



Le processus d'adaptation : Une démarche scientifique pour traduire le test psychométrique

Mémoire doctoral

François René de Cotret

**Doctorat en psychologie
Docteur en psychologie (D. Psy.)**

Québec, Canada

Le processus d'adaptation

Une démarche scientifique pour traduire le test psychométrique

Mémoire doctoral

François René de Cotret

Sous la direction de :

Yvan Leanza, directeur de recherche

Résumé

Le présent mémoire doctoral porte sur la transposition du test psychométrique dans une nouvelle langue/culture, ce que j'ai décidé de nommer, en référence aux plus récentes avancées sur le sujet, le *processus d'adaptation*. La traduction à proprement parler représente une fraction du processus, lequel vise la démonstration de l'équivalence entre le test original et le test adapté. Les matériaux de base dudit processus sont les stratégies d'adaptation (ou simplement stratégies), qui servent à encadrer celui-ci, mais, surtout, à lui conférer sa légitimité scientifique.

Dans le premier article, il est question des stratégies d'adaptation qui font consensus dans la littérature scientifique. Les 17 stratégies d'adaptation consensuelles ont été ciblées par le biais d'une analyse de contenu, puis réparties en cinq phases : la Traduction initiale, la Révision interne, la Révision externe, l'Évaluation quantitative (Testing) et la Supra phase, qui chapeaute les autres phases du processus d'adaptation.

Le deuxième article porte sur les différentes manières de rapporter, dans une publication scientifique, comment un processus d'adaptation a été effectué. La *Grille d'évaluation de la description du processus d'adaptation* (G-PA) a tout d'abord été élaborée par le biais des 17 stratégies d'adaptation consensuelles. La G-PA a permis de catégoriser les descriptions tirées d'une recension de la littérature, soit 131 descriptions de processus d'adaptation effectuées au Québec de 2000 à 2014. La longueur et la complexité des descriptions ainsi que le nombre de stratégie(s) qu'elles comportent varient énormément. De l'analyse du corpus découle trois types de description – Libre, Positiviste et Intégrative – selon ce qui est mis de l'avant pour légitimer le processus d'adaptation.

À la lumière des résultats obtenus, il est possible de conceptualiser le processus d'adaptation en tant que démarche scientifique à part entière. Or, une telle démarche ne saurait servir uniquement à mesurer l'équivalence entre la version originale et la version traduite du test. Les résultats illustrent que la démarche vise, dans un premier temps, à négocier le sens de l'équivalence pour ensuite en extraire une mesure. Lors des trois premières phases du processus d'adaptation, différents experts réfléchissent sur l'équivalence des versions et en coconstruisent le sens. Ce n'est qu'ensuite, lors du Testing, qu'il est possible de s'avancer à propos d'une mesure ou, plutôt, de la mesure de l'équivalence. Par la force des choses, la recherche de l'équivalence en deux temps implique des approches scientifiques distinctes. La légitimité du processus d'adaptation repose sur ces approches, qui doivent être exposées et différencierées du moment que la connaissance scientifique est visée, ce qui est fait dans la section Conclusion.

Table des matières

Résumé.....	III
Table des matières	IV
Liste des tableaux	VI
Liste des figures	VII
Épigraphie	VIII
Remerciements.....	IX
Avant-propos.....	X
Introduction	1
1 – Problématique	3
1.1 Qu'est-ce que la traduction?	3
1.2 Comment traduire?	6
1.3 Qui traduit?	11
1.4 Le test psychométrique comme type de texte	14
1.5 Le processus d'adaptation.....	19
2 – Article I (PA-TH).....	29
Résumé	29
Abstract.....	30
Method.....	36
Literature review.....	36
Content analysis.....	36
Results	38
Phase 1: Initial Translation.....	42
Phase 2: Internal Review.....	45
Phase 3: External Review.....	46
Phase 4: Testing.....	47
Supra Phase	48
Discussion.....	49
References	53

3 – Article II (PA-AP).....	58
Résumé	58
Abstract.....	59
Method.....	62
Results	63
AP-G criteria.....	69
Phase 1 – Initial Translation.....	71
Phase 2 – Internal Review.....	71
Phase 3 – External Review.....	71
Phase 4 – Testing.....	71
Supra Phase – Documentation.....	72
Numeric variables.....	72
Neglected strategies.....	74
Limitations.....	75
Discussion.....	76
Three types of descriptions to legitimize PA-APs.....	77
<i>Free descriptions.</i>	77
<i>Positivist descriptions.</i>	78
<i>Integrative descriptions.</i>	79
References	82
Conclusion	93
Étudier la traduction	95
La traduction littérale et la traduction inversée	98
L’importance de la théorie.....	99
Réviser le cursus universitaire.....	100
Références.....	105
Annexe A : AP-G Development Steps.....	119
Annexe B : Typical examples of PA-AP description	129
Annexe C : Grille d’évaluation de la description du Processus d’Adaptation	132
Annexe D: Adaptation Process description assessment Grid	134

Liste des tableaux

Tableau 1. Comparaison du positivisme et du constructivisme sur trois plans.....26

Article I

Table 1. List of texts in the corpus, with year of publication, authors (and organisations, where applicable), text type, and citations count.....40

Table 2. The 17 adaptation strategies (S) which reached consensus (%), regrouped into five phases.....41

Article II

Table 1. Journal articles by alphabetical order of authors, and journals in which they were published.....64

Table 2. Academic publications (PhD and MA theses and/or dissertations) by alphabetical order of author, and academic affiliation.....66

Table 3. Percentage (in decimal number) of descriptions per criteria met, by group of descriptions.....70

Table 4. Values of the three numeric variables, by group of descriptions.....72

Table 5. Fractions and percentages (in decimal numbers) of journal articles with direct mention of PA-APs, by group of descriptions.....74

Liste des figures

Figure 1. Schématisation des deux postulats unissant le test psychométrique au construit .	19
Figure 2. Schématisation du <i>decentering</i>	20
Figure 3. Schématisation de la traduction inversée en tant que technique de mesure.....	20
Figure 4. Schéma des cinq phases du processus d'adaptation.....	93

Article I

Figure 1. Knowledge transfer of the Adaptation Process.....	32
Figure 2. Example of the development of a consensual adaptation strategy.....	38
Figure 3. Search strategy for the literature review.....	39

Article II

Figure 1. Two ways of addressing the adaptation process.....	60
Figure 2. Number of descriptions according to the number of AP-G criteria they meet.....	69

Épigraphe

*Some cross-cultural researchers have even suggested
that a high percentage of the research in their fields
is flawed to the point of being invalid
because of poorly adapted tests.*
Hambleton, Merenda & Spielberger, 2005, p. 4

Remerciements

Je remercie les deux organismes subventionnaires qui ont investi dans mon projet. Le Conseil de la vie française en Amérique m'a accordé une bourse de recherche lors de ma deuxième année doctorale. Trois ans plus tard, le Centre interuniversitaire d'études sur les lettres, les arts et les traditions m'a accordé deux bourses, une d'incitation à la publication ainsi qu'une bourse d'excellence de troisième cycle.

Je remercie les membres de mon laboratoire de recherche, Psychologie et Cultures, pour leur soutien et nos échanges. Une mention spéciale à Jean-Philippe pour l'éclair de génie qu'il a partagée lors d'une de mes présentations!

Je remercie mon directeur de recherche, Yvan Leanza, pour son travail régulier ainsi que son dévouement. Il aura été fort agréable de discuter de mon mémoire avec lui, mais aussi de questions périphériques, de la politique à la recherche interculturelle en passant par l'épistémologie. Il a par ailleurs été à mes côtés lors de moments plus difficiles, ce qui m'a aidé à maintenir le cap. Idem pour la membre de mon comité de mémoire doctoral, Aline Francoeur, que je remercie. Elle a investi temps et énergie dans mon projet, en particulier pour le volet traduction/ traductologie ainsi que pour la traduction des deux articles à proprement parler, faite en collaboration avec John O'Meara. Par ailleurs, c'est principalement grâce à Aline que j'ai pu aller présenter mon projet en Espagne en 2014 lors de la neuvième Conférence de l'International Test Commission.

Je remercie mon ami et confrère doctoral Olivier pour nos discussions sur le processus d'adaptation, qui ont entre autres conduit à la typologie telle que présentée dans le deuxième article. Je remercie mon ami Louis, qui a passé plusieurs heures à lire mon mémoire et à en discuter avec moi.

Je remercie finalement ma famille, ma conjointe et les autres personnes qui me sont chères. Au fil des années, elles m'ont appris à me connaître.

Avant-propos

Les deux articles du mémoire doctoral, que j'ai rédigés et dont je suis l'auteur principal, ont été traduits par Aline Francoeur et révisés par John O'Meara, Ph.D. Mon directeur de recherche, Yvan Leanza, et la membre de mon comité de mémoire doctoral, Aline Francoeur, m'ont par ailleurs conseillé et accompagné lors de la rédaction.

Introduction

Des professeurs et doctorants en psychologie que je côtoie, tout le monde sans exception peut me parler de la traduction de tests psychométriques. Plusieurs d'entre eux ont déjà participé ou même dirigé une telle activité, qui est monnaie courante au Québec, seule province canadienne unilingue francophone. Quand ils en parlent, plusieurs précisent qu'il faut maintenant dire « adaptation » au lieu de « traduction ». Leur discours n'est d'ailleurs pas orienté sur la traduction à proprement parler, mais sur les moyens mis en place pour l'encadrer. Ce discours gravite autour d'expressions telles que « traduction inversée », « approche par comité », « prétest » et « validation ».

Mon cursus universitaire en psychologie, à l'image de celui de mes collègues, contraste avec cette réalité, dans laquelle nous sommes impliqués de près ou de loin dans la traduction/adaptation de plusieurs tests psychométriques. De fait, aucun cours ni bloc de cours offert n'a comme objectif de réfléchir à la traduction/adaptation de tests. Si la plupart de mes collègues peuvent me nommer un texte sur le sujet, à savoir Vallerand (1989), il est rare qu'ils en connaissent d'autres. Ce sont plutôt les expériences personnelles qui nourrissent la discussion, qui est – il ne faut pas se méprendre – peu fréquente et généralement de courte durée.

Au fil des années, je me suis demandé comment il se faisait que le discours entourant la traduction/adaptation de tests prenne si peu de place en psychologie alors que tant de résultats de recherche proviennent de tests adaptés. En consultant la littérature scientifique, je me suis rendu compte qu'aucun terme ne faisait consensus pour circonscrire cette pratique. Mes travaux sur le sujet me permettent aujourd'hui de parler du *processus d'adaptation*, choix que j'explique dans quelques pages. Ce qu'il faut savoir pour l'instant, c'est que le processus d'adaptation consiste en une traduction sur laquelle s'agglutine une série de mécanismes visant à transformer en un savoir scientifique la réponse à la question suivante : est-ce que le texte traduit est équivalent au texte original? La légitimité du test adapté, mais également du processus d'adaptation sous-jacent, passe par la démonstration de cette équivalence.

Le présent mémoire doctoral offre l'analyse de ce que divers chercheurs proposent pour arriver à une telle démonstration. L'objectif général était d'offrir une organisation conceptuelle de la pratique en question, aussi bien au niveau de la théorie (Article I) que de la pratique en contexte québécois (Article II).

Le premier article porte sur le discours qu'entretiennent les scientifiques par rapport au processus d'adaptation. Il s'agissait de circonscrire de leurs propos une définition prototypique du processus d'adaptation ainsi que des composantes qui lui confèrent sa légitimité scientifique. On passe ensuite, au deuxième article, de la théorie à la pratique, en explorant les différentes manières par lesquelles des chercheurs ayant effectué un processus d'adaptation décrivent leur démarche dans leur publication scientifique. L'objectif était de déterminer les éléments les plus souvent mis de l'avant pour légitimer la démarche entreprise.

La section Problématique vise d'une part à familiariser le lecteur avec les rudiments de la traduction vue en tant que pratique mais aussi en tant qu'objet d'étude. D'autre part, elle situe la pertinence de réfléchir à propos des mécanismes « objectivant » du processus d'adaptation. Nourrie par les deux articles du présent mémoire, cette réflexion ouvre, dans la section Conclusion, sur la définition du processus d'adaptation et sur les considérations épistémologiques que celle-ci soulève.

1 – Problématique

Le processus d’adaptation noue une relation aussi étroite que particulière avec la traduction, ce pour quoi il est impératif de se pencher sur celle-ci pour considérer celui-là. D’abord, je définis sur trois plans ce que représente l’acte de traduire : d’un point de vue théorique (Qu’est-ce que la traduction?), d’un point de vue technique (Comment traduire?) et d’un point de vue personnel (Qui traduit?). La sous-section « Qui traduit? » couvre par ailleurs la compréhension, par le traducteur, de ce qui doit être traduit, en ce qui nous concerne, le test psychométrique, introduit à la sous-section suivante. Je poursuis ensuite avec le processus d’adaptation.

1.1 Qu’est-ce que la traduction?

Bien que la pratique de la traduction existe depuis fort longtemps, ce n'est que depuis le milieu du XX^e siècle qu'elle est traitée comme une discipline universitaire. À partir de cette époque, des théoriciens ont étudié le phénomène de la traduction pour en faire ressortir des principes et des théories, ce qui constitue l'objet d'étude de la traductologie. La traductologie peut ainsi être considérée comme une praxéologie, c'est-à-dire une discipline qui prend pour objet une pratique.

L’objet d’étude de la traductologie. Pour toute praxéologie, la définition de l’objet d’étude est fonction de la vision de chaque théoricien. Pym (2009) examine la traduction par le biais d’une série de paradigmes (selon le sens général donné par Kuhn, 2012[1962]) non unifiés ni unificateurs qui supposent des définitions distinctes de la traduction. Chacun de ces paradigmes gravite autour d’une notion fondamentale : l’équivalence naturelle, l’équivalence dynamique, la fonction, la description, l’incertitude, la localisation et la traduction culturelle. J’avais développé les sept principes dans une version antérieure du mémoire doctoral sans que cela n’ajoute une compréhension accrue du processus d’adaptation. Par souci de concision, seulement les trois premières notions, largement sollicitées lors du processus d’adaptation, seront abordées.

Du latin *aequus* (égal) et *valeo* (valeur), le terme *équivalence* signifie dans le domaine de la traduction l’égalité des valeurs entre le texte source et le texte cible : « The term roughly assumes that, on some level, a source text and a translation can share the same value, and

that this assumed sameness is what distinguishes translations from all other kinds of texts » (Pym, 2009, p. 6). Il est question d'une notion fort simple, qui, pourtant, se complexifie du moment qu'on l'explore, avec son lot de définitions et de théories.

Selon Pym (2009), l'équivalence constitue les fondations de la traductologie. Il la divise en deux : l'équivalence naturelle et l'équivalence directionnelle. L'équivalence naturelle sous-entend qu'un texte en langue A peut être traduit en langue B puis retraduit en langue A sans que des différences soient observées entre les deux versions en langue A. Selon cette conception, le traducteur s'évertue à trouver la *bonne* traduction : « Natural equivalence is presumed to exist between languages or cultures prior to the act of translating » (p. 7).

L'équivalence directionnelle sous-entend quant à elle une relation asymétrique entre les langues, « where the creation of an equivalent by translating one way does not imply that the same equivalence will also be created when translating the other way » (p. 27). L'adjectif directionnel réfère à l'unilatéralité de l'équivalence, du texte source vers le texte cible. Une telle traduction résulte d'une série de décisions prises par le traducteur; tout compte fait, deux traducteurs n'ont aucune chance de produire la même traduction pour un texte donné.

Reconnu pour avoir défini la traduction par l'équivalence, Nida (1964) divise également cette notion générale en deux. Il explique que le choix d'une équivalence dépend de l'effet recherché chez le lecteur cible. La traduction formelle, qui vise l'équivalence formelle, met l'accent sur le message, aussi bien dans la forme que dans le contenu : « In such a translation one is concerned with such correspondences as poetry to poetry, sentence to sentence, and concept to concept. » (p. 159). L'équivalence formelle sous-entend une traduction littérale pour laquelle la syntaxe demeure pratiquement inchangée.

En comparaison, l'équivalence dynamique vise à ce que la relation (dynamique) entre le message et lecteur cible soit équivalente à la relation entre le message et le lecteur source. En favorisant le sens du texte (sémantique), l'équivalence dynamique vise essentiellement l'équivalence de la réponse du lecteur, pour qui la lecture, adaptée à sa culture, devrait être plus naturelle, fluide : « A translation of dynamic equivalence aims at complete naturalness of expression, and tries to relate the receptor to modes of behavior relevant within the

context of his own culture; it does not insist that he understand the cultural patterns of the source-language context in order to comprehend the message » (p. 159). L'équivalence dynamique demeure un idéal en cela que la réponse ne sera jamais identique, « for the cultural and historical settings are too different » (1969, p. 24).

Matthews (2003) explique que les travaux de Nida, démodés selon plusieurs traductologues contemporains, « ont ouvert la voie à de nouvelles formes de traduction; des formes qui s'éloignent de la traduction mot à mot » (p. 61). Une de ces formes, popularisée après Nida, correspond aux théories fonctionnalistes, aujourd'hui la forme la plus populaire chez les théoriciens (Gutt, 2000). Les théories fonctionnalistes stipulent que la finalité (*function*) du texte est une propriété qui importe autant sinon plus que la recherche d'équivalence.

Pym (2009) regroupe les théories fonctionnalistes sous l'appellation du paradigme de la finalité, popularisé dans les années 1980. Une théorie phare de ce paradigme est celle du skopos, un terme grec signifiant justement « finalité ». Selon cette théorie, l'évaluation de la traduction passe notamment par la compréhension des besoins du récepteur de la traduction, qui peut être le lecteur cible ou le demandeur de la traduction, par exemple. En définitive, selon le paradigme de la finalité, le contexte influence la traduction du texte source vers le texte cible, qui ne sont plus considérés en vase clos.

Un changement de paradigme n'implique pas nécessairement que la notion sous-jacente à l'un soit condamnée par l'autre. Il serait plutôt question d'une progression qui ouvre la porte à plus de possibilités. Ainsi, sans la rejeter, la théorie du skopos minimise l'importance de l'équivalence. En effet, cette théorie favorise plutôt la finalité du texte cible; celui-ci n'est plus uniquement déterminé par la finalité du texte source, mais également par sa propre finalité, qui peut être distincte : « The translator is thus placed in a social situation where they [sic] have to negotiate with all the parties involved, including the client, the reader or end-user, and the author » (p. 61). On parlerait ainsi d'une négociation qui s'opère entre le traducteur et son client afin de convenir de la manière de traduire un texte selon la finalité envisagée.

L'orientation de la traduction. Les notions qui viennent d'être présentés influent différemment sur l'importance accordée aux textes source et cible. Pour comprendre ce

phénomène, il importe d'introduire une autre notion, l'orientation, propre au descriptiviste Toury.

Toury (2012) mentionne que si les années 1970 étaient dominées par des traductions orientées vers la langue source, par la suite, la langue cible a été de plus en plus considérée : « Most translation scholars, while not abandoning the seemingly safe haven of the source text, have at least come to integrate more and more target-bound considerations into their reasoning » (p. 18). Il distingue ce faisant la traduction orientée vers le texte source de celle orientée vers le texte cible. La première se plie aux normes de la langue et de la culture source et la deuxième, à celles de la langue et de la culture cible. Nous y reviendrons.

1.2 Comment traduire?

Les théories de la traduction, dont je viens de faire un bref survol, n'offrent pas au traducteur de moyens concrets pour faire son travail. Pour y arriver, il peut toutefois avoir recours à des stratégies et à des procédés de traduction.

Stratégies et procédés de traduction. Comme l'explique Delisle (2003, p. 60), « la stratégie de traduction oriente la démarche globale du traducteur à l'égard d'un texte particulier à traduire et se distingue des décisions ponctuelles comme l'application des divers procédés de traduction ». Par exemple, un traducteur peut adopter comme stratégie de favoriser la traduction littérale (équivalence formelle) pour un texte en entier afin que le lecteur puisse être mis en contact avec la culture source (voir Nida, 1964). En comparaison, les procédés portent sur des segments de texte et visent à respecter les particularités idiomatiques de la langue cible.

Considéré comme un ouvrage pionnier dans les rangs des traducteurs, l'ouvrage de Vinay et Darbelnet (1977[1954]) présente sept procédés de traduction, hiérarchisés selon la difficulté d'utilisation, du plus facile au plus difficile. Il s'agit de l'emprunt, du calque, de la traduction littérale, de la transposition, de la modulation, de l'équivalence et de l'adaptation.

Présentées dans les prochains paragraphes, les définitions desdits procédés sont tirées du glossaire de Delisle (2003, pp. 25-69), à l'exception de la *traduction littérale*, qui provient directement de l'ouvrage de Vinay et Darbelnet (1977[1954], pp. 48-49). Selon Delisle, la traduction littérale représente une stratégie de traduction, car elle vise l'ensemble du texte et non une unité de traduction plus restreinte. Il précise par ailleurs que la *traduction mot à mot* est un genre de traduction littérale qui consiste « à transposer dans le texte [cible] les éléments du texte [source] sans en changer l'ordre » (p. 65). Vinay et Darbelnet ne font pas cette distinction; ils considèrent ces deux termes comme des synonymes, choix auquel je me rallie. Je nuance l'emploi en précisant s'il est question d'une stratégie ou d'un procédé au besoin.

I - L'emprunt. Procédé de traduction qui consiste à conserver dans le texte [cible] un mot ou une expression appartenant à la langue [source], soit parce que la langue [cible] ne dispose pas d'une correspondance lexicalisée, soit pour des raisons d'ordre stylistique ou rhétorique. Note : Un emprunt s'écrit généralement en romain ou en italique selon qu'il est intégré ou non dans la langue emprunteuse.

Ex. 1 : Un pub.

Ex. 2 : Des lunettes *new look*.

Ex. 3 : Un curriculum vitae.

Ex. 4 : *Ad hoc*.

Ex. 5 : Troïka.

II - Le calque. Procédé de traduction qui consiste à transposer dans le texte [cible] un mot ou une expression du texte [source] dont on traduit littéralement le ou les éléments.

Ex. 1 : *Skyscraper* = gratte-ciel.

Ex. 2 : To double-click = double-cliquer.

III - La traduction littérale (*mot à mot*). Désigne le passage de la langue [source] à la langue [cible] aboutissant à un texte à la fois correct et idiomatique sans que le traducteur ait eu à se soucier d'autre chose que des servitudes linguistiques.

Ex. : *I left my spectacles on the table downstairs* = J'ai laissé mes lunettes sur la table en bas.

Mais si, une fois ce procédé atteint, la traduction littérale est reconnue inacceptable par le traducteur, il faut recourir à [un autre procédé]. Par inacceptable, nous entendons que le message, tel qu'il se laisse rédiger littéralement, a) donne un autre sens; b) n'a pas de sens; c) est impossible pour des raisons structurales; d) ne correspond à rien dans la métalinguistique de la langue d'arrivée; ou e) correspond bien à quelque chose, mais non pas au même niveau de langue.

IV - La transposition (recatégorisation). Procédé de traduction qui consiste à établir une équivalence par un changement de catégorie grammaticale. Note : Le terme transposition, applicable à un trop grand nombre de procédés de traduction, n'a pas été retenu dans ce sens.

Ex. 1 : *Three ways to make life easier.* = Trois façons de se simplifier la vie.

Ex. 2 : *Keep refrigerated.* = Garder au réfrigérateur.

V - La modulation. Procédé de traduction qui consiste à restructurer un énoncé du texte [source] en faisant intervenir un changement de point de vue ou d'éclairage par rapport à la formulation originale, ce qui se produit, notamment, quand on emploie la partie pour le tout, l'abstrait pour le concret, l'actif pour le passif.

Ex. : *The party leadership convention is over but a far harder test faces the new leader: the challenge of running the country.* = La course à la direction du parti est terminée, mais une épreuve beaucoup plus difficile attend le nouveau chef : diriger le pays.

VI - L'équivalence. Procédé de traduction qui consiste à rendre une expression figée de la langue [source] par une expression figée qui, bien que renvoyant à une représentation différente dans la langue [cible], exprime la même idée.

Ex. 1 : *One bitten, twice shy* = chat échaudé craint l'eau froide.

Ex. 2 : *You could have knocked me down with a feather* = Les bras m'en sont tombés.

Ex. 3 : *She let the cat out of the bag* = Elle a vendu la mèche.

VII - L'adaptation. Procédé de traduction qui consiste à remplacer une réalité socioculturelle de la langue [source] par une réalité propre à la socioculture de la langue [cible] convenant au public cible du texte d'arrivée.

Niveaux de difficulté et d'interprétation. Delisle (2003) explique que certains segments textuels sont plus faciles à traduire que d'autres et propose trois *niveaux d'interprétation* des mots en contexte : le report, la remémoration et la création discursive.

Le *report*, le niveau le plus facile, touche les segments « qui échappent presque complètement à l'analyse du sens » (p. 166). Il s'agit principalement des noms propres, des nombres, des dates, des termes techniques ou scientifiques monosémiques, des titres de revues, de certains symboles (%., S, etc.), des codes postaux, des formules mathématiques, etc.

La *remémoration* vise les segments qui sont généralement consignés dans les dictionnaires : « Une fois l'acception isolée, le traducteur réactive sa mémoire et peut compter trouver en langue d'arrivée un mot ou une expression utilisé spontanément par les usagers de cette langue pour désigner la même réalité dans la même situation de communication » (p. 167).

Le niveau le plus difficile, la *création discursive*, nécessite « de trouver une formulation *non donnée d'avance* qui recouvre le sens du passage original » (p. 168). Voici un des exemples donnés par Delisle (p. 169) : *When passions flow, the first casualty is perspective* = Le déchaînement des passions obscurcit le jugement.

Traduction directe ou oblique. Finalement, l'usage d'un procédé de traduction ou d'un niveau d'interprétation exige de choisir entre deux directions possibles, soit la traduction directe ou la traduction oblique.

La traduction directe sous-entend une symétrie entre les deux langues : « [I]l peut arriver que le message en langue [source] se laisse parfaitement transposer dans le message en langue [cible], parce qu'il repose soit sur des catégories parallèles (parallélisme structural), soit sur des conceptions parallèles (parallélisme métalinguistique) » (Vinay & Darbelnet, 1977[1954], p. 46). Ce type de traduction est mis à profit lors d'un report ou d'une

remémoration. Les quatre premiers procédés de traduction présentés plus haut (emprunt, calque, traduction littérale et recatégorisation) constituent des traductions directes.

On a recours à la traduction oblique lorsque la transposition ne peut se faire directement : « [I]l se peut aussi que le traducteur constate dans la langue [cible] des trous ou “lacunes”, qu'il faudra combler par des moyens équivalents, l'impression globale devant être la même pour les deux messages » (p. 46). C'est le cas lors d'une création discursive. Les trois derniers procédés de traduction présentés plus haut (modulation, équivalence et adaptation) constituent des formes de traductions obliques.

La subjectivité de la traduction. Tous les auteurs abordés précédemment s'accordent pour dire que le travail du traducteur ne pourrait être réduit à l'utilisation de divers procédés appliqués systématiquement. Introduite par Delisle, la création discursive ouvre d'ailleurs la porte à l'importance... de la créativité.

La créativité, ou encore la part de subjectivité propre à l'acte de traduire, est décrite par Wilss (1990), qui souligne, ce faisant, l'impossibilité de trouver *la* bonne traduction dans certains cas. Son propos conclut la présente sous-section tout en introduisant la prochaine, portant sur le traducteur, et, en particulier, sur son rôle :

Translators often find themselves in a situation that forces them to progressively filter out, through comprehensive information-processing, norm-deviant variants from the possibilities available in their more or less open horizon of action. They then must select from the acceptable variants the one that most closely corresponds to their conception of an adequate target text. The decision *for* a variant signifies, of course, also a decision *against* all other possible variants, some of which might, under certain circumstances, be just as acceptable. (p. 29)

1.3 Qui traduit?

Comme Bolaños-Medina et González-Ruiz (2012) l'écrivent, « translation is not just a transcoding process, but a form of human action » (p. 734). Contrairement à l'ordinateur, le traducteur possède une connaissance du monde qui précise le sens des segments de traduction. Delisle (2003) parle de *compléments cognitifs*, soit « des connaissances extralinguistiques mobilisées au moment de la recherche de l'équivalence qui sont indispensables à la constitution du sens » (p. 185).

L'exemple suivant, fourni par Bédard (référence secondaire de Delisle, 2003, p. 186), renseigne sur la pertinence des compléments cognitifs. La traduction-machine de *Time flies like an arrow* peut donner *Le temps vole comme une flèche*, *Les mouches du temps aiment une flèche* ou *Chronométrez les mouches comme une flèche*. Or, contrairement au traducteur, l'ordinateur n'est pas (encore) capable de trancher; les trois possibilités demeurent syntaxiquement plausibles étant donné que *time*, *flies* et *like* peuvent être traités comme des verbes. Dans ce cas-ci, les compléments cognitifs permettent au traducteur de rejeter sans trop de difficulté les deux possibilités loufoques. Ceci dit, les exemples avec un niveau de complexité supérieur abondent dans les livres pédagogiques de traduction; être humain et parler deux langues ne suffit pas pour traduire convenablement.

Les compléments cognitifs se multiplient au cours de la vie, riche d'expérience de toute sorte. À propos de la traduction en particulier, il y a au moins trois aspects qui favorisent cette richesse et accroissent la qualité de ce que le traducteur produit : la formation universitaire, l'expérience en traduction et la compréhension du texte à traduire.

Formation universitaire. Vreck (1998) mise sur la formation des traducteurs pour que ceux-ci prennent d'abord conscience « qu'il est rare qu'un mot, quelle que soit la langue, ait un sens unique aux contours bien définis », pour ensuite « [les] persuader, d'une part, qu'un mot ne s'analyse jamais seul et, d'autre part, que la grammaire n'est pas une science abstraite destinée à empoisonner [leurs] plus belles années, mais une quantité limitée et répertoriée d'outils porteurs et faiseurs de sens » (p. 50). Il s'agit, pour le traducteur en herbe, de se familiariser avec la logique sous-jacente à la langue à l'aide d'exercices pratiques et de remises en question.

L'auteure avance que le procédé de traduction littérale est fréquent chez les traducteurs qui n'ont pas encore appris à prendre une distance par rapport à la structure de surface de la langue : « [O]n peut espérer que l'étudiant qui aura appris à se détacher des mots qu'il a sous les yeux et à se préoccuper des signifiés sous-jacents aura le réflexe, dès que la traduction littérale d'une structure de surface ne lui offrira pas une solution satisfaisante, de se tourner vers cet au-delà du texte pour en faire son point de départ » (p. 56).

En cette matière, Delisle (2003) spécifie que la formation universitaire permet à l'étudiant d'apprendre à dissocier les langues, qui ont des structures non parallèles, en les comparant, mais aussi en étant confronté à des difficultés récurrentes de traduction. « Apprendre à traduire, écrit-il, c'est apprendre à aborder un texte de manière “raisonnée”, c'est découvrir progressivement tout ce qu'exige le transfert du sens d'un texte d'une langue dans une autre [...] » (p. 17).

À ces propos s'ajoutent ceux de Darbelnet (1966), qui constate que la traduction littérale est employée à outrance, parfois même dans les rangs des traducteurs professionnels. Le défi est d'autant plus grand que le traducteur doit souvent travailler rapidement pour satisfaire les exigences de son employeur, et la traduction littérale est rapide : « L'ignorance, ou plutôt une imparfaite compréhension des processus de la traduction, et la hâte sont les deux ennemis de la qualité dans ce domaine. Cela il faut l'enseigner, textes en mains, et en proposant aux élèves des exercices courts, précis et gradués » (p. 158).

Une traduction littérale peut être correcte, sans pour autant être idiomatique, ce que Darbelnet (p. 157) souligne avec quelques exemples éloquents :

Ainsi on parle des « octrois » fédéraux sans doute parce que *to grant* peut vouloir dire *octroyer*, et qu'on est passé du verbe au substantif. On a ainsi généralement une traduction qui ne vaut que dans certains cas et on l'a fait aux dépens de *subventions*. Le même processus est reconnaissable à propos de « maîtriser » qui semble rendre parfaitement *to master*. Mais la similitude de forme n'entraîne pas forcément l'identité de sens ou de valeur. Il est exact que *to hit*

signifie *frapper*, mais nous sommes *heurtés* et non pas « frappés » par une autre voiture.

Le traducteur doit également maîtriser les langues source et cible, en particulier la langue cible, qui constitue généralement sa langue maternelle. Darbelnet (p. 159) met toutefois en garde contre la croyance voulant qu'on excelle dans notre langue maternelle :

L'illusion vient de ce qu'on la parle couramment et sans effort. Dès qu'on est obligé de rédiger, on s'aperçoit qu'on hésite sur beaucoup de points et qu'il est souvent difficile d'exprimer toute sa pensée.
[...] L'expérience montre qu'il faut observer les faits de langue et y réfléchir si l'on veut vraiment avoir bien en main cet outil indispensable qu'est la langue maternelle, langue de pensée et non pas seulement de communication.

Il propose comme solution d'intégrer à l'enseignement de la traduction une formation parallèle à propos de la grammaire, de la lexicologie, etc., afin de parfaire la connaissance du système linguistique qu'est la langue maternelle.

Expérience en traduction. La formation académique va de pair avec une autre caractéristique du traducteur, l'expérience en traduction. Selon le sens commun, la qualité d'une traduction varie proportionnellement à l'expérience du traducteur. Ceci dit, les études en traductologie appliquée permettent d'en apprendre davantage sur cette relation. Bolaños-Medina et González-Ruiz (2012) se réfèrent à quelques-unes d'entre elles pour dresser une série de constats.

Premièrement, les traducteurs expérimentés ne visent pas *la* bonne traduction, mais la production d'un texte qui est adéquat et suffisant pour une situation communicationnelle particulière. De plus, ils auraient de la facilité à jongler avec les différents principes, stratégies et procédés de traduction.

Les auteurs mentionnent également que plus les traducteurs gagnent en expérience, plus leurs prises de décision deviennent des automatismes, ce qui leur permet de diriger leur attention sur des aspects ignorés préalablement. Ainsi, « more controversial aspects of a

particular commission are identified and cost-effectiveness is improved since time and efforts are devoted to the resolution of key problems » (p. 732).

L'efficacité des traducteurs d'expérience viendrait aussi du fait qu'ils (a) analysent des segments de traduction plus longs que ceux qu'analysent les traducteurs ayant moins d'expérience; (b) favorisent principalement le sens des segments, soit l'idée sous-jacente; (c) prennent en considération la stylistique, i.e. l'étude des moyens qu'emploie la langue pour exprimer la pensée ainsi que « des procédés d'écriture caractéristiques d'un domaine professionnel donné » (définition de Valentine & Aubin, 2004, p. 2); et (d) ont à leur disposition un plus grand inventaire de possibles quand ils traduisent.

Compréhension du texte. Il est fréquent d'entendre que le traducteur se base sur le sens des segments de traduction. Cette prise de position, déplore Ballard (1998), « amène à ne pas parler de la lecture en tant que telle comme phase de l'opération de traduction » (p. 27). La lecture du texte conduit à une interprétation qui aboutit à la construction de sens, « qui sera le sens [que le traducteur] attribue au texte et qui pourra différer plus ou moins de l'intention de l'auteur et de l'interprétation d'autres lecteurs ».

Comprendre un texte, affirme Vreck (1998), exige de la personne qui traduit des connaissances « qui doivent s'ajouter aux significations de la langue originale » (p. 65). Ces connaissances extralinguistiques, ou compléments cognitifs, pour reprendre la formulation de Delisle, touchent entre autres le domaine de spécialité auquel le texte se rapporte. De la sorte, la connaissance d'un ensemble de concepts à la base du test psychométrique est déterminante pour arriver à traduire ce type de texte.

1.4 Le test psychométrique comme type de texte

Bolaños-Medina et González-Ruiz (2012) considèrent le test psychométrique comme un type de texte propre (*textual genre*), soit « a class of communication event that takes place in a given communicative situation with a particular purpose and which presents a characteristic pattern of textual conventions in terms of schematic structure, style, content and intended audience » (p. 718). Comme pour tout autre type de texte, l'efficacité de la communication y est essentielle.

Il comporte une série de questions auxquelles le lecteur, appelé participant, est invité à répondre selon certaines modalités, dépendant du type de test. Le test autorapporté (*self-administered*) est le type de test le plus fréquent, car il est économique, complété par le participant seul, sans l'aide d'un tiers, tel un expérimentateur.

Le test psychométrique se divise en plusieurs éléments textuels distincts. Il y a tout d'abord le questionnaire, qui comporte deux parties : les consignes et les items. Sauf exception, les consignes sont formulées de manière concise et claire dans un langage simple :

Instructions are given in a straightforward manner, reflecting the main characteristics of common technical discourse, such as the use of imperative, present tenses and descriptive verbs, in an attempt at an impersonal and objective style. Syntax is plain, frequently comprising several simple independent sentences. Frequent use of general vocabulary, and a tendency to exclude technique terms and acronyms which could create ambiguity, are also common. Examples of how to complete the forms correctly are usually given. (Bolaños-Medina & González-Ruiz, 2012, p. 720)

Les items sont constitués d'une question et d'une modalité de réponse telle que vrai/faux ou des choix multiples de type Likert. Comme les consignes, les items sont rédigés dans un style concis et un registre relativement informel ainsi qu'avec des expressions précises afin que les participants, pour qui la maîtrise de la langue varie, comprennent facilement et sans ambiguïté de quoi il est question. Les réponses sont inscrites sur une feuille-réponse, sur laquelle le participant est généralement invité à fournir certaines informations biographiques telles que son âge, son sexe ou encore son niveau de bilinguisme.

Finalement, le test peut venir avec un manuel servant entre autres à informer l'expérimentateur sur la marche à suivre pour effectuer la passation du test ainsi que sur les qualités psychométriques du test en question et la manière d'interpréter les réponses données par les sujets.

