

The Operation Procedure of Equivalence of Measuring Instruments in Cross-Cultural Researches: What Have We Learnt from the Positive Psychology?*

Wenjie Duan^{1,2}, Yu Bai³, Samuel M. Y. Ho⁴, Xiaoqing Tang^{1,2}

¹School of Culture and Social Development, Southwest University, Chongqing

²Research Society of Cultural Development in Chongqing Municipal, Chongqing

³Institute of Psychology, Chinese Academy of Science, Beijing

⁴Department of Applied Social Studies, City University of Hong Kong, Hong Kong

Email: wjtuan@gmail.com

Received: Feb. 28th, 2012; revised: Mar. 6th, 2012; accepted: Mar. 15th, 2012

Abstract: With the logical starting point of Combined Etic-Emic Approach, and the guidance of rules related to equivalence in cross-cultural researches, this paper effectively integrates the relevant rules about measuring tool translation and modification developed by World Health Organization (WHO), expands the basic steps of traditional method of Back-translation, proposes a “Trilateral Translation Procedure”, and attains the reference materials from the perspectives of quantitative and qualitative, finally produces a measurement tool equivalence operation procedure of cross-cultural research including 7 steps.

Keywords: Cross-Cultural Research; Measurement Equivalence; Etic Approach; Emic Approach; Positive Psychology; Character Strength

跨文化研究中测量工具等价性操作流程初探： 我们从积极心理学中学到了什么？*

段文杰^{1,2}, 白羽³, 何敏贤⁴, 唐小晴^{1,2}

¹西南大学文化与社会发展学院, 重庆

²重庆市文化发展研究会, 重庆

³中国科学院心理研究所, 北京

⁴香港城市大学应用社会学系, 香港

Email: wjtuan@gmail.com

收稿日期: 2012年2月28日; 修回日期: 2012年3月6日; 录用日期: 2012年3月15日

摘要: 本文以兼顾文化共通性与文化特殊性研究范式为逻辑起点, 以跨文化研究中有关等价性原则为指导, 有效整合世界卫生组织关于测量工具翻译和修订基本指导方针的相关内容, 扩展传统回译法的基本步骤, 提出“三方翻译程序”, 同时从定性和定量两个角度获得参考资料, 最终整合出含有7个步骤的跨文化研究测量工具等价性操作流程。

关键词: 跨文化研究; 测量等价性; 文化共通性; 文化特殊性; 积极心理学; 性格优势

1. 引言

*资助信息: 本文为国家社科基金项目(11BSH037)跨文化测量理论研究阶段性成果。

心理学在中国经过30多年的高速发展取得了不少成就, 我国研究者在国际顶级刊物上发表了很高水平论文, 其应用研究也为我国的社会经济建设做出

了贡献，但是也应该看到与发达国家相比，我国心理学研究在整体水平和规模上还有很大差距(杨玉芳, 孙健敏, 2011)。从历史角度看，西方心理学较早成为科学的学科，并在一定程度上被看作是世界心理学的主流，因此在一定时期内，我国研究者还应该借鉴西方心理学领域的优秀理论和研究成果。但是，随着跨文化心理学研究的深入，研究者逐渐认识到根植于西方社会文化的心理学理论、范式和测量工具不能很好的适用于其他种族和文化，东西方最根本的差异来自文化与思维两方面。举例来说，很多研究表明东方文化强调集体主义(Collectivism)，而西方文化强调个人主义(Individualism)(Hofstede, 1980; Oyserman, Coon, & Kimmelmeier, 2002)；东方个体的思维遵循简单辩证主义(Dialectical Socialist)，而西方个体更强调逻辑推理思维(Peng & Nisbett, 1999)。积极心理学是世界心理学界相对年轻的领域，其强劲的发展势头引起了国内外研究者的普遍关注，在非西方学者积极引入、研究并发展相关理论的同时，共同面对一个最主要的困境：如何于非西方文化中正确地测量积极心理建构(Positive Psychological Constructs)？其中最基础、最关键，研究最多的问题是测量工具的测量等价性(Berry, 1980; Leung, Leung, & Cheung, 2010)。

