}$  (Field, 2013). La matrice de corrélation entre les 29 énoncés restants est une matrice appropriée à des AFE puisque le test de Bartlett est significatif ( $p < ,001$ ) et l'indice KMO est largement supérieur à ,6 (KMO = ,9). Enfin, 6. les éléments de la diagonale de la matrice de corrélation anti-image sont  $> ,8$ .

### 4.2.2. *Extraction des facteurs*

Les méthodes d'extraction de Kaiser et de Cattell et enfin Horn ont donné respectivement 4, 3 et 3 facteurs. Étant donné que les deux derniers critères ont donné le même résultat et que le critère de Horn figure parmi les critères d'extraction les plus rigoureux, trois facteurs ont été extraits. Ces facteurs expriment 48,05 % de la variance totale ce qui est 21,95 % inférieur à ce qui a été obtenu par le biais de la première démarche.

### 4.2.3. *Rotation des axes*

L'interprétation de la première solution sans rotation était difficile. Ceci explique notre recours à la rotation des axes tel que recommandé par Pett et al. (2003). Le modèle théorique du QTI suppose l'indépendance de deux facteurs ce qui amène les chercheurs à choisir la rotation *Varimax*. Néanmoins, la nature exploratoire des AF nécessite la justification de l'indépendance des facteurs. C'est la raison pour laquelle nous avons procédé par deux types de rotation : *Varimax* et *Oblimin*. Nous avons examiné par la suite la corrélation entre les facteurs pour conserver la solution la plus appropriée (tableau 4).

#### 4.2.4. Interprétation de la solution factorielle

Comme le montre le tableau 3, la première solution avec la rotation *Varimax* renferme 22 énoncés et la deuxième solution avec la rotation *Oblimin* en compte autant. La seule différence entre les deux solutions est l'énoncé 23 qui est complexe seulement avec la rotation *Varimax*. L'examen de la matrice de corrélation factorielle (Tableau 4) justifie l'existence des corrélations faibles à modérées entre les trois facteurs ( $.21 < r < .46$ ). Par ailleurs, la solution avec la rotation *Varimax* doit être rejetée étant donné que ce type de rotation exige l'indépendance entre les facteurs trouvés. La solution avec la rotation *Oblimin* qui autorise la corrélation entre les variables (Field, 2013) sera plus appropriée dans notre cas. Sur cette base, nous retenons la solution avec la rotation *Oblimin* qui comprend trois facteurs : facteur 1 (19, 21, 27, 5, 3, 13, 11, 29, 9, 1, 17, 15, 31, 7, 25), facteur 2 (24, 16, 8 et 32) et facteur 3 (2, 10 et 26). Les trois facteurs retenus (tableau 3) respectent les critères cités ci-dessus à savoir la suppression des énoncés complexes (énoncés barrés), toutes les communalités sont  $> 0,4$  et chaque facteur est formé par au moins trois énoncés.

**Table 3 :** Comparaison entre les solutions factorielles après rotation *Varimax* et *Oblimin*

Matrice factorielle après la rotation <i>Varimax</i>				Matrice factorielle après la rotation <i>Oblimin</i>			
Item	Facteur 1	Facteur 2	Facteur3*	Item	Facteur 1	Facteur 2	Facteur3*
E19	<b>0,773</b>	0,002	0,034	E19	<b>0,773</b>	0,002	0,034
E21	<b>0,697</b>	-0,041	-0,021	E21	<b>0,697</b>	-0,041	-0,021
E27	<b>0,693</b>	-0,019	0,076	E27	<b>0,693</b>	-0,019	0,076
E5	<b>0,679</b>	0,024	-0,039	E5	<b>0,679</b>	0,024	-0,039
E3	<b>0,657</b>	-0,006	-0,103	E3	<b>0,657</b>	-0,006	-0,103
E13	<b>0,655</b>	-0,040	-0,016	E13	<b>0,655</b>	-0,040	-0,016
E11	<b>0,635</b>	-0,109	-0,029	E11	<b>0,635</b>	-0,109	-0,029
E29	<b>0,635</b>	-0,002	-0,106	E29	<b>0,635</b>	-0,002	-0,106
E9	<b>0,617</b>	-0,116	0,220	E9	<b>0,617</b>	-0,116	0,220
E1	<b>0,606</b>	-0,116	0,160	E1	<b>0,606</b>	-0,116	0,160
E17	<b>0,555</b>	-0,173	0,098	E17	<b>0,555</b>	-0,173	0,098
E15	<b>0,543</b>	0,155	-0,090	E15	<b>0,543</b>	0,155	-0,090
E31	<b>0,540</b>	0,151	-0,023	E31	<b>0,540</b>	0,151	-0,023
E23	<b>0,522</b>	-0,048	-0,257	E23	<b>0,522</b>	-0,048	-0,257
E7	<b>0,511</b>	0,018	-0,199	E7	<b>0,511</b>	0,018	-0,199
<del>E25</del>	<del><b>0,481</b></del>	<del><b>0,289</b></del>	<del><b>0,131</b></del>	E24	0,110	<b>0,680</b>	0,015
E16	-0,106	<b>0,628</b>	0,069	E16	0,044	<b>0,659</b>	0,002
E8	-0,186	<b>0,623</b>	0,109	E8	-0,044	<b>0,631</b>	0,029
E32	-0,190	<b>0,515</b>	0,014	E32	-0,086	<b>0,516</b>	-0,062
<del>E14</del>	<del><b>0,323</b></del>	<del><b>0,453</b></del>	<del><b>0,382</b></del>	<del>E14</del>	<del>-0,208</del>	<del><b>0,394</b></del>	<del><b>0,301</b></del>
<del>E6</del>	<del><b>0,333</b></del>	<del><b>0,443</b></del>	<del><b>0,356</b></del>	<del>E6</del>	<del>-0,225</del>	<del><b>0,381</b></del>	<del><b>0,273</b></del>
<del>E28</del>	<del>0,301</del>	<del>0,397</del>	<del>0,367</del>	<del>E28</del>	<del>-0,198</del>	<del><b>0,339</b></del>	<del><b>0,295</b></del>
E26	-0,098	0,064	<b>0,602*</b>	E26	-0,022	0,014	<b>0,602*</b>
E2	-0,029	-0,022	<b>0,596*</b>	E10	0,061	0,007	<b>0,579*</b>
E10	-0,014	0,036	<b>0,563*</b>	<del>E12</del>	<del>-0,202</del>	<del><b>0,310</b></del>	<del><b>0,465</b></del>
<del>E12</del>	<del><b>0,318</b></del>	<del><b>0,383</b></del>	<del><b>0,534</b></del>	<del>E4</del>	<del>-0,235</del>	<del><b>0,273</b></del>	<del><b>0,400</b></del>
<del>E4</del>	<del><b>0,334</b></del>	<del><b>0,349</b></del>	<del><b>0,471</b></del>	<del>E30</del>	<del><b>0,255</b></del>	<del><b>0,284</b></del>	<del><b>0,338</b></del>
<del>E30</del>	<del><b>0,348</b></del>	<del><b>0,360</b></del>	<del><b>0,415</b></del>	E2	0,034	-0,063	<b>0,616*</b>

