

应用 Rasch 模型分析 Rosenberg 自尊量表

高爽 张向葵*

(东北师范大学心理学院, 长春 130024)

摘要:采用 Rosenberg 自尊量表(RSES)对 425 名在校大学生进行施测,应用项目反应理论的 Rasch 模型对项目指标进行分析及 DIF 检验。结果表明,Rosenberg 自尊量表具有单维性,量表的信度为 0.84;除项目 8 以外,其他项目拟合指标良好,较适用来区分中等及偏低自尊水平的个体,项目功能差异检验发现在项目 1 和项目 5 上存在 DIF,表现为男生自尊水平要高于女生。相对于经典测量理论,应用 Rasch 模型分析 Rosenberg 自尊量表具有优势,为进一步的完善和使用该自尊量表提供依据。

关键词:自尊;Rasch 模型;项目反应理论;DIF 检验

中图分类号:B841.2

文献标识码:A

文章编号:1003-5184(2018)05-0445-06

1 引言

自尊,是个体生活中的基本需要,是个体人格的核心因素之一。自尊定义强调自尊应该从自我概念(Self-concept)结构中分离出来,自尊是在表达情感,或者评估自我概念结构中,强调人们如何感受自我(Leary & Baumeister, 2000)。自尊不仅影响着人们的日常行为,如低自尊被证实与焦虑、抑郁、自杀、酗酒和暴力等消极行为相联系;还影响人们的情绪及生活满意度,与个体的心理健康密切相关(Baumeister, Campbell, Krueger, & Vohs, 2003; Baumeister, 2010; Sowislo & Orth, 2013; 高爽, 张向葵, 徐晓林, 2015)。对于自尊的研究,一直以来就受到研究者的重视,而有关自尊的测量亦是研究者关注的问题之一,使用何种测量工具能够测量出个体更真实的自尊水平,何种测量工具具有更好的测量学意义,这些都会影响研究者在自尊领域研究中的准确性与客观性。有关自尊的测量,目前使用最广泛的工具是由 Rosenberg 于 1965 年编制的自尊量表(简称 RSES),具有广泛性和代表性(Schmitt & Allik, 2005)。Rosenberg 提出该量表所测量的项目为单一维度,目的在于测量个体的整体自尊(Rosenberg, 1995)。以往的一些研究对 Rosenberg 自尊量表究竟是单因素的模型还是两因素的模型进行了验证并提出了自己的见解,如 Roth 等(2008)认为 RSES 之所以被抽取两个维度,完全是人为造成的——试题陈述方式的差异,但多数研究已证明

RSES 为测量单一维度(整体自尊)的有效工具(Chao, Vidacovich, & Green, 2016; Quilty, Oakman, & Risko, 2006)。

项目反应理论(Item Response Theory, 简称 IRT)的发展得益于传统经典测量理论(Classical Test Theory, 简称 CTT)的不足与局限,尤其体现在以下两方面:一是项目的参数指标具有样本依赖性,估计值依赖于样本,并随着样本能力不同而有所变化;二是 CTT 的前提假设是所有样本具有相同的测量误差,这会忽视样本在能力上的个体差异,导致估计不精确(漆书青,戴海崎,丁树良, 2002)。相对于经典测量理论,项目反应理论能有效地评定被试的特质水平(杨业兵等, 2008)。鉴于此,项目反应理论(IRT)在克服经典测量理论(CTT)不足的基础上发展出新的测量方法,并广泛应用在人格、教育及能力等测量中。Rasch 模型(Rasch Model)以数据与模型的拟合为前提,对个体能力值和项目难度进行对数转换,建立一个等距的刻度衡量测试题的难度和个体能力,通过这种方法可以克服传统测量方法中对样本和测验的依赖(Wright, 2000; 刘昊, 刘肖岑, 冯晓霞, 2013)。Rasch 模型是目前 IRT 领域中最简化的模型(reduced model),需要估计的参数最少,因而参数估计稳定性及精度往往比复杂的模型(如专家提到的用于人格测验的 GCUM 模型等)更高(Al-Owaidha, 2007; 晏子, 2010),因此为了得到更精确的参数估计结果选用 Rasch 模型的重要考量之一。