Que mesure le test psychométrique? L'American Education Research Association, l'American Psychological Association et le National Council on Measurement in Education définissent le test psychométrique comme « un instrument ou un processus d'évaluation dans lequel un échantillon de comportements d'une personne, dans un domaine précis, est obtenu et, par la suite, évalué et corrigé en utilisant un processus standardisé » (2003[1999], p. 3). Cette définition sous-entend l'idée de mesure d'un échantillon de comportements, ce qui est clairement indiqué dans la définition d'un autre ouvrage de référence : « le test psychométrique est essentiellement une mesure objective [...] d'un échantillon de comportements... » (Anastasi, 1994 [1982], p. 26).

Or, les comportements mesurés ne sont pas une finalité en soi. Ceux-ci sont censés livrer suffisamment d'information à propos « du concept ou de la caractéristique que le test est censé mesuré » (American Educational Research Association *et al.*, 2003[1999], p. 6). Par souci de simplicité, je vais employer le terme *construit* pour me référer aussi bien au concept qu'à la caractéristique que le test est censé mesuré. Explicité par Cronbach et Meehl (1955), ce terme est réservé au monde de la recherche, contrairement au terme *concept*, ce pour quoi je favorise celui-là à celui-ci.

Dans l'ouvrage de l'American Psychological Association *et al.*, la définition du test psychométrique et l'explication à propos de ce à quoi réfèrent les comportements mesurés par le test sont abordées dans des sous-sections distinctes. En juxtaposant les segments correspondants, on obtient une définition à la fois simple et décisive : *le test psychométrique, censé mesurer un construit, mesure des comportements*. En effet, le test à lui seul ne pourrait servir à mesurer un construit, immatériel. Pour arriver à lier ce que le test mesure et ce qu'il est censé mesurer, il faut adhérer à deux postulats.

Postulat concept-comportements. Le premier postulat se formulerait ainsi : admettre comme vrai que le construit peut être mesuré par l'entremise de comportements. Ce principe, qui s'est popularisé en psychologie avec la réflexion de Tolman dans les années 1930, perdure dans le temps bien qu'il ne soit plus abordé comme tel dans la majorité des cursus du domaine (Green, 1992).

Tolman avançait qu'il était nécessaire de recourir à l'opérationnalisme pour arriver à mesurer un construit. Par exemple, pour mesurer l'anxiété, un phénomène immatériel, il faut l'opérationnaliser, c'est-à-dire le définir selon des opérations qui permettent sa mesure. L'anxiété pourrait être définie comme l'élévation du taux de cortisol dans le sang et du rythme cardiaque.

On pourrait aussi la définir comme un score élevé à un test psychométrique comme le State Trait Anxiety Inventory, qui comporte des items tels que *I feel secure*, *I have disturbing thoughts* et *I tire quickly*. Bien que ces deux définitions opérationnelles de l'anxiété soient distinctes, on considère qu'elles réfèrent (à peu près) au même construit. La polysémie du construit rend (à peu près) légitime l'existence des méta-analyses en psychologie : parce qu'ils mesurent (à peu près) le même construit, les différents tests psychométriques recensés sont (carrément) traités sur le même pied d'égalité. La polysémie sert également lors de la validation d'un test. Il est fréquent de comparer les résultats d'un test aux résultats d'autres tests psychométriques censés mesurer (à peu près) le même construit. Dans ce cas, on parle de validité convergente.

L'opérationnalisme provient d'une époque à laquelle le bémorisme gagnait ses lettres de noblesse. Gardner (1993[1985]) résume ce courant en mentionnant qu'il importait pour un bémoriste d'« employer des méthodes d'observations, utilisables et quantifiables par tout scientifique » et de « travailler uniquement sur des comportements » (p. 24). En rendant le construit observable par l'entremise de comportements, l'opérationnalisme rend légitime l'utilisation du test psychométrique, du moins, si on considère, comme l'American Educational Research Association *et al.* (2003[1999]), que le test mesure des comportements.

Or, avec la majorité des tests psychométriques, on ne mesure pas directement des comportements. Le test psychométrique autorapporté, par exemple, consiste à recueillir les réponses que le sujet donne à une série de questions. Ce sont ces questions et ces réponses, des items, qui réfèreraient aux « comportements mesurés ».

Postulat comportements-items. Le deuxième postulat se formulerait ainsi : admettre comme vrai que les réponses au test psychométrique réfèrent à des comportements.

Postuler un tel lien entre les items et les comportements minimise pourtant l'influence que peut avoir le participant dans l'équation.

Pour comprendre le rôle que le participant peut avoir lors de la lecture de l'item, mettons en opposition deux items fictifs :

1. Je vais sur Facebook a) 0 à 2 heures par semaine, b) 2 à 4 heures par semaine, c) 4 à 6 heures par semaine, et d) plus de 6 heures par semaine.
2. Je vais sur Facebook a) 0 à 4 heures par semaine, b) 4 à 8 heures par semaine, c) 8 à 12 heures par semaine, et d) plus de 12 heures par semaine.

Pour un comportement donné chez le participant, ici le temps passé sur Facebook par semaine, la réponse obtenue à ces items peut pourtant varier. Dans cet exemple, les choix de réponse influenceront probablement le participant, qui aura tendance à se positionner par rapport au nombre d'heures passées sur Facebook qu'il considère normal. Le premier item suggère que passer plus de six heures par semaine sur Facebook n'est pas normal – car il s'agit d'une réponse à l'extrême des choix possibles –, contrairement au deuxième item, qui suggère le contraire. Et même si les choix de réponses n'avaient pas d'influence sur le participant, sa réponse devra être une approximation du temps réel moyen passé sur Facebook par semaine, à moins qu'il comptabilise ce genre de données ou qu'il n'utilise pas du tout Facebook, ce qui est peu probable.

Ce qu'il faut retenir de cet exemple est que le participant a un rôle actif dans l'interprétation de l'item, et qu'il est impossible d'ignorer cet intermédiaire lorsqu'on tente de lier les items aux comportements.

En résumé, il faut adhérer à au moins deux postulats pour arriver à considérer le test psychométrique comme un outil de mesure scientifique. Chacun de ces deux postulats est illustrés par une flèche sur la Figure 1; ils permettent de lier le construit au test par le biais des comportements. Les deux postulats vont de pair avec le paradigme positiviste, présenté dans la prochaine sous-section.

Figure 1. Schématisation des deux postulats unissant le test psychométrique au construit



1.5 Le processus d'adaptation

Je situe la naissance du processus d'adaptation en 1970 avec la publication de l'article de Brislin intitulé *Back-Translation for Cross-Cultural Research*. Déterminer un temps zéro est nécessairement réducteur, mais fort utile à la fois. La Révolution française n'a pas réellement commencé la journée de la prise de la Bastille ni même en 1789; il s'agit d'un symbole. Idem pour l'article de Brislin, qui traite de la possibilité d'organiser la traduction à l'aide de la démarche scientifique.

La stratégie popularisée par Brislin se nomme *decentering*¹ et consiste en une série de traductions inversées² (*back translations* ou *back-translations*). Au moins deux traducteurs sont impliqués dans une traduction inversée. Le premier produit un texte cible en partant du texte source initial. Le deuxième, qui n'a pas accès au texte source, traduit le texte cible vers la langue source. C'est de ce retour du texte vers la langue source qu'on tire le terme traduction inversée.

Une fois que le deuxième traducteur a terminé, il est possible de traduire ce texte à nouveau vers la langue cible, et ainsi de suite (voir Figure 2). Cela permettrait en théorie de se départir des expressions et des tournures de phrase qui se transposent mal d'une langue à l'autre, afin que seuls les concepts disponibles dans les deux langues subsistent. Brislin qualifie ces concepts d'étiques (*etic concepts*), en opposition aux concepts émiques, propres à chaque culture. Les deux versions produites sont par conséquent supposées mesurer le même construit : « “Decentering” means that the research project is not centered around

¹ Rarement traduit en français, le terme *decentering* pourrait être traduit par *décentrement*, bien que je n'ai pas rencontré cette traduction dans mes lectures. J'ai donc décidé de m'en tenir au terme anglais.

² Le terme *back-translation* est traduit de différentes manières à l'écrit : *traduction inversée*, *traduction renversée*, *traduction à l'envers*, *traduction-retraduction*, *traduction à rebours*, *contre-traduction*, *traduction rétroaction*, *rétro traduction*, *rétro-traduction*, *rétrotraduction*, *retraduction*, *méthode en double-aveugle*, *double traduction*, etc. Je n'ai pas trouvé une formulation en français qui puisse rendre l'idée avec suffisamment de précision, ce pourquoi je m'en tiens au choix le plus populaire et, à mon sens, intuitif : *traduction inversée*.

any one culture or language. Instead, the idiosyncrasies of each language under study contribute to the final version of the questionnaire » (Brislin, 1976, p. 222).

Figure 2. Schématisation du *decentering*

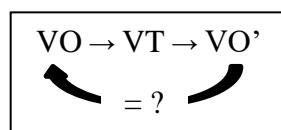
Texte initial (A) → Traduction en langue cible (B) → Traduction en langue source (A')
→ Traduction en langue cible (B') → Traduction en langue source (A'')

L'essor de la traduction inversée. Il est prévu que le *decentering* soit mis à contribution lorsqu'un test psychométrique est élaboré dans deux langues en même temps. Selon Brislin, cette technique ne pourrait ni ne devrait être utilisée du moment que l'objectif du chercheur est d'adapter un test déjà existant : « The result of decentering contrasts with the awkward, stilted versions common when material in one language is taken as the final content that must be translated with minimal change into another language » (p. 222).

Faisant l'éloge de la traduction inversée, plusieurs chercheurs ont pourtant proposé des processus d'adaptation visant à adapter un test déjà existant (p. ex. Bradley, 1994; Butcher, 1996; Guillemin, Bombardier & Beaton, 1993; Prieto, 1992; Vallerand, 1989). Dans ces processus d'adaptation, on ne traite pas de *decentering*, mais spécifiquement de traduction inversée. L'objectif n'est pas d'élaborer conjointement un test en deux langues/cultures, mais plutôt d'assurer que le test traduit (version traduite, VT) est équivalent au test original (version originale, VO).

De technique de traduction, la traduction inversée proposée par Brislin est devenue une technique de mesure. On pouvait dès lors considérer l'équivalence des versions à la lumière de la similitude entre le test original et le test original retraduit (voir Figure 3).

Figure 3. Schématisation de la traduction inversée en tant que technique de mesure



Publié en 1994, l'article de Geisinger intitulé *Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments* marque un tournant dans le domaine de l'adaptation de tests, tel que mentionné par l'auteur : « Back translation was once the technique of choice in test translation » (p. 306). À partir de cette époque, plusieurs auteurs ont souligné les dérives possibles de la traduction inversée, dont certaines avaient déjà été relevées par Brislin d'ailleurs (p. ex., Gierl, Rogers & Klinger, 1999; Hambleton & Kanjee, 1995; Harkness & Schoua-Glusberg, 1998; Tassé & Craig, 1999; Van Widenfelt, Treffers, De Beurs, Siebelink & Koudijs, 2005).

Entre autres, Gierl *et al.* (1999) estiment qu'il est hasardeux de postuler l'équivalence entre la VT et la VO après avoir comparé la VO et la VO' : « [T]here is no assurance that the findings in the source language generalize to the target language because the source-to-target language translation is not directly assessed » (p. 8). Un tel raccourci sous-entend une symétrie entre les langues, qui, bien qu'elle soit envisageable dans une certaine mesure, comme l'expliquent Vinay et Darbelnet (1977[1954]) (voir la section Introduction), ne pourrait l'être pour un texte en entier.

La traduction inversée employée en tant qu'instrument de mesure illustre le besoin pour les chercheurs de posséder des moyens pour démontrer l'équivalence des versions. Dans les années 1990, d'autres moyens ont été popularisés, ce qui sera abordé dans l'article I.

Cet essor concorde avec le glissement du terme traduction vers le terme adaptation. Avant les années 1990, les chercheurs du domaine de la psychologie ou de domaines connexes référaient à la traduction pour parler de la transposition d'un test psychométrique d'une langue et/ou d'une culture à l'autre. Par la suite, certains auteurs ont insisté sur la distinction entre *traduction* et *adaptation*, dont Geisinger (1994) : « In recent years, the term *test adaptation* frequently has replaced test translation. This shift in terminology documents the adaptations in references to culture, in content, and in wording that are needed in addition to simple translation in revising a test » (p. 305).

Une distinction est ainsi faite entre la « simple » traduction du test et les phases subséquentes visant à assurer la qualité de la traduction par le biais de la révision et de la

validation. Intitulé *Adapting educational and psychological tests for cross-cultural assessment*, le livre phare de Hambleton, Merenda et Spielberger (2005) offre une définition claire de ce qu'est l'adaptation de tests psychométriques, en mettant justement l'accent sur les « *activities* » autres que la traduction, terme absent de la définition d'ailleurs :

Test adaptation includes all the activities from deciding whether or not a test could measure the same construct in a different language and culture, to selecting translators, to deciding on appropriate accommodations to be made in preparing a test for use in a second language, to adapting the test and checking its equivalence in the adapted form. (p. 4)

Comme cette définition le suggère, le terme *adaptation* en est venu à référer à l'adaptation du test en général. Or, il peut également référer au procédé de traduction éponyme présenté précédemment à la sous-section « Comment traduire? », comme Geisinger (2003) le précise : « When translators find functionally equivalent behaviors to use to replace those that do not generalize across cultures, they are adapting, rather than translating [...] » (p. 99). Le terme *adaptation* peut également référer à un troisième sens, c'est-à-dire au résultat de l'opération d'adaptation; par exemple, on peut parler d'une adaptation pour désigner un test qui a été adapté.

Par souci de clarté, je propose le terme *processus d'adaptation de tests psychométriques* (ou simplement *processus d'adaptation*) pour référer à l'ensemble des « *activities* » effectuées pour arriver à transposer un test psychométrique d'une langue et/ou d'une culture à l'autre.

Les « *activities* » nommées par Hambleton *et al.* (2005) remplissent trois rôles. Elles visent a) à déterminer si le construit existe dans la culture cible, b) à produire une traduction le cas échéant et c) à assurer l'équivalence entre cette traduction et le test original. À mon sens, l'adaptation d'un test débute quand la décision de l'adapter a été prise, ce pour quoi je ne traiterai pas du point a). De surcroît, déterminer l'existence d'un construit dans une culture donnée nécessite le recours à un test psychométrique, ce qui implique tout d'abord de

procéder aux points b) et c). Je ne développerai pas davantage sur cette circularité, qui remet en doute la pertinence même du test psychométrique; le problème est documenté depuis des décennies (voir Green, 1992).

En définitive, le processus d'adaptation réfère au passage de la version originale d'un test psychométrique (VO) à une version adaptée (VA) en passant par une ou plusieurs versions traduites (VT). La VT devient VA lorsque le processus d'adaptation est terminé. Celui-ci se conclut lorsque l'équivalence entre la VO et la VT est démontrée. En ce sens, le processus d'adaptation vise la production d'un savoir scientifique et doit être qualifié de démarche scientifique. Cette démarche est pourtant aussi délicate que son objet d'étude, l'équivalence des versions. Dans la littérature scientifique, l'équivalence des versions ne réfère pas à l'équivalence du texte source et du texte cible telle que traitée dans la section Introduction. Nous reviendrons sur cette polarisation théorique dans quelques paragraphes.

Définir l'équivalence des versions. En 1970, Brislin soulignait le défi que représente le processus d'adaptation, avançant que celui-ci serait simple « if there were some criteria for translation quality or translation equivalence against which a given study could be gauged » (p. 190). Quelque 40 ans plus tard, Angel (2013) arrivait à la même conclusion : « Despite an increase in the attention devoted to instrument development and translation into multiple languages, formal criteria for assessing the adequacy and quality of translated protocols are lacking » (p. 224).

Nombreux sont les chercheurs qui ont proposé des types d'équivalence afin d'offrir de tels critères. Méthot (1998) rapporte que les dizaines d'auteurs recensées abordent l'équivalence conceptuelle ou une variante de celle-ci, « supportant ainsi l'idée que cette forme d'équivalence est capitale » (p. 19) et « [doit être] évaluée en tout premier lieu » (p. 22).

Herdman, Fox-Rushby et Badia (1997) constatent également la prépondérance de ce type d'équivalence à la suite d'une revue de littérature. Dans les 66 études multinationales qu'ils ont répertoriées, on mentionne 19 types d'équivalence, les plus populaires étant l'équivalence conceptuelle (30 % du temps), l'équivalence sémantique (12 %), l'équivalence fonctionnelle (8 %), l'équivalence métrique (8 %), l'équivalence des échelles

(8 %), l'équivalence technique (6 %) et l'équivalence opérationnelle (6 %). Or, un autre de leur constat touche à la polysémie de ces termes, en particulier à propos de l'équivalence conceptuelle :

Our literature review revealed that there is a distinct lack of clarity and a considerable amount of confusion surrounding the way in which various types of equivalence are defined [...]. This is particularly true of conceptual equivalence. (p. 243)

Ils posent l'hypothèse que la polysémie relative aux types d'équivalence découle de différentes conceptions théoriques (*theoretical frameworks*) auxquelles adhèrent les chercheurs. Elles se divisent en trois selon eux :

- la conception absolutiste, selon laquelle on peut transposer les concepts d'une culture à l'autre avec la certitude que celle-ci n'a qu'une influence minimale sur ceux-là – le test psychométrique est ainsi imposé à l'autre culture;
- la conception universaliste, selon laquelle on peut également transposer les concepts d'une culture à l'autre, mais en gardant à l'esprit l'influence de la culture, entre autres en adaptant les tests psychométriques minutieusement;
- et la conception relativiste, selon laquelle tout concept représente une construction sociale dépendante d'un contexte et d'une culture donnée – le test ne peut être adapté selon cette conception.

Ces différentes conceptions théoriques viendraient teinter la manière de définir l'équivalence, et devraient être considérées afin qu'une systématisation des types d'équivalence puisse avoir lieu. Selon mes recherches, personne n'a effectué ce travail ni une normalisation ou une uniformisation terminologique des divers types d'équivalence. Il est par ailleurs possible que ce soit un défi théorique insurmontable, comme le souligne Angel (2013, p. 228) :

Translation may always and in principle remain an art. It is possible and even likely that the nature of language itself places limits on the extent to which complete equivalence can be achieved through even

the most rigorous techniques. As in the world of sub-atomic physics we may have to accept the reality that a degree of uncertainty is inevitable in all human communication and that there may exist limits to how scientific translation can become.

Il existe pourtant une définition opérationnelle de l'équivalence sur laquelle les chercheurs semblent : « It has become customary to treat equivalence from a measurement level perspective » (van de Vijver & Tanzer, 2004, p. 121). On peut ainsi définir l'équivalence par la similarité des réponses entre les items de la VT et ceux de la VO (Hambleton & Patsula, 1999, p. 8). Selon cette définition implicite, il y a équivalence lorsque les styles de réponses obtenues lors de la passation de la VT correspondent aux styles de réponses obtenues lors de la passation de la VO. Dans cette perspective, il est possible de faire passer la VT à un échantillon de sujets jugé représentatif de la population cible, et ensuite de comparer leurs réponses, à l'aide d'analyses statistiques, à celles obtenues par un échantillon de sujets de la population source ayant complété la VO. Il est aussi possible de faire passer la VT et la VO à un échantillon de sujets jugés bilingues et d'évaluer la symétrie de leurs réponses, toujours à l'aide d'analyses statistiques.

Ce faisant, on « observe » l'équivalence par le biais d'un instrument de mesure, le test psychométrique. Il est pourtant possible de l'« observer » sans instrument. Afin de comprendre cette distinction et ses ramifications, je vais recourir à deux paradigmes, soit le positivisme et le constructivisme.

Positivisme et constructivisme. Au risque d'être (trop) réducteur, je définis le positivisme et le constructivisme en les mettant en opposition sur trois plans. J'ai tiré cette idée, que je résume dans le Tableau 1, de Nguyễn-Duy et Luckerhoff (2006, pp. 7-11).

Tableau 1*Comparaison du positivisme et du constructivisme sur trois plans*

	Positivisme	Constructivisme
Ontologie	Universaliste <i>Une seule réalité existe</i>	Relativiste* <i>Plusieurs réalités coexistent</i>
Épistémologie	Objectiviste <i>Le chercheur observe l'objet</i>	Subjectiviste* <i>Le sens émerge de la relation entre l'observateur et le sujet</i>
Méthode d'analyse des données	Expérimentale <i>Instrumentale</i> <i>Opérationnelle</i> <i>Hypothético-déductive</i>	Naturaliste* <i>Description détaillée des sujets et des phénomènes considérés dans leur contexte naturel</i> <i>Principalement inductive</i>

*Terminologie tirée de Nguyen-Duy et Luckerhoff (2006, pp. 7-11)

Le positivisme postule la possibilité d'une connaissance dite positive du monde au sens d'une connaissance vraie et objective. Selon celui-ci, il est possible d'observer l'objet d'étude (épistémologie objectiviste) par le biais de comportements qu'il faut mesurer (méthode d'analyse des données expérimentale). Lors d'un processus d'adaptation, l'équivalence des versions se définirait donc par des items qui se « comportent » (*function*) de manière semblable (Hambleton & Patsula, 1999, p. 8). Cela implique d'utiliser deux versions d'un même test, et de comparer les résultats obtenus à ceux-ci.

Procéder de la sorte exige de considérer une seule réalité (ontologie universaliste), ce qui est sous-entendu par l'utilisation du test psychométrique. Tel qu'expliqué précédemment, le test psychométrique est considéré comme un instrument qui permet de mesurer (observer) un construit appartenant à une seule réalité dans laquelle baigne l'ensemble des individus de la population cible. Ceci dit, les deux versions du test représentent deux langues/cultures distinctes auxquelles sont rattachées différentes manières de concevoir le monde qui nous entoure. Comme le souligne Angel (2013, p. 228), « the need for translation brings us face to face with some basic epistemological issues and the potential limits of objectivity when we are dealing with culturally constructed subjective reality ».

Nous l'avons vu à la sous-section « Qui traduit? », le traducteur doit comprendre le texte à traduire. Il doit certes le comprendre pour lui-même, mais aussi être capable de se mettre à la place du lecteur source : comment la VO est-elle comprise par le lecteur source? Il doit ensuite arriver à *faire comprendre* le texte source au lecteur cible, ce qui n'est pas une mince tâche, comme l'explique Skibińska (2007) :

Non seulement [le traducteur] interprète, ou pèse, le contenu pragmatico-sémantique du texte original, mais il doit aussi l'évaluer du point de vue de la réception du texte traduit : percevoir les écarts culturels entre les récepteurs de l'original et ceux de la traduction, dresser un inventaire des difficultés qui en résultent et adopter des solutions conformes aux possibilités perceptives et aux attentes des lecteurs virtuels de celle-ci. Dans cette « négociation », il doit peser et estimer la valeur des éléments de l'œuvre originale pour voir lesquels doivent être rendus impérativement dans la traduction, lesquels peuvent être modifiés et lesquels peuvent disparaître sans que cela nuise à l'intégrité de « l'édifice » qu'il est en train de construire. (p. 7)

Cette gymnastique mentale implique que le traducteur soit capable de se faire « une image, consciente ou inconsciente, de son lecteur à lui : il traduit à l'intention de ce lecteur, il se met à sa place, il crée et recrée ce lecteur au fil de son travail » (Hewson, 1995, cité dans Skibińska, 2007, p. 6). Il doit aussi s'être fait, au préalable, une image du lecteur source. Le traducteur entre dès lors en relation, non pas avec ces deux lecteurs, mais avec l'image qu'il se fait d'eux. C'est en se mettant à la place de l'un et ensuite de l'autre qu'il arrive à estimer, comme l'avance Nida (1964) et le rappelle Skibińska (2007), l'équivalence. Finalement, les six principes qui ont suivi l'équivalence naturelle (Pym, 2009) témoignent du besoin, pour le traducteur, de considérer autre chose que l'égalité des valeurs au sens strict, telle la fonction du texte cible, qui peut différer de celle du texte source, par exemple si le demandeur (celui qui a initié le processus d'adaptation) veut ajouter ou modifier des items au test pour une raison particulière.

Ainsi donc, traduire un test psychométrique impliquerait de tenir compte de différentes réalités (ontologie relativiste), ce qui ouvre la porte à une épistémologie subjectiviste et à une méthode d'analyse des données naturaliste. En tant que démarche scientifique, le processus d'adaptation posséderait conséquemment un objet d'étude polymorphe. L'équivalence serait, d'une part, relative en cela qu'elle dépendrait de la manière de comprendre et de faire comprendre le texte et, d'autre part, absolue car mesurée selon la méthode expérimentale.

Les deux articles qui suivent abordent ce problème, qui s'inscrit dans un contexte plus large touchant à la légitimité du processus d'adaptation et du test adapté.

Le premier article circonscrit un portrait prototypique du processus d'adaptation, ce qui permet de schématiser celui-ci afin de se l'approprier et donc de savoir plus clairement de quoi il est question. En particulier, l'étude porte sur le discours qu'entretiennent les chercheurs à propos du processus d'adaptation dans la littérature scientifique en général et, en particulier, sur ce qui permet de légitimer le processus d'adaptation et par voie de conséquence le test adapté.

Le deuxième article offre une typologie qui permet d'organiser les différentes manières de décrire dans une publication scientifique le processus d'adaptation effectué. Il s'agit de répondre à une question semblable à celle posée au premier article, mais cette fois-ci en regardant dans la pratique plutôt que dans la littérature : qu'est-ce que les chercheurs mettent de l'avant pour légitimer le processus d'adaptation qu'ils ont effectué?

2 – Article I

Les stratégies consensuelles pour traduire et adapter les tests psychométriques : Comprendre et schématiser le Processus d'Adaptation

François RENÉ DE COTRET¹

Yvan LEANZA, Ph.D.¹

Aline FRANCOEUR, Ph.D.²

Andrée-Anne BEAUDOIN-JULIEN¹

¹École de psychologie

²Département des langues, de linguistique et de traduction

Université Laval

Québec, Canada

Résumé :

La traduction ou l'adaptation d'un test psychométrique pour une population dont la langue ou la culture diffère de celle de la population pour laquelle il a été élaboré à l'origine est un processus qui comporte plusieurs phases et stratégies. Les experts soulignent qu'il n'y a pas de consensus scientifique sur la meilleure façon de mener ce processus, ce que l'article dément. Une recension de la littérature a été effectuée afin de répertorier un nombre significatif de publications décrivant le processus d'adaptation et ces publications ont été soumises à une analyse de contenu afin de déterminer quelles stratégies étaient consensuelles. Au total, 34 publications ont été répertoriées et de nombreux points de consensus ont été trouvés dans les stratégies recommandées. Le processus d'adaptation a été divisé en cinq phases et en 17 stratégies d'adaptation consensuelles. Une telle schématisation souligne les consensus existant et facilite le transfert de connaissances. Elle porte également à réfléchir sur l'influence prédominante du constructivisme sur le processus d'adaptation du moment que la connaissance scientifique est visée.

Mots-clés : traduction, testing, pretesting, equivalence culturelle, ITC

Consensual strategies for translating and adapting psychometric tests: Understanding and schematizing the Adaptation Process

François RENÉ DE COTRET¹

Yvan LEANZA, Ph.D.¹

Aline FRANCOEUR, Ph.D.²

Andrée-Anne BEAUDOIN-JULIEN¹

¹School of Psychology

²Department of Languages, Linguistic, and Translation

Laval University

Québec, Canada

Abstract:

Translating or adapting a psychometric test for a population whose language/culture differs from that of the population for which it was originally developed is a process that involves multiple phases and strategies. Experts have pointed out that there is no scientific consensus on the best way to conduct this process. **Objectives:** This article aims to demonstrate that there are indeed strategies that are agreed upon when it comes to test adaptation. **Method:** A literature review was conducted to identify a representative number of publications describing the adaptation process. These publications were then submitted to content analysis in order to determine which strategies were consensual. **Results:** A total of 34 publications were reviewed, and many points of consensus were found in the strategies recommended. The adaptation process was separated into five phases and 17 adaptation strategies. **Discussion:** The determination of consensus among experts with regard to the adaptation process answers the need for providing researchers with the scientific bases that are required to establish which strategies can be used and to facilitate knowledge transfer. It also offers a schematization of the adaptation process that calls attention to the predominant influence of constructivism on this particular process.

Keywords: Translation, testing, pretesting, guidelines, cultural equivalence, ITC

Adapting a psychometric test³, i.e. transposing it into a language/culture different from the one it was originally developed for, involves multiple steps. In the passage from the original version (OV) to the adapted version (AV), a series of translated versions (TV) are produced, followed by various stages of evaluation and review. Hence the term *adaptation process* to describe all the *strategies* (or adaptation strategies) used to make this operation possible.

When it comes to adapting a test, the very choice of strategies is rooted in a legacy which is mainly theoretical. Since the early 1970s, significant works proposing strategies for conducting, overseeing and improving the adaptation process have been published. Since these publications are based on scientific knowledge, we have decided to call them “PA-THs” (for “processes of adaptation - theoretical”).

There are also other publications that offer a concrete description of a specific adaptation process, i.e. what was actually done to adapt a specific test. We decided to call these applied processes “PA-APs” (for “processes of adaptation - applied”). Consequently, PA-THs contain a set of recommended strategies for adapting tests, while PA-APs refer to actual applications of strategies that may or may not correspond to those recommended in PA-THs.

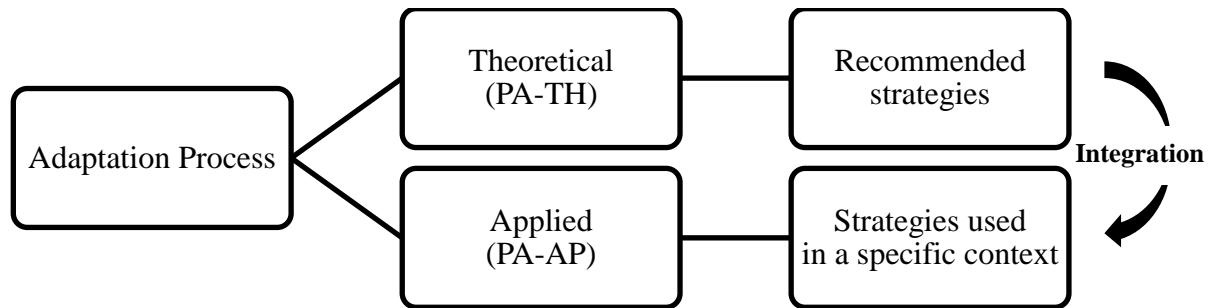
In other words, a PA-TH is a reflection supported by scientific knowledge that results in ready-to-use adaptation strategies, while a PA-AP represents a method which results in a specific adapted test.

This study looks at the knowledge transfer and, in particular, the organization and synthesis of knowledge on the adaptation process. From this standpoint, PA-THs and their recommended adaptation strategies constitute organized knowledge that can be integrated by users carrying out PA-APs. Knowledge transfer from PA-THs to PA-APs implies the ability to choose from among numerous recommended strategies the ones that are

³ The term *test* is used in the paper in a broad sense to include any psychometric instrument, as defined by the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). *Construct* is used as an umbrella term for everything being measured: constructs, aptitudes, etc.

appropriate in the context of adapting a specific test. Figure 1 provides a schematic view of this explanation.

Figure 1. Knowledge transfer of the Adaptation Process



The knowledge transfer can be conceptualized and defined in different ways. The simplest models suggest that knowledge travels linearly from producers to users, while others consider the interaction between the two actors involved in the process within the social context in which it occurs (Fournier 2011). Every model, however, is based on the idea that knowledge must be organized in such a way as to permit users to integrate it (Ward, House & Hamer 2009).

As part of a broader social dynamic, the integration of knowledge is unpredictable: it can be problematic, requiring “significant unlearning as part of the re-ordering of knowing” (Davies, Nutley & Walter 2008 : 189). The challenge is even greater in social sciences, which produce forms of knowledge that are not strictly defined: “What counts as social research moves beyond just those data produced by academic researchers to include, for example, local research and evaluation reports and the findings from audit” (p. 189).

Brislin’s PA-TH and the problematic integration of back-translation. Brislin (1970) is one of the first works—and certainly the most cited of the period—that recommends a series of adaptation strategies based on the scientific literature as well as on the results of a field study. Although Brislin’s reflection is based on cross-cultural studies, it ties in with other domains such as translation studies (e.g., Nida 1964), statistics (e.g., Winer 1962) and sociology (e.g., Schuman 1966).

At the end of his 1970 article, which is a full-fledged PA-TH, Brislin recommends seven clearly identified adaptation strategies, all of which are formulated and qualified in the 20-page article. As he mentions (1970 : 214), his PA-TH is based on “an ideal situation” because it is theoretical; it is not applied to a specific context like PA-APs: “[M]ost research projects would have to change the procedure in response to a given problem.”

Popularized by Brislin (1970), the adaptation strategy commonly called *back-translation* involves the participation of at least two translators, “one translating from the source to the target language, the second blindly translating back from the target to the source.” This strategy allows the OV to be compared to the retranslated OV, hence the name “back-translation.” Brislin (1970) makes no claim, however, that back-translation is an effective strategy for measuring equivalence (the quality of the translation) between the OV and the TV. On the contrary, he issues several caveats, one of which is that the adaptation process “cannot depend solely on the back-translation” (p. 213), which should be used to modify the TV as well as the OV if equivalence is the goal:

[The concomitant modification of the TV and the OV] contrasts with the awkward, stilted versions common when material in one language is taken as the final content that must be translated with minimal change into another language. (Brislin 1976 : 222)

Brislin’s PA-TH (1970) is often cited, however, in publications describing a PA-AP to justify the use of back-translation as a basic measurement of version equivalence. René de Cotret, Francoeur and Leanza (manuscript in preparation) have identified this reference in 14% of the 131 PA-APs in their study. Here are two examples taken from descriptions of PA-APs:

Questionnaires were translated from English using the back-translation method to ensure content and conceptual equivalence when used with a French-Canadian population (Brislin, 1970; Geisinger, 1994). (Brousseau 2011 : 21)

The scales were translated from English into French using the ‘back-translation’ method (Brislin, 1970), which allowed us to correct the inconsistencies that had appeared during the translation process. (French Bourgeois 2011 : 33) (translation)

Moreover, in over half (54%) of the 131 PA-APs analyzed by René de Cotret *et al.* (manuscript in preparation), back-translation was used as a basic measurement of version equivalence, making it the most popular adaptation strategy in the sample. Back-translation is still widely used today for this purpose, even though Brislin and several other authors have provided rational or empirical arguments for *not* using it to assess version equivalence (e.g., Behr, 2016; Gierl, Rogers & Klinger, 1999; Hambleton & Kanjee, 1995; Harkness & Schoua-Glusberg, 1998; Tassé & Craig, 1999; van Widenfelt, Treffers, De Beurs, Siebelink & Koudijs, 2005).

Looking for a consensual synthesis. With respect to the organization and synthesis of knowledge, Davies *et al.* (2008, p. 189) explain that it is “unclear who has the appropriate perspective to carry out such integrative tasks” and that “drawing on diverse sources prompts new methodological challenges for transparent synthesis.”

The most impressive synthesis comes without doubt from the International Test Commission (ITC). In the 1990s, the ITC created a committee of experts in the adaptation process to formulate “a set of guidelines describing recommended practices in test translations” (van de Vijver & Hambleton 1996 : 11). The first edition of the *ITC Guidelines for Translating and Adapting Tests* was published in 1996; these guidelines are seen as “a thorough approach to the adaptation of psychological instruments” (Bolaños-Medina & González-Ruiz 2012 : 721).

The ITC Guidelines propose general rules, however, and users are free to choose “the best technical procedures available to respect these rules” (Gregoire & Hambleton 2009 : 76). Jeanrie and Bertrand (1999 : 282) suggest that the ITC Guidelines are rather abstract in nature, and probably “more adapted for academics than for potential [PA-AP] users.” As highlighted by Rios and Sireci (2014), there is an important gap between the ITC Guidelines and the adaptation practices observed since 1996 in intercultural studies: of the

28 studies published post-ITC, Rios and Sireci found that “none cited the ITC Guidelines or publications that have promoted them” (p. 304).

Two other initiatives that appeared after 2000 proved unsuccessful. In these cases, the objective was not a synthesis of expert points of view on the adaptation process, but the integration of empirical data on the subject. After reviewing 31 PA-THs, Epstein *et al.* (2015 : 437) concluded that there was “no universal consensus among researchers” and that it was impossible to accurately determine which of the “strategies relevant to translation and adaptation of questionnaires across cultures” were the most efficient. As they put it (p. 439), “the lack of [quantitative] evidence limits the possibilities of clear-cut recommendations.” Their conclusion confirmed earlier findings by Acquadro, Conway, Hareendran and Aaronson (2008 : 518).

In spite of the results of these two quantitative studies, researchers seem to agree on certain strategies proposed in the ITC Guidelines. The synthesis proposed by the ITC does not offer the transparency of scientific methods, however. A qualitative approach would improve the transparency of the synthesis of this knowledge and possibly facilitate its transfer by specifying recommended adaptation strategies available in the literature and agreed on by experts in the field.

From a qualitative research perspective, it is possible to organize and draw firm conclusions from the views expressed by people with diverse backgrounds and interests on a specific topic. “Qualitative research typically includes [...] discerning the perspectives of these people, or what is often referred to as the actor’s point of view” (Sandelowski 2004 : 893). With regard to test adaptation, the “actors” (or experts) are researchers who share theoretical views and propose PA-THs.

The main goal of the research was to determine the existence of any consensual adaptation strategies in the PA-THs available in the scientific literature in order to facilitate knowledge transfer from PA-THs to PA-APs. The research questions were as follows: (Q1) Are there recommended adaptation strategies that are agreed upon in the literature? If so, (Q2) what are they? (Q3) Is it possible to map the adaptation process by categorizing these strategies?

