

# 项目表述方法对中文 Rosenberg 自尊量表因子结构的影响

王孟成<sup>1</sup>, 蔡炳光<sup>2</sup>, 吴 艳<sup>3</sup>, 戴晓阳<sup>1</sup>

(1. 深圳大学 师范学院, 深圳 518060; 2. 电白县林头中学, 茂名 525424; 3. 华南师范大学 心理应用研究中心, 广州 510631)

**摘 要:** 该研究探索项目表述方法对量表因子结构的影响, 在多特质—多方法 (MTMM) 矩阵框架下对关联特质—相关模型 (The correlated trait—correlated method, CTCM) 和关联特质—相关独特性模型 (The correlated trait—correlated uniqueness, CTCU) 等 8 个相关方法效应模型进行比较。结果发现, 导致中文 Rosenberg 自尊量表两因子结构的主要是由项目表述的方法效应引起的。因此, 在编制量表过程中应考虑方法效应对因子结构的影响。

**关键词:** 中文 Rosenberg 自尊量表; 项目表述方法; 因子结构; 方法效应

**中图分类号:** B841.2

**文献标识码:** A

**文章编号:** 1003—5184(2010)03—0063—06

## 1 引言

共同方法偏差 (Common Method Bias) 是指由测量的方法而不是所测量的构念 (Construct) 引起的变异, 即同样的数据来源或评分者、同样的测量环境、项目语境以及项目本身特征等相同方法造成的预测变量与效标变量之间虚假的共变关系<sup>[1,2]</sup>。共同方法偏差被当作测量中系统误差的主要来源之一, 在实证研究中常常会混淆研究结果以至于误导结论<sup>[3]</sup>。共同方法偏差的来源很多, Podsakoff 等 (2003) 系统总结了共同方法偏差的种类和来源, 结果达 25 种。反向记分项目 (Reverse—Coded Items) 或负性词语表述 (Negatively Worded Items) 也是经常出现的一种共同方法偏差来源<sup>[1~4]</sup>。因为在编制或修订测验条目时, 测量学家常常建议采用反向记分项目或负性词语表述, 这样可以避免被试由于形成反应定势而对测验效度产生影响。但实证研究则发现采用反向记分项目也会产生方法效应, 当反向记分项目被正向表述后此效应便随之消失<sup>[4]</sup>。

控制共同方法偏差的方法可以分成两大类<sup>[1,2]</sup>: 程序控制 (procedural remedy) 和统计控制。前者是在研究设计阶段找出预测源和效标变量测评的共同之处, 然后通过研究设计来消除和减小其影响。统计控制的方法也有多种, 目前最常用的方法是在多特质—多方法模型 (MTMM) 中采用验证性

因素分析 (CFA) 来进行<sup>[1,2]</sup>, 通常采用的有关联特质—相关方法 (The Correlated Trait—Correlated Method, CTCM) 和关联特质—相关独特性 (The Correlated Trait—Correlated Uniqueness, CTCU) 两种分析方法。

在 CTCM 中将观测方差分解为特质、方法和误差方差三个部分图 1 左图, 这也是传统 MTMM。其中的特质方差一般来说是研究要考察的对象; 方法方差部分是由使用同一种方法而产生的方法效应方差, 换句话说, 是由于同一种方法而产生的独特性, 而这种独特性来自于独立的方法因子; 误差方差则为测量误差造成的变异。CTCM 多用于一个或多个潜在的方法变量对所有以相同方式呈现的项目 (如: 反向记分) 都可能产生影响的情景。由于 CTCM 模型存在识别上的问题<sup>[5]</sup>, 所以研究者提出了 CTCU 模型图 1 右图。在 CTCU 模型中假设不存在独立的方法因子, 而是认为方法效应存在于独特性彼此相关之中。例如, 一份量表中所有反向条目都存在“反向内容表述”这样一个独特性, 在 CTCM 中每个条目的独特性都是由同一个方法因子造成的, 而在 CTCU 中则认为这些反向条目间的独特性是彼此相关的。两个模型间最大的区别是 CTCM 中存在独立的方法效应因子, 而 CTCU 模型中则没有。由于 CTCU 模型中没有独立的方法效应因子, 使得需要估计的参数要少 (以图 1 为例,