* 通讯作者:张向葵, E-mail: zhangxk@nenu.edu.cn。

此外,与其它 IRT 模型相比(如 2PLM, 3PLM 或 GGUM), Rasch 模型有其自身独特的优势,即项目难度参数 b 是等距量表(interval scale)。近年来,已有越来越多研究者将 Rasch 模型应用在人格领域。巫博瀚等(2013)采用 Rasch 模型针分析 Rosenberg 自尊量表发现该量表较适用于中等自尊水平的被试。Chao(2016)采用 Rasch 模型对非裔美国大学生的自尊测量进行分析发现,该量表具有单维性并适用于自尊水平较高个体的测量。目前,国内多数研究采用经典测量理论的方式对 Rosenberg 自尊量表进行分析与报告。田录梅(2006)研究发现, Rosenberg 自尊量表具有较高的信效度,但被试对项目 8 的理解与西方存在差异,建议将此改为正向计分或删除。鉴于前述 IRT 所弥补经典测量理论在测量上的不足与局限性,考虑到 Rasch 模型的特点以及在 Likert 量表应用上的多种优势,采用 Rasch 模型对 RSES 进行分析,探讨该量表的项目特征和适用个体等。此外,由于该量表对不同文化条件下的群体的适用性问题,已有研究探讨中美不同文化背景下自尊量表上的差异功能检验(Song, Cai, Brown, & Grimm, 2011),但不同性别类型被试在 Rosenberg 自尊量表上的差异功能检验的相关研究相对较少,随着公平教育、性别平等教育等理念的大力提倡,对性别差异的研究就变得更加关键。

综上,研究旨在通过项目反应理论对 Rosenberg 自尊量表进行项目参数估计以及差异项目功能分析,并应用拟合指标查找异常的项目或个体,以期为进一步完善和使用该量表提供依据。

2 研究方法

2.1 研究对象

样本选取自吉林省一所师范类大学和一所综合性大学学生进行测量,共 440 人进行施测,回收问卷 435 份,问卷回收率为 98.9%;其中有效问卷 425 份,问卷有效率为 97.7%。被试的平均年龄为 21.76 ± 2.31 岁,男生 191 人,占 44.9%;女生 234 人,占 55.1%。

2.2 研究工具

Rosenberg 自尊量表由 Rosenberg 于 1965 年编制而成,是个体对自己整体自尊的自我报告测量工具,是目前自尊研究领域中使用最广泛的工具。该量表包含 10 个项目,其中有 5 个项目为正向计分题,如项目 7:整体而言,我对自己感到很满意;另外 5 个项目为反向计分题,如项目 10:我有时认为自己

一无是处。该量表为 Likert 四点评分,采用 1-4 评分方式,理论分数范围是 10~40 分,得分愈高,表明个体的自尊水平越高。

2.3 统计方法

Rasch 模型是一种单参数的项目反应模型,在 Likert 量表分析中具有客观等距的优势(刘昊等, 2013)。考虑到 Rosenberg 自尊量表为 Likert 式单维度量表,并在数据分析中通过单维检验亦可证实,采用 Rasch 模型对 Rosenberg 自尊量表进行分析,以此评估项目单维性、信度、难度及模型拟合度,包括各项项目的信息量。进一步进行项目功能差异检验(Differential Item Functioning, 简称 DIF),以探讨自尊量表每个项目在性别上是否存在 DIF。

采用 SPSS21.0 进行数据的前期整理及单维性检验,应用 Conquest2.0 软件进行模型拟合的数据处理,采用 R 3.0.1 软件中的“Lordif”程序包进行 DIF 分析。

3 研究结果

3.1 单维性检验

采用 Rasch 模型对量表进行编制和修订时有一个前提:量表具有单维性(Roth, Decker, Herzberg, Brähler, 2008)。若第一因子特征根与第二因子特征根的比值接近或大于 3,则说明该量表具有单维性(Hambleton & Swamniathan, 1985)。对 Rosenberg 自尊量表的数据进行主成分分析发现(见表 1),第一因子的特征根与第二个因子特征根的比值为 3.426,说明数据基本满足单维性的要求,因此适合 Rasch 模型分析。

表 1 单维性检验

| 因子 | 第一因子特征根 | 第二因子特征根 | 两特征根比值 |
|----------------|---------|---------|--------|
| Rosenberg 自尊量表 | 4.294 | 1.253 | 3.426 |