Barré : énoncés complexes supprimés de l'analyse.

Table 4: Matrice de transformation factorielle

Facteur	1	2	3
1	0,800	-0,470	-0,374
2	0,570	0,397	0,720
3	0,190	0,789	-0,585

#### 4.2.5. La consistance interne

Rappelons que la solution avec la rotation *Oblimin* a été conservée. La consistance interne entre les énoncés de chacun de ses trois facteurs a été évaluée à l'aide de l'alpha de Cronbach. Un alpha de,9 a été observé entre les 16 énoncés du premier facteur (19, 21, 27, 5, 3, 13, 11, 29, 9, 1, 17, 15, 31, 23, 25 et 7). Un coefficient alpha de,7 a été observé respectivement entre les quatre énoncés (8, 16, 24, et 32) du deuxième facteur et les trois énoncés (2, 10 et 26) du troisième. La suppression d'autres énoncés n'améliore pas les valeurs trouvées. Pour les trois facteurs, la corrélation complète des énoncés corrigés montre des valeurs  $>.4$ . Ceci indique l'existence d'une bonne consistance entre les énoncés de chaque facteur nous permettant de créer trois variables sur SPSS associées à chaque d'entre eux.

#### 4.2.6. Corrélation entre les variables

Pour nommer ces variables, nous nous sommes basés sur le circomplexe de Leary (1957). Le premier construit sera nommée coopération-soumission (CS) vu qu'il regroupe les énoncés situés à droite de l'axe de soutien et que le nombre d'énoncés de la permissivité (4 énoncés : 5, 7, 15 et 31) est plus élevé que celui de leadership (3 énoncés : 9, 17 et 1). L'exigence (E) (3 énoncés : 2, 10 et 26) et l'incertitude (I) (4 énoncés : 8, 16, 24 et 32) correspondent respectivement au deuxième et au troisième construits.