目前以文化视角发展心理理论或开发心理测量工具的方法主要有三种：文化共通性方式(Etic Approach, 也称为“客位研究”)、文化特殊性方式(Emic Approach, 也称为“主位研究”)和兼顾文化共通性与文化特殊性方式(Combined Etic-Emic Approach)。Headland(1990)发现自文化共通性和文化特殊性概念提出以来，在各个学科的质化研究进程中都起到了重要的推动作用。文化共通性方式强调文化的普适性，认为根植于西方文化的心理理论及心理测量工具具有一定的普遍性(Universal)和可概括性(Generalizable)，因此持这种观点的研究者通常对在西方文化下开发的心理测量工具进行简单翻译后，就直接运用于非西方文化下目标群体的测量，而忽视了在测验具体项目、维度、及结果解释等方面可能会存在的文化差异(Cheung, 2004)；而文化差异性方式则强调文化的独特性，认为基于西方文化的心理学理论及心理测量工具只适用于该文化的成员，如果研究者想在其他文化下发展同一心理概念的理论或测量工具，则必须抛弃西方文化下原有的理论和假设，立足本土通过调查获得

的数据来发掘相关概念或范式(Lett, 1990)，世界范围内不少研究者在早期采用此方法开发了具有本土文化特征的人格测量工具(Cheung, Cheung, Leung, Ward, & Leong, 2003)，同时这种方法也是扎根理论¹(Grounded Theory)的核心内容。

兼顾文化共通性与文化特殊性的方式融合了上述两种研究范式，目的在于获得一个更丰富、更综合、更平衡的兼顾普遍性和特殊性理论或构建，认为通过在现有的西方心理学理论或模型中增加特定文化因素可以发展出具有真正普遍性的心理学理论(Cheung, van de Vijver, & Leong, 2011)。Ho 和 Cheung(2007)以此种研究方式，在西方“主观幸福感”理论和测量工具中增加了具有东方文化特色的“人际间主观幸福感”维度，结果表明该方式较大的提高了“主观幸福感”的生态效度，即同时具有广泛的普遍性和本土的敏感性。张妙清等著名华人心理学家(2011)较早在国内人格研究领域采用此方法以文化包容的态度研究人格心理，编制了“跨文化(中国人)个性评估量表”(The Cross-Cultural (Chinese) Personality Assessment Inventory, CPAI)，十多年跨文化研究结果表明该工具具有较高的信效度。可见兼顾文化共通性与文化特殊性的方式能够弥合主流心理学与本土心理学之间的沟壑，同时弥补只关注普遍性或只关注特殊性研究方法的缺陷(Morris, Leung, Ames, & Lickel, 1999)。Cheung 等(2011)认为兼顾文化共通性与文化特殊性的应用方式主要包括：1) 同时使用文化共通性与文化特殊性的测量工具；2) 反复地收集数据来不断修订具有文化共通性或文化特殊性的测量工具；3) 使用综合方法，如以具有文化普遍性的测量工具为主，通过对具有文化特殊性群体的访谈收集相关信息补充到原有工具中。无论哪种具体方式，其最终目的都是提高相关理论和工具的生态效度。

尽管“兼顾文化共通性与文化特殊性方式”已经被国外许多研究者采用，但国内不少应用工作者在跨文化研究中就该如何保证测量工具等价性这一问题

¹ 扎根理论(Grounded Theory)研究法是由芝加哥大学的 Barney Glaser 和哥伦比亚大学的 Anselm Strauss 两位学者共同发展出来的一种研究方法。研究者在研究开始之前一般没有理论假设，直接从实际观察入手，从原始资料中归纳出经验概括，然后上升到理论。这是一种从下往上建立实质理论的方法，即在系统收集资料的基础上寻找反映社会现象的核心概念，然后通过这些概念之间的联系建构相关的社会理论。