3.2 信度

Rasch 模型计算个体信度(Person separation reliability)是指由个体所产生“真实”变异与总变异的比,目的在于考察受试者在项目评定上的可靠性程度(刘昊等, 2013)。在整个测验水平上定义针对评价单个被试的信度概念,就是测验信息函数,测验信息函数就是测验所含项目的信息函数的累加(Bond & Fox, 2015)。Rasch 模型测量的整体信度通过个体层面的解释率计算获得,数值范围在 0~1 之间。一般而言,信度指标在 0.7 以上是可接受的,高于 0.8 为较好(Bond & Fox, 2006)。经计算获得,本测试的信度值为 0.84。

3.3 难度

从表 2 中的项目估计值发现,各项目的难度分布在 -1.006 与 1.497 之间(平均难度设定为 0),图 1 表示被试的自尊水平与项目难度的对应关系。结合表 1 难度估计值,由图 1 可知,Rosenberg 自尊量表的项目对于中等及偏低水平自尊的被试提供的信息量最大,但不适用于用来评定自尊水平较高的被试。

3.4 模型拟合度

在 Rasch 模型中,加权残差均方(Infit MNSQ)和残差均方(Outfit MNSQ)常用来评价项目的 χ^2 拟合指标,其中 Infit MNSQ 则是加权(以方差为加权系数)后的残差均方,Outfit MNSQ 是残差的均方。通过 Rasch 模型分析发现,Rosenberg 自尊量表中的第 8 题的 Infit MNSQ 及 Outfit MNSQ 值分别为 1.51 和 1.55,根据 Wright 和 Linacre(1994)的建议:凡是 MNSQ > 1.4 或 < 0.6 时,即可以认为该题的拟合效果较差。可见,除第 8 题外,其它项目的拟合效果均良好,具体见表 2。

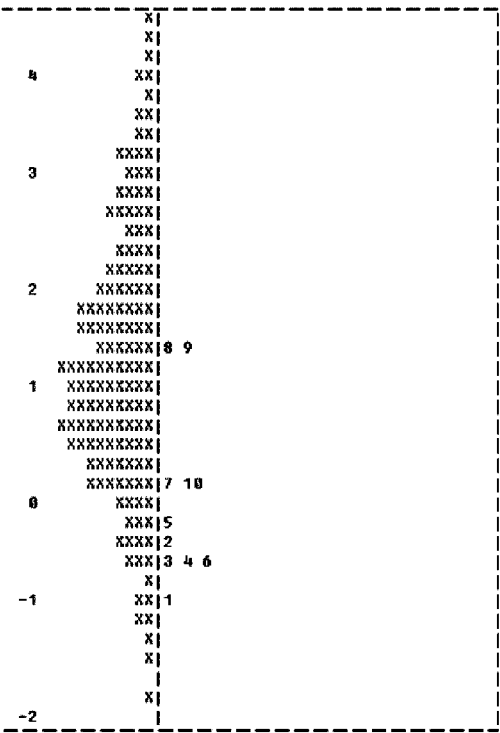


图 1 项目难度与自尊特质分布图

表 2 模型拟合度参数

| 项目 | 难度估计值 | 标准误 | Infit | | | Outfit | | | T | |
|----|---------|-------|-------|------|------|--------|------|------|------|------|
| | | | MNSQ | CI | T | MNSQ | CI | | | |
| 1 | -1.006 | 0.064 | 0.85 | 0.87 | 1.13 | -2.4 | 0.84 | 0.87 | 1.13 | -2.4 |
| 2 | -0.405 | 0.064 | 0.90 | 0.86 | 1.14 | -1.4 | 0.91 | 0.87 | 1.13 | -1.4 |
| 3 | -0.531 | 0.064 | 0.85 | 0.86 | 1.14 | -2.2 | 0.86 | 0.87 | 1.13 | -2.1 |
| 4 | -0.616 | 0.064 | 0.82 | 0.86 | 1.14 | -2.7 | 0.86 | 0.87 | 1.13 | -2.2 |
| 5 | -0.272 | 0.063 | 1.05 | 0.86 | 1.14 | 0.8 | 1.05 | 0.87 | 1.13 | 0.7 |
| 6 | -0.472 | 0.064 | 0.62 | 0.86 | 1.14 | -6.3 | 0.62 | 0.87 | 1.13 | -6.4 |
| 7 | 0.196 | 0.063 | 0.82 | 0.86 | 1.14 | -2.7 | 0.82 | 0.87 | 1.13 | -2.7 |
| 8 | 1.357 | 0.061 | 1.51 | 0.87 | 1.13 | 6.6 | 1.55 | 0.87 | 1.13 | 6.9 |
| 9 | 1.497 | 0.061 | 1.04 | 0.87 | 1.13 | 0.7 | 1.10 | 0.87 | 1.13 | 1.4 |
| 10 | 0.251 * | 0.189 | 1.37 | 0.86 | 1.14 | 4.8 | 1.38 | 0.87 | 1.13 | 4.9 |

