

# 社会称许性反应的测量与控制\*

潘逸沁<sup>1</sup> 骆方<sup>1,2</sup>

(<sup>1</sup>北京师范大学心理学部)(<sup>2</sup>未来教育高精尖创新中心, 北京 100875)

**摘要** 社会称许性反应是影响人格测验效度最常见的一种反应偏差, 从 20 世纪 30 年代至今, 测量学家一直努力揭示社会称许性反应的性质和内部结构, 并开发社会称许性量表对其直接测量。本文综述了社会称许性量表的发展脉络, 指出了至今遗留的问题。并且, 从事前控制和事后控制的角度, 对当前主要的控制社会称许性反应的方法进行了论述, 指出了社会称许性量表的效度对事后控制方法的重要影响, 并给出了一些如何控制社会称许性反应的指导建议。

**关键词** 社会称许性反应; 自我欺骗; 印象管理

**分类号** B841

20 世纪 30 年代, 在人格测验相继开发和流行的同时, 人们已经发现被试在完成人格问卷时, 倾向于按照社会公认的标准选择答案, 并不关心这个答案是否符合自己的真实行为。心理测量学家认为这是一种反应偏差(Response Bias)——为了避免社会压力而夸大反应的倾向(Kelly, Miles, & Terman, 1936)。到了 60 年代相关的探讨有了突破, 其中以 Edwards 的研究最受瞩目。他对明尼苏达测验多项人格问卷(Minnesota Multiphasic Personality Inventory, MMPI)正常人群的测验结果进行探索性因素分析, 发现所有项目在第一个因子(没有进行因子旋转)上都有较大的因子载荷, 而且载荷的大小与项目受社会称许的程度高度相关。于是 Edwards 正式提出“社会称许性反应(Socially Desirable Response, SDR)”的概念, 并对其界定了操作性定义, 即“被试对人格测验项目作出符合社会期望的作答的反应趋势”(Edwards & Walker, 1961)。

## 1 社会称许性反应的控制

由于社会称许性反应会污染测验结果, 大大降低测验分数的真实性(Krumpal, 2013; Perinelli

& Gremigni, 2016), 所以, 测量学家在 1939 年已经开始研究如何控制社会称许性反应。Nederhof (1985)把控制社会称许性反应的方法分为事前控制和事后控制两类。

### 1.1 事前控制

事前控制是指一些研究者试图在编制测量工具时或者在施测过程中通过一些方法对社会称许性反应进行控制, 来阻止受测者的歪曲反应发生。一种常见的事前控制方法是编制迫选测验, 比如在 NRC's Committee on Selection and Training of Aircraft Pilots 的年鉴报告上, Rethlingshafer 认为“编制一个迫选测验, 可以排除被试认出‘好’答案的可能性, 使该测验每个题目的各选项具有相同的表面效度”。

迫选测验要求被试在两个或多个社会称许性程度相等的选项之间进行选择。例如, 一个人格测验的题组要求被试在下列两项中选择符合自己的一项: A. 周边的人都很喜欢我; B. 我做事情很认真。这两个选项的社会称许性程度相同, 但是描述的是不同的人格特质。如果被试选择“A”, 则宜人性维度计 1 分, 责任感维度计 0 分; 如果被试选择“B”, 则宜人性维度计 0 分, 责任感维度计 1 分。研究者由此控制社会社会称许性反应, 更为准确地测量宜人性和责任感。但是, 无论被试怎样作答, 每个题组的满分都是 1 分, 整个测验的得分对所有被试而言也是个常数(王珊, 骆方, 刘

收稿日期: 2017-02-20

\* 北京市教育科学“十三五”规划项目(CADA17078)。

通讯作者: 骆方, E-mail: luof@bnu.edu.cn

红云, 2014)。这种所有被试的测验总分都相同的分数被称为自模式数据。迫选测验提供了一种使得人格测验的虚假反应降到最小化的可能, 但是自模式数据不能进行方差分析, 因素分析等处理, 而且不能进行个体之间的比较, 因而测量学家并不推荐使用迫选量表。

近几年, 一些测量学家基于被试对迫选测验的作答过程建构了一些项目反应理论 (Item Response Theory, IRT) 模型, 比如, Brown 等人提出了 Thurstone IRT Model。这些精细的模型摒弃了传统的计分方式, 可以估计被试潜在的人格特质水平, 从而规避了自模式数据带来的一系列问题 (Brown, 2016; Brown & Maydeu-Olivares, 2011, 2013; Guenole, Brown, & Cooper, 2016; Hontangas et al., 2015; Wang, 2015)。但是, 这些模型必须有大量数据做支持才能对被试的特质水平准确估计, 而且这些模型的适用性还有待进一步的检验。

