

Method Effects in Rosenberg Self Esteem Scale and Their Correlates

Qingke Guo, Yang Hu

Department of Psychology, Shandong Normal University, Jinan Shandong
Email: guoqingke@163.com

Received: Dec. 7th, 2015; accepted: Dec. 18th, 2015; published: Dec. 24th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

There is an ongoing debate on the use of negatively worded items in Rosenberg Self Esteem Scale (RSES) and other personality measures. This study was designed to explore the following questions: Do the five positively and five negatively worded RSES items measure one factor or two oblique factors? Do negative wording effect and positive wording effect exist in Chinese culture? If the answer is yes, to what extent do they affect structural and criterion related validity of RSES? Do negative and positive wording effect abstracted from different personality measures exhibit congruent validity? Can they be considered response style or meaningful personality variables? What features can be observed in Chinese participants that differentiate them from Western ones? 525 undergraduates were recruited and were administered RSES, BFI, STAI, CES-D, GHQ-12, IRI, WLEIS, GWB, and MCSD. The results showed: the positively and negatively worded items in RSES perform similar psychometric properties; Method effect can be observed both in positively and negatively worded items, which are detrimental to the construct validity of RSES; When one trait factor and two method factors are defined in CTCM, close model data fit result in; In RSES the size of positive method effect is much smaller than negative method effect; Method effects in RSES and other personality measures are similar in nature; Modest bias is confirmed in the present study, Chinese people show stronger tendency to admit negative aspects of their self-concepts. As a conclusion, balanced use of positively and negatively worded items is encouraged. Advanced statistic method should be employed to control method factors and estimate the true relationships between constructs measured by different personality measures. Chinese people are dialectical and are more modest when assessing their self-esteem.

Keywords

Rosenberg Self Esteem Scale, Method Effects, Big Five Personality Models, Culture Difference

Rosenberg自尊量表中的方法因素及其相关物

郭庆科, 胡 杨

山东师范大学心理学院, 山东 济南

Email: guoqingke@163.com

收稿日期: 2015年12月7日; 录用日期: 2015年12月18日; 发布日期: 2015年12月24日

摘 要

Rosenberg自尊量表的因素结构仍然是学术界争论的焦点问题。本研究拟解决如下问题: 量表测量的是单维还是双维结构, 中国被试中是否存在正向和反向方法因素, 这两种方法因素在多大程度上影响了测验的结构效度和效标关联效度, 不同测量工具中的方法因素是否同质, 是否可解释为一种稳定的反应风格或人格变量, 中国人的自尊量表得分有哪些不同于西方人的独有特征等。对525名大学生被试测试了自尊、大五人格、一般健康等人格量表, 结果发现自尊量表中的正向和反向题有相似的心理测量学性能, 测量相同特质。正向和反向方法因素在中国被试中都存在, 但没有破坏自尊的单维结构。反向方法因素对测验结果的影响更强。不同测量工具中的方法因素具有同质性, 属于系统误差, 具有高抑郁、高神经质等特征的人容易受方法因素的影响。中国被试在自尊上表现出了谦虚、含蓄, 能辩证地认识积极和消极自我等东方文化的特质。

关键词

Rosenberg自尊量表, 方法因素, 大五人格模型, 文化差异

1. 前言

1.1. 关于 Rosenberg 自尊量表因素结构的争议

自尊是积极心理学中的重要概念, 被认为是消极情绪的缓冲机制, 与身心健康、幸福感、甚至寿命等都有密切关系(Urbán, Szigeti, Kökönyei, & Demetrovics, 2014)。一般自尊测量中最常用的量表之一是Rosenberg自尊量表(RSES)。量表中有5个正向题和5个反向题。RSES最初被设计为单维结构, 但探索性和验证性因素分析却经常发现正向和反向题测量的是两种不同的自尊因素(5个正向题测量的因素被命名为正向自尊, 5个反向题测量的因素被命名为反向自尊)(Ang, Neubronner, Oh, & Leong, 2006; Boduszek, Hyland, Dinghra, & Mallet, 2013)。有人甚至将自尊区分为自我胜任感和自我喜爱两个维度(Schmitt & Allik, 2005)。但近来除在一些特殊群体(如少年犯)中的研究发现二因素模型最能拟合数据外(Boduszek et al., 2013), RSES在少年至成年的正常人群中都测量单维的自尊结构。