The answers to these questions led to some significant epistemological considerations which will be addressed at the end of the Discussion section.

Method

Literature review. The first author conducted a review in April 2016 to identify PA-THs in articles, book chapters, etc. that present theoretical and methodological aspects of test adaptation that result in recommended adaptation strategies. The review included texts in French or English published since 1970, with the exception of those (a) restricted to statistical analysis, (b) consisting essentially of a summary of previously published material, or (c) referring to tests that were not self-reported (for example, tests for people with hearing disabilities).

Under the supervision of a librarian, a traditional keyword search was used in PsycNet and PubMed (“test or questionnaire or instrument” combined with “adaptation or translation”). The search was limited to the “Abstract” field in PsycNet, and the “Title/Abstract” field in PubMed to exclude the tremendous number of texts which contained the keywords but did not meet the main criterion, which was to describe a PA-TH. Secondly, a footnote chasing (White 2009) was performed by going through the references of the texts kept from the previous step.

After this “upstream review,” a “downstream review” was performed using Google Scholar to access texts citing those retained from the previous two steps, as “the web search based Google Scholar can provide a more extensive picture of scientific activity that covers a broader scale of scientific output than traditional databases” (Aaltojarvi, Arminen, Auranen & Pasanen 2008 : 6).

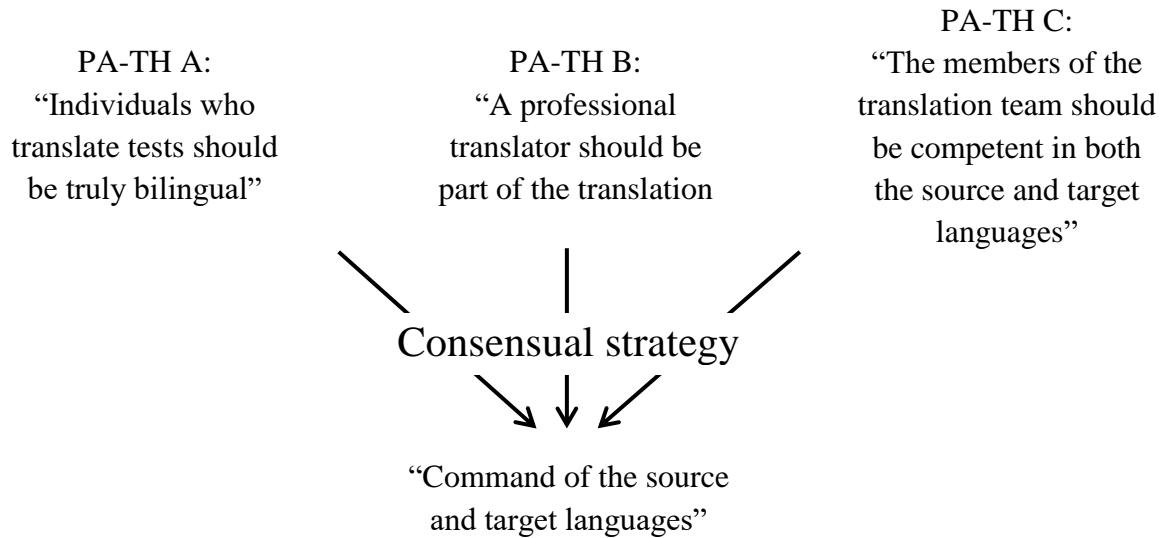
Content analysis. Content analysis was used to determine consensual adaptation strategies in the PA-THs found. This is a method known “for making replicable and valid inferences from texts to the contexts of their use” (Krippendorff 2004 : 18). As explained by Krippendorff, “the content analyst views data as representations not of physical events but of texts [...] that are created to be seen, read, interpreted, and acted on for their meanings, and must therefore be analyzed with such uses in mind” (p. XIII).

Each PA-TH used in the study describes adaptation strategies, some of which are described in a few words, others in lengthy paragraphs. The first task was to identify each PA-TH's adaptation strategies, which were then compared to find a way to categorize them based on a set of rules, creating categories which a) took into account the largest number of strategies, b) represented those strategies in a relevant way, c) were homogeneous and d) were mutually exclusive. After a first set of categories was established, it was applied to the strategies found in each PA-TH. Changes were then made to improve the categorization, and so on. After working on his own, the first author then discussed his results with the other two authors in order to produce a final list of categories. Once the definitive list was established, the three authors calculated the number of strategies in each category and interpreted the results.

If a given adaptation strategy was mentioned more than once in a PA-TH, it was counted as a single occurrence. So, even if a strategy like "individuals who translate tests should be truly bilingual" was mentioned several times in the same PA-TH, it still counted as a single instance. The reason for this was simple: our goal was not to establish the relative significance of a specific strategy in each particular PA-TH, but rather in all the PA-THs, in order to determine whether some strategies were agreed upon by the various PA-THs.

We can take this example one step further. If a second PA-TH recommended that "a professional translator should be part of the translation team," and a third argued that "the members of the translation team should be competent in both the source and target languages," all three occurrences were grouped under the same category, namely, "Command of the source and target languages." If these three PA-THs represented the majority (more than 50%) of the corpus, the conclusion would be that there was consensus on the fact that people who translate tests should be fluent in languages involved. Although the comments in the three PA-THs do not appear to be consensual when considered separately, consensus is created when the comments are grouped under a single category as shown in Figure 2.

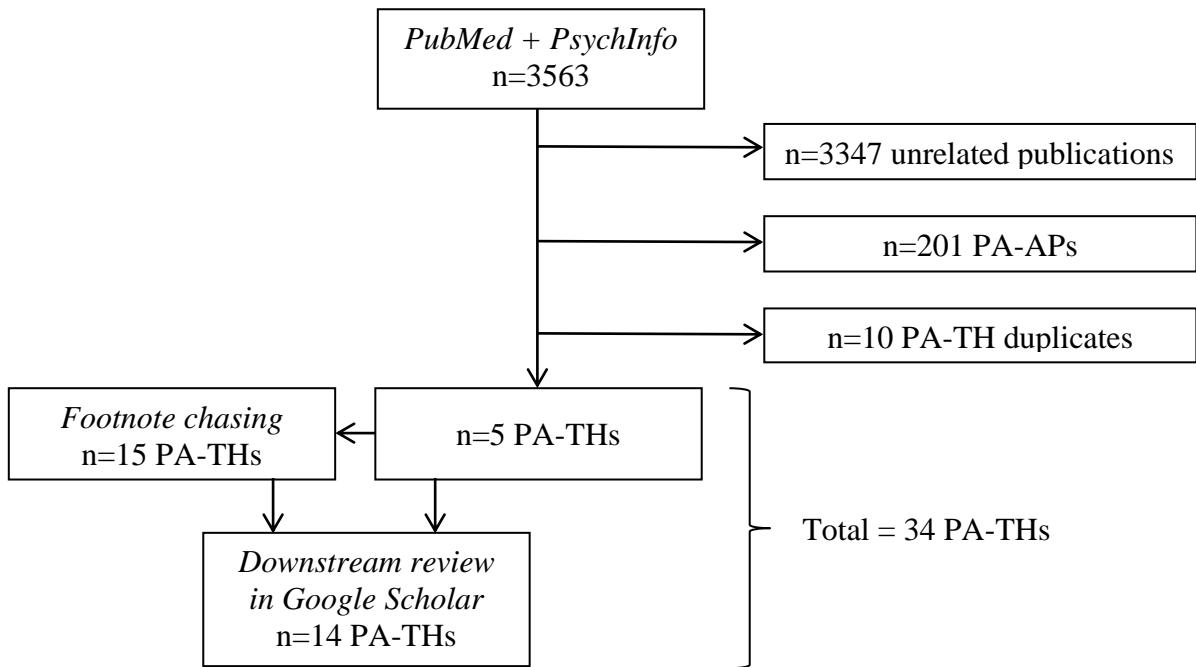
Figure 2. Example of the development of a consensual adaptation strategy by considering the strategies described in PA-THs



Results

Of the 216 texts returned from the keyword search in PsycNET and PubMed, 201 did not describe a PA-TH but a PA-AP, so they were rejected. Among the remaining 15 texts, seven referred to the ITC Guidelines (transformed into a PA-TH by Hambleton & Patsula 1999) and three, to a PA-TH proposed by Butcher (1996), so they were also excluded. Thus, at the end of this first step, only five PA-THs remained. Fifteen more were found by following up the references in those remaining five. Finally, Google Scholar was used to access texts that cited those identified in the previous two steps, resulting in an additional 14 PA-THs (Figure 3).

Figure 3. Search strategy for the literature review



In all, 34 PA-THs that met the criteria were identified: 31 in English, and three in French. As shown in Table 1, a majority of the texts in the corpus are peer-reviewed articles, three were not officially published, most were published in the 2000s, and nine refer to PA-THs that are used by organizations. Different fields are also represented: intercultural research (14 texts), psychology (7), medicine (4), education sciences (3), nursing (2), linguistics (2), translation studies (1) and statistics (1).

Table 1

List of texts in the corpus, with year of publication, authors (and organisations, where applicable), text type, and citations count

Year	Author(s)/Organisation	Text type
1970	Brislin	Scientific article
1976	Spielberger & Sharma	Book chapter
1979	Chapman & Carter	Scientific article
1989	Vallerand	Scientific article
1991	Bracken & Barona	Scientific article
1992	Prieto	Scientific article
1993	Guillemin, Bombardier & Beaton	Scientific article
1994	Bradley	Book chapter
1994	Geisinger	Scientific article
1994	Sperber	Scientific article
1996	Butcher	Book chapter
1996	Maxwell (IEA)*	Research report **
1999	Hambleton & Patsula (ITC)*	Scientific article
1999	Jeanrie & Bertrand (ITC)*	Scientific article
1999	Tassé & Craig	Book chapter
2000	Beaton, Bombardier, Guillemin & Ferraz (AAOS)*	Scientific article
2000	Carlson	Scientific article
2000	Vissandjée & Dupéré	Scientific article
2001	Jones, Lee, Phillips, Zhang & Jaceldo	Scientific article
2002	Massoubre, Lang, Jaeger <i>et al.</i>	Scientific article
2005	Chàvez & Canino (HSRI)*	Governmental doc. **
2005	Eremenco, Cella & Arnold	Scientific article
2005	Pan & de la Puente (U.S. Bureau of the Census)	Research report
2005	Van Widenfelt, Treffers, de Beurs <i>et al.</i>	Scientific article
2006	Wang, Lee & Fetzer	Scientific article
2006	Weidmer, Hurtado, Weech-Maldonado <i>et al.</i> (AHRQ)*	Research report
2007	Cha, Kim & Erlen	Scientific article
2007	Weeks, Swerissen & Belfrage	Scientific article
2008	Willis, Brick, Norberg <i>et al.</i>	Conference abstract
2009	World Health Organization (WHO)*	Website **
2009	Solano-Flores, Backhoff & Contreras-Nino	Scientific article
2011	Harkness	Book chapter
2013	Arffman (ODCE / IEA)*	Scientific article
2013	Ercikan & Lyons-Thomas	Book chapter

*IEA (International Association for the Evaluation of Educational Achievement), ITC (International Test Commission), AAOS (American Association of Orthopaedic Surgeons), HSRI (Human Services Research Institute), AHRQ (Agency for Healthcare Research and Quality), WHO (World Health Organization), OECD (Organisation for Economic Co-operation and Development).

**Nonpublished material.

Each of the 34 PA-THs contains an average of 13 adaptation strategies (range=7-22). The vast majority (96%) of these strategies were broken down into 17 consensual strategies. The consensual strategies were regrouped into five meta-categories called “Phases of the adaptation process.” As shown in Table 2, there are nine primary strategies, some of which are further broken down into secondary strategies (a total of eight).

The strategy called *Working in a multidisciplinary setting* covers three phases and takes different forms, which is why it counts as three distinct primary strategies. This exception will be explained in the following paragraphs, as will the reason why no strategies were listed for the Testing phase.

Table 2

The 17 adaptation strategies (S) which reached consensus (%), regrouped into five phases

Phase	Primary strategy	Secondary strategy
1: Initial Translation	S-1.1: Working in a multidisciplinary setting (91%)*	S-1.1.1: Command of both languages (100%)
		S-1.1.2: Understanding of the construct (81%)
		S-1.1.3: Experience in both cultures (72%)
	S-1.2: Producing at least two translated versions (66%)	n/a.
2: Internal Review	S-1.3: Providing translation guidelines (63%)	n/a.
	S-1.4: Translating all the components of the test (63%)	n/a.
		S-2.1.1: Back-translation (68%)
3: External Review	S-2.1: Overseeing the internal review (97%)	S-2.1.2: Other review strategies (65%)
	S-2.2: Working in a multidisciplinary setting (85%)	S-2.2.1: Additional qualifications (59%)
4: Testing	S-3.1: Overseeing the external review (81%)	S-3.1.1: Interview with subjects (56%)
	S-3.2: Working in a multidisciplinary setting (53%)**	S-3.1.2: Other review strategies (56%)
	n/a.***	n/a.
Supra Phase	S-SP: Documenting the adaptation process (53%)	n/a.

*This strategy applies to the first three phases, for which secondary strategies and level of consensus may vary. **The percentage varies from Phase II, but the setting is the same (the review committee).

*** Testing strategies is not detailed, as previously noted.

Phase 1: Initial Translation. Once it has been decided that the original version of a test is to be adapted, the first phase of the process is the translation per se (OV → TV). When more than one TV is produced, the reconciliation of the independent versions is carried out either in this first phase or in the next; there was no consensus on this point.

S-1.1: Working in a multidisciplinary setting. This primary strategy—the most frequently mentioned—allows for the contribution of people with various areas of expertise. It applies to the first three phases of the process, although the setting in Phase I (a translation team or independent translators) must be distinguished from the one in Phases II and III (a review committee⁴). The review committee may or may not include the translators who participated in Phase I; there is no consensus on this point in the corpus.

S-1.1.1: Command of the source and target languages. There is unanimity on the fact that people who translate tests should know both the source and target languages. Half of the PA-THs (50%) recommend including at least one professional translator on the translation team. At the other extreme, about a quarter of the texts (22%) recommend that people who translate tests should be truly *fluent* or *bilingual*. The rest (28%) are somewhere in the middle, either recommending that the translators be native speakers of the target language and/or specifying what they mean by “bilingual.” For example, Jeanrie and Bertrand (1999 : 280) refer to “functionally bilingual” translators, meaning that they are “able to conduct professional activities in both languages,” while Butcher (1996) suggests that people who translate tests should have lived in each country involved for at least five years.

S-1.1.2: Understanding of the construct. A majority of authors (81% of the texts) make the minimal recommendation that the translators be familiar with the construct being assessed and aware that the test items are meant to measure this construct. Translators should be “sensible of meaning in a psychological or educational measurement device” (Chapman & Carter 1979 : 75), have “some knowledge of the construct being measured” (Hambleton & Patsula 1999 : 5), or have a knowledge of the object of the test and be able to identify the concept behind each item (Massoubre, Lang, Jaeger, Jullien & Pellet 2002). Forty-two

⁴ The terms *review committee* and *committee approach* should not be confused. The latter may refer to translators only (e.g., Brislin, 1970; Butcher, 1996; Cha, Kim & Erlen, 2007), reviewers only (e.g., Guillemin *et al.*, 1993; Vallerand, 1989; Weidmer, Hurtado, Weech-Maldonado, Ngo-Metzger & Bogen, 2006), or both translators and reviewers (e.g., Hambleton & Patsula, 1999; Tassé & Craig, 1999; Weeks *et al.*, 2007).

percent of the texts recommend that at least one translator have considerable expertise in the construct, meaning that they should be a psychologist/psychiatrist (Spielberger & Sharma 1976), a test developer (Tassé & Craig 1999), or someone “teaching and learning in the subject for educational achievement tests” (Ercikan & Lyons-Thomas 2013 : 553).

S-1.1.3: Experience in the source and target cultures. This secondary strategy, represented in 72% of the texts, was the most difficult to capture because of the vague way in which the notion of “culture” is raised by the authors. In half of the texts that address the issue, a lack of information about what exactly “knowledge of the culture” entails makes it impossible to understand what exactly is being recommended. For example, Geisinger (1994 : 306) refers to translators who are “extremely knowledgeable about both cultures,” and Pan and De La Puente (2005 : A2) highlight the importance of using “culturally appropriate expressions” when translating a test.

The other half of the texts state that the translators should live or have lived in the country or region where the AV will be used. Arffman (2013), for example, refers to “national translators,” insisting on the importance of taking their extra-linguistic knowledge into consideration. This notion of “extra-linguistic knowledge” is developed in Vissandjée and Dupéré (2000), who argue that being culturally competent does not only imply an awareness and understanding of the conceptual universe of the culture being studied but also the ability to use this understanding.

Some authors recommend harmonizing the TV to ensure that the terminology used is understood by different populations sharing the same language (Eremenco, Cella & Arnold 2005; Harkness 2011); others consider that “cultural authenticity” is better (Ercikan & Lyons-Thomas 2013) unless the test aims at intercultural comparisons.

S-2.2.1: Additional qualifications. More than half of the texts (59%) recommend that the reviewers have a more thorough and/or more diverse expertise than the translators. What this means, essentially, is that the members of the review committee should be experts both in the construct and in test development. “[K]nowledge of survey measurement goals, questionnaire design and survey methods [...] extends beyond the level of such knowledge generally represented by the initial translating team” (Willis *et al.* 2008 : 5). It is

recommended in 34% of the texts that at least one of the translators understand the logic of test development. Bradley (1994 : 44) argues that the translator “needs an appreciation of the requirements and pitfalls of item construction and questionnaire design.”

This being said, many authors, including Solano-Flores, Backhoff and Contreras-Niño (2009), insist on the participation of test developers and psychometricians in Phases I and II. Only Tassé and Craig (1999 : 175) advocate for more competencies on the part of the translator(s).

Some authors also refer to review committee members’ command of the source and target languages. In 44% of the texts, it is proposed that the review committee include at least one professional translator or linguist. Interestingly, the majority of the authors who discuss command of languages in Phases II and III are referring to a professional translator or linguist. In comparison, even though all the authors insist on a command of the source and target languages in Phase I, only half of them advocate for the inclusion of a professional translator.

S-1.2: Producing at least two TVs. Two-thirds of the texts recommend that at least two TVs be produced. According to Vallerand (1989 : 665), this serves to minimize linguistic, psychological and comprehension biases. Having multiple TVs enables the review committee to establish comparisons and detect “discrepancies that may reflect more ambiguous wording in the original or discrepancies in the translation process” (Beaton, Bombardier, Guillemin & Ferraz 2000 : 3188). Arffman (2013) does not totally share these views. She argues that using a single TV “may be at least as effective, if not even more effective than double translation,” (p. 7) although she admits that more research is needed on this matter. In her opinion, as soon as texts are merged, consistency and coherence can be affected.

S-1.3: Providing translation guidelines. Almost two-thirds of the texts (63%) recommend that translators be given precise instructions, although opinions vary on how detailed those instructions should be. Half of the cases suggest that a series of rules similar to Brislin’s be provided (1970, 1980) for phrasing items: “Use short, simple sentences of less than 16 words,” “Use the active rather than the passive voice,” etc. Not surprisingly, back-

translation, which is facilitated by such rules, is also recommended in those cases. The other half of the cases prefer to suggest that background information be provided on the scope and purpose of the translation. For example, Jeanrie and Bertrand (1999) recommend that translators be given “written information concerning the kind of translation that [is] expected,” while Pan and De La Puente (2005) think they should be told who the target audience is, and be briefed on the type of questionnaire involved, the media used, and the level of formality expected in the target language. Translation guidelines can even be in the form of a comprehensive document (Maxwell 1996) or based on scientific literature (Arffman 2013).

S-1.4: *Translating all the components of the test.* Sixty-three percent of the texts recommend that not only the items be translated, but all of the material related to the test as well: instructions, answer sheets, scoring manual, etc.

Phase 2: Internal Review. The Internal Review of the TV(s) is based on the opinion of experts. After the Internal Review, only one TV remains.

S-2.1: *Overseeing the Internal Review.* Virtually all the authors (97% of the texts) recommend the implementation of at least one strategy to oversee the Internal Review. Back-translation is suggested for this purpose in 68% of the texts.

S-2.1.1: *Back-translation.* Half of the authors who recommend back-translation are wary of its excessive use. Hambleton and Patsula (1999 : 5) explain that it can “be revealing of poor translations but without a focus on the target language version of the test, problems in the adaptation can be missed” (p. 5). Beaton *et al.* (2000 : 3188) note that back-translation “is only one type of validity check, highlighting gross inconsistencies or conceptual errors in the translation.”

S-2.1.2: *Other review strategies.* Oversight strategies other than back-translation are mentioned in 65% of the texts. Some suggest basing the review on the notion of “equivalence,” which, as discussed by Herdman, Fox-Rushby, and Badia (1997), is a rather ambiguous one. For example, Jeanrie and Bertrand (1999) propose a rating scale that “focuses on content, conceptual, and linguistic types of equivalence,” while Beaton *et al.*

(2000) take four levels of equivalence into consideration: semantic, idiomatic, experiential, and conceptual. Others propose using rating scales that do not directly involve the notion of equivalence. Spielberger and Sharma (1976) suggest that reviewers rate each item based on whether they consider the translation to be “good,” “satisfactory,” or “unsatisfactory.” Maxwell (1996) proposes a combination of “severity codes” ranging from 1 (serious error) to 4 (acceptable adaptation), and 14 “type codes” such as “A. Spelling, B. Grammar, C. Vocabulary” to indicate, if applicable, the kind of change needed. The theory of test translation error proposed by Solano-Flores *et al.* (2009) draws largely on the idea of rating items based on such dimensions such as spelling, grammar, meaning, or level of difficulty.

Phase 3: External Review. The purpose of the External Review is to seek input on the TV from representatives of the target population and revise the TV accordingly. This phase is generally called the *pretest*, a term that actually obscures the objective of external review, which is to give subjects and researchers a chance to exchange views on the TV. Another term used is *field testing*, which can be confusing since it can also refer to the fourth phase described below. For these reasons, the term *External Review* was chosen, which reinforces the contrast with the Internal Review, and echoes the term *external validity*. While external validity is concerned with generalizing results and answering a question like “Does the opinion of the experts correspond to that of the target population?”, internal validity is meant to control the variables that can affect results. In the context of test adaptation, the Internal Review seeks to structure, or even measure, the translation process.

S-3.1: Overseeing the External Review. In 81% of the PA-THs, at least one strategy is suggested to oversee the External Review. The strategies proposed involve asking individuals from the target population to read the TV, or administering the TV to a limited number of subjects from the target population.

S-3.1.1: Interview with subjects of the target population. One of the consensual secondary strategies (56%) is to interview subjects from the target population during or after administration of the TV. This is often referred to as the *probe technique*, although the term *cognitive interview* seems to be more widely used in recent studies. This type of interview asks respondents, item by item, “what they thought the question was asking, whether they

could repeat the question in their own words, what came to their mind when they heard a particular phrase or term,” as well as “how they [chose] their answer” (World Health Organization 2009 : n.p.).

S-3.1.2: Other review strategies. The most popular strategy in this category is to administer the TV and the OV to bilingual subjects who are able to judge their equivalence; this can involve statistical analysis. Guillemin *et al.* (1993 : 1424) state that it is effective to “detect possible discrepancies,” and “help pinpoint any inadequacy of the final version with the cultural context.”

Another popular strategy is to ask subjects to circle ambiguous items or to use Likert scales. According to Vallerand (1989), items that obtain an average quality score of four or less out of seven should be modified. Sperber, Devellis and Boehlecke (1994) propose that participants from the source population be asked to formally compare the OV with the back-translation based on two measurements (comparability of language and similarity of interpretability), using “Likert scales ranging from 1 (extremely comparable/extremely similar) through 4 (moderately comparable/moderately similar) to 7 (not at all comparable/not at all similar)” (p. 506).

The third most popular strategy is the use of focus groups. Chávez and Canino (2005 : 17) suggest conducting two focus group sessions for each culturally diverse group involved, the first “to discuss the construct that is supposed to be measured by the instrument” and the second “to [go] over every item contained in the instrument in detail.” Instead of submitting the TV to potential subjects, Bradley (1994) proposes that it be submitted to potential *users*, i.e. clinicians and examiners. She argues (p. 46) that “[i]f the questionnaire does not suit their needs and preferences it is unlikely to be used,” adding that such a method enables those potential users to detect items that are not clear to them.

Phase 4: Testing. The Testing phase consists in the administration of the TV to dozens or hundreds of subjects of the target population, followed by statistical analysis and data interpretation. It provides insights on the psychometric properties of the TV, and is ultimately used to establish the equivalence “from a measurement level perspective” (van de Vijver & Tanzer 2004 : 121) between the OV and the TV. When the TV is judged *not*

significantly different from the OV, it gains the status of adapted version (AV) (TV→AV) and, consequently, of a legitimate test, like the OV. If there are significant differences, the TV goes through the first three phases again, with a focus on translated items that are statistically problematic.

The Testing phase is presented as the achievement of the adaptation process in the vast majority of the corpus. Since our study focuses on the previous phases in the process, the strategies used in this phase are not detailed here and were not included in the content analysis. Testing is not proper to the process of adaptation, in the sense that every test—adapted or not—goes through this phase. For a survey of testing strategies, see Geisinger (2013) and Hambleton, Merenda and Spielberger (2005).

Supra Phase. The fifth phase of the adaptation process is called the Supra Phase because it applies to the previous phases. In this phase, the sole strategy that reached consensus was called “S-SP: Documenting the process” after Hambleton and Patsula (1999), Harkness (2011), Pan and De La Puente (2005) and Willis *et al.* (2008).

The majority of the PA-THs (53%) insist that PA-APs should detail the strategies applied so that readers can: (a) evaluate the methodological rigour and validity of the process, (b) easily replicate it, and/or (c) be aware of problems encountered. In fact, some PA-THs in the corpus are taken from texts that offer comprehensive descriptions of a PA-AP (e.g., Jeanrie & Bertrand 1999; Sperber *et al.* 1994; Wang, Lee & Fetzer 2006), which shows that documenting PA-APs can give rise to theoretical insights and significant discussion. As noted by Van Widenfelt, Treffers, De Beurs, Siebelink and Koudijs (2005 : 145), however, “the current common practice” of describing PA-APs is “to give a minimal report of procedures (i.e., ‘a translated version was used’; ‘a back-translation was conducted’).”

According to the World Health Organization (2009), “[a]ll the cultural adaptation procedures should be traceable through the appropriate documents,” and researchers should minimally provide (a) the TV(s) and back-translation (if applicable), (b) a summary of the experts’ recommendations, (c) a summary of problems found during the External Review and the solutions proposed, and (d) the AV.

Discussion

The answers to the research questions gave rise to one corollary, related to the fact that the objective of the phases of the adaptation process prior to the Testing phase is not to establish equivalence but rather to encourage discussion among the actors involved in negotiating the meaning of the equivalence. These two distinct ways of addressing equivalence call for an examination of the strategies that are considered valid in the literature—not only for the purpose of adapting tests, but also to establish the validity of the adaptation process per se.

The hypothesis has been confirmed, and the research questions have been answered: during the course of this research, it was determined that there are indeed strategies for test adaptation that are agreed on in the literature. The result is a straightforward schematization of the adaptation process (see Table 2) that should facilitate knowledge transfer from PA-THs to PA-Aps. It should also allow information that would otherwise be difficult to discern to be accounted for and used.

Looking at the diversity of strategies recommended provided us with a prototypical image of the adaptation process. Even though none of the PA-THs we reviewed combined the 17 consensual strategies or proposed a five-phase process similar to the one described here, all of them tend in that direction. Obviously, this image—and each strategy taken individually—does nothing more than show a certain direction that PA-APs could take. None of the consensual strategies could be considered mandatory according to the results since they are theoretical and therefore acontextual. Consequently, the consensual strategies highlighted should not be seen as principles that have to be strictly followed, but rather as a range of tools available to researchers who wish to use strategies that are considered scientifically legitimate.

Some of the 17 consensual strategies are subject to interpretation, and their applicability is limited in one way or another. There is no guarantee that a strategy used in one particular adaptation process will be implemented in the same way during another adaptation process. For example, *Producing translation guidelines* is a strategy that leaves room for methodological freedom: while one researcher might prepare write several pages of

guidelines and provide multiple examples based on previous PA-Aps, another might do no more than just give the translators a few simple instructions. The 17 strategies highlighted go only as far as the current consensual knowledge, and their limitations in terms of applicability should not be seen as a methodological flaw. To return to the example above, there is no consensus on how translation guidelines should be produced.

This does not imply, however, that precise criteria should be issued with regard to strategy implementation. The study results show that the quality of an adapted version depends primarily on the first three phases of the adaptation process and these phases cannot be overly operationalized. These crucial points are addressed in the last subsection.

It is important at this point to say a few words about one of the adaptation strategies that does not achieve consensus: *Ensuring that construct equivalence exists in the language and cultural groups of interest* (Hambleton & Patsula 1999 : 4). We believe the reason for this lack of consensus to be that this strategy represents a preliminary step in the adaptation process as conceptualized and broken down into five phases in the vast majority of the PA-THs in the study (see Table 2). It is our opinion, however, that construct equivalence between two languages/cultures should not just be proposed, but duly evaluated as recommended by Hambleton and Patsula (1999).

Keyword issues. At another level, the study shows that keywords such as *translation*, *adaptation* and *test* are not precise enough to find relevant PA-THs in the literature. Only five PA-THs were found using those keywords, and they were lost in a sea of PA-APs, which are far more numerous in the literature. Part of the problem is that all of these terms are very polysemous. In this respect, promoting the use of distinct terms (e.g., adaptation process, PA-TH and PA-AP) could contribute to the recognition of a major phenomenon in psychology and in other scientific fields where tests are widely used, as it would facilitate knowledge transfer to improve practices.

A process of deconstruction. Most of the 17 consensual strategies identified in the study imply that individuals construct their own reality based on who they are and what they experience. In the first three phases of the adaptation process, translators, reviewers and subjects from the target population are asked to share their views on the test being adapted.

Who they are, where they come from, and so on, will necessarily be reflected in their views. The advantage of working in a multidisciplinary setting—the most frequent strategy mentioned in the corpus—is clear: it implies a plurality of expertise (and thus of experiences). The production of at least two TVs follows the same logic: the respective views of each translator or subgroup of translators are considered before being contrasted in order to stimulate reflection. This logic can even be applied to back-translation. Van Widenfelt *et al.* (2005 : 142) express this clearly: “We advocate to primarily use the back-translation procedure in an earlier phase of the translation process as an important feedback method [which] often leads to additional discussion...”

Many PA-THs recommend specifying who participated in Phases I, II and III of the process. Where does each translator come from? Has s/he ever translated tests? Does s/he have a degree in translation, and does s/he work as a professional translator? To what extent is a given reviewer familiar with the construct being assessed? Is s/he a clinician? Who is this particular representative of the target population, and where does s/he come from? Why was s/he chosen? The Supra Phase aims at documenting this kind of information in order to make it available (and enhance the credibility of the process).

To sum up, the OV as an original instrument must be meticulously deconstructed and reconstructed by a diversity of individuals before it reaches the last phase of the adaptation process. The translators, reviewers and representatives of the target population all contribute in their own way: “I translated it this way because...,” “Would it be better to use the term X instead of Y?”, “The phrasing of this item is not clear; it makes me think of...,” and so on. By raising such issues, the reader freely interprets the items, and is encouraged to share their interpretation with other individuals who are also participating in the process.

This being said, a drastic change takes place in Phase IV, when the reader’s interpretation of the test is influenced, even predetermined, by the answers at their disposal, which are the same for all readers. This is when the reader’s interpretation yields quantitative results, through some sort of funnel effect. At this stage, the reader becomes the object of the test. S/he is no longer the interpreter of the test: it is s/he who is being “interpreted” by the test. This reversal in the reader-test relationship ultimately brings back the objectivity which the

test was supposed to have from the outset, i.e. as the OV. Although the scientific credibility of the process relies on this reversal, the Testing phase does not aim to evaluate the adaptation process but rather to establish the equivalence between the TV and the OV. Once the equivalence has been established, the conclusion is that the adaptation process was valid.

The results of our study offer an alternative, supporting the view that it is possible to directly judge the scientific credibility of the adaptation process. Because they are recommended by a majority of experts in the field, the 17 consensual strategies identified through content analysis provide such credibility, in the sense that the scientific community recognizes their use. Prior to our study, none of those strategies had a clear status because of the apparent lack of consensus (Acquadro *et al.* 2008; Epstein *et al.* 2015). They would go virtually unnoticed compared to the Testing phase, which is recognized as being rigorous and transparent. Those qualities have now been extended to the whole adaptation process through the 17 consensual strategies, most of which imply a plurality of views, a quality proper to qualitative research and an essential part of “deconstructing” any OV and “reconstructing” it into an AV.