## 1.2 事后控制

事后控制是指作假发生后, 采用一定的技术手段, 对作假进行识别和控制。常见的事后控制方法是把社会称许性反应视为一个有意义的心理结构, 编制有效的量表对其直接测量, 然后再采用校正或识别的方法, 从人格测验分数中去除社会称许性反应的影响。

### 1.2.1 校正技术

在校正技术中, 研究者首先使用社会称许性量表测量个体的作假动机, 获得作假变量; 然后把社会称许性视为抑制变量处理, 校正人格分数上的夸大分数, 获得近似真实分数。

最为常见的校正方法是偏控制法。偏控制法是对作假变量对人格分数的回归, 残差即近似真实分数。例如, 在 Whitaker, Long, Petróczi 和 Backhouse (2014) 的研究中, 4.9% 的被试在印象管理量表 (社会称许性量表的一个子量表) 上的得分过高, Whitaker 等人利用分层回归控制了社会称许性反应的影响, 而不是删除这些被试。Miller 和 Gallagher (2016) 也将被试在印象管理量表上的得分作为一个变量放置于分层回归模型中。

Sjöberg (2015) 利用实验研究证明了这一方法的有效性。研究中存在两组被试, 第一组被试的作答中包含较高水平的社会称许性反应, 而另一组被试的作答是匿名的, 因此社会称许性反应水平较低。Sjöberg 基于社会称许性量表的得分, 利

用回归模型移除了人格问卷得分中社会称许性反应的影响。通过比较移除影响前两组被试问卷得分的差异和移除影响后两组被试问卷得分的差异, Sjöberg 发现社会称许性量表可以移除 90% 社会称许性反应的影响。

然而, 也有很多研究探讨了这种方法的局限性。Ellingson, Sackett 和 Hough (1999) 发现基于偏控制法得到的校正分数, 不能够代表个体在作假情境下的真实分数; Barrick 和 Mount (1996) 研究结果发现, 偏控制的方法导致了预测效度的轻微缩减, 说明校正移走了一些特质成分。这可能是因为社会称许性反应与量表项目, 测量环境以及被试个人动机等因素存在交互作用, 偏控制法没有考虑到这些因素的影响 (Leite & Cooper, 2010; Ziegler, MacCann, & Roberts, 2012)。此外, 作假后的测验分数是作假效应与真实分数的叠加, 作假动机相同的个体, 如果其原有水平较低, 那么测验分数升高的幅度较大; 如果其原有水平较高, 那么测验分数升高的幅度就较小。个体的真实水平制约着作假对测验分数的作用, 诚实分数和作假的关系不是乘法或者加法原则可以解释的, 因而不能通过偏控制, 把测验分数中的作假效应去掉, 而得到真实分数。

### 1.2.2 识别技术

在识别技术中, 研究者将在作假变量上做出异常反应的被试判定为作假者, 从样本中剔除, 例如, 对作假变量划分界线, 将得分高于分界线的被试判定为作假者。Lambert, Arbuckle 和 Holden (2016) 在比较 MCS-D 和 BID-R 量表识别作假者的效果时, 利用判别分析, 将社会称许性反应较为明显的被试判定为作假者。还有一些研究基于这个技术, 使用 IRT 来识别人格测验中的作假。这种数据驱使的事后识别作假个体的方法, 认为作假者的反应模型不同于正常被试的反应模型。因此, 研究者可以利用混合 Rasch 模型 (Mixed Rasch Model, MRM), 依据受测者的反应模式是否类似, 把个体归到不同的类别, 把反应模式属于异常类别的被试判定为作假被试 (Ziegler & Kemper, 2013)。

这种方法的优点在于, 它将控制社会称许性反应的任务缩小至识别重度作假者, 所识别出的被试均为倾向于选择极端答案的被试, 与其他被试的得分或反应模式有较大差异。然而, 识别技

术的最大问题是,误判率与击中率是一对不可调和的矛盾体。当识别的标准比较严格时,误判率很低,然而击中率也会降低,很多作假者会成漏网之鱼;当识别的标准比较宽松时,击中率会明显增加,然而一些受测者可能被误判为作假者。