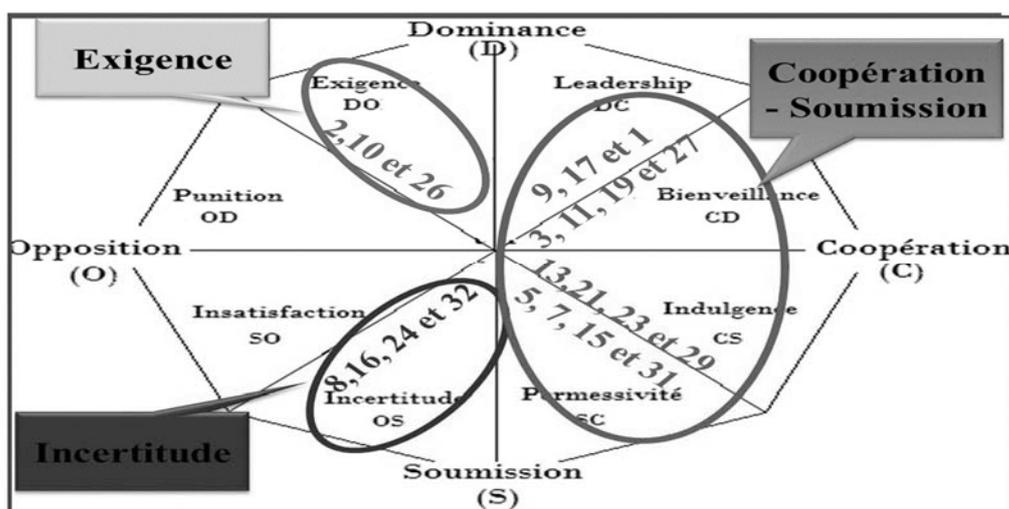


Figure 1: Solution factorielle à trois facteurs

Le test de Kolmogorov-Smirnov pour les trois variables trouvées (coopération-soumission ( $,09$  ;  $p < ,001$ ), exigence ( $,21$  ;  $p < ,001$ ) et incertitude ( $,21$  ;  $p < ,001$ ) est significatif dénotant une fois de plus une distribution des données non conforme au principe de normalité. Ceci explique notre recours au coefficient de Spearman pour étudier le lien entre les trois variables.

Comme le montre le tableau 6, des corrélations négatives et significatives sont observées entre les variables 1 et 2 ( $r_{1s} = -,19$  ;  $p < ,01$ ) et les variables 1 et 3 ( $r_{2s} = -,32$  ;  $p < ,01$ ) tandis que la corrélation entre les variables 2 et 3 ( $r_{3s} = ,11$  ;  $p < ,01$ ) est statistiquement positive. Les coefficients de détermination des trois couples de variables sont faibles ( $r_{1s}^2 = 0,04$  ;  $r_{2s}^2 = ,1$  ;  $r_{3s}^2 = ,01$ ).

**Tableau 6.** *Corrélation entre les trois variables créées.*

		Corrélations			
		Variables	1	2	3
Rho de Spearman	1		1,00	-0,19**	-0,32**
	2		-0,19**	1,00	0,11**
	3		-0,32**	0,11**	1,00

\*\* La corrélation est significative à  $p = ,01$  (bilatéral).

#### 4.2.7. Des qualités psychométriques insuffisantes

Pour conclure, contrairement aux résultats obtenus par la première démarche d'AFE, ceux obtenus par l'utilisation de la démarche rigoureuse nous ont permis d'obtenir trois facteurs démontrant une bonne consistance interne, mais ne reflétant pas la structure théorique bidimensionnelle du QTI.

### 4.3. Discussion des résultats

Cet article fait partie des rares études qui ont critiqué les démarches de validation du QTI. Il s'inscrit aussi une continuité des études ayant critiqué l'usage non approprié des AFE dans la validation des questionnaires psychométriques traduits de l'anglais au français à partir de l'exemple du QTI.

Les résultats produits dans cette étude montrent qu'un même ensemble de données recueillies par un même questionnaire peut démontrer des qualités psychométriques suffisantes et non suffisantes selon les choix et le degré de rigueur des stratégies d'analyse factorielle choisies. Par ailleurs, une attention particulière doit être accordée aux démarches et aux étapes d'analyses factorielles pour produire des résultats rigoureux.

Si l'on s'appuie sur les résultats de la première démarche d'AFE qui ne suit ni l'ordre des étapes ni les seuils recommandés par les experts, nous pourrions soutenir que le QTI possède une validité et une fidélité suffisantes pour être utilisé dans le contexte particulier de l'ÉPS au Québec. D'ailleurs, ce résultat converge avec ce qui a été reporté dans la plupart des démarches de validation précédentes du QTI que ce soit au Canada (Lapointe et Legault, 1999) ou dans les autres pays du monde (Passini, Molinari et Speltini, 2015; Sivan et al., 2014). Il sera tout à fait légitime de s'interroger sur le degré de rigueur des démarches de validations factorielles utilisées dans les études antérieures (choix et justification de type de rotation, de nombre de facteurs à extraire, etc.).

Afin de trouver une réponse à cette question, nous précisons que den Brok et al. (2003) sont les premiers, à notre connaissance, à avoir signalé la présence de quelques faiblesses dans les démarches

choisies antérieurement pour valider le QTI. Cependant, ils n'ont cité que quelques exemples d'une façon superficielle. Très récemment, Ben Alaya et al. (2018) ont poursuivi dans cet axe de recherche d'une manière plus approfondie. C'est en analysant 23 articles publiés dans plusieurs revues qu'ils ont pu critiquer les différents protocoles des AFE utilisées pour valider le QTI. Leur étude a permis de conclure qu'aucune démarche ne respecte toutes les prémisses et les exigences des AFE. Les choix des critères ou des stratégies visant à valider le questionnaire semblent fondés sur une rationalité faible et sont peu explicités. À titre d'exemple, les chercheurs supposent sans la vérifier l'existence de huit échelles et ils effectuent leurs AF sur ces échelles et non pas sur les énoncés (ex. Telli et al., 2007; Wei et al., 2009). Est-ce que ces échelles mesurent bien ce qu'elles sont censées mesurer ? Le mesurent-elles avec précision ? Comment sait-on que des énoncés ne posent pas de menaces à la validité des construits ?