References

- Aaltojärvi, I., Arminen, I., Auranen, O., Pasanen, H. M.: Scientific productivity, web visibility and citation patterns in sixteen Nordic sociology departments. *Acta Sociol.* 51(1), 5–22 (2008)
- Acquadro, C., Conway, K., Hareendran, A., Aaronson, N.: Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health.* 11(3), 509–521 (2008)
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education: Standards for educational and psychological testing: American Educational Research Association (1999)
- Arffman, I.: Problems and issues in translating international educational achievement tests. *Educational Measurement: Issues and Practice.* 32(2), 2–14 (2013)
- Behr, D.: Assessing the use of back translation: the shortcomings of back translation as a quality testing method. *Int J Soc Res Methodol.* 20(6), 573–584 (2016)
- Beaton, D. E., Bombardier, C., Guillemin, F., Ferraz, M. B.: Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 25(24), 3186–3191 (2000)
- Bolaños-Medina, A., González-Ruiz, V.: Deconstructing the Translation of Psychological Tests. *Meta: Translators' Journal.* 57(3), 715–739 (2012)
- Bradley, C.: Translation of questionnaires for use in different languages and cultures. In: Bradley, C. (ed.) *Handbook of Psychology and Diabetes: a guide to psychological measurement in diabetes research and practice*, pp. 423. Harwood Academic Publisher, Swiss (1994)
- Brislin, R. W.: Back-translation for cross-cultural research. *J Cross Cult Psychol.* 1(3), 185–216 (1970)
- Brislin, R. W.: Comparative research methodology: Cross-cultural studies. *Int J Psychol.* 11(3), 215–229 (1976)
- Brousseau, M.: He said, she said: An examination of sexual coercion from the perspective of mixed-sex couples. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada (2011)
- Butcher, J. N.: Translation and adaptation of the MMPI-2 for international use. In: Butcher, J. N. (ed.) *International adaptation of the MMPI-2: research and clinical application*, pp. 26–43). University of Minnesota Press, MN (1996)

- Cha, E. S., Kim, K. H., Erlen, J. A.: Translation of scales in cross-cultural research: Issues and techniques. *J Adv Nurs.* 58(4), 386–395 (2007)
- Chapman, D. W., Carter, J. F.: Translation procedures for the cross-cultural use of measurement instruments. *Educ Eval Policy Anal.* 1(3), 71–76 (1979)
- Chávez, L. M., Canino, G.: Toolkit on translating and adapting instruments. www.hsri.org (2005). Accessed June 2018
- Davies, H., Nutley, S., Walter, I.: Why ‘knowledge transfer’ is misconceived for applied social research. *J Health Serv Res Policy.* 13(3), 188–190 (2008)
- Epstein, J., Santo, R. M., Guillemin, F.: A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol.* 68(4), 435–441 (2015)
- Ercikan, K., Lyons-Thomas, J.: Adapting tests for use in other languages and cultures. In: Geisinger, K. F., Bracken, B. A., Carlson, J. F., Hansen, J.-I. C., Kuncel, N. R., Reise, S. P., Rodriguez, M. C. (eds.) (2013). *APA handbooks in psychology. APA handbook of testing and assessment in psychology, Vol. 3: Testing and assessment in school psychology and education*, pp. 545–569. American Psychological Association, Washington DC (2013)
- Eremenco, S. L., Cella, D., Arnold, B. J.: A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof.* 28(2), 212–232 (2005)
- Fournier, M.: La mobilisation des connaissances: synthèse des réflexions sur les pratiques à l’AETMIS. www.inesss.qc.ca (2011), Accessed March 2019.
- French Bourgeois, L.: L’influence des normes sociales sur le comportement : un processus médié par l’élaboration de stratégies. (M. Sc.), Université de Montréal, Montreal (2011)
- Geisinger, K. F.: Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychol Assess.* 6(4), 304 (1994)
- Geisinger, K. F.: Testing and Assessment in Cross-Cultural Psychology. In Graham, J. R., Naglieri, J. A. (eds.) *Handbook of psychology - Assessment Psychology*, pp. 114–139. John Wiley & Sons Inc., Hoboken (2013)
- Gierl, M. J., Rogers, W. T., Klinger, D.: Using statistical and judgmental reviews to identify and interpret translation DIF. In annual meeting of the National Council on Measurement in Education, Montreal (1999)

- Gregoire, J., Hambleton, R. K.: Advances in test adaptation research: a special issue. *Int J Test*, 9(2), 75–77 (2009)
- Guillemain, F., Bombardier, C., Beaton, D.: Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*, 46(12), 1417–1432 (1993)
- Hambleton, R. K., Kanjee, A.: Increasing the validity of cross-cultural assessments: Use of improved methods for test adaptations. *Eur J Psychol Assess*, 11(3), 147–157 (1995)
- Hambleton, R. K., Merenda, P. F., Spielberger, C. D.: Adapting educational and psychological tests for cross-cultural assessment. Psychology Press, East Sussex (2005)
- Hambleton, R. K., Patsula, L.: Increasing the validity of adapted tests: myths to be avoided and guidelines for improving test adaptation practices. *J Appl Test Tech*, 1(1), 1–13 (1999)
- Harkness, J. A.: Translation. In Survey Research Center (ed.) *Guidelines for best practice in cross-cultural surveys*, 3rd ed., pp. 1–35. University of Michigan, MI (2011)
- Harkness, J.A., Schoua-Glusberg, A.: Questionnaires in translation. In: Harkness, J. (ed.) *Cross-culturasl survey equivalence*, pp. 87–126. Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA- (1998)
- Herdman, M., Fox-Rushby, J., Badia, X.: ‘Equivalence’ and the translation and adaptation of health-related quality of life questionnaires. *Qual Life Res*, 6(3), 237–247 (1997)
- Jeanrie, C., Bertrand, R.: Translating tests with the International Test Commission's guidelines: Keeping validity in mind. *Eur J Psychol Assess*, 15(3), 277–283 (1999)
- Krippendorff, K.: Content analysis - An introduction to its methodology. Sage, New York (2004)
- Massoubre, C., Lang, F., Jaeger, B., Jullien, M., Pellet, J.: [The translation of questionnaires and of tests: techniques and problems]. *Can J Psychiatry*, 47(1), 61–67 (2002)
- Maxwell, B.: Translation and cultural adaptation of the survey instruments. In: Martin, M. O., Kelly, D. L. (eds.) *Third international mathematics and science study (TIMSS) technical report*, pp. 159–169. Boston College, MA (1996)
- Nida, E.: Toward a science of translating. Brill, Leiden (1964)

- Pan, Y., De La Puente, M.: Census Bureau guideline for the translation of data collection instruments and supporting materials: Documentation on how the guideline was developed. US Bureau of the Census. www.census.gov (2005). Accessed April 2016
- René de Cotret, F., Francoeur, A., Lanza, Y.: How to describe the adaptation process: A typology proposal and review of adaptation practices in Quebec from 2000 to 2014. Manuscript in preparation (2019)
- Rios, J. A., Sireci, S. G.: Guidelines versus practices in cross-lingual assessment: A disconcerting disconnect. *Int J Test.* 14(4), 289–312 (2014)
- Sandelowski, M.: Qualitative Research. In Lewis-Beck, M. S., Bryman, A. E., Liao, T. F. (eds.) *The SAGE Encyclopedia of Social Science Research Methods*, Vol. 1, pp. 893–894. SAGE, Thousand Oaks (2004)
- Schuman, H.: The random probe: technique for evaluating the quality of closed questions. *Am Sociol Rev.* 31, 218–222 (1966)
- Solano-Flores, G., Backhoff, E., Contreras-Niño, L. Á.: Theory of test translation error. *Int J Test.* 9(2), 78–91 (2009)
- Sperber, A. D., Devellis, R. F., Boehlecke, B.: Cross-cultural translation methodology and validation. *J Cross Cult Psychol.* 25(4), 501–524 (1994)
- Spielberger, C. D., Sharma, S.: Cross-cultural measurement of anxiety. In: Spielberger, C. D., Diaz-Guerrero, R. (eds.) *Cross-cultural anxiety*. John Wiley & Sons, New York, Vol. 1, pp. 13–25 (1976)
- Tassé, M. J., Craig, E. M.: Critical issues in the cross-cultural assessment of adaptive behaviour. In: Schalock, R. L. (ed.) *Adaptive behavior and its measurement: Implications for the field of mental retardation*, pp. 161–184. American Association on Mental Retardation, Washington, DC (1999)
- Vallerand, R. J.: Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques : Implications pour la recherche en langue française. *Can Psychol.* 30(4), 662 (1989)
- van de Vijver, F., Hambleton, R. K.: Translating tests. *Eur Psychol.* 1(2), 89–99 (1996)
- van de Vijver, F., Tanzer, N. K. Bias and equivalence in cross-cultural assessment: An overview. *Eur Rev Appl Psychol.* 54(2), 119–135 (2004)
- van Widenfelt, B. M., Treffers, P. D., De Beurs, E., Siebelink, B. M., & Koudijs, E.: Translation and cross-cultural adaptation of assessment instruments used in

psychological research with children and families. *Clin Child Fam Psychol Rev.* 8(2), 135–147 (2005)

Vissandjée, B., Dupéré, S.: Culture, migration et instrument de mesure : défis incontournables. *Journal of International Migration and Integration.* 1(4), 477–492 (2000)

Wang, W. L., Lee, H. L., Fetzer, S. J.: Challenges and strategies of instrument translation. *West J Nurs Res.* 28(3), 310–321 (2006)

Ward, V. L., House, A. O., Hamer, S.: Knowledge brokering: exploring the process of transferring knowledge into action. *BMC Health Serv Res.* 9(1), 12 (2009)

Weeks, A., Swerissen, H., Belfrage, J.: Issues, challenges, and solutions in translating study instruments. *Eval Rev.* 31(2), 153–165 (2007)

Weidmer, B., Hurtado, M., Weech-Maldonado, R., Ngo-Metzger, Q., Bogen, K.: Guidelines for translating CAHPS surveys. Agency for Healthcare Research and Quality. www.ahrq.gov (2006). Accessed April 2016

White, H. D.: Scientific communication and literature retrieval. In: Cooper, H., Hedges, L. V., Valentine, J. C. (eds.) *The handbook of research synthesis and meta-analysis*, Second Edition, pp. 51–71. Russell Sage Foundation, New York (2009)

Willis, G., Brick, P. D., Norberg, A., Stark, D. S., Kudela, M. S., Forsyth, B. H., . . . Lawrence, D.: Evaluation of a multi-step survey translation process. Paper presented at the Proceedings of the International Conference on Survey Methods in Multinational, Multiregional, and Multicultural Contexts (3MC), Berlin (2008)

Winer, B.: *Statistical principles in experimental design*. McGraw-Hill, New York (1962)

World Health Organization: Process of translation and adaptation of instruments. www.who.int (2009). Accessed April 2016

3 – Article II

Comment décrire le processus d'adaptation effectué : Proposition d'une typologie et analyse de 131 descriptions publiées au Québec entre 2000 à 2014

François RENÉ DE COTRET¹

Yvan LEANZA, Ph.D.¹

Aline FRANCOEUR, Ph.D.²

Andrée-Anne BEAUDOIN-JULIEN¹

¹École de psychologie

²Département des langues, de linguistique et de traduction

Université Laval

Québec, Canada

Résumé :

Adapter d'un test est un processus qui implique une série de stratégies visant la transposition d'une culture/langue à une autre. De tous les tests adaptés chaque année, seule une fraction donne lieu à des publications scientifiques. L'objectif de l'étude était de déterminer à l'aide d'une nouvelle catégorisation comment ces processus d'adaptation sont décrits dans la littérature et ce qui est mis de l'avant pour légitimer le processus d'adaptation. Une recension a été effectuée et 131 descriptions des processus d'adaptation réalisés au Québec entre 2000 et 2014 ont été identifiées. Ces descriptions ont ensuite été analysées avec la Grille d'évaluation de la description du processus d'adaptation (G-PA). Les deux tiers des 131 descriptions à l'étude font référence à peu ou pas de stratégies faisant consensus dans la littérature et comptent entre un et 250 mots. En comparaison, certaines descriptions font référence à plusieurs stratégies d'adaptation, leur longueur moyenne est de plus de 1 500 mots et elles comprennent des compléments théoriques et des références à des publications à propos de l'adaptation des tests. La G-PA permet d'avoir un portrait détaillé des différents types de descriptions effectuées au Québec. Elle met également en lumière trois types de descriptions – libres, positivistes et intégratives – utilisées par les chercheurs pour légitimer les tests qu'ils ont adaptés.

Mots-clés : traduction inversée, Vallerand, ITC, Hambleton, recherche interculturelle

**How to describe the adaptation process:
A typology proposal and review of adaptation practices
in Quebec from 2000 to 2014**

François RENÉ DE COTRET¹

Aline FRANCOEUR, Ph.D.²

Yvan LEANZA, Ph.D.¹

¹School of Psychology

²Department of Languages, Linguistic, and Translation

Laval University

Québec, Canada

Abstract:

Adapting a test is a process that entails a series of strategies aimed at transposing it from one culture/language to another. Of all the tests being adapted every year, only a fraction gives rise to scientific publications, many of which address in a few words only the adaptation strategies that were used. **Objective:** The objective of the study was to determine with a new categorisation how these adaptation processes are described in the literature and what is put forward in these descriptions to legitimize the adaptation process.

Method: A literature review was conducted, and 131 descriptions of adaptation processes realized in Quebec between 2000 and 2014 were identified. They were then organized according to the adaptation strategies that create consensus among researchers (René de Cotret, Leanza, Francoeur & Beaudoin-Julien, 2019), and which are the basis of the categorisation proposed, the Adaptation Process description assessment grid (AP-G).

Results: Two thirds of the 131 descriptions under study refers to no or few adaptation strategies and are between one and 250 words in length. In comparison, some descriptions refer to several adaptation strategies, their average length is more than 1,500 words, and they include theoretical developments and references to publications about test adaptation.

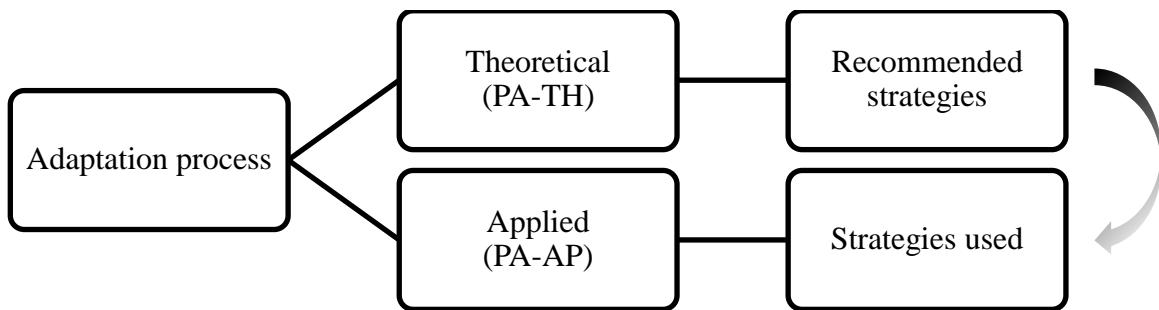
Discussion: The AP-G allows for a complete picture of the different kinds of descriptions found in the literature. It also sheds light on three types of descriptions—free, positivist, and integrative—used by researchers to legitimize the tests they have adapted. The AP-G is an easy-to-use tool for researchers to review the description of some adaptation processes performed. The grid can also be used by publishers and universities to establish strict criteria for such descriptions: what is acceptable and what is not?

Keywords: back-translation, Vallerand, ITC, Hambleton, cross-cultural research

Adapting a psychometric test⁵ consists in transposing it from the language/culture of a population for which it was originally created to the language/culture of another population. It corresponds to a process, since the passage from the original version (OV) to the adapted version (AV) involves a series of translated versions (TVs), which are then reviewed and adjusted. In fact, it is a scientific process, in that it aims at producing knowledge: it confirms the equivalence between the TV and the OV, which in turn allows the TV to gain the status of AV.

As explained by René de Cotret *et al.* (2019), this process can be addressed in theoretical or in concrete terms. In the first case, strategies are recommended for adapting tests; in the second, actual strategies are applied to adapt a specific test, and they may or may not be inspired by recommended strategies. The abbreviations PA-TH (process of adaptation—theoretical) and PA-AP (process of adaptation—applied) are proposed to refer to these two cases. Figure 1 shows the potential influence of PA-THs on PA-APs via recommended strategies.

Figure 1. Two ways of addressing the adaptation process



PA-APs are carried out in various contexts. Some are headed by international agencies with enormous resources such as the Organisation for Economic Cooperation and Development (OECD), which manages the PISA programme, and the International Association for the Evaluation of Educational Achievement (IEA), in charge of the Trends in International

⁵ We use the term *test* in a broad sense to include any psychometric instrument as defined by the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). As for *construct*, it is an umbrella term encompassing what is being measured by tests: constructs, aptitudes, etc.

Mathematics and Science Study (TIMSS). Others are performed on a small scale by research groups or individuals with more limited means.

The latter situation complicates research on PA-APs. First, a significant number of small-scale PA-APs fall into the “homemade translations” category and go unnoticed.

Secondly, PA-APs discussed in academic publications are not necessarily described in detail, making it hard for readers to figure out how the adaptation process was actually conducted (Brislin, 1970; Maneesriwongul & Dixon, 2004; Prieto, 1992; Rios & Sireci, 2014).

As noted by van Widenfelt, Treffers, De Beurs, Siebelink, and Koudijs (2005, p. 145), “[t]hough in psychology journals there is generally a uniform way to present a method section, when presenting information on the translation and adaptation of an instrument there is no uniformity.” Guillemin, Bombardier, and Beaton (1993, p. 1428), in their literature review on cross-cultural adaptation of health-related quality of life measures, further comment: “[T]he replicability of the adapted instrument, i.e. the clarity of the presentation and the thoroughness of the directions provided for its use [...] was addressed in none of the papers.”

Thirdly, strategies used for adapting tests differ tremendously (Acquadro, Conway, Hareendran, & Aaronson, 2008; Guillemin *et al.*, 1993; Maneesriwongul & Dixon, 2004; Vallerand, 1989).

How to study PA-APs. To study PA-APs, researchers might conduct a huge field study, asking about what was done exactly to those primarily concerned, the researchers responsible to lead each PA-AP under study. It is, however, possible to focus on how researchers describe, in a paper, how they conducted their PA-AP. These descriptions, as van Widenfelt *et al.* (2005) posited, are part of a scientific inquiry and should be address in the scientific literature.

Three teams of researchers elaborated ways to categorise PA-APs (Guillemin *et al.*, 1993; Maneesriwongul & Dixon, 2004; Rios & Sireci, 2014). They all consider their method

limited by the fact that it does not assess PA-APs *per se* but their description, often laconic or even missing, in the literature.

These categorisations are normative as they aim to determine which adaptation strategies are best suited to the adaptation process. However, the theoretical back-ground of these normative initiatives is not based on scientific data, which may lead to inconsistencies. As example, even though Maneesriwongul and Dixon (2004)'s categorisation implies that back-translation is an effective strategy, both the ODCE and the IEA use PA-APs that do not involve it (Arffman, 2013; Maxwell, 1996). As a matter of fact, Epstein, Santo, and Guillemin concluded in 2015 that it was not possible to clearly establish which adaptation strategies, among those recommended in the literature, are more efficient than others, because "the lack of evidence limits the possibilities of clear-cut recommendations" (p. 439).

Yet, on the basis of analyzing 34 PA-THs, René de Cotret *et al.* (2019) uncovered a set of adaptation strategies on which a majority of experts or teams of experts agree. They consist in 17 consensual strategies, grouped into the five phases of the adaptation process: Initial Translation, Internal Review, External Review, Testing, and the Supra Phase. Based on content analyses, these consensual strategies provide a scientific frame of reference on which a new categorisation was elaborated, the Adaptation process description assessment grid (AP-G) (Appendices C and D).

The goal of the study was twofold: 1) to determine how PA-APs are described in the literature; 2) to determine how authors of PA-AP descriptions legitimize tests that they have adapted.

Method

A literature review was conducted in May and June 2014 to find scientific texts (journal articles and academic publications such as PhD and MA theses) dealing with English-to-French adaptation of psychometric tests for the French-speaking population of Quebec, Canada.

The texts included in the review had to: (a) contain the terms “adaptation” and/or “traduction/translation” in their abstract, method section, or appendices; (b) refer directly (in the abstract or objective) or indirectly (in the body of the text, without previous mention) to a test that had been adapted from English to French; (c) and have been published since 2000. Some texts meeting those criteria were excluded, for two reasons: they did not refer to the translation/adaptation of psychometric tests (the term “translation” was used in another sense, for instance to refer to the way a psychological problem *translates* into symptoms) or they dealt exclusively with the testing of the TV.

Journal articles were found through PsycNet, PubMed, Psychology and Behaviorial Sciences Collection, and Web of Science with the following keywords: translation OR traduction OR adaptation OR validation AND French OR français AND Quebec. As for academic publications, they were found via the online library catalogue of each of the nine Quebec universities with graduate programs in psychology, with the help of a librarian in each case.

The AP-G was then used to organize and assess the descriptions found through the literature review. The AP-G development steps are provided in Appendix A.

Results

A total of 698 texts were found, among which 284 were journal articles and 414 academic publications. After reading the abstracts and/or the whole texts, 131 texts were kept: 52 articles and 79 academic publications. In 90% of the cases, these texts discuss self-reported psychometric tests. The other 10% covers structured or semi-structured tests, or tests completed by a third party.

The articles come from 28 journals (Table 1), with one—the *Canadian Journal of Behavioural Science* (16 articles)—being particularly well represented. As for the academic publications (Table 2), they come from eight of the nine Quebec universities with a graduate program in psychology.⁶

⁶ None comes from the Université du Québec en Outaouais.

Table 1

Journal articles by alphabetical order of authors, and journals in which they were published

Author(s) (year of publication)	Journal
Beaupré <i>et al.</i> (2002)	Canadian Journal of Psychiatry
Béland <i>et al.</i> (2012)	L'Encéphale
Birmes <i>et al.</i> (2005)	European Psychiatry
Brown <i>et al.</i> (2001)	Canadian Family Physician
Brunet <i>et al.</i> (2003)	Canadian Journal of Psychiatry
Cordeau and Dubé (2008)	Les cahiers internationaux de psychologie sociale
Côté <i>et al.</i> (2013)	Journal de thérapie comportementale et cognitive
Coutu <i>et al.</i> (2002)	Canadian Journal of Behavioural Science
Couture <i>et al.</i> (2010)	European Eating Disorders Review
Daigneault <i>et al.</i> (2013)	Child Abuse & Neglect
De Benedictis <i>et al.</i> (2012)	Journal of Psychiatric and Mental Health Nursing
Demers <i>et al.</i> (2002)	Quality of Life Research
Deshaises <i>et al.</i> (2009)	Canadian Journal of Behavioural Science
DiTommaso <i>et al.</i> (2007)	Social Behavior and Personality
Fikretoglu <i>et al.</i> (2006)	Canadian Journal of Psychiatry
Filiatrault <i>et al.</i> (2014)	Disability and Rehabilitation
Forget and Robichaud (2002)	Canadian Journal of Behavioural Science
French <i>et al.</i> (2005)	Canadian Journal of Behavioural Science
Giguère <i>et al.</i> (2006)	Canadian Journal of Behavioural Science
Grégoire <i>et al.</i> (2013)	Canadian Journal of Behavioural Science
Hébert <i>et al.</i> (2007)	Measurement and Evaluation in Counseling and Dev.
Husky <i>et al.</i> (2004)	Canadian Journal of Psychiatry
Julien <i>et al.</i> (2008)	Journal of Anxiety Disorders
Koski <i>et al.</i> (2010)	Dement Geriatric Cognitive Disorders
Lachapelle <i>et al.</i> (2000)	Revue francophone de la déficience intellectuelle
Lafontaine and Lussier (2003)	Canadian Journal of Behavioural Science

Laliberté and Earls (2007)	Revue québécoise de psychologie
Lauzier (2010)	Revue québécoise de psychologie
Lauzier and Haccoun (2010)	Journal of Behavioural Science
Lauzier <i>et al.</i> (2013)	Revue européenne de psychologie appliquée
Laverdière <i>et al.</i> (2010)	European Journal of Psychological Assessment
Leblanc <i>et al.</i> (2007)	Revue québécoise de psychologie
Leboeuf and Losier (2012)	Revue européenne de psychologie appliquée
Lecavalier and Tassé (2001)	Revue francophone de la déficience intellectuelle
Lemelin <i>et al.</i> (2007)	Canadian Journal of Behavioural Science
Lepage <i>et al.</i> (2009)	Canadian Journal of Behavioural Science
Martin and Gosselin (2012)	Canadian Journal of Behavioural Science
Méthot and Morin (2004)	Revue francophone de la déficience intellectuelle
O'Rourke and Cappeliez (2003)	Canadian Journal of Behavioural Science
Roberge <i>et al.</i> (2003)	Canadian Journal of Behavioural Science
Paquette <i>et al.</i> (2004)	Santé mentale au Québec
Plante <i>et al.</i> (2012)	Behavioral Sciences and the Law
Poirier <i>et al.</i> (2004)	Journal of Nervous and Mental Disease
Poulin and Desrosiers (2010)	Canadian Journal of Occupational Therapy
Radomsky <i>et al.</i> (2006)	Canadian Journal of Behavioural Science
Ricard-St-Aubin <i>et al.</i> (2010)	Revue européenne de psychologie appliquée
Richer <i>et al.</i> (2012)	Accident Analysis and Prevention
Seidah <i>et al.</i> (2002)	Canadian Journal of Behavioural Science
Bouffard <i>et al.</i> (2002)	Canadian Journal of Behavioural Science
Trottier <i>et al.</i> (2008)	Canadian Journal of Behavioural Science
Turgeon and Chartrand (2003)	Psychological Assessment
Turgeon <i>et al.</i> (2013)	Journal de thérapie comportementale et cognitive

Table 2

*Academic publications (PhD and MA theses and/or dissertations)
by alphabetical order of author, and academic affiliation*

Author (year of publication)	Academic affiliation
Achim (2006)	Université de Montréal
Aimé (2002)	Université Laval
Audet-Lapointe (2004)	Université de Montréal
Beaulieu (2012)	Université de Montréal
Bélanger (2011)	Université Laval
Belleville (2006)	Université Laval
Bergeron (2006)	Université de Sherbrooke
Boislard Pépin (2010)	Université du Québec à Montréal
Boisseau (2010)*	Université du Québec à Montréal
Boisvert (2006)	Université du Québec à Chicoutimi
Bouchard (2007)	Université du Québec à Chicoutimi
A. Bouffard (2010)	Université Laval
Bougie (2005)	Université McGill
Brassard (2003)	Université du Québec à Chicoutimi
Brin (2011)	Université Laval
Brousseau (2011)	Université du Québec à Montréal
Charbonneau (2009)	Université du Québec à Montréal
Chevrier (2009)	Université du Québec à Montréal
Cordeau (2007)	Université du Québec à Trois-Rivières
Cossette (2008)	Université du Québec à Montréal
A. Couture (2014)	Université Laval
Da Silva Luis (2014)	Université Laval
Daley (2007)	Université Laval
D'Amico (2010)	Université du Québec à Montréal

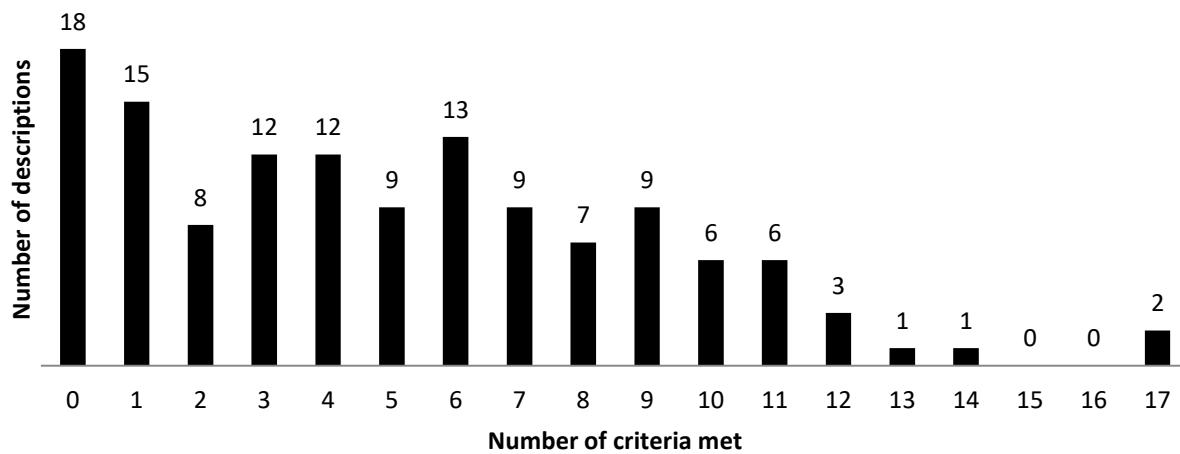
Daniel (2008)	Université du Québec à Montréal
J. Delisle (2011)	Université du Québec à Montréal
Denis (2009)	Université du Québec à Montréal
Dubé (2011)	Université de Sherbrooke
Emard (2010)	Université du Québec à Montréal
Engelberg (2003)	Université McGill
Finn (2012)	Université du Québec à Montréal
Fortin (2009)	Université de Montréal
French Bourgeois (2011)	Université de Montréal
Gagné (2010)	Université de Montréal
C. Gagnon (2013)	Université Laval
M. Gagnon (2010)	Université du Québec à Chicoutimi
V. Gagnon (2004)	Université du Québec à Chicoutimi
Gilbert (2011)	Université Laval
Grimard (2011)*	Université Laval
Hovington (2002)	Université du Québec à Chicoutimi
Johnson (2008)	Université de Montréal
Lacasse (2004)	Université McGill
Lafrenière (2013)	Université du Québec à Montréal
Lamoureux-Hébert (2006)	Université du Québec à Montréal
Landry (2010)	Université du Québec à Montréal
L'Archevêque (2009)	Université du Québec à Montréal
Laverdière (2009)	Université Laval
Lavergne (2003)	Université Laval
Lavigne (2011)	Université du Québec à Montréal
Lekes (2012)	Université McGill
Lemieux (2009)	Université Laval
Lessard (2012)	Université du Québec à Montréal

Mayer-Renaud (2008)	Université de Montréal
Ménard (2008)	Université de Montréal
Morin (2007)	Université Laval
Orendain Tremear (2010)	Université Laval
Paradis (2012)	Université du Québec à Montréal
Payer (2012)	Université de Montréal
Pelletier (2007)	Université Laval
Perozzo (2007)	Université de Montréal
Plante-Jean (2012)	Université Laval
Potvin (2011)	Université du Québec à Montréal
Poulin-Vallières (2004)	Université du Québec à Chicoutimi
Quesnel (2007)	Université Laval
Rabiau (2006)	Université McGill
M. Roberge (2007)	Université du Québec à Montréal
Sergerie (2008)	Université du Québec à Montréal
Simard (2008)	Université Laval
Taillon (2011)	Université du Québec à Montréal
Temimi (2011)	Université de Montréal
Thibaudeau (2006)	Université du Québec à Montréal
Thibault (2008)	Université du Québec à Montréal
Toledano (2006)	Université de Concordia
C. Tremblay (2001)	Université de Montréal
S. Tremblay (2009)	Université de Montréal
Turcotte (2013)	Université du Québec à Chicoutimi
Vaillancourt (2009)	Université du Québec à Montréal

* Two distinct PA-APs described in the same publication.

AP-G criteria. As shown in Figure 2, two thirds of the descriptions under study meet six or less criteria, and 18 descriptions meet none. In comparison, four descriptions (3%) meet more than half of the 25 criteria. Typical examples of descriptions according to the number of criteria they meet are provided in Appendix B.

Figure 2. Number of descriptions according to the number of AP-G criteria they meet



Regrouping the descriptions according to the number of criteria they meet allowed us to establish the percentage of criteria mentioned in each of the groups. Because of their limited number, descriptions that meet 13 criteria or more were combined into one group. Table 3 illustrates the variation of each criterion by group of descriptions.

Table 3

Percentage (in decimal number) of descriptions per criteria met, by group of descriptions

Phase	Criteria	Descriptions grouped according to the total number of criteria that are met													
		0	1	2	3	4	5	6	7	8	9	10	11	12	13+
I	More than one translator (translation team)	0	0	0	.17	.25	.67	.62	.56	.71	.78	1	1	1	1
	At least one language expert	0	0	.13	.25	.17	.11	.15	0	.14	.33	.50	.33	.33	.50
	Good command of the SL and TL	0	0	0	.08	0	.11	.15	.11	.43	.11	.50	.33	.67	1
	At least one translator familiar with the construct	0	.19	.13	.17	.33	.67	.46	.89	.86	.78	.67	1	1	1
	At least one translator familiar with test development	0	0	0	.17	.17	.67	.46	.89	.57	.78	.33	.83	.33	.75
	At least one translator with experience in both cultures	0	0	0	0	0	0	0	0	0	0	.17	0	0	0
	Production of two or more TVs	0	.06	0	.17	.17	.33	.38	.22	.29	.56	.83	.83	1	.75
II	Provision of translation guidelines	0	0	0	0	0	0	0	0	0	0	0	0	.33	.25
	Translation of all the components	0	0	0	0	0	0	0	.11	.29	.33	.33	.50	.67	.75
	More than one reviewer (review committee)	0	0	.25	.33	.42	.22	.69	.89	.71	.78	.83	1	1	1
	At least one language expert	0	.13	0	0	0	0	.23	0	.29	.11	.17	.50	.67	0
	Mastery of the construct/concept by at least one reviewer	0	0	.13	.25	.50	.22	.69	.89	.71	.78	.83	1	1	1
	At least one reviewer with expertise in test development	0	0	0	0	.08	0	.23	0	.14	.22	.17	0	0	.25
	At least one reviewer with experience in both cultures	0	0	0	0	0	0	0	0	0	0	.17	0	0	0
III	Use of back-translation as an overseeing procedure	0	.47	.38	.58	.83	.78	.77	.78	.86	.89	.67	.67	.67	1
	Use of at least another overseeing procedure	0	0	0	0	0	0	0	0	0	.22	.17	.50	.67	.75
	External review was performed by the review committee	0	0	0	0	0	0	0	0	0	.22	.17	.33	0	.50
	Individual debriefing interviews	0	0	0	0	0	0	0	.11	0	.11	0	0	0	.50
	Written evaluations of the TV	0	0	0	0	0	.11	.23	.11	.14	.22	.50	.83	0	.25
	Focus groups	0	0	0	0	0	0	0	.11	.14	0	0	0	0	.25
	Other overseeing procedures	0	0	0	.08	.25	0	.07	.11	.14	.44	.50	0	.67	.25
IV	Validation of the TV via statistical analysis	0	.20	.50	.67	.67	.89	.84	.78	.88	1	1	1	1	1
	Changes to the TV following statistical analysis	0	0	0	0	0	0	0	.11	.14	.11	0	.14	.33	.75
SP	Account of translation/adaptation problems	0	0	.13	0	0	.11	0	.22	.29	.11	.33	.33	.67	.75
	Provision of the OV and AV	0	0	.25	.08	.17	.11	0	.11	.14	.11	.17	0	0	1

Note. Percentages of 50% and more are in bold to facilitate comparisons.

Phase 1 – Initial Translation. As shown in Table 3, in most descriptions that satisfy five or more criteria, three criteria related to the Initial Translation Phase are met: More than one translator, Being familiar with the construct, and Being familiar with test development. In comparison, five criteria are more rarely satisfied —Good command of the SL and TL, At least one language expert, Production of two or more TVs, Provision of translation guidelines, and Translation of all the components of the test. When these criteria are satisfied, it is in descriptions with 10 or more criteria. As for the criterion Experience in the source and target cultures, it is very rarely met.

Phase 2 – Internal Review. Use of back-translation as an overseeing procedure is a dominant criterion: it is met by a majority of descriptions containing three or more criteria. Contrastingly, only descriptions meeting 11 or more criteria mention At least another overseeing procedure in a majority of the cases. As a result, only seven out of the 79 descriptions where back-translation is mentioned refer to another overseeing procedure.

As in Phase I, the Experience in the source and target cultures criterion is rarely met here. The criterion At least one language expert is met in a majority of descriptions containing 11 or 12 criteria, but in none of those with 13 or more criteria. Being familiar with test development is mentioned in a minority of descriptions but reaches 25% in the descriptions with 13 or more criteria. Finally, the criteria More than one reviewer and Mastery of the construct are met in a majority of descriptions with six or more criteria.

Phase 3 – External Review. The criteria related to this phase are either absent or marginal in the descriptions that meet eight or fewer criteria. However, a majority of descriptions containing nine or more criteria mentions different forms of overseeing procedures for the external review. The criteria Clear mention that the external review was performed by the review committee and Individual debriefing interviews are either not met or rarely met in most of the descriptions, except for those with 13 or more criteria, in which they appear in a majority.

Phase 4 – Testing. Testing is mentioned in the majority of the descriptions, except for those containing no or only one criteria. However, in only one group—13 or more criteria—does the criterion Changes to the TV following statistical analysis appear in a majority of the descriptions.

Supra Phase – Documentation. The two criteria of the Supra Phase are mentioned in a majority of descriptions only in the case of those that satisfy the greater number of criteria.

Numeric variables. Table 4 proposes an overview of the results with regard to the three numeric variables, by group of descriptions.

Table 4

Values of the three numeric variables, by group of descriptions

Group	Description length		Theoretical development			References	
	Average length	Extent	Number	Average length	Extent	Average number	Extent
0	5	1 – 27	0	-	-	0.05	0 - 1
1	25	3 – 106	0	-	-	0.40	0 - 1
2	47	19 – 83	0	-	-	0.38	0 - 1
3	42	11 – 113	0	-	-	0.17	0 - 1
4	69	17 – 157	0	-	-	1.00	0 - 3
5	82	29 – 174	0	-	-	0.33	0 - 1
6	116	49 – 210	2	607	66 – 1 147	1.15	0 - 6
7	148	76 – 271	0	-	-	0.88	0 - 1
8	201	102 – 368	0	-	-	0.57	0 - 1
9	305	64 – 865	0	-	-	1.44	0 - 7
10	261	170 – 430	2	1 451	125 – 2 777	2.33	0 - 8
11	372	168 – 611	3	1 390	42 – 3 767	5.67	2 - 11
12	365	328 – 408	1	287	-	3.00	1 - 5
13+	1 560	406 – 3 035	3	1 333	249 – 3 016	3.25	2 - 5

Note. Average length and extent is a number of word(s) or reference(s).

Description Length. Descriptions that meet none of the dichotomic criteria of the assessment grid have an average length of five words. In comparison, the average length of descriptions that meet 13 or more criteria is over 1,500 words (many paragraphs). There is a high correlation between the number of criteria met in a description and the length of the description ($r(131) = .65, p < .001$).

Additionally, there are significant intergroup variations with regard to length (see Table 4). The lowest threshold is 170 words in descriptions that meet ten criteria, and it goes up to 406 words in descriptions that meet 13 or more criteria. One description clearly stands out in this latter group, with more than 3,000 words.

Theoretical Development. Descriptions meeting less than ten criteria rarely include theoretical developments, whereas almost half (47%) of the descriptions meeting ten or more criteria do so.

In the case of descriptions meeting 13 or more criteria, 75% include some theoretical development.

The value of these developments varies considerably and correlates poorly with the number of criteria that are satisfied ($r(131) = .32$, $p < .001$). This being said, if we consider only the presence or absence of theoretical developments, the correlation is moderate ($r(131) = .49$, $p < .001$).

References. The number of references to the adaptation process and the presence of theoretical developments correlate moderately with the number of dichotomic criteria met ($r(131) = .54$, $p < .001$). The correlation remains the same if we consider the presence or absence of references in the description only ($r(131) = .54$, $p < .001$). Only one description that includes some theoretical development has no references.

Half of the texts (51%) in the corpus contain references to the adaptation process. These 67 texts contain a total of 136 citations from 49 references (extent = 1-11 ; average = 2). The most cited authors/groups of authors are Vallerand (54x)⁷, the ITC (23x)⁸, Brislin (18x)⁹, Geisinger (5x)¹⁰, Behling and Law (2000) (4x), Spielberger and Sharma (1976) (3x), Tassé and Craig (1999) (3x), Arnold and Matus (2000) (2x), Van Widenfelt *et al.* (2005) (2x), and Sperber, Devellis, and Boehlecke (1994) (2x). They represent 117 of the 136 citations (86%).

References to Vallerand span the whole corpus. We even found a description that met none of the dichotomic criteria but still mentions Vallerand (1989). In comparison, references to the ITC are found 95% of the time in descriptions that contain nine or more criteria. No particular pattern was noticed with regard to the other most cited authors/groups of authors.

⁷ Vallerand (1989) (52x), and Vallerand and Halliwell (1983) (2x).

⁸ van de Vijver and Hambleton (1996) (4x), Hambleton and de Jong (2003) (3x), Hambleton and Kanjee (1995) (3x), Hambleton (2001) (3x), Jeanrie and Bertrand (1999) (2x), Hambleton (1993) (1x), Hambleton (1994) (1x), van de Vijver and Leung (1997) (1x), Hambleton and Patsula (1999) (1x), Hambleton and Patsula (2000) (1x), Hambleton (2005) (1x), Sireci, Patsula, and Hambleton (2005) (1x), and van de Vijver and Poortinga (2005) (1x).

⁹ Brislin (1970) (8x), Brislin, Lonner, and Thorndike (1973) (5x), Brislin (1986) (2x), Brislin (1976) (1x), Brislin (1980) (1x), and Sinaiko and Brislin (1973) (1x).

¹⁰ Geisinger (1994) (4x), and Geisinger (2003) (1x).

Types of Text and Ways of Mentioning PA-APs. In 57% of the texts under study, PA-APs are mentioned directly. It is the case in 51 of the 52 journal articles (98%), and in 24 of the 79 academic publications (30%). Table 5 summarizes the results.

Table 5

Fractions and percentages (in decimal numbers) of journal articles with direct mention of PA-APs, by group of descriptions

Group of descriptions	Journal articles	Direct reference
0	0/18 (0)	0/18 (0)
1	3/15 (.20)	3/15 (.20)
2	5/8 (.63)	5/8 (.63)
3	4/12 (.33)	6/12 (.50)
4	3/12 (.25)	5/12 (.42)
5	5/9 (.56)	6/9 (.67)
6	7/13 (.54)	9/13 (.69)
7	3/9 (.33)	7/9 (.78)
8	5/7 (.71)	6/7 (.88)
9	6/9 (.67)	9/9 (1)
10	4/6 (.67)	6/6 (1)
11	4/6 (.67)	6/6 (1)
12	2/3 (.67)	3/3 (1)
13+	1/4 (.25)	3/4 (.75)

Descriptions that contain no criteria or one criterion come exclusively from academic publications in which PA-APs are mentioned indirectly. However, the type of text correlates poorly with the number of criteria met ($r(131) = .32, p < .001$). By contrast, there is a strong correlation between the way PA-APs is mentioned (directly or indirectly) and the number of criteria met ($r(131) = .60, p < .001$).