如何认识自尊量表中的方法因素(method factor, MF), 即题目陈述方向带来的反应偏差? Marsh (1996)首先发现, 在CFA分析中引入与反向题目有关的MF后, RSES测量了单维的自尊结构。此后的研究进一步发现, 当正向题目MF与反向题目MF都包括在模型中时, 模型拟合才更理想(Marsh, Scalas, & Nagengast, 2010; Quilty, Oakman, & Risko, 2006), 说明正向题与反向题都存在MF。相比之下, 一般认为反向题目带来的方法偏差更大, 而正向题目带来的MF较少(Huang & Dong, 2012)。但也有研究发现正向

题目的 MF 更大。Wang, Siegal, Falck 和 Carlson (2001)在吸毒人群中发现正向 MF 的影响远远大于反向 MF。而 Wu (2008)在台湾被试中的研究发现正向与反向自尊题目对测验结构效度的影响程度相当,因此他们建议中国文化中的自尊测量应该平衡使用正向和反向题。

Schmitt 和 Allik (2005)发现正向与反向自尊量表都与神经质和外向性等人格变量存在非常一致的相关,说明二者测量的内容相同,从反面支持了自尊的单维结构。同时也发现了反向题目偏差,即在反向题上得分显著低于正向题。他们还发现了文化因素对自尊量表的影响。

1.2. 方法因素的性质及其相关物

题目陈述效应在不同人群中的表现形式和效应并不一致,这与被试如何理解问卷题目有关(Schmitt & Allik, 2005; Wang et al., 2001)。Marsh (1996)发现与反向题目有关的人为因素(artifacts)在低言语能力的人中作用较明显,在高言语能力的人中则作用降低。Tomás, Oliver, Galiana, Sancho 和 Lila (2013)发现 RSES 等自尊量表中的 MF 与焦虑水平显著负相关。这是因为焦虑的人会对威胁性的和不确定性的信息优先注意,所以对反向题目有赞同偏好。

更多学者则倾向于将 MF 视为一种稳定的反应风格,甚至是有用的人格变量(Urbán et al., 2014)。RSES 中的反向 MF 与其他量表(心理控制源量表等)中的 MF 存在一致性,且具有跨年龄和跨性别的稳定性(Lindwall et al., 2012; Urbán et al., 2014)。有研究发现自我意识强、害怕负面评价的人更少地表现出与反向题目有关的 MF。回避动机则与反向 MF 负相关(可以推论高回避动机的人更愿意赞同反向题)。大五人格中的责任感、情绪稳定性也与反向 MF 表现出正相关(即责任感强、情绪稳定的人更不愿意赞同反向题)(Quilty, Oakman, & Risko, 2006)。Lindwall et al. (2012)进一步发现反向 MF 与呈抑郁负相关而与生活满意度呈正相关。说明高抑郁、低生活满意度的人更倾向于赞同反向题目。正向 MF 则与生活满意度负相关。但 MF 与社会赞许性(SDR)不具有相关性,说明是一种不同于 SDR 的系统误差。

1.3. 中国人的自尊特点

Schmitt 和 Allik (2005)发现 RSES 的正向与反向自尊维度与外部效标的相关相当一致,但 Ang 等(2006)在新加坡学生中发现二维自尊模型更能拟合数据,且正向自尊能预测掌握目标导向(mastery goal orientation)和自我效能感,而反向自尊则能够预测破坏行为。Kim, Peng 和 Chiu (2008)在中国被试也发现正向自尊题目能预测坚持性、心理坚强、寻求挑战和达成挑战性目标的行为,反向题则不能。这种不一致很可能是由于中国人与西方人相比更不需要保持自我的统一性(Boucher, Peng, Shi, & Wang, 2009)。另有研究发现中国被试中反向自尊与抑郁正相关,而正向自尊则与抑郁没有相关(Supple, Su, Plunkett, Peterson, & Bush, 2013)。这暗示中国被试中可能存在二元自尊结构,两种自尊有不同的心理测量学性质。

多数学者承认的另一个事实是中国人在自尊量表上的得分低于西方人(Spencer-Rodgers, Peng, Wang, & Hou, 2004; Kim, et al., 2008)。其原因很可能是中国人拥有辩证的自我观。中国人相信任何事物都有正反两面。因此既能接纳积极自我又能接纳消极自我。表现在自尊量表上就是既赞同正向题又赞同反向题。西方人则追求自我的统一性,其在正向题和反向题上的得分高度一致。正是因为中国人对反向记分自尊题目的肯定倾向导致其自尊量表得分低于西方人。Kim 等人(2008)的研究证明了这一特点,同时发现中国被试在正向和反向自尊题目上得分的一致性低于北美人。