D'une façon plus générale, l'utilisation non appropriée des AFE ne concerne pas seulement le QTI. Des problèmes similaires ont été observés dans d'autres domaines dont la psychologie, la communication et l'éducation (Ben Alaya et al., 2018).

DeVellis (2016) par exemple a montré que les chercheurs omettaient la plupart du temps des étapes dans la construction de leurs échelles et supposent que les énoncés de ces dernières conduisent à la formation de construits appropriés. Selon cet auteur, des échelles produites de cette façon ne partagent pas obligatoirement une cause commune ce qui peut aboutir à des conclusions erronées et à une déformation de la réalité étudiée. Ceci rejoint notre première interrogation quant à la légitimité de commencer la démarche de validation du QTI assumant l'existence des huit échelles.

Plus spécifiquement, en éducation, domaine de notre étude, (Bourque et al., 2006) ont montré jusqu'à quel point les exigences des AFE ne sont pas respectées. Selon cette étude, la plupart des chercheurs ne vérifient pas les prémisses des AFE, ne justifient pas leurs recours à un type particulier de rotation ni à un nombre déterminé de facteurs à extraire, etc. Les travaux réalisés avec le QTI ne semblent pas être l'exception.

Dans cet article, nous avons présenté un exemple concret montrant qu'une démarche avec des critères minimaux a permis de vérifier les qualités psychométriques de la version courte canadienne-française du QTI et d'obtenir un portrait conforme à sa théorie sous-jacente. Cependant, en utilisant la même base des données constituée à l'aide du questionnaire, les résultats de la deuxième démarche, plus rigoureuse celles-là, montrent plutôt que l'application des critères précis ne permet pas d'obtenir une solution factorielle acceptable. Plus précisément, en suivant toutes les étapes et en adoptant tous les seuils recommandés par les spécialistes des AFE, le QTI ne démontre plus des qualités psychométriques suffisantes pour être utilisé dans le contexte de l'ÉPS québécois. La comparaison entre les résultats de deux démarches montre bel et bien qu'en suivant les recommandations et les exigences des AF, les énoncés qui posent des problèmes de multi-colinéarité, de singularité ou de complexité disparaissent d'une étape à une autre. C'est la raison pour laquelle la structure factorielle initiale peut s'en trouver modifiée, voire fragilisée ceci rendant la validation du questionnaire plus incertaine. Ainsi, la sélection de seuils moins contraignants ou de stratégies insuffisamment justifiées peut conduire à la remise en cause des résultats publiés par ces instruments. D'ailleurs, cet article montre que, contrairement à ce qui est publié par Legault et Lapointe (1999), la version courte du QTI composée de 32 énoncés est structurellement déformée et ne permet pas d'étudier les comportements d'interaction des enseignants d'une manière totalement compatible à ses fondements théoriques. Ceci soulève plusieurs questions. Jusqu'à quel point peut-on se fier aux résultats publiés par cette version de QTI ? Serait-il possible de remettre en question aussi les autres versions de ce questionnaire malgré leur large diffusion ou s'agit-il plutôt un problème isolé concernant la version abrégée de ce questionnaire ?

Le retour sur les démarches de validation des autres versions de QTI pourra former une première piste pour répondre à ces questions. Il est clair que dans la plupart des cas, sous prétexte de l'espace limité des articles, les chercheurs ne fournissent pas assez de détails sur les démarches adoptées. Par exemple (Ben Alaya et al., 2018) constate que seulement trois articles sur les 23 analysés fournissent des tableaux complets de leurs solutions factorielles. Ces mêmes auteurs ont aussi montré que sur les 23 articles analysés, seulement deux se sont basés sur le test de Bartlett, le déterminant et l'indice KMO pour justifier leur recours aux AF. Or rappelons que ce sont des conditions qui en cas d'insuffisance peuvent justifier l'arrêt immédiat des procédures d'AF. À titre d'exemple, le test non significatif de Bartlett indique l'absence de combinaison entre les énoncés (matrice identité) ce qui rend le recours aux AFE inutile (Pett et al., 2003). De même, une valeur de KMO  $< 0,6$  indique que pour l'ensemble des énoncés analysés, il est impossible de construire des facteurs latents (Bourque et al., 2006). C'est ce qui signifie que le manque d'information rend difficile l'appréciation des résultats publiés.

Rappelons que les AFE sont des démarches statistiques qui doivent être effectuées itérativement jusqu'à l'atteinte de la solution factorielle la plus adéquate (Beavers et al., 2013). Ceci nécessite une bonne maîtrise des aspects statistiques de la part des chercheurs, mais aussi de celle des évaluateurs mandatés par les revues arbitrées. Cette connaissance peut limiter la marge de subjectivité du chercheur présentée par les experts en méthodologie comme étant une faiblesse des AF: « *A commonly cited limitation of exploratory factor analysis (EFA) is its level of subjectivity stemming from the many methodological decisions a researcher must make to complete a single analysis, with the accuracy of the results largely dependent upon the quality of these decisions* » (Beavers et al. 2013).