并不十分清楚，大部分研究仍采取传统的翻译 - 回译 (Translation and Back-translation) 和验证性因素分析两个步骤来达到等价性的目的，但这实际上并不完整而且缺乏相应指导原则确保等价性的实现。因此本文在兼顾文化共通性与文化特殊性方式的指导下，综述跨文化研究中的主要等价性原则，结合世界卫生组织 (WHO) 提出的测量工具翻译和适应基本步骤，以及相关的统计策略，给出一个实用的跨文化测量工具等价性操作流程，以有利于帮助研究者正确掌握保证测量等价性的基本方法。

2. 跨文化测量中的等价性

尽管在跨文化研究领域的热度持续增加，但进行跨文化研究中测量工具的修订或适应性标准程序远不及在单文化 (Mono-cultural) 研究中完善。目前很多研究都把相同的工具运用于不同的文化群体，同时常常假设这些不同文化下群体之间的结果具有可比性。然而，每种文化环境都是由多种因素、过程和特质交汇而成的混合体，测量工具中的条目对不同文化的个体来说可能具有完全不同的含义，这给跨文化研究者带来了极大的困扰。在跨文化测量中，最简单的方法就是直接翻译在西方文化下已经成型的测量工具，并施测于非西方文化群体。然而关注跨文化研究工具可比性的研究者就此提出了一系列质疑 (Kankaraš & Moors, 2010)，其中最关键的问题包括：翻译质量 (Equality of Translations)、文化相关性 (Cultural Relevance)、构念的测量等价性 (Measurement Equivalent of the Constructs) 和工具的结构效度 (Validity of the adapted instrument) (Hui & Triandis, 1985; Lonner & Berry, 1986)。

跨文化测量中的等价性是指在不同的文化群体中，对同一心理概念、测量工具、观测变量等的解释具有一致性或较大的重合性。Johnson (1998) 曾通过文献回顾筛选了 50 多种测量等价性的类型，而通常被跨文化研究者采用最多的有四种：语言等价性 (Linguistic Equivalence)、功能等价性 (Functional Equivalence)、概念等价性 (Conceptual Equivalence) 和度量标准等价性 (Metric Equivalence)。

2.1. 语言等价性

语言等价性是指心理测量工具中所使用的词语在

不同的文化中具有相同的含义，即保持测量条目的语言精准性。这是在本文化中翻译其他文化下开发的心理测量工具首先会碰到的问题，正如上文所述，简单的翻译是不可取的。经过多年的发展，Brislin (1970) 所确立的回译法 (Back-translation) 被看作是跨文化研究领域测量工具翻译的标准化程序。其基本做法是要求一位双语者 (如同时精通中英文) 将一个英文原版量表翻译成中文，再由另一位双语者把翻译好的量表回译为英文；接着将译回的英文量表与原版量表进行比对，发现其中不准确及被曲解的部分并进行修正。很明显，该过程的最主要目的是让原版工具在目标国家或文化中在语言上具有精准的等价性，那么经过该标准程序翻译好的测量工具是否就具有跨文化等价性呢？

答案是不一定的，以笔者对性格优势的测量工具优势行动价值问卷 (the Values in Action Inventory of Strengths, VIA-IS) (段文杰等, 2011; 段文杰, 李婷婷, 张永红, 2011) 的有关研究为例进行说明。性格优势被定义为“在个体思想、情绪和行为中表现出来的一组积极特质” (Peterson & Seligman, 2004), Peterson (2004) 等确定了全世界所普遍具有的 24 种性格优势 (具有文化共通性)，同时，这 24 种性格优势在理论上被认为可以聚合为 6 大美德 (Steger, Hicks, Kashdan, Krueger, & Bouchard, 2007)。Peterson and Seligman (2004) 根据其分类法开发了测量性格优势的工具——优势行动价值问卷 (the Values in Action Inventory of Strengths, VIA-IS)，它包括对应于 24 种性格优势的 24 个分量表，每个分量表 10 个题项，共 240 题。目前为止，关于 VIA-IS 的研究主要围绕两个争论焦点，第一是其因素结构的跨文化不变性，第二是在不同文化环境内运用时每个题目测量其对应性格优势的适合性。事实上，我们的研究发现这两个争论焦点的核心就是“跨文化测量等价性”，通过一些合理的操作可以使 VIA-IS 在不同文化中具有一定等价性或可比性。具体来说，在问卷的翻译过程中，一些题目含有的单词“Pride”被翻译为“骄傲”，这在“回译法”整个操作流程中并没有出现任何异议，但当我们将翻译好的问卷进行初测与认知访谈时，却发现很多学生将“骄傲”理解为负面的含义 (如“傲慢”)，造成结果偏差，于是我们将“骄傲”一词替换为“自豪”，既符合了原文的含义，又避免了中国被试的误解。又如题目中的