Neglected strategies. Two criteria from the assessment grid were met by only one of the 131 descriptions under study. They are criteria #6 and #14, both related to experience in the source and target cultures (in the case of the translators in #6, and the reviewers in #14). These criteria are particularly relevant in the context of adapting tests for cross-cultural comparisons and/or when the OV is from a source language/culture that bears little resemblance to the target language/culture. In the context of the study, though, most of the 131 descriptions, including the one which satisfies these two criteria, are about tests that are not meant to be used for cross-cultural comparisons and originate from the United States.

Also, the criterion related to translation guidelines is seldom mentioned: only two descriptions, both meeting 12 or more criteria, mention it (see Arffman 2013 with regard to the relevance of such guidelines).

Limitations. One possible limitation is the fact that some of the descriptions under study come from the same author/co-author: Lauzier (three), Boisseau (two), Cordeau (two), Demers (two), Grimard (two) and Laverdière (two). It was decided to keep them because they do not refer to the same PA-AP and also because they do not meet the same number of criteria.

The number of criteria met should be seen as an indication, and not in absolute terms. One could imagine two descriptions, each meeting four criteria, although in one case all the criteria would refer to the first phase of the adaptation process, whereas in the second case each criterion would refer to a different phase of the process.

Our assessment grid includes only two criteria relative to the overseeing of the Internal Review, compared to four for the External Review. However, this has no impact on the results, since none of the 131 descriptions refers to more than one overseeing method (other back-translation) for the internal review. Moreover, this is due to theoretical limitations that are proper to the Internal Review, such as the vagueness of the concept of equivalence (see Herdman *et al.*, 1997). Such limitations are not present in the context of the External Review, where the overseeing procedures differ mainly in terms of the mode of communication between the assessor and the subject.

As explained in the Method section, the corpus of 131 descriptions was used to create the AP-G, and then assessed with the grid, something that could be seen as a limitation. Yet, the 25 criteria of the assessment grid are primarily based on test adaptation strategies that reach a consensus among researchers, as established through an analysis of 34 PATHs, most of which were published outside Quebec and Canada (René de Cotret *et al.*, 2019). These strategies were operationalized through the use of the 131 descriptions; no new strategies were established on the basis of the 131 descriptions. Additionally, from the number ($n=131$) and variety of the descriptions in our corpus, it is believed that the strategies they contain are representative of typical strategies from other descriptions of PA-APs.

Discussion

The AP-G was applied to a set of 131 PA-AP descriptions. Compared to the other PA-AP categorisations (Guillemin *et al.*, 1993; Maneesriwongul & Dixon, 2004; Rios & Sireci, 2014), the AP-G can be applied to short descriptions such as those frequently found in the literature. Short descriptions of PA-APs imply that adapted versions of psychometric tests are actually being used in the same way as the corresponding original versions, even though no information—or barely any—is provided to legitimize their validity. It was important to develop a tool that could be used with such descriptions, so that they could be taken into consideration.

As a matter of fact, two thirds (88) of the 131 PA-AP descriptions studied meet a maximum of six criteria. Of these 88 descriptions, 18 contain less than five words and meet none of the assessment criteria, meaning that it is either impossible or very difficult to know how the adaptation process was done in these particular cases.

By contrast, 3% of the descriptions meet more than half of the 25 criteria. Descriptions that satisfy the greater number of criteria tend to: (a) refer to more phases of the adaptation process; (b) contain more words; (c) mention PA-APs directly; (d) include theoretical developments; (e) refer to authors/groups of authors who wrote about test adaptation, (f) among which is the ITC. Even if only 6% of the texts under review refer to the ITC, this result counterbalances Rios and Sireci (2014), who found no reference to the ITC in their literature review.

Basically half (49%) of the 131 descriptions contain no reference related to test adaptation. Describing a PA-AP without citing any author or group of authors who wrote about test adaptation gives the impression that there is no theory on this topic or that, if there is, it is not necessary to rely on it when adapting a test.

Of the 25 assessment criteria, only two are met by a majority of descriptions: Use of back-translation as an overseeing procedure (Phase II) and Validation of the TV via accepted statistical analysis methods (Phase IV). Only 5% of the descriptions (mostly with ten criteria or more) mention another overseeing procedure in Phase II. Yet, since the end of the 1990s, a majority of experts have recommended not using back-translation alone (e.g. Hambleton &

Patsula, 1999; Harkness, 2011; Pan & De La Puente, 2005; Solano-Flores *et al.*, 2009; Tassé & Craig, 1999; Van Widenfelt *et al.*, 2005).

With regard to Phase IV (Testing), only 5% of the descriptions (mostly with 13 criteria or more) state that changes were made to the TV following statistical analysis. This result is in line with René de Cotret *et al.* (2019)'s argument that testing should be distinguished from the first three phases of the adaptation process. While these phases involve combining various forms of expertise and reviewing procedures towards a single end, that of translating and adapting the OV, the fourth phase—testing—essentially aims at bringing the TV to the same level of objectivity as the OV. Theoretically, the testing phase also allows for identifying problematic items and making changes to the TV accordingly. However, our results show that such changes are rarely mentioned in PA-AP descriptions.

Lastly, as shown in Table 5, descriptions that meet the greater number of criteria were found in journal articles and academic publications. Two of the four descriptions with 13 or more criteria came from theses dealing specifically with PA-APs. These theses (Bergeron, 2006; Dubé, 2011) are proof that PA-APs can be the focus of major research projects realized over many years and detailed into 200 pages or so.

Three types of descriptions to legitimize PA-APs. The main goal of our research was to determine, from the study of PA-AP descriptions, how researchers legitimize their own adapted versions of tests. From the 131 descriptions analyzed in Study 2, a convenient typology of three types of descriptions came up, the Typology of Adaptation Process Descriptions, which regroups three types of description: free, positivist, and integrative. By contrast with free descriptions, positivist and integrative descriptions are meant to convince readers that both versions of a test are equivalent; in that way, they are scientific demonstrations. The three types of descriptions do not directly match the number of criteria met, although this number remains relevant.

Free descriptions. Some authors describe PA-APs in a minimalist, elusive way, leaving readers in the dark as to the legitimacy of the process, and consequently of the AV. Readers have no other option, in such cases, than to trust the author of the publication. Because expressions such as *free translation* and *homemade translation* are used by some of these authors to refer to their PA-AP, the term *free descriptions* was chosen. As a general rule, free descriptions do not meet

any of the criteria from the assessment grid, suggesting that none of the adaptation strategies recommended in the literature were taken into account.

Free descriptions are not meant to demonstrate that the AV of a test is equivalent to its OV. Instead, they suggest that the translation is good *a priori*. Besides, there is no need in such descriptions to persuade the readers of the equivalence between the two versions, because it is implied that readers are already convinced. The resources put into the adaptation process are considerably reduced in such context. Results show that free descriptions were found only in academic publications. PhD and MA theses or dissertations are not usually directly concerned with the adaptation process, but the results they present depend on it; in other words, in such cases, the adaptation process is a means, not an end.

This being said, expressions such as *free translation* or *homemade translation* do not relate well with the adaptation process. There is nothing wrong with the fact that free translation can be applied to an expression or a sentence. However, conducting an adaptation process has nothing to do with translating an expression. When one speaks of the adaptation process as a free translation, it implies a shift from a scientific approach to a mathematical operation, a conversion or, more precisely, a geometric translation which can be done by any *free* translator. Translation would thus be a matter of mere transliteration. Frontiers between languages would be blurred, allowing for a *free* passage from a language to another. In fact, it would be inappropriate to call it a science approach, since it would be a question of techniques, without any theoretical background. It would fit the absolutist world described by Berry, Poortinga, Breugelmans, Chasiotis and Sam (2011), a hypothetical world in which culture would have only a negligible influence—or none whatsoever—on behaviours.

Positivist descriptions. Some authors describe PA-APs in terms of quantitative data, thus the expression *positivist descriptions*. Knowledge would originate from *measuring* the phenomenon under study, which, in turn, would transform into a positivist data. Such descriptions mainly consist in providing figures to the readers in order to convince them that the adaptation process is legitimate. Few words are needed in such a case. Positivist descriptions satisfy some of the assessment criteria in the grid, generally less than ten.

Back-translation and Testing are the main overseeing strategies mentioned in positivist descriptions, and their purpose is to present translation as a series of numeric clues. Without defining what equivalence means, the idea put forward is that a back-translated version that is similar to the OV informs on equivalence. Testing is addressed in the same way: the similarity of the two versions arises, in that case, from the answers provided by the subjects to whom the OV and TV were submitted. Quantitative approaches, when used in such a way, are a means to artificially circumvent epistemological questions regarding the concept of equivalence (Herdman, Fox-Rushby & Badia, 1997).

It is no surprise if those strategies—Back-translation and Testing—are the only two criteria that are satisfied by a majority of the descriptions in our corpus: quantitative approaches largely dominate in psychology. Yet those two strategies can serve another logic, typical of integrative descriptions.

Integrative descriptions. Some authors use both quantitative and qualitative data to describe PA-APs. Their descriptions are referred to as *integrative* because, among other things, they integrate two ways of conceptualizing the notion of equivalence.

On the one hand, they imply that equivalence can arise from measuring the similarity of answers provided to the TV and the OV. On the other hand, they imply that equivalence arises from meaning-negotiation proper to several consensual strategies that rely on collecting multiple informed opinions. Translators and reviewers (language experts, psychologists, experts of the construct, authors of the OV, etc.) as well as subjects enrolled in Phase III of the adaptation process, who are experts in the views expressed by the target population, all engage in the enterprise, which also needs to be properly overseen so that the reflections and discussions are directed towards the same goal.

Descriptions that meet the highest number of criteria bring out the diversity of views: command and expertise in the language, the plurality of individuals taking part in the process, production of several TVs, etc. In many cases, Back-translation is not used to measure the equivalence between two versions of a test, but rather as a means to negotiate meaning. The same applies to the testing phase, which can be used to improve the TV in light of the results obtained.

With regard to Phase III, Interviews with subjects (probe technique, or cognitive interview) are mentioned only in descriptions that meet 13 or more criteria, in a majority of them. Because it implies a discussion between the assessor and the subject, this overseeing procedure, compared to written evaluations of the TV, actively encourages meaning-negotiation. Focus groups, by fostering different points of view, also inspire debate and dialogue. However, results show that they are relatively rarely used. Bergeron (2006) provides insightful observations on how to integrate them into the adaptation process.

In the first book on qualitative approach published by the APA, Eisner (2003) discusses two essentials. The first is the level of factual detail provided about the phenomena under study: “Authors of qualitative studies increase their credibility when a reader comes to believe that its author has done the necessary to understand the situation addressed” (p. 22). The second is the researchers’ ability to see what there is to be seen, in the same way clinicians can see and decipher subtleties in their patients’ demeanour. This requires a certain degree of sensitivity: “The ability to provide a credible interpretation requires a grasp of the context in which an action occurs” (p. 23).

To adhere to Eisner’s principles in the context of a PA-AP, researchers must mobilize the knowledge embodied by the translators, reviewers, and subjects who participate in the first three phases of the process. The contribution of these people is in fact essential for data interpretation since no single researcher combines the expertise in the various fields (translation, linguistics, psychology, psychometry, etc.) that are involved in the adaptation process. Furthermore, simply seeking the input of the various experts relative to each phase, in the same way a researcher/observer would do, is not enough. Any researcher leading a PA-AP must discuss with the different experts involved. This is crucial in order to grasp the phenomena under study in its entirety, to be able, consequently, to determine which final version of the TV will go through Phase IV, and to convince readers, through the PA-AP description provided, of the credibility of the process.

Finally, these descriptions are *integrative* also in that they are the only ones that openly integrate translation into the scientific process. Incidentally, theoretical developments are particularly meaningful in these descriptions, as they support the view that translation is a theoretical object

which can be approached from many angles. It is indeed the main focus of study in translation studies.

Moreover, the vast majority of the 131 descriptions under review appear in the Method section of journal articles or academic publications, while the Results section is dedicated to presenting the results of statistical analysis. Such an editorial decision suggests that the end result of a PA-AP is not the AV *per se*, but the psychometric qualities of the AV. This in turn suggests that what lies at the core of the adaptation process is not translation but signs showing that the translation is “good”.

Yet, both descriptions that meet 17 criteria in the corpus overlap the Method and Results sections of the text in which they appear. Demers *et al.* (2002), for instance, list in their Results section the French and English items of the test they adapted, as well as problematic items that needed to be revised and their new phrasing. Dubé (2011) includes the same elements in the Results section, but also the main considerations that emerged from the external review. Readers are referred to two appendices containing: (a) the two TVs and the two back-translations produced in phase I; (b) the TV that went through phase II; (c) input received in phase III; (d) changes made to the TV after phase III; (e) the TV that went through phase IV.

References

- Achim, J. (2006). *Pour une exploration de la perspective de la maternité dans un contexte de troubles des conduites alimentaires*. Université de Montréal, Montréal, Québec, Canada.
- Acquadro, C., Conway, K., Hareendran, A., & Aaronson, N. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health*, 11(3), 509-521.
- Aimé, A. (2002). *Analyse des aspects psychosociaux reliés aux troubles des conduites alimentaires*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Arffman, I. (2013). Problems and issues in translating international educational achievement tests. *Educational Measurement: Issues and Practice*, 32(2), 2-14.
- Arnold, B. R., & Matus, Y. E. (2000). Test translation and cultural equivalence methodologies for use with diverse populations. *Handbook of multicultural mental health: Assessment and treatment of diverse populations*, 121-136.
- Audet-Lapointe, M. (2004). *Le processus d'accommadation (coping) de l'enfant et de l'adolescent confrontés à des situations stressantes : stress de la vie quotidienne ou leucémie*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- Beaulieu, G. (2012). *Développer le soutien aux besoins psychologiques des employés : évaluation des effets affectifs, comportementaux et motivationnels d'une formation destinée aux gestionnaires*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- Beaupré, M. G., King, S., Bauer, I., Hess, U., Debruille, J. B., & Corrigan, P. W. (2002). Validation of French- and English-Canadian Versions of the Social Cue Recognition Test. *Can J Psychiatry*, 47(1), 81-85.
- Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions* (Vol. 133): Sage.
- Béland, M., Chabot, K., Goulet Gervais, L., Morin, A. J. S., & Gosselin, P. (2012). Évaluation de la peur de l'accouchement. Validation et adaptation française d'une échelle mesurant la peur de l'accouchement. *L'Encéphale*, 38, 336-344.
- Bélanger, M. (2011). *Efficacité du programme mieux-être "wellness" sur la gestion du poids, l'observance à la médication et la qualité de vie chez les personnes ayant un trouble psychotique ou de l'humeur*. (D.Psy.), Université Laval, Québec, Québec, Canada.
- Belleville, G. (2006). *Évaluation d'une bibliothérapie de l'insomnie avec sevrage des hypnotiques assisté chez des adultes souffrant d'insomnie*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Bergeron, C. (2006). *Traduction et validation du Worry About Victimization (WAV) auprès d'une population âgée francophone*. (M.A.), Université de Sherbrooke, Sherbrooke, Canada.
- Berry, J. W., Poortinga, Y. H., Breugelmans, S. M., Chasiotis, A., & Sam, D. L. (2011). *Cross-cultural psychology: Research and applications (3rd edition)*: Cambridge University Press.
- Birmes, P., Brunet, A., Benoit, M., Defer, S., Hatton, L., Sztulman, H., & Schmitt, L. (2005). Validation of the Peritraumatic Dissociative Experiences Questionnaire self-report version in two samples of French-speaking individuals exposed to trauma. *European Psychiatry*, 20, 145-115#.
- Boislard Pépin, M. (2010). *Précocité sexuelle et comportement sexuel à risque à l'adolescence : Étude longitudinale des facteurs individuels familiaux, dans le groupe d'amis et contextuels associés*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.

- Boisseau, É. (2010). *La théorie de l'esprit chez les schizophrènes paranoïdes*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Boisvert, N. (2006). *Validation de l'Échelle de motivation du client face à la thérapie (EMCT), version française du Client Motivation for Therapy Scale (CMOTS)*. (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Bouchard, J. (2007). *Validation de la version française du Aggression Questionnaire auprès de deux échantillons : étudiants universitaires (étude 1) et adultes non-recrutés en milieu universitaire (étude 2)*. (Maîtrise en psychologie), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Bouffard, A. (2010). *Variabilités familiale et individuelle de la sensibilité paternelle et maternelle : corrélats contextuels, individuels et liés à l'enfant*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Bouffard, T., Seidah, A., McIntyre, M., Boivin, M., Vezneau, C., & Cantin, S. (2002). Mesure de l'estime de soi à l'adolescence : version canadienne-française du Self-Perception Profile for Adolescents de Harter. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(3), 158-162.
- Bougie, E. (2005). *The Cultural Narrative of Francophone and Anglophone Quebecers and their perceptions of temporal relative deprivation: Links with esteem and well-being*. (Ph.D.), Université McGill, Montréal, Québec, Canada.
- Brassard, A. (2003). *Syndrome d'apnées du sommeil et déficits cognitifs associés chez des patients avec ischémie cérébrale transitoire*. (M.S), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Brin, J. (2011). *Adaptation et validation française du Narcissistic Personality Inventory*. (D.Psy.), Université Laval, Québec, Canada.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of cross-cultural psychology*, 1(3), 185-216.
- Brislin, R. W. (1976). Comparative Research Methodology: Cross-Cultural Studies. *International Journal of Psychology*, 11(3), 215-229. doi:10.1080/00207597608247359
- Brislin, R. W. (1980). Cross-cultural research methods *Environment and culture* (pp. 47-82): Springer.
- Brislin, R. W. (Producer). (1986). The wording and translation of research instruments. *Field methods in cross-cultural research*.
- Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). *Cross-Cultural: Research Methods* (Vol. 11): Year Book Medical Pub.
- Brousseau, M. (2011). *He said, She said : An examination of sexual coercion from the perspective of mixed-sex couples*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Brown, J. B., Schmidt, G., Lent, B., Sas, G., & Lemelin, J. (2001). Épreuves de validation et de fiabilité d'un instrument de mesure français. *Canadian Family Physician*, 47, 988-995.
- Brunet, A., St-Hilaire, A., Jehel, L., & King, S. (2003). Validation of a French Version of the Impact of Event Scale-Revised. *Can J Psychiatry*, 48(1), 56-61.
- Charbonneau, S. (2009). *L'estimation cognitive : Analyse des fonctions cognitives sous-jacentes et étude de l'impact du vieillissement normal et de la démence de type Alzheimer*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Chevrier, N. (2009). *Adaptation québécoise de l'Oldenberg Burnout Inventory (OLBI)*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.

- Cordeau, D. (2007). *Simplicité volontaire et bien-être psychologique - Chapitre 2 : Traduction et validation québécoise de la Mesure de caractère mercantile*. (Ph.D.), Université du Québec à Trois-Rivières, Trois-Rivières, Québec, Canada.
- Cordeau, D., & Dubé, M. (2008). L'échelle de simplicité volontaire : une validation en français. *Les cahiers internationaux de psychologie sociale*, 79(3), 33-46.
- Cossette, M. (2008). *Styles de régulation émotionnelle des employés de service : Antécédents organisationnels, motivation et conséquences sur la santé psychologique*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Côté, G., Gosselin, P., & Dagenais, I. (2013). Évaluation multidimensionnelle de la régulation des émotions : propriétés psychométriques d'une version francophone du Difficulties in Emotion Regulation Scale. *Journal de thérapie comportementale et cognitive*, 23, 63-72.
- Coutu, S., Dubeau, D., Provost, M. A., Royer, N., & Lavigne, S. (2002). Validation de la version française du questionnaire Coping with Children's Negative Emotions Scale-CCNES. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(4), 230-234.
- Couture, A. (2014). *Étude pilote d'implantation et d'efficacité du CRAFT au Québec*. (D.Psy.), Université Laval, Québec, Québec, Canada.
- Couture, S., Lecours, S., Beaulieu-Pelletier, G., Philippe, F. L., & Strychar, I. (2010). French Adaptation of the Eating Disorder Recovery Self-Efficacy Questionnaire (EDRSQ): Psychometric Properties and Conceptual Overview. *European Eating Disorders Review*, 18(3), 234-243.
- D'Amico, É. (2010). *Dévoilement de l'orientation sexuelle à la famille d'origine et adaptation des jeunes gais, lesbiennes et bisexuels*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Da Silva Luis, R. (2014). *Adaptation et validation du Pathological Narcissism Inventory (PNI)*. (D. Psy.), Université Laval, Québec, Canada.
- Daigneault, I., Dion, J., Hébert, M., McDuff, P., & Collin-Vézina, D. (2013). Psychometric properties of the Child and Youth Resilience Measure (CYRM-28) among samples of French Canadian youth. *Child Abuse & Neglect*, 37, 160–171.
- Daley, M. (2007). *The impact and economic costs of insomnia : health-care utilisation, work function and accidents*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Daniel, C. (2008). *Évaluation des effets d'un programme de développement de l'empathie chez des enfants présentant des difficultés relationnelles*. (D.Psy.), Université du Québec à Montréal, Montréal, Québec, Canada.
- De Benedictis, L., Dumais, A., Stafford, M. C., Côté, G., & Lesage, A. (2012). Factor analysis of the French version of the shorter 12-item Perception of Aggression Scale (POAS) and of a new modified version of the Overt Aggression Scale (MOAS). *Journal of Psychiatric and Mental Health Nursing*, 19, 875–880.
- Delisle, J. (2011). *Le logiciel SAVR comme outil descriptif des fonctions exécutives de haut niveau dans le Trouble déficitaire de l'attention avec hyperactivité chez l'adulte*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Demers, L., Monette, M., Descent, M., Jutai, J., & Wolfson, C. (2002). The Psychosocial Impact of Assistive Devices Scale (PIADS): Translation and preliminary psychometric evaluation of a Canadian-French version. *Quality of Life Research*, 11, 583-592.
- Denis, P. (2009). *Étude de la capacité de l'aptitude cognitive générale (ACG) évaluée avec et sans contrainte temporelle et des facettes de la personnalité à prédire différentes*

- dimensions de la performance en emploi dans un contexte de sélection de personnel.*
 (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Deshaises, C., Toupin, J., & Déry, M. (2009). Validation de l'Échelle d'évaluation des traits antisociaux. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 41(1), 45-50.
- DiTommaso, E., Turbide, J., Poulin, C., & Robinson, B. (2007). L'Échelle de solitude sociale et émotionnelle (ÉSSÉ): A French-Canadian Adaptation of the Social and Emotional Loneliness Scale for Adults. *Social Behavior and Personality*, 35(3), 339-350.
- Dubé, S. (2011). *Validation transculturelle en langue française du Political Skill Inventory*.
 (Ph.D.), Université de Sherbrooke, Sherbrooke, Canada.
- Durkeim, E. (1895[1992]). *Les règles de la méthode sociologique*. Paris, France: PUF, Quadrige.
- Eisner, E. W. (2003). On the art and science of qualitative research in psychology. In P. M. Camic, J. E. Rhodes, & L. Yardley (Eds.), *Qualitative research in psychology: Expanding perspectives in methodology and design* (pp. 17-29). Washington, DC, USA: American Psychological Association.
- Emard, M. (2010). *Les contributions parentales à la compréhension des émotions des enfants en début de scolarisation*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Engelberg, M. (2003). *Binge Antecedents and Consequences in Bulimic Syndromes: An Examination of the Roles of Dietary Restraint, Affect, and Dissociation*. (Ph.D.), Université McGill, Montréal, Québec, Canada.
- Epstein, J., Santo, R. M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of clinical epidemiology*, 68(4), 435-441.
- Fikretoglu, D., Brunet, A., Poundja, J., Guay, S., & Pedlar, D. (2006). Validation of the Deployment Risk and Resilience Inventory in French-Canadian Veterans: Findings on the Relation Between Deployment Experiences and Postdeployment Health. *Can J Psychiatry*, 51(12), 755-763.
- Filiatrault, J., Demers, L., Parisien, M., Lorthios-Guilledroit, A., Kaegi, C., Ménard, I., . . . St-Laurent, C. (2014). Development and validation of a French Canadian version of the falls Behavioral (FaB) Scale. *Disabil Rehabil*, 1-6.
- Finn, K. (2012). *Le perfectionnisme et sa relation avec le contingentement universitaire et l'anxiété de performance*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Forget, H., & Robichaud, C. (2002). Validation d'une traduction française d'un instrument de mesure des réponses affectives à la première relation sexuelle avec pénétration. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(3), 151-157.
- Fortin, M. (2009). *Prévention de la récidive et non-adhérence aux programmes d'intervention obligatoires chez les contrevenants de la conduite avec facultés affaiblies par l'alcool*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- French Bourgeois, L. (2011). *L'influence des normes sociales sur le comportement : un processus médié par l'élaboration de stratégies*. (M. Sc.), Université de Montréal, Montréal, Québec, Canada.
- French, D. J., Noël, M., Vigneau, F., French, J. A., Cyr, C. P., & Thomas Evans, R. (2005). L'Échelle de dramatisation face à la douleur PCS-CF. Adaptation canadienne en langue

- française de l'échelle « Pain Catastrophizing Scale ». *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 37(3), 181-192.
- Gagné, J. (2010). *L'association entre la psychopathie et les pratiques parentales dans une population non criminelle*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- Gagnon, C. (2013). *Traduction et validation du O'Brien Multiphasic Narcissism Inventory (OMNI)*. (D.Psy.), Université Laval, Québec, Canada.
- Gagnon, M. (2010). *Évaluation du perfectionnisme prescrit par autrui chez des patients souffrant de maladies cardiovaskulaires du Saguenay-Lac-St-Jean et de ses relations avec l'activité physique, les habitudes alimentaires et le bien-être psychologique*. (D.Psy.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Gagnon, V. (2004). *Étude comparative des stratégies d'adaptation de rumination et de distraction chez les adultes et les adolescents dépressifs*. (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological assessment*, 6(4), 304.
- Geisinger, K. F. (2003). Testing and Assessment in Cross-Cultural Psychology. In J. R. Graham & J. A. Naglieri (Eds.), *Handbook of psychology - Assessment Psychology* (pp. 650). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Giguère, J., Fortin, C., & Sabourin, S. (2006). Déterminants de la persistance conjugale chez des personnes vivant en première et en seconde union conjugale. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 38(2), 185-199.
- Gilbert, E. (2011). *La mémoire épisodique et le fonctionnement social dans les troubles du spectre autistique*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Grégoire, S., Baron, L., Ménard, J., & Lachance, L. (2013). The Authenticity Scale: Psychometric Properties of a French Translation and Exploration of Its Relationships With Personality and Well-Being. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, Advance online publication. doi:10.1037/a0030962
- Grimard, F. (2011). *L'anxiété sociale et la victimisation par les pairs à l'adolescence*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Guillemain, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of clinical epidemiology*, 46(12), 1417-1432.
- Hambleton, R. K. (1993). *Technical standards for translating tests and establishing test score equivalence*. Paper presented at the 101st Annual Convention of the American Psychological Association. Toronto, Ontario, Canada.
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal*, 229-240.
- Hambleton, R. K. (2001). The next generation of the ITC Test Translation and Adaptation Guidelines. *European Journal of Psychological Assessment*, 17(3), 164.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (Vol. 1, pp. 17-38). Mahwah, NJ: Erlbaum.
- Hambleton, R. K., & de Jong, J. H. (2003). Advances in translating and adapting educational and psychological tests. *Language testing*, 20(2), 127-134.

- Hambleton, R. K., & Kanjee, A. (1995). Increasing the validity of cross-cultural assessments: Use of improved methods for test adaptations. *European Journal of Psychological Assessment*, 11(3), 147.
- Hambleton, R. K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Association of Test Publishers*, 1(1), 1-13.
- Hambleton, R. K., & Patsula, L. (2000). Adapting Tests for Use in Multiple Languages and Cultures. Laboratory of Psychometric and Evaluative Research Report.
- Harkness, J. A. (2011). Translation. In S. R. Center (Ed.), *Guidelines for best practice in cross-cultural surveys (3rd ed.)*. (pp. 1-35). Ann Arbor, MI: University of Michigan: Institute for Social Research. Retrieved from www.ccsq.isr.umich.edu.
- Hébert, M., Parent, N., & Daignault, I. V. (2007). The French-Canadian Version of the Self-Report Coping Scale: Estimates of the Reliability, Validity, and Development of a Short Form. *Measurement and Evaluation in Counseling and Development*, 40, 2-15.
- Herdman, M., Fox-Rushby, J., & Badia, X. (1997). 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research*, 6(3), 237-247.
- Hovington, C. (2002). *Les inondations de juillet 1996 au Saguenay : Les effets psychologiques durables chez les adultes jeunes et âgés*. (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Laval.
- Husky, M. M., Grondin, O. S., & Compagnone, P. D. (2004). Validation de la version française du questionnaire de Sociotropie-Autonomie de Beck et collègues. *Can J Psychiatry*, 49(12), 851-858.
- Jeanrie, C., & Bertrand, R. (1999). Translating tests with the International Test Commission's guidelines: Keeping validity in mind. *European Journal of Psychological Assessment*, 15(3), 277.
- Johnson, K. (2008). *L'orientation vers les buts : le contraste de deux modèles*. (M. Sc.), Université de Montréal, Montréal, Québec, Canada.
- Julien, D., Careau, Y., O'Connor, K. P., Bouvard, M., Rhéaume, J., Langlois, F., . . . Cottraux, J. (2008). Specificity of belief domains in OCD: Validation of the French version of the Obsessive Beliefs Questionnaire and a comparison across samples. *Journal of Anxiety Disorders*, 22, 1029-1041.
- Koski, L., Xie, H., Konsztowicz, S., & Tetteh, R. (2010). French-English Cross-Linguistic Comparison and Diagnostic Impact of the AD-8 Dementia Screening Questionnaire in a Geriatric Assessment Clinic. *Dement Geriatr Cogn Disord*, 29, 265-274.
- L'Archevêque, A. (2009). *Étude exploratoire des contextes d'accès à la parentalité et des facteurs contribuant à l'intégration identitaire chez les pères gais*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Lacasse, A. (2004). *Peer sexual harassment among adolescents: A developmental perspective*. (Ph.D.), Université McGill, Montréal, Québec, Canada.
- Lachapelle, Y., Boisvert, D., & Leclerc, D. (2000). La traduction et la validation transculturelle de l'Échelle d'autodétermination de l'ARC. *Revue francophone de la déficience intellectuelle*, 70-74.
- Lafontaine, M. F., & Lussier, Y. (2003). Structure bidimensionnelle de l'attachement amoureux : Anxiété face à l'abandon et évitement de l'intimité. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 35(1), 56-60.

- Lafrenière, M. (2013). *Le rôle modérateur de la passion dans la relation entre la performance/le rehaussement du soi et l'adaptation psychologique*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Laliberté, S., & Earls, C. M. (2007). Traduction et validation en français du Situaiton Humor Response Questionnaire et du Coping Humor Scale. *Revue québécoise de psychologie*, 28(2), 213-232.
- Lamoureux-Hébert, M. (2006). *Les besoins de soutien des personnes présentant un retard mental manifestant des comportements problématiques*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Landry, T. (2010). *La dyspareunie chez les adolescents : Prévalence, caractéristiques et corrélats biopsychosociaux*. (Ph.D.), Université du Québec à Montréal, Montréal.
- Lauzier, M. (2010). Adaptation et validation canadienne-française de l'Échelle des conceptions implicites de l'intelligence (ECII). *Revue québécoise de psychologie*, 31(3), 189-206.
- Lauzier, M., Barrette, J., Lemyre, L., & Corneil, W. (2013). Validation francophone de l'Inventaire des facilitateurs de l'apprentissage organisationnel (IFAO) : étude comparative de l'apprentissage perçu au niveau des individus, des groupes et de l'organisation. *Revue européenne de psychologie appliquée*, 63, 353-362.
- Lauzier, M., & Haccoun, R. R. (2010). Validation canadienne-française de l'Échelle des styles d'orientation des buts (ESOB). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 42(2), 127-133.
- Laverdière, O. (2009). *La pathologie de la personnalité : contributions du tempérament, de l'attachement, des traumas et médiation de la compréhension émotionnelle - Premier article : French adaptation of the Mini-IPIP: A short measure of the Big Five*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Laverdière, O., Diguer, L., Gamache, D., & Evans, D. E. (2010). The French Adaptation of the Short Form of the Adult Temperament Questionnaire. *European Journal of Psychological Assessment*, 26(3), 212-219.
- Lavergne, A. G. (2003). *Les déterminants de l'intensité du deuil à la suite de la perte d'un animal de compagnie - Validation d'un instrument et étude de corrélats*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Lavigne, G. (2011). *L'influence des orientations du besoin d'affiliation sur l'adaptation intrapersonnelle et interpersonnelle*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Leblanc, G., Marchand, A., Khouzam, C., & Guay, S. (2007). Validation canadienne-française du Questionnaire sur la peur de l'intimité. *Revue québécoise de psychologie*, 28(2), 277-292.
- Leboeuf, G., & Losier, G. F. (2012). L'échelle d'orientation contingente extrinsèque (ÉOCE) : Traduction et validation francophone du Extrinsic Contingency Focus Scale (ECFS). *Revue européenne de psychologie appliquée*, 62, 137-145.
- Lecavalier, L., & Tassé, M. J. (2001). Traduction et adaptation transculturelle du Reiss Screen for Maladaptive Behavior. *Revue francophone de la déficience intellectuelle*, 12(1), 31-44.
- Lekes, N. (2012). *Self-growth, close relationships, and community contribution: Exploring the development of intrinsic value priorities and their influence on well-being*. (Ph.D.), Université McGill, Montréal, Québec, Canada.
- Lemelin, J. P., Tarabulsy, G. M., Provost, M. A., Fournier, M., Robitaille, J., Hémond, I., & Tessier, R. (2007). Le Questionnaire d'évaluation du comportement de l'enfant : version

- canadienne française du Toddler Behavior Assessment Questionnaire. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 39(4), 291-300.
- Lemieux, R. (2009). *Étude de relations entre la dissociation, le fonctionnement réflexif et les difficultés d'adaptation chez des enfants abusés sexuellement*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Lepage, J. F., Lortie, M., Taschereau-Dumouchel, V., & Théoret, H. (2009). Validation of French-Canadian Versions of the Empathy Quotient and Autism Spectrum Quotient. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 41(4), 272-276.
- Lessard, M. (2012). *Interventions psychologiques brèves du Trouble panique chez les individus qui se présentent au service d'urgence pour des douleurs thoraciques non cardiaques : Efficacité et recension des écrits*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: a methods review. *Journal of advanced nursing*, 48(2), 175-186.
- Martin, A., & Gosselin, P. (2012). Propriétés psychométriques de l'adaptation francophone d'une mesure de symptômes des troubles anxieux auprès d'enfants et d'adolescents (SCARED-R). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 44(1), 70-76. doi:10.1037/a0023103
- Maxwell, B. (1996). *Translation and cultural adaptation of the survey instruments*. Retrieved from <http://files.eric.ed.gov/fulltext/ED406418.pdf#page=159>.
- Mayer-Renaud, J. (2008). *Contribution de l'intimité à l'effet du soutien social dans la relation entre l'alexithymie et la dépression*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- Ménard, J. (2008). *Authenticité et bien-être au travail : étude auprès de gestionnaires*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- Méthot, S., & Morin, D. (2004). Traduction d'un instrument de dépistage des symptômes affectifs et anxieux chez les adultes ayant un retard mentale. *Revue francophone de la déficience intellectuelle*, 15(1), 57-73.
- Morin, M. (2007). *La créativité dans l'activité mentale du thérapeute*. (M. Ps.), Université Laval, Québec, Québec, Canada.
- O'Rourke, N., & Cappeliez, P. (2003). Validation d'une mesure de réponses biaisées à propos de la relation conjugale : l'Échelle d'embellissement conjugal. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 35(1), 45-49.
- Orendain Tremear, M. (2010). *Les relations d'objet, la dualité anaclitique-introjective et leur relation avec les symptômes dépressifs*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Pan, Y., & De La Puente, M. (2005). *Census Bureau guideline for the translation of data collection instruments and supporting materials: Documentation on how the guideline was developed*. Retrieved from www.census.gov.
- Paquette, D., Laporte, L., Bigras, M., & Zoccolillo, M. (2004). Validation de la version française du CTQ et prévalence de l'histoire de maltraitance. *Santé mentale au Québec*, 29(1), 201-220.
- Paradis, A. (2012). *Maltraitances dans l'enfance, problèmes interpersonnels, et violence dans les relations de couple : Différences de trajectoires entre hommes et femmes*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.