也有人认为中国人既倾向于赞同反向题,又在正向题上存在谦虚偏差(Wu, 2008; Kim, et al., 2008)。如 Farh 和 Cheng (1997)发现台湾公司职员的自我绩效评估分数显著低于上级对他们的评价,这种效应在正向题目中表现得尤为明显。下属的自我绩效评估中包括 2 个因素(分别载荷于正向题目和反向题目),而用同一量表对上司的评估却得到了单维结构。可能是中国文化中的谦逊价值观造成了这一结果,而且正

向题目会扩大这一趋势。因此他们建议反向题目更符合中国被试。

1.4. 问题的提出

关于 RSES 中的 MF 问题是几十年来争论的热点, 而且波及到其他自陈量表。很多重要的理论性问题仍然没有得到解决。而 MF 在中国被试中的表现形式和特点尤其值得研究。概括起来, 本研究拟解决的问题有: 1) 自尊量表中的 MF 在中国被试中是否存在。2) 在承认正向和反向 MF 都存在的前提下, 测验的单维因素结构是否被破坏。3) 如果正向和反向 MF 都存在, 哪种 MF 所带来的偏差更强? 4) 正向和反向 MF 会不会对题目所测量的内容产生实质性影响? 5) 正向和反向 MF 在其他自陈式量表中是否也存在? 6) 正向和反向 MF 与人格量表间有怎样的关系, 是不是有用的人格变量? 7) 中国人在自尊量表上的得分有哪些不同于国外的特点?

2. 研究方法

2.1. 测量工具

本研究中的测量工具都有中文修订版且信效度经过了检验。Rosenberg 自尊量表(RSES)。包含 10 个测量一般自尊的题目, 4 级记分, Cronbach α 系数为 0.885。自尊对人的身心健康领域的人格变量有广泛影响(Schmitt & Allik, 2005; Urbán, et al., 2014)。因此除自尊量表外, 按照本研究的目的, 通过参阅文献, 我们选择了可作为自尊测量效标的人格测量工具。包括 BFI-44, Cronbach α 系数为 0.680; 一般健康问卷 GHQ-12, Cronbach α 系数为 0.829; 状态 - 特质焦虑问卷 STAI, 状态焦虑和特质焦虑量表的 Cronbach α 系数分别为 0.921、0.884; 流调中心用抑郁量表 CES-D, Cronbach α 系数为 0.800; 中文版人际反应指针 IRI-C, Cronbach α 系数为 0.744; WLEIS 情商量表, Cronbach α 系数为 0.876; 总体幸福感量表 GWB, Cronbach α 系数为 0.851; 社会期望量表 MCS-D, Cronbach α 系数为 0.690。

2.2. 被试和程序

采用整群抽样的方式从东部某省份一所综合性大学中选取 525 名大学生(男 239 人, 女 286 人)为被试, 以团体施测的方式实施测验。每位被试都得到礼物作为补偿。平均年龄 20.13 岁, 标准差 1.97 岁。该校生源来自全国各地, 因此样本具有较好的代表性。

2.3. 统计分析方法

历史上关于 RSES 的研究模型多达 10 个。考虑到有的模型本身没有实际价值, 本研究只采纳了 1、2、3 因素模型, 并依据 CTCM 方法设置研究模型。CTCM 可以估计 MF 的大小及与其他量表中特质因素和方法因素的关系, 这是 CTCU 所做不到的。研究中用 SPSS18.0 进行初步的数据处理和信度分析, 用 AMOS17.0 进行模型验证。

3. 结果

3.1. 自尊与其他人格测量的信度

首先计算出自尊与其他人格测量正反向陈述分量表及总量表的信度(表 1)。情商量表中只有正向题, MCS-D 量表的因素结构在中国还没有得到很好的确认, 因此本研究只计算了它们的总分和总信度。

表中列出了本研究中所用的人格量表的信度(alpha 系数)。可以看出各量表的总体信度都在可接受的范围内, 保证了研究结论的可靠性。总体上看反向记分量表的信度略低于正向量表, 但差异不明显。而反向自尊量表的信度甚至超过了正向自尊量表。