“Spirituality”翻译成“精神性”并不存在任何问题，同样在后来的测试与访谈中，我们发现学生根本无法理解什么是“精神性”，这种文化上的差异是无法通过语言等价性手段解决的，后来，我们将“Spirituality”意译为“信念或信仰”，其认同度得到了大幅提升。因此，我们还需同时考虑以下几个方面。

2.2. 功能等价性

功能等价性是指在不同的文化环境或群体中的，有两种及两种以上相同或相似的行为对待同一问题或状态具有功能上的相关性。也就是说，当某种构念在不同的文化中与其对应的行为完全不同时，功能等价性问题必须认真考虑。仔细分析 VIA-IS 中的条目，会发现原文所描绘的行为在我们中国人看来非常奇怪，尤其是内地居民。例如，原题“I never tell outsiders things bad about my team”用来测量“公民行为”(Citizenship)这一优势，而身处集体主义文化下的我们却会将其看作很普通的社会期望，因此，该题就丧失了其在西方文化下应有的敏感度。同样，“I practice my religion”也不适合测量中国人的“精神性”优势，因为多数内地居民都是无神论者。不仅如此，一些条目所描绘的行为在我们看来甚至是不符合社会规范的，如“When I hear people say something mean, I protest”，因为中国社会相比独特性而言更强调社会的一致性与和谐性(Kim & Markus, 1999)。那么究竟该如何处理类似的问题？是修改还是直接删去还得根据不同情况具体分析。

另一种情况是通过检验具有功能等价性的测量工具与测量工具的关系来考量其功能等价性，例如考察一个新的抑郁测量工具与贝克抑郁量表之间在不同文化下的关系，当这种关系并不一致时，就需要对新的抑郁测量工具进行修订(Leong et al., 2010)。目前我们正在以这种方式进行研究以期进一步提高 VIA-IS 中文版的功能等价性。

2.3. 概念等价性

概念等价性与上述两种等价性相比，更为抽象和复杂，它是指一组行为所代表的概念在不同的文化中具有相同或相似的内涵。与功能等价性相比，概念等价性关注的层次好比是结构方程模型中的潜在变量，而功能等价性关注的层次则是观测变量。为了评估概

念等价性，Usunier(1998)推荐研究者在翻译过程中依靠目标语言使用去中心化策略(Decentering)：首先研究者假定一个适用于所有不同文化的概念区域(Conceptual Area)，熟悉目标文化的母语人士在这个区域内填充符合该概念区域的概念；然后，由一个跨文化研究团队共同筛选一些在目标文化内最为经常使用的概念，并对其进行回译；最后，该团队就可确定具有文化共通性和特殊性的概念维度，在不同文化中出现频率较高的概念往往意味着较低的文化特殊性。其中很重要的一点是，虽然一些概念会出现在不同的文化环境中，但其构成的概念维度却往往是不同的(Usunier, 1998)。

