

VERS UNE MÉTHODOLOGIE DE VALIDATION TRANS-CULTURELLE DE QUESTIONNAIRES PSYCHOLOGIQUES: IMPLICATIONS POUR LA RECHERCHE EN LANGUE FRANÇAISE

ROBERT J. VALLERAND
*Laboratoire de psychologie sociale
Département de psychologie
Université du Québec à Montréal*

RÉSUMÉ

La validation d'instruments de mesure traduits de l'anglais au français est un problème méthodologique qui préoccupe un nombre grandissant de chercheurs francophones. Parallèlement, les chercheurs en psychologie sont confrontés à l'absence d'une méthodologie systématique de validation trans-culturelle. Le présent article a pour but de combler cette lacune en proposant une méthodologie de validation trans-culturelle d'instruments psychologiques fondée sur les écrits pertinents émanant du secteur de la psychologie trans-culturelle. Ainsi, une approche systématique contenant une série d'étapes menant à la réalisation de traductions fidèles et valides en langue française est proposée. Ces étapes sont: (1) la préparation d'une version préliminaire, (2) l'évaluation et la modification de cette version préliminaire, (3) l'évaluation de la version expérimentale par un prétest, (4) l'évaluation de la validité concomitante et de contenu, (5) l'évaluation de la fidélité, (6) l'évaluation de la validité de construit, et (7) l'établissement de normes. En se référant à des recherches récentes issues du domaine de la psychologie sociale, des applications de la méthodologie proposée sont présentées.

L'utilisation de questionnaires s'avère un aspect méthodologique très important pour la recherche en psychologie. En effet, un fort pourcentage des chercheurs dans ce secteur utilisent dorénavant les questionnaires psychologiques¹, que ce soit dans les secteurs de la psychologie clinique, sociale, du développement, ou de la personnalité. À titre d'exemple, au cours des années 1982 à 1986, pas moins de 60% des articles publiés dans la *Revue Canadienne des Sciences du Comportement* a fait usage de questionnaires psychologiques (si on incluait les autres types de questionnaires dans les statis-

tiques, le pourcentage atteindrait facilement les 90%).

Malheureusement, le chercheur francophone, comparativement au chercheur anglophone, est quelque peu désavantagé dans les recherches nécessitant l'utilisation de questionnaires psychologiques. En effet, ce dernier jouit d'un éventail étendu de tests psychologiques portant sur de nombreux thèmes. Face à une telle situation, trois options s'offrent au chercheur francophone, soit: (1) utiliser les tests originaux et se limiter à effectuer de la recherche auprès de la population anglophone; (2) développer de nouveaux instruments en français et ainsi étudier

¹Les termes "tests", "questionnaires", "instruments de mesure" et "instruments psychologiques" sont utilisés interchangeablement dans le présent article. Ils correspondent à des mesures d'une dimension psychologique définie obtenues par écrit et nécessitant des preuves de leur fidélité et de leur validité statistiques. La définition du terme "questionnaire psychologique" exclut donc les échelles d'un seul item et des questions portant sur des variables socio-démographiques. Des échelles de personnalité (e.g., le 16PF de Cattell, Eber, & Tatsuoka, 1970), d'émotions (e.g., Izard, Dougherty, Bloxom, & Kotsch, 1974), ou d'attributions (e.g., Russell, 1982) représentent des exemples d'instruments psychologiques considérés dans le présent document.

la population canadienne-française; ou (3) valider en français des instruments psychologiques déjà existant en anglais et les utiliser avec la population francophone.

Alors que la première option offre certes des résultats immédiats en recherche qui ne sont pas à négliger, il semble, par contre, que le processus de validation d'instruments de mesure psychologique en langue française réponde à un besoin différent mais tout aussi important. En effet, un tel processus de validation permet de rapprocher les recherches effectuées sur les courants dominants en Amérique du Nord à celles réalisées dans les milieux francophones. Ceci offre l'avantage supplémentaire de tester la validité externe des résultats des publications scientifiques anglophones dans la culture canadienne-française. En plus, cette approche permet de développer un éventail d'instruments psychologiques pouvant être utilisés en recherche auprès de sujets francophones. Dans un tel cadre de validation d'instruments en français, cette seconde option, soit le développement d'un nouvel instrument, n'est recommandée que lorsqu'aucun instrument fidèle et valide n'existe déjà en langue anglaise. En effet, il semble inapproprié d'ajouter à la confusion en recherche en créant des instruments similaires pour mesurer le ou les mêmes concepts dans les deux langues. Donc, dans l'éventualité où un tel instrument existe déjà en anglais, la troisième option, soit celle d'effectuer une validation trans-culturelle en français du test en question, est recommandée. C'est dans le cadre de cette troisième et dernière option que s'inscrit le présent document².

Au cours des dernières années au Canada, un nombre grandissant d'études ont été conduites afin de valider en français différents questionnaires psychologiques anglais. Malgré cet intérêt évident pour la validation trans-culturelle d'instruments psychologiques, il est important de noter qu'au meilleur de notre connaissance, il ne semble exister aucune méthodologie de validation trans-culturelle articulée à cet égard. Il est également possible de remarquer une très grande diversité dans les procédures utilisées, le tout allant de la simple traduction à des procédures de validation trans-culturelle des plus complètes.

²Même si le présent article traite de validation trans-culturelle d'instruments de recherche dans un contexte de recherche, la présente méthodologie peut être également utilisée avec des instruments utilisés en intervention ou dans un cadre vocationnel.

Des méthodologies de qualité variable risquent fort de mener à des versions en français de qualités inégales. Une traduction d'un instrument psychologique ne demeure que cela, une traduction, et n'assure en rien la validité et la fidélité de l'instrument dans la culture canadienne-française. Il devient donc important qu'une méthodologie adéquate soit utilisée dans le processus de validation trans-culturelle d'instruments psychologiques.

À la lumière des présentes considérations, le but du présent article est de proposer une série d'étapes systématiques menant à la validation, en langue française, de questionnaires psychologiques originalement validés en langue anglaise. En se basant sur les travaux et méthodes de recherche trans-culturelle, sept (7) étapes sont proposées pour la validation trans-culturelle de questionnaires psychologiques. Ces étapes sont: (1) la préparation d'une version préliminaire; (2) l'évaluation et la modification de la version préliminaire; (3) l'évaluation de la clarté des questions par des membres de la population cible dans un prétest; (4) l'évaluation de la validité concomitante et de contenu du questionnaire; (5) l'évaluation de la fidélité test-retest et de la consistance interne de l'instrument; (6) l'étude de la validité de construit du questionnaire; et (7) l'établissement de normes. La Figure 1 présente cette séquence impliquant les différentes étapes à franchir afin d'assurer une version valide et fidèle de l'instrument en français. Il est possible de remarquer qu'à chacune des différentes étapes, le chercheur doit évaluer la qualité de l'échelle à l'aide de diverses techniques et méthodes et décider si l'instrument est prêt à passer à la phase suivante ou s'il est préférable de réévaluer l'échelle en question. Généralement, lorsqu'une réévaluation de l'échelle s'impose, il y a retour à l'étape précédente. Par contre, des problèmes sérieux peuvent mener à une réévaluation plus profonde de l'échelle et à l'occasion, à un retour aux étapes initiales. Cette série de décisions doit être judicieuse car elle peut mener à des conséquences importantes sur le contenu de la dernière version de l'instrument ainsi que pour les recherches qui découleront de son utilisation.

Pour nos présents besoins, nous utiliserons des exemples tirés de récentes recherches ayant pour but de valider en français des questionnaires d'origine américaine. Ces études représentent les articles publiés ou soumis pour fins de publication depuis 1982 dans la *Revue canadienne des*