Table 1. Reliability of the positive and negative subscales in RSES and other psychological measures
表 1. 自尊等量表正反向陈述分量表及总量表的信度

| | 自尊 | 一般健康 | 状态焦虑 | 特质焦虑 | 抑郁 | 同情 | 情商 | 主观幸福 | 外向 | 宜人 | 责任 | 神经质 | 开放 | MCSD |
|------|------|------|------|------|------|------|-------|------|------|------|------|------|------|------|
| 正向量表 | 0.81 | 0.72 | 0.92 | 0.83 | 0.91 | 0.77 | | 0.61 | 0.67 | 0.65 | 0.69 | 0.62 | 0.67 | |
| 反向量表 | 0.88 | 0.76 | 0.90 | 0.88 | 0.75 | 0.45 | | 0.83 | 0.73 | 0.58 | 0.70 | 0.67 | 0.56 | |
| 总量表 | 0.89 | 0.83 | 0.78 | 0.74 | 0.90 | 0.80 | 0.876 | 0.84 | 0.77 | 0.73 | 0.76 | 0.71 | 0.68 | 0.70 |

3.2. 自尊与其他人格测量间的相关

然后根据研究目的计算出了正向自尊、反向自尊与可能的效标变量(根据题目陈述方向也划分了正向和反向分量表)间的相关(表 2)。

正向与反向自尊如果与其他人格测量间存在一致性的相关,则证明正向与反向自尊其实是一种自尊。为更详细地评估 MF,我们对每种人格量表都区分出了正向与反向两种形式。这里的考虑是:如果 MF 对测验影响不大,则用正向与反向题目测量的特质与用正向与反向题目测量的特质间的相关应该相当一致。从结果中可以看出正向自尊、反向自尊与外部效标间的相关相当一致,说明题目陈述方向没有改变自尊量表题目所测量的内容。支持自尊量表的单维结构。总体自尊与 SDR 呈负相关,说明自尊测量较少受社会赞许性的影响。

从另一角度分析,如果 MF 的作用强大,则题目陈述方向相同的两个特质得分间的相关就更大。即正向自尊与正向健康由于共享了正向 MF,其相关应大于正向自尊与反向健康间的相关。同样反向自尊与反向健康由于共享了反向 MF,其相关也应该更大。从上表的结果中可以看出正向自尊、反向自尊与其他人格量表正反两种得分间的相关并未表现出明显差异,说明正向与反向 MF 并没有导致测验题目背离所测量的内容。这一结论适用于本研究中的所有人格量表。

正向自尊与反向自尊量表的差值也是个重要的分析指标。差值为正说明被试更愿意赞同正向自尊题目,差值为负说明被试更愿意赞同反向自尊题目。这不是一种有价值的人格倾向?从结果中可以看出正反向自尊的差值与抑郁、神经质、同情心呈负相关。这说明越情绪化的人越愿意赞同反向题。这与先前的结论(Lindwall et al., 2012)非常一致。

3.3. 方法因素对模型拟合的影响

为探讨 MF 是否影响自尊结构,设置了如下模型(表 3)。包括单维自尊模型 M1a、双维自尊模型 M1b、含有正向 MF 的自尊模型 M1c、含有反向 MF 的自尊模型 M1d、含有正向和反向 MF 的自尊模型 M1d。同时再将自尊、一般健康、状态焦虑量表的题目放在一个模型中估计,产生了三因素模型 M2a(模型中只设置了自尊、一般健康、状态焦虑 3 个特质)、M2b(设置了正向 MF)、M2c(设置了反向 MF)。通过模型比较可判断自尊量表、一般健康量表、状态焦虑量表中的 MF 是否有相同的性质。

从表 3 中可见所有单因素和二因素模型都不符合模型拟合严格标准(Marsh, Hau, & Wen, 2004)。而引入 MF 后模型拟合得到了明显提高。但只有在正向和反向 MF 同时纳入模型后单维量表的模型拟合才接近或超过模型拟合的严格标准。说明正向和反向 MF 在不同量表中都存在,且反向 MF 对测验结构的破坏力更大。