目前,无论是在人格测验的研究领域还是实践领域,人们基本上还都是采用事后控制方法来应对社会称许性反应。事前控制方法虽然在近些年的发展有所突破,但仍然仅是社会称许性反应主流研究中的一些点缀,文献相对较少,同时由于操作方法过于复杂,也难以在实践中推广使用。

然而,事后控制方法如前面所论述的,无论是识别技术还是校正技术都有一些局限性。实际上,社会称许性量表能否有效地测量社会称许性反应是事后控制方法的关键。如果研究者对社会称许性反应的性质和结构没有清晰准确的认知,所开发的社会称许性量表没有良好的信度和效度,基于该量表测得的社会称许性反应分数进行事后控制,是存在较大风险的。而且,目前流行使用的社会称许性量表不下10个,研究者使用不同的社会称许性量表可能获得不同的研究结果,这也是事后控制方法的效度研究结论并不一致的主要原因。

## 2 社会称许性反应的测量

准确测量社会称许性反应的前提是能够准确理解社会称许性反应的本质,然后依据社会称许性反应的心理结构设计社会称许性量表。这个编制社会称许性量表的过程不是一蹴而就的过程,将近半个世纪以来,很多著名的测量学家对社会称许性反应的心理结构进行了探讨,开发了相应的社会称许性量表。下面基于这些量表的理论基础和编制方式进行分类综述。

### 2.1 无结构的量表

早期的测量学家认为社会称许性反应仅是一种反应偏差,对其内在的结构和发生机制并不关注,代表性的人物是Edwards。但是他们编制量表时却十分谨慎,一个标准化的编制流程是:(1)对大量的人格测验项目所描述的行为进行社会称许程度的评定;(2)选择称许度最高的项目组合成社会称许性量表。合理性在于,承认高称许性项目和否认低称许性项目的被试,完成人格测验时很可能是依据社会称许性而不是测验内容本身做判断。1957年Edwards采用这种方法从MMPI中挑

选了37个项目组成了Edwards SD量表。

另外一种对社会称许性反应的测量操作是“角色扮演”,其操作逻辑是:让两组被试完成同一个人格测验,但是要求一组被试“夸大反应,尽可能表现出最好的形象”,对另一组要求“尽可能诚实回答,结果是匿名的”。然后把两组分数差别大的项目挑选出来,因为被试在这些项目上的夸大反应非常明显。Wiggins Sd量表就是用这种方法建构的,33个项目均来自MMPI。Graham (2000)认为Wiggins Sd量表是一个非常好的量表,应该作为MMPI-2的一个效度量表使用。

这两种操作看起来都是可行的,研究者都假设被试在这些特殊项目上的得分越高,其社会称许性反应越强。然而,它们的代表性量表——Edwards SD量表和Wiggins Sd量表的相关却很低(Gómez, Crespo, & Tobón, 2009), Jiménez-Gómez, Sánchez-Crespo 和 Ampudia-Rueda (2013)比较了Edwards SD量表和Wiggins Sd量表,发现Wiggins Sd量表的诊断准确性高于Edwards SD量表。这说明它们测量了不同的东西,而并非都是社会称许性反应,或者各自测量了社会称许性反应的一部分,早期的社会称许性量表因此受到了强烈的批评。

### 2.2 模糊结构的量表

早期的社会称许性量表的项目多数选自MMPI的临床量表,带有明显的病态特征,同时这种纯粹经验性的量表编制方法也受到了批评。一些研究者认为Edwards SD量表的项目虽然有很高的称许性,比如项目“我经常是快乐的”,但是如果被试回答“是”,并不一定是虚假反应,也可能被试确实有这样的真实表现,因而Edwards SD量表测量的不完全是社会称许性反应,还有被试的人格特质。因而,他们强调社会称许性量表的项目不仅需要具有强烈的“社会称许”意涵,而且要尽量不含有人格特质信息,这样才能更加纯粹地测量出社会称许性反应这种反应偏差(Paulhus, 1998)。

最早的例子是Hartshorne和May的说谎量表,项目描述的都是“得到社会广泛的赞同,但是很少有人做到”的行为。这种项目很少包含人格特质信息,因为很少有人具有这种行为,如果被试承认这种行为,说明其没有真实回答,只是在夸大反应。MMPI和艾森克人格问卷(Eysenck Personality