CTCU 要比 CTCM 节俭 3 个自由度),特别是在方 法和特质因子较多时比 CTCM 更容易识别<sup>[5,19]</sup>。

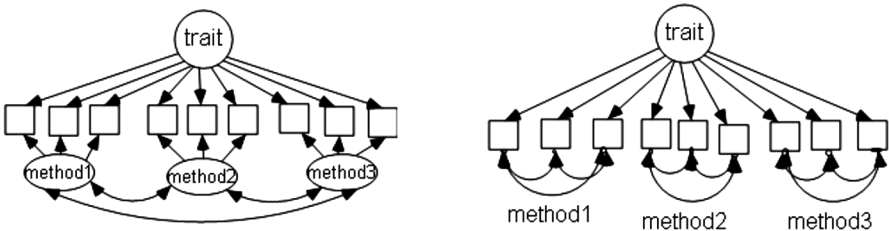


图 1 CTCM 和 CTCU 示意图

长久以来,RSES 被认为是单维度的结构<sup>[7,8]</sup>,但也有研究发现其包含正性自尊和负性自尊两个因子<sup>[9,10]</sup>。然而,方法学角度的研究则一致认为这种两因子结构并非量表内容所致,而是因为方法效应的影响<sup>[11-14]</sup>。Greenberger 将 RSES 的文字表达方式变换成全正性表达或全负性表达方式后,测量结果呈现单因子结构,而原量表正、负性表达各占一半的测试结果则显示出两因子结构。Tomás 和 Oliver(1999)在 MTMM 框架下通过 CFA 发现存在方法效应的模型比不存在方法效应的模型拟合得更好<sup>[13]</sup>。

前期就中文 Rosenberg 自尊量表的研究发现其第 8 题存在翻译上的缺陷<sup>[15,16]</sup>,而最近一项专门针对条目 8 问题的文章在对第 8 题的译文做了修改后探索性因素分析仍得到了两个因子结构<sup>[16]</sup>。究竟中文 RSES 所测量的整体自尊就是由自我肯定和自我否定两个方面构成,还是像国外研究所发现的那

样是由于方法效应引起的呢?

2 研究方法

2.1 被试

来自广东的一所中学的 545 名初中生,其中男生 227 人,女生 318 人;初一 186 人,初二 181 人,初三 178 人。年龄在 11~18 岁之间,平均年龄 15.4,标准差 1.2 岁。

2.2 工具

由于 Rosenberg 自尊量表的大陆译本第 8 题存在翻译上的不恰当<sup>[15,17]</sup>,所以采用台湾 98 年的译本<sup>[18]</sup>。中文 Rosenberg 自尊量表(RSES)由 10 个条目组成,按原量表的设计,其中有 5 个正向评价即 1,2,4,6,7 题和 5 个负向评价即 3,5,8,9,10 题,采用 1~4 级评分,从很不同意到很同意。分值越高表明自尊水平越高。

3 结果

3.1 量表项目的描述性统计分析

表 1 各条目的描述统计量

条目	M	SD	SK	KU	条目	M	SD	SK	KU
T1	2.92	0.70	-0.54	0.64	T6	2.90	0.72	-0.08	-0.51
T2	2.76	0.71	-0.11	-0.24	T7	2.87	0.78	-0.16	-0.58
T3	2.28	0.84	0.50	-0.31	T8	2.82	0.89	0.32	-0.65
T4	3.07	0.71	-0.41	-0.01	T9	3.08	0.85	0.55	-0.42
T5	2.28	0.73	-0.15	-0.24	T10	3.20	0.78	0.70	-0.07

注:SK 为偏态系数,KU 为峰态系数。

表 1 结果表明,被试在 10 个条目上的得分并没有出现天花板效应和地板效应,各条目的偏态分布和峰态系数均在 -1 到 1 之间,说明数据符合多元正态分布,采用 ML 方法进行 CFA 参数估计可以得到稳健的估计值。