Parallèlement à l'aspect statistique, les utilisateurs des AFE sont, comme nous l'avons déjà mentionné, aussi appelés à bien maîtriser les fondements théoriques de leur questionnaire (Beavers et al., 2013). À titre d'exemple, le QTI est basé sur la communication interpersonnelle de Watzlawick, Beavers et Jackson (1967) et sur le modèle en circomplexe de Leary (1957). Ce dernier modèle suppose l'indépendance des dimensions de soutien et de contrôle (Wubbels et Levy, 1993). Sur cette base, il devient légitime de prioriser la rotation orthogonale qui conserve l'indépendance des facteurs (*Varimax*) (Osborne et Costello, 2009). Cependant ce choix nécessite une compréhension connaissance de la rotation orthogonale et de la différence entre celle-ci et la rotation oblique. Étant en phase exploratoire, la décision de conserver la solution factorielle avec la rotation *Varimax* dépend de l'interprétation de la matrice de transformation factorielle qui nous renseigne sur le lien entre les facteurs. Si la matrice démontre l'existence d'une corrélation significative entre les facteurs, le recours à la rotation oblique sera plus indiqué même si le modèle théorique suppose l'indépendance des facteurs.

À force d'ignorer sciemment ou non les prémisses et les exigences des AFE, n'y a-t-il pas un risque que les chercheurs publient des résultats peu fiables ? Pour le QTI par exemple, on peut distinguer au moins deux versions (ex. version de 48, version de 64 énoncés) entre la version originale de 77 énoncés et celle de 32 énoncés utilisée dans cet article (Passini et al., 2015). Si chaque démarche de validation est entachée par quelques lacunes méthodologiques et statistiques, les nouvelles versions restent-elles toujours fiables ? Il pourrait même devenir légitime de se demander s'il est toujours question des mêmes construits mesurés.

## 5. CONCLUSION ET RECOMMANDATIONS

Dans cet article nous avons présenté un exemple concret qui montre jusqu'à quel point le choix et les étapes effectués dans une démarche d'AFE peuvent influencer les résultats trouvés. Ce propos ne signifie pas qu'il faut arrêter d'utiliser les AFE, mais plutôt être méthodique dans tout le processus.

Dans le but de surmonter les problèmes liés au recours aux AFE et d'aider les chercheurs à profiter des avantages et de l'utilité de ces analyses, nous proposons quelques recommandations pour les futures recherches. D'ailleurs nous avons essayé dans cet article, à travers l'exemple du QTI, d'offrir aux chercheurs francophones une démarche détaillée des AFE qui comporte toutes les étapes, tous les détails et les seuils recommandés par les spécialistes de ces analyses.

En effectuant des AFE, le chercheur est appelé à adopter une froide neutralité et à accepter le résultat même s'il ne correspond pas à ce qui est attendu.

L'utilisateur des AFE doit justifier chaque décision en se référant aussi bien aux exigences statistiques des AFE qu'aux aspects théoriques des questionnaires.

Pour effectuer des AFE, au moins un membre de l'équipe de recherche doit avoir une formation approfondie en méthodologie et en statistiques en plus de son domaine d'étude. Ceci permet à l'équipe de respecter toutes les étapes et les exigences de ce type d'analyse. Également, au moins un membre du comité d'évaluation des revues doit posséder les connaissances méthodologiques et statistiques suffisantes pour pouvoir juger de la qualité des publications soumises. Ceci permet de diminuer la diffusion de travaux comportant des usages non appropriés des AF dans les démarches de validation que ce soit du QTI ou d'autres questionnaires psychométriques.

D'une façon plus générale, la production de connaissances valides et fiables porteuses de progrès sur les plans scientifique, social et professionnel est une responsabilité partagée par les chercheurs et les éditeurs des revues. Le chercheur est responsable d'élaborer le meilleur projet dont il est capable, de le mener d'une façon acceptable et d'en interpréter les résultats de la manière la plus impartiale possible. Cela suppose du chercheur qu'il s'interroge sur ses propres limites et qu'il n'entreprenne pas de travaux pour lesquels lui ou les membres de son équipe n'ont pas les ressources pour le mener à bien. Les éditeurs de revue et les comités de rédaction ont aussi leur part de responsabilité. Ils doivent s'assurer de la qualité des articles soumis ce qui suppose de solliciter les personnes les plus aptes à arbitrer les textes qui leur sont soumis.

Le chercheur doit être conscient de l'effet des spécificités sociales et culturelles sur les construits produits. Autrement dit, le questionnaire psychométrique est administré par des êtres humains à des êtres humains. Ces derniers ont leurs propres expériences, cultures, croyances, etc. En conséquence, il est tout à fait plausible de trouver qu'un questionnaire valide dans un contexte (pays, catégorie d'âge, sexe...) ne l'est pas dans un autre. Nous pouvons ici citer l'étude menée par Chandon et Bartikowski (2010) sur les facteurs qui risquent d'influencer la transposition des questionnaires d'un pays à un autre. Cette étude a démontré l'effet des facteurs culturels sur la qualité psychométrique des questionnaires (Davis, Douglas et Silk, 1981). À titre d'exemple, les chinois attribuent une signification socio-culturelle positive au chiffre (3) comme porteur du chance et négative au chiffre (4). Ceci peut influencer leurs choix dans les échelles de type Likert. Chandon et Bartikowski (2010) traitent le phénomène des choix les « plus socialement désirables » en précisant qu'il y a plusieurs personnes qui ne cochent pas la réponse la plus réelle mais plutôt la plus prestigieuse socialement. Parallèlement aux facteurs reliés à la culture, plusieurs autres facteurs internes ou externe aux répondants peuvent influencer les réponses sur un questionnaire et par ailleurs sa validité et fidélité. Citons comme exemple la motivation et les préférences du répondant.