将 M2b、M2c 与 M2a 比较,发现其拟合是相当理想的。因此可以认为自尊、一般健康、状态焦虑量表中的正向、反向 MF 具备同质性,是性质相同的反应偏差。

Table 2. Correlation between positive and negative RSES subscales and criterions
表 2. 正向自尊、反向自尊与可能的效标变量间的相关

| | 全量表 | | | | 正向量表 | | | | 反向量表 | | | |
|------|--------|--------|--------|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 正向自尊 | 反向自尊 | 总自尊 | 差值 | 正向自尊 | 反向自尊 | 总自尊 | 差值 | 正向自尊 | 反向自尊 | 总自尊 | 差值 |
| 一般健康 | 0.46* | 0.46* | 0.52* | 0.05 | 0.33 | 0.43 | 0.42 | -0.06 | 0.46 | 0.38 | 0.47 | 0.14 |
| 状态焦虑 | -0.42* | -0.44* | -0.47* | -0.03 | -0.36 | -0.44 | -0.44 | 0.03 | -0.35 | -0.31 | -0.37 | -0.09 |
| 特质焦虑 | -0.52* | -0.51* | -0.57* | -0.07 | -0.53 | -0.44 | -0.54 | -0.17 | -0.32 | -0.44 | -0.42 | 0.08 |
| 情商 | 0.31* | 0.41* | 0.40* | -0.06 | | | | | | | | |
| 主观幸福 | 0.40* | 0.40* | 0.45* | 0.06 | 0.42 | 0.39 | 0.45 | 0.09 | 0.30 | 0.34 | 0.36 | 0.00 |
| 外向 | 0.35* | 0.29* | 0.36* | 0.10* | 0.34 | 0.36 | 0.39 | 0.03 | 0.26 | 0.11 | 0.21 | 0.18 |
| 宜人 | 0.26* | 0.31* | 0.32* | -0.01 | 0.19 | 0.34 | 0.29 | -0.13 | 0.25 | 0.18 | 0.25 | 0.10 |
| 责任 | 0.33* | 0.32* | 0.37* | 0.05 | 0.29 | 0.32 | 0.34 | 0.00 | 0.29 | 0.25 | 0.31 | 0.08 |
| 神经质 | -0.44* | -0.36* | -0.45* | -0.14* | -0.44 | -0.32 | -0.43 | -0.18 | -0.31 | -0.32 | -0.35 | -0.03 |
| 开放 | 0.20* | 0.28* | 0.26* | -0.06 | 0.19 | 0.32 | 0.28 | -0.10 | 0.09 | 0.05 | 0.08 | 0.06 |
| 抑郁 | -0.52* | -0.43* | -0.53* | -0.15* | -0.49 | -0.39 | -0.49 | -0.17 | -0.43 | -0.47 | -0.50 | -0.02 |
| 同情 | 0.00 | 0.13* | 0.07 | -0.14* | -0.02 | 0.13 | 0.06 | -0.17 | 0.02 | -0.01 | 0.00 | 0.03 |
| SDR | -0.15* | -0.09 | -0.14* | -0.08 | | | | | | | | |

注: * $p < 0.05$, ** $p < 0.01$ 。

Table 3. Fitness statistics of the hypothesized models
表 3. 研究模型的拟合优度统计量

| | χ^2 | df | CFI | TLI | RMSEA | SRMR | AIC |
|-----|----------|------|------|------|-------|-------|---------|
| M1a | 509.63 | 35 | 0.90 | 0.87 | 0.17 | 0.081 | 549.63 |
| M1b | 329.82 | 34 | 0.94 | 0.92 | 0.13 | 0.061 | 371.82 |
| M1c | 189.78 | 30 | 0.96 | 0.95 | 0.10 | 0.049 | 239.78 |
| M1d | 138.82 | 30 | 0.98 | 0.96 | 0.086 | 0.036 | 188.82 |
| M1e | 74.95 | 24 | 0.99 | 0.98 | 0.066 | 0.027 | 136.92 |
| M2a | 5643.49 | 816 | 0.92 | 0.91 | 0.11 | 0.079 | 5817.49 |
| M2b | 4175.62 | 792 | 0.93 | 0.92 | 0.094 | 0.076 | 4397.62 |
| M2c | 3820.02 | 792 | 0.93 | 0.93 | 0.089 | 0.074 | 4042.02 |

3.4. 方法效应的大小及与人格的相关

一般用特质因素和方法因素对应的因子载荷的平方和表示特质效应和方法效应。结果发现自尊量表反向题目方法载荷的平均值(0.48)明显大于正向题目(0.21)。说明反向 MF 对测验结果的影响更大。因子载荷的平方和也是反向题目(1.32)大于正向题目(0.30)。这与先前研究很一致。