3.2 验证的模型

研究共比较了以下 8 个模型(见图 2),其中 M1 表示 10 个条目都负荷在一个整体自尊因子上,该模型为原量表的理论模型;M2 表示 5 条负性记分项目负荷在负性自尊因子上,5 条正向记分项目负荷在正性自尊因子上,该模型是众多探索性因素分析

所得到的模型;M3、M4 和 M5 均是 CTCU 模型,分别表示负性表述条目上存在共同方法误差、正向表述条目上存在共同方法误差和正负向表述中均存在共同方法误差;M6、M7 和 M8 是 CTCM 模型,分别

表示的一个方法效应因子影响 5 条负向表述条目,一个方法效应因子影响 5 条正向表述条目和两个相关的方法效应因子分别影响 5 条正/负向表述条目。

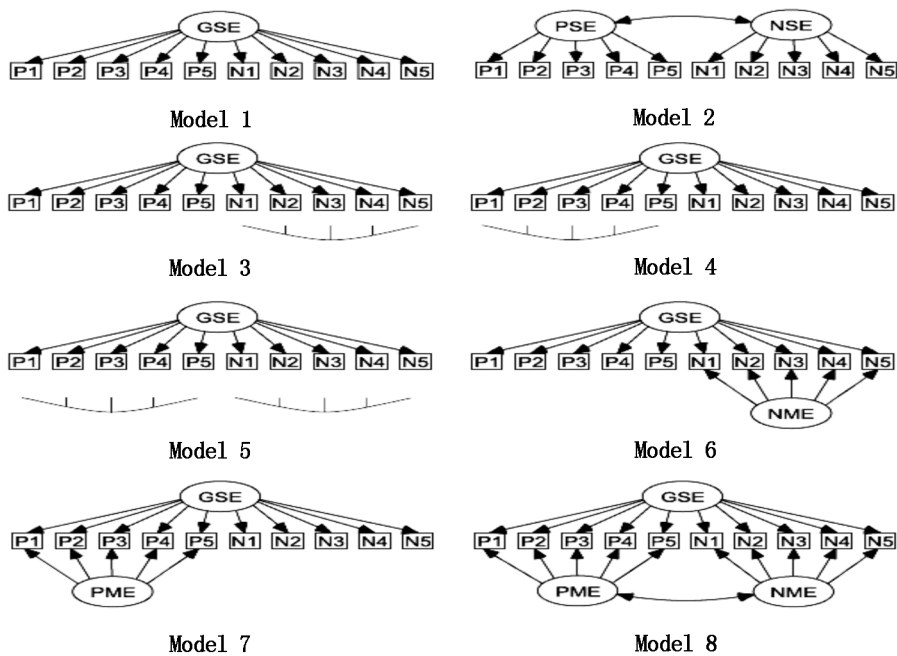


图 2 模型示意图

注:GSE:globe self-esteem,PME:positive method effect,NME:negative method effect,下同

3.3 模型拟合结果

采用三类拟合指标来判断因子模型拟合数据的恰当性:1)绝对拟合指数有: $\chi^2/df$ 、GFI、AGFI、RMSEA、ECVI;2)相对拟合指数有:CFI、NFI、NNFI;3)简约拟合指数采用信息指数 AIC(Akaike Information Criterion),在多个模型比较时其指数越小模型拟合越好,所以用其评价非嵌套模型的优劣<sup>[19,20]</sup>。但是单个的 AIC 指数包含一些来源不清的恒定因素,并且也受样本容量的影响,所以研究者提出了一个修正: $\Delta AIC$ ,其值为备选模型的 AIC 值减去多个模型中拟合最好的模型的 AIC 值,所以

AIC 的值始终为正值。当 $\Delta AIC \leq 2$ 时,备选模型受到最大的支持;当 $4 \leq \Delta AIC \leq 7$ 时,备选模型的支持力度减少;当 $\Delta AIC > 10$ 时,备选模型不再受到支持<sup>[20]</sup>。所以研究也通过此指标来比较和选择最优模型。