Il y a des élèves qui n'aiment pas lire les questions et qui choisissent arbitrairement les extrémités ou le milieu des échelles (Baumgartner et Steenkamp, 2001). D'autres ne comprennent pas bien le sens des énoncés car le même mot peut avoir une signification différente selon le vécu, le niveau intellectuel ou l'origine du répondant (Chandon et Bartikowski, 2010).

Dumas (2009) fait partie des rares chercheurs ayant opté pour une stratégie d'analyse alternative dans le but de contourner certaines des limites des stratégies d'analyse habituellement préconisées liées notamment à la nature des données, à l'irrégularité des distributions et à la difficulté de confirmer les construits. Il a utilisé une méthode, l'analyse des correspondances multiples (ACM), qui appartient à la famille des analyses factorielles. L'ACM s'inscrit dans le cadre d'une approche descriptive, exploratoire et multi-dimensionnelle des données (Debois, 2008). Partant des énoncés modalisés plutôt que de construits supposés sous-jacents, elle a l'avantage d'être peu influencée par la normalité de la distribution des variables et la non linéarité des rapports entre elles. L'ACM permet de décrire les patrons d'organisation des modalités des énoncés du questionnaire étudiés selon leur proximité et leur distance sur un plan souvent bidimensionnel (Greenacre, 2006). Cette méthode permet de visualiser graphiquement les ressemblances et les disparités entre les répondants de même que les associations entre les différentes modalités de réponses. Ces patrons d'associations peuvent éventuellement permettre d'inférer l'existence de construits latents. L'analyse des données avec cette méthode peut s'avérer porteuse de significations utiles aux chercheurs qui s'intéressent à la dimension relationnelle de l'environnement d'apprentissage de même qu'aux enseignants qui cherchent des manières d'obtenir des rétroactions provenant des élèves à propos de leurs interventions et de leurs manières d'agir et d'être avec eux. Le chercheur intéressé par cette méthode peut trouver dans le mémoire de Dumas accessible en ligne une référence intéressante des exigences liées à ce type d'analyse ainsi qu'une illustration exhaustive de la façon dont il peut être appliqué avec le QTI.

## REFERENCES

- Barats, C. (2016). *Manuel d'analyse du web-2e éd.: En Sciences Humaines et Sociales*. Armand Colin.
- Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J. et Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical assessment, research et evaluation*, 18(6), 1-13.
- Ben Alaya, I., Grenon, V., Bali, N., et Desbiens, J. F. (2018). Analyse critique des démarches de vérification des qualités psychométriques du Questionnaire for Teacher Interaction (QTI). *e-JIREF*, 4(2), 3-26.
- Bourque, J., Poulin, N. et Cleaver, A. F. (2006). 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.
- Cattell, R. (1966). Validation and intensification of the sixteen personality factor questionnaire. *Readings in Clinical Psychology*, 12, 241-254.
- Davis, H. L., Douglas, S. P. et Silk, A. J. (1981). Measure unreliability: a hidden threat to cross-national marketing research? *The Journal of Marketing*, 45(2), 98-109.
- Desbois, D. (2008). L'analyse des correspondances multiples « à la hollandaise » : introduction à l'analyse d'homogénéité. *Revue Modulad*, 38, 194-244.
- den Brok, P., Brekelmans, M., Levy, J. et Wubbels, T. (2002). Diagnosing and improving the quality of teachers' interpersonal behaviour. *International journal of educational management*, 16(4), 176-184.
- den Brok, P., Levy, J., Wubbels, T. et Rodriguez, M. (2003). Cultural influences on students' perceptions of videotaped lessons. *International Journal of Intercultural Relations*, 27(3), 355-374.
- DeVellis, R. F. (2016). *Scale Development: Theory and Applications*. 4 ed. 26: Los Angeles: SAGE Publications. <https://books.google.ca/books?id=48ACCwAAQBAJ> (Consulté le 12/1/2017).
- Dumas, D. (2009). Perceptions des comportements interpersonnels d'enseignantes et enseignants et d'enseignantes d'éducation physique et à la santé par des élèves du secondaire. Unpublished Mémoire de maîtrise en sciences et en kinanthropologie, Université de Sherbrooke, Sherbrooke.
- Fabrigar, L. R., Visser, P. S. et Browne, M. W. (1997). Conceptual and methodological issues in testing the circumplex structure of data in personality and social psychology. *Personality and Social Psychology Review*, 1(3), 184-203.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. et Strahan Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. et Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, 4(3), 272.
- Ferguson, E. et Cox, T. (1993). Exploratory factor analysis: A users' guide. *International Journal of Selection and Assessment*, 1(2), 84-94.
- Field, A. (2009). *Discovering statistics using IBM SPSS statistics* (3th edition ed.). Los Angeles : SAGE.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th edition ed.). Los Angeles : SAGE.
- Fisher, D. L., Fraser, B. J. et Rickards, T. W. (1997). Gender and Cultural Differences in Teacher-Student Interpersonal Behavior. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).