- Payer, M. (2012). *L'état de stress post-traumatique-complexe : un concept novateur à explorer.* (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- Pelletier, O. (2007). *Troubles de la personnalité et jeu pathologique : comorbidité et prédicteurs d'abandon du traitement.* (Ph.D.), Université Laval, Québec, Québec, Canada.
- Perozzo, C. (2007). *Le changement de l'identité sociale : une question de discrimination ou de privation relative?* (M.Sc.), Université de Montréal, Montréal, Québec, Canada.
- Plante-Jean, M. (2012). *Facteurs prédictifs de la dépression chez les gens souffrant d'anxiété sociale.* (Ph.D.), Université Laval, Québec, Québec, Canada.
- Plante, N., Daigle, M. S., Gaumont, C., Charbonneau, L., Gibbs, J., & Barriga, A. (2012). Validation of the 'How I Think Questionnaire' in a Population of French-speaking Adolescents with Externalizing Behaviors. *Behavioral Sciences and the Law*, 30, 196–210.
- Poirier, S., Bureau, V., Lehoux, C., Bouchard, R. H., Maziade, M., Pelletier, S., . . . Roy, M. A. (2004). A Factor Analysis of the Strauss and Carpenter Revised Outcome Criteria Scale: A Validation of the French Translation. *The Journal of Nervous and Mental Disease*, 192(12), 864-867.
- Potvin, M. (2011). *Évaluation écologique et rééducation de la mémoire prospective chez des patients ayant subi un traumatisme craniocérébral.* (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Poulin-Vallières, I. (2004). *Étude sur le comportement, le sommeil et la mémoire chez les jeunes âgés entre 7 et 15 ans, victimes d'un traumatisme crânien léger.* (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Poulin, V., & Desrosiers, J. (2010). Validation française de questionnaires de participation sociale auprès d'aînés en situation d'incapacités. *Canadian Journal of Occupational Therapy*, 77(3), 159-166.
- Prieto, A. J. (1992). A method for translation of instruments to other languages. *Adult Education Quarterly*, 43(1), 1-14.
- Quesnel, C. (2007). *Évaluation longitudinale du fonctionnement cognitif de femmes traitées pour un cancer du sein non métastatique.* (Ph.D.), Université Laval, Québec, Québec, Canada.
- Rabiau, M. (2006). *Balancing Temptations and Health Goals: The Role of Compensatory Health Beliefs.* (Ph.D.), Université McGill, Montréal, Québec, Canada.
- Radomsky, A. S., Ashbaugh, A. R., Saxe, M. L., Ouimet, A. J., Golden, E. R., Lavoie, S. L., & O'Connor, K. P. (2006). Psychometric Properties of the French and English Versions of the Social Phobia Inventory. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 38(4), 354-360.
- René de Cotret, F., Leanza, Y., Francoeur, A., & Beaudoin-Julien, A. (2019). Consensual strategies for adapting psychometric tests: Understanding and schematization of the adaptation process. Manuscript submitted for publication.
- Ricard-St-Aubin, J. S., Philippe, F. L., Beaulieu-Pelletier, G., & Lecours, S. (2010). Validation francophone de l'Échelle des émotions différentielles IV (EED-IV). *Revue européenne de psychologie appliquée*, 60, 41-53.
- Richer, I., & Bergeron, J. (2012). Differentiating risky and aggressive driving: Further support of the internal validity of the Dula Dangerous Driving Index. *Accident Analysis and Prevention*, 45, 620– 627.
- Rios, J. A., & Sireci, S. G. (2014). Guidelines versus practices in cross-lingual assessment: A disconcerting disconnect. *International Journal of Testing*, 14(4), 289-312.

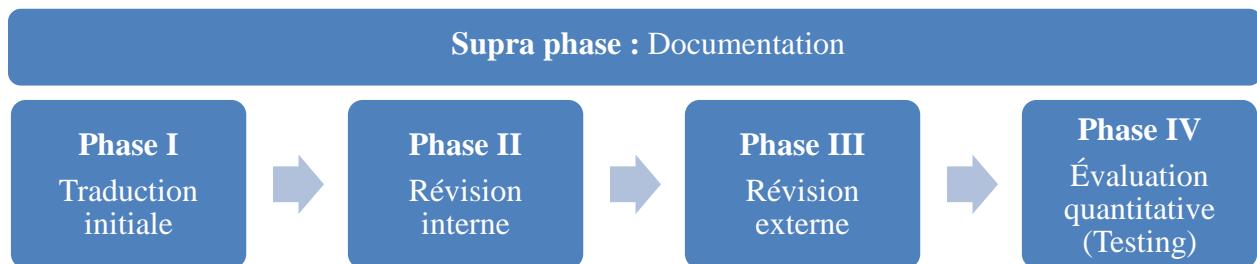
- Roberge, M. (2007). *État de stress aigu et état de stress post-traumatique après un infarctus du myocarde : Prévalence et facteurs associés*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Roberge, P., Marchand, L., Grenier, S., & Marchand, A. (2003). Validation canadienne-française de l'Échelle d'évaluation de la panique et de l'agoraphobie. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 35(1), 61-66.
- Seidah, A., Bouffard, T., & Vezeau, C. (2002). La mesure du perfectionnisme : validation canadienne-française du Positive and Negative Perfectionism Scale. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(3), 168-171.
- Sergerie, M. (2008). *Analyse dyadique des relations cybermentorales de la ressource Academos*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Simard, S. (2008). *Vers une conceptualisation multidimensionnelle de la peur de la récidive du cancer : évaluation, nature des pensées intrusives et comorbidité psychiatrique*. (Ph.D.), Université Laval, Québec, Québec, Canada.
- Sinaiko, H. W., & Brislin, R. W. (1973). Evaluating language translations: Experiments on three assessment methods. *Journal of Applied Psychology*, 57(3), 328.
- Sireci, S. G., Patsula, L., & Hambleton, R. K. (2005). Statistical methods for identifying flaws in the test adaptation process. *Adapting educational and psychological tests for cross-cultural assessment*, 93-115.
- Solano-Flores, G., Backhoff, E., & Contreras-Niño, L. Á. (2009). Theory of test translation error. *International Journal of Testing*, 9(2), 78-91.
- Sperber, A. D., Devellis, R. F., & Boehlecke, B. (1994). Cross-cultural translation methodology and validation. *Journal of cross-cultural psychology*, 25(4), 501-524.
- Spielberger, C. D., & Sharma, S. (1976). Cross-cultural measurement of anxiety *Cross-cultural anxiety* (Vol. 1, pp. 13-25).
- Taillon, A. (2011). *Peur d'une dysmorphie corporelle : réflexion critique sur les traitements psychologiques et évaluation d'une thérapie basée sur les inférences*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Tassé, M. J., & Craig, E. M. (1999). Critical issues in the cross-cultural assessment of adaptive behavior *Adaptive behavior and its measurement: Implications for the field of mental retardation* (pp. 161-184).
- Temimi, N. (2011). *Le rôle des dimensions culturelles de Hofstede en regard du fonctionnement interne des équipes de travail pluriculturelles*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.
- Thibaudeau, P. (2006). *L'attachement parental et l'ajustement psychosocial des adolescents : validation d'un instrument de mesure et étude du rôle médiateur de l'estime de soi*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Thibault, P. (2008). *Perception d'authenticité du sourire : Perspective interculturelle et développementale*. (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Toledano, R. (2006). *Development of the Sexual Arousal and Desire Inventory (SADI): A Multidimensional Scale of Subjective Sexual Arousal and Desire in Men and Women*. (Ph.D.), Université Concordia, Montréal, Québec, Canada.
- Tremblay, C. (2001). *Identité ethnique et maturité vocationnelle à l'adolescence : perspective développementale, contextuelle et interculturelle*. (Ph.D.), Université de Montréal, Montréal, Québec, Canada.

- Tremblay, S. (2009). *Engagement du père lors de la transition à devenir parent : rôle des cognitions paternelles et maternelles.* (Ph.D.), Université Laval, Québec, Québec, Canada.
- Trottier, C., Mageau, G., Trudel, P., & Halliwell, W. R. (2008). Validation de la version canadienne-française du Life Orientation Test-Revised. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 40(4), 238-243.
- Turcotte, S. (2013). *Traduction et validation du Relationship Patterns Questionnaire (RPQ) et étude pilote de ses liens avec le Personality Organization Diagnostic Form (PODF).* (D.Psy.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Turgeon, L., & Chartrand, É. (2003). Reliability and Validity of the Revised Children's Manifest Anxiety Scale in a French-Canadian Sample. *Psychological Assessment*, 15(3), 378-383. doi:10.1037/1040-3590.15.3.378
- Turgeon, M. E., Forget, J., & Bégin, J. (2013). Validation francophone et québécoise de la sous-échelle Weight-Related Teasing du Perception of Teasing Scale chez les enfants. *Journal de thérapie comportementale et cognitive*, 23, 52-62.
- Vaillancourt, L. (2009). *L'inconfort spatio-moteur et les attitudes envers la pratique d'activités physiques dans le Trouble panique avec ou sans agoraphobie.* (Ph.D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- 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.
- Vallerand, R. J., & Halliwell, W. R. (1983). Vers une méthodologie de validation transculturelle de questionnaires psychologiques: Implications pour la psychologie du sport. *Canadian Journal of Applied Sport Sciences*, 8(1), 9-18.
- van de Vijver, F., & Hambleton, R. K. (1996). Translating tests. *European psychologist*, 1(2), 89-99.
- van de Vijver, F., & Leung, K. (1997). *Methods and data analysis for cross-cultural research* (Vol. 1): Sage.
- van de Vijver, F., & Poortinga, Y. H. (2005). Conceptual and methodological issues in adapting tests. *Adapting educational and psychological tests for cross-cultural assessment*, 39-63.
- Van Widenfelt, B. M., Treffers, P. D., De Beurs, E., Siebelink, B. M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical child and family psychology review*, 8(2), 135-147.

Conclusion

Le processus d'adaptation constitue une démarche scientifique ayant comme objet l'équivalence des textes cible et source, à savoir la VT et la VO. Cette recherche de l'équivalence passe par divers mécanismes qui s'agglutinent sur la traduction. L'analyse de contenu (Article I) a permis de circonscrire les mécanismes faisant consensus dans la littérature scientifique, à savoir 17 stratégies d'adaptation. Ces stratégies constituent les matériaux de base du processus d'adaptation et peuvent être regroupées en cinq phases, tel que présenté ci-dessous (Figure 7).

Figure 4. Schéma des cinq phases du processus d'adaptation



Nous avons vu dans le premier article que la légitimité du processus d'adaptation passe par lesdites stratégies. Or, celles-ci ne se situent pas toutes sur le même plan ontologique, ce qui n'est pas sans incidence. La plupart des stratégies relatives aux Phases I-II-III sous-entendent que chaque individu qui participe au processus d'adaptation construit activement sa propre réalité selon qui il est et ce qu'il expérimente. La diversité des points de vue est essentielle à ces phases, car elle permet la négociation du sens : le sens que l'équivalence peut prendre est négocié entre les différents individus et est relatif à cette négociation. La traduction inversée, qu'on pourrait qualifier de stratégie positiviste à la suite de la lecture de l'Article II, sert pourtant la négociation du sens lorsqu'elle est employée en concomitance avec d'autres stratégies d'encadrement de la révision interne.

En comparaison, à la Phase IV, on ne considère qu'une seule réalité afin de pouvoir utiliser la VT en tant que test psychométrique, sans quoi il serait impossible de comparer les résultats des sujets. Ainsi, à la Phase IV, il est question d'une ontologie universaliste, alors qu'aux trois autres phases, nous avons affaire à une ontologie relativiste.

De cette transition ontologique découle une transition épistémologique ainsi qu'une transition de la méthode d'analyse des données (voir le Tableau 1 en sous-section 1.5 et les paragraphes qui suivent). Aux Phases I-II-III, le chercheur est participant alors qu'en Phase IV, il est observateur et emploie un instrument pour mesurer l'objet d'étude, qui est ensuite traité par le biais d'analyses statistiques.

Également, nous avons constaté que dans la quasi-totalité des descriptions examinées pour le deuxième article, les stratégies décrites aux premières phases du processus d'adaptation sont uniquement traitées en tant que méthode, non en tant que résultat. Ce qui est produit à ces phases représente pourtant un résultat.

Prenons la Phase III en exemple. Un consensus existe dans la littérature scientifique à savoir qu'il faille effectuer une révision externe, par exemple en sous-questionnant les sujets à l'oral (*probe technique*). Employer cette stratégie accroît vraisemblablement la légitimité du processus d'adaptation. Or, quelle est l'utilité de recourir à une méthode si on n'en communique pas le résultat? À la Phase III, le résultat n'est pas dichotomique comme en Phase IV, car la méthode d'analyse ne pourrait être la même. Le résultat en Phase III ne pourrait pas être : « Les sujets confirment que la VT est claire ». La Phase III ne sert pas à confirmer ou à infirmer une hypothèse, mais plutôt à en apprendre davantage sur le phénomène étudié en ayant une discussion avec des sujets. Ce sont les points saillants de ces discussions qui devraient faire office de résultats. Idem aux Phases I et II. Aux trois premières phases, l'équivalence des versions n'est pas mesurée, mais négociée. La capacité du chercheur à identifier et à décrire les points saillants de cette négociation agira comme critère de crédibilité de la démarche, que le lecteur aura tôt fait de juger.

En discussion de l'Article II, nous avons posé l'existence de trois types de description, à savoir les descriptions libres, positivistes et intégratives. Une description libre ne fournit pas de résultats à propos du processus d'adaptation; son auteur ne cherche pas à convaincre le lecteur de la légitimité du test adapté. En comparaison, dans une description positiviste, on met de l'avant des résultats quantitatifs. Cette description est un excellent exemple d'une analyse des données instrumentale pour laquelle les Phases I-II-III sont au service du Testing. Finalement, dans une description intégrative, des résultats qualitatifs s'ajoutent aux résultats quantitatifs et peuvent même se situer dans la section Résultats de la publication. En outre, le Testing n'est pas

seulement employé pour confirmer l'équivalence des versions, mais aussi pour cibler les segments de traduction à réviser. Dans la description intégrative, plus longue, la documentation de la démarche est détaillée.

Selon les résultats obtenus avec la G-PA, il est possible de conclure que la majorité des 131 descriptions analysées s'éloignent passablement des stratégies consensuelles, et donc qu'il y a un décalage entre ce qui est recommandé dans la littérature (Article I) et l'information publiée à propos de ce qui est effectué sur le terrain (Article II). Cela n'est pas étonnant : plusieurs des stratégies consensuelles sous-entendent une ontologie relativiste, ce qui contraste avec le positivisme, dominant dans le domaine de la psychologie et des domaines connexes.

Il s'agit d'un contraste, mais pas d'une opposition. Comme le rappelle Eisner (2003), le terme *empirique* vient du latin *empiria*, qui signifie *expérimenter* : « Empirical research always refers to phenomena that can be experienced, and to be experienced the senses must be engaged with qualities, even when the qualities in question are imaginatives » (p. 20). Il précise qu'autant les recherches quantitatives que qualitatives décrivent des qualités, mais que la différence entre les deux réside dans les moyens utilisés pour se les représenter. L'équivalence des versions peut, d'une part, être représentée en termes de quantité, particulièrement lors de la Phase IV. À cette phase ultime, l'expérience des sens est artificielle, cloisonnée dans un texte aux réponses prédéterminées. Ce serait justement ces réponses, dites objectives car elles n'impliquent pas une réflexion de la part de l'évaluateur pour en comprendre le sens, qui permettraient de décrire l'équivalence en termes de magnitude.

D'autre part, l'équivalence peut être représentée directement en termes de qualité, en examinant comment les différents acteurs impliqués dans les Phases I-II-III expérimentent et comprennent le texte et l'équivalence des textes. Les mots ne sont jamais neutres : ils suscitent des réactions, intellectuelles et émotives; ces témoignages permettent de qualifier l'équivalence, et, ultimement, de déterminer quels segments de traduction composeront la VT qui ira en Phase IV.

Étudier la traduction

Dans la section Problématique, j'ai présenté des stratégies et procédés de traduction. Ceux-ci sont connus des personnes ayant étudié la traduction à l'université, mais, sauf exception, pas de

celles ayant étudié en psychologie ou dans un domaine connexe. Une reconnaissance des stratégies et procédés de traduction dans ces domaines-ci pourrait faciliter le processus d'adaptation. En particulier, on pourrait considérer les trois niveaux de difficulté et d'interprétation proposés par Delisle (2003) afin d'étudier si certains segments de traduction n'aboutissent pas à l'équivalence. Les segments de traduction identifiés comme des reports ou des remémorations pourraient également occuper une place moindre lors de la révision. Il est par ailleurs envisageable que la création discursive suscite des discussions déterminantes quant à la négociation du sens lors de la révision. Il serait aussi possible de tenir compte de la direction de la traduction de chaque segment, soit directe ou oblique.

Je note au passage le parallèle entre le terme *stratégie de traduction* utilisé par Delisle (2003) (sous-section 1.2) et le terme *stratégie d'adaptation* que je propose dans mon mémoire doctoral. La stratégie de traduction vise l'intégralité d'un texte au même titre que la stratégie d'adaptation vise l'intégralité d'un test. D'une manière ou d'une autre, la stratégie ne se situe jamais au niveau du segment de traduction, mais à un niveau supérieur. Ainsi, en transposant la définition de Delisle (p. 60), on obtient une définition tout à fait valable : « [Les stratégies d'adaptation] oriente[nt] la démarche globale du traducteur à l'égard d'un texte en particulier à traduire et se distingue[nt] des décisions ponctuelles comme l'application des divers procédés de traduction ».

Tel que mentionné à la sous-section 1.3, les recherches en traductologie appliquée illustrent des différences importantes entre les traducteurs débutants et les traducteurs expérimentés. L'expérience permet au traducteur de jongler plus aisément avec les principes, stratégies et procédés de traduction, ce qui débouche sur un plus grand inventaire de possibles. L'expérience tend également à permettre une analyse de segments de traduction plus longs et à automatiser les prises de décision des traducteurs, qui peuvent dès lors rediriger leur attention sur des aspects tels que le construit sous-jacent au texte. Elle permet également de ne pas recourir à la traduction littérale là où elle n'est pas indiquée et de favoriser la stylistique afin que la lecture du texte cible demeure fluide et cohérente.

Considérer l'expérience en traduction des traducteurs occasionnerait d'autres perspectives de recherche. Par exemple, on pourrait étudier l'influence de l'expérience en traduction sur la gravité des erreurs de traduction selon la stratégie de révision interne de Solano-Flores *et al.* (2009). On pourrait également étudier l'influence de l'expérience en traduction sur le type de

procédé de traduction favorisé en se référant à la typologie de Vinay et Darbelnet (1977[1954]). Dans ce cas, il serait également possible de regrouper les procédés selon qu'ils sont directs ou obliques : par exemple, les segments de traduction sont-ils plus souvent obliques ou directs quand le traducteur a moins d'expérience? On pourrait faire de même avec la typologie de Delisle (2003) : est-ce que l'expérience en traduction a une influence sur les niveaux d'interprétation favorisés, soit le report, la remémoration et la création discursive? Des questions plus larges pourraient aussi être envisagées : qu'en est-il de la vitesse d'exécution? Est-il plus rentable d'investir dans un traducteur professionnel? Est-il rentable de ne pas l'impliquer dans la révision par la suite?

Également, il ne faut pas perdre de vue que la traduction de tests psychométriques se distingue de plusieurs autres types de traduction en cela que le contexte, qui donne un sens aux segments de traduction, en particulier les items, n'y est pas explicite. Il est envisageable que les traducteurs non professionnels, par exemple des psychologues, soient plus efficaces que les traducteurs professionnels quand il s'agit d'accéder au sens de certains items, car ils comprennent plus amplement la logique sous-jacente au test psychométrique.

Voici, pour illustrer la compétence respective de ces traducteurs, mais également leur complémentarité, quatre des items du Narcissistic Personality Inventory (NPI) traduits par une équipe de quatre traducteurs amateurs, dont moi (Brin, 2011) :

I like to be complimented → J'aime être complimenté

I really like to be the center of attention → J'aime être le centre de l'attention

I will be a success → Je deviendrai un succès

I have a strong will of power → J'ai une forte volonté de pouvoir

Il est question de quatre traductions directes, bien que la troisième et la quatrième auraient nécessité une traduction oblique. La deuxième, par ailleurs, aurait dû tenir compte du mot *really*. Le traducteur professionnel n'aurait pas fait ces erreurs. Celui-ci aurait pourtant avantage à travailler avec un spécialiste du domaine de la psychologie. Par exemple, il est pertinent de savoir que l'item *I have a strong will of power* représente un des critères diagnostiques du Trouble de la personnalité narcissique du DSM-III formulé ainsi : « Preoccupation with fantasies

of unlimited success, power, brilliance, beauty, or ideal love ». La connaissance de l'évolution du NPI peut également servir. À l'origine (Raskin & Hall, 1979), les items du NPI contenaient des couples d'affirmations. Le participant devait s'identifier à l'une des deux affirmations, par exemple entre *I have a strong will to power* et *Power for its own sake doesn't interest me*. En ayant en tête les critères diagnostiques du DSM-III et la version originale du NPI, il serait facile de trancher entre plusieurs traductions si l'une d'elles était *Je suis assoiffé de pouvoir*, qui correspondrait à ce que l'item *I have a strong will to power* est censé mesurer.

Tout comme la psyché, le langage n'appartient pas seulement aux spécialistes : chaque individu le possède. La psyché et le langage peuvent pourtant être traités en tant que systèmes complexes. Les psychologues et les linguistes s'entendent sur ce point à propos de leur objet d'étude respectif. Au même titre qu'une formation universitaire rigoureuse est entendue pour l'apprenti psychologue, il serait approprié que l'importance de la formation universitaire pour l'apprenti traducteur soit considérée dans le processus d'adaptation de tests psychométriques. Indiquer la formation universitaire du traducteur et les certifications qu'il possède, le cas échéant, serait approprié. Également, le nombre d'année(s) d'expérience, l'expérience en traduction de tests psychométriques, les expériences connexes et le principal gagne-pain du traducteur, si ce n'est pas la traduction, pourraient être précisés.

La traduction littérale et la traduction inversée

Tel que mentionné à la sous-section 1.2, la traduction littérale peut être aussi bien une stratégie qu'un procédé de traduction. En tant que stratégie, elle vise le texte cible dans son entièreté alors que, en tant que procédé, elle vise un segment de traduction. Lors du processus d'adaptation, la traduction littérale en tant que procédé peut être indiquée, contrairement à la stratégie. De fait, certains segments de traduction se prêtent bien à une traduction littérale, par exemple, le premier des quatre items donnés en exemple précédemment : *I like to be complemented* traduit par *J'aime être complimenté*. Il est pourtant risqué de penser qu'un test puisse être traduit littéralement dans son intégralité. Nida (1964) mentionne que cette stratégie permet au lecteur cible de se mettre dans la peau du lecteur source, pour que celui-là comprenne comment celui-ci a pu comprendre, voire expérimenter le texte. Or, adapter un test implique de recréer chez le lecteur cible la réponse hypothétique que le texte source suscite chez le lecteur source. Pour ce faire, la

traduction doit conséquemment être orientée vers le texte cible, ce qui s'oppose à une stratégie de traduction littérale.

La traduction inversée nécessite le recours à la traduction littérale ou aux autres procédés de traduction directe décrits par Darbelnet (1966), soit l'emprunt et le calque. Tel qu'indiqué à la sous-section 1.2, la traduction directe sous-entend une symétrie entre deux langues. C'est cette symétrie qui permet aux tenants de la traduction inversée de concevoir que la langue source peut être transposée en langue cible pour ensuite être ramenée en langue source sans engendrer d'aspérité significative. Or, prises en tant que systèmes, les langues ne sont pas symétriques; chacune d'elle ne découpe pas l'expérience vécue de la même manière, ce qui explique l'existence de la traduction oblique par ailleurs. Ainsi, la traduction inversée ne pourrait être qu'une technique partiellement efficace pour adapter un test, ce pour quoi, à mon sens, la communauté scientifique recommande qu'elle ne soit jamais employée seule.

L'importance de la théorie

Toute démarche scientifique repose sur une arborescence théorique, sans quoi les connaissances produites ne pourraient être qualifiées de scientifiques. Ainsi, effectuer un processus d'adaptation implique de bâtir un cadre théorique justifiant les stratégies utilisées et la démarche dans son ensemble. Un tel cadre supporte les observations faites lors du processus, qui, une fois mentionnées dans la description, permettent au lecteur de conclure que la VA est équivalente à la VO. Exemple : en tant que tel, produire plusieurs VT ne permet pas de conclure en une équivalence plus grande entre la VA et la VO. Or, cela peut être le cas si le chercheur se réfère, dans le complément théorique, à des publications scientifiques dans lesquelles on avance ce fait. Les trois types de descriptions proposés dans le deuxième article offrent des exemples, et surtout des contre-exemples, en cette matière.

Élaborée pour les besoins de l'Article II, la G-PA constitue un outil facile d'utilisation pour le chercheur ou le clinicien qui veut examiner la description de certains processus d'adaptation effectués. La grille peut aussi être utilisée par les éditeurs et les universités afin d'établir des critères stricts par rapport à de telles descriptions : qu'est-ce qui est acceptable et qu'est-ce qui ne l'est pas? Les trois types de description, proposés également à l'article II, complètent la G-

PA. Ceci dit, parce qu’elles se situent aux antipodes de la démarche scientifique, les descriptions libres devraient être écartées d’entrée de jeu.

Réviser le cursus universitaire. Finalement, offrir aux étudiants en psychologie une formation de troisième cycle portant sur le processus d’adaptation apparaît nécessaire compte tenu de la complexité et de l’importance du phénomène que mon mémoire doctoral met en lumière. Inspirée de celui-ci, la formation pourrait être divisée en deux segments de trois heures et être incorporée à un cours existant.

Premier bloc. Essentiellement théorique, le premier bloc serait destiné à informer l’étudiant à propos du processus d’adaptation. Qu’est-ce que le processus d’adaptation? Depuis quand adapte-t-on des tests, et pourquoi? À quel point est-ce pertinent d’étudier le processus d’adaptation? De quelle manière peut-on le schématiser? Qu’est-ce qu’une traduction inversée, un prétest, un testing, ...?

À ces questions de base s’ajoutent plusieurs questions périphériques telles que la différence entre un traducteur professionnel et un traducteur amateur : par exemple, sur quelles bases peut-on distinguer l’un de l’autre? Et pourquoi ces deux types de traducteurs se complètent-ils lorsqu’il est question d’adapter un test? Il serait aussi question de distinguer l’acte de traduire et la traduction en tant que discipline, régie par des théories; la traductologie appliquée permet entre autres de cibler, à l’aide de méthodes scientifiques, quels facteurs influencent la traduction.

L’étudiant devrait aussi arriver à distinguer le processus d’adaptation réalisé par un chercheur de la description que celui-ci en fait. Introduite dans le corps du deuxième article, la G-PA a été élaborée spécialement pour différencier les multiples descriptions de processus d’adaptation.

Finalement, le dernier segment du premier bloc consisterait à distinguer, d’une part, les stratégies du processus d’adaptation destinées à favoriser la négociation de l’équivalence et, d’autre part, les stratégies destinées à la mesure de l’équivalence. Il faudrait, pour y arriver, aborder l’épineuse question des paradigmes en science. L’aborder par le biais des stratégies faciliterait la tâche en cela que celles-ci sont concrètes, appliquées, en comparaison aux paradigmes eux-mêmes. Faire le parallèle avec les méthodes quantitatives et qualitatives rendrait également l’expérience beaucoup plus concrète, et, de surcroît, pertinente : l’objectif général de la

formation serait d'amener l'étudiant à comprendre que le processus d'adaptation puise sa légitimité aussi bien dans le positivisme (méthodes quantitatives) que dans le constructivisme (méthodes qualitatives).

Deuxième bloc. Le deuxième bloc consisterait essentiellement en des exercices pratiques, pour permettre à l'étudiant de consolider les principes théoriques vus dans le premier bloc aussi bien que pour le préparer à adapter des tests de manière aussi efficace que professionnelle.

Un premier exercice servirait à le familiariser avec la recherche par mots-clés dans les bases de données. Quels mots-clés utiliser lorsqu'on veut obtenir de l'information à propos du processus d'adaptation, soit les PA-TH? Et lesquels utiliser lorsqu'on cherche un test adapté en particulier, soit un PA-AP? Les tentatives de l'étudiant vont lui permettre de réfléchir sur la démarche et de discuter avec le groupe à propos de ce qui semble être les méthodes et les mots-clés les plus et les moins efficaces.

Un autre exercice viserait à ce que l'étudiant détermine le type d'un certain nombre de descriptions de processus d'adaptation. De longueur variée, les descriptions véhiculent les présupposés de leur auteur à propos du processus d'adaptation : juge-t-il la traduction bonne a priori? Considère-t-il que l'équivalence soit uniquement une question de mesure? Y a-t-il une place pour la discussion et la subjectivité dans le processus décrit, à savoir : est-ce que l'équivalence a été négociée dans une certaine mesure? J'estime que ces présupposés sont d'une importance capitale le moment venu de juger de la légitimité des tests adaptés, ce sur quoi la grande majorité des doctorants en psychologie, une fois diplômés, auront à trancher. Ils devront le faire, d'une part, en tant qu'utilisateurs de tests, mais aussi, d'autre part, en tant que chercheurs. Ils devront également déterminer à quels présupposés ils adhèrent; en effet, plusieurs auront la responsabilité de diriger au moins un processus d'adaptation pendant leur carrière.

Le dernier exercice serait de planifier un processus d'adaptation de A à Z. Une mise en contexte serait offerte à l'étudiant ou au groupe d'étudiants. Il serait question de déterminer quelles stratégies utilisées et de défendre ses choix. Pour y arriver, l'étudiant ou le groupe d'étudiants devrait répondre à une série de questions telles que :

- Prévoit-on recruter des traducteurs ou utiliser uniquement les membres de l'équipe de chercheur? Pourquoi?
- S'il y a un recrutement de traducteur, prévoit-on que ceux-ci soient des traducteurs professionnels ou amateurs? Pourquoi?
- Combien de versions traduites prévoit-on produire et pourquoi?
- Veut-on procéder à une traduction inversée? Le cas échéant, que vise-t-on avec cette stratégie?
- Prévoit-on entrer en contact avec le créateur du test original? Si oui, quel rôle aura-t-il dans le processus d'adaptation?
- Qui formera le comité de révision et pourquoi? Est-ce qu'il incorporera un ou plusieurs des traducteurs et pourquoi?
- Le test adapté servira-t-il à faire des comparaisons entre différentes populations? Si oui, que prévoit-on faire lors du processus d'adaptation par rapport à ce besoin? Si non, quelles données seront employées pour analyser les résultats des futurs participants de la population cible?
- Prévoit-on publier le processus d'adaptation, et, le cas échéant, pour quelle(s) raison(s) et dans quelle revue?
- Évidemment, un budget pourrait être alloué pour ajouter plus de réalisme à la réflexion. Quelles dépenses favoriser et pourquoi? Est-ce qu'engager un traducteur professionnel importe plus que de faire une révision externe? Pourrait-on recruter des étudiants au baccalauréat en traduction au lieu d'un traducteur professionnel? Est-ce que la révision externe pourrait se faire via internet et pourquoi? Combien de temps prévoit-on investir dans la révision interne et combien de rencontres sont prévues à cet égard?

Fondements empiriques de la psychologie. J'estime que les deux blocs de trois heures décrits pourraient être intégrés au cours obligatoire *Fondements empiriques de la psychologie* offert à l'Université Laval aux doctorants en psychologie.

Quand j'ai suivi le cours en 2013, il était essentiellement question de la pratique de la psychologie basée sur les données probantes. Comme je l'ai expliqué précédemment en me basant sur les observations d'Eisner (2003), une connaissance empirique réfère à l'expérience

sensible : « [T]o be experienced the senses must be engaged with qualities, even when the qualities in question are imaginatives » (p. 20). Sont imaginaires, et donc indirectes les qualités qui découlent d'une mesure, tels les résultats à un test psychométrique. Il faut, pour que celles-ci demeurent valides, justifier la pertinence du test utilisé et connaître les mécanismes théoriques qui leur permettent d'être effectives. Pour ainsi dire, les fondements empiriques de la psychologie, de même que ceux de toute science, sont soit directement soit indirectement basés sur l'expérience des sens.

Ceci dit, avancer que les données probantes sont un fondement empirique revient à emprunter un raccourci où le moyen et le résultat peuvent être confondus. Le scientifique fait confiance aux données probantes (le résultat) dans la mesure où il fait confiance aux instruments (le moyen) qui en permettent l'observation. Une des exigences de l'Ordre des psychologues du Québec est que les doctorants doivent acquérir « les bases scientifiques de la discipline » (2015, p. 15). Or, comprendre le test psychométrique (bases scientifiques) ne réfère pas à l'utilisation ni à l'interprétation des résultats et des données dites probants (bases techniques). Une telle compréhension se situe en amont de la technique, à la jonction de la psychométrie, de la psychologie cognitive, de la linguistique et de la philosophie des sciences, un carrefour où les paradigmes méritent d'être discutés.

Finalement, l'Ordre avance que le futur psychologue « doit être en mesure d'appliquer de façon rigoureuse et cohérente les méthodes de recherche de façon à apporter une contribution à la pratique professionnelle » (2015, p. 15). Parce que le processus d'adaptation rallie méthodes quantitatives et qualitatives, il m'apparaît comme un excellent vecteur d'apprentissage et de discrimination de la méthode. Nous l'aurons compris, ces différentes méthodes sous-entendent des ontologies distinctes. Définir des ontologies à un étudiant et les comparer entre elles constitue une tâche ardue. Certes, les ontologies comme les paradigmes sont des concepts complexes. En réponse à cet épineux problème, je partage la citation de Stephen Jay Gould (1996, p. 20), qui a certainement appris à mettre l'accent sur le particulier pour embrasser le général :

If I have learned one thing as a monthly essayist for more than twenty years, I have come to understand the power of treating generalities by particulars. It is no use writing a book on “the meaning of life” (though we

all long to know the answers to such great questions, while rightly suspecting that true solutions do not exist!). But an essay on “the meaning of 0.400 hitting in baseball” can reach a genuine conclusion with surprisingly extensive relevance to such broad topics as the nature of trends, the meaning of excellence, and even (believe it nor not) the constitution of natural reality. You have to sneak up on generalities, not assault them head-on. One of my favorite lines, from G. K. Chesterton, proclaims: “Art is limitation; the essence of every picture is the frame.”

Références

- Achim, J. (2006). *Pour une exploration de la perspective de la maternité dans un contexte de troubles des conduites alimentaires*. Université de Montréal, Montréal, Québec, Canada.
- Acquadro, C., Conway, K., Hareendran, A., & Aaronson, N. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health*, 11(3), 509-521.
- Aimé, A. (2002). *Analyse des aspects psychosociaux reliés aux troubles des conduites alimentaires*. (Ph. D.), Université Laval, Québec, Québec, Canada.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*: American Educational Research Association.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2003[1999]). *Normes de pratique du testing en psychologie et en éducation [Standards for educational and psychological testing]* (G. Sarrazin, Trans.): American Educational Research Association.
- Anastasi, A. (1994 [1982]). *Introduction à la psychométrie [Psychological Testing, 5th edition]* (F. Gagné, Trans.). Montréal, Québec, Canada: Guérin universitaire.
- Angel, R. J. (2013). After Babel: Language and the fundamental challenges of comparative aging research. *Journal of cross-cultural gerontology*, 28(3), 223-238.
- Arffman, I. (2013). Problems and issues in translating international educational achievement tests. *Educational Measurement: Issues and Practice*, 32(2), 2-14.
- Arnold, B. R., & Matus, Y. E. (2000). Test translation and cultural equivalence methodologies for use with diverse populations. *Handbook of multicultural mental health: Assessment and treatment of diverse populations*, 121-136.
- Audet-Lapointe, M. (2004). *Le processus d'accommodation (coping) de l'enfant et de l'adolescent confrontés à des situations stressantes : stress de la vie quotidienne ou leucémie*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Ballard, M. (1998). Les "mauvaises lectures" : Étude du processus de compréhension (chapitre 2). In J. Delisle & H. Lee-Jahnke (Eds.), *Enseignement de la traduction et traduction dans l'enseignement* (pp. 242). Ottawa, Canada: Les Presses de l'Université d'Ottawa.
- Bardin, L. (1977). *L'analyse de contenu*. Paris, France: Presses universitaires de France.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186-3191.
- Beaulieu, G. (2012). *Développer le soutien aux besoins psychologiques des employés : évaluation des effets affectifs, comportementaux et motivationnels d'une formation destinée aux gestionnaires*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Beaupré, M. G., King, S., Bauer, I., Hess, U., Debruille, J. B., & Corrigan, P. W. (2002). Validation of French- and English-Canadian Versions of the Social Cue Recognition Test. *Can J Psychiatry*, 47(1), 81-85.
- Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions* (Vol. 133): Sage.
- Béland, M., Chabot, K., Goulet Gervais, L., Morin, A. J. S., & Gosselin, P. (2012). Évaluation de la peur de l'accouchement. Validation et adaptation française d'une échelle mesurant la peur de l'accouchement. *L'Encéphale*, 38, 336-344.