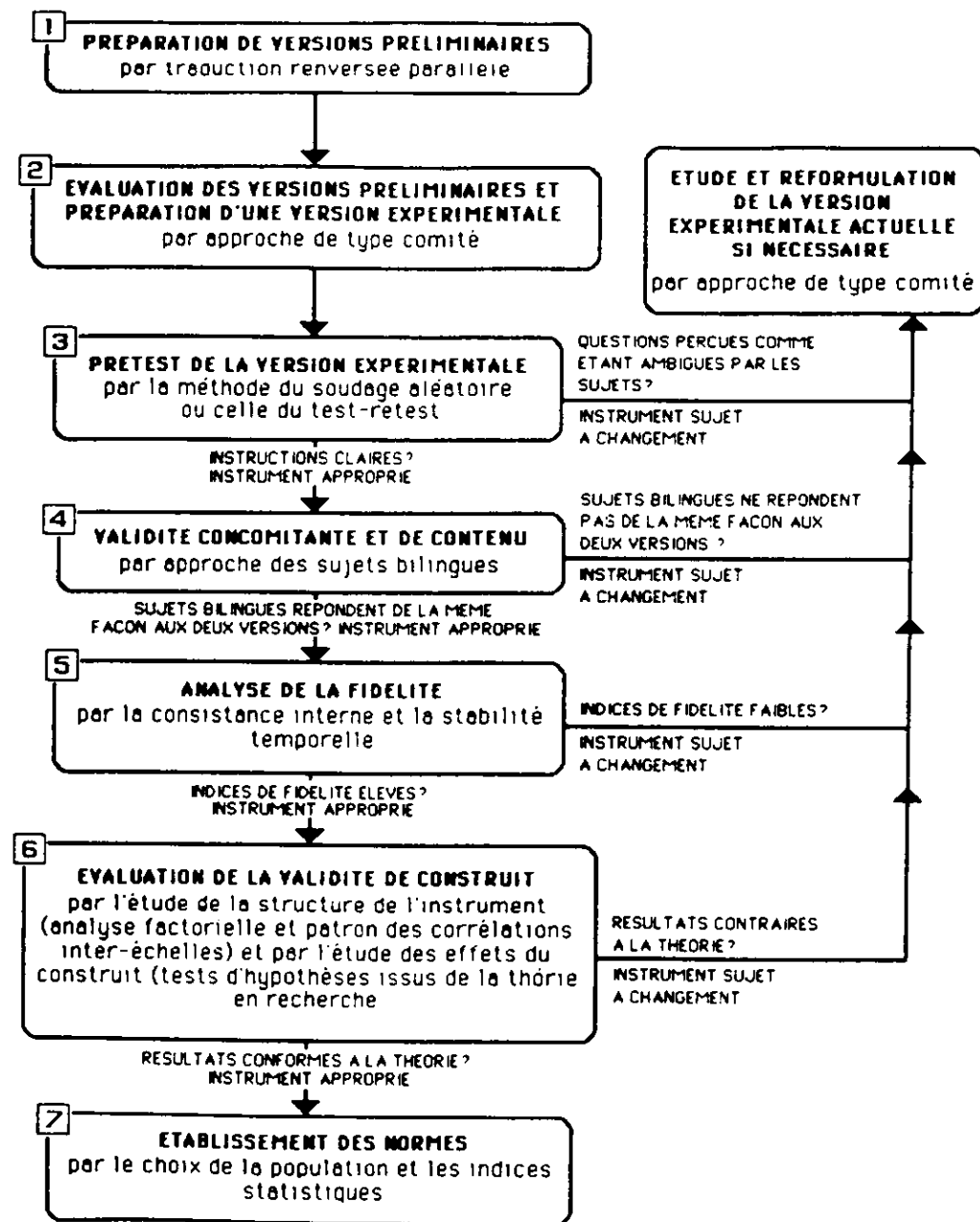


FIGURE 1

Présentation séquentielle des différentes phases systématiques de la méthodologie de validation trans-culturelle de tests psychologiques.

sciences du comportement et la *Revue québécoise de psychologie*. Ces différentes recherches ainsi que les étapes suivies dans le processus de validation de l'instrument en français sont présentées dans le Tableau 1. Bien que le Tableau 1 ne prétend pas être un compte-rendu complet des recherches sur la validation trans-culturelle d'instruments psychologiques en français, un tel tableau présente un compte-rendu révélateur et fidèle des travaux réalisés dans ce secteur. On comprendra que le Tableau 1 n'a pas pour but de rapporter les résultats de toutes les études ayant mis en français des questionnaires psychologiques anglais, mais plutôt de faire ressortir les tendances méthodologiques dans ce secteur de recherche au Canada français. Le Tableau 1 est organisé en fonction des différentes étapes proposées. On discutera maintenant de chacune de ces étapes dans l'ordre suggéré ci-dessus.

Version Préliminaire

La première étape dans le processus de validation trans-culturelle d'un questionnaire psychologique consiste en la préparation d'une version préliminaire du questionnaire original dans la langue cible (Spielberger & Sharma, 1976). De nombreuses méthodes peuvent alors être utilisées pour obtenir une traduction préliminaire adéquate. Parmi ces techniques, notons la méthode de la traduction traditionnelle, la méthode du comité et la méthode de la traduction inversée.

Bien que chacune de ces méthodes puisse être utilisée selon les besoins et les circonstances (Brislin, Lonner, & Thorndike, 1973), on doit s'empêcher de noter que les deux premières techniques ne sont pas optimales pour développer une version préliminaire d'un instrument psychologique. Ceci est particulièrement le cas de la traduction traditionnelle où le chercheur prépare seul une traduction en français de l'instrument de mesure psychologique. Les biais linguistique, psychologique et de compréhension du chercheur risquent alors de mener à certaines divergences entre la version originale et celle obtenue en français. Ces biais risquent d'être moins importants si cette technique est suivie de celle de type comité où il y a discussion de groupe sur le contenu de la traduction. Il demeure tout de même possible que les biais du chercheur persistent, même en présence du comité. Donc, même si la technique de traduction traditionnelle est très utilisée (pas moins de 7 études sur 12 l'ont utilisée, voir Tableau 1) et même si elle implique la par-

ticipation de traducteurs professionnels cette première technique n'est pas recommandée.

La seconde technique, celle du comité (e.g., Spielberger & Sharma, 1976), est très utile car, justement, elle offre une certaine garantie contre les biais du chercheur seul. En plus, lorsqu'il est possible d'obtenir la participation du chercheur qui a créé l'instrument original, cette technique offre en plus l'assurance que chacun des items sera bien compris par les chercheurs intéressés à valider en français l'instrument en question. À cause de tous ces avantages évidents, l'approche de type comité doit faire partie d'une méthodologie de validation trans-culturelle d'instruments psychologiques. Par contre, malgré ses différents avantages, la méthode du comité se prête mal à la réalisation d'une première traduction de l'instrument psychologique, compte tenu de la difficulté de traduire avec efficacité, un instrument dans son entier en contexte de groupe. L'approche par comité trouve son utilité surtout dans l'analyse de la version préliminaire de l'instrument. C'est donc à ce rôle que la méthode par comité sera confinée dans le cadre de la présente méthodologie.

Selon l'avis de plusieurs chercheurs en psychologie trans-culturelle (e.g. Brislin, 1986; Brislin et al., 1973), la troisième technique, soit celle de la traduction inversée ("back-translation"), semble idéale pour établir une version préliminaire de l'instrument psychologique. Les avantages de la traduction inversée sont au moins de deux ordres. Premièrement, ce type de traduction implique au moins deux personnes et généralement quatre. Il y a donc moins de chances que les biais d'un seul chercheur se retrouvent dans la version préliminaire. Le deuxième avantage est que, contrairement à la simple traduction traditionnelle, la traduction inversée permet de vérifier la justesse de la traduction en comparant cette dernière avec la version originale de l'instrument. Il devient alors relativement facile de déceler les lacunes découlant de la traduction de l'instrument (Brislin, 1986). Il est donc suggéré que la première étape de la validation trans-culturelle d'un instrument psychologique soit de réaliser une ou plusieurs traductions inversées de l'instrument.

La technique de la traduction inversée peut s'effectuer de plusieurs façons différents (voir Brislin, 1986). En général, cependant, cette technique nécessite que la version originale de l'instrument en question soit traduite en français par un individu bilingue. Par la suite, cette traduction

TABLEAU 1
Tests psychologiques mis en français et utilisés pour fins d'exemple

TESTS PSYCHOLOGIQUES	VERSION PRÉLIMINAIRE			COMITÉ	PRÉ-TEST	VALIDITÉ: CONTENU + CONCOMITANTE	FIDÉLITÉ		VALIDITÉ DE CONSTRUIT			NORMES							
	Traduction tradit.	Services prof.	Traduction rev.				S	P	KR	test-retest	Struct.	Diff. corr.	Effets	Pop.	M	E	P	Z	T
1. Questionnaire de dépression de Beck (1978) (Bourique & Beaudette, 1982)	x		x				.92	4 mois r = .82	x										
2. Échelle d'auto-évaluation de l'intensité de la dépression (Gauthier, Morin, Thériault, & Lawson, 1982)	x						.82	3 mois r = .75	x										
3. Questionnaire d'excitation sexuelle (Trodol, Campbell, & Daigle, 1982)																			
4. Questionnaire d'inféction sexuelle (Trodol & Dufort, 1984)	x						.74- .75												
5. Échelle d'ajustement, diadique (Baillargeon, Dubois, & Marinette, 1986)	x						.61- .91												
6. Lieu de contrôle tridimensionnel (IPAH) (Jurass, 1987)	x						.26- .93 .57- .73												
7. Échelle de conscience de soi (Cyr, Bouchard, LeComte, Lalonde, & Valiquette, 1987)	x		x				.72 .81	2 sem. r = .75- .95	x										
8. Échelle des orientations générales à la causalité (Vallerand, Lacouture, Blais, & Deci, 1987)							.63- .76	5 sem. r = .58- .71											
9. Échelle révisée de la conscience de soi (Pelletier & Vallerand, sous presse)	x						.73- .82	4 sem. r = .78- .86	x										
10. Échelle de monitoring de soi (Colavecchio, Vallerand, Blais, & Pelletier, 1988)							.74	5 sem. r = .75											
11. Échelle de satisfaction de vie (Blais, Vallerand, Pelletier & Brière, 1989)							.79 .84	2 mois r = .64	x										
12. Questionnaire de satisfaction du consommateur de services psychologiques (Sabourin, Pérusse & Gendreau, sous presse)	x						.88												

française sera remise à un second individu bilingue qui la traduira en anglais. Cette dernière traduction doit se faire sans l'aide de la version originale. En d'autres termes, il s'agit d'essayer de reproduire la version originale du test en se servant uniquement de la version française. La justesse de cette dernière version s'évalue par le degré avec lequel elle aura permis de reproduire fidèlement la version originale. Il est proposé que ce jugement soit effectué lors de la phase d'évaluation de la version préliminaire réalisée par la méthode du comité.

Probablement à cause de ces avantages, de plus en plus de chercheurs francophones utilisent la traduction inversée pour produire une version préliminaire de l'instrument désiré. Par exemple, au Canada, au cours des dernières années, la traduction inversée fut utilisée dans la validation des instruments suivants: l'Échelle de conscience de soi (Cyr, Bouchard, Lecomte, Lalonde, & Valiquette, 1987), l'Échelle des orientations générales à causalité (ÉOGC) (Vallerand, Lacouture, Blais, & Deci, 1987), l'Échelle du monitoring de soi (Colavecchio, Vallerand, Blais, & Pelletier, 1988) et l'Échelle de satisfaction de vie (Blais, Vallerand, Pelletier, & Brière, 1989)³ (voir le Tableau 1).