ECVI 为复核效度期望值,同 AIC 一样用于评价多个竞争模型的优劣,其值越小越好。竞争模型的评价和选择中采用多个指标是明智的,研究以绝对拟合指数和相对拟合指数均达到建议值,同时与其它模型比较有更小的 ECVI 和 AIC 的模型为最优模型。

表 2 模型拟合指数统计表

指标	$\chi^2$	$df$	$\chi^2/df$	NFI	NNFI	CFI	GFI	AGFI	RMSEA	ECVI	AIC	$\Delta AIC$
M1	209.947	35	5.998	0.812	0.790	0.837	0.916	0.868	0.096	0.459	249.947	162.064
M2	75.278	34	2.214	0.933	0.949	0.961	0.972	0.955	0.047	0.216	117.278	29.395
M3	51.401	25	2.056	0.954	0.956	0.975	0.982	0.960	0.040	0.205	111.401	23.518
M4	45.986	25	1.839	0.959	0.965	0.980	0.983	0.962	0.039	0.195	105.986	18.108
M5	23.214	16	1.451	0.979	0.981	0.993	0.992	0.971	0.029	0.186	101.214	13.331
M6	57.887	30	1.930	0.948	0.961	0.974	0.979	0.962	0.041	0.198	107.887	20.004
M7	72.420	30	2.414	0.935	0.941	0.960	0.974	0.952	0.051	0.225	122.420	34.537
M8	25.833	24	1.078	0.977	0.997	0.998	0.991	0.979	0.012	0.162	87.883	0

表 2 的结果总体表明,存在方法效应的模型比不存在方法效应的模型拟合数据要好,其中 M8 即存在两个相关的方法效应因子分别影响 5 条正/负向表述条目的模型拟合数据最好,各项拟合指数最好。M5 正负向表述均存在共同方法误差的 CTCU 模型拟合的也不错,多项拟合指标与 M8 相当,但是 ECV1 和  $\Delta$ AIC 均不支持 M5 (ECVI = 0.186 vs 0.162,  $\Delta$ AIC = 13.331 > 10)。

3.4 方法效应分析

表 3 呈现了 M5 和 M8 的条目的因素负荷。在 M5 中 5 条正向表述项目在整体自尊因子上的负荷均高于 5 条负向表述项目的负荷,但 10 条目均达到了显著水平 ( $p < 0.05$ ), 5 条正向表述项目间的误差相关均不显著 ( $p > 0.05$ ), 5 条负向表述项目间的误差相关均达到了显著水平 ( $p < 0.05$ ), 此结果说明正向表述的方法效应不显著,而负向表述的方法效应显著。M8 的因素负荷结果与 M5 的不同,整体来说,正向表述的条目在整体自尊因子上的负荷要比负向表述项目的因素负荷稍高。正向表述的条目在正方法效应因子上的负荷小于负向表述条目在负方法效应因子上的负荷,说明负方法效应比正方法效应明显,但正向表述的条目在正方法效应因子上的负荷也多达到显著水平 (P2 除外,  $p < 0.05$ )。

表 3 M5 和 M8 的条目因子负荷

	M5		M8	
	GSE	GSE	PME	NME
P1	0.757	0.548	0.259	0
P2	0.679	0.799	0.078	0
P3	0.611	0.335	0.477	0
P4	0.551	0.237	0.584	0
P5	0.564	0.437	0.289	0
N1	0.384	0.379	0	0.433
N2	0.327	0.342	0	0.284
N3	0.230	0.165	0	0.422
N4	0.397	0.313	0	0.736
N5	0.372	0.309	0	0.645

4 讨论

在编制测验项目时心理测量学家常常建议采用反向记分项目来平衡由于全部正向表述而引起的反应定势,但是实证研究却发现这种反向记分表述却会带来量表结构的不稳定,关于 RSES 的因子结构与方法效应的研究是这一情况的典型案例。

关联特质—相关方法 (CTCM) 模型和关联特质—相关独特性 (CTCU) 模型是解决方法效应最有效最常用的方法 [1,2,6]。研究共对 8 个相关模型进行了分析比较,结果表明:第一,在探索性因素分析中被广泛支持的 2 因子模型各拟合指标均达到了传统标准,但在与其他竞争模型的比较中则并未表现出优势。第二,存在方法效应的模型比不存在方