- Fortin, L., Plante, A. et Bradley, M. (2011). Recension des écrits sur la relation enseignant-élève. *Chaire de recherche de la Commission scolaire de la Région-de-Sherbrooke sur la réussite et la persévérance scolaire. Document téléaccessible à l'adresse* <[http://www.csrs.qc.ca/fileadmin/user\\_upload/Page\\_Accueil/Enseignants/Fenetre\\_pedagogique/PEPS/Relation-maitre-eleve.pdf](http://www.csrs.qc.ca/fileadmin/user_upload/Page_Accueil/Enseignants/Fenetre_pedagogique/PEPS/Relation-maitre-eleve.pdf)>.
- Goh, S. C. et Fraser, B. J. (1996). Validation of an elementary school version of the Questionnaire on Teacher Interaction. *Psychological Reports*, 79(2), 515-522.
- Greenacre, M. (2006). From Simple to Multiple Correspondence Analysis. In M. Greenacre, J. Blasius. *Communication présentée Multiple correspondence analysis and related methods*, New York : Chapman and Hall/CRC, Taylor and Francis Group.
- Hair, J. F. (2006). *Multivariate data analysis* (6e éd.). Pearson international edition: Upper Saddle River NJ.
- Harman, H. H. (1976). *Modern factor analysis* (3rd. ed.). Chicago: The University of Chicago Press.
- Henderson, D., Fisher, D. et Fraser, B. (2000). Interpersonal behavior, laboratory learning environments, and student outcomes in senior biology classes. *Journal of Research in Science Teaching*, 37(1), 26-43.
- Henson, R. K., Capraro, R. M. et Capraro, M. M. (2001). Reporting Practice and Use of Exploratory Factor Analysis in Educational Research Journals. Communication présentée dans: Annual Meeting of the Mid-South Educational Research Association, Little Rock, AZ.
- Henson, R. K. et 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.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement*, 20, 141-151.
- Lapointe, J. et Legault, F. (1999). Version francophone du Questionnaire on Teacher Interaction en contexte québécois. *Mesure et évaluation en éducation*, 22(2-3), 1-19.
- Leary, T. F. (1957). *Interpersonal Diagnosis of Personality: A Functional Theory and Methodology for Personality Evaluation*. New York: Ronald Press Company. <https://books.google.ca/books?id=G2SvI6bVQj4C> (consulté le 29/02/15).
- Nunnally, J. (1967). *Psychometric theory*. New York: McGraw-Hill.
- Nunnally, J. (1978). *Psychometric theory*, 2d ed. New York: McGraw-Hill.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior research methods, instruments, et computers*, 32(3), 396-402.
- Osborne, J. W. et Costello, A. B. (2009). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pan-Pacific Management Review*, 12(2), 131-146.
- Park, H. S., Dailey, R. et Lemus, D. (2002). The use of exploratory factor analysis and principal components analysis in communication research. *Human Communication Research*, 28(4), 562-577.
- Passini, S., Molinari, L. et Speltini, G. (2015). A validation of the Questionnaire on Teacher Interaction in Italian secondary school students : the effect of positive relations on motivation and academic achievement. *Social Psychology of Education*, 18(3), 547-559.

- Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of consumer research*, 21(2), 381-391.
- Pett, M. A., Lackey, N. R. et Sullivan, J. J. (2003). *Making sense of factor analysis : The use of factor analysis for instrument development in health care research*. London: New-Delhi : Sage.
- Russell, D. W. (2002). In search of underlying dimensions : The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and social psychology bulletin*, 28(12), 1629-1646.
- Sivan, A., Chan, D. W. et Kwan, Y. W. (2014). Psychometric Evaluation of the Chinese Version of the Questionnaire on Teacher Interaction (C-QTI) in Hong Kong. *Psychological reports*, 114(3), 823-842.
- Stafford, J. et Bodson, P. (2006). *L'Analyse Multivariée Avec SPSS*. Sainte-Foy : Presses de l'Université du Québec.
- Steiger, J. H. (1990). Some additional thoughts on components, factors, and factor indeterminacy. *Multivariate Behavioral Research*, 25(1), 41-45.
- Tabachnick, B. G., Fidell, L. S. et Osterlind, S. J. (2001). *Using multivariate statistics* (5th ed.). Boston ; Montreal. <http://tocs.ulb.tu-darmstadt.de/135813948.pdf> (Consulté le 13/08/2016).
- Telli, S., den Brok, P. et Cakiroglu, J. (2007). Students' perceptions of science teachers' interpersonal behaviour in secondary schools : Development of a Turkish version of the Questionnaire on Teacher Interaction. *Learning Environments Research*, 10(2), 115-129.
- Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques : Implications pour la recherche en langue française. *Canadian Psychology/Psychologie Canadienne*, 30(4), 662.
- Watzlawick, P., Bavelas, J. B. et Jackson, D. D. A. (1967). *Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes*. New York : Norton and Company.
- Wei, M., Den Brok, P. et Zhou, Y. (2009). Teacher interpersonal behaviour and student achievement in English as a Foreign Language classrooms in China. *Learning Environments Research*, 12(3), 157-174.
- Wubbels, T., Creton, H. A. et Hooymayers, H. P. (1985). Discipline Problems of Beginning Teachers, Interactional Teacher Behaviour Mapped Out. Communication présentée lors de : The Annual Meeting of the American Educational Research. Retrieved 31 mars - 4 avril.
- Wubbels, T. et Levy, J. (1993). *Do you know what you look like ? Interpersonal relationships in education*. London; Washington, D.C. : Psychology Press.
- 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. <http://www.tqmp.org/RegularArticles/vol09-2/p079/p079.pdf> (consulté le 1/07/2017).