Peterson(2004)在编制 VIA-IS 时，通过整理、阅读不同文化背景下的经典的哲学著述来寻找具有文化共通性的概念，即后来的性格优势。以此为依据编写的问卷具有较高的表面效度，但是我们的研究发现其结构效度并不理想(段文杰，白羽等，2011)，很显然在编制过程中他们没有考虑到“在不同文化中，构成同一概念的概念维度可以千差万别”。后来通过打包策略(Item Parceling)验证了 24 个性格优势的一阶相关模型在中国大学生群体中信效度良好(段文杰，白羽等，2011)，但由于采取了题目打包策略，即在牺牲功能等价性的基础上提高了概念等价性，以后的研究应该采用更为合理的策略同时提高功能等价性和概念等价性。因此，为了保证同一个概念在不同的文化中具有较高的文化共同性，研究者需要极其仔细、认真地选择不同文化所侧重的概念维度。

2.4. 度量标准等价性

度量标准等价性比较好理解，是指心理测量工具在不同文化或群体间测量所得的几组分数具有可比性。一旦对某概念进行测量的工具在不同的文化环境内存在差异，那从这个工具所得到的研究结论即是不可靠的，这就好比用刻度不同的尺子量取物件而后进行比对一样。Kwok Leung and Zhou(2010)总结了在跨文化研究中检验度量标准等价性的三种常用方法，包括方差分析(Van de Vijver & Leung, 2000)、项目反应理论(Item Response Theory, IRT)(Maurer, Raju, & Collins, 1998)和验证性因素分析(Confirmatory Factor Analysis, CFA)。国内研究者最为常用的还是使用验证性因素分析进行度量标准等价性分析，Vandenberg and Lance

(2000)和 Byrne et al.(2009)等人在回顾了有关研究后提出包含 8 个基本步骤的操作流程，具体包括：1) 检验不同样本中观测变量的方差协方差矩阵是否相等；2) 检验不同样本中因素结构是否相同；3) 通过检验不同样本的因素载荷确定度量标准等价性；4) 通过检验题项截距的差异性来评价截距等价性(Scalar Equivalence)；5) 检验每个变量的唯一性；6) 比较不同样本的因子方差；7) 比较不同样本的因子协方差以确定跨样本的对等性；8) 检验不同样本因子均值是否相等。

验证性因素分析归根到底是结构方程模型(Structural Equation Modeling, SEM)的一种特殊形式，结构方程模型是建立在多重回归和因素分析之上的比较灵活的建模技术，研究者可以通过此来探析一组变量之间的内在关系。近年来，以结构方程模型为统计技术的心理学研究增长迅速(MacCallum & Austin, 2000)。在进行跨文化研究时，结果可在两个层次上进行比较，一是群体文化层次，二是个体文化层次。文化层次的建构往往是个体反应的聚合，能够在文化层级上进行比较；而来自不同文化的个体也可以以心理理论为基础进行比较。从已有研究来看，这两种分别在文化和个体上的分析方法经常被独立使用(K. Leung, 1989)，然而，目前已经可能用多层次分析法(Multilevel Analysis)同时检验两个层次(van de Vijver & Leung, 1997)。这种方法目前已被越来越多的跨文化研究者采用(Fischer, Ferreira, Assmar, Redford, & Harb, 2005)，国内研究者最近对有关方法进行了详细介绍(方杰，邱皓政，张敏强，2011)。

3. WHO 对于测量工具翻译和修订的基本指导方针

世界卫生组织(2011)为了规范跨文化研究中测量工具的使用及提高相应的测量等价性，根据若干 WHO 的研究重新优化传统的回译法(Back-translation)，于 2011 年发布了含有四个步骤的指导方针，用于将英文测量工具修订为符合目标文化的其他语言版本，具体包括：

1) 正向翻译(Forward Translation)，即由一位熟知英语文化的专业人士(其母语是目标文化中的主语言)先将英文测量工具翻译为目标语言。译者必须将翻译的重点放在概念等价性上，而不是字面上的直译，同时考虑用语的自然性与可接受性；句子要尽可能的简单、清晰、