此外,图 2 表示自尊量表所有项目组成的总测验信息曲线,该曲线反映的是自尊量表所有项目作为一个整体对具有不同特质水平的被试所提供精确评估的程度。其中横坐标代表项目的难度,对应着被试的特质水平,每一个刻度代表一个 logit 单位,纵坐标表示信息量大小,即 Fisher 信息函数(Bond & Fox,2015)。整体来说,自尊估计值范围在 0 ~ -2 之间,所能提供的测量精确性最高,对于自尊水平中等及偏低的被试所提供的信息量最大。

3.5 性别的项目功能差异检验

为确保测验的公平性,进一步进行项目功能差异检验,项目功能差异检验是指来自两个不同团体但心理特质水平相等的人选择某项目同一选项的概

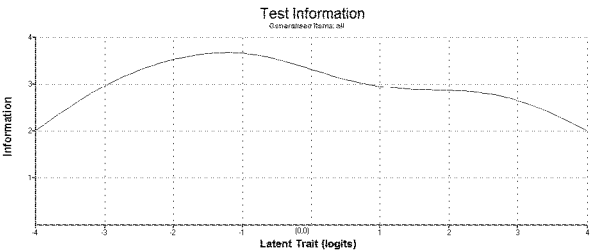


图 2 测验信息曲线

率不等(涂冬波,戴海琦,2007),即项目对于不同团体的人具有不同的测量功能。通过 Lord 卡方检验法计算项目参数差异的大小来判断该题目是否具有项目功能差异,采用 R 软件中程序包“Lordif”进行 DIF 分析,对不同性别的被试在各项目上的差异进

行项目功能差异检验 (Choi, Gibbons, & Crane, 2011), χ^2_{13} 为总 DIF 检验指标, 小于 0.05 为存在 DIF, 其中项目 10 为参照题, 一般不予考虑 (Choi, Gibbons, & Crane, 2011)。根据结果, 由表 3 可知,

表 3 项目功能差异分析检验

| 题目 | no. cat | χ^2_{12} | χ^2_{23} | χ^2_{13} |
|---------------------------|---------|---------------|---------------|---------------|
| 1 我认为自己是个有价值的人,至少与别人不相上下。 | 3 | 0.009 | 0.457 | 0.025 |
| 2 我觉得我有许多优点。 | 3 | 0.697 | 0.021 | 0.064 |
| 3 总的来说,我倾向于认为自己是一个失败者。 | 3 | 0.020 | 0.919 | 0.066 |
| 4 我做事可以做的和大多数人一样好。 | 3 | 0.112 | 0.876 | 0.280 |
| 5 我觉得自己没有什么值得自豪的地方。 | 3 | 0.010 | 0.046 | 0.005 |
| 6 我对自己持有一种肯定的态度。 | 3 | 0.664 | 0.968 | 0.909 |
| 7 整体而言,我对自己感到满意。 | 3 | 0.083 | 0.475 | 0.172 |
| 8 我要是能更看得起自己就好了。 | 4 | 0.616 | 0.495 | 0.698 |
| 9 有时候我的确感到自己很没用。 | 4 | 0.045 | 0.842 | 0.132 |
| 10 我有时认为自己一无是处。 | 4 | 0.244 | 0.028 | 0.046 |