Bien que ces différentes recherches utilisèrent la traduction inversée, bien peu utilisèrent deux traductions inversées parallèles. Cette dernière approche est suggérée car elle permet d'éviter certains biais qui seraient dus à des caractéristiques personnelles des deux individus choisis pour réaliser la traduction inversée. Seules les études de Colavecchio et al. (1988), de Vallerand et al. (1987) et de Blais et al. (1989) ont utilisé deux traductions inversées parallèles. Par exemple, en ce qui concerne l'ÉOGC (Vallerand et al., 1987), la version originale fut remise à deux chercheurs bilingues en psychologie sociale afin qu'ils la traduisent en langue française. Par la suite, ces deux traductions françaises furent remises à autres chercheurs bilingues afin qu'ils essaient de reproduire la forme anglaise de l'ÉOGC à partir des traductions obtenues. La séquence apparaissait donc comme suit:

Traduction Inversée #1: ÉOGC anglais →
ÉOGC français — ÉOGC anglais
Individu 1 — Individu 3
Traduction Inversée # 2: ÉOGC anglais →
ÉOGC français — ÉOGC anglais
Individu 2 — Individu 4

Deux formes françaises de l'ÉOGC furent ainsi obtenues (voir Vallerand et al., 1987). Ces versions préliminaires français de l'ÉOGC purent ainsi être évaluées plus en détail par un comité d'experts lors de la phase d'évaluation afin qu'une version expérimentale française de l'ÉOGC soit constituée. Comme nous le verrons ci-dessous, cette procédure procure des avantages certains comparativement à une traduction traditionnelle ou encoure à une seule traduction inversée.

Évaluation des Versions Préliminaires

L'évaluation minutieuse de la version (ou des versions) préliminaire(s) dans le langage cible représente une phase très importante du processus de validation trans-culturelle d'un questionnaire psychologique. En effet, cette seconde phase d'évaluation permet d'étudier la valeur des versions préliminaires de sorte à obtenir une seule version française expérimentale de l'instrument.

Alors que cette phase d'évaluation se fait parfois par le chercheur seul, une telle pratique est à déconseiller dans la mesure où elle pourrait mener à certains biais linguistique, psychologique et même de compréhension. Par contre, une approche du type comité semble permettre une évaluation plus objective et précise des versions préliminaires (Spielberger & Sharma, 1976). Celle-ci est donc suggérée comme stratégie de base pour produire une version française expérimentale de l'instrument (Vallerand & Halliwell, 1983).

Comme le démontre le Tableau 1, l'approche de type comité semble relativement populaire auprès des chercheurs francophones. En effet, cette approche fut utilisée dans pas moins de 7 recherches sur 12. Il semble important de noter les différents types de comités utilisés dans les procédures de validation trans-culturelle. Un premier type de comité comprend uniquement le groupe de chercheurs eux-mêmes. Le second comprend les chercheurs ainsi que ceux qui ont participé à la préparation de la version préliminaire. Enfin, dans le troisième type de comité, s'ajoute un expert linguistique en français ainsi que l'auteur du questionnaire

³Ces divers instruments représentent des validations des questionnaires suivants: The Self-Consciousness Scale (Fenigstein, Scheier, & Buss, 1975), The General Causality Orientations Scale (Deci & Ryan, 1985), The Self-Monitoring Scale (Snyder, 1974) et The Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin (1985).

original (voir le Tableau 1 pour les différents types de comités utilisés).

Il semble important de noter que ces trois types de comité n'offrent pas tous des avantages équivalents. En effet, dans le cas du premier type de comité formé uniquement des chercheurs, il est important de souligner qu'il est sujet aux biais des chercheurs seuls, énumérés précédemment. En plus, en étant à l'abri de conseils extérieurs, il se peut fort bien que les chercheurs aient alors de la difficulté à voir certains changements à effectuer. Ce premier type de comité n'est donc pas recommandé.

Le second type de comité représente déjà une amélioration comparativement au premier puisqu'il incorpore dans le comité d'autres personnes, soit celles qui ont travaillé à l'élaboration des traductions inversées parallèles. Il devient alors possible pour le comité d'offrir des opinions autres que celles des chercheurs seuls. En plus, les personnes impliquées dans le processus de traduction inversée faisant partie du comité, ce dernier est donc en mesure de saisir pourquoi certains termes anglais furent traduits (ou retraduits) de telle ou telle façon. Enfin, plus il y a de personnes impliquées dans le comité (pourvu que tous aient droit égal de parole) moins il y a de chances qu'un biais évident provenant d'une des versions préliminaires soit incorporé dans la version expérimentale de l'instrument.

Malgré ses divers avantages, ce second type de comité comporte tout de même certaines lacunes. Premièrement, il subsiste un risque que la version expérimentale issue des travaux du comité soit quelque peu déficiente sur le plan linguistique. Deuxièmement, sans la présence de l'auteur de questionnaire psychologique original, il y a toujours risque d'erreur de compréhension de la part des membres du comité. C'est justement ces deux sources de problèmes que cherche à solutionner le troisième type de comité.

Celui-ci comprend, outre les chercheurs, les personnes ayant participé aux traductions inversées parallèles, un expert linguiste ainsi que l'auteur de l'instrument psychologique original (Vallerand & Halliwell, 1983). En plus, on peut y ajouter des chercheurs dans le secteur de recherches liées à vocation de l'instrument en question (Spielberger & Sharma, 1976). Cette situation devrait assurer que la version expérimentale issue des travaux du comité représentera bien le construit psychologique désiré et qu'elle sera rédigée dans un français à l'épreuve de toutes failles linguistiques. Ce type

de comité fut utilisé avec succès, entre autre, dans l'étude de validation trans-culturelle de l'ÉOGC (Vallerand et al., 1987).

L'évaluation effectuée par le comité doit se faire avec minutie, et de façon systématique. En termes plus concrets, cela signifie que l'évaluation de la (ou des) version(s) préliminaire(s) se fait à deux niveaux. Premièrement, chacun des items issus des deux traductions inversées anglaises est comparé aux items de la version originale anglaise. Lorsque les items provenant des traductions inversées et de la version anglaise originale sont identiques, les items en français ont satisfait un premier critère. En effet, ils ont permis un retour fidèle en langue originale et ceci représente un premier prérequis.

Un deuxième niveau d'analyse auquel le comité doit s'attarder consiste en l'étude des termes techniques utilisés dans les traductions françaises pour véhiculer les différents sens reliés au contenu psychologique. Ceci est très important. En effet, vu le vocabulaire nuancé de la langue française, il est possible que deux termes, intimement reliés mais légèrement différents, permettent des traductions renversées identiques. Ceci explique pourquoi les items français, même s'ils ont permis une traduction inversée fidèle, doivent être passés au crible par les experts linguistes et les chercheurs. Ainsi les différences d'intensité ou d'expression des énoncés pourront être respectées.

Il est à noter que si la forme française de certains items n'est pas satisfaisante, le comité se réserve alors le droit d'effectuer les changements jugés nécessaires. Cependant, ces changements, pour être incorporés dans la version expérimentale française de l'instrument psychologique, devront respecter certaines règles bien précises proposées par Spielberger et Sharma (1976, pp. 15-17), à savoir que: (1) lorsque les traductions littérales ne sont pas possibles, la version expérimentale doit tenter de dégager le sens de l'item original et non le «mot pour mot»; (2) les propriétés psycholinguistiques uniques de la langue cible doivent être utilisées autant que possible dans la formulation de la version expérimentale. Une très grande attention devra être accordée aux propriétés de la langue française telle que parlée au Canada. Ne pas tenir compte de ce dernier point pourrait résulter en la construction d'un instrument présenté dans la langue cible mais n'ayant que très peu de sens pour les membres de la population canadienne-française; (3) lorsqu'il y a mésentente entre les

TABLEAU 2

Exemples de l'item 10 des versions française et anglaise du General Causality Orientation Scale

Item 10: Anglais

You are embarking on a new career. The most important consideration is likely to be:

- | | | | | | | | | |
|--------|---|---------------|---|---|-------------------|---|---|-------------|
| (I) a) | Whether there is someone to bail you out if you get overwhelmed | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | very unlikely | | | moderately likely | | | very likely |
| (A) b) | How interested you are in that kind of work | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | very unlikely | | | moderately likely | | | very likely |
| (C) c) | Whether there are good possibilities for advancement | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | very unlikely | | | moderately likely | | | very likely |

Item 10: Français

Vous entreprenez une nouvelle carrière. Votre considération la plus importante sera probablement:

- | | | | | | | | | |
|--------|--|-----------------|---|---|---------------------|---|---|---------------|
| (I) a) | S'il a quelqu'un pour vous sortir du pétrin si vous êtes submergé(e) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | Très improbable | | | Modérément probable | | | Très probable |
| (A) b) | A quel point vous êtes intéressé(e) par ce genre de travail | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | Très improbable | | | Modérément probable | | | Très probable |
| (C) c) | S'il y a de bonnes possibilités de promotion | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | Très improbable | | | Modérément probable | | | Très probable |

Note: Tiré de Vallerand, Lacouture, Blais, et Deci (1987). Reproduit avec la permission de la Société canadienne de psychologie. I = Échelle impersonnelle, A = Échelle d'autonomie, C = Échelle de contrôle.

membres du comité sur certains items, des formes alternatives pour ces items devront être incluses dans la version expérimentale; et enfin (4) la version expérimentale française doit emprunter le même format de présentation et les mêmes directives que celles utilisées dans la version originale. L'importance de ce dernier point a été démontrée par Converse et Presser (1986) qui ont montré que des présentations matérielles différentes du même instrument peuvent mener à des résultats différents. Donc, afin d'assurer une validité trans-culturelle de l'instrument, la version française devra représenter les mêmes modalités de présentation et les mêmes directives que celles de la version originale.

Illustrons ces lignes directrices par quelques exemples. Dans le cas de la validation trans-culturelle de l'ÉOGC (Vallerand et al., 1987), le comité décida de traduire l'expression «bail

you out» par «vous sortir du pétrin» car cette dernière expression bien québécoise correspond bien à l'idée de l'item anglais. Ainsi, la règle numéro 2 ci-dessus fut respectée. D'autre part lors de la validation de l'Échelle du monitoring de soi (Colavecchio et al., 1988) les membres du comité ne s'entendaient pas sur la version française de l'item 9 («I am not particularly good at making other people like me»). Donc, en accord avec la règle numéro 4 ci-dessus, deux versions de l'item en française furent incluses dans la version expérimentale. Les évaluations subséquentes (étapes 3 et 4 - prétest et évaluation des validités concomitante et de construit) permirent de retenir le meilleur des deux items («Il m'est difficile de faire en sorte que les autres m'aiment»). Enfin, de façon conforme avec la règle numéro 4 ci-dessus, Vallerand et al. (1987) ont utilisé une formule de présentation de la version française de

l'ÉOGC équivalente à celle de la version originale (voir Tableau 2). De telles procédures sont importantes car elles contribuent à l'équivalence des deux versions de l'instrument.