简洁，避免长句、复句和从句；以最普通的受众为翻译对象，避免使用术语、俗语、习语和方言。此外还要仔细衡量对任何群体可能有冒犯性的词语。

2) 专家组回译(Expert panel Back-translation)。一般来讲，专家组要包括该领域的专家，具有测量工具开发、翻译和修订经验的研究者。他们的主要职责是发现翻译过程中出现的不合适表达、概念和翻译差异，同时提供可能的解决方案。在专家组完成校正后，由另一位英语母语的非专业领域内独立翻译者将此版本重新翻译为英文，尔后两个英文版测量工具之间出现的任何差异都需要再次提交给专家组进行讨论。这个过程可以反复出现直到获得一个比较满意的版本。

3) 前测和认知访谈(Pre-testing and Cognitive Interviewing)。在用于目标文化或群体之前，非常有必要在目标群体中选择具有代表性的样本进行前测和访谈。WHO 建议对于每个部分至少需要 10 名在性别、年龄、经济社会地位上相匹配的对象进行测试和深度访谈，同时制定一个比较系统的访谈提纲，问题要包括对字、词、句各个层次的理解。对任何存在异议或疑问的地方要进行详细的访问，并要求参与者提供一些他们认为更好理解、更为常用的词语。

4) 最终版本(Final version)。根据上面所有步骤反馈的结果进行最后的修订，完成最终版本，并对该过程中所有的使用的和获得的资料信息进行存档。

4. 建议的操作流程

根据前面的相关讨论和我们的实践经验，提出一个在跨文化研究中提高测量工具等价性的基本操作流程(图 1)，并就其特点进行补充说明。

第一，将 WHO 重新优化的传统回译法(Back-translation)扩展为“三方翻译程序”。以第三方身份出现的专家小组具有较高的独立性和权威性，能够进一步提高专家小组在翻译过程中的指导作用，对在翻译过程中出现的任何差异问题进行论证并提出解决方案。

第二，在回译过程中，明确了由非专业领域内双语者进行。绝大多数心理测量工具的施测群体都是普通大众，他们对心理学专业术语不甚了解，如他们对“人格”的理解与心理学专业工作者的理解可谓是大相径庭，因此由非专业领域内人士进行回译可以缩小专业与非专业之间的差异。

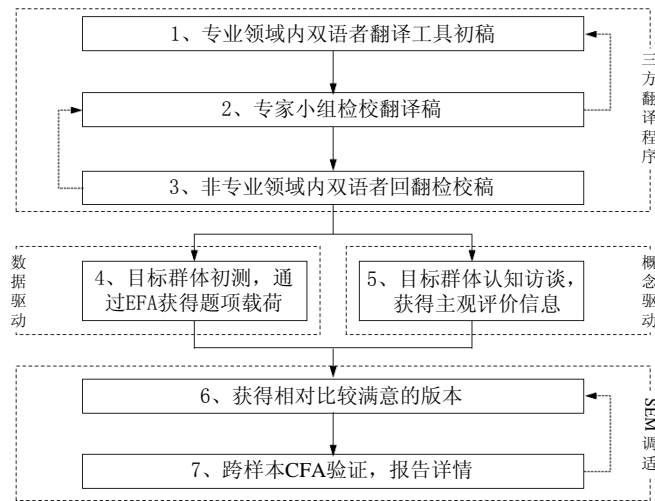


Figure 1. Operation flow chart of improving measurements equivalent
图 1. 提高测量工具等价性的基本操作流程图

第三，将前测与认知访谈相分离，前者重点放在以数据驱动获得客观数据资料，后者重点放在以概念驱动获得主观认知资料。我们在 VIA-IS 修订过程中发现，大多数情况下从这两个途径获得的资料具有一致性，即载荷较低的题项往往被访谈者的认同程度也比较低，此时可以采取删去题项的策略；反之亦然。但有一特殊情况值得注意，理论上单维的结构却探索出了两个或两个以上的维度。对这种情况，我们建议首先对取样样本进行检验，看是否存在取样偏差；排除取样偏差后，以验证性因素分析进行单维再验证，不能单靠探索性因素分析检验量表的单维性(Gerbing & Anderson, 1988)；拟合效果如果确实不好，说明这一心理构念的“概念等价性”出现了问题，此时结合认知访谈获取的资料和载荷确定需要保留的题项。