Questionnaire, EPQ)中的说谎量表(L 量表)就是据此开发的。

最为著名的 Marlowe-Crowne SD 量表(简称 MCSD)采用的编制程序最为复杂, 选编的项目是(1)受社会称许, 但是不经常有的行为(比如, 我从来不会掩饰自己的错误); 或者(2)不受社会称许, 但是常有的行为(比如, 我有时候想偷看别人的日记); (3)较少的心理病理暗示(比如, 项目“男性: 我总希望我是女的; 女性: 我从不因为我是女的而遗憾”带有明显的性别认同障碍含义)。该量表规定个体对不可能的美德肯定, 或者否定人们共有的陋习, 都计 1 分。被试在整个量表上的得分越高, 说明其社会称许性反应越强。

MCSD 选择项目的标准, 成为以后社会称许性项目撰写的重要参考规范。比如, Schinka, Kinder 和 Kremer (1997)给大五人格测验编制了正向表现管理量表(PPM)。PPM 的项目来自大五人格测验, 是基于内容分析(比如, 项目描述了不常见的美德和常见的缺点)和经验分析(比如, 被试得分比较极端的项目)挑选出来的。但是, 由于 MCSD 量表较长, 有 33 道题目, 所以有一些研究者尝试编制简短版的 MCSD 量表(Sârbescu, Costea, & Rusu, 2012; Vésteinsdóttir, Reips, Joinson, & Thorsdottir, 2015)。例如, Vésteinsdóttir, Reips, Joinson 和 Thorsdottir (2017)利用项目反应理论(IRT)和认知访谈(Cognitive Interviews, CogI)删去测量学属性较差的项目, 将剩余的 23 个项目组成简短版的 MCSD 量表。

MCSD 量表是至今应用最为广泛的社会称许性量表(Vu, Tran, Pham, & Ahmed, 2011), 有很多实证研究来支持其有效性。例如, Greenberg 和 Weiss (2012)的研究表明, 13 题版本的 MCSD 能够有效地测量执法人员的社会期许效应。Vésteinsdóttir 等人(2015)测量了 MCSD 在网络在线调查中的信度, 结果表明, 不论是完整版的 MCSD 还是简短版的 MCSD 在网络在线调查中都有较好的信度。

目前, MCSD 量表受到的批评主要有两类。其一是 MCSD 测量的可能不是反应偏差, 而是一种心理特质。McCrae 和 Costa (1983)研究表明, MCSD 的一些高分者声称自己具有的大量优秀的特质, 经他们的配偶证实确实如此, 而非夸大反应。另外, MCSD 量表可以预测一些工作的绩效,

社会称许性反应得分高的个体, 也会有较高的工作绩效。如果控制了社会称许性反应得分会削弱人格测验的预测效度。这表明 MCSD 量表至少测量了一部分的心理特质, 其实这种关于社会称许性反应是人格特质还是反应偏差的争论, 起始于 Block (1965), 他认为 Edward SD 量表(也包括 MMPI 的第一个因子)代表了一种积极的人格特征, 叫做自我复原力(Ego-resiliency)。其二是 MCSD 量表并非单维结构, 测量数据显示其结构比较复杂, 与其假设的单维结构并不一致, 即结构效度不佳(Leite & Beretvas, 2005)。例如, 一项来自马来西亚的研究发现, 使用英文版的 MCSD 和中文版的 MCSD 测量的社会称许性因子具有两维结构, 而且在一个因子上不能满足测量不变性(Kurz, Drescher, Chin, & Johnson, 2016)。

### 2.3 双成分结构的量表

很早研究者就已经发现社会称许性反应有两个不同的成分—— $\alpha$  和  $\gamma$ , 这个命名并没有明确的意义(Wiggins, 1964)。 $\alpha$  因素由 Edwards SD 量表标识,  $\gamma$  因素由 Wiggins Sd 量表标识。这两个因素的心理学意义是什么呢? 直到 1978 年 Sackeim 和 Gur 才对其进行了清晰的界定,  $\alpha$  是自我欺骗: 一些个体报告了不真实的自我描述, 但是他们确信自己就是这样的, 这是无意识中发生的夸大反应;  $\gamma$  是他人欺骗: 一些人为了提高测验分数而故意歪曲反应, 从而愚弄他人。