Suite à cette deuxième étape, une version expérimentale de l'instrument en question devrait être prête. Bien que certains chercheurs se limitent à cette seconde étape de validation, il est important de souligner que s'arrêter ici ne garantit pas la validité et la fidélité de l'instrument. Plus encore, s'arrêter après la seconde étape ne permet pas de juger du contenu du questionnaire tel qu'il peut apparaître aux yeux des sujets. Bon nombre de recherches démontrent en effet que la perception de l'expérimentateur et celle du sujet ne concordent pas toujours (voir McAuley & Gross, 1983; Russell, 1982). Il devient donc important de vérifier comment l'instrument est perçu par des membres de la population cible. Cette étape de vérification se fait dans le cadre d'un prétest de la version expérimentale.

Prétest

Le but du prétest est de déterminer si les items qui composent la version expérimentale du questionnaire sont clairs, rédigés sans ambiguïté et ce, dans un langage qui s'apparente bien à celui de la population cible. Généralement, deux techniques peuvent être utilisées très avantageusement lors du prétest: le sondage aléatoire et la méthode du «test-retest».

La première technique de prétest, le sondage aléatoire («random-probe technique», Schuman, 1966), nécessite qu'un intervieweur choisisse au hasard une série d'items provenant du questionnaire à être évalué et demande à des sujets ($N = 20$) aléatoirement choisis d'y répondre. Puis, l'interviewer demande aux sujets d'expliquer chacune des réponses des items désignés au hasard. Si les sujets ne peuvent clairement répondre aux questions de l'interviewer, l'item est jugé ambigu et devient sujet à changement.

La deuxième technique de prétest est celle du test-retest (Mitchell, 1966) où deux interviewers questionnent les mêmes sujets ($N = 20$) à des moments différents. Si les questions du test sont claires, les réponses obtenues par les interviewers devraient être similaires.

Bien qu'utiles, ces deux techniques sont quelque peu onéreuses. En effet, elles demandent beaucoup de temps et nécessitent, selon le cas, la présence d'un ou même deux expérimentateurs.

Par souci d'économie, Vallerand et al. (1987), dans leur étude sur l'ÉOGC, ont demandé à des sujets de répondre au questionnaire de l'ÉOGC et d'encercler toute expression ou item ambigu. Les items encadrés avec régularité devaient être reformulés.

Une autre variante de l'approche précédente consiste à demander à différents sujets de lire chacun des items de l'échelle en question et de juger leur clarté sur une échelle de 1 à 7. Les items ayant obtenu des cotes moyennes faibles (4 et moins, par exemple) sont alors sujets à changement.

Les propos des sujets concernant les items problématiques sont importants car ils éclaireront les discussions et travaux des membres du comité cherchant à reformuler les items problématiques. Il devient donc essentiel que les sujets aient l'opportunité d'indiquer la nature de l'ambiguïté de l'item (ou des items) en question. Ceci peut être fait par écrit ou encore en le signifiant verbalement à l'expérimentateur qui en prendra bonne note (voir Blais, Vallerand, Brière, Gagnon, & Pelletier, 1987).

Si après le prétest, il devient évident que certains items de la version française ne sont pas clairs, le questionnaire devra alors être modifié avec l'aide d'experts linguistes lors d'une nouvelle réunion du comité. Les diverses informations fournies par les sujets lors du prétest faciliteront la tâche du comité dans la reformulation des items problématiques. Par contre si tous les items sont jugés clairs, nous pouvons alors procéder à l'évaluation de la validité concomitante et de contenu de la version française de l'instrument.

Evaluation de la validité concomitante et de contenu

La validité de contenu («content validity») réfère au jugement subjectif, et non quantitatif, à l'effet que le test en question mesure effectivement ce qu'il est sensé mesurer (e.g., Anastasi, 1976). Par contre, la validité concomitante («concurrent validity») est démontrée empiriquement lorsqu'un test est fortement corrélé avec un critère mesurant le(s) même(s) concept(s) (Allen & Yen, 1979).

En ce qui concerne la traduction d'instruments psychologiques, dans la mesure où le chercheur démontre empiriquement que la version en français est étroitement reliée à la version originale (validée en anglais), il devient alors possible de

conférer à la fois validité concomitante et validité de contenu à la version française (Spielberger & Sharma, 1976). Dans une telle optique, le recours à des sujets bilingues est particulièrement indiqué.

Deux procédures peuvent être utilisées afin d'étudier la validité concomitante et de contenu à l'aide des sujets bilingues. La première consiste à administrer la forme originale et la version expérimentale aux sujets bilingues afin d'étudier le patron des corrélations entre les deux versions. Les corrélations peuvent être effectuées entre le pointage total des versions en français et en anglais ou encore pour chacun des items français et anglais. Des corrélations positives élevées indiquent qu'il existe congruence entre les deux versions.⁴

Une seconde procédure avec les sujets bilingues consiste à demander à ces derniers de répondre à l'un de quatre questionnaires préparés à l'aide des versions en anglais et/ou en français et de comparer les réponses des sujets aux différents types de questionnaires (Vallerand & Halliwell, 1983).

Une seule étude à notre connaissance a utilisé cette technique dans des procédures de validation trans-culturelle de questionnaires psychologiques en français. Dans cette étude (Colavecchio et al., 1988), quatre groupes ($N = 20$) de sujets bilingues répondirent à l'Échelle de monitoring de soi (EMS) soit en français et/ou en anglais, mais dans un ordre contrebalancé. Les sujets bilingues furent donc aléatoirement assignés à l'un des quatre groupes suivants: le premier groupe répondit seulement à la version originale du EMS et le second seulement à la version expérimentale française du EMS; le troisième groupe répondit à la première moitié du EMS en anglais et la seconde moitié de EMS en français. Finalement, le dernier groupe répondit à la première moitié du EMS en français et la seconde moitié du EMS en anglais. À partir des réponses de chacun des groupes, une analyse de variance fut réalisée sur les pointages totaux des sujets. Afin que les versions anglaises et françaises soient jugées équivalentes, aucune

⁴La validité concomitante peut être démontrée à l'aide d'autres échelles en français et non uniquement en corrélant les versions française et anglaise de l'instrument. Par contre, le but de la présente méthodologie étant de valider en français la version originale d'un questionnaire donné, il semble que la technique des sujets bilingues offrent des avantages certains, notamment de démontrer clairement l'équivalence trans-culturelle entre les deux versions.

différence significative ($p < .10$) ne devait exister entre les quatre groupes sur le pointage total des échelles. C'est ce que les résultats indiquèrent.

Une autre technique statistique servant à mesurer l'équivalence des items, plus sévère que la précédente, consiste à effectuer des tests-t indépendants sur chacun des items anglais et français (Triandis & Davis, 1965). Des tests-t non-significatifs ($p > .10$) indiquent l'équivalence des items alors que des résultats significatifs suggèrent une non équivalence. Dans ce dernier cas, une révision de l'item de langue française est alors jugée nécessaire.

L'utilisation des sujets bilingues exige que ces derniers soient bien identifiés. Afin de déterminer le degré de bilinguisme des sujets, la méthode de Gonzalez-Reigosa (1976) telle que modifiée par Vallerand et Halliwell (1983) est recommandée. Il s'agit de demander à chacun des sujets d'indiquer sur une échelle de 1 (très peu) à 4 (parfaitement) à quel point il (ou elle) peut lire, écrire, comprendre une conversation et s'exprimer et ce, en français et en anglais. Le degré de bilinguisme est obtenu en additionnant les pointages sur les quatre dimensions pour chaque langue. Plus les pointages totaux sont élevés pour les langues françaises et anglaises, plus l'individu est jugé comme étant bilingue. Pour fins de validation trans-culturelle, il est conseillé de conserver uniquement les sujets ayant un pointage de 12 et plus pour chacune des langues française et anglaise. Bien que cette technique puisse être teintée de subjectivité de la part du répondant, elle s'avère beaucoup moins onéreuse que l'administration d'un test d'habileté dans la seconde langue d'intérêt. Par contre, lorsque possible, cette dernière approche est conseillée.

Une autre utilité de la présente étape d'étude de la version en français consiste à étudier, s'il y a lieu, lequel des items ajoutés par le comité lors de cas de mésentente s'apparente le plus à l'item original. On se souviendra que dans l'étude de Colavecchio et al. (1988) portant sur l'EMS, deux versions de l'item 9 avaient été insérées dans la version expérimentale française. Les analyses statistiques basées sur les réponses des sujets bilingues ont permis de trancher en faveur de la version la plus près de l'item original. Ainsi, la présente phase d'analyses trans-culturelles permet d'obtenir des réponses quantifiées à des questions que le comité répondrait de façon très subjective, et peut-être même

erronée, dans la phase précédente. Cette troisième étape joue donc un rôle très important dans le processus de validation trans-culturelle de tests psychologiques.

Lorsque toutes les analyses statistiques effectuées indiquent une équivalence entre les versions anglaise et française de l'instrument, il est permis de conclure qu'il existe validité de contenu et validité concomitante pour la version française de l'instrument (Spielberger & Sharma, 1976). On peut donc soutenir que la version française de l'instrument mesure le même contenu que l'instrument original. Ceci représente une étape importante dans le processus de validation trans-culturelle d'un instrument. Cependant, bien que de tels résultats nous rassurent quant au contenu de l'instrument, ils ne nous renseignent pas quant aux caractéristiques psychométriques de la version française elle-même. En effet, trois autres étapes doivent être effectuées avant que l'on puisse recommander l'utilisation de la version française de l'instrument en question. La première porte sur l'analyse de la fidélité, la seconde sur la validité de construit de l'instrument et la troisième sur l'établissement des normes de l'instrument. Nous traiterons de ces trois étapes dans les prochaines sections de l'article.