- Bélanger, M. (2011). *Efficacité du programme mieux-être "wellness" sur la gestion du poids, l'observance à la médication et la qualité de vie chez les personnes ayant un trouble psychotique ou de l'humeur.* (D.Psy.), Université Laval, Québec, Québec, Canada.
- Belleville, G. (2006). *Évaluation d'une bibliothérapie de l'insomnie avec sevrage des hypnotiques assisté chez des adultes souffrant d'insomnie.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Bergeron, C. (2006). *Traduction et validation du Worry About Victimization (WAV) auprès d'une population âgée francophone.* (M.A.), Université de Sherbrooke, Sherbrooke, Canada.
- Berry, J. W., Poortinga, Y. H., Breugelmans, S. M., Chasiotis, A., & Sam, D. L. (2011). *Cross-cultural psychology: Research and applications (3rd edition)*: Cambridge University Press.
- Birmes, P., Brunet, A., Benoit, M., Defer, S., Hatton, L., Sztulman, H., & Schmitt, L. (2005). Validation of the Peritraumatic Dissociative Experiences Questionnaire self-report version in two samples of French-speaking individuals exposed to trauma. *European Psychiatry*, 20, 145-115#.
- Boislard Pépin, M. (2010). *Précocité sexuelle et comportement sexuel à risque à l'adolescence : Étude longitudinale des facteurs individuels familiaux, dans le groupe d'amis et contextuels associés.* (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Boisseau, É. (2010). *La théorie de l'esprit chez les schizophrènes paranoïdes.* (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Boisvert, N. (2006). *Validation de l'Échelle de motivation du client face à la thérapie (EMCT), version française du Client Motivation for Therapy Scale (CMOTS).* (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Bolaños-Medina, A., & González-Ruiz, V. (2012). Deconstructing the Translation of Psychological Tests. *Meta: Journal des traducteurs/Translators' Journal*, 57(3), 715-739.
- Bouchard, J. (2007). *Validation de la version française du Aggression Questionnaire auprès de deux échantillons : étudiants universitaires (étude 1) et adultes non-recrutés en milieu universitaire (étude 2).* (Maîtrise en psychologie), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Bouffard, A. (2010). *Variabilités familiale et individuelle de la sensibilité paternelle et maternelle : corrélats contextuels, individuels et liés à l'enfant.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Bouffard, T., Seidah, A., McIntyre, M., Boivin, M., Vezeau, C., & Cantin, S. (2002). Mesure de l'estime de soi à l'adolescence : version canadienne-française du Self-Perception Profile for Adolescents de Harter. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(3), 158-162.
- Bougie, E. (2005). *The Cultural Narrative of Francophone and Anglophone Quebecers and their perceptions of temporal relative deprivation: Links with esteem and well-being.* (Ph. D.), Université McGill, Montréal, Québec, Canada.
- Bradley, C. (1994). Translation of questionnaires for use in different languages and cultures. In C. Bradley (Ed.), *Handbook of Psychology and Diabetes: a guide to psychological measurement in diabetes research and practice* (pp. 423). Swiss: Harwood Academic Publisher.

- Brassard, A. (2003). *Syndrome d'apnées du sommeil et déficits cognitifs associés chez des patients avec ischémie cérébrale transitoire*. (M.S), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Brin, J. (2011). *Adaptation et validation française du Narcissistic Personality Inventory*. (D.Psy.), Université Laval, Québec, Canada.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of cross-cultural psychology*, 1(3), 185-216.
- Brislin, R. W. (1976). Comparative Research Methodology: Cross-Cultural Studies. *International Journal of Psychology*, 11(3), 215-229. doi:10.1080/00207597608247359
- Brislin, R. W. (1980). Cross-cultural research methods *Environment and culture* (pp. 47-82): Springer.
- Brislin, R. W. (Producer). (1986). The wording and translation of research instruments. *Field methods in cross-cultural research*.
- Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). *Cross-Cultural: Research Methods* (Vol. 11): Year Book Medical Pub.
- Brousseau, M. (2011). *He said, She said : An examination of sexual coercion from the perspective of mixed-sex couples*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Brown, J. B., Schmidt, G., Lent, B., Sas, G., & Lemelin, J. (2001). Épreuves de validation et de fiabilité d'un instrument de mesure français. *Canadian Family Physician*, 47, 988-995.
- Brunet, A., St-Hilaire, A., Jehel, L., & King, S. (2003). Validation of a French Version of the Impact of Event Scale-Revised. *Can J Psychiatry*, 48(1), 56-61.
- Butcher, J. N. (1996). Translation and adaptation of the MMPI-2 for international use. In J. N. Butcher (Ed.), *International adaptation of the MMPI-2: research and clinical application* (pp. 26-43). Minneapolis, Minn., USA: University of Minnesota Press.
- Carlson, E. D. (2000). A case study in translation methodology using the Health-Promotion Lifestyle Profile II. *Public health nursing*, 17(1), 61-70.
- Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: Issues and techniques. *Journal of advanced nursing*, 58(4), 386-395.
- Chapman, D. W., & Carter, J. F. (1979). Translation procedures for the cross cultural use of measurement instruments. *Educational Evaluation and Policy Analysis*, 1(3), 71-76.
- Charbonneau, S. (2009). *L'estimation cognitive : Analyse des fonctions cognitives sous-jacentes et étude de l'impact du vieillissement normal et de la démence de type Alzheimer*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Chávez, L. M., & Canino, G. (2005). *Toolkit on translating and adapting instruments*. Retrieved from http://www.hsri.org/files/uploads/publications/PN54_Translating_and_Adapting.pdf.
- Chevrier, N. (2009). *Adaptation québécoise de l'Oldenborg Burnout Inventory (OLBI)*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Cordeau, D. (2007). *Simplicité volontaire et bien-être psychologique - Chapitre 2 : Traduction et validation québécoise de la Mesure de caractère mercantile*. (Ph. D.), Université du Québec à Trois-Rivières, Trois-Rivières, Québec, Canada.
- Cordeau, D., & Dubé, M. (2008). L'échelle de simplicité volontaire : une validation en français. *Les cahiers internationaux de psychologie sociale*, 79(3), 33-46.
- Cossette, M. (2008). *Styles de régulation émotionnelle des employés de service : Antécédents organisationnels, motivation et conséquences sur la santé psychologique*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.

- Côté, G., Gosselin, P., & Dagenais, I. (2013). Évaluation multidimensionnelle de la régulation des émotions : propriétés psychométriques d'une version francophone du Difficulties in Emotion Regulation Scale. *Journal de thérapie comportementale et cognitive*, 23, 63-72.
- Coutu, S., Dubeau, D., Provost, M. A., Royer, N., & Lavigne, S. (2002). Validation de la version française du questionnaire Coping with Children's Negative Emotions Scale-CCNES. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(4), 230-234.
- Couture, A. (2014). *Étude pilote d'implantation et d'efficacité du CRAFT au Québec*. (D.Psy.), Université Laval, Québec, Québec, Canada.
- Couture, S., Lecours, S., Beaulieu-Pelletier, G., Philippe, F. L., & Strychar, I. (2010). French Adaptation of the Eating Disorder Recovery Self-Efficacy Questionnaire (EDRSQ): Psychometric Properties and Conceptual Overview. *European Eating Disorders Review*, 18(3), 234-243.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281.
- D'Amico, É. (2010). *Dévoilement de l'orientation sexuelle à la famille d'origine et adaptation des jeunes gais, lesbiennes et bisexuels*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Da Silva Luis, R. (2014). *Adaptation et validation du Pathological Narcissism Inventory (PNI)*. (D. Psy.), Université Laval, Québec, Canada.
- Daigneault, I., Dion, J., Hébert, M., McDuff, P., & Collin-Vézina, D. (2013). Psychometric properties of the Child and Youth Resilience Measure (CYRM-28) among samples of French Canadian youth. *Child Abuse & Neglect*, 37, 160-171.
- Daley, M. (2007). *The impact and economic costs of insomnia : health-care utilisation, work function and accidents*. (Ph. D.), Université Laval, Québec, Québec, Canada.
- Daniel, C. (2008). *Évaluation des effets d'un programme de développement de l'empathie chez des enfants présentant des difficultés relationnelles*. (D.Psy.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Darbelnet, J. (1966). Réflexions sur la formation générale du traducteur. *Meta: Journal des traducteurs/Translators' Journal*, 11(4), 155-160.
- De Benedictis, L., Dumais, A., Stafford, M. C., Côté, G., & Lesage, A. (2012). Factor analysis of the French version of the shorter 12-item Perception of Aggression Scale (POAS) and of a new modified version of the Overt Aggression Scale (MOAS). *Journal of Psychiatric and Mental Health Nursing*, 19, 875-880.
- Delisle, J. (2003). *La traduction raisonnée : Manuel d'initiation à la traduction professionnelle de l'anglais vers le français* (2e édition). Ottawa, Canada: Les Presses de l'Université d'Ottawa.
- Delisle, J. (2011). *Le logiciel SAVR comme outil descriptif des fonctions exécutives de haut niveau dans le Trouble déficitaire de l'attention avec hyperactivité chez l'adulte*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Demers, L., Monette, M., Descent, M., Jutai, J., & Wolfson, C. (2002). The Psychosocial Impact of Assistive Devices Scale (PIADS): Translation and preliminary psychometric evaluation of a Canadian-French version. *Quality of Life Research*, 11, 583-592.
- Denis, P. (2009). *Étude de la capacité de l'aptitude cognitive générale (ACG) évaluée avec et sans contrainte temporelle et des facettes de la personnalité à prédire différentes dimensions de la performance en emploi dans un contexte de sélection de personnel*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.

- Deshaises, C., Toupin, J., & Déry, M. (2009). Validation de l'Échelle d'évaluation des traits antisociaux. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 41(1), 45-50.
- DiTommaso, E., Turbide, J., Poulin, C., & Robinson, B. (2007). L'Échelle de solitude sociale et émotionnelle (ÉSSE): A French-Canadian Adaptation of the Social and Emotional Loneliness Scale for Adults. *Social Behavior and Personality*, 35(3), 339-350.
- Dubé, S. (2011). *Validation transculturelle en langue française du Political Skill Inventory*. (Ph. D.), Université de Sherbrooke, Sherbrooke, Canada.
- Durkeim, E. (1895[1992]). *Les règles de la méthode sociologique*. Paris, France: PUF, Quadrige.
- Eisner, E. W. (2003). On the art and science of qualitative research in psychology. In P. M. Camic, J. E. Rhodes, & L. Yardley (Eds.), *Qualitative research in psychology: Expanding perspectives in methodology and design* (pp. 17-29). Washington, DC, USA: American Psychological Association.
- Emard, M. (2010). *Les contributions parentales à la compréhension des émotions des enfants en début de scolarisation*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Engelberg, M. (2003). *Binge Antecedents and Consequences in Bulimic Syndromes: An Examination of the Roles of Dietary Restraint, Affect, and Dissociation*. (Ph. D.), Université McGill, Montréal, Québec, Canada.
- Epstein, J., Santo, R. M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of clinical epidemiology*, 68(4), 435-441.
- Erçikan, K., & Lyons-Thomas, J. (2013). Adapting tests for use in other languages and cultures. In K. F. Geisinger (Ed.), *Handbook of Testing and Assessment in Psychology: Testing and Assessment in School Psychology and Education* (Vol. 3, pp. 545-569).
- Eremenco, S. L., Cella, D., & Arnold, B. J. (2005). A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Evaluation & the health professions*, 28(2), 212-232.
- Fikretoglu, D., Brunet, A., Poundja, J., Guay, S., & Pedlar, D. (2006). Validation of the Deployment Risk and Resilience Inventory in French-Canadian Veterans: Findings on the Relation Between Deployment Experiences and Postdeployment Health. *Can J Psychiatry*, 51(12), 755-763.
- Filiatrault, J., Demers, L., Parisien, M., Lorthios-Guilledroit, A., Kaegi, C., Ménard, I., . . . St-Laurent, C. (2014). Development and validation of a French Canadian version of the falls Behavioral (FaB) Scale. *Disabil Rehabil*, 1-6.
- Finn, K. (2012). *Le perfectionnisme et sa relation avec le contingentement universitaire et l'anxiété de performance*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Forget, H., & Robichaud, C. (2002). Validation d'une traduction française d'un instrument de mesure des réponses affectives à la première relation sexuelle avec pénétration. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(3), 151-157.
- Fortin, M. (2009). *Prévention de la récidive et non-adhérence aux programmes d'intervention obligatoires chez les contrevenants de la conduite avec facultés affaiblies par l'alcool*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.

- Franzosi, R. (2004). Content Analysis. In M. S. Lewis-Beck, A. E. Bryman, & T. F. Liao (Eds.), *The SAGE Encyclopedia of Social Science Research Methods* (Vol. 1, pp. 187-189). Thousand Oaks: SAGE Publications, Inc.
- French Bourgeois, L. (2011). *L'influence des normes sociales sur le comportement : un processus médié par l'élaboration de stratégies*. (M. Sc.), Université de Montréal, Montréal, Québec, Canada.
- French, D. J., Noël, M., Vigneau, F., French, J. A., Cyr, C. P., & Thomas Evans, R. (2005). L'Échelle de dramatisation face à la douleur PCS-CF. Adaptation canadienne en langue française de l'échelle « Pain Catastrophizing Scale ». *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 37(3), 181-192.
- Gagné, J. (2010). *L'association entre la psychopathie et les pratiques parentales dans une population non criminelle*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Gagnon, C. (2013). *Traduction et validation du O'Brien Multiphasic Narcissism Inventory (OMNI)*. (D.Psy.), Université Laval, Québec, Canada.
- Gagnon, M. (2010). *Évaluation du perfectionnisme prescrit par autrui chez des patients souffrant de maladies cardiovasculaires du Saguenay-Lac-St-Jean et de ses relations avec l'activité physique, les habitudes alimentaires et le bien-être psychologique*. (D.Psy.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Gagnon, V. (2004). *Étude comparative des stratégies d'adaptation de rumination et de distraction chez les adultes et les adolescents dépressifs*. (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Gardner, H. (1993[1985]). *The mind's new science: A history of the cognitive revolution [Histoire de la révolution cognitive : la nouvelle science de l'esprit]* (J.-L. Peytavin, Trans.). Paris, France: Payot.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological assessment*, 6(4), 304.
- Geisinger, K. F. (2003). Testing and Assessment in Cross-Cultural Psychology. In J. R. Graham & J. A. Naglieri (Eds.), *Handbook of psychology - Assessment Psychology* (pp. 650). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Gierl, M. J., Rogers, W. T., & Klinger, D. A. (1999). Using statistical and judgmental reviews to identify and interpret translation differential item functioning. *Alberta Journal of Educational Research*, 45(4), 353.
- Giguère, J., Fortin, C., & Sabourin, S. (2006). Déterminants de la persistance conjugale chez des personnes vivant en première et en seconde union conjugale. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 38(2), 185-199.
- Gilbert, E. (2011). *La mémoire épisodique et le fonctionnement social dans les troubles du spectre autistique*. (Ph. D.), Université Laval, Québec, Québec, Canada.
- Green, C. D. (1992). Of Immortal Mythological Beasts Operationism in Psychology. *Theory & Psychology*, 2(3), 291-320.
- Gregoire, J., & Hambleton, R. K. (2009). Advances in test adaptation research: a special issue. *International Journal of Testing*, 9(2), 75-77.
- Grégoire, S., Baron, L., Ménard, J., & Lachance, L. (2013). The Authenticity Scale: Psychometric Properties of a French Translation and Exploration of Its Relationships With Personality and Well-Being. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, Advance online publication.
- doi:10.1037/a0030962

- Grimard, F. (2011). *L'anxiété sociale et la victimisation par les pairs à l'adolescence*. (Ph. D.), Université Laval, Québec, Québec, Canada.
- Guillemain, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of clinical epidemiology*, 46(12), 1417-1432.
- Gutt, E.-A. (2000). Translation as interlingual interpretive use. *The translation studies reader*, 376-396.
- Hambleton, R. K. (1993). *Technical standards for translating tests and establishing test score equivalence*. Paper presented at the 101st Annual Convention of the American Psychological Association. Toronto, Ontario, Canada.
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal*, 229-240.
- Hambleton, R. K. (2001). The next generation of the ITC Test Translation and Adaptation Guidelines. *European Journal of Psychological Assessment*, 17(3), 164.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. Merenda, & C. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (Vol. 1, pp. 17-38). Mahwah, NJ: Erlbaum.
- Hambleton, R. K., & de Jong, J. H. (2003). Advances in translating and adapting educational and psychological tests. *Language testing*, 20(2), 127-134.
- Hambleton, R. K., & Kanjee, A. (1995). Increasing the validity of cross-cultural assessments: Use of improved methods for test adaptations. *European Journal of Psychological Assessment*, 11(3), 147.
- Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*: Psychology Press.
- Hambleton, R. K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Association of Test Publishers*, 1(1), 1-13.
- Hambleton, R. K., & Patsula, L. (2000). Adapting Tests for Use in Multiple Languages and Cultures. Laboratory of Psychometric and Evaluative Research Report.
- Harkness, J. A. (2011). Translation. In S. R. Center (Ed.), Guidelines for best practice in cross-cultural surveys (3rd ed.). (pp. 1-35). Ann Arbor, MI: University of Michigan: Institute for Social Research. Retrieved from <http://www.ccsg.isr.umich.edu>.
- Harkness, J. A., & Schoua-Glusberg, A. (1998). Questionnaires in translation. *ZUMA-Nachrichten Spezial*, 3(1), 87-127.
- Hébert, M., Parent, N., & Daignault, I. V. (2007). The French-Canadian Version of the Self-Report Coping Scale: Estimates of the Reliability, Validity, and Development of a Short Form. *Measurement and Evaluation in Counseling and Development*, 40, 2-15.
- Herdman, M., Fox-Rushby, J., & Badia, X. (1997). 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research*, 6(3), 237-247.
- Hovington, C. (2002). *Les inondations de juillet 1996 au Saguenay : Les effets psychologiques durables chez les adultes jeunes et âgés*. (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Laval.
- Husky, M. M., Grondin, O. S., & Compagnone, P. D. (2004). Validation de la version française du questionnaire de Sociotropie-Autonomie de Beck et collègues. *Can J Psychiatry*, 49(12), 851-858.

- Jay Gould, S. (1996). *The Mismeasure of Man, Revised & Expanded*: W. W. Norton & Company.
- Jeanrie, C., & Bertrand, R. (1999). Translating tests with the International Test Commission's guidelines: Keeping validity in mind. *European Journal of Psychological Assessment*, 15(3), 277.
- Johnson, K. (2008). *L'orientation vers les buts : le contraste de deux modèles*. (M. Sc.), Université de Montréal, Montréal, Québec, Canada.
- Jones, P. S., Lee, J. W., Phillips, L. R., Zhang, X. E., & Jaceldo, K. B. (2001). An adaptation of Brislin's translation model for cross-cultural research. *Nurs Res*, 50(5), 300-304.
- Julien, D., Careau, Y., O'Connor, K. P., Bouvard, M., Rhéaume, J., Langlois, F., . . . Cottraux, J. (2008). Specificity of belief domains in OCD: Validation of the French version of the Obsessive Beliefs Questionnaire and a comparison across samples. *Journal of Anxiety Disorders*, 22, 1029-1041.
- Kaplan, A. (1943). Content analysis and the theory of signs. *Philosophy of Science*, 10(4), 230-247.
- Koller, M., Kantzer, V., Mear, I., Zarzar, K., Martin, M., Greimel, E., . . . Kulić, D. (2012). The process of reconciliation: evaluation of guidelines for translating quality-of-life questionnaires. *Expert review of pharmacoeconomics & outcomes research*, 12(2), 189-197.
- Koski, L., Xie, H., Konsztowicz, S., & Tetteh, R. (2010). French-English Cross-Linguistic Comparison and Diagnostic Impact of the AD-8 Dementia Screening Questionnaire in a Geriatric Assessment Clinic. *Dement Geriatr Cogn Disord*, 29, 265-274.
- Kuhn, T. S. (2012[1962]). *The structure of scientific revolutions*: University of Chicago press.
- L'Archevêque, A. (2009). *Étude exploratoire des contextes d'accès à la parentalité et des facteurs contribuant à l'intégration identitaire chez les pères gais*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Lacasse, A. (2004). *Peer sexual harassment among adolescents: A developmental perspective*. (Ph. D.), Université McGill, Montréal, Québec, Canada.
- Lachapelle, Y., Boisvert, D., & Leclerc, D. (2000). La traduction et la validation transculturelle de l'Échelle d'autodétermination de l'ARC. *Revue francophone de la déficience intellectuelle*, 70-74.
- Lafontaine, M. F., & Lussier, Y. (2003). Structure bidimensionnelle de l'attachement amoureux : Anxiété face à l'abandon et évitement de l'intimité. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 35(1), 56-60.
- Lafrenière, M. (2013). *Le rôle modérateur de la passion dans la relation entre la performance/le rehaussement du soi et l'adaptation psychologique*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Laliberté, S., & Earls, C. M. (2007). Traduction et validation en français du Situational Humor Response Questionnaire et du Coping Humor Scale. *Revue québécoise de psychologie*, 28(2), 213-232.
- Lamoureux-Hébert, M. (2006). *Les besoins de soutien des personnes présentant un retard mental manifestant des comportements problématiques*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Landry, T. (2010). *La dyspareunie chez les adolescents : Prévalence, caractéristiques et corrélats biopsychosociaux*. (Ph. D.), Université du Québec à Montréal, Montréal.
- Lauzier, M. (2010). Adaptation et validation canadienne-française de l'Échelle des conceptions implicites de l'intelligence (ECII). *Revue québécoise de psychologie*, 31(3), 189-206.

- Lauzier, M., Barrette, J., Lemyre, L., & Corneil, W. (2013). Validation francophone de l'Inventaire des facilitateurs de l'apprentissage organisationnel (IFAO) : étude comparative de l'apprentissage perçu au niveau des individus, des groupes et de l'organisation. *Revue européenne de psychologie appliquée*, 63, 353-362.
- Lauzier, M., & Haccoun, R. R. (2010). Validation canadienne-française de l'Échelle des styles d'orientation des buts (ESOB). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 42(2), 127-133.
- Laverdière, O. (2009). *La pathologie de la personnalité : contributions du tempérament, de l'attachement, des traumas et médiation de la compréhension émotionnelle - Premier article : French adaptation of the Mini-IPIP: A short measure of the Big Five.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Laverdière, O., Diguer, L., Gamache, D., & Evans, D. E. (2010). The French Adaptation of the Short Form of the Adult Temperament Questionnaire. *European Journal of Psychological Assessment*, 26(3), 212-219.
- Lavergne, A. G. (2003). *Les déterminants de l'intensité du deuil à la suite de la perte d'un animal de compagnie - Validation d'un instrument et étude de corrélats.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Lavigne, G. (2011). *L'influence des orientations du besoin d'affiliation sur l'adaptation intrapersonnelle et interpersonnelle.* (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Leblanc, G., Marchand, A., Khouzam, C., & Guay, S. (2007). Validation canadienne-française du Questionnaire sur la peur de l'intimité. *Revue québécoise de psychologie*, 28(2), 277-292.
- Leboeuf, G., & Losier, G. F. (2012). L'échelle d'orientation contingente extrinsèque (ÉOCE) : Traduction et validation francophone du Extrinsic Contingency Focus Scale (ECFS). *Revue européenne de psychologie appliquée*, 62, 137-145.
- Lecavalier, L., & Tassé, M. J. (2001). Traduction et adaptation transculturelle du Reiss Screen for Maladaptive Behavior. *Revue francophone de la déficience intellectuelle*, 12(1), 31-44.
- Lekes, N. (2012). *Self-growth, close relationships, and community contribution: Exploring the development of intrinsic value priorities and their influence on well-being.* (Ph. D.), Université McGill, Montréal, Québec, Canada.
- Lemelin, J. P., Tarabulsy, G. M., Provost, M. A., Fournier, M., Robitaille, J., Hémond, I., & Tessier, R. (2007). Le Questionnaire d'évaluation du comportement de l'enfant : version canadienne française du Toddler Behavior Assessment Questionnaire. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 39(4), 291-300.
- Lemieux, R. (2009). *Étude de relations entre la dissociation, le fonctionnement réflexif et les difficultés d'adaptation chez des enfants abusés sexuellement.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Lepage, J. F., Lortie, M., Taschereau-Dumouchel, V., & Théoret, H. (2009). Validation of French-Canadian Versions of the Empathy Quotient and Autism Spectrum Quotient. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 41(4), 272-276.
- Lessard, M. (2012). *Interventions psychologiques brèves du Trouble panique chez les individus qui se présentent au service d'urgence pour des douleurs thoraciques non cardiaques :*

- Efficacité et recension des écrits.* (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: a methods review. *Journal of advanced nursing*, 48(2), 175-186.
- Martin, A., & Gosselin, P. (2012). Propriétés psychométriques de l'adaptation francophone d'une mesure de symptômes des troubles anxieux auprès d'enfants et d'adolescents (SCARED-R). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 44(1), 70-76. doi:10.1037/a0023103
- Massoubre, C., Lang, F., Jaeger, B., Jullien, M., & Pellet, J. (2002). [The translation of questionnaires and of tests: techniques and problems]. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, 47(1), 61-67.
- Matthews, G. G. (2003). *L'équivalence en traduction juridique: analyse des traductions au sein de l'Accord de libre-échange nord-américain (ALENA)*. Université Laval.
- Maxwell, B. (1996). *Translation and cultural adaptation of the survey instruments*. Retrieved from <http://files.eric.ed.gov/fulltext/ED406418.pdf#page=159>.
- Mayer-Renaud, J. (2008). *Contribution de l'intimité à l'effet du soutien social dans la relation entre l'alexithymie et la dépression*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Ménard, J. (2008). *Authenticité et bien-être au travail : étude auprès de gestionnaires*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Méthot, N. (1998). *L'établissement d'une méthode jugementale de traduction de tests*. (Master), Université Laval, Québec, Canada.
- Méthot, S., & Morin, D. (2004). Traduction d'un instrument de dépistage des symptômes affectifs et anxieux chez les adultes ayant un retard mentale. *Revue francophone de la déficience intellectuelle*, 15(1), 57-73.
- Morin, M. (2007). *La créativité dans l'activité mentale du thérapeute*. (M. Ps.), Université Laval, Québec, Québec, Canada.
- Nguyễn-Duy, V., & Luckerhoff, J. (2006). Constructivisme/positivisme: où en sommes-nous avec cette opposition? *Les actes*, 4.
- Nida, E. A. (1964). *Toward a science of translating: with special reference to principles and procedures involved in Bible translating*: Brill Archive.
- Nida, E. A. (1969). Science of translation. *Language*, 483-498.
- O'Rourke, N., & Cappeliez, P. (2003). Validation d'une mesure de réponses biaisées à propos de la relation conjugale : l'Échelle d'embellissement conjugal. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 35(1), 45-49.
- Ordre des psychologues du Québec. (2015). *Manuel d'évaluation des programmes de doctorat en psychologie*. Retrieved from
- Orendain Tremear, M. (2010). *Les relations d'objet, la dualité anaclitique-introjective et leur relation avec les symptômes dépressifs*. (Ph. D.), Université Laval, Québec, Québec, Canada.
- Pan, Y., & De La Puente, M. (2005). *Census Bureau guideline for the translation of data collection instruments and supporting materials: Documentation on how the guideline was developed*. Retrieved from <https://www.census.gov/srd/papers/pdf/rsm2005-06.pdf>.
- Paquette, D., Laporte, L., Bigras, M., & Zoccolillo, M. (2004). Validation de la version française du CTQ et prévalence de l'histoire de maltraitance. *Santé mentale au Québec*, 29(1), 201-220.

- Paradis, A. (2012). *Maltraitances dans l'enfance, problèmes interpersonnels, et violence dans les relations de couple : Différences de trajectoires entre hommes et femmes.* (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Payer, M. (2012). *L'état de stress post-traumatique-complexe : un concept novateur à explorer.* (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Paynter, R. A. (2009). Evidence-based research in the applied social sciences. *Reference Services Review*, 37(4), 435-450.
- Pelletier, O. (2007). *Troubles de la personnalité et jeu pathologique : comorbidité et prédicteurs d'abandon du traitement.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Perozzo, C. (2007). *Le changement de l'identité sociale : une question de discrimination ou de privation relative?* (M.Sc.), Université de Montréal, Montréal, Québec, Canada.
- Plante-Jean, M. (2012). *Facteurs prédictifs de la dépression chez les gens souffrant d'anxiété sociale.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Plante, N., Daigle, M. S., Gaumont, C., Charbonneau, L., Gibbs, J., & Barriga, A. (2012). Validation of the 'How I Think Questionnaire' in a Population of French-speaking Adolescents with Externalizing Behaviors. *Behavioral Sciences and the Law*, 30, 196–210.
- Poirier, S., Bureau, V., Lehoux, C., Bouchard, R. H., Maziade, M., Pelletier, S., . . . Roy, M. A. (2004). A Factor Analysis of the Strauss and Carpenter Revised Outcome Criteria Scale: A Validation of the French Translation. *The Journal of nervous and mental disease*, 192(12), 864-867.
- Potvin, M. (2011). *Évaluation écologique et rééducation de la mémoire prospective chez des patients ayant subi un traumatisme craniocérébral.* (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Poulin-Vallières, I. (2004). *Étude sur le comportement, le sommeil et la mémoire chez les jeunes âgés entre 7 et 15 ans, victimes d'un traumatisme crânien léger.* (M.S.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Poulin, V., & Desrosiers, J. (2010). Validation française de questionnaires de participation sociale auprès d'aînés en situation d'incapacités. *Canadian Journal of Occupational Therapy*, 77(3), 159-166.
- Prieto, A. J. (1992). A method for translation of instruments to other languages. *Adult Education Quarterly*, 43(1), 1-14.
- Pym, A. (2009). *Exploring translation theories.* New York, United States: Routledge.
- Quesnel, C. (2007). *Évaluation longitudinale du fonctionnement cognitif de femmes traitées pour un cancer du sein non métastatique.* (Ph. D.), Université Laval, Québec, Québec, Canada.
- Rabiau, M. (2006). *Balancing Temptations and Health Goals: The Role of Compensatory Health Beliefs.* (Ph. D.), Université McGill, Montréal, Québec, Canada.
- Radomsky, A. S., Ashbaugh, A. R., Saxe, M. L., Ouimet, A. J., Golden, E. R., Lavoie, S. L., & O'Connor, K. P. (2006). Psychometric Properties of the French and English Versions of the Social Phobia Inventory. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 38(4), 354-360.
- Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological reports*.
- René de Cotret, F., Lanza, Y., & Francoeur, A. (2016). *How to adapt psychometric tests: Strategies recommended by the scientific community. Manuscrit soumis pour publication.*
- Ricard-St-Aubin, J. S., Philippe, F. L., Beaulieu-Pelletier, G., & Lecours, S. (2010). Validation francophone de l'Échelle des émotions différentielles IV (EED-IV). *Revue européenne de psychologie appliquée*, 60, 41-53.

- Richer, I., & Bergeron, J. (2012). Differentiating risky and aggressive driving: Further support of the internal validity of the Dula Dangerous Driving Index. *Accident Analysis and Prevention*, 45, 620– 627.
- Rios, J. A., & Sireci, S. G. (2014). Guidelines versus practices in cross-lingual assessment: A disconcerting disconnect. *International Journal of Testing*, 14(4), 289-312.
- Roberge, M. (2007). *État de stress aigu et état de stress post-traumatique après un infarctus du myocarde : Prévalence et facteurs associés*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Roberge, P., Marchand, L., Grenier, S., & Marchand, A. (2003). Validation canadienne-française de l'Échelle d'évaluation de la panique et de l'agoraphobie. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 35(1), 61-66.
- Sandelowski, M. (2004). Qualitative Research. In M. S. Lewis-Beck, A. E. Bryman, & T. F. Liao (Eds.), *The SAGE Encyclopedia of Social Science Research Methods* (Vol. 1, pp. 893- 894). Thousand Oaks: SAGE Publications, Inc.
- Seidah, A., Bouffard, T., & Vezeau, C. (2002). La mesure du perfectionnisme : validation canadienne-française du Positive and Negative Perfectionism Scale. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(3), 168-171.
- Sergerie, M. (2008). *Analyse dyadique des relations cybermentoriales de la ressource Academos*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Simard, S. (2008). *Vers une conceptualisation multidimensionnelle de la peur de la récidive du cancer : évaluation, nature des pensées intrusives et comorbidité psychiatrique*. (Ph. D.), Université Laval, Québec, Québec, Canada.
- Sinaiko, H. W., & Brislin, R. W. (1973). Evaluating language translations: Experiments on three assessment methods. *Journal of Applied Psychology*, 57(3), 328.
- Sireci, S. G., Patsula, L., & Hambleton, R. K. (2005). Statistical methods for identifying flaws in the test adaptation process. *Adapting educational and psychological tests for cross-cultural assessment*, 93-115.
- Skibińska, E. (2007). La retraduction, manifestation de la subjectivité du traducteur. *Doletiana: revista de traducció, literatura i art*(1), 000-000.
- Solano-Flores, G., Backhoff, E., & Contreras-Niño, L. Á. (2009). Theory of test translation error. *International Journal of Testing*, 9(2), 78-91.
- Sperber, A. D., Devellis, R. F., & Boehlecke, B. (1994). Cross-cultural translation methodology and validation. *Journal of cross-cultural psychology*, 25(4), 501-524.
- Spielberger, C. D., & Sharma, S. (1976). Cross-cultural measurement of anxiety *Cross-cultural anxiety* (Vol. 1, pp. 13-25).
- Taillon, A. (2011). *Peur d'une dysmorphie corporelle : réflexion critique sur les traitements psychologiques et évaluation d'une thérapie basée sur les inférences*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Tassé, M. J., & Craig, E. M. (1999). Critical issues in the cross-cultural assessment of adaptive behavior *Adaptive behavior and its measurement: Implications for the field of mental retardation* (pp. 161-184).
- Temimi, N. (2011). *Le rôle des dimensions culturelles de Hofstede en regard du fonctionnement interne des équipes de travail pluriculturelles*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Thibaudeau, P. (2006). *L'attachement parental et l'ajustement psychosocial des adolescents : validation d'un instrument de mesure et étude du rôle médiateur de l'estime de soi*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.

- Thibault, P. (2008). *Perception d'authenticité du sourire : Perspective interculturelle et développementale*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Toledano, R. (2006). *Development of the Sexual Arousal and Desire Inventory (SADI): A Multidimensional Scale of Subjective Sexual Arousal and Desire in Men and Women*. (Ph. D.), Université Concordia, Montréal, Québec, Canada.
- Toury, G. (2012). *Descriptive Translation Studies and beyond: Revised edition* (Vol. 100): John Benjamins Publishing.
- Tremblay, C. (2001). *Identité ethnique et maturité vocationnelle à l'adolescence : perspective développementale, contextuelle et interculturelle*. (Ph. D.), Université de Montréal, Montréal, Québec, Canada.
- Tremblay, S. (2009). *Engagement du père lors de la transition à devenir parent : rôle des cognitions paternelles et maternelles*. (Ph. D.), Université Laval, Québec, Québec, Canada.
- Trottier, C., Mageau, G., Trudel, P., & Halliwell, W. R. (2008). Validation de la version canadienne-française du Life Orientation Test-Revised. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 40(4), 238-243.
- Turcotte, S. (2013). *Traduction et validation du Relationship Patterns Questionnaire (RPQ) et étude pilote de ses liens avec le Personality Organization Diagnostic Form (PODF)*. (D.Psy.), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.
- Turgeon, L., & Chartrand, É. (2003). Reliability and Validity of the Revised Children's Manifest Anxiety Scale in a French-Canadian Sample. *Psychological assessment*, 15(3), 378-383. doi:10.1037/1040-3590.15.3.378
- Turgeon, M. E., Forget, J., & Bégin, J. (2013). Validation francophone et québécoise de la sous-échelle Weight-Related Teasing du Perception of Teasing Scale chez les enfants. *Journal de thérapie comportementale et cognitive*, 23, 52-62.
- Vaillancourt, L. (2009). *L'inconfort spatio-moteur et les attitudes envers la pratique d'activités physiques dans le Trouble panique avec ou sans agoraphobie*. (Ph. D.), Université du Québec à Montréal, Montréal, Québec, Canada.
- Valdivia Vazquez, J. A. (2014). *Test adaptation activities across languages and cultures*. (Ph. D.), Washington State University, Pullman, Washington, United States.
- Valentine, E., & Aubin, M.-C. (2004). *Stylistique différentielle et traduction*. Montréal, Québec, Canada: Sodilis.
- 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.
- Vallerand, R. J., & Halliwell, W. R. (1983). Vers une méthodologie de validation transculturelle de questionnaires psychologiques: Implications pour la psychologie du sport. *Canadian Journal of Applied Sport Sciences*, 8(1), 9-18.
- Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests. *European Psychologist*, 1(2), 89-99.
- Van de Vijver, F., & Leung, K. (1997). *Methods and data analysis for cross-cultural research* (Vol. 1): Sage.
- Van de Vijver, F., & Poortinga, Y. H. (2005). Conceptual and methodological issues in adapting tests. *Adapting educational and psychological tests for cross-cultural assessment*, 39-63.

- Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 54(2), 119-135.
- Van Widenfelt, B. M., Treffers, P. D., De Beurs, E., Siebelink, B. M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical child and family psychology review*, 8(2), 135-147.
- Vinay, J.-P., & Darbelnet, J. (1977[1954]). *Stylistique comparée du français et de l'anglais : Méthode de traduction (nouvelle édition revue et corrigée)* (Vol. 1). Montréal, Québec, Canada: Beauchemin.
- Vissandjée, B., & Dupéré, S. (2000). Culture, migration et instrument de mesure: défis incontournables. *Journal of International Migration and Integration/Revue de l'intégration et de la migration internationale*, 1(4), 477-492.
- Vreck, F. (1998). Les préalables de la traduction (chapitre 3). In J. Delisle & H. Lee-Jahnke (Eds.), *Enseignement de la traduction et traduction dans l'enseignement* (pp. 242). Ottawa, Canada: Les Presses de l'Université d'Ottawa.
- Wang, W. L., Lee, H. L., & Fetzer, S. J. (2006). Challenges and strategies of instrument translation. *West J Nurs Res*, 28(3), 310-321. doi:10.1177/0193945905284712
- Weeks, A., Swerissen, H., & Belfrage, J. (2007). Issues, challenges, and solutions in translating study instruments. *Evaluation Review*, 31(2), 153-165.
- Weidmer, B., Hurtado, M., Weech-Maldonado, R., Ngo-Metzger, Q., & Bogen, K. (2006). *Guidelines for translating CAHPS surveys*. Retrieved from http://www.ahrq.gov/cahps/surveys-guidance/helpful-resources/resources/cahpsGuidelines_Translation.html.
- Willis, G., Brick, P. D., Norberg, A., Stark, D. S., Kudela, M. S., Forsyth, B. H., . . . Lawrence, D. (2008). *Evaluation of a multi-step survey translation process*. Paper presented at the Proceedings of the International Conference on Survey Methods in Multinational, Multiregional, and Multicultural Contexts (3MC), Berlin, Germany.
- Wilss, W. (1990). Cognitive aspects of the translation process. *Language & Communication*, 10(1), 19-36.
- World Health Organization. (2009). Process of translation and adaptation of instruments. Retrieved from http://www.who.int/substance_abuse/research_tools/translation/en/

Annexe A

AP-G Development Steps

The first draft of the Adaptation process description assessment grid (AP-G) corresponded broadly to the 17 consensual strategies determined by René de Cotret *et al.* (2018). These strategies were subsequently refined on the basis of PA-AP descriptions found in the literature. The result is a list of concrete assessment criteria that can be applied to assess PA-AP descriptions.