法效应的模型拟合要好,这一结果与 Marsh, Tomas & Oliver, Corwyn 的结果是一致的 [12,13,21],说明方法效应可能是导致模型拟合不佳的重要原因。第三,通过对 3 个 CTCM 模型和 3 个 CTCU 模型的比较发现:CTCU 模型与 CTCM 模型拟合结果差不多,但其中以同时存在正负向方法效应因子的模型拟合最好。M5 和 M8 分别是各模型系列中拟合最好的模型,但 M8 要比 M5 拟合的结果要稍好,这与 Bachman, Marsh, Quilty 等的研究结果有所不同 [10,12,14],在他们的研究中均发现 CTCU 模型要比 CTCM 模型拟合的好。但也有研究与此的结果一致,如 Corwyn (2000) [21] 的研究发现在青少年和成人中 M8 都是拟合数据最好的模型。另外,就 CTCU 模型与 CTCM 模型哪个更适合解决此类问题, Lance 等 (2002) [22] 在一篇比较两种方法的文章中指出:这两种方法都有其优缺点,但当 CTCM 能得到识别时是最优的方法。Conway 等 (2004) 通过模拟研究发现 CTCU 模型在某些情况下会得到有偏的参数估计,并建议优先考虑 CTCM 模型 [23]。

通过对 M5 和 M8 的方法效应分析不难发现,不论是独立的方法效应因子还是相关误差都会对条目在原因子上的负荷产生影响。正向表述的条目在整体自尊因子上的负荷比负向表述条目要高,而在各自的方法效应因子上的负荷则出现相反的结果,这与 Corwyn 的研究结果是一致的 [21],这说明负向表述比正向表述的方法学效应更强些。负向表述的测验项目更容易产生方法效应,究其原因可能是由于应答者的反应定势所致 [12]。

另外在模型比较方面,对于嵌套模型可以通过似然比来进行检验,但对于非嵌套模型,采用常用的几种拟合指数有时候很难区分出较有优势的模型,这里也遇到了此类问题,如 M5 和 M8 在常用的相对拟合指数和绝对拟合指数等指标上并没有明显的差异,有的指标 M5 要优于 M8 (如 NFI, GFI), 但通过 ECVI, AIC 和  $\Delta$ AIC 则可以有效的区分出较优的模型。因此,在模型比较中采用常用的拟合指数有时候并不能给出多少有用的信息,所以建议在模型比较和选择的研究中采用 ECVI, AIC 和  $\Delta$ AIC 等指标。

5 结论

5.1 中文 RSES 的两因子模型各拟合指标均达到了传统标准,但不如其它竞争模型。

5.2 假设存在方法效应的模型比假设不存在方法效应的模型拟合要好。

5.3 CTCM 模型和 CTCU 模型拟合的都很好,但 CTCM 模型拟合的更好。

5.4 造成中文 RSES 的两因子结构是由于项目表述方法或方法效应引起的,中文 RSES 实质上是个单因子结构的量表,在今后的研究中将其作为整体自尊评定工具是合适的。