Analyse de la fidélité

La fidélité d'un test psychologique concerne la précision de l'instrument peu importe ce qu'il mesure (e.g., Nunnally, 1970). Un test fidèle mesure toujours le construit psychologique de la même façon. Il est fiable. Plusieurs indices de fidélité sont utilisés dans la littérature scientifique. Ils ne seront pas tous repris dans cette section. Le lecteur intéressé pourra trouver dans Anastasi (1976) et Cronbach (1984) une riche source d'informations à ce sujet. Les recherches récentes sur la validation de questionnaires psychologiques et la validation trans-culturelle de mesures psychologiques soulignent l'utilisation de deux types d'indices, soit les indices de consistance interne ou d'homogénéité et ceux de stabilité temporelle.

Lorsque le test psychologique est mesuré sur une échelle continue (telle une échelle de style Likert, ce qui est généralement le cas de la plupart des échelles contemporaines), l'analyse de l'alpha de Cronbach est recommandée.⁵ Une telle analyse permet d'évaluer le degré avec lequel les items corrélerent entre eux, ou si l'on veut, à quel point les items sont homogènes tout en tenant compte du nombre d'items dans le test en question. La formule est la suivante:

$= Np / [1 + p(N-1)]$ où N correspond au nombre d'items dans l'échelle et p représente la corrélation interitem moyenne.

La valeur de l'alpha peut varier entre 0 et 1.00. Plus les pointages sont élevés et plus l'instrument est jugé comme possédant un niveau élevé de consistance interne. Par contre, un pointage trop élevé (plus de .90 par exemple) démontre une certaine redondance dans les items et peut laisser supposer que les items de l'échelle en question mesurent un aspect très (trop) restreint du concept visé (McIver & Carmines, 1981). Des valeurs variant entre .70 et .85 sont donc généralement préférables. Une inspection du Tableau 1 révèle que l'analyse de consistance interne est faite dans la plupart des études (11/12) portant sur la validation trans-culturelle d'instruments en français. En plus, on peut remarquer que les indices de consistance se situent généralement près des valeurs recommandées.

La deuxième mesure de fidélité qui nous concerne porte sur les propriétés de stabilité temporelle de l'instrument. Les propriétés de stabilité temporelle s'évaluent en demandant à des sujets de répondre à l'instrument à deux reprises avec un écart de temps relativement important entre les deux passations de l'instrument. Un écart d'environ 1 mois s'avère généralement suffisant. Si l'instrument est fidèle, les pointages des sujets devraient demeurer stables même après qu'un certain laps de temps se soit écoulé entre les deux passations de l'instrument⁶. Des indices de corrélation test-retest de l'ordre de .60 et plus pour une telle période, s'avèrent satisfaisants. Le Tableau 1 révèle des corrélations test-retest généralement très acceptables. On peut

⁵Par contre, lorsque l'échelle de mesure n'est pas continue mais plutôt dichotomique, il est alors préférable d'utiliser la formule 20 de Kuder-Richardson. Cette formule est pratiquement la même que celle de l'alpha de Cronbach si ce n'est que le nombre de personnes qui ont "passé" (ou bien répondu) aux items est pris en considération au lieu de la somme des variances de l'écart-type des items du test (voir Nunnally, 1970).

⁶Cette statistique ne s'applique qu'aux instruments qui sont sensés mesurer des construits relativement stables sur une base temporelle. Certains instruments mesurent les fluctuations situationnelles — par exemple le Causal Dimension Scale, de Russell (1982). Il n'est alors donc pas approprié de mesurer un indice de stabilité temporelle dans de tels cas. Voir Vallerand et Richer (1988) pour une analyse factorielle confirmatoire du Causal Dimension Scale.

également remarquer que ces indices semblent diminuer progressivement avec le temps. Enfin, il est possible de noter que cet indice de stabilité temporelle n'est malheureusement pas toujours rapporté.

Une approche intéressante est celle de Haccoun (1987) qui permet de vérifier la validité concomitante et de contenu ainsi que la fidélité test-retest de l'instrument dans une seule étape. Haccoun propose que l'on demande à un même groupe de sujets bilingues de répondre aux versions française et anglaise de l'instrument à deux reprises avec un intervalle de temps entre les deux administrations. Il devient alors possible de vérifier les relations entre les formes originale et traduite de l'instrument ainsi que de comparer l'ampleur des corrélations test-retest pour les versions anglaise et française de l'instrument. Ces dernières corrélations devraient être similaires. Donc, dans l'ensemble, l'approche proposée par Haccoun se veut économe et fort utile.

Deux points s'imposent concernant les indices de fidélité obtenus dans le cadre de la présente étape. Premièrement, les pointages obtenus avec les versions françaises peuvent difficilement excéder les valeurs de la version originale. Ceci se comprend par le fait que le but de la démarche de validation trans-culturelle est de reproduire de façon aussi juste que possible chacun des items de la version originale en français. Lorsque ce but est atteint on devrait donc obtenir des indices très près de ceux obtenus par la version originale. Dans la majorité des cas, la version dans la langue seconde n'étant pas identique à l'originale, les indices seront un peu plus faibles que ceux de cette dernière. Dans l'ensemble, et de façon générale, nous sommes donc en droit à nous attendre que les indices de fidélité de la version originale imposent une limite supérieure aux indices pouvant être obtenus par la version validée en français. Donc, le standard de comparaison concernant la qualité de la fidélité de la version traduite devrait inclure non seulement les pointages optimaux escomptés, mais également les indices obtenus par la version originale de l'instrument. Ceci permettrait de vérifier à quel point les propriétés psychométriques de la version en langue française s'apparentent à celles de la version originale.

Un second point d'importance concernant les analyses de fidélité obtenues avec la version canadienne-française consiste dans le fait qu'elles fournissent des informations supplémentaires

concernant les items qui sont sujets à changement. Dans ce contexte, les analyses de consistance interne sont très utiles. En effet, en plus de fournir l'indice de l'alpha de Cronbach, l'analyse de consistance interne ajoute également d'autres indices tels la corrélation entre chaque item et le pointage total de l'échelle (corrige en enlevant l'item en question) et l'alpha de Cronbach si chacun des items est enlevé de l'échelle. Il est ainsi possible d'identifier les items qui ne cadrent pas avec les autres, de comparer ces items avec les énoncés originaux et de les modifier ultérieurement en comité.

Par exemple, dans la première étude de validation de l'instrument de mesure du Lieu de contrôle tridimensionnel (IPAH), Jutras (1987) avait obtenu un indice alpha de .26 pour l'échelle I (Lieu interne). En analysant tous les indices mentionnés ci-dessus, il fut observé qu'un certain nombre d'items étaient problématiques. Cette constatation mena à des changements ultérieurs ainsi qu'à une seconde étude qui permit d'augmenter la consistance interne de l'échelle de .26 à .57. Bien que cet indice soit légèrement faible, il est tout de même comparable à celui de .64 de la version originale (Levenson, 1973).

Il est à retenir que même après avoir démontré la validité de contenu et la validité concomitante de la version traduite lors de l'étape précédente, des modifications d'items peuvent s'imposer lors de l'analyse de la fidélité de l'instrument. Ainsi, s'il devient évident à ce stade-ci de la démarche de validation trans-culturelle que certains items ou encore que l'échelle dans son ensemble ne se comporte pas comme prévu, alors un retour à une approche de comité s'impose afin de pouvoir résoudre les problèmes (voir Figure 1). Dans plusieurs cas, une phase d'expérimentation supplémentaire devrait permettre de résoudre les problèmes associés à l'échelle. Par contre, si les résultats à cette étape-ci démontrent la fidélité de l'instrument de mesure, il devient alors possible de passer à l'étude de sa validité de construit.

Evaluation de la Validité de Construit

L'évaluation de la validité de construit ("construct validity") de la version traduite d'un test psychologique consiste à vérifier si celle-ci permet de bien mesurer le construit tel que défini par son cadre théorique spécifique. Cette évaluation vise aussi à s'assurer que la version traduite est assez sensible pour détecter les effets du construit hypothétique propre à la théorie, de

concert avec les hypothèses émises par cette dernière (Novick, 1985). L'importance d'une telle évaluation tient du fait qu'il est possible qu'un questionnaire possède une validité concomitante et de contenu sans pour autant manifester une validité de construit. Il devient donc important de démontrer que la théorie qui sous-tend l'instrument psychologique est valide dans la culture canadienne-française, sinon l'utilité du questionnaire et la validité de la théorie sous-jacente à ce dernier seraient remises en question. L'évaluation de la validité de construit d'un questionnaire psychologique peut se faire à trois niveaux, soit au niveau de la structure de construit lui-même, au niveau des relations entre les différents construits inhérents à la structure théorique ainsi qu'en ce qui concerne les conséquences et corrélats du ou des construits.

Structure du Construit Psychologique

Au premier niveau d'analyse, l'évaluation de la validité de construit d'un questionnaire psychologique consiste à démontrer que ce dernier est bel et bien formulé de façon conséquente à la théorie qui le sous-tend. Ainsi, si une théorie sur l'amour propose qu'il existe trois types d'amour (Sternberg, 1986), alors un questionnaire qui propose de mesurer le construit d'amour tel qu'élaboré par le théoricien devrait être formé de trois échelles mesurant les trois types d'amour en question. Dans ce contexte, l'utilisation de l'analyse factorielle peut s'avérer fort pertinente pour tester la validité de construit de la version traduite en français. En effet, dans la mesure où ce type d'analyse permet d'étudier les regroupements entre les items, il devient possible de vérifier si les résultats démontrent bel et bien la présence de trois facteurs correspondant aux types d'amours proposés par la théorie. Donc, l'utilisation de l'analyse factorielle peut s'avérer fort utile dans une démarche cherchant à établir la validité de construit⁷ d'une version traduite en français (voir Kim & Mueller, 1978 et Briggs & Cheek, 1986 pour une discussion plus approfondie concernant les applications de l'analyse factorielle dans le cadre du développement d'instruments de mesure).