The categorisation proposed is unidimensional: PA-AP descriptions are sorted out according to the number of consensual adaptation strategies they mention. There are 25 criteria, underlined in what follows and numbered as they are listed in the grid. The criteria are distributed according to the phase of the adaptation process. There are also three numeric variables related to the supra phrase. Those 28 elements constitute the AP-G.

Phase I - Initial Translation

This first phase of the adaptation process is divided into four primary strategies, for a total of nine assessment criteria addressing translation from the source language (SL) to the target language (TL). Back-translation is not addressed in this phase, since the resources involved do not directly affect the TV.

First primary strategy: Working in a multidisciplinary setting. Six assessment criteria are connected with this strategy.

Being multidisciplinary in nature, the work involved in the first phase of the adaptation process implies the contribution of several people, thus the criterion More than one translator (translation team) (1). This criterion is met if it can be concluded from PA-AP descriptions that at least two people translated from the SL to the TL. However, if one of the two people translated only from the TL to the SL (back-translation), the criterion is not met. The wording in some descriptions can make it difficult to determine if this criterion is satisfied or not. For instance, it is not always clear if the pronoun “we” refers to more than one person. From a sentence such as “Nous avons traduit en français tous les instruments utilisés dans le cadre de l’étude de validation originale,” (Brown *et al.* (2001, p. 989), it is not possible to establish if more than one translator was involved.

A variety of experience and expertise is involved in the translation team, among which is that of language, represented by two criteria. One refers to the contribution of language *experts*, that is, translators, linguists or terminologists who work as such in a professional setting: At least one language expert (2). However, stating that one is a language expert is not enough, because it does not address directly one’s command of the SL and/or TL; hence the next criteria below. Richer

and Bergeron (2012, p. 622), for example, mention that a professional translator participated in the initial translation but say nothing about his/her command of the languages involved: “The instrument was first translated from English to French by the authors of the paper and a professional translator.” Such a description does not satisfy this criterion, nor does the fact that a language expert performed a back-translation.

The other criteria related to language, Good command of the source and target languages by the translators (3), is met when descriptions are detailed enough to demonstrate that the translators are proficient in the SL and TL. Saying that the translators are “bilingual” or “completely bilingual” is not sufficient in this respect. Additional information is needed: the translator’s native language is French, s/he works with the two languages involved, her/his language proficiency level is certified, etc. A description such as Bergeron’s (2006, p. 59) meets this criterion: “... traduite en français de manière indépendante par trois chercheurs francophones et bilingues...” Another example is Lauzier and Haccoun (2010, p. 129): “Il est à noter que les traducteurs ayant participé au processus de traduction ont tous réussi avec succès les examens d’évaluation de la langue seconde établis par le Centre de psychologie du personnel de la Commission de la fonction publique du Canada.”

An understanding of the construct being measured is also recommended, thus the criterion At least one translator familiar with the construct/concept (4). This translator can be a researcher, a clinician or another expert in the area of the construct/concept. Those participating in the Initial Translation phase should at least have been briefed on the construct/concept, so that they can perceive how it influences the wording of the test items.

Familiarity with test development is also recommended: At least one translator familiar with test development (5). We assume that researchers, research assistants, psychologists, and psychometricians have that kind of expertise and are familiar with test item writing, which is not necessarily the case for professionals working in a mental health setting. Chevrier (2009, p. 79) provides a satisfying description: “Le troisième [traducteur] est un expert qui est familier avec les normes psychométriques...” On the other hand, L. Turgeon and Chartrand (2003, p. 379) description lacks precision: “... an expert committee of four members working in the field.”

The last recommendation from the scientific community with regard to this phase of the adaptation process is an experience in the source and target cultures: At least one translator with experience in the source and target cultures (6). To be consistent with this criterion, PA-AP descriptions must provide information about the kind of contact the translation team has had with the cultures involved: one translator has been living or lived in the source culture for a certain number of years, another has been living or lived in the target culture, etc. An example is found in Chevrier (2009, p. 79): “Le premier individu évolue dans l’environnement source, le deuxième évolue dans l’environnement cible.”

Second primary strategy: Producing at least two TVs. One criterion corresponds to this strategy: Production of two or more TVs (7). To fulfil it, PA-AP descriptions must clearly state that at least two TVs were indeed produced in the course of the adaptation process. Descriptions such as Lemelin *et al.* (2007, p. 294)—“Les items ont d’abord été traduits en français par les quatrième, cinquième et sixième auteurs de cette étude”—are not sufficient, because they are not specific enough. The same applies to Daigneault *et al.* (2013, p. 163): “The scale was first independently translated in French. This initial version was submitted to two reviewers.”

Third primary strategy: Providing translation guidelines. One criterion is associated with this strategy: Provision of translation guidelines (8). For it to be met, PA-AP descriptions must mention that some guidance was given to the translators: general translation rules were proposed, information about the target audience was shared, etc. Despite its seeming simplicity, the following comment by Johnson (2008, p. 28) is sufficient: “... traduire l’item anglais en français en respectant le mieux possible le sens de la phrase originale.”

Fourth primary strategy: Translating all the components of the test. This last strategy of the Initial Translation phase is an assessment criterion in itself: Translation of all the components of the test (9). It implies that psychometric tests are complete texts: in addition to items, they include a title, instructions, scoring manuals, answer sheets, etc. PA-AP descriptions do not have to mention all those elements: they only have to make it clear that the adaptation process was not limited to the test items. For example, Demers *et al.* (2002, p. 585) state that “the questionnaire items, instructions, and introductory text” were translated, and Richer and Bergeron (2012, p. 622) explain that “[a]n initial version of the translated instrument was administered to a small group of drivers to ensure that the instructions and various items were clear.” These two examples are consistent with this criterion.

Phase II - Internal Review

Two primary strategies are related to the Internal Review, for a total of seven criteria in our assessment grid.

Fifth primary strategy: Working in a multidisciplinary setting. Five criteria stem from this strategy to assess descriptions of the review committee. Three are similar to Phase I criteria #1, #2, and #6. Here we have More than one reviewer (review committee) (10), At least one language expert in the review committee (11), and At least one reviewer with experience in the source and target cultures (14).

The other two criteria imply additional qualifications on the part of the review committee compared with the translation team. That explains why having a good command of the source and target languages is not a relevant criterion here whereas being a language expert is. Along the same line, the review committee needs to be more than simply familiar with the construct, thus the criterion Mastery of the construct/concept by at least one reviewer (12). At least one

member of the review committee must be an expert in the area of the construct, such as the expert mentioned by Birmes *et al.* (2005, p. 147): “some linguistic revisions were made by a French psychiatrist with extensive expertise in PTSD research and treatment...” Another example comes from Lecavalier and Tassé (2001, p. 35): “La version préliminaire du premier comité a été donnée à un deuxième comité indépendant formé d'un traducteur et de deux psychologues qui avaient une formation dans le domaine du retard mental.” Obviously, authors of articles or academic publications dealing partially or totally with the construct being measured are considered specialists of this construct. Thus, when Lauzier *et al.* (2013, p. 356) state that their committee included five members, i.e. the authors themselves plus a professional linguist, they satisfy this criterion.

Lastly, there has to be At least one reviewer with expertise in test development (13). This person could be the author of the OV, the author of another test, or a psychometrician. Authors of journal articles or academic publications do not *a priori* qualify as experts in test development. For this criterion to be met, PA-AP descriptions must demonstrate that someone in the review committee had such expertise. According to this criterion, a description such as Lauzier and Haccoun (2010, p. 129), who state that their review committee was composed of five people, “soit les deux spécialistes en apprentissage ayant pris part au processus de traduction, deux professeurs d'université (les auteurs) et une traductrice professionnelle” is not sufficient. Contrastingly, Birmes *et al.* (2005, p. 147) provide a satisfying description when they write: “This draft was back translated from French to English by a professional translator, and approved by [the creator of the original test].” The same applies to Laliberté and Earls (2007, p. 216): “Une copie des versions traduites a également été envoyée à un des auteurs des versions originales (Rod A. Martin, qui est bilingue) pour recueillir ses commentaires.” Another example is Pelletier (2007, p. 41): “La traduction a été vérifiée par deux psychologues indépendants. Une des deux psychologues pratiquait dans le domaine des [troubles de la personnalité] et offrait des formations sur l'administration du SCID-II au Québec.”

Sixth primary strategy: Overseeing the internal review. The scientific community recommends that PA-AP descriptions refer to at least one overseeing procedure in the context of the Internal Review. *Overseeing* is the right word here, because the review committee uses strict, specific criteria. As explained by René de Cotret *et al.* (2018), the scientific community also mentions that when back-translation is used to this effect, it should be as a complement to another procedure.

The criterion Use of back-translation as an overseeing procedure (15) refers to the application of one or more back-translations to assess, by inference, the extent to which the TV and OV correspond. For example, C. Gagnon (2013, p. 21) explains that “lorsque les items de la version traduite en anglais et ceux de la version originale étaient identiques, les items en français étaient conservés”. Another example is provided by Lemelin *et al.* (2007, p. 294), who states: “[U]ne

comparaison des items retraduits et des items originaux a permis d'identifier les items dont la traduction en français était imparfaite.”

Back-translation is not the only possible overseeing measure, thus the criterion Use of at least another overseeing procedure (apart from back-translation) (16). Another way to proceed would be to compare the TV and OV items on the basis of different types of equivalence. Demers *et al.* (2002, p. 585), for example, used conceptual equivalence: “[The reviewers] were asked to examine the discrepancies and determine, in light of the conceptual equivalence with the source items and the author’s original purpose, what should be the most accurate and satisfying phrasing.” Laverdière *et al.* (2010) used a rating scale developed by Jeanrie and Bertrand (1999), which allows for the review of each item on the basis of three types of equivalence (content-based, conceptual, and linguistic).

Dimensions such as spelling, grammar, level of language, etc., can also intervene to establish if the translation is of good quality and if some items have been mistranslated. The theory of test translation error by Solano-Flores, Backhoff, and Contreras-Niño (2009) is along those lines. There were no examples of the sort in our corpus.

Dimensions others than those mentioned above can also be taken into account. Johnson (2008, p. 28) explains that reviewers can “inscrire le pourcentage (de 0% aucunement fidèle à 100% parfaitement fidèle) correspondant au respect du sens de l’item anglais par rapport à l’item français”. As for Trottier *et al.* (2008, p. 239), they state: “Finalement, le comité d’experts a évalué la validité de contenu de la version française en confirmant que ses items semblaient mesurer véritablement les mêmes aspects de l’optimisme que la version originale de l’instrument.”

Evaluating technical terms (Vallerand (1989) is another way to proceed, especially to complement the back-translation procedure. This strategy was applied by Dubé (2011, p. 47), who performed “l’étude des termes techniques utilisés dans les traductions françaises [...] pour éviter que des termes étroitement reliés mais légèrement dissemblables expliquent des traductions inversées identiques”.

To demonstrate that the internal review was overseen, authors of PA-AP descriptions must go beyond mere allusion to this review. For example, L. Turgeon and Chartrand (2003, p. 379) does confirm that there was an internal review, but not that it was overseen: “The first version of the translation was revised by a bilingual person independent of the committee and then was approved by the committee”. In the same way, stating that two or more reviewers agree on something, or, as in Delisle (2011), that a 96% agreement rate was reached between two reviewers, does not mean that the reviewing process was overseen. Also, involving outside reviewers in the Internal Review does not qualify as an overseeing procedure. Thus, when Achim (2006, p. 150) writes that “[q]uelques modifications mineures ont enfin été apportées au niveau de syntaxe, à la suggestion de deux psychologues ayant procédé à une évaluation

minutieuse de la version française”, her description does not meet this criterion. The description provided by Méthot and Morin (2004, p. 60) is not any better: “Deux rencontres entre les deux comités ont eu lieu pour discuter des versions traduites.”

Phase III - External Review

This third phase of the adaptation process is divided into two primary strategies, for a total of five assessment criteria.

Seventh primary strategy: Working in a multidisciplinary setting. This setting refers to the review committee as described in Phase II. What matters here, in Phase III, is to include a Clear mention that the external review was performed by the review committee (17). From our results, this criterion is rarely fulfilled, and it is generally difficult to know if the external review was done by the entire review committee, a single reviewer or someone outside the committee. When Hébert *et al.* (2007, p. 5) write, for example, that “[r]esults showed all items were correctly classified”, they do not provide sufficient information. Tassé and Craig (1999), on their part, used a PA-TH that is rather an exception, since two review committees were recommended. This is also the case for Méthot and Morin (2004, p. 60): “Les membres du premier comité se sont rencontrés à nouveau pour évaluer les informations obtenues lors [du prétest].”

Eighth primary strategy: Overseeing the external review. Four assessment criteria stem from this strategy. They correspond to different data collection procedures used as a pre-test.

First, PA-AP descriptions might mention that Individual debriefing interviews (18) were conducted during or after the pre-test. This procedure, usually called *probe technique* or *cognitive interview*, is helpful to determine if items are ambiguous or controversial, if the test is too long, etc. Specific translation problems can be identified during the interviews, on the basis of the subjects’ comments. Dubé (2011, p. 48) is consistent with this criterion: “Lorsqu’indiqué, les personnes ont été questionnées sur la nature de l’ambiguïté de l’item concerné. Les personnes ont également été invitées à émettre leurs suggestions d’amélioration.” Another example is Demers *et al.* (2002, p. 585): “An investigator was present for debriefing.”

Another way to proceed is through Focus groups (20), which implies that several subjects are interviewed at the same time. Focus groups are useful to collect the opinion of small groups of individuals who share similar characteristics. The excerpts below correspond to examples of descriptions meeting this criterion:

Pour produire une dernière version en intégrant les commentaires du groupe d’experts, l’équipe de recherche a procédé à l’analyse de toutes les suggestions émises et elle a modifié les énoncés des questions qui étaient plus problématiques aux trois sites de l’étude. Pour s’y faire, l’équipe a regroupé les commentaires des participants des trois villes pour chacune des questions.

Ensuite, nous avons analysé les convergences et les divergences des suggestions pour chaque énoncé. (Bergeron, 2006, p. 66)

Deux focus groups ont contribué à vérifier la compréhension des items de la version française adaptée de l'échelle. Le premier groupe était formé de quatre étudiants universitaires âgés entre 19 et 22 ans. Le deuxième groupe était formé de quatre personnes âgées entre 42 et 53 ans. Les deux groupes ne trouvèrent aucune difficulté de compréhension dans le questionnaire modifié. (Cordeau, 2007, p. 20)

Subjects can also be asked to evaluate the quality of the TV, thus the criterion Written evaluations of the TV (19). In this context, subjects will circle ambiguous items, or use Likert scales to rate the clarity of each item. They might also be asked to comment or propose translations of their own. Here is an example, taken from Plante-Jean (2012, p. 94): “[Les sujets] devaient évaluer sur une échelle allant de 1 (pas clair du tout) à 7 (extrêmement clair) la clarté de chacun des items des deux questionnaires. Dans le cas où un item apparaissait ambigu (score de 4 et moins), il était demandé au participant de faire des commentaires ou propositions pour en améliorer la compréhension.” Another example is provided by Gagné (2010, p. 33): “Les participants ont répondu au questionnaire et ont inscrit des commentaires lorsqu'un item n'était pas clairement expliqué.”

A variety of other procedures can be used to collect data; this is reflected in the criterion Other overseeing procedures (21). The following paragraphs provide examples from our corpus.

The review committee could rely on subjects who know both the source and target languages. These bilingual subjects, as they are usually referred to, could be asked to compare—explicitly or implicitly—the TV and OV. This corresponds to the procedure described by Filiatrault *et al.* (2014, p. 2): “Participants were asked to compare each item of the French and original versions of the FaB and indicated if they thought there were discrepancies between versions.” Bilingual subjects could also be asked to complete both versions of the test, in which case the comparison would be done via statistical analysis of the tests results.

Another example comes from Lauzier and Haccoun (2010, p. 129), who state that subjects were asked if the TV items were phrased in a clear and unambiguous language, without specifying if the subjects' input was verbal or written. Ménard (2008, p. 56) is a similar case: “[A] pilot test of the French translated questionnaires has been made with 5 managers to make sure items were adequately formulated and easy to understand.”

In other cases, preliminary statistical analysis is done to establish if the answers provided by the subjects correspond to what is expected. For instance, Gagné (2010, p. 33) mentions that “l'analyse de la cohérence interne de cette version combinée [...] a permis d'identifier les items auxquels des changements devaient être apportés”, while Lavergne (2003, p. 65) explains: “La

moyenne obtenue dans l'étude pilote était de 86.4, ce qui est légèrement plus élevé que la moyenne de 81.68 obtenue par Archer et Winchester en 1994, mais se situe toujours dans la catégorie d'un deuil moyennement intense.”

Phase IV – Testing

Only two assessment criteria are related to this phase. In fact, we consider testing as a principle: all psychometric tests, even those which are not adapted, go through this phase. There is abundant literature on testing and on the variety of statistical analysis that can possibly be used. We suggest reading Rios and Sireci (2014), who propose ten assessment criteria related to this phase.

Our first criterion, Validation of the TV via accepted statistical analysis methods (22), is fulfilled when PA-AP descriptions clearly state that the TV was administered to dozens or hundreds of participants, and that statistical analysis was performed afterwards to measure its psychometric value. The second criterion, Changes to the TV following statistical analysis (23), is met when one clearly mentions, for instance, that items were removed or re-translated in light of the results of statistical analysis.

Supra Phase – Documentation

This primary strategy aims at documenting the adaptation process so that: (a) those reading PA-AP descriptions are able to assess its rigour and validity; (b) it can easily be replicated by other researchers, and (c) the scientific community is made aware of the problems encountered. This is a *supra* phase, in the sense that it encompasses all the other phases of the adaptation process, which should all be documented. To some extent, each of the criteria previously mentioned provides information on the specific PA-AP being addressed in the description. In fact, the more criteria a description meets, the more information it contains on the PA-AP itself. Nevertheless, four additional criteria should be considered with regard to this Supra Phase.

First, an Account of the translation/adaptation problems (24) should be provided and ideally discuss how they were solved. The account should contain relevant information about the problems. Statements such as “[t]his step led to the adjustment of four items” (Grégoire *et al.* (2013, p. 3) or “[d]e légères corrections furent apportées en suivant un processus décisionnel par consensus” (Johnson (2008, p. 29) are not enough to satisfy this criterion, nor are comments to the effect that the translation/adaptation process went smoothly, without any problem. In our opinion, adapting a test cannot be problem-free. The following excerpt from Demers *et al.* (2002, p. 588) corresponds to an account that provides relevant information and therefore meets this criterion:

For the remaining items ($n = 12$), either one or both back-translated versions did not match with the source item. Discussions within the review committee

led to the acceptance of one of the French options for the following eight items: adequacy, efficiency, expertise, skillfulness, sense of power, sense of control, eagerness to try new things, and ability to take advantage of opportunity. The accurate meanings of the items adequacy and skillfulness were, however, difficult to translate because the French word ‘aptitude’ could theoretically apply to both. To help resolve the issue, the participating author of the tool provided a glossary of the terms (unpublished material) that had been developed to help administer the PIADS when an interviewer is being [sic] present. Accordingly, adequacy, which refers to being capable of handling life situations and handling little crisis, could be translated by ‘se sentir à la hauteur’. [...]

PA-AP descriptions can also contain the OV and/or AV—or a link leading to an online version—, thus the criterion Provision of the OV and AV (25). The AV is in itself a research result, and as such should be made available to the research community. Researchers willing to appreciate the quality of the AV of a test would greatly benefit from having access to both the OV and AV, in order to compare them. In the same way, they would gain from seeing the TVs produced in the first three phases of the adaptation process as well as the back-translations, if any. However, giving access to the TVs and back-translations is not among the strategies on which a consensus is reached, so we did not make a criterion out of it.

In addition to the 25 criteria presented up to now, three numeric variables are included in our assessment grid; they represent information sources correlated with the number of assessment criteria that are met (see Study II).

The first variable touches on the number of references related to the adaptation process. Descriptions such as Achim (2006, p. 46), who brings up using “la méthode de la double traduction (Vallerand, 1989)”, are in line with this criterion. In this precise case, referring to Vallerand (1989) would count as one reference. In the case of Demers *et al.* (2002, p. 584), whose translation was based on “the first steps of Vallerand’s methodology [6] and on Guillemin *et al.*’s [7] guidelines for cross-cultural adaptation and validation of health questionnaire”, this would count as two references. References that justify the use of certain strategies, such as in Lamoureux-Hébert (2006, p. 57), also count: “Comme le recommandent Massobré *et al.* (2002), cette méthode propose...” When secondary references are quoted, only the main reference is counted. For example, when Lauzier and Haccoun (2010, p. 129) explain that they followed “les recommandations émises par Spielberger et Sharma (1976; cités dans Vallerand, 1989) sur la préparation de la forme expérimentale ont été respectées”, this counts as one reference only. References to statistical analysis are not taken into account, because they do not address exclusively the adaptation process.

The second and third numeric variables refer to the number of words in the description and the number of words in additional theoretical developments.¹¹ As previously mentioned, our grid is meant explicitly to assess PA-AP descriptions. However, these descriptions are sometimes supplemented by theoretical developments that either precede them or are included in the theoretical framework. We decided to consider the descriptions and theoretical developments separately. Besides, we noticed in the descriptions that the parts addressing phase IV of the adaptation process were generally much longer than those addressing phases I, II, and III. Therefore, only the latter were taken into consideration.

Other variables

Two more variables were used in the context of Study II, although they are not included in our assessment grid. The first is the type of text: journal articles or academic publications (PhD theses/dissertations or MA theses/dissertations). The second is relative to the way PA-APs are discussed in the texts, either directly or indirectly. We consider that PA-APs are discussed directly: 1) when they are mentioned in the abstract and/or objectives; 2) when the abstract and/or objectives discuss the validation of the AV, and the PA-AP itself is discussed elsewhere in the text. We consider that PA-APs are discussed indirectly in any other cases, for instance, when they are merely mentioned in the body of the text or in an appendix.

¹¹ Counting words manually can prove difficult when they amount to hundreds or thousands. One can copy and paste the description into a word processor to get an automatic word count, or, depending on how the data will be used, simply count the number of lines or paragraphs.

Annexe B

Typical examples of PA-AP description according to the number of criteria met on the AP-G

Description that meets none of the criteria:

Une traduction libre de ce questionnaire est utilisée dans la présente recherche (voir Appendice H). (Brassard, 2003, p. 76)

Description that meets four criteria:

Ce questionnaire a été traduit de l'anglais au français, puis soumis à un groupe de quatre professeurs de psychologie du travail et des organisations, Canadiens bilingues, de l'Université de Montréal, qui ont examiné les deux versions française et anglaise du questionnaire, et ce, afin de s'assurer de la qualité de la traduction. (Behling et Law, 2001). Suite aux commentaires des professeurs, certains items ont été ajustés. (Temimi, 2011, p. 65)

Description that meets eight criteria:

L'échelle originale a d'abord été traduite par la technique du back translation selon le protocole proposé par Bradley (1994) : pour la version anglaise vers la version française par l'auteur de la présente étude, puis par une psychologue parfaitement bilingue pour la version anglaise vers la version française. Cette psychologue est une experte du discours psychanalytique de langue anglaise et a plusieurs expériences dans la traduction d'outils psychométriques de l'anglais vers le français. La version anglaise ainsi obtenue a été comparée par l'auteur à la version anglaise originale et les différences ont été identifiées avec un très bas seuil de tolérance. Sur les 34 items, 26 ont immédiatement été jugés équivalents, n'ayant qu'un ou deux mots différents mais étant entièrement synonymes d'une version à l'autre, et six autres items étaient parfaitement exacts. Deux items (items 23 et 29) ont fait l'objet de brèves discussions, que les juges ont conclues en constatant l'équivalence des items en cause. Plusieurs efforts pour contacter l'auteure de l'outil original et solliciter son évaluation de la forme anglaise de l'ÉAC sont demeurés vains." (Côté *et al.*, 2013, p. 57)

Description that meets 12 criteria:

The design used in this study is based on the guidelines proposed by Vallerand [16] for cross-cultural adaptation of questionnaires. In keeping with these guidelines, the original version of the FaB was translated in French by one professional translator and two bilingual health professionals (parallel translations) working in the field of geriatrics and whose mother tongue was

French. As shown in Figure 1, the two translated versions were subsequently translated in English (parallel backward translations) by a second professional translator and another team of bilingual health professionals whose mother tongue was English. This backward translation step increases the quality of the adaptation process by ensuring the intercultural equivalence between the original and translated versions of a questionnaire [16,17]. All versions of the FaB (two French Canadian versions, two backward translated English versions and the original version) were examined by an expert committee. This committee was composed of three researchers (JF, LD and MP), one research professional, one professional translator and two health professionals involved in the translation process. After comparing the five versions of the FaB, the committee agreed on an optimal French version for each item of the FaB.

This led to the elaboration of a pre-experimental version of the FaB-FC, which was pilot tested among six bilingual community-dwelling seniors, recruited among the research participants' database of the Institut universitaire de gériatrie de Montréal and among acquaintances. Participants were asked to compare each item of the French and original versions of the FaB and indicated if they thought there were discrepancies between versions. They were also asked to report on any ambiguous or incomprehensible terms or expressions in the French version of the FaB. Minor changes were brought to the questionnaire and involved making some items more precise (#1, #22, #26), more concise (#3, #28), substituting a word (#4, #24) and removing a word (#15). This pilot test led to the elaboration of a final version of the FaB-FC. (Filiatrault *et al.*, 2014, p. 2)

Theoretical development (typical example):

[...] Cette technique consiste à faire traduire un texte par plusieurs traducteurs, qui n'ont pas pris connaissance des traductions faites par les autres (Lauzier & Haccoun, 2008).

Certaines études ont employé la technique de la traduction renversée (back-translation), mais plusieurs chercheurs l'ayant utilisée ont constaté qu'il était bien difficile d'obtenir une équivalence parfaite entre la version retraduite et la version d'origine (voir Rybowiak, Garst, Frese & Batinic, 1999). En outre, la technique TPA présente certains avantages comparativement à la technique de traduction renversée (Usunier, 1992). Premièrement, la technique de la traduction renversée repose sur l'hypothèse qu'il existe un langage universel, lequel serait exempt de toute variabilité culturelle et contextuelle. Cela peut entraîner un sentiment erroné de confiance chez le chercheur, en produisant une équivalence lexicale trompeuse, car apparente. Cette lacune peut toutefois être comblée par un processus de « recadrage », qui fait appel à de multiples traductions simultanées (Sarazin, 2003), comme l'exige la technique de la

TPA. Deuxièmement, la technique de la traduction renversée se centre de façon quasi exclusive sur le principe d'équivalence des termes, pour lesquels il existerait une traduction équivalente dans toutes les langues. Haccoun (1987) illustre bien la lacune de ce principe à l'aide de l'exemple suivant: *My supervisor treats me fairly* (item original) – *Mon superviseur me traite de façon équitable* (traduction) – *My boss treats me equitably* (traduction renversée). L'exemple montre clairement la non-équivalence des termes pouvant appartenir à chacune des versions anglaises (*supervisor* ≠ *boss*; *fairly* ≠ *equitably*). Pour ces diverses raisons, nous croyons que la traduction parallèle effectuée par de multiples traducteurs est préférable. (Lauzier & Haccoun, 2010, pp. 128-129)

Annexe C : Grille d'évaluation de la description du Processus d'Adaptation (G-PA)

Grille d'évaluation de la description du processus d'adaptation (G-PA)

Phase I – Traduction initiale Version originale (VO) → Version(s) traduite(s) (VT)	
1. Plus d'un traducteur forme le groupe de traduction	<input type="checkbox"/>
2. Présence d'au moins un traducteur professionnel / linguiste / terminologue	<input type="checkbox"/>
3. Maîtrise des langues cible et source du/des traducteur(s)	<input type="checkbox"/>
✓ = « De langue maternelle française », « travaille dans les deux langues », etc.	
✗ = « Bilingue », « complètement/vraiment bilingue », etc.	
4. Au moins un traducteur familier avec le construit/concept mesuré par le test	<input type="checkbox"/>
✓ = Un chercheur/clinicien/professionnel du domaine ou quelqu'un qui a reçu des explications à propos du construit/concept mesuré par le test.	
5. Au moins un traducteur familier avec la construction de tests	<input type="checkbox"/>
✓ = Un chercheur / assistant de recherche / psychologue / psychométricien ou, minimamente, quelqu'un familier avec la rédaction d'items.	
✗ = « Professionnel en santé mental », « expert », etc.	
6. Au moins un traducteur ayant l'expérience des cultures cible et source	<input type="checkbox"/>
✓ = Un contact significatif avec les deux cultures (p. ex. il habite là-bas).	
✓ = Un traducteur a l'expérience de la culture cible et un autre, de la culture source.	
✗ = On mentionne uniquement que le traducteur est « biculturel ».	
7. Au moins deux VT ont été produites	<input type="checkbox"/>
8. Consignes de traduction	<input type="checkbox"/>
→ Règles de traduction fournies, informations données sur le lecteur cible, etc.	
9. Traduction de l'intégralité du test	<input type="checkbox"/>
→ Titre, instructions, items, échelles, feuille-réponse, etc.	
N.B. Les neufs critères de la phase I réfèrent à une traduction de la langue source vers la langue cible, non pas à une traduction inversée.	
Phase II – Révision interne de la ou des VT	
10. Plus d'un réviseur forme le comité de révision	<input type="checkbox"/>
11. Présence d'au moins un traducteur professionnel / linguiste / terminologue	<input type="checkbox"/>
12. Au moins un réviseur maîtrise le construit/concept mesuré par le test	<input type="checkbox"/>
✓ = Spécialisé dans le domaine où le construit/concept est utilisé.	
✗ = Ses activités professionnelles ne sont pas liées au construit/concept.	
13. Au moins un réviseur ayant une expertise de la construction de tests	<input type="checkbox"/>
✓ = Auteur de la VO, auteur d'un autre test, psychométricien, etc.	
14. Au moins un réviseur ayant l'expérience des cultures cible et source	<input type="checkbox"/>
15. La traduction inversée est employée en tant qu'encadrement de la révision interne	<input type="checkbox"/>
16. Au moins un type d'encadrement (excluant la traduction inversée) est employé	<input type="checkbox"/>
✓ = La révision des termes techniques de Vallerand (1989), le recours à différents types d'équivalence comme ceux proposés par Jeanrie & Bertrand (1999), l'évaluation des erreurs de traduction (Solano-Flores et al., 2006), etc.	
✗ = Une révision libre, un accord interjuge, le recours à des évaluateurs externes, etc.	

Phase III – Révision externe (avec prétest) de la VT

17. Mention explicite que le comité de révision a effectué la révision externe
18. Sous-questionnement des sujets à l'oral pendant ou après le prétest
✓ = Probe technique/cognitive interview, débriefing, etc.
19. Les sujets évaluent par écrit la clarté de la VT
✓ = En encerclant les éléments confus, avec des échelles Likert, etc.
20. Focus group
21. Autre type d'encadrement de la révision externe employé : _____

Phase IV – Testing VT → Version adaptée (VA)

22. Utilisation d'analyses statistiques reconnues pour légitimer la VT
23. Les résultats des analyses statistiques servent à modifier la VT

Supra phase – Documentation du processus

24. Compte rendu des problèmes de traduction/adaptation
✓ = On mentionne quels éléments ont causé problème et/ou ce qui a été fait pour y pallier.
✗ = On mentionne, par exemple, que trois items ont causé problème, sans les identifier, ou qu'aucun item n'a causé problème lors processus d'adaptation.

25. Disponibilité de la VO et de la VA dans le texte ou lien internet fourni

Nombre de référence(s) liée(s) au processus d'adaptation : _____ réf.(s)

Nombre de mot(s) de la description : _____ mot(s)

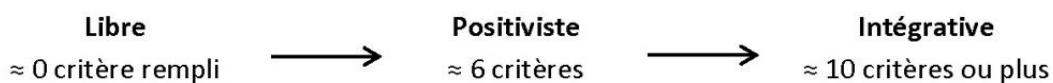
→ En excluant le segment de la description portant sur la phase IV,
généralement surreprésentatif.

Nombre de mot(s) du complément théorique : _____ mot(s)

Nombre de critère(s) dichotomique(s) rempli(s) : _____ / 25

René de Cotret, Lanza & Francoeur (2016)®

Interprétation. Le nombre de critère(s) rempli(s) donne une idée générale du type de description (libre, positiviste ou intégrative, cf. René de Cotret, Lanza et Francoeur (2016) pour des explications détaillées), bien qu'un examen attentif soit nécessaire. En effet, il importe de déterminer ce qui est mis de l'avant dans la description pour convaincre le lecteur de la légitimité du processus d'adaptation effectué. Est-ce que la description est laconique, démunie de tels arguments? Met-elle de l'avant essentiellement des données quantitatives? Laisse-t-elle au lecteur le soin de juger de ladite légitimité en lui fournissant un maximum d'informations, aussi bien quantitatives que qualitatives? Encerclez le type de description qui correspond le mieux :



Annexe D: Adaptation Process description assessment Grid (AP-G)

PA-AP Description Assessment Grid

Phase I – Initial Translation Original Version (OV) → Translated Version(s) (TV)	
1. More than one translator (translation team)	<input type="checkbox"/>
2. At least one language expert	<input type="checkbox"/>
3. Good command of the source and target languages by the translators ✓ = "mother tongue is the target language", "works in both languages", etc. ✗ = "bilingual", "completely/truly bilingual", etc.	<input type="checkbox"/>
4. At least one translator familiar with the construct/concept ✓ = one researcher/clinician/expert in the area of the construct or someone who was briefed on the construct/concept being measured.	<input type="checkbox"/>
5. At least one translator familiar with test development ✓ = one researcher/research assistant/psychologist/psychometrician, or minimally, someone who is familiar with test item writing. ✗ = « professionnal working in a mental health setting », « expert », etc.	<input type="checkbox"/>
6. At least one translator with experience in the source and target cultures ✓ = a significant contact with both cultures (e.g. translator lives in the target culture). ✓ = one translator has experience in the target culture, another, in the source culture. ✗ = translator is « bicultural ».	<input type="checkbox"/>
7. Production of two or more TVs	<input type="checkbox"/>
8. Provision of translation guidelines → general translation rules are proposed, info about the target audience is shared, etc.	<input type="checkbox"/>
9. Translation of all the components of the test → title, instructions, items, scales, answer sheets, etc.	<input type="checkbox"/>
Note. The above criteria refer to translation from the source language to the target language. They do not take back-translation into account.	
Phase II – Internal Review of the TV(s)	
10. More than one reviewer (review committee)	<input type="checkbox"/>
11. At least one language expert in the review committee	<input type="checkbox"/>
12. Mastery of the construct/concept by at least one reviewer ✓ = expert in the area of the construct/concept ✗ = professional activities not related to the construct/concept	<input type="checkbox"/>
13. At least one reviewer with an expertise in test development ✓ = author of the OV, author of another test, psychometrician, etc.	<input type="checkbox"/>
14. At least one reviewer with experience in the source and target cultures	<input type="checkbox"/>
15. Use of back-translation as an overseeing procedure	<input type="checkbox"/>
16. Use of at least another overseeing procedure (apart from back-translation) ✓ = review of technical terms as suggested by Vallerand (1989), use of different types of equivalence as proposed by Jeanrie & Bertrand (1999), evaluation of translation errors (Solano-Flores et al., 2006), etc. ✗ = free translation, agreement between reviewers, external reviewers, etc.	<input type="checkbox"/>

Phase III – External Review (with pretest) of the TV	
17. Clear mention that the external review was performed by the review committee	<input type="checkbox"/>
18. Individual debriefing interviews ✓ = Probe technique/cognitive interview, debriefing, etc.	<input type="checkbox"/>
19. Written evaluations of the TV ✓ = circling ambiguous items, Likert scales, etc.	<input type="checkbox"/>
20. Focus groups	<input type="checkbox"/>
21. Other overseeing procedures _____	<input type="checkbox"/>
Phase IV – Testing TV → Adapted Version (AV)	
22. Validation of the TV via accepted statistical analysis methods	<input type="checkbox"/>
23. Changes to the TV following statistical analysis	<input type="checkbox"/>
Supra phase – Documenting the process	
24. Account of translation/adaptation problems ✓ = problems are listed and/or solutions are described ✗ = e.g. mention to the effect that three items were problematic, but items not identified; mention to the effect that there were not problems during the adaptation process	<input type="checkbox"/>
25. Provision of the OV and AV	<input type="checkbox"/>
Number of references related to the adaptation process:	_____ refs
Number of words in the description: → parts addressing phase IV of the adaptation process are excluded, because they are usually overrepresented.	_____ words
Number of words in additional theoretical developments:	_____ words
Number of dichotomic criteria met: _____ / 25	

René de Cotret, Leanza & Francoeur (in press)[©]

Interpretation. The number of criteria met gives a general idea of the type of description (free, positivist, or integrative), although a closer look is needed. In fact, it is important to determine the type of arguments put forward in the description to convince the reader of the legitimacy of the process. Is the description laconic, without justifications? Does it contain mostly quantitative data? Does it allow the reader to assess the legitimacy of the process, by providing a maximum of information, both quantitative and qualitative? Circle the type of description which corresponds the best to the description:

