

الإطار المفاهيمي والمنهجي لتكييف الاختبارات والمقاييس النفسية عبر الثقافات

A conceptual and methodological framework for adapting psychological tests and measures across cultures

نسيمة علي تودرت *

جامعة الجزائر 2 (الجزائر)، nassima.alitoudert@univ-alger2.dz

تاريخ الاستلام: 2023/10/17 تاريخ القبول: 2023/10/21 تاريخ النشر: 2023/11/21

Abstract:

ملخص:

The article discusses the problem of adapting and using psychological tests across cultures. Psychological tests have become used throughout the Arab environment, but due to the poor transfer process that was conducted for them; they suffer from several problems and shortcomings. Most of them give false answers, and most of their content does not agree with the culture of society. The shortcomings in it make it vulnerable to cultural biases, which are considered the biggest threat to the validity of the adapted tests. In this article, we have tried to emphasize the conceptual framework and methodological issues in adapting psychological tests in other languages and cultures and to suggest procedures for verifying their equivalence. We first touched on some important concepts, such as the concept of cross culture, the concept of standardization, the difference between translation and adaptation, as well as an explanation of the concept of bias and its types (construct bias, method bias, and item bias) and the concept of equivalence. We also touched on the methodological framework, distinguishing various methods that can help identify bias and consequences of cross-cultural equivalence in test scores, in addition to methods that help avoid these consequences, by presenting the instructions of the International Testing Commission (ITC) that focus on procedures for developing and adapting the test. As well as procedures for applying and administering the test, and finally procedures for interpreting the results. We also presented instructions from the University of Minnesota that call for studying the equivalence of the translated test at the level of linguistic equivalence, structural equivalence, psychometric equivalence, and finally psychological equivalence.

يتناول المقال إشكالية تكييف الاختبارات النفسية واستخدامها عبر الثقافات. حيث أن الاختبارات النفسية أصبحت مستعملة في كامل البيئة العربية، لكن نظرا لسوء عملية النقل التي أجريت لها جعلها تعاني من عدة مشكلات ونقائص فأغلبها تعطي إجابات مزيفة ومعظمها لا يتفق محتواها مع ثقافة المجتمع. فالنقائص الموجودة فيها جعلها عرضة إلى الانحيازات الثقافية والتي تعتبر أكبر مهدد لصدق الاختبارات المكيفة. وقد حاولنا في هذا المقال التأكيد على الإطار المفاهيمي والقضايا المنهجية في تكييف الاختبارات النفسية في اللغات والثقافات الأخرى واقتراح إجراءات التحقق من تكافؤها. فتطرقنا أولا إلى بعض المفاهيم المهمة مثل مفهوم العبر ثقافي ومفهوم التقنين والفرق بين الترجمة والتكييف وكذا شرح مفهوم الانحياز وأنواعه (انحياز البنية، وانحياز المنهج، وانحياز البند) ومفهوم التكافؤ. كل هذه المفاهيم تعتبر مهمة لفهم الإجراءات المنهجية لعملية التكييف عبر الثقافات. تطرقنا أيضا إلى الإطار المنهجي فقمنا بتمييز مختلف الأساليب التي يمكن أن تساعد في تحديد التحيز والعواقب للتكافؤ بين الثقافات في درجات الاختبار. بالإضافة إلى الأساليب التي تساعد على تجنب هذه العواقب، وذلك من خلال عرض تعليمات اللجنة الدولية للاختبارات (ITC) التي تركز على إجراءات تطوير وتكييف الاختبار وكذا إجراءات تطبيق وإدارة الاختبار وأخيرا إجراءات تفسير النتائج. عرضنا أيضا تعليمات جامعة Minnesota التي تحث على دراسة تكافؤ الاختبار المترجم على مستوى التكافؤ اللغوي، والتكافؤ البنوي، والتكافؤ السيكومتري، وأخيرا التكافؤ السيكولوجي.

Keywords: Test Adaptation; Equivalence; Bias; Cross culture; International Test Commission (ITC).

كلمات مفتاحية: تكييف الاختبار؛ التكافؤ؛ الانحياز؛ عبر ثقافي؛ اللجنة الدولية للاختبارات (ITC).

*المؤلف المرسل

1. مقدمة

إن الوطن العربي اليوم بحاجة إلى الاستفادة من خبرة الآخرين في ميدان تطوير الاختبارات النفسية بما يلائم ظروفه وثقافته، ولذلك نجد أن معظم هذه الدول لا تعد اختبارات النفسية، لأن ذلك يتطلب جهداً علمياً فائقاً، فتلجأ إلى تعديل الاختبارات التي ظهرت في الدول الغربية المتقدمة، ويتطلب هذا القيام ببحوث علمية حول هذه الاختبارات كونها تمثل فئة خاصة من البحث العلمي في ميدان القياس النفسي والعقلي والتربوي تعرف ببحوث التكيف والتقنين، وذلك بهدف أن تصبح هذه الاختبارات أكثر ملائمة للظروف والثقافة الجديدة. وعندما نحلل حركة الاختبارات النفسية في العالم العربي بوجه عام، نجد أن أغلب هذه الاختبارات لا تتناسب مع مختلف نواحي الحياة النفسية والتربوية والمهنية، فمعظمها مترجم وفقط أو أعيدت استخراج خصائصها السيكمومترية على الأكثر، وحتى الاختبارات التي يدعى أصحابها أنها مقننة لا تستوفي كامل شروط التقنين. وبما أن بحوث تكيف الاختبارات تندرج ضمن منظور الدراسات عبر الثقافية (Cross-cultural studies)، التي تحاول المقارنة ما بين الثقافات في أكثر من مجال، سوف نوضح في هذا المقال أهمية تكيف الاختبارات النفسية بدل من ترجمتها وتقنيها مع التأكيد على الاعتبارات النظرية والمنهجية التي لا بد مراعاتها في عملية تكيف الاختبارات النفسية عبر الثقافات.

أحد الأهداف المهمة لدراسات علم النفس عبر الثقافات هو تقييم عالمية أو خصوصية المفهوم النفسي أو شرح الاختلافات والتشابهات في سلوكيات الأشخاص من ثقافات مختلفة (Poortinga, & Malpass, 1986) والافتراض الأساسي للبحث عبر الثقافات هو أن هناك "هوية ذات أبعاد dimensional identity" (Frijda, & Jahoda, 1966) وهو ما يشجع الباحثين على وصف الثقافات في نظام تصنيفي متطابق. ويمكن تأكيد هذا الافتراض من خلال إنشاء التكافؤ بين الثقافات (Berry, 1980) لأداة القياس. وبدون وجود دليل على التكافؤ على عدة مستويات مختلفة، لا يمكن تفسير أي اختلافات ملحوظة بين الثقافات بشكل موثوق. حيث أنه قبل أن يتم تطبيق أداة ما بنجاح في ثقافة أخرى، يجب حل المشكلات المتعلقة بتطبيق الاختبارات النفسية عبر الثقافات. قبل البدء في تعديل الاختبار، إذ يجب على المترجم التعامل مع عدد من المسائل الأولية. على سبيل المثال، يجب تحديد ما إذا كانت مهمة إجراء الاختبار المطلوبة ذات صلة بالثقافة المستهدفة وما إذا كانت الاستجابات للبنود من المحتمل أن تكون متكافئة عبر الثقافات. وفي هذا السياق أشار (Lonner, 1986) إلى أنه يجب على مكيف الاختبار عبر الثقافات أن يكون على دراية بالاختلافات الثقافية المحتملة في مجموعات الاستجابة، ويجب أن يكون مستعد لتقييم الاختلافات المحتملة والتعويض عنها. على سبيل المثال، في بعض الثقافات، قد يجيب الأفراد ببساطة على الأسئلة بدافع الأدب ثم يقدمون فقط الإجابات المرغوبة اجتماعياً. والبعض الآخر، الغير متعودون على الإجابة على الاختبارات النفسية، قد يستجيبون بطريقة مهمة. هنا قد يفترض مكيف الاختبار بسداجة أن المتقدمين للاختبار في الثقافات الأخرى يشعرون بنفس الارتياح أمام تلك المقاييس الغربية (Draguns, 1984).

2. الإطار المفاهيمي

1.2. مفهوم عبر-الثقافي أو بين-الثقافي Cross-cultural / Interculturel

ظهر مصطلح العبر-ثقافي (Cross-Cultural) إلى الوجود في ثلاثينات القرن العشرين على يد عالم الاجتماع George Murdock (1985-1987) إثر الأبحاث والدراسات الإحصائية المقارنة التي أجراها على المجتمعات المختلفة.

والتي أظهرت له أشكالاً متعددة من الفروق الثقافية بين هذه المجتمعات. ومن هذا المنطلق أصبح موضوع العبر-ثقافي نقطة انطلاق للعديد من الدراسات التي أجريت لتوضيح أثر هذه الفروق في الجوانب السلوكية والاستعدادات الفطرية والنفسية والاجتماعية وغيرها، عندما يحدث نوع من التواصل بين شخصين أو مجموعة من الأشخاص ينتمون لثقافات مختلفة، أو بمعنى آخر تحديد الدور الذي تلعبه هذه الفروق الثقافية في عملية الاتصال بين الثقافات المختلفة. وقد اكتسبت الأبحاث والدراسات المتعلقة بالفروق الثقافية بين المجتمعات المختلفة مزيداً من الأهمية بعد الحرب العالمية الثانية خاصة في الخمسينات والستينات.

وقد أوضح (De-Carlo, 1998, p. 40) أن مفهوم عبر الثقافي (Interculturel) له أصل فرنسي؛ ففي سياق مهاجري سنوات السبعينات بفرنسا الذين عانا أطفالهم من صعوبات مدرسية ظهرت فكرة أن الفوارق ليست معرقلات؛ إنما مصدر إثراء متبادل وذلك في حالة تجنيدها. وهو نفس الوضع الذي وجد بكندا حسب ما أشار إليه (Ferréol & 175, p. 175). وتري (Jucquois, 2003, p. 49) أن ارتباط كلمة «عبر» (Inter) بمصطلح عبر الثقافي (Interculturel) تتضمن علاقة أو تحديدا ما يرتبط بالآخر. ف"عبر الثقافي" يأخذ بعين الاعتبار مجموع التفاعلات القائمة بين الأفراد وجماعات الانتماء، بمعنى مواجهات الهوية فهو لا يتماشى مع واقع موضوعي؛ إنما هو بصلة ذاتية تتموضع في فضاء وزمن محددين. فهو عملية تحليل تعطي للموضوع المناقش صفة عبر الثقافية. ومن جهته، اعتبر (Hofstede, 1991) «عبر الثقافي» سيرورات ديناميكية وليدة التفاعلات ما بين الثقافات؛ أي أنها تعد نمطا خاصا من التفاعلات والتبادلات العلائقية التي تحدث حينما تدخل ثقافات مختلفة في اتصال وكذا مجموع التغيرات والتحويلات التي تفرزها. وأوضح Hofstede أنها مجموع العمليات النفسية والعلائقية والجماعية والمؤسسية، التي تحدث بفضل تفاعلات الثقافات فيما بينها في إطار علاقات متبادلة ومن منظور المحافظة على هوية ثقافية نسبية للشركاء في العلاقة.

وتلقي المجتمعات الغربية المعاصرة اهتماما كبيرا على ظاهرة البينثقافية حالياً، خاصة بسبب المشكلات الناجمة عن تزايد معدلات الهجرة إلى أوروبا وإلى أمريكا وحاجة هذه المجتمعات للمهاجرين بسبب تناقص عدد السكان وانخفاض معدلات الإنجاب في معظم مجتمعات شمال غرب أوروبا. كما أن صعوبة انخراط عدد لا يستهان به من هؤلاء المهاجرين – إن لم نقل العدد الأكبر من المهاجرين القادمين من بلاد العالم الإسلامي – في المجتمعات الأوروبية جعل هذه المجتمعات تنتبه للخطر الذي يحيط بها، إذ لم تقم بتسيخ ظاهرة البينثقافية في المجتمع وتوضيح الجوانب المتعلقة بـ "التعدد الثقافي" (Multiculturalism) و "التنوعية الثقافية" (Cultural Pluralism).

2.2. الفرق بين علم النفس الثقافي وعلم النفس العبرثقافي

يهتم علم النفس الثقافي بدراسة تأثير الثقافة على البناء النفسي للإنسان في حين يدرس علم النفس عبر الثقافي سلوكية معينة عبر ثقافات عديدة. كما يختلف علم النفس الثقافي عن علم النفس عبر الثقافي بكون الثقافة في الأخير وسيلة لاختبار عالمية العملية السيكلولوجية واختلافها بين الثقافات، بدلا من تحديد أثر السلوكيات الثقافية على العمليات السيكلولوجية.

إلى جانب ذلك نجد أن علم النفس الثقافي يرفض فكرة العالمية، أما علم النفس عبر الثقافي فيرى أن النفس البشرية لا يمكن فصلها عن العناصر الثقافية التي تنفرد بها كل منها. من ناحية أخرى، يختلف علم النفس الثقافي عن علم النفس بين الثقافات في أن علماء النفس بين الثقافات يستخدمون بشكل عام الثقافة كوسيلة لاختبار عالمية العمليات النفسية بدلاً من تحديد كيفية تشكيل الممارسات الثقافية المحلية للعمليات النفسية. و تخبرنا أبحاث علم النفس الثقافي عن العديد من المجالات داخل علم النفس، و هي تشمل علم النفس الاجتماعي وعلم النفس النمائي وعلم النفس المعرفي. ومع ذلك، يميل المنظور النسبي لعلم النفس الثقافي إلى الصدام مع وجهات النظر العالمية الشائعة في معظم مجالات علم النفس.

يرى Richard Shweder، أن النتيجة الرئيسية للطريقة التعميمية لعلم النفس بين الثقافات هي الفشل المتكرر فيما يتعلق بتكرار النتائج المخبرية الغربية في الإعدادات الغير غربية. وفي المقابل يمكن القول أن الهدف الرئيسي من علم النفس الثقافي هو امتلاك ثقافات متباينة تحدد النظريات النفسية الأساسية من أجل تحسين أو توسيع هذه النظريات، وهذا لتصبح أكثر ملائمة للتوقعات والأوصاف والتفسيرات الخاصة بكل السلوكيات البشرية وليس السلوكيات الغربية فقط. كما يمكن تعريف علم النفس الثقافي بأنه مجال في علم النفس والذي يفترض فكرة أن الثقافة والعقل لا ينفصلان، وأنه ربما تكون النظريات النفسية المتأصلة في ثقافة واحدة محدودة في قابليتها للتطبيق، وذلك عند تطبيقها على ثقافة مختلفة. وكما كتب Shweder وهو أحد الأنصار الأساسيين لهذا المجال: "إن علم النفس الثقافي هو دراسة طريقة تنظيم التقاليد الثقافية والممارسات الاجتماعية وتحديددها وتحويلها للنفس البشرية، مما يؤدي إلى وجود وحدة نفسية للجنس البشري أقل من الاختلافات العرقية في العقل والنفس والعاطفة" (Shweder, 1999, p. 72).

3.2. ترجمة الاختبارات Translation

يعرف Patsula و Hambleton (1999) الترجمة العلمية للاختبارات النفسية على أنها عملية الوصول إلى مفردات وتعابير متعادلة لغويا وثقافيا ونفسيا، للغة وثقافة أخرى انطلاقا من لغة وثقافة معينة، أو هي ترجمة الكلمات وتعابير مكونة من عدة كلمات بذات التواتر في اللغتين مع الأخذ بعين الاعتبار اللهجات الموجودة داخل اللغة الواحدة (p. 22).

4.2. التكيف Adaptation

يعرف Hambleton التكيف بأنه كل الأنشطة والعمليات بما فيها تقرير ما إذا كان باستطاعة الاختبار تقدير بنية الاختبار نفسها في لغة وثقافة أخرى، مع محافظة الاختبار على بنيته الأصلية، في معادلة صيغتين من اختبار واحد في ثقافتين مختلفتين لغويا وثقافيا ونفسيا والتي تشمل الاختلافات اللغوية الثقافية، قواعد الدرجات، لغة التعليمات، ألفة الاختبار، تكافؤ البنود والبنية الاختبار، إدارة الاختبار، تكافؤ المنهج، صدق الاختبار المكيف، خصوصية المجموعة المستهدفة ومعايير الاختبار (Hambleton, & Bollwark, 1991, p. 4).

5.2. الفرق بين الترجمة والتكيف

هناك فرق كبير بين مصطلح الترجمة ومصطلح التكيف، ففي مجال التقييم عبر الثقافات، يعتبر المصطلح تكيف Adaptation أجدر بالترجيح من مصطلح ترجمة Translation كونه لا يدل ضمنا على ترجمة حرفية. و تعتبر

عملية تكيف الاختبارات بصورة نموذجية أكثر مرونة، حيث تسمح باستعضاضات لفظية أكثر تعقيدا بحيث يكون المعنى المقصود مصونا عبر اللغات (Geisinger, 1994, p. 305). حيث أنه من الأفضل استخدام عملية التكيف بدل من الترجمة مع الاختبارات المنقولة ثقافيا، فعملية الترجمة هي جزء فقط من عملية التكيف، عند إعداد اختبار تم إعداده للاستخدام في لغة وثقافة واحدة للتطبيق في لغة وثقافة أخرى. أين يتضمن تكيف الاختبارات النفسية كل الأنشطة بدءا من تقرير ما إذا كان باستطاعة الاختبار تقدير بنية الاختبار ذاتها في لغة وثقافة أخرى، إلى اختيار المترجمين، وإلى تقرير التكيف المناسب الذي يجب القيام به لإعداد الاختبار من أجل الاستعمال في لغة أخرى (Hambleton & Bollwark, 1991, p. 5). وعلى هذا الأساس يمكن القول أن مصطلح التكيف مناسب أكثر من مصطلح الترجمة، وذلك أن المترجمين يحاولون الحصول على مفاهيم، و مفردات و تعابير متعادلة ثقافيا، و نفسيا ولغويا مع تلك المترجم منها، وبذلك تأخذ الترجمة العلمية للاختبارات النفسية أبعادا أكثر من ترجمة محتويات الاختبار حرفيا.

6.2. التقنين Standardization

التقنين بالمعنى الواسع للكلمة هو تحديد وضبط كل ما من شأنه أن يؤثر في أداء المفحوص، وبهذا المعنى يعتمد التقنين على رسم خطة شاملة وواضحة ومحددة لجميع خطوات بناء الاختبار وإجراءاته وطريقة تطبيقه وتصحيحه وتفسير درجاته، وتحديد السلوك المطلوب من المفحوص والشروط المحيطة به في أثناء تطبيق الاختبار، بالإضافة إلى وجود معايير لتفسير الدرجات.

بمعنى آخر، هو العملية التي يتم من خلالها التحكم في العوامل غير المناسبة التي يمكن أن تؤثر في عملية القياس وذلك من خلال تخفيض أخطاء القياس إلى حدها الأدنى عن طريق اختبار عينة ممثلة لمجتمع الدراسة يطبق عليها اختبار؛ ومن ثم توحيد فقراته وإجراءات تطبيقه وتصحيحه بشكل يوفر للاختبار خصائص سيكومترية تتفق مع خصائص الاختبار الجيد، ومن ثم توفير المعايير المناسبة لتفسير الدرجات الخام (زيدان، 1979، ص. 68).

7.2. الانحياز والتكافؤ Bias and Equivalence

ترتبط كلمة الانحياز بوجود عوامل ضارة (مزعجة) Nuisance factors بالقياس. إذ يعتبر القياس منحازا إذا تأثرت نتائج نسختين مختلفتي اللغة لأداة قياس ما بشكل مختلف بسبب مصدر تباين غير مرغوب فيه. لا يعد الانحياز صفة جوهرية للاختبار وإنما هي نتيجة لتطبيق هذا الاختبار على مجموعة خاصة بغية الوصول إلى هدف محدد. ويشير الانحياز إلى جميع أنواع العوامل المزعجة التي تعيق تفسير اختلاف الدرجات بين مجموعة وأخرى، ويفهم الانحياز بشكل أفضل من خلال القابلية للتعميم.

أما مصطلح التكافؤ فيرتبط ذهنيا بقياس أوجه الاختلاف بين الأعراف وما ينتج عن ذلك يسمى انحياز، أي أن البند أو الأداة المنحازة ستعطي درجات غير متكافئة، كما يشير عدم التكافؤ إلى عدم إمكانية المقارنة بين الدرجات، وعملا بهذا العرف فإننا نستعمل عدم التكافؤ كصفة مميزة لدرجات الاختبار التي تأثرت بفعل الانحياز الثقافي (Van de Vijver, & Poortinga, 2005, pp. 41-42).

إن التكافؤ والانحياز مرتبطان بشكل وثيق، حتى أن بعض المؤلفين يستخدمون هذين المصطلحين للدلالة على المعنى ذاته، بعد التعرف على هذه المصطلحات أصبح من الممكن دراسة العلاقات بشكل مفصل (Van de Vijver, & Poortinga, 2005, p. 49). حيث أصبح التكافؤ كمؤشر يشير إلى عدم إمكانية المقارنة بين الدرجات، وعملاً بهذا العرف فإننا نستعمل عدم التكافؤ كصفة مميزة لدرجات الاختبار التي تأثرت بفعل الانحياز الثقافي، وفي نفس الوقت لا يمكن اعتبار الانحياز كصفة جوهرية للإخبار، وإنما هي نتيجة لتطبيق هذا الاختبار على مجموعة معينة من أجل الوصول إلى هدف محدد، ويشير الانحياز إلى جميع أنواع العوامل المزججة التي تعيق تفسير اختلاف الدرجات بين مجموعة وأخرى، ويفهم الانحياز بشكل أفضل من خلال القابلية للتعميم.

يمكن تعريف الانحياز على أنه توافق غير متوازن بين ميدان الملاحظات وعالم التعميم، ومثال على ذلك ما افترضه أننا طبقنا اختبار الامتداد الزمني (اختبار يقيس سعة الذاكرة قصيرة المدى) على أطفال أمريكيين وأطفال من ريف إفريقيا، لم يحصلوا على تعليم جيد في المدرسة، ونجد أن الأطفال الأمريكيين قد حصلوا على درجة أفضل. فلو فسرت درجات الاختبار على أنها تمثل عدد الأرقام التي يمكن أن يحتفظ بها أطفال المجموعتين في الذاكرة قصيرة المدى فإن الاختبار سوف يكون غير منحاز، وفي المقابل إذا اعتبرت درجات سعة الأرقام على أنها تمثل قدرة الذاكرة قصيرة المدى -وهو التفسير الشائع- فإن الاختبار سوف يصبح منحازاً. فالحقائق في علم النفس عبر الثقافي تؤكد أن سعة الذاكرة قصيرة المدى شديدة التباين من ثقافة إلى أخرى، كما أن الاختلافات الملحوظة في درجات اختبار سعة الأرقام سوف لن تتكرر لو استعملنا أدوات قياس تستعمل فيها مثيرات (بنود) تتميز بأكثر صدق بيئي بالنسبة لأطفال الريف الإفريقي. ففي دراستنا للتحيز عادة ما يطرح السؤال حول التغيرات التي من شأنها تفسير الفروق الملحوظة بين الثقافات (هل هي التعليمية، أم المثير، أم إجراءات الإجابة، أم المعاينة، أم التطبيق، أم طريقة التصحيح؟ كما أن اختلاف الحافز بين الثقافات له تأثير شامل على الدرجات في كثير من الاختبارات (Van de Vijver, & Poortinga, 2005, p. 42).

1.7.2. أنواع الانحياز التي تؤثر في التقييم عبر الثقافات

على الرغم من أن الانحياز قد ينشأ من عدة مصادر، إلا أن (Van de Vijver, & Tanzer, 1997) حددا ثلاثة أنواع من مصادر الانحياز التي يمكن أن تؤثر في التقييم عبر الثقافات، و عبر اللغات هي: الانحياز في البنية، الانحياز في المنهج والانحياز في بنود الاختبار. وقد ناقشوا أيضاً تأثير نوعين من التكافؤ على الأبحاث والتقييم عبر الثقافات والمتمثلة في: تكافؤ البنية وتكافؤ وحدة القياس. (Cook et al. 2005, p. 173).

1.1.7.2. انحياز البنية

يقصد بهذا النوع من الانحياز، ذلك التباين الموجود في بنيات المفاهيم بين مجموعات الثقافات المختلفة، مثل العناصر التي تشكل البنية (السلوك، والاتجاهات، والمعايير السلوكية) والتي تكون غير متطابقة بين مجموعات الثقافات المختلفة. فأنصار نسبية المواقف الثقافية كـ Sinha (1997)؛ وعلم النفس الثقافي كـ Kule (1996)؛ Miller (1997) يميلون إلى الاعتقاد أن انحياز البنية هو قاعدة وليس استثناء في علم النفس عبر الثقافي (Van de Vijver, & Poortinga, 2005, p. 42).

وقد عرف (Van de Vijver, & Tanzer, 1997) انحياز البنية بأنها تحدث إذا كانت البنية المقاسة غير متماثلة عبر المجموعات الثقافية، وقد استعملوا كمثال اختبارات الذكاء الغربية. وقد اعتبر هؤلاء الكتاب بشكل أساسي أن انحياز البنية ليس مصطلحا ينطبق على أداة ما ولكنه ينطبق على خصائص مقارنة عبر الثقافات، حيث أشاروا إلى أن الأداة التي تكشف الانحياز في مقارنة موضوعات بين اليابانية والألمانية قد لا تكشف الانحياز في مقارنة بين الألمانية والهولندية. (Cook, et al, 2005, p. 173). ومن الأمثلة على انحياز البنية نجد مفهوم الذكاء، حيث تميل معظم اختبارات الذكاء إلى استخدام تعريف ضمني للذكاء يتألف من المحاكاة والتفكير المنطقي (كما هو الحال في اختبارات Raven) وبدرجة أقل من المعارف المكتسبة و الذاكرة (كما هو الحال في سلم Wechsler للذكاء عند الأطفال وعند البالغين) (Sternberg, 1981). Conway, Keton & Bernstein, 1981). إلا أن الدراسات في الأوساط غير الغربية بينت أن المفهوم الشائع للذكاء أوسع من ذلك ويتضمن بعض المظاهر الاجتماعية، مثل ما قالت الأمهات في كينيا، أن الطفل الذي هو الطفل المطيع الذي لا يسبب مشاكل.

وقد بينت كذلك دراسات أخرى مماثلة كدراسة Serpell (1993) في زمبيا؛ ودراسة Kashiwagi و Azuma (1987) في اليابان أن صفات الشخص الذي تتجاوز ميدان المدرسة الذي يُعتمد عادة في الولايات المتحدة الأمريكية وأوروبا، وكمثال آخر على اختلاف في مضمون الذكاء يمكن أن يشاهد في أعمال Ho (1996) حول طاعة الوالدين في الصين، الذي بيّن أنه بالمقارنة مع الغرب يميل الصينيون إلى تطبيق تعريف أوسع للذكاء، وأن الطاعة والاحترام للوالدين هي عناصر موجودة أيضا في الغرب، إلا أن المفهوم الصيني لاحترام الوالدين يشمل أيضا العناية المادية بالوالدين عندما يتقدمان في السن ويحتاجان إلى مساعدة (as cited in Van de vijver, & Poortinga, 2005, pp. 42-44).

2.1.7.2. انحياز المنهج

هو مصطلح عام لنوع ثان من الانحياز الذي يتضمن كل التحولات المزعجة الناتجة عن عوامل خاصة بالمنهج وقد ابتكر هذا المصطلح لأن هذه العوامل تذكر عادة في القسم الخاص بالمنهج في الدراسات التجريبية (الإمبريقية). وهناك نوعان من انحياز المنهج هي:

- **انحياز الأداة:** يتضمن هذا النوع من الانحياز خصائص الأداة التي لا علاقة لها بهدف الدراسة، إلا أنها مع ذلك تسبب اختلافات في درجات الاختبار. ونجد أن السبب الأكثر شيوعا لانحياز الأداة يكون في الاختبارات الذهنية وهذا ما يعرف بتألف الأشخاص مع المنبهات والاستجابات (شكل الإجابة)، وقد تم توضيح ذلك في دراسة Serpell (1979) الذي اهتم بالمهارات الإدراكية/الحسية عند الأطفال البريطانيين والزامبيين، وقد طلب منهم استنساخ صور أشخاص باستخدام الورق وقلم الرصاص، والأسلاك المعدنية (التي هي شائعة في زمبيا). فكما هو متوقع تفوق الأطفال البريطانيون في الرسوم المعمولة بالورق وقلم الرصاص بينما حصل الزامبيون على درجات أعلى بشكل واضح في الرسوم المعمولة بالأسلاك المعدنية، في حين لم نسجل أي اختلاف باستعمال أدوات ووسائل غير معروفة في كلا البلدين، والتفسير هنا يعود بالدرجة الأولى إلى الألفة للأداة أو التألف مع الإجابة (as cited in Van de vijver, & Poortinga, 2005, p. 45). وقد استخدم (Van de Vijver, & Tanzer, 1997) في مناقشة انحياز الأداة "الإثارة المألوفة Stimulus familiarity" كمثال

معروف جدا. وقد توسعوا في ذلك المثال بالاستشهاد بأعمال Triandis و Hiu (1989) اللذين توصلا إلى أن الأمريكيين من أصول لاتينية يميلون إلى اختيار الحد الأقصى في مدرج قياس الشخصية ذي خمس درجات بنسبة أكبر من الأمريكيين البيض. واكتشف Triandis و Hiu أن الارتباك الحاصل بسبب تفضيل الخيار قد اختفى عندما استعمل مقياس ذو عشر درجات (as cited in Cook, et al, 2005, p. 173).

■ **الانحياز في التطبيق:** يقصد بهذا النوع من الانحياز اختلاف الدرجة الناجمة عن التعليمات وغير ذلك من مشكلات التواصل بين الفاحصين والمفحوصين، وتحدث هذه المشكلات بشكل خاص حسب Gass و Varonis (1991) عندما يستخدم الفاحصون لغة غير لغتهم الأم، فقد يكون فقدان المعلومات اللازمة ناتج عن عدم القدرة على التعبير عن الأفكار بلغة ثانية. كما أن عدم معرفة ثقافة المفحوصين قد يؤدي إلى انتهاك القواعد المحلية للمجاملة. وقد عالج المختصون في علم النفس والباحثون في الثقافة تأثير مميزات الشخص الذي يقوم بتطبيق الاختبار (كالجنس، العمر، العرق) على نتائج القياس. وفي مراجعته للدراسات السيكلوجية الخاصة بتأثير جنس المختبر على أداء الأطفال في اختبارات الذكاء استنتج Jensen (1980) أن الدراسات المناسبة قليلة، فلا توجد دراسات تعالج قضية عرق مفحوص/فاحص مع ذلك فإن البيانات المتوفرة لا تشير إلى أن عرق المختبر ولا جنسه ذو أهمية كبيرة، كما درس الباحثون في الثقافات ما يدعى بنظرية الإذعان (Deference theory)، فقد وجد كل من Cohen, Cotter و Coulter (1982) أن الأشخاص أكثر ميلا لإبداء مواقف إيجابية في ثقافة معينة عندما تجرى مقابلتهم من قبل أفراد تلك الجماعة. ومع ذلك فإن دراسة Senger (1989) قد ركزت على أهمية التأثيرات الراجعة لصفات الباحث الذي يجري المقابلة، أين تبدو النتائج متضاربة في مختلف الدراسات (as cited in Van de vijver, & Poortinga, 2005, p. 45).

كما اعتبر بعض الكتاب أيضا موضوع الانحياز في التطبيق كتهديد لصدق تفسير درجات الاختبارات المكيفة. على سبيل المثال ناقش Geisenger (1994) حقيقة أن الاختلافات الثقافية لمجموعات قومية تتغير حسب مستوى معرفتهم ببنود بنية الاختبار المختلفة، واقترح استخدام عدد كاف من التمارين العلمية لجعل البنية الجديدة مألوفة إلى الطلاب الممتحنين. ويمكن القول أنه قد ينتج عن انحياز المنهج اختلاف الدرجات داخل المجموعات نفسها وليس له علاقة بالبنية وإنما يعود إلى خطأ في القياس (as cited in Cook, et al. 2005, p. 173).

3.1.7.2. انحياز البند

يقصد بهذا النوع من الانحياز بالأخطار التي تؤثر في صحة البنود فقط، بينما يتناول انحياز البنية والمنهج المظاهر العامة للاختبار، وقد استخدم مصطلح انحياز البند لأول مرة من طرف Hilton و Cleary (1968). وبعد أكثر من ثلاثة عقود من التطور المهم في القياسات النفسية لاكتشاف الموضوعات غير السوية تم استبدال هذا المصطلح بـ "اختلاف عمل البند" لأنه كان هناك شعور بأن انحياز البند يتضمن معنى الابتعاد عن المعيار الأوروبي/الأمريكي الذي كان وما يزال أكثر مجموعة مستخدمة في الأبحاث بالولايات المتحدة. إلا أنه تم التمسك بالمصطلح الأصلي لأنه يؤكد على العلاقة الوثيقة مع الأنماط الأخرى من الانحياز ويقصد به تهديد لصحة البند ويمنع المقارنة المباشرة للدراجات (as cited in Van de vijver, & Poortinga, 2005, p. 47).

وقد ناقش Tanzer و Van de Vijver (1997) انحياز البند كونه عامل مهم يسبب الإرباك في دراسات عبر الثقافات، وقد وضعوا لائحة بالأسباب التالية لانحياز البند في تلك الدراسات وهي: "الترجمة السيئة للبند" حيث تكون البنود غامضة؛ عوامل ضارة بالقياس (حيث يسبب البند آثاراً إضافية أو قدرات)؛ المواصفات الثقافية (كالاختلافات الثانوية في معنى المفاهيم وملائمة محتوى البند) (as cited in Cook, et al. 2005, p. 174).

ولعل من أهم أسباب انحياز البند هو الترجمة السيئة واختلاف المعاني الضمنية للكلمات والمثال على ذلك واستناداً إلى قاموس Webster الأمريكي فإن العدوانية Aggressiveness مثلاً تتجلى في كمونها تصميم واضح واستعداد للخصام، بينما قاموس Oxford البريطاني يعطي المعنى الأول للعدوانية بأنها عمل أو ممارسة الهجوم دون تريض وبشكل خاص في البدء بالحرب أو النزاع، ويأتي المعنى السابق في القاموس الأمريكي في المرتبة الثالثة، ومن هنا فإنه من المهم الإشارة إلى أن مفهوم العدوانية في اللغات الأخرى كالألمانية والفرنسية والهولندية وغيرهم من اللغات هي أقرب في معناها إلى التعريف البريطاني وليس الأمريكي.

ولتحليل مصطلح انحياز البند نقوم بتقسيم مجموعتين من ثقافتين مختلفتين بناءً على درجة إجابتهن على الاختبار في عدة مجموعات، فمثلاً بالنسبة إلى الثقافة الأولى نضع في المجموعة (1) جميع الأفراد الذين تحصلوا على الدرجة (10) وفي المجموعة (2) جميع الأفراد الذين تحصلوا على الدرجة (12) وهكذا بالنسبة لكامل الدرجات وبنفس الطريقة مع مجموعة الثقافة الثانية، وتسمى هذه العملية بطريقة التكييف الشرطي Conditioning، ثم نقوم بمقارنة بين الثقافتين بالنسبة إلى الأفراد الذين لهم درجات متماثلة وننظر إلى إجاباتهم بالنسبة إلى البند الذي نتأكد من انحيازه. وبصورة أدق، يمكن القول أن البند منحاز إذا كان الأشخاص الذين لديهم المستوى نفسه في الدرجة الكلية بالنسبة إلى الاختبار ولم يحصلوا على نفس الدرجة المتوقعة منهم في الإجابة على البند (Holland, & Thayer, 2000).

وقد وصف Mellenbergh (1982) نوعين من انحياز البند، هما الانحياز المتماثل Uniform والانحياز غير المتماثل Nonuniform، إذ يصف الفرق بينهما، في أن البند المنحاز بشكل متسق إذا كان الاختلاف في مستوى الأداء ثابتاً في جميع المستويات تقريباً وهو يقصد بذلك أنه في كل مجموعة من الدرجات يكون أداء أفراد الثقافة الأولى أفضل من أداء أفراد الثقافة الثانية بنفس المقدار، أما الانحياز غير المتسق إذا كان حجم الفارق يختلف بشكل مضطرب بين المجموعات (مجموعات الدرجات). ومثال ذلك ما أشار إليه Mazor و Clauser (1998) في أن عدد الأمريكيين الذين يعرفون اسم رئيسهم أقل من عدد الهولنديين الذين يعرفون اسم الملكة في بلادهم، إلا أن هذا الفرق يتضاءل تدريجياً عند المجموعات ذات الدرجات الأولى (as cited in Van de vijver, & Poortinga, 2005, p. 48).

4.1.7.2. إنحياز العينة

إن جميع أشكال الانحياز التي ذكرناها حتى الآن تتناول التفاوت في تقييم درجات الاختبار على البشر وعلى مجال معين من السلوك في الثقافات المختلفة. إلا أن نمطاً واحداً من الانحياز لم نذكره وهو انحياز العينة، حيث يتمثل هذا النمط من الانحياز في عدم قابلية العينات للمقارنة بين ثقافات مختلفة، وهذا يتعلق بتباين العينات في قدرتها على تمثيل ثقافة السكان الذي تؤخذ منهم العينات. ففي كثير من الدراسات تم أخذ العينات من طلاب الجامعات إلا أن الدخول إلى

الجامعات في بعض البلدان يعتمد بالدرجة الأولى على الأداء الدراسي، في حين يكون الوضع الاقتصادي والاجتماعي للأهل في بلدان أخرى هو المقياس للدخول إلى الجامعة، وهنا تظهر اختلافات في نظام العينات فيما يتعلق بمواصفات الخلفيات مما يعلل أي تباين يلاحظ في الدرجات. بالإضافة إلى الصفات الثقافية للمجموعة المعينة.

إن أحد الشواهد الموثقة في الدراسات عبر الثقافات تتعلق بالنتائج المعرفية للأمية، ففي كثير من الدراسات في هذا الخصوص يقارن بين الأميين وغير الأميين بناء على الانتساب إلى المدرسة، لأن الانتساب إلى المدرسة والتعلم يعني معرفة القراءة والكتابة، ولكن في الدراسة التي أجراها كل من Scribner و Cole (1981) في ليبيريا ودراسة Bennett (1991) في كندا، وجدوا أن القراءة والكتابة تكتسب عن طريق التعلم غير الرسمي (خارج المدرسة)، ومن الجدير بالاهتمام أن كلا الدراستين وجدت أن النتائج المعرفية لعدم التعليم المدرسي كانت ضئيلة (as cited in Van de vijver, & Poortinga, 2005, pp. 51-52).

3. الإطار المنهجي

من المؤكد أن الكتابات التقنية لتوجيه عملية التكيف غير كاملة (من وجهة النظر التقديرية) ومتفرقة في كثير من المطبوعات العالمية، التقارير والكتب، ولم يكن هناك أي مصدر كامل يستطيع الممارسون الرجوع إليه لبعض النصائح ولم تتكون مجموعة من الإرشادات لتكييف الاختبار (Hambleton, 1994; Van de vijer & Hambleton, 1996) ولم تكن طرق القياسات المعقدة (نماذج بنود الاستجابة ونماذج توازن التركيبات) التي تساعد في إنشاء عملية معادلة الدرجات التي تم الحصول عليها في الاختبارات المكيفة لاستخدامها في اللغات والثقافات المختلفة معروفة للباحثين الذين يقومون بتكييفها حتى وقت قريب، لكن الوضع قد تحسن منذ أوائل التسعينيات من القرن الماضي.

في الواقع كان الهدفان من مؤتمر اللجنة الدولية للاختبارات International Test Commission (ITC) الذي عقد في جورج تاون في الولايات المتحدة (1999) أولاً: جمع باحثين من العالم لتبادل المعرفة والخبرة في تكييف الاختبار. وثانياً: قدم النسخ النهائية للدليل الموجز لتكييف الاختبارات الذي أشرفت عليه الهيئة العالمية للاختبارات (ITC) (Hambleton, 2004, p.19).

حيث قامت اللجنة الدولية للاختبارات (ITC) في (2017) بإعداد مجموعة من الخطوط الرئيسية لتكييف الاختبارات (Hambleton, 1994; Van de vijer & Hambleton, 1996)، والذي يشار إليه بـ "دليل الهيئة الدولية للاختبارات لتكييف الاختبارات"، حيث ساهمت ثمانية (08) هيئات في إعداد الدليل العالمي لتكييف الاختبارات وهي: اللجنة الدولية للاختبار (ITC)، والجمعية الأوروبية للتقويم النفسي (EAPA)، ومجموعة ناشري الاختبارات الأوروبية (ETPG)، والجمعية الدولية النفسية عبر الثقافات (IACCP)، والجمعية الدولية لعلم النفس التطبيقي (IAAP)، والجمعية الدولية لتقويم الأداء التربوي (IAA)، والجمعية الدولية للاختبارات اللغوية (ILTA)، والاتحاد الدولي للعلوم النفسية (luPSYs).

فدليل تكييف الاختبار منظم في أربعة أقسام: (1) المحتوى، (2) تطوير وتكييف الاختبار، (3) إدارة وتطبيق الاختبار، (4) تفسير وتوثيق الدرجات النهائية. كان رأي اللجنة التي أعدت الدليل أنه سيكون مريحاً أكثر في الاستخدام إذا نُظم في فئات ذات هدف أساسي. أين تناولت الخطوات الرئيسية في منشأ المحتوى تكافؤ المفهوم في لغة المجموعة المستخدمة للاختبار. حيث أقرت الهيئة الدولية للاختبارات التعريف التالي لدليل تكييف الاختبارات "إن دليل تكييف الاختبارات هو مزاولة مهنة تعد مهمة لإدارة وتقويم التكييف أو تطور مواز للاختبارات النفسية والتربوية للاستخدام في مجتمعات مختلفة". فالخطوات الرئيسية المقدمة من الهيئة الدولية للاختبارات يمكن تلخيصها في الخطوات العريضة الموجودة في التقرير الأخير للجنة (ITC, 2017) (as cited in Hambleton, 2005, p. 20).

1.3. تعليمات اللجنة الدولية للاختبارات (ITC) في عملية تكييف الاختبارات عبر الثقافات

إن عملية تطوير وتكييف الاختبارات النفسية تبدو عملية معقدة نظراً للاختلافات الثقافية واللغوية بالنسبة إلى العينات المستهدفة وما يقابلها من عينات أصلية في عملية التكييف والتطوير، الأمر الذي يستوجب من مطوري ومكفي الاختبارات النفسية إتباع التعليمات التي جاءت في دليل الهيئة الدولية للاختبارات (ITC) في (2017) وهي كما يلي:

1.1.3. تطوير الاختبار وتكييفه

- يجب على الذين يقومون بعملية التطوير والناشرين التأكد من أن عملية التكييف تأخذ بعين الاعتبار الاختلافات اللغوية الثقافية للمجموعات المقصودة.
- يجب على الذين يقومون بعملية التطوير والناشرين إقامة الأدلة بأن اللغة المستخدمة في تعليمات الاختبار، وإرشادات الدرجات والبنود، مناسبة للغة وثقافة جميع المجموعات التي ستختبر.
- يجب على المطورين والناشرين إقامة دليل على أن اختيار أسلوب الاختبار، وهيكل البنود، وقواعد الاختبار مألوفة للمجموعة المقصودة.
- يجب على المطورين والناشرين إقامة دليل على أن محتوى البنود والمواد المنبهة في الاختبار مألوفة بالنسبة إلى المجموعة التي ستقوم بالاختبار.
- يجب على المطورين والناشرين جمع دليل النقد المعرفي، لغوي، نفسي، لتحسين دقة عملية التطوير وجمع دليل على تكافؤ كل النسخ في الثقافات المختلفة.
- يجب على المطورين والناشرين التأكد من أن خطة جمع المعطيات تسمح باستخدام أساليب إحصائية مناسبة لإقامة تكافؤ البند والبنية والمنهج في نسخ الاختبار في لغات مختلفة.
- يجب على المطورين والناشرين استخدام أساليب إحصائية مناسبة لكي يستطيعوا:
 - إقامة التكافؤ في لغة نسخ الاختبار.
 - التعرف على العناصر التي يمكن أن تحدث مشكلات أو تكون غير مناسبة لإحدى المجموعات المشاركة.
- يجب على المطورين والناشرين توفير معلومات عن صدق الاختبار المكيف للمجموعة المقصودة.
- يجب على المطورين والناشرين توفير دليل إحصائي عن تكافؤ البنود لكل المجموعات المقصودة.
- يجب عدم ربط بنود الاختبار المكيف غير المتكافئ للمجموعة المقصودة مع الدرجات العامة للاختبار.

على كل حال يمكن أن تكون تلك البنود مفيدة لإعطاء تقرير عن الدرجات في كل مجموعة على حدا.

2.1.3. تطبيق وإدارة الاختبار

- ✓ يجب أن تكون البيئة التي تؤثر في إدارة الاختبار متشابهة إلى أقصى حد مع المجموعات التي أجرت الاختبار.
- ✓ يجب على المطورين والناشرين محاولة توقع المشكلات التي يمكن حدوثها واتخاذ جميع الإجراءات المناسبة لمعالجة هذه المشكلات، وذلك بإعداد مواد وإرشادات مناسبة لمطقي الاختبار.
- ✓ يجب على المطورين والناشرين أن يدركوا العناصر المتعلقة المواد المحفزة، اجراءات التطبيق وطرق الاستجابة التي قد تخفض من صدق الاستنتاجات التي تم الحصول عليها من الدرجات.
- ✓ يجب أن تكون لغة إرشادات وتعليمات المطبقين في اللغة المستهدفة هي نفسها في اللغة الأصلية، مع ضرورة العمل على التقليل من المتغيرات الغير مرغوب فيها لدى المجموعات.
- ✓ يجب أن يعطي دليل الاختبار وصفا دقيقا لكل أوجه الاختبار وطريقة تطبيقه التي تتطلب الدقة في تطبيق الاختبار في محيط ثقافي جديد.
- ✓ يجب ألا يكون المطبقين فضوليين، ويجب أن تكون العلاقة بين المطبقين والذين يجرون الاختبار قليلة إلى أدنى حد، ويجب إتباع القواعد الواضحة التي جرى وصفها في دليل الاختبار.

3.1.3. تفسير النتائج

- عند تكييف الاختبار للاستخدام في مجموعة أخرى، يجب توثيق التغيرات مع الدليل الذي يدعم تكافؤ النسخة المكيفة للاختبار.
- يجب ألا تؤخذ اختلافات الدرجات لنماذج المجموعات التي قامت بالاختبار كقيمة ظاهرية. فعلى الباحث مسؤولية إقامة دليل على معنى الاختلافات من الأدلة التجريبية.
- يمكن إقامة المقارنات لدى المجموعات فقط على مستوى الثوابت التي أقيمت للمقياس الذي تنقله الدرجات.
- يجب على المطورين والناشرين توفير معلومات محددة عن الطرق التي يمكن أن تؤثر فيها المفاهيم الاجتماعية/ الثقافية والبيئية للمجموعات على أداء الاختبار، كما يجب اقتراح إجراءات لتحليل تلك التأثيرات في تفسير النتائج (Hambleton, Merenda, & Spielberger, 2004, pp. 22-24).

2.3. تعليمات جامعة Minnesota في ترجمة وتكييف اختبارات الشخصية

اتبعت أغلبية الدراسات حول ترجمة و تكييف استبيان (MMPI-2) إجراءات منهجية اعتمدت فيها على اقتراحات باحثين ك (Brislin, Lonner, & thorndiker, 1973 ; Butcher, & pancheri, 1976 ; Draguns, 1984 ; Brislin, 1986)، والتي تبنتها مطبعة جامعة Minnesota التي تعتبر الناشر الرسمي لاستبيان (MMPI-2) وصاحبة كل الحقوق.

حيث ميّز هؤلاء الباحثين أربعة مستويات من التكافؤ التي يجب أن تؤخذ بعين الاعتبار عند ترجمة وتكييف استبيان (MMPI-2) عند استعماله عبر الثقافات هي: التكافؤ اللغوي والتكافؤ البنيوي والتكافؤ السيكمومتري وأخيرا التكافؤ السيكلوجي (Butcher, et al 2006, pp. 506-516).

1.2.3. دراسة التكافؤ اللغوي Linguistic Equivalence

المستوى الأول في دراسة تكافؤ نسخة مترجمة مع النسخة الأصلية هو التكافؤ اللغوي، أي نفس معنى البنود والتعليمات (Butcher, 1996). وقد ميّز Sechrest وآخرون (1972) عدة أنواع في التكافؤ اللغوي، حيث نجد تكافؤ المفردات (Vocabulary equivalence) وتكافؤ التعبيرات (Idiomatic equivalence) والتكافؤ النحوي (Grammatical-syntactical equivalence).

وحتى نضمن التكافؤ في المفردات لا بد على المترجمين اعتماد لغة سهلة (لغة كل يوم) وليس بالضرورة لغة القاموس (Dictionary language). أما بالنسبة للتكافؤ في التعبير فهو من المستويات الصعبة، إذ لا بد من إيجاد تعبيرات خاصة بالثقافة المستهدفة والتي تؤدي نفس المعنى في الثقافة الأصلية. أما التكافؤ النحوي فهو تحدي متعلق بإيجاد أشكال لأفعال وأسماء وحروف مناسبة تؤدي نفس المعنى. ولا بد من الابتعاد عن الجمل النادرة الاستعمال والكلمات المجردة (Sechrest, et al., 1972, p. 44). بصفة عامة، يقترح أغلبية الباحثين استعمال الترجمة الحرفية والغير الحرفية في نفس الوقت، من أجل تحقيق معنى البند، وهذه الطريقة سوف تثقل ربما الاختبار لكن هذا أفضل من أن نفقد معنى البند كله (Butcher, et al 2006, p. 517).

أما بالنسبة للتقنية التي يجب استعمالها لضمان التكافؤ اللغوي حسب (Butcher, 1996) نجد أولاً: الترجمة الأولية أو المباشرة لبنود الاستبيان الأصلي إلى اللغة المستهدفة. ثانياً: الترجمة العكسية أو الرجعية (Backtranslation) من طرف مختصين مستقلين من اللغة المستهدفة إلى اللغة الأصلية، كما تم ذلك في الترجمة العالمية لاستبيان (MMPI-2). فالترجمة الأولى من اللغة الأصلية إلى اللغة المستهدفة تتم من طرف مجموعة من المترجمين الذين يكونون ثنائيي اللغة (Bilingual) متمكنين من اللغتين.

ويقترح (Butcher, Lim, & Nezami, 1998) استعمال مترجمون مستقلون، و بعد الترجمة يتم مناقشة البنود بينهم لكن لا بد أن تتوفر لديهم تجربة كافية في كلا اللغتين، كأن يعيش كل مترجم على الأقل 05 سنوات في الثقافة الأصلية للاختبار، و هي الولايات المتحدة الأمريكية فيما يخص (MMPI-2).

أما المرحلة الثانية من منهجية التكييف والتي تحقق التكافؤ اللغوي والتي اقترحت من طرف العديد من الباحثين كدراسة Rosen (1958)؛ Brislin (1970)؛ Werner و Campbell (1970)؛ Sechrest وزملائه (1972)؛ Brislin وزملائه (1973)؛ Butcher و Pancheri (1976)؛ ودراسة Bracken و Barona (1991) وهي الترجمة العكسية للبنود، لكن من طرف مترجم آخر مستقل عن الأولين، و الهدف هو التحقق من أن ترجمة البنود اختلفت في المعنى ومنه يمكن الكشف عن البنود التي كانت صعبة الترجمة، و بالتالي يعاد ترجمتها من جديد.

فهذه الترجمة المباشرة والترجمة العكسية سوف تعاد حتى يقلص الاختلاف في المعنى بين النسخة الأصلية والنسخة المترجمة. وقد أظهرت التجربة العالمية في ترجمة وتكييف استبيان (MMPI-2) حسب (Butcher, 1996) أن ما

نسبته من 10% إلى 15% من البنود يعاد ترجمتها بعد الترجمة الأولى، وذلك حتى يصل المترجمون إلى اتفاق حول البنود التي يكون فيها الإشكال، وعليه تعد هذه المرحلة من أهم مراحل أي مسعى للترجمة (Butcher, et al., 2006, p. 512).

2.2.3. دراسة تكافؤ البنية Construct Equivalence

يعرف هذا النوع من التكافؤ أيضا بتكافؤ المفهوم (Conceptual equivalence)، ما يعني أن المفاهيم المستعملة في النسخة المترجمة يجب أن تُشابه المفاهيم المستعملة في النسخة الأصلية (Butcher, et al., 2006, p. 512). ولتحقيق هذا التعميم للبنية من عينة إلى عينة أخرى ومن ثقافة إلى ثقافة أخرى، لابد من قيام بدراسة ثنائية اللغة عبر نموذج تطبيق وإعادة التطبيق (Bilingual Test-Relest). حيث تتطلب هذه الدراسة حسب (Butcher, Nezami, & Exner, 1998) عينة من أشخاص ثنائيي اللغة، أين يكونون متمكنين من اللغتين (الأصلية والمستهدفة) ويكونون قد عاشوا في الثقافة الأصلية بما فيه الكفاية، فنطبق عليهم النسختين فيجيبون على النسخة الأصلية (الإنجليزية) والنسخة المترجمة باحترام فارق زمني يتراوح بين أسبوع إلى أسبوعين بين التطبيق الأول والثاني. أين لا يمكن وفقا لـ (Hulin, 1987) تطبيق الاستبيان المترجم مرتين على عينتين منفصلتين من أحادي اللغة (Monolingual) من أجل تقييم تكافؤ الترجمة، حتى ولو توفرت نفس الخصائص في العينتين، مع تصميم التطبيق وإعادة التطبيق (Test-retest) على عينة ثنائية اللغة، أين يتم حساب الفرق كما نقوم به في أي تصميم من هذا النوع. لكن، على اعتبار أن التطبيق الأول هو الإجابة على الاستبيان باللغة الأصلية والتطبيق الثاني يكون بالإجابة على الاستبيان المترجم، نقوم بتحليل المعطيات بنفس الطريقة التي نقوم بها في التصميم تطبيق وإعادة تطبيق. وتبعاً لذلك فإن النتائج سوف تساعدنا في القرار ما إذا كان يمكن اعتبار النسختين متكافئتين (Butcher, Derken, Sloore, & Sirigatti, 2003). ومن المهم جداً أن تُقام فيما بعد دراسة أخرى مع أحادي اللغة في إطار تطوير المعايير الخاصة بالثقافة المستهدفة وفي إطار دراسات الصدق (Butcher et al., 2006, p. 513).

3.2.3. دراسة التكافؤ السيكومتري Psychometric Equivalence

يهدف هذا النوع من التكافؤ إلى دراسة ما إذا كان الاستبيان يتميز بنفس الخصائص السيكومترية في الثقافات المختلفة (Butcher, & han, 1996, p. 45). ونتحقق من هذا عبر فحص البنية الداخلية للاستبيان وذلك من خلال التحليل العاملي (Factor Analysis)، وعبر دراسة نمط الإجابة على البنود أو التصديق على البنود (Endorserment Item patterns) بين عينات متشابهة، مثلاً عينة طلبة جامعيين من ثقافتين مختلفتين. حيث أنهما نظرياً عينتين متشابهتين، ويمكن أن يكون لدهما نفس نمط الإجابة عبر الثقافات.

نهتم في التكافؤ السيكومتري بما يسمى بالتكافؤ العددي (Scalar Equivalence)، أي فكرة أن نتائج أداتين متشابهتين يكون لهما نفس مستوى الصعوبة في نفس السلم أو المقياس. مثلاً إذا كانت الدرجة المعيارية T لسلم الانطوائية الاجتماعية (Si) في اللغة الأصلية هي (80) ووجد نفس مستوى الانطوائية الاجتماعية (T=80) في اللغة المستهدفة، إذن يمكن القول أن التكافؤ العددي قد تحقق. حيث نحقق التكافؤ العددي عبر تطبيق الاختبار على عينتين متشابهتين في ثقافتين مختلفتين، وذلك من أجل تحديد ما إذا كانت الدرجات المتحصل عليها في نفس السلالمة بين العينتين متقاربة.