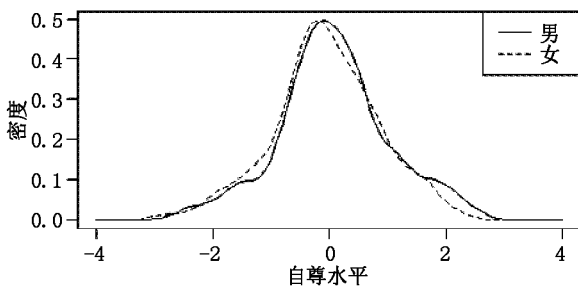


图 3 自尊水平 θ 值的分布

4 讨论

4.1 针对拟合指标不佳的项目的分析

采用 Rasch 模型对 Rosenberg 自尊量表的特性进行分析,由模型拟合度的指标发现,项目 8(我要是能更看得起自己就好了)的 Infit MNSQ 及 Outfit MNSQ 值分别为 1.51 和 1.55,拟合度较差,该项目为量表中的反向计分题,但在陈述措辞方式上同时包含了正向(更看得起自己)与反向(要是能)两层含义,这可能导致被试误解该项目的原意。国内学者田录梅(2006)在对 Rosenberg 自尊量表的优点与不足进行论述,尤其针对项目 8 进行鉴别与处理,在意义理解上由于我国文化注重谦虚的态度,因此会将项目 8 理解为正向,而给西方认为的消极负向,并进一步发现,将项目 8 若按正向题计分,其鉴别力会达到显著水平。巫博瀚等人(2013)在对 Rosenberg 自尊量表进行修订时发现项目 8 拟合度不佳,在删除项目 8 后,被试在其余 9 个项目上均具有较好的模型拟合。造成这种差异的原因很大程度上是由于文化背景的不同,Lehman 等人认为文化是某一特定群体共享,并不同于其他群体的一系列行为准则与认知 (Lehman, Chiu, & Schaller, 2004)。由此可知,

项目 1 和 5 存在 DIF,即在这两个项目上,性别差异导致自尊水平不同。不同性别的被试在自尊水平上的分布见图 3。由图 3 发现,不同性别被试的自尊水平参数 θ 值均分布在 $[-4, 4]$ 。

在不同的文化背景下,个体的自尊水平既具有相似的成分,又表现出个体差异(蔡华俭,丰怡,岳曦彤,2011)。蔡华俭等(2007)在一项对中西方自尊进行比较的研究中发现,中国人在自尊的情感成分上与西方相似,在自尊的总体和认知成分上要低于西方水平。相对而言,西方文化背景更为强调个人主义,而东方文化强调集体主义,因此会将谦虚内敛作为积极品质,从而导致在自尊量表项目 8 上表现出差异。据此,以后的研究可考虑是否将项目 8 进行正向修改或调整,以获得更客观的自尊测量结果。

4.2 关于 Rosenberg 自尊量表适合人群的探讨

对 Rosenberg 自尊量表的测验信息曲线进行分析发现,该量表对自尊中等及偏低特质水平被试能提供较为精准的评估。Halama 和 Biesiad (2006)的研究发现,Rosenberg 自尊量表对中等及偏低水平的被试提供较多的信息量,对自尊水平较高的被试提供的信息量较少,结果与此相一致。一般而言,高自尊个体具有更好的心理适应性和社会适应性,但有关高自尊者个体之间存在着质的差异,包括防御性自尊、不稳定的自尊、自恋等(田录梅,张向葵,2006)。通过项目难度及信息曲线图发现,Rosenberg 自尊量表对自尊水平较高的个体评估相对不够适合,这可能是由于高自尊个体的异质性所导致,个体无法将其他成分与自尊水平相分离。Deci 和 Ryan (1995)将自尊区分为相倚性高自尊和真正的高自尊,相倚性自尊是指对自己的感受来自或取决于符合某些优秀标准或不辜负某些人际或内心的期望。由此可见,在某种程度上,通过测量获得的自尊分数具有不同的可能性,如自我欺骗、印象管理及防御性地保持一种高自尊的感觉等,倾向于对自己有

利的评价,这也从另一角度解释自尊异质对测量结果造成偏差影响的原因。未来研究宜设计一些更为客观反映个体高自尊水平的测量项目,更好地完善测量的精确性。

4.3 关注 Rosenberg 自尊量表在性别上的 DIF 检验

项目功能差异检验是一种考察测验项目对于来自不同团体的被试是否具有相同功能的分析方法,通过观察项目特征曲线来解释同一项目在不同被试组之间是否存在差异(Wang, 2004; Crane & Belle, 2006; Choi, Gibbons, & Crane, 2011)。针对被试在自尊项目上的 DIF 检验发现,项目 1(我认为自己是个有价值的人,至少与别人不相上下)和项目 5(我觉得自己没有什么值得自豪的地方)具有 DIF,倾向于对男生有利,这说明对于总分相同的被试而言,男生倾向在这两个项目上有较高的得分。造成这种差异的原因可能是由于相对于女生,男生较不容易受外在评价的影响,将自尊水平与能力相结合,较少受社会关系影响。此外,由 DIF 检验的 θ 值分布可知,男生自尊水平高于女生。Kling 等(1999)关于自尊性别差异的元分析结果发现,男性自尊水平要高于女性。伍秋萍和蔡华俭(2006)的一项元分析也发现同样结果,男性自尊略高于女性,DIF 检验结果与此相一致。社会文化因素对性别差异造成一定的影响,而个体在自尊水平上的性别差异与其社会性习得性别角色和刻板印象密切相关(Wood & Eagly, 2002)。鉴于此,建议以后在使用该量表的同时,注意考虑项目 1 与项目 5 在性别上造成的差异,以保证测量的客观性。

5 结论

采用 Rasch 模型分析 Rosenberg 自尊量表,结果发现:项目 8(我要是能更看得起自己就好了)的表述内容有必要进行调整;在被试群体上,考虑到个体高自尊的异质性,需要进一步对高自尊水平个体的测量进行甄别与开发;考虑性别差异时,项目 1(我认为自己是个有价值的人,至少与别人不相上下)和项目 5(我觉得自己没有什么值得自豪的地方)需要调整,以期个体自尊水平的测量提供更为精准和客观的估计。

参考文献

- 蔡华俭,丰怡,岳曦彤.(2011). 泛文化的自尊需要:基于中国人的研究证据. *心理科学进展*, 19(1), 1-8.
- 高爽,张向葵,徐晓林.(2015). 大学生自尊与心理健康的元分析——以中国大学生为样本. *心理科学进展*, 23(9), 1499-1507.
- 刘昊,刘肖岑,冯晓霞.(2013). 应用 rasch 模型测试和分析

- 儿童入学准备状态. *心理科学*, 36(2), 484-488.
- 漆书青,戴海崎,丁树良.(1998). *现代教育与心理测量学原理*. 南昌:江西教育出版社.
- 田录梅.(2006). Rosenberg(1965)自尊量表中文版的美中不足. *心理学探新*, 26(2), 88-91.
- 田录梅,张向葵.(2006). 高自尊的异质性研究述评. *心理学进展*, 14(5), 704-709.
- 涂冬波,戴海琦.(2007). 项目反应理论下 likert 型量表的 dif 检测方法初探. *江西师范大学学报(自然科学版)*, 31(3), 311-315.
- 巫博瀚,赖英娟,施庆麟.(2013). “Rosenberg 自尊量表”之试题衡鑑:评等量尺模型的应用. *测验学刊*, 60(2), 263-290.
- 伍秋萍,蔡华俭.(2006). 对中国人自尊性别差异的元分析. *中国社会心理学会 2006 年学术研讨会论文集*.
- 晏子.(2010). 心理科学领域内的客观测量——rasch 模型之特点及发展趋势. *心理科学进展*, 18(8), 1298-1305.
- 杨业兵,苗丹民,田建全,肖利军,孙茜,洪霞.(2008). 应用项目反应理论对《中国士兵人格问卷》的项目分析. *心理学报*, 40(5), 611-617.
- Al - Owidha, A. A. (2007). *A comparison of the Rasch model and the three - parameter logistic model applied to the quantitative subtest of the General Aptitude Test, Saudi Arabia* (Unpublished Doctoral dissertation). University of Denver, Colorado, USA.
- Baumeister, R. F. (2010). The self. In R. F. Baumeister & E. J. Finkel (Eds.), *Advanced social psychology: The state of science* (pp. 139-175). New York: Oxford University Press.
- Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self - esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, 4(1), 1-44.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York, NY: Routledge.
- Cai, H. J., Brown, J. D., Deng, C., & Oakes, M. A. (2007). Self - esteem and culture: Differences in cognitive self - evaluations or affective self - regard? *Asian Journal of Social Psychology*, 10(3), 162-170.
- Chao, C. L., Vidacovich, C., & Green, K. E. (2016). Rasch analysis of the rosenberg self - esteem scale with african americans. *Psychological Assessment*, 29(3), 329-342.
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif: An r package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and monte carlo simulations. *Journal of Statistical Software*, 39(8), 1-30.
- Crane, P. K., & Belle, G. V. (2006). Differential item functioning analysis with ordinal logistic regression techniques. dif - detect and difwithpar. *Medical Care*, 44(44), 115-123.
- Deci, E. L., & Ryan, R. M. (1995). Human autonomy: The basis