## EXTENDED ABSTRACT

Psychometric questionnaires are among the most commonly tools in social sciences, especially in education (Barats, 2016). The advent of different statistical software packages which allow the analysis of data collected through these tools has contributed to the use of questionnaires in several fields (e. g. psychology, social sciences, education, etc.) throughout the world (Bourque *et al.*, 2006). In order to be able to use a questionnaire in a given context, a researcher should ensure that it can offer satisfactory psychometric qualities, particularly in terms of validity and reliability. Experts in methodology and statistics differentiated between many types of validity (criterion validity, predictive validity, construct validity) and reliability (internal consistency, temporal consistency, etc.) (Field, 2013). Construct validity holds a key role in the validation procedures of questionnaires in social sciences. It assesses their ability to measure what they claim to measure, namely the construct (Field, 2009).

The exploratory factor analysis (EFA) is considered one of the most used extraction methods to validate construct of questionnaires in humanities and social sciences (Osborne & Costello, 2009). However, most procedures for verifying construct validity of questionnaires translated in French are tainted by statistical and methodological weaknesses (Ben Alaya, Grenon, Desbiens & Bali, 2018). More specifically, several studies based on exploratory factor analyses (EFA) are marked by controversial interpretations of this type of analyses (Bourque, Poulin & Cleaver, 2006). This may call into question the credibility of some questionnaires, and of results arising from them.

In this paper, the *Questionnaire on Teacher Interaction* (QTI) was chosen because of its great worldwide diffusion. Indeed, since 1985, it has been used in more than 1300 scientific publications and translated in more than 20 languages. Also, it has been the focus of validation procedures in more than 40 countries (den Brok *et al.*, 2003a), over different school levels (from primary school to university) and in various school subjects (e. g. mathematics, science, English, etc.) (Sivan & Chan, 2013). These validation procedures explain the existence of several different versions of the QTI, ranging from the Dutch original version including 77 items (Wubbels, 1985) to a short Canadian version with only 32 items (Lapointe & Legault, 1999).

These factor validation procedures have been the subject of a widespread criticism about many versions of the QTI (Ben Alaya *et al.*, 2018). This paper follows on the heels of this criticism, but it focuses only on the short version developed by Lapointe & Legault (1999). The latter was used by other Francophone researchers including Dumas (2009), as well as by Soussi and Nidegger (2014), etc.

The present study is a secondary analysis of existing data collected by Dumas in 2009 in Physical and Health Education. These data were obtained between November 2004 and December 2005 from a nonprobability sample comprised of 2261 Secondary 1 to 5 students (60.5% of boys and 39.5% of girls). A detailed description can be found under the address: (<http://savoirs.usherbrooke.ca/handle/11143/868>). Analyses were performed on the anonymized data encoded in the database. SPSS software (version 19.0) was used to doing statistical analyses.

We compared the results of two distinct EFA approaches: the first arose from previous studies which validated the QTI; the second systematically followed the steps and recommendations of EFA specialists (e. g. Fergusson & Cox, 1993, Field, 2013; Osborne & Costello, 2009). This close comparison allowed us to show just how decisions taken by analysts may influence the quality of the factor solution found. It also casts doubt on the robustness of the abridged version of the QTI as well as on the robustness of the previously adopted validation procedures. Finally, it leads us to reflect on the credibility of all

other versions that were used in research and in many scientific studies throughout the world. Among other things, with this study, we were able to cast a critical eye on the stringency of results published for the last 30 years, using the QTI to describe student-teacher relationship.

Last, we showed that the QTI, a questionnaire widely distributed and used in Educational Sciences, can present insufficient psychometric qualities according to the characteristics of the methodological approach taken to carry out the verifications. Thus, the adopted procedures to run the same checks on any other questionnaire could potentially present the same problems. The production of valid and reliable knowledge about studied facts is a responsibility shared by all members of research communities, especially by researchers and scientific committees of journals. We invite those people to tighten use and publication of EFA approach standards through all means available in order to fill the identified gaps and improve the quality of scientific productions. These precautions will allow for the production of knowledge in which different categories of users can have faith since it will reflect, with minimal bias, what it is supposed to reflect.