إن الإجراءات المستعملة من أجل فحص مستوى التكافؤ السيكومتري عبر الثقافات، تتضمن دراسات التكافؤ معتمدة على تقنيات التحليل العاملي (Factor Analytic techniques) ونظرية الاستجابة للفقرة (IRT). والمرحلة الإضافية والنهائية لضمان التكافؤ السيكومتري تبقى تطوير المعايير الخاصة بالثقافة المستهدفة (Butcher et al 2006 p 513).

1.3.2.3 التحليل العاملي Factorial Analysis

إن التحليل العاملي هي التقنية التي تفحص التشابه والتطابق بين العوامل الداخلية من أجل تقييم ما إذا كان للاختبار نفس البنية بين عينات من ثقافات مختلفة. إذ ينبغي إجراء مقارنات بين العوامل باستخدام مؤشرات موضوعية بدلا من الفحص الذاتي الغير موضوعي لعوامل التشبع (Factor loadings) (Brislin, et al, 1973 ; Butcher,) (Pancheri, 1976) & إذ أن هناك عدة طرق متاحة من أجل هذا الهدف وهي تضم:

أ. معامل التطابق أو التجانس (C.C) Congruence Coefficient

وهو يمثل مؤشر اتفاق بين عاملين مشبعين قيمته بين 1 + (اتفاق وتطابق تام) و 1 - (تطابق عكسي) مرورا ب 0 (عدم وجود تطابق) ويعتبر مؤشر Toker (1951) من أكثر العوامل استعمالا لأنه يحسب بطريقة بسيطة (as cited in Butcher, & Han, 1996, p. 51).

ب. ارتباطات نتائج العامل (FSC) Factor Score Correlations

يسمى كذلك معامل القابلية للمقارنة Comparability Coefficient وحسب Evereh و Entekin (1980) فإن هذا المؤشر يقيس عامل التشابه أو التطابق (Similarity) ويتمثل في الارتباط بين نوعين من الدرجات العاملية. فهو يميز بين عاملين هما: عامل التشابه (Similarity Factor) وعامل التطابق المشبع (Factor loading similarity)، حيث أن معامل التطابق (C.C) يقيس عامل التشابه المشبع أما ارتباطات نتائج العامل (FSC) فهي تقيس عامل التشابه (as cited in Butcher, & Han, 1996, p. 53).

ج. التحليل العاملي التوكيدي Confirmatory Factor Analysis

تستعمل هذه الطريقة لأجل اختبار الفرضيات المتعلقة بوجود أو عدم وجود علاقة بين المتغيرات والعوامل الكامنة، كما يستخدم التحليل العاملي التوكيدي كذلك في تقييم قدرة نموذج في المقارنة بين نماذج العوامل بهذا المجال. ولاستعمال هذه الطريقة في البحوث عبر الثقافية، يحتاج الباحث أولا معلومات حول عدد العوامل التي تظهر في الثقافة الأصلية ومعلومات حول العلاقة بين هذه العوامل المشبعة، و حجم هذه العوامل نفسها (Butcher, et al, 2006, p. 514).

2.3.2.3 نظرية الإجابة للفقرة Theory Item Response

وهي من النظريات المستعملة في دراسة التكافؤ السيكومتري في الاختبارات المترجمة، وبصفة عامة يستعمل النموذج اللوغاريتمي ثنائي المعلم Two –Parameter Model (Butcher, & Han, 1996). وتفترض هذه النظرية أن السمة المقاسة هي قدرة معينة أو خاصية من خصائص الفرد الذي يختبر بها، بحيث أنه توجد علاقة منتظمة بين مستويات

السمة المقاسة لدى أفراد مختلفين وبين احتمالات الاستجابة الصحيحة لفقرات مختلفة، لذلك فإن هذه النماذج تعد نماذج احتمالية Probabilistic Models، وتستند جميعها إلى دالة الترجيح اللوغاريتمي Logistic Function بدلا من دالة الكثافة الاحتمالية الاعتدالية Normal Density Function التي تستند عليها النظرية التقليدية (علام، 2000، ص. 686).

3.3.2.3. تطوير معايير خاصة بالثقافة المستهدفة Development of a culturally appropriate normative base

عندما يترجم (MMPI-2) إلى لغة أخرى على سبيل المثال تستعمل المعايير الأمريكية، وذلك إذا كانت معطيات العينتين المعياريتين لهما نفس الخطأ المعياري للقياس. لكن إذا كانت الثقافتين مختلفتين في طريقة الإجابة على بنود الاختبار، فلا بد من تطوير معايير جديدة كما كان الحال بالنسبة للثقافة الفرنسية والاسبانية والشيلية والهولندية والمكسيكية (Avila-Espada, & al., 1996 ; Gillet et al., 1996 ; Lucio-G & Reyes-Lagunes, 1996). لذلك، لابد من اختيار عينة سوية وغير مَرَضِيَّة ممثلة للمجتمع الأصلي للثقافة المستهدفة وكذلك عينة أخرى من الأفراد يكون معروف عنهم أنهم مضطربون نفسيا (Butcher et al, 2006, p. 515).

4.2.3. دراسة التكافؤ السيكولوجي Psychological equivalence

بالرغم من الترجمة الجيدة للبنود من اللغة الأصلية إلى اللغة المستهدفة، وبالرغم من احترام البنود للمعنى اللغوي والبنية الداخلية وامتلاكها لنفس الخصائص السيكمترية إلا أنها يمكن ألا تحقق التكافؤ السيكولوجي.

فالتكافؤ السيكولوجي يرجع إلى التشابه أو التطابق في المعنى أو في الأهمية الثقافية لبنود الاختبار. ويسمى كل من (Sechrest et al., 1972)؛ (Butcher, 1996a) هذا النوع من التكافؤ بالتكافؤ التجريبي Experimental equivalence أو الجمل والمصطلحات التي لها نفس التجارب في الثقافة الأصلية والثقافة المستهدفة، مثلا البند "أحب أن أكون بائع زهور" لا يمكن ترجمته في نفس المعنى السيكولوجي لغير الأمريكيين، لأن محلات بيع الأزهار غير موجودة في العديد من بلدان العالم (Sechrest et al., 1972). فالمعنى السيكولوجي لشخص له اهتمام ببيع الأزهار سوف يكون تذوق واهتمام بالجمال والإبداع وهذا يمكن أن يعكس اهتمامات أنثوية في ثقافة أو اهتمام ببيع الزهور في ثقافة أخرى.

هناك مسألة أخرى لها علاقة بالتكافؤ السيكولوجي، هي التكافؤ الوظيفي Functional equivalence والتي تعني أن وظيفة السلوك في ثقافة ما هي مكافئة لوظيفة سلوك في ثقافة أخرى مختلفة حتى ولو أن السلوكات ذاتها يمكن أن تكون مختلفة: مثلا الابتسامة Smiling فهي تدل على الإحراج Embarrassment في العديد من الثقافات الآسيوية بينما في الثقافات الغربية، فالابتسامة تدل على الابتهاج Pleasure (Butcher & Han, 1996) لهذا، فإن للابتسامة في الثقافات الآسيوية وظيفة تميل أكثر إلى تجنب النظر إلى الآخر، بينما في الثقافة الأمريكية فهي تميل إلى الاحمرار خجلا.

لذلك يكون تقييم التكافؤ السيكولوجي عبر الثقافات للاختبارات الموضوعية لقياس الشخصية عن طريق دراسات الصدق Validation research كما تم ذلك في التكييف العالمي لاختبار (MMPI-2) (Butcher, Derksen, et al., 2003)، إذ تستلزم هذه الإجراءات تطبيق الاستبيان المترجم على مجموعات معروفة في الثقافة المستهدفة وتشبه نفس

المجموعات في الثقافة الأصلية، وذلك من أجل فحص وجود تشابهات في أنماط الإجابة وهل هناك ارتباط بين النتائج عبر الثقافات، مثلاً: طلبة جامعيين، مرضى عقليين، نزلاء السجون.

نظرياً، يجب أن تعطي المجموعات المتشابهة نفس نمط الإجابة إزاء الاختبار. فإذا حققت النسختين للاختبار (الأصلية والمترجمة) نفس التنبؤ فيما يخص أفراد هذه المجموعات عبر الثقافات، فهذا يدل على أن الصدق الخارجي قد أثبت غالباً ما تستعمل مجموعات الطلبة في دراسات الصدق، لأنها مجموعات متوفرة بالنسبة للباحث وسهل الوصول إليها، إضافة إلى أن الطلبة متعودون في الإجابة على الاختبارات ويمكن تشكيل عينة معيارية مشابهة لعينة الطلبة الأمريكيين.

4. خاتمة

مع تزايد الاهتمام بالأبحاث عبر الثقافات، هناك حاجة متزايدة للمعايير والممارسات الصحيحة لترجمة الاختبارات والمقاييس النفسية. فتطوير اختبارات مقبولة نفسياً لمجموعة ثقافية أخرى تتطلب دائماً جهداً أكبر مقارنة بالترجمة الحرفية التي غالباً ما تكون ممارسة شائعة، ويمكن أن تكون الترجمات مهددة بمصادر التحيز المختلفة. لذلك، من المهم أن تخضع ترجمات الاختبار لتقييم دقيق لتحديد التعديلات المعيارية التي يجب إجراؤها للثقافة الجديدة، ويجب أن يتضمن البحث دراسة العينات السوية والمرضية لتحديد ما إذا كانت العناصر والمقاييس تعمل على النحو المطلوب. فبمجرد ترجمة الاختبارات، هناك العديد من الوسائل المفيدة للتحقق من صحتها في التكيف بين الثقافات. أولاً، يجب أن يُظهر الاختبار المترجم والمكيّف أنه يقيس في الثقافة المستهدفة نفس التركيبات بنفس الطريقة التي يقيس بها في الثقافة الأصلية. كما يجب أن يكون قد أثبت صحة المحتوى، حيث قد يكون التحقق من صحة المحتوى كافياً لبدء تطوير اختبار في ثقافة جديدة، ولكن هناك حاجة إلى معلومات إضافية لضمان تكافؤ الاختبار في الثقافة الجديدة. قد يكون المنهج الثاني للتحقق من صحة المقياس في ثقافة جديدة هو ضمان صحة العوامل، أي تحديد ما إذا كانت بنية المقاييس متكافئة في الثقافة المستهدفة (أي أن العناصر والمقاييس تحافظ بشكل عام على نفس العوامل في المجموعة المستهدفة) كما هو الحال في المجموعة الأصلية؛ أخيراً، تتضمن استراتيجية التحقق الأكثر أهمية هي ضمان الصلاحية التنبؤية، أي تحديد ما إذا كانت المقاييس تتنبأ بالفعل بالسلوك المتوقع في المجموعة المستهدفة.