Sackeim 和 Gur 选择有理论基础的项目分别编制了自我欺骗量表和他人欺骗量表。他们认为性和侵犯的念头广泛存在, 但是往往被否认。如果被试对具有一些冲突性质的内容(比如, 你是否曾经想杀过人)反应过度, 则说明其具有自我欺骗的倾向。对于测量他人欺骗, 他们编制了具有高社会称许意涵的项目, 内容非常夸大和炫耀(比如, 我总是捡起垃圾), 如果被试过度宣称自己具有这些值得赞美的行为, 则一定涉及有意识的粉饰。此外,  $\alpha$  和  $\gamma$  成分都被称为“欺骗”, 是因为都有夸大的特点, 为了确保这种夸大的倾向能由项目充分测量到, 他们推荐使用这种计分方式, 即: 在一个 7 点量表上, 被试如果对正向项目反应为“6”或“7”计 1 分; 对负向项目反应为“1”或“2”计 1 分, 其他反应都计“0”分, 因而被试得分越高, 说明其自我欺骗或者他人欺骗程度越高(Sackeim & Gur, 1978)。

Paulhus 使用大样本被试和大量的社会称许性量表重新进行了因素分析,结果探索出了两个主要的因子,落在第一个因子上的量表包括 Edwards 的 SD 量表、MMPI 的 K 量表、Byrne 的 Repression-Sensitization 量表、Sackeim 和 Gur 的自我欺骗量表。落在第二个因子上的量表包括 Wiggins Sd 量表、Eysenck 的说谎量表、Sackeim 和 Gur 的他人欺骗量表。其他的量表包括 Marlowe-Crowne 量表、好印象量表和 MMPI 的 L 量表,在两个因子上都有较高的载荷。据此,Paulhus 于 1984 年正式提出双成分模型,其理论解释与 Sackeim 和 Gur 一致,但是更加细致和有影响力。他把听起来有些冒昧的维度名字“他人欺骗”,改为“印象管理”,并定义为:一些作答者有意修改答案来传递较好的社会形象的反应趋势(Paulhus, 1984)。自我欺骗则反映了一定程度的非病态自恋和对自我了解的缺乏,类似一种人格特质。自我欺骗与心理适应、乐观、自尊及一般的胜任感等人格概念有本质上的联系,与心理健康、自尊及各种各样的认知偏差量表之间有显著的相关(Paulhus, 1991)。

这种对自我欺骗和印象管理的区分,可以解释社会称许性量表测量的是心理特质还是反应偏差之争。目前看来,应用任何测量自我欺骗的社会称许性量表(如 Edwards SD 量表、K 量表,及 Marlowe-Crowne 量表)对社会称许性反应加以控制,都会降低与自欺反应有联系的人格测验(如焦虑感、成就动机、支配性、幸福感、控制感知、自尊)的预测效度(Paulhus, 1991)。这种社会称许性反应与人格内容变量密不可分,应该不要加以控制才对。相反的,在某些情境下的印象管理是反应偏差则应该加以控制(Paulhus, 1986)。也就是说,印象管理在概念上与要测量的心理特质无关,但仍对此特质的自陈得分产生作用,应该对之加以控制。

Paulhus (1991)编制了 40 个项目社会称许性平衡量表(BIDR 量表),它包括两个子量表——自我欺骗量表和印象管理量表,这是对 Sackeim 和 Gur 的量表的重大改编。具体如下:(1)所有的项目被改写成陈述句,不再是疑问句的形式,比如项目“别人尊重你对你来说是不是很重要?”改成“如果别人碰巧不喜欢我,那也是可能的”。(2)措辞为特质肯定(如,我很和蔼)和否定(如,我不

和蔼)的项目被去掉,全部改为行为描述,比如项目“我是诚实的”被改成“迫不得已的时候我会撒谎”。(3)每个量表中正向(受称许,比如,我从不遮掩自己的错误)和负向(不受称许,比如,我在背后说过朋友的坏话)陈述的数量是相同的。

Paulhus (1998)年指出,已有 40 多个研究证明了 BIDR 量表的结构效度。印象管理量表的合理性在于项目是如此夸大和透明,只有有意识的自我提升才能获得极端的高分。提供给被试的作答指导语对该量表的影响非常大,比如:如果告诉被试这是一个选拔性的测试,目的是挑选优秀的人才,那么被试在该量表上的得分将大大提高。该量表对指导语的高利害性描述的敏感性超过了大五人格测验中的任何一个维度,这也说明该量表对说谎反应的测量是敏感的。自我欺骗量表可以预测各种自我欺骗偏差,比如说,事后聪明偏差。