En observant le Tableau 1, il est possible de remarquer que l'utilisation de l'analyse factorielle est de plus en plus utilisée dans

l'évaluation de la validité de construit des versions françaises d'instruments psychologiques. En effet, pas moins de 8 études ont utilisé cette technique statistique. L'utilisation de ce type d'analyse statistique est probablement due d'une part à l'évolution des connaissances et habiletés en statistiques chez les chercheurs en psychologie. L'analyse factorielle fait maintenant partie de l'arsenal régulier du chercheur. Il devient donc relativement aisé de l'utiliser en recherche de validation trans-culturelle. D'autre part, l'utilisation accrue de l'analyse factorielle peut également s'expliquer par le fait que lors des procédures de validation des versions originales elles-mêmes, les chercheurs utilisent l'analyse factorielle. Ainsi, il est logique d'essayer de reproduire les résultats originaux avec la version canadienne-française de l'instrument en utilisant la même technique statistique.

Il semble important de noter brièvement qu'il existe deux grandes classes d'analyses factorielles: les analyses factorielles exploratoire et confirmatoire. Les chercheurs utilisant l'analyse factorielle *exploratoire* soumettent les données au logiciel en essayant de reproduire la structure originale. Ils observent alors le nombre de facteurs obtenus ainsi que les saturations des items sur les différents facteurs. Dans la mesure où le nombre de facteurs obtenus et la saturation des items correspondent aux résultats originaux, ceci soutient la dimension structurelle de la validité de construit de l'instrument. Par contre, si les résultats ne coïncident pas avec les résultats de la version originale, il faut retourner en comité et reformuler les items problématiques.

C'est ce type d'analyse que Cyr et al. (1987) ont utilisé afin d'étudier la validité de construit de la version française de l'Échelle de conscience de soi (Fenigstein, Scheier, & Buss, 1975). Il est intéressant de noter que les résultats d'une première analyse factorielle n'avaient supporté que partiellement la théorie tripartite de la conscience de soi (conscience de soi privée et publique, et anxiété sociale) de Fenigstein et al. (1975). Cyr et al. ont alors procédé à des changements sur un item (l'item 2). Les résultats d'une seconde analyse factorielle devaient donner raison aux chercheurs: la structure à trois facteurs et les saturations des items sur les trois facteurs correspondaient à la structure factorielle originale.

D'autre part, l'utilisation de l'analyse factorielle *confirmatoire* permet de vérifier statistiquement si la forme traduite correspond à la

version originale. Ceci peut se faire de deux façons. Dans un premier temps, le chercheur peut soumettre à l'analyse confirmatoire de type LISREL (Jöreskog et Sörbom, 1984 - ou autres), les réponses à la version française en spécifiant le nombre de facteurs qui devraient être obtenus ainsi que la position de chacun des énoncés sur les divers facteurs. Divers indices d'adéquation sont alors obtenus supportant ou infirmant la structure postulée. Par exemple, Blais et al. (1989) ont testé la présence d'un seul facteur regroupant tous les énoncés de l'Échelle de satisfaction de vie. Les résultats ont confirmé la structure unidimensionnelle de l'instrument.

Dans un second temps, l'analyse factorielle confirmatoire peut être utilisée afin de comparer directement l'équivalence des matrices de covariance sur lesquelles reposent les structures factorielles des versions originale et traduite. Le chercheur utilisant l'approche confirmatoire peut donc évaluer statistiquement si les structures factorielles issues des versions originale et traduite sont équivalentes. En plus, dans la mesure où les deux versions ne seraient pas équivalentes, certains indices statistiques indiquent au chercheur les items sujets à modification. Ces informations peuvent donc être utilisées par la suite par le comité dans sa démarche de reformulation des items insatisfaisants.

Au meilleur de notre connaissance, une seule étude de validation trans-culturelle en français a fait usage de cette seconde possibilité offerte par l'analyse factorielle confirmatoire. Il s'agit de l'étude de Pelletier et Vallerand (sous presse), sur la validation trans-culturelle en français de l'Échelle révisée de la conscience de soi (Scheier & Carver, 1985; voir aussi Pelletier & Vallerand, 1987 pour d'autres résultats avec cet instrument). Dans cette étude, Pelletier et Vallerand ont comparé la matrice de covariance de l'étude de validation de Scheier et Carver (1985) obtenue avec des sujets américains à celle obtenue avec des sujets québécois. Les données furent soumises au logiciel LISREL (Jöreskog & Sörbom, 1984) afin d'évaluer l'équivalence des deux matrices de covariance. Les résultats de l'analyse confirmatoire révélèrent que les deux matrices n'étaient pas significativement différentes l'une de l'autre. Ces résultats supportèrent donc l'équivalence des deux matrices de covariance démontrant que la structure factorielle des versions originale et traduite pouvait être jugée équivalente. Il va sans dire que ce type d'analyse factorielle confirmatoire représente un atout précieux pour le cher-

cheur intéressé à la validation trans-culturelle d'instruments psychologiques. Il apparaît souhaitable qu'elle soit de plus en plus utilisée en recherche trans-culturelle.

Relations entre les différents construits inhérents au modèle théorique

En plus de démontrer que les items du questionnaire se regroupent sur les facteurs ou sous-échelles de la façon postulée par la théorie, il devient également important de démontrer que les facteurs ou sous-échelles sont également reliés entre eux de façon conforme à la théorie. (Bien sûr, ceci ne s'applique pas dans le cadre d'échelles unidimensionnelles). Ceci se vérifie en évaluant les corrélations entre les sous-échelles du questionnaire. Une telle vérification est particulièrement importante si des corrélations négatives sont postulées entre certaines sous-échelles et des corrélations positives sont prévues entre d'autres sous-échelles. Des corrélations en accord avec la théorie apportent un soutien additionnel pour la composante structurelle de la validité de construit de la version traduite de l'instrument.

De telles analyses de corrélations ont été réalisées avec succès par Cyr et al. (1987) et Vallerand et al. (1987) dans leurs études de validation trans-culturelle des Échelles de conscience de soi et de l'ÉOGC, respectivement. Ces corrélations étaient toutes en accord avec celles obtenues avec les versions originales (voir Deci & Ryan, 1985 pour l'ÉOGC, et Fenigstein, Scheier & Buss, 1975 pour l'Échelle de la conscience de soi).

Effets ou corrélats du construit psychologique

Tel qu'indiqué précédemment, le troisième niveau d'analyse de la validité de construit d'un instrument porte sur les conséquences ou corrélats du construit psychologique. Il s'agit d'évaluer si l'instrument servant à mesurer le construit psychologique produit des effets conformes aux hypothèses issues de la théorie. Ceci se fait généralement en réalisant une étude avec la version traduite. Deux types d'études peuvent être réalisées, soit des reproductions intégrales d'études déjà réalisées avec la version originale, soit la réalisation de nouvelles études. Le premier type d'études représente un prérequis essentiel à toute tentative de validation trans-culturelle de questionnaire alors que le second, tout en n'étant pas toujours essentiel, peut apporter un éclairage nouveau sur les propriétés de

⁷Anastasi (1976) et Allen et Yen (1979) utilisent le terme "validité factorielle" pour décrire ce type de validité de construit étudiée à l'aide de l'analyse factorielle.

l'instrument dans la culture canadienne-française.

Le Tableau 1 révèle qu'un nombre restreint d'études (seulement 5 études) ont analysé la composante "effet" de la validité de construit de la version canadienne-française de l'instrument en question. L'étude de Colavecchio et al. (1988) représente un exemple d'une étude dans laquelle les auteurs ont tenté de reproduire des résultats rapportés par Snyder (1974) dans sa validation de l'Échelle de monitoring de soi (habileté à se présenter sous différents angles et à agir de façon conséquente aux situations). Snyder avait alors démontré que des acteurs obtiennent des pointages plus élevés sur l'échelle que des étudiants universitaires. Colavecchio et al. ont reproduit ces résultats avec exactitude. Ces résultats supportent donc cet aspect de la validité de construit de la version canadienne-française de l'Échelle de monitoring de soi.

Parfois, les résultats ne mènent pas à une reproduction des résultats originaux. De tels résultats pourraient renvoyer à l'une des trois possibilités suivantes: a) il y a une faiblesse dans l'étude en question, b) un manque de validité du questionnaire de langue française est évident ou c) le questionnaire est adéquat mais le construit psychologique ne produit pas les mêmes effets avec la culture canadienne-française qu'avec la culture américaine. C'est dans un tel contexte qu'une seconde étude peut s'avérer utile. En effet, cette dernière peut permettre de trancher entre les trois hypothèses. On ne peut donc pas négliger l'importance du contexte expérimental dans le processus de validation trans-culturelle d'instruments psychologiques.

Le second type d'étude de validation de construit, soit la réalisation de nouvelles études, ne devrait être réalisé que lorsque les résultats avec le premier type d'études se sont avérés concluants. Dans un tel cas, il devient alors possible d'étudier en profondeur, et de façon originale, la validité du construit dans une nouvelle culture. Les résultats de telles études offrent deux avantages non négligeables. Premièrement, ils permettent une analyse plus fouillée de la validité de construit de l'instrument en question dans la culture cible. Et, deuxièmement, ils viennent ajouter aux nouvelles connaissances scientifiques dans le secteur en question. Par contre, il demeure important de ne pas brusquer les choses et de s'assurer que la reproduction d'études (le premier type d'études de validité de construit) a bel et bien été réalisée préalablement. En effet, sans ce premier

type d'études, on ne pourrait tirer avec certitude les implications de résultats opposés à ceux postulés par la théorie en regard de la validité de la version française de l'instrument.

Lorsque la version canadienne-française de l'instrument psychologique démontre des propriétés psychométriques adéquates concernant la fidélité et les trois types de validité de construit, l'instrument peut alors être utilisé en recherche. Toutefois, la normalisation de l'instrument offrirait des avantages supplémentaires à son utilisateur.

Etablissement des normes

Etablir les normes d'un instrument psychologique représente la dernière étape du processus de validation trans-culturelle. L'importance de cette étape ne doit pas être sous-estimée. En effet, comme le mentionne Anastasi (1976), en l'absence de normes, un score brut d'un individu sur un test psychologique donné ne veut rien dire. Ainsi, il importe, en premier lieu, de faire un choix des statistiques qui permettent de comparer l'individu avec un groupe de référence approprié. Généralement, les moyennes, écarts types, rang percentile ainsi que les scores z et T devraient suffire.

Les moyennes et écarts types sont utiles car ils permettent, entre autre, de comparer les versions originale et canadienne-française de l'instrument. Il devient alors possible d'établir si les sujets canadiens-français se distinguent ou non des sujets anglophones sur les moyennes des deux versions. Il est à noter que des différences dans les moyennes n'infirment pas nécessairement la version en langue française. Il se peut fort bien que ces différences soient dues à l'échantillon choisi. D'autre part, il est également possible que la culture canadienne-française se distingue sur la dimension psychologique en question. Par contre, des différences énormes entre les deux versions du questionnaire pourraient indiquer que certains problèmes se posent au niveau de la version française de l'instrument. Il devient également intéressant de comparer les pointages des hommes et des femmes des deux cultures. Bien qu'il est possible, voire courant, d'obtenir des différences entre les deux sexes, il devient intéressant de vérifier si ces différences se retrouvent dans les deux cultures. Enfin, les écarts types permettent de vérifier le niveau de dispersion des pointages des sujets et de les comparer à celui des pointages de la version originale

de l'instrument. Normalement, les écarts types devraient être similaires et d'ampleur modérée.

Le rang percentile représente un autre indice normatif d'importance. Un score percentile d'une personne indique le pourcentage de personnes dans le groupe critère qui ont obtenu un pointage inférieur à celui de cet individu. Par exemple, si un individu a obtenu un score brut de 28 sur un questionnaire et que ceci correspond à un rang percentile 97, cela signifie que 97% de la population à laquelle il est comparé a obtenu un pointage plus faible que lui sur cette dimension psychologique. Les rangs percentiles permettent donc de comparer les sujets à un groupe de référence.

Bien qu'ils offrent une information utile, les rangs percentiles sont quelque peu limités car ils n'offrent aucune indication relative à la distance entre les pointages. Les scores standardisés, par contre, permettent de situer les individus sur une échelle à intervalle. Ceux-ci sont de différents ordres. Il y a entre autre ce qu'on appelle les scores z et les scores T . Les scores z sont obtenus à l'aide de la formule suivante:

$$z = \frac{X - M}{S}$$

où z = score standardisé, X = un score brut donné, M = la moyenne du groupe de comparaison, et s = l'écart type du groupe de comparaison. Quant aux scores T , ce sont des scores z transformés en multipliant les scores z obtenus par 10 et en additionnant 50. Ceci situe ainsi la moyenne à 50 et l'écart type à 10. La formule est la suivante:

$$T = 50 + 10z$$

La majorité des chercheurs semblent préférer les scores T aux scores z . Nous croyons que ceci est purement matière d'opinion. L'objectif demeure toujours de fournir l'information concernant le groupe de comparaison le plus clairement possible.

Gauthier, Morin, Thériault, et Lawson, (1982), dans leur validation de la version canadienne-française du *Beck Depression Inventory* (Beck, Ward, Mendelson, Moch, & Erbauch, 1961) ont présenté des normes fort complètes. En effet, les scores bruts, les scores T , les percentiles ainsi que les scores z sont tous présentés. Il devient alors possible de retirer les informations précises et pertinentes à partir de telles normes. Ceci est tout aussi vrai pour le chercheur à la recherche de sujets ayant des pointages extrêmes sur le questionnaire en question

que pour l'intervenant intéressé à comparer le pointage d'un client à celui de la population générale afin de mieux saisir l'ampleur du problème de celui-ci.

Le choix des groupes de comparaison représente la deuxième décision qui doit être prise lors de cette étape d'établissement de normes. Quels seront les groupes qui serviront de critères de comparaison? La réponse à cette question est directement liée à la vocation de l'instrument psychologique. Si le test psychologique a pour but de mesurer des tendances *générales* chez les individus (e.g., questionnaire d'estime personnelle), alors des groupes composés seulement d'étudiants ne peuvent pas servir d'échantillon critère. Par contre, si nous désirons mesurer les perceptions que les étudiants ont d'eux-mêmes dans différents domaines de leur vie, alors seuls des étudiants devraient servir de groupe critère. Il en va de même pour les instruments d'ordre clinique (e.g., instruments mesurant la dépression, les problèmes chez le couple, etc.). En effet, afin que l'interprétation de la comparaison individus-normes soit vraisemblable, il faut que l'unité de comparaison soit la même.

Dans leur étude de validation trans-culturelle du *Beck Depression Inventory*, Gauthier et al. (1982) ont choisi de rapporter les normes obtenues auprès de la population générale. Ceci représente un excellent choix de population. En effet, des normes basées sur des étudiants universitaires ne sauraient être utiles aux cliniciens intéressés à utiliser l'instrument dans leur pratique ou dans des travaux scientifiques réalisés auprès de la population générale.

Donc clairement, les instruments psychologiques mesurant certaines caractéristiques dans des domaines précis nécessitent l'utilisation de populations spécifiques pour générer des normes. Pour fins de comparaisons à but multiples, il est suggéré que, lorsque possible, les normes soient établies en fonction de différents groupes d'âges, du sexe et même le cas échéant, pour les sous-groupes spécifiques auxquels l'instrument est destiné.

Nous espérons que cette section a su mettre en évidence l'importance de produire des normes pour une version canadienne-française d'un instrument psychologique. Puisqu'il est peu indiqué de se servir des normes anglophones pour fins de comparaison, le chercheur soucieux de rendre service à la collectivité scientifique francophone saisira bien l'importance de développer des normes complètes. Malheureusement, une analyse du

Tableau 1 révèle qu'une telle pratique est peu courante. En effet, on peut noter que les normes rapportées se limitent principalement aux moyennes et aux écarts types. Il est à espérer que dans un avenir prochain des normes plus complètes seront rapportées dans les procédures de validation trans-culturelle d'instruments de mesure.

Une fois les normes de l'instrument établies, le chercheur peut alors diffuser les résultats des différentes étapes du processus de validation trans-culturelle qu'il a poursuivies, et rendre l'instrument psychologique disponible aux chercheurs de son secteur d'intérêt. Ceci devrait par la suite favoriser un essor de la recherche et une saine coopération entre les chercheurs de ce secteur.

Conclusions

L'objet du présent document est de formuler une méthodologie de validation trans-culturelle de tests psychologiques. On y propose sept étapes distinctes: (1) la préparation minutieuse d'une version préliminaire; (2) l'évaluation de cette version préliminaire et la préparation d'une version expérimentale; (3) l'évaluation de la clarté des questions de la version expérimentale dans un prétest; (4) l'évaluation de la validité concomitante et de contenu du questionnaire; (5) l'évaluation de la fidélité; (6) l'étude de la validité de construit de l'instrument psychologique dans le langage cible; et (7) l'établissement des normes de l'instrument. On accorde à chacune de ces étapes, une attention toute spéciale à l'utilisation d'indices statistiques afin d'éclairer le cheminement du chercheur dans un processus qui risque toujours d'être trop subjectif.

Tout au cours de cette présentation, on présente des exemples afin de faciliter la compréhension des différentes méthodes proposées. L'utilisation de tels exemples permet d'explicitier le comment et le pourquoi des différentes techniques proposées. De plus, on fait ressortir certaines erreurs qui peuvent survenir lors de la validation trans-culturelle d'instruments psychologiques.

En terminant, il nous semble indiqué de faire trois considérations concernant le processus de validation trans-culturelle de tests psychologiques. Premièrement, une telle méthodologie, contrairement à une simple traduction, est très exigeante. En effet, elle implique un processus dynamique dans lequel la nouvelle version de

l'instrument est soumise à des évaluations (et selon le cas, des changements) d'étape jusqu'à ce que la preuve soit faite que cette version représente un instrument psychologique fidèle et valide en français. La validation trans-culturelle d'instruments psychologiques doit donc être considérée avec autant d'égards que la validation de tests originaux. En effet, elle demeure avant tout la démonstration qu'un instrument donné est bel et bien fidèle et valide tel que transposé en langue française. Une simple traduction de l'instrument ne pourrait donc suffire.

Notre seconde considération est à l'effet que sauter des étapes ne permet pas toujours de sauver du temps. Bien au contraire! Le chercheur qui n'a pas pris le temps d'effectuer un prétest ou qui n'a pas étudié la validité de contenu de l'instrument à l'aide de sujets bilingues risque de se rendre compte dans phases ultérieures que certains items sont problématiques. Tôt ou tard il risque d'avoir à reformuler les items fautifs et à recommencer là où les problèmes se sont manifestés. Avoir à recommencer à un stade avancé demande souvent plus de temps et d'énergie que la réalisation complète des démarches initiales qui ont été négligées. Certains problèmes du genre ont été présentés en exemple dans le texte. On a pu voir que les coûts d'une telle stratégie (e.g., reprendre une étude afin de réaliser une analyse factorielle sur une version révisée de l'instrument) sont souvent énormes. D'autres coûts importants passent, par contre, inaperçus lorsque les changements nécessaires ne sont pas effectués. En effet, l'échelle risque d'être utilisée en recherche alors que son équivalence à la version originale n'a pas été établie, pouvant produire ainsi des résultats erronés et trompeurs.

Enfin, troisième considération, nous exprimons l'espoir qu'à la lumière de la méthodologie présentée dans le présent article, le lecteur saura dorénavant se méfier des questionnaires traduits à la sauvette. En effet, si des changements s'imposent après plusieurs étapes dans le processus de validation trans-culturelle, nous sommes en droit de questionner la valeur d'instruments qui n'en sont rendus qu'aux premières étapes d'un tel processus. Si à l'avenir, le lecteur sait se montrer plus critique dans son analyse des traductions et validations d'instruments psychologiques qu'il rencontre dans la littérature scientifique, nous estimons que cet article aura atteint l'un de ses buts principaux.

ABSTRACT

The cross-cultural validation of psychological questionnaires from English to French is a methodological problem of interest to an increasing number of French-Canadian researchers. In addition, no systematic cross-cultural validation methodology is available to psychologists interested in performing such a task. The purpose of this paper is to fill this gap by proposing a cross-cultural validation methodology allowing validation in French of English psychological scales. This methodology comprises seven (7) steps: (1) the preparation of a preliminary version, (2) the evaluation and modification of the preliminary French version, (3) the evaluation of the experimental French version in a pretest, (4) the evaluation of the content and concomitant validity of the French version, (5) the evaluation of the reliability of the French version, (6) the evaluation of the construct validity of the French version, and (7) the preparation of norms for the French version of the instrument. The application of the proposed methodology is exemplified with recent social psychological research.

Références

- Allen, M.J., & Yen, W.M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.
- Anastasi, A. (1976). *Psychological Testing* (4th edition). New York: MacMillan.
- Baillargeon, J., Dubois, G., & Marineau, R. (1986). Traduction française de l'Echelle d'Ajustement Dyadique. *Revue Canadienne des Sciences du Comportement*, 18, 25-34.
- Beck, A.T., Ward, G.H., Mendelson, M., Nock, J.E., & Endaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Blais, M.R., Vallerand, R.J., & Brière, N.M., Gagnon, A., & Pelletier, L.G. (1989). Significance, structure and gender differences in life domains of college students. Manuscript soumis pour fins de publication.
- Blais, M.R., Vallerand, R.J., Pelletier, L.G., & Brière, N.M. (1989). L'Echelle de Satisfaction de Vie: Validation Canadienne-Française du "Satisfaction with Life Scale". *Revue Canadienne des Sciences du Comportement*, 21, 210-223.
- Bourque, P., & Beaudette, D. (1982). Etude psychométrique du questionnaire de dépression de Beck auprès d'un échantillon d'étudiants universitaires francophones. *Revue Canadienne des Sciences du Comportement*, 14, 211-218.
- Briggs, S.R., & Cheek, J.M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54, 106-148.
- Brislin, R.W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1, 185-216.
- Brislin, R.W. (1986). The wording and translation of research instruments. In W. Lonner & J. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137-164). Beverly Hills, CA: Sage.
- Brislin, R.W., Lonner, W.J., & Thorndike, R.M. (1973). *Cross-cultural research methods*. New York: John Wiley & Sons.
- Cattell, R.B., Eber, H.W., & Tatsuoka, M.M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
- Colavecchio, P., Vallerand, R.J., Blais, M.R., & Pelletier, L.G. (1988). L'Echelle Révisée du Monitorage de Soi: Validation canadienne-française du "Revised Self-Monitoring Scale". Manuscript inédit, Université de Québec à Montréal.
- Converse, J.M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire*. Monographie de la série "Quantitative applications in the social sciences", 07-063. Beverly Hills, CA: Sage.
- Cronbach, L.J. (1984). *Essentials of psychological testing* (4e ed.). New York: Harper & Row.
- Cyr, M., Bouchard, M.-A., Lecomte, C., Lalonde, F., & Valiquette, C. (1989). Analyse psychométrique d'une adaptation en langue française de l'Echelle de Conscience de Soi. *Revue Canadienne des Sciences du Comportement*, 19, 287-298.
- Deci, E.L., & Ryan, R.M. (1985). The "General causality Orientations Scale": Self-determination in personality. *Journal of Research in Personality*, 19, 109-134.
- Diener, E., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, 49, 71-76.
- Fenigstein, A., Scheier, M.F., & Buss, A.H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 43, 522-527.
- Gauthier, J., Morin, C., Thériault, F., & Lawson, J.S. (1982). Adaptation française d'une Mesure d'Auto-évaluation de l'Intensité de la Dépression. *Revue Québécoise de Psychologie*, 3, 13-27.
- Gonzalez-Reigosa, F. (1976). The anxiety-arousing effect of taboo work in bilinguals. In C.D. Spielberger and R. Diaz-Guerrero (Eds.), *Cross-cultural anxiety*. Washington: Hemisphere.
- Haccoun, R.R. (1987). Une nouvelle technique de vérification de l'équivalence de mesures psychologiques traduites. *Revue Québécoise de Psychologie*, 8, 30-38.
- Hui, C.N., & Triandis, H.C. (1985). Measurement in cross-cultural psychology: A review of comparison strategies. *Journal of Cross-Cultural Psychology*, 16, 131-152.
- Hulin, C.L., Drasgow, F., & Komokar, J. (1982). Application of item response theory to analysis of attitude scale translations. *Journal of Applied Psychology*, 67, 818-825.
- Izard, C.E., Dougherty, F.E., Bloxom, B.M., & Kotsch, W.E. (1974). *The Differential Emotions Scale: A method of measuring the subjective experience of discrete emotions*. Manuscript inédit, Université de Vanderbilt.
- Jöreskog, K.G., & Sörbom, D. (1984). *LISREL VI*. Chicago: National Educational Resources.

- Jutras, S. (1987). L'IPAH, Version canadienne-française de l'Echelle de Levenson mesurant le Lieu de Contrôle Tridimensionnel. *Revue canadienne des sciences du comportement*, 19, 74-85.
- Kim, J.-O., & Mueller, C.W. (1978). *Factor analysis: Statistical methods and practical issues* (Monographie de la série: "Quantitative Applications in Social Sciences", no. 14). Beverly Hills, CA: Sage.
- Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology*, 41, 397-404.
- McAuley, E., & Gross, J.B. (1983). Perception of causality in sport: An application of the Causal Dimension Scale. *Journal of Sport Psychology*, 5, 72-76.
- McIver, J.P., & Carmines, E.G. (1981). *Unidimensional scaling* (Monographie de la série: "Quantitative Applications in Social Sciences", no. 24). Beverly Hills, CA: Sage.
- Mitchell, R. (1966). *The problems and possibilities of measuring social attitudes in African social surveys*. Communication présentée à la 9e conférence annuelle de l'Association pour les Etudes Africaines, Bloomington, Indiana.
- Novick, J.R. (1985). *Standards for educational and psychological testing* (préparé par l'American Education Research Association, l'American Association, et le National Council on Measurement in Education). Washington, DC: The American psychological Association.
- Nunnally, J.C., Jr. (1970). *Introduction to psychological measurement*. New York: McGraw-Hill.
- Pelletier, L.G., & Vallerand, R.J. *The Revised Self-Consciousness Scale: A French-Canadian replication*. Manuscrit inédit, Université de Québec à Montréal.
- Pelletier, L.G., & Vallerand, R.J. (sous presse). *Validation canadienne-française de l'Echelle Révisée de la Conscience de Soi*. Manuscrit soumis pour fins de publication.
- Russell, D. (1982). The Causal Dimension Scale: A measure of how individuals perceive causes. *Journal of Personality and Social Psychology*, 42, 1137-1145.
- Sabourin, S., Pérusse, D., & Gendreau, P. (sous presse). Les qualités psychométriques de la version canadienne-française du Questionnaire de satisfaction du consommateur de services psychothérapeutiques (QSC-8 et QSC-18B). *Revue Canadienne des Sciences du Comportement*.
- Scheier, M.F., & Carver, C.S. (1985). The Self-Consciousness Scale: A revised version for use with general populations. *Journal of Applied Social Psychology*.
- Schuman, H. (1966). The random probe: A technique for evaluating the quality of closed questions. *American Sociological Review*, 31, 218-222.
- Snyder, M. (1974). The self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30, 526-537.
- Spielberger, C.D. & Sharma, S. (1976). Cross-cultural measurement of anxiety. In C.D. Spielberger and R. Diaz-Guerrero (Eds.), *Cross-cultural anxiety* (pp. 13-25). Washington: Hemisphere.
- Sternberg, R.J. (1986). A triangular theory of love. *Psychological Review*, 93, 119-135.
- Triandis, H.C., & Davis, E. (1965). *Some methodological problems concerning research on negotiations between monolinguals*. Rapport technique No. 28, Group Effectiveness Research Laboratory, Urbana, Illinois.
- Trudel, G., Campbell, M., & Daigle, R. (1982). Etude psychométrique du Questionnaire d'Excitation Sexuelle. *Revue Québécoise de Psychologie*, 3, 83-87.
- Trudel, G., & Dufort, F. (1984). Le Questionnaire d'Interaction Sexuelle: Une analyse psychométrique. *Revue Québécoise de Psychologie*, 5, 27-37.
- Vallerand, R.J., Blais, M.R., Gagnon, A., Brière, N., & Pelletier, L.G. (1987). *Les domaines importants de la vie des jeunes étudiants du niveau collégial: Etude exploratoire*. Données brutes disponibles au Laboratoire de Psychologie Sociale, Université du Québec à Montréal.
- Vallerand, R.J., & Halliwell, W.R. (1983). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la Psychologie du sport. *Canadian Journal of Applied Sport Sciences*, 9-18.
- Vallerand, R.J., Lacouture, Y., Blais, M.R., & Deci, E.L. (1987). L'Echelle des Orientations Générales à la Causalité: Validation canadienne-française du General Causality Orientations Scale. *Revue Canadienne des Sciences du Comportement*, 19, 1-15.
- Vallerand, R.J., & Richer, F. (1988). On the use of the Causal Dimension Scale in a field setting: A test with confirmatory factor analysis in success and failure situations. *Journal of Personality and Social Psychology*, 54, 704-712.