المراجع

برمدا، هالة. (2006). *تكييف الاختبارات التربوية والنفسية للتقييم عبر الثقافات*. [رونالد. هامبلتون، بيتر. ميريندا، وتشارلز. سبيليرغر: الكتاب] (ط1). مكتبة العبيكان. (المرجع الأصلي منشور في 2004).

زيدان، محمد. (1979). *معاجم المصطلحات النفسية و التربوية* (ط1). دار الشر والتوزيع.

علام، صلاح الدين. (2000). *القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة* (ط1). دار الفكر العربي.

- اللجنة الدولية للاختبارات (ITC). (2017). مبادئ ITC الإرشادية لترجمة ونقل الاختبارات لثقافة جديدة. [هشام. جادالرب، ميسون. النصر، فجر. العيسى، نور. الجاسر، وسمية. العجوي: المترجمون] (الإصدار الثاني). <https://www.intestcom.org/upload/media-library/guideline-test-adaptation-2-arabic-1629138769gRcyM.pdf>
- Abdallah-Preteille, M. (1999). *L'éducation interculturelle* (Ed.). Paris: PUF.
- Avila-Espada, A., Jimenez-Gomez, F., & Butcher, J. N. (1996). The Castilian version of the MMPI-2 in Spain: Development, adaptation, and psychometric properties. In J. N. Butcher, (1stEd.), *International Adaptations of the MMPI-2: research and clinical applications* (pp. 305–328). USA: University of Minnesota Press.
- Ben-Porath, Y. S. (1990). Cross-cultural assessment of personality: The case for replicatory factor analysis. In J. N. Butcher & C. D. Spielberger (Vol 8.), *Advances in personality assessment* (pp. 27–48). NY: Lawrence Erlbaum Associates, Inc.
- Berry, J. W. (1980). Acculturation as Varieties of Adaptation. In A.M. Padilla, (Ed.), *Acculturation: Theory, Models and Some New Findings* (pp. 9-25). Boulder, CO: Westview.
- Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Ed.), *Field methods in cross-cultural research* (pp. 137–164). NY: Sage Publications, Inc.
- Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). *Cross-cultural research methods* (Ed.). NY: John Wiley & Sons.
- Butcher, J. N. & Han, K. (1996). Methods of establishing cross-cultural equivalence. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 44-66). Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N. (1996a). *International adaptations of the MMPI-2: Research and clinical applications* (Ed.). Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N. (1996b). Translation and adaptation of the MMPI-2 for international use. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 26-43). Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N. (1996c). Understanding abnormal behavior across cultures: The use of objective personality assessment methods. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 3-25). Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N. and Pancheri, P. (1976). *Handbook of cross-national MMPI research* (Ed.). Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N., Graham, J. R., Kamphuis, J. H., & Rouse, S. V. (2006). Evaluating MMPI-2 Research: Considerations for Practitioners. In J. N. Butcher (Ed.), *MMPI-2: A practitioner's guide* (pp. 15–38). USA: American Psychological Association. <https://doi.org/10.1037/11287-002>

- Butcher, J. N., Lim, J., & Nezami, E. (1998). Objective study of abnormal personality in cross-cultural settings: The Minnesota Multiphasic Personality Inventory (MMPI-2). *Journal of Cross-Cultural Psychology*, 29(1), 189-211. <https://doi.org/10.1177/0022022198291010>
- Butcher, J. N., Nezami, E., & Exner, J. (1998). Psychological assessment of people in diverse cultures. In S. S. Kazarian & D. R. Evans (Eds.), *Cultural clinical psychology: Theory, research, and practice* (pp. 61–105). United Kingdom: Oxford University Press.
- Butcher, J., Derksen, J., Sloore, H., & Sirigatti, S. (2003). Objective personality assessment of people in diverse cultures: European adaptations of the MMPI-2. *Behaviour Research and Therapy*, 41(7), 819-840. [https://doi.org/10.1016/S0005-7967\(02\)00186-9](https://doi.org/10.1016/S0005-7967(02)00186-9)
- Cook, A., Spinazzola, J., Ford, J., Lanktree, C., Blaustein, M., Cloitre, M., & Van der Kolk, B. (2005). Complex trauma. *Psychiatric annals*, 35(5), 390-398. <https://psychrights.org/research/Digest/CriticalThinkRxCites/cook.pdf>
- De-Carlo, M. (1988). *L'interculturel* (Ed.). Paris: Clé internationale.
- Draguns, J. G. (1984). Assessing mental health disorders across cultures. In P. B. Pedersen, N. Sartorius, & A. J. Marsella (Ed.), *Mental health services: The cross-cultural context* (pp. 31 -58). Beverly Hills, CA: Sage.
- Ferréol, G., Jucquois, G. (2003.). *Dictionnaire de l'altérité et des relations interculturelles* (Ed.). Paris, Armand Colin, coll.
- Frijda, N., & Jahoda, G. (1966). On the scope and methods of cross-cultural research. *International Journal of Psychology*, 1(2), 109–127. <https://doi.org/10.1080/00207596608247118>
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological assessment*, 6(4), 304-312. <https://doi.org/10.1037/1040-3590.6.4.304>
- Gillet, I., Simon, M., Guelfi, J. D., Brun-Eberentz, A., Monier, C., Seunevel, F., & V-Svarna, L. (1996). The MMPI-2 in France. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 395-415). Minneapolis, MN: University of Minnesota Press.
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10(3), 229–244. <https://psycnet.apa.org/record/1995-39303-001>
- Hambleton, R. K. (2005). Issues, Designs, and Technical Guidelines for Adapting Tests into Multiple Languages and Cultures. In K. Hambleton, P. Merenda, & C. Spielberger (1stEd.), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment* (pp.3-38).Mahwah, NJ: Lawrence Erlbaum. <https://doi.org/10.4324/9781410611758>
- Hambleton, R. K., & Bollwark, J. (1991). Adapting Tests for Use in Different Cultures: Technical Issues and Methods. *The Education Resources Information Center (ERIC)*. 1-44. <https://files.eric.ed.gov/fulltext/ED337481.pdf>
- Hambleton, R. K., & Patsula, L. (1998). Adapting tests for use in multiple languages and cultures. *Social indicators research*, 45, 153-171. <https://doi.org/10.1023/A:1006941729637>

- Hambleton, R.K., Merenda, P.F., & Spielberger, C.D. (2004). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment* (1stEd.). Psychology Press. <https://doi.org/10.4324/9781410611758>
- Hofstede, G. (1991). *Cultures and Organizations: Software of the Mind* (Ed.). London: McGraw-Hill.
- Holland, P. W., & Thayer, D. T. (2000). Univariate and bivariate loglinear models for discrete test score distributions. *Journal of Educational and Behavioral Statistics*, 25(2), 133–183. <https://doi.org/10.2307/1165330>
- Lucio, E. & Reyes-Lagunes, I. (1996). The Mexican version of the MMPI-2 in Mexico and Nicaragua: Translation, adaptation, and demonstrated equivalency. In J. N. Butcher (Ed), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 265-283). Minneapolis, MN: University of Minnesota Press
- Nunnally, J.C. (1978) : *Psychological Theory* (2ndEd). New York : McGraw-Hill.
- Poortinga, Y. H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology*, 24(6), 737–756. <https://doi.org/10.1080/00207598908247842>
- Poortinga, Y. H., & Malpass, R. S. (1986). Making inferences from cross-cultural data. In W. J. Lonner & J. W. Berry (Ed.). *Field methods in cross-cultural research* (pp. 17–46). Sage Publications, Inc.
- Sechrest, L., Fay, T. L., & Hafeez Zaidi, S. M. (1972). Problems of translation in cross-cultural research. *Journal of Cross-Cultural Psychology*, 3(1), 41–56. <https://doi.org/10.1177/002202217200300103>
- Shweder, R. A. (1999). Why cultural psychology?. *Ethos*, 27(1), 62-73. <https://doi.org/10.1525/eth.1999.27.1.62>
- Singer, E., & Presser, S. (Eds.). (1989). *Survey research methods: A reader* (Ed.). University of Chicago Press.
- Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of personality and social psychology*, 41(1), 37-55. <https://doi.org/10.1037/0022-3514.41.1.37>
- Van de Vijver, F. J. R., & Poortinga, Y. H. (2005). Conceptual and methodological issues in adapting tests. In R.K. Hambleton, P. Merenda, & C. Spielberger (Ed.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 39- 63). Hillsdale, NJ: Lawrence Erlbaum
- Van de Vijver, F. J., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European review of applied psychology*, 47(4), 263-279. https://pure.uvt.nl/ws/files/225989/26727_11858.pdf
- Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1(2), 89-99. <http://dx.doi.org/10.1027/1016-9040.1.2.89>