参考文献

1 Podsakoff P M, MacKenzie S B, Lee J Y, et al. . Common method biases in behavioral research: A critical review of



- the literature and recommended remedies. *Journal of Applied Psychology*, 2003, 88: 879—903.
- 2 周浩, 龙立荣. 共同方法偏差的统计检验与控制方法. *心理科学进展*, 2004, 12(6): 942—950.
  - 3 Bagozzi R P, Yi Y. Multitrait - multimethod matrices in consumer research. *Journal of Consumer Research*, 1991, 17: 426—439.
  - 4 Greenberger E, Chuansheng C, Dmitrieva J, et al. . Item—wording and the dimensionality of the Rosenberg Self—Esteem Scale: do they matter? *Personality and Individual Differences*, 2003, 35: 1241—1254.
  - 5 Marsh H W. Confirmatory factor analyses of multitrait—multimethod data: Many problems and a few solutions. *Applied Psychological Measurement*, 1989, 13: 335—361.
  - 6 Bagozzi R P. Assessing construct validity in personality research: Applications to measures of self—esteem. *Journal of Research in Personality*, 1993, 27: 49—87.
  - 7 Schmitt D P, Jüri Allik. 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*, 2005, 89(4): 623—642.
  - 8 Rosenberg M. *Society and the adolescent self—image*. Princeton University Press, Princeton, NJ, 1965.
  - 9 杨烨, 王登峰. Rosenberg 自尊量表因素结构的再验证. *中国心理卫生杂志*, 2007, 21(9): 603—605.
  - 10 Bachman J, O'Malley P. Self—concepts, self—esteem, and educational experiences: the frog pond revisited (again). *Journal of Personality and Social Psychology*, 1986, 50: 35—46.
  - 11 Dunbar M, Ford G, Hunt K, et al. . Question wording effects in the assessment of global self—esteem. *European Journal of Psychological Assessment*, 2000, 16(1): 13—19.
  - 12 Marsh H W. Positive and negative self—esteem: substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 1996, 70: 810—819.
  - 13 Tomas J M, Oliver A. Rosenberg's Self—Esteem Scale: Two factors or method effects. *Structural Equation Modeling*, 1999, 6(1): 84—98.
  - 14 Quilty L C, Oakman J M, Risko E. Correlates of the Rosenberg Self—Esteem Scale method effects. *Structural Equation Modeling*, 2006, 13(1): 99—117.
  - 15 申自力, 蔡太生. Rosenberg 自尊量表中文版条目 8 的处理. *中国心理卫生杂志*, 2008, 22(9): 661—663.
  - 16 田录梅. Rosenberg(1965)自尊量表中文版的美中不足. *心理学探新*, 2006, 2: 88—91.
  - 17 季益富, 于欣. 自尊量表. 见: 汪向东, 王希林, 马弘. 编著. *心理卫生评定量表手册*. 中国心理卫生杂志, 1999 (增刊): 318—320.
  - 18 杨宜音, 张志学译. *性格与社会心理测量总揽*. 台北: 远流出版社, 1998.
  - 19 侯杰泰, 温忠麟, 成子娟. *结构方程模型及其应用*. 北京: 教育科学出版社, 2004.
  - 20 Burnham K P, Anderson D R. *Model selection and multimodel inference: A practical information—theoretic approach 2<sup>nd</sup>*. New York: Springer, 2002.
  - 21 Corwyn R F. The factor structure of global self—esteem among adolescents and adults. *Journal of Research in Personality*, 2000, 34: 357—379.
  - 22 Lance C E, Noble C L, Scullen S E. A Critique of the correlated trait—correlated method and correlated uniqueness models for multitrait—multimethod data. *Psychological Methods*, 2002, 7(2): 228—244.
  - 23 Conway J M, Lievens F, Scullen S E, et al. . Bias in the correlated uniqueness model for MTMM Data. *Structural Equation Modeling*, 2004, 11(4): 535—559.

## The Factor Structure of Chinese Rosenberg' Self—esteem Scale Affected by Item Statement Method

Wang Mengcheng<sup>1</sup>, Cai Bingguang<sup>2</sup>, Wu Yan<sup>3</sup>, Dai Xiaoyang<sup>1</sup>

(1. Department of Psychology, Shenzhen University, Shenzhen 518060; 2. Lintou middle school, Maoming 525424;

3. Center for Studies of Psychological Application, South China Normal University, Guangzhou 510631)

**Abstract:** To explore the factor structure of Chinese Rosenberg' self—esteem affected by item statement method. The study compare “The Correlated Trait—Correlated Method, CTCM” with “The Correlated Trait—Correlated Uniqueness, “CTCU” models in the framework of MTMM. The results indicated that the two factor structure of Chinese Rosenberg' self—esteem scale mainly caused by method effects. More attentions should be paid to method effects when establish a psychological scale.

**Key words:** Chinese Rosenberg' self—esteem scale; item statement method; factor structure; method effects