第四，除了目标群体明确的工具外，我们建议对一般性心理测量工具进行跨样本验证性因素分析，以获得测量工具比较准确的生态效度。同时比照验证性因素分析的结果对测量工具进行微调，从而获得更好的版本。此外，在不同的步骤中也较好的突出了不同的等价性基本原则，即在“三方翻译程序”中重点体现“语言等价性”，在“初测与访谈”中重点突出“功能和概念等价性”，而在“SEM 调试”中主要确保度量标准等价性。

5. 小结

近年来，在国内心理学杂志上发表的测量工具类

文章大多没有从兼顾文化共通性与文化特殊性的视角出发，在具体的修订过程中也没有同时考虑到语言、功能、概念及度量标准等方面的等价性。因此，本文在作者以往研究经验的基础上，以兼顾文化共通性与文化特殊性的研究范式为逻辑起点，整合世界卫生组织关于测量工具翻译和修订基本指导方针的有关内容，将传统双方参与的“回译法”扩展为“三方翻译程序”，重点强调专家小组在翻译过程中的指导作用，在确保“语言等价性”的同时兼顾“概念等价性”；同时从定性(认知访谈)和定量(群体初测)两个角度考量测量工具的“概念等价性”和“功能等价性”，为工具的修订提供参考框架；最后通过跨样本的 CFA 获得心理测量工具在本文文化中的生态效度。本文所提出的含有 7 个步骤的操作流程为整合主流心理学和本土心理学架设了一座桥梁，不仅能够极大的方便心理学应用工作者的推进心理科学实践，还能够在更大范围上促进跨文化研究领域测量的实践与发展。

参考文献 (References)

- 段文杰, 白羽, 张永红, 唐小晴, 王志章, 李婷婷(2011). 优势行动价值问卷(VIA-IS)在中国大学生中的适用性研究. *中国临床心理学杂志*, 4期, 473-475, 478.
- 段文杰, 李婷婷, 张永红(2011). 优势行动价值问卷及其应用研究进展. *中国临床心理学杂志*, 2期, 205-208.
- 范为桥, 张妙清, 张建新, 张树辉(2011). 兼顾文化共通性与特殊性的人格研究: CPAI 及其跨文化应用. *心理学报*, 12期, 1418-1429.
- 方杰, 邱皓政, 张敏强(2011). 基于多层结构方程模型的情境效应分析——兼与多层线性模型比较. *心理科学进展*, 2期, 284-292.