- for true self – esteem. In M. Kemis (Ed.), *Efficacy, Agency, & Self – Esteem* (pp. 31 – 49). Plenum Press.
- Gray – Little, B. , Williams, V. S. L. , & Hancock, T. D. (1997). An item response theory analysis of the rosenberg self – esteem scale. *Personality & Social Psychology Bulletin*, 23(5), 443 – 451.
- Hambleton, R. K. , & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer – Nijhoff.
- Kling, K. C. , Hyde, J. S. , Showers, C. J. , & Buswell, B. N. (1999). Gender differences in self – esteem: A meta – analysis. *Psychological Bulletin*, 125(4), 470 – 500.
- Leary, M. R. , & Baumeister, R. F. (2000). The nature and function of self – esteem: Sociometer theory. *Advances in Experimental Social Psychology*, 32(00), 1 – 62.
- Lehman, D. R. , Chiu, C. Y. , & Schaller, M. (2004). Psychology and culture. *Annual Review of Psychology*, 55, 689 – 714.
- Quilty, L. C. , Oakman, J. M. , & Risko, E. (2006). Correlates of the rosenberg self – esteem scale method effects. *Structural Equation Modeling A Multidisciplinary Journal*, 13(1), 99 – 117.
- Rosenberg, M. , Schooler, C. , Schoenbach, C. , & Rosenberg, F. (1995). Global self – esteem and specific self – esteem: Different concepts, different outcomes. *American Sociological Review*, 60(1), 141 – 156.
- Roth, M. , Decker, O. , Herzberg, P. Y. , & Brähler, E. (2008). Dimensionality and norms of the rosenberg self – esteem scale in a german general population sample. *European Journal of Psychological Assessment*, 24(3), 190 – 197.
- 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 & Social Psychology*, 89(4), 623 – 42.
- Song, H. , Cai, H. , Brown, J. D. , & Grimm, K. J. (2011). Differential item functioning of the rosenberg self – esteem scale in the us and china: Measurement bias matters. *Asian Journal of Social Psychology*, 14(3), 176 – 188.
- Sowislo, J. F. , & Orth, U. (2013). Does low self – esteem predict depression and anxiety? A meta – analysis of longitudinal studies. *Psychological Bulletin*, 139(1), 213 – 240.
- Tomas, J. M. , & Oliver, A. (1999). Rosenberg’s Self – Esteem Scale: Two factors or method effects. *Structural Equation Modeling*, 6(1), 84 – 98.
- Wang, W. C. (2004). Effects of anchor item methods on the detection of differential item functioning within the family of rasch models. *Journal of Experimental Education*, 72(3), 221 – 261.
- Wood, W. , & Eagly, A. H. (2002). A cross – cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, 128(5), 699 – 727.
- Wright, B. D. , & Linaere, J. M. (1994). Reasonable mean – square fit value. *Rasch Measurement Transactions*, 8, 370.
- Wright, B. D. , & Mok, M. M. C. (2000). Rasch models overview. *Journal of Applied Measurement*, 1(1), 83 – 106.

Use Rasch Model to Analyze the Rosenberg Self – Esteem Scale

Gao Shuang Zhang Xiangkui

(Department of Psychology, Northeast Normal University, Changchun 130024)

Abstract: The study aimed to use the Rasch Model in analyzing the Rosenberg Self – Esteem Scale (RSES) to explain its model fitness of each item, explored the subjects individual differences of DIF, and it provided a basis for effectiveness and usage in RSES. 425 undergraduates (mean age = 21.71 years old, 191 males) were sampled from Northeast China, they were measured by the Rosenberg Self – Esteem Scale (RSES) which containing 10 items. The data was collected by SPSS and analyzed through the software Conquest2.0 and R package. The results showed the Rosenberg Self – Esteem Scale’s reliability was .84, its item difficulty distribution was from –1.006 to 1.497, more items difficulty levels indicated that the test is relatively easy for participants, the RSES can discriminated the examines with moderate and lower self – esteem rather than the upper. All except item 8 have acceptable model fit indices, the mean square of both Infit and Outfit MNSQ of item 8 in RSES were 1.51 and 1.55. Differential item functioning in gender was item 1 and 5, it showed that the male had higher self – esteem than female. All of these findings may provide implications of using Rasch Model for the Rosenberg Self – esteem Scale development and evaluation, it also support discussions and suggestions for future research.

Key words: self – esteem; Rasch model; item response theory; differential item functioning