在分离出自尊量表中的 MF 后我们分析了 MF 与不同人格变量的关系。具体做法是将 MF 设置为自变量, 将人格量表得分设置为因变量, 在 AMOS17.0 软件中进行单自变量、单因变量的路径分析。以便计算出的路径系数不受其他变量影响。计算结果见表 4。

从表 4 中可见正向与反向 MF 都与人格变量之间存在一定相关, 相关的方向与国外研究基本一致。抑郁、一般健康、神经质、特质焦虑与反向 MF 的相关相当可观, 说明高抑郁、心理健康状况差、情绪不稳定、焦虑的被试更倾向于赞同反向题目。正向 MF 则与情商、一般健康、经验开放性等有相当可观的负相关, 说明有这些特点的人更不愿意赞同正向自尊题目。这看起来矛盾, 却揭示了另一个事实: 情商高、心理健康、经验开放的人更能辩证地认识自我, 对自己的优点并不急于承认, 这恰恰是一种成熟的自信。

3.5. 中国被试在自尊量表上的得分模式

本研究发现大学生在正向自尊量表上的得分虽然高于反向自尊量表($p < 0.00$), 但差异(1.26)却没有想象的大, 与美国被试(1.2)和英国被试(1.1)非常接近。这一结果与先前研究有一定出入(Schmitt & Allik, 2005; Wu, 2008)。中国被试的正向和反向自尊量表得分都低于美国被试, 但没有证据表明中国人在自尊量表上的低分是由于中国人更愿意赞同反向题目。在 Schmitt 和 Allik 的样本中香港被试正向和反向自尊量表上的均分差值是 1.8, 台湾则是 2.3。Schmitt 和 Allik (2005)的研究是 10 年前进行的, 10 年前香港和台湾人在自尊量表上的得分模式更像欠发达国家的被试。而本研究的结果却更接近世界发达国家。这说明中国人寻求自我统一性的需求增强了。

但中国被试在正向和反向自尊量表上的得分仍然低于美国等西方国家。这部分可以用 Farh 和 Cheng (1997)的观点来解释: 中国被试在正向自尊题目上可能有低估自己的倾向。其实这正是谦虚偏差的表现。即一般人不会直接承认自己的优点(即赞同正向自尊量表上的题目), 这才体现出含蓄和教养。这与先前研究是一致的。即中国被试在自尊量表上存在谦虚偏差, 能辩证地认识积极和消极自我。

4. 讨论

4.1. 自尊量表中反向题目的使用

测验中是否该使用反向题目仍然是目前心理学界争论的热点。有人认为反向题目能起到认知减速器的作用, 能促使被试在完全搞懂题目的意义后再作答。正向与反向题目的平衡使用可以抵消默认等反应偏差, 相加得到的总分就较少包含测量误差, 也就能更准确地计算与其它变量的关系(Podsakoff, MacKenzie, & Podsakoff, 2012)。

本研究从正向与反向自尊量表的信度、与外部效标的相关等方面都没有发现正向和反向自尊量表在信效度上的差异。而且有足够证据表明正向和反向自尊量表测量了同一种特质。尽管自尊量表中反向方法因素解释的总变异量大于正向方法因素, 但二者对模型拟合的贡献又难分伯仲。只有同时纳入两种 MF 时模型拟合才能达到满意的水平。这些结果表明自尊量表及本研究所涉及的其他量表中反向题目的心理测量学性能并不比正向题目差。两类题目都会带来测量误差, 都应该控制。综合以上分析我们赞同中国人的人格测量中应平衡使用正向和反向题目的结论(Wu, 2008)。

Table 4. Correlation between method factors and personality measures

表 4. 方法因素与人格的相关

| | 一般健康 | 状态焦虑 | 特质焦虑 | 抑郁 | 幸福感 | 情商 | 外向 | 宜人 | 责任 | 神经质 | 开放 | 同情心 |
|------|--------|-------|--------|--------|-------|--------|--------|--------|--------|--------|--------|--------|
| 正向方法 | -0.26* | -0.11 | -0.10 | 0.25* | 0.06 | -0.35* | -0.18* | -0.13* | -0.13* | -0.06 | -0.24* | -0.11 |
| 反向方法 | 0.32* | 0.01 | -0.21* | -0.37* | 0.20* | 0.14* | 0.04 | 0.20* | -0.07 | -0.28* | 0.04 | -0.16* |