- 杨玉芳, 孙健敏(2011). 心理学的学科体系和方法论及其发展趋势. *中国科学院院刊*, 6期, 611-619.
- 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 Press.
- Brislin, R. (1970). Back translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1, 185-216.
- Byrne, B. M., Oakland, T., Leong, F. T. L., Van de Vijver, F. J. R., Hambleton, R. K., Cheung, F. M., et al. (2009). A critical analysis of cross-cultural research and testing practices: Implications for improved education and training in psychology. *Training and Education in Professional Psychology*, 3, 94-105.
- Cheung, F. M. (2004). Use of Western and Indigenously Developed Personality Tests in Asia. *Applied Psychology: An International Review*, 53, 173-191.
- Cheung, F. M., Cheung, S. F., Leung, K., Ward, C., & Leong, F. (2003). The English version of the Chinese personality assessment inventory. *Journal of Cross-Cultural Psychology*, 34, 433-452.
- Cheung, F. M., Van de Vijver, F. J., & Leong, F. T. (2011). Toward a new approach to the study of personality in culture. *American Psychologist*, 66, 593-603.
- Duan, W. J., Bai, Y., Tang, X. Q., Wang, Z. Z., Zhang, Y. H., & Engineering Information, I. (2011). *Construct validity of the values in action inventory of strengths in Chinese culture context-based on a Chinese sample of undergraduate students*. Paper presented at the Conference on Psychology and Social Harmony. Retrieved from <Go to ISI>://WOS:000298569200047
- Duan, W. J., Tang, X. Q., Wang, Z. Z., Zhang, Y. H., & Cceoc. (2011). *Investigation and analysis of character strengths of Chinese college students a cross-cultural perspective*. Paper presented at the Conference on Creative Education. Retrieved from <Go to ISI>://WOS:000297713000132
- Fischer, R., Ferreira, M., Assmar, E. M. L., Redford, P., & Harb, C. (2005). Organisational behaviour across cultures: Theoretical and methodological issues for developing multi-level frameworks involving culture. *Journal of Cross Cultural Management*, 5, 27-48.
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25, 186-192.
- Headland, T. N. (1990). A dialogue between Kenneth Pike and Marvin Harris on emics and etics. In T. N. Headland, K. L. Pike & M. Harris (Eds.), *Emics and etics: The insider/outsider debate* (Vol. *Frontiers of anthropology*). Newbury Park, Calif: Sage Publications.
- Ho, S. M. Y., & Cheung, M. W. L. (2007). Using the Combined Etic-Emic approach to develop a measurement of interpersonal subjective well-being in Chinese populations. In A. D. Ong. & M. V. Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 139-152). New York: Oxford University Press.
- Hofstede, G. (1980). *Culture's Consequences: International differences in work-related values*. Beverly Hills, CA: Sage.
- Hui, C. H., & Triandis, H. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology*, 16, 131-152.
- Johnson, T. P. (1998). *Approaches to equivalence in cross-cultural and cross-national survey research*. *ZUMA-Nachrichten Spezial*, 3, 1-40.
- Kankaraš, M., & Moors, G. (2010). Researching measurement equivalence in cross-cultural studies. *Psihologija*, 43, 121-135.
- Kim, H., & Markus, H. R. (1999). Deviance or uniqueness, harmony or conformity? A cultural analysis. *Journal of Personality and Social Psychology*, 77, 785-800.
- Leong, F. T., Leung, K., & Cheung, F. M. (2010). Integrating cross-cultural psychology research methods into ethnic minority psychology. *Cultural Diversity and Ethnic Minority Psychology*, 16, 590-597.
- Lett, J. (1990). Emics and etics: Notes on the epistemology of anthropology. In T. N. Headland, K. L. Pike, & M. Harris (Eds.), *Emics and etics: The insider/outsider debate* (Vol. *Frontiers of anthropology*). Newbury Park, Calif: Sage Publications.
- Leung, K. (1989). Cross-cultural differences: Individual-level vs culture-level analysis. *International Journal of Psychology*, 24, 703-719.
- Leung, K., & Zhou, F. (2010). Cross-cultural research methods: Review and prospect. *Acta Psychologica Sinica*, 42, 41-47.
- Lonner, W. J., & Berry, J. W. (1986). *Field methods in cross-cultural research*. Beverly Hills, CA: Sage.
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modelling in psychological research. *Annual Review of Psychology*, 51, 201-226.
- Maurer, T. J., Raju, N. S., & Collins, W. C. (1998). Peer and subordinate performance appraisal measurement equivalence. *Journal of Applied Psychology*, 83, 693-702.
- Morris, M. W., Leung, K., Ames, D., & Lickel, B. (1999). Views from inside and outside: Integrating Emic and Etic Insights about Culture and Justice Judgment. *The Academy of Management Review*, 24, 781.
- Organization, W. H. (2011). Process of translation and adaptation of instruments. http://www.who.int/substance_abuse/research_tools/translation/en/
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, 128, 3-72.
- Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, 54, 741-754.
- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Washington DC: American Psychological Association.
- Steger, M. F., Hicks, B. M., Kashdan, T. B., Krueger, R. F., & Bouchard, T. J. (2007). Genetic and environmental influences on the positive traits of the values in action classification, and biometric covariance with normal personality. *Journal of Research in Personality*, 41, 524-539.
- Usunier, J. C. (1998). *International and cross-cultural management research*. London: Sage.
- Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Van de Vijver, F. J. R., & Leung, K. (2000). Methodological issues in psychological research on cultures. *Journal of Cross-Cultural Psychology*, 31, 33-51.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-70.