注: * $p < 0.05$ 。

4.2. 方法因素的性质

本研究发现不同人格量表中的 MF 具备同质性, MF 与多个人格特质存在不容忽视的相关, 说明 MF 不是随机误差而是系统误差, 甚至是有用的人格变量。即不论测量的是哪些人格特质, 只要采用反向(或正向)陈述的形式, 有某些人格特征的人(如抑郁、神经质等)都会做出有规律性的、异常的应答模式。从被试在反向和正向陈述题目上得分的差异也能揭示出关于被试人格的有用信息。本研究发现正向和反向 MF 都与社会期望性量表(MCSD)存在低相关, 这与以往研究是一致的。说明 MF 与社会赞许性不是同一种测验偏差。结合本研究 and 以往研究的发现, 我们认为在抑郁、神经质、焦虑、回避动机、自我意识等方面表现出高分特点的被试容易受 MF 的影响, 因此设计测验题目时要考虑这些被试的特点, 对他们的测验分数要采用某种手段进行校正。

4.3. 中国人的自尊特征

Schmitt 和 Allik (2005)的研究发现一个国家的文化特质会影响人们对反向自尊量表题目的理解, 导致自尊量表得分的跨文化比较出现困难。几乎所有国家都存在反向题偏差, 即正向自尊量表得分大于反向自尊量表。而且二者的差值越大测验信度越低。发达国家的差值都不大。比如差值最小的是德国(0), 然后是其他欧美国家和日本等。差值最大的是墨西哥(3.4)和玻利维亚(3.3), 非洲国家(如民主刚果为 3.4)也比较大。集体主义强的国家趋中反应倾向明显, 而且总分较低, 比如日本(0.8)。尽管日本人的自尊得分低于美国人, 自尊量表的心理测量学性能却与美国人相近。自尊量表的编写和标准化是在美国进行的, 但却适用于与美国文化异质的日本。这显示了心理机制的普适性。但对非洲国家则不然。

本研究发现中国被试的自尊得分模式与 10 年前的香港和台湾有较大差异, 但却更接近 Schmitt 和 Allik (2005)研究中的发达国家。即中国人寻求自我统一性的需求增强了, 另外在自尊量表的信效度、因素结构、正向与反向量表的相关上中国被试也接近发达国家。这应该能折射出几十年来中国人社会文化心态的变化。

同时中国被试在正向和反向自尊量表上的得分仍然体现了东方文化的特点。谦虚、含蓄, 不直接承认自己的优点, 也不掩饰自己的缺点, 能辩证地认识自我的积极面和消极面等。

5. 结论

1) 自尊量表中的正向和反向题测量相同的因素, 在信效度等心理测量学指标上没有实质性差异, 因此应该平衡使用。而且这一结论可以推广到本研究所涉及到的人格测量工具。

2) 正向和反向 MF 在中国被试中都存在, 但却没有破坏自尊的单维结构。用 CTCM 等技术将 MF 分离出来后就能估计自尊与其他研究变量间的真实关系。

3) 自尊及其他自陈式量表中的 MF, 可视为一种与测量工具有关的系统误差, 或有用的人格变量。在抑郁、神经质、焦虑、回避动机、自我意识等方面表现出高分特点的被试容易受 MF 的影响。在解释测验结果时要考虑被试的人格特点, 必要时应采用某种手段进行校正。

4) 对自尊量表进行跨文化比较时要考虑文化背景。中国被试在自尊量表上的得分体现了东方文化的特点, 如谦虚、含蓄, 能辩证地认识自我的积极面和消极面等。这一结论似乎也可以推论到其他自陈式人格测量。

参考文献 (References)

Ang, R., Neubronner, M., Oh, S., & Leong, V. (2006). Dimensionality of Rosenberg's Self-Esteem Scale among Normal-Technical Stream Students in Singapore. *Current Psychology: Developmental, Learning, Personality, Social*, 25, 120-131.

- <http://dx.doi.org/10.1007/s12144-006-1007-3>
- Boduszek, D., Hyland, P., Dinghra, K., & Mallet, J. (2013). The Factor Structure and Composite Reliability of the Rosenberg Self-Esteem Scale among Ex-Prisoners. *Personality and Individual Differences*, 55, 877-887. <http://dx.doi.org/10.1016/j.paid.2013.07.014>
- Boucher, H. C., Peng, K., Shi, J., & Wang, L. (2009). Culture and Implicit Self-Esteem: Chinese Are “Good” and “Bad” at the Same Time. *Journal of Cross-Cultural Psychology*, 40, 24-45. <http://dx.doi.org/10.1177/0022022108326195>
- Farh, J. L., & Cheng, B. S. (1997). Modesty Bias in Self-Ratings in Taiwan: Impact of Item Wording, Modesty Values, and Self-Esteem. *Chinese Journal of Psychology*, 39, 103-118.
- Kim, Y. H., Peng, S., & Chiu, C. Y. (2008). Explaining Self-Esteem Differences between Chinese and North Americans: Dialectical Self (vs. Self-Consistency) or Lack of Positive Self-Regard. *Self and Identity*, 7, 113-128. <http://dx.doi.org/10.1080/15298860601063437>
- Lindwall, M., Barkoukis, V., Grano, C., Lucidi, F., Raudsepp, L., Liukkonen, J., & Thøgersen-Ntoumani, C. (2012). Method Effects: The Problem With Negatively Versus Positively Keyed Items. *Journal of Personality Assessment*, 94, 196-204. <http://dx.doi.org/10.1080/00223891.2011.645936>
- Marsh, H. W. (1996). Positive and Negative Global Self Esteem: A Substantively Meaningful Distinction or Artifacts? *Journal of Personality and Social Psychology*, 70, 810-819. <http://dx.doi.org/10.1037/0022-3514.70.4.810>
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler’s (1999) Findings. *Structural Equation Modeling*, 11, 320-341. http://dx.doi.org/10.1207/s15328007sem1103_2
- Marsh, H. W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal Tests of Competing Factor Structures for the Rosenberg Self-Esteem Scale: Traits, Ephemeral Artifacts, and Stable Response Styles. *Psychological Assessment*, 22, 366-381. <http://dx.doi.org/10.1037/a0019225>
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. *Annual Review of Psychology*, 63, 539-569. <http://dx.doi.org/10.1146/annurev-psych-120710-100452>
- Quilty, L. C., Oakman, J. M., & Risko, E. (2006). Correlates of the Rosenberg Self-Esteem Scale Method Effects. *Structural Equation Modeling*, 13, 99-117. http://dx.doi.org/10.1207/s15328007sem1301_5
- Schmitt, D. P., & Allik, J. (2005). Simultaneous Administration of the Rosenberg Self-Esteem Scale in 53 Nations: Exploring the Universal and Culture-Specific Features of Global Self-Esteem. *Journal of Personality and Social Psychology*, 89, 623-642. <http://dx.doi.org/10.1037/0022-3514.89.4.623>
- Spencer-Rodgers, J., Peng, K., Wang, L., & Hou, Y. (2004). Dialectical Self-Esteem and East-West Differences in Psychological Well-Being. *Personality and Social Psychology Bulletin*, 30, 1416-1432. <http://dx.doi.org/10.1177/0146167204264243>
- Tomás, J. M., Oliver, A., Galiana, L., Sancho, P., & Lila, M. (2013). Explaining Method Effects Associated With Negatively Worded Items in Trait and State Global and Domain-Specific Self-Esteem Scales. *Structural Equation Modeling: A Multidisciplinary Journal*, 20, 299-313. <http://dx.doi.org/10.1080/10705511.2013.769394>
- Urbán, R., Szigeti, R., Kökönyei, G., & Demetrovics, Z. (2014). Global Self-Esteem and Method Effects: Competing Factor Structures, Longitudinal Invariance, and Response Styles in Adolescents. *Behavior Research Methods*, 46, 488-498. <http://dx.doi.org/10.3758/s13428-013-0391-5>
- Wang, J., Siegal, H. A., Falck, R. S., & Carlson, R. G. (2001). Factorial Structure of Rosenberg’s Self-Esteem Scale among Crack-Cocaine Drug Users. *Structural Equation Modeling*, 8, 275-286. http://dx.doi.org/10.1207/S15328007SEM0802_6
- Wu, C. H. (2008). An Examination of the Wording Effect in the Rosenberg Self-Esteem Scale among Culturally Chinese People. *The Journal of Social Psychology*, 148, 535-552. <http://dx.doi.org/10.3200/SOCP.148.5.535-552>