Barrick 和 Mount (1996)是第一批应用 Paulhus 的框架,检验作假是否会削弱大五人格结构的预测效度的研究者。Rosse, Stecher, Miller 和 Levin (1998)强烈推荐使用 BIDR 来测量反应偏差,认为仅有印象管理可以评估有意的夸大反应偏差,并且指出以前的研究没有区分两种类型的社会称许性反应,才会出现不一致的研究结果。Lanyon 和 Carle (2007)的研究也表明,不论是对社会称许效应,还是对印象管理和自我欺骗拔高, BIDR 都有较好的测量表现。

Paulhus 的一系列研究影响力非常大,奠定了其本人在社会称许性研究领域的领军地位。目前, BIDR 得到了非常广泛的应用,其势头已经超过了老牌的 MCSD 量表。

#### 2.4 具有层级结构的 SD 量表

一些研究者采用实验研究来检验 BIDR 量表的结构效度。他们操作了两种实验条件:“诚实”条件:在指导语中要求被试“尽可能诚实回答”;“装好”条件:在指导语中要求被试“尽可能的夸大反应,表现出一个形象最佳的你”。根据 Paulhus 的理论可以假设,在“诚实”条件下被试只可能发生自我欺骗;在“装好”条件下被试才会发生强烈的印象管理。而且自我欺骗在无意识中发生,印象管理则在意识层面之中,二者在本质上不存在关联。

然而, Holden, Book, Edwards, Wasylkiw 和

Starzyk (2003)发现“诚实”条件下, 被试在自我欺骗和印象管理两个量表上的得分相关为 0.36 (Paulhus 报告的相关为 0.20~0.36), 而在“装好”条件下, 二者相关为 0.80。这与 Furnham, Petrides 和 Spencer-Bowdage (2002)使用因素分析得到的结果类似。很显然, 自我欺骗量表在“装好”条件下, 同印象管理量表一样测量了有意识作假的成分; 而在“诚实”条件下, 被试在印象管理上仍然有较高的得分, 说明也测量到了一些无意识的成分。可见, 自我欺骗和印象管理量表能否独立测量无意识和有意识的社会称许性反应, 受外部条件(测验情景)的影响, 不是由其项目本身的性质决定的。

Paulhus 本人在 1998 年对自我欺骗也有了新的认识。他使用自我标准残差(Self-Criterion Residual, SCR, 是指被试自评的人格分数控制了其知情人的他评分数后的残差, 直接反映了个体的自我知觉偏离真实自我的程度)进行因素分析, 并综合人格特质和防御机制领域的研究成果, 提出自我欺骗有两个成分, 自我取向偏差和道德取向偏差。其根源可以追溯到人类的两种基本的价值观, 代理(Agency)和公共性(Communion)。代理是针对个体(包括个人努力, 个体发展和个人成就)的积极价值观, 公共性是针对人际关系(包括亲密关系, 利他关系甚至整个社会关系)的积极价值观。而自我取向偏差是指个体夸大自己的社会和智力特征, 在支配性、无畏性、情绪稳定性, 智力和创造力等特质上表现卓越, 具有“超人”的品质。道德取向偏差, 是把自己视为一个社会良好的人, 严格遵守道德准则, 否定反社会行为, 声称具备怡人的, 有责任感的, 抑制性的, “圣人”般的特质。

2002 年, Paulhus 在此研究的基础上正式提出了社会称许性反应的层级结构(Paulhus, 2002)。首先, 他质疑了 1984 年自己提出  $\alpha$  和  $\gamma$  因素的本质区别就是意识水平的观点, 认为  $\alpha$  和  $\gamma$  因素的区别在于内容,  $\alpha$  对应着自我取向偏差,  $\gamma$  对应着道德取向偏差。 $\alpha$  和  $\gamma$  因素在无意识和意识层面都会发生, 即自我欺骗包括两个维度  $\alpha$  和  $\gamma$ , 印象管理也包括同样的两个维度  $\alpha$  和  $\gamma$ , 这是一个层级嵌套结构。其次, 他也指出 BIDR 量表的自我欺骗量表只测量了  $\alpha$  因素, 印象管理量表则测量了  $\alpha$  和  $\gamma$  两种因素; 同时这两个量表对无意识和有意识的

夸大反应的测量也是没有清楚区分的, 两个量表都同时测量了无意识和有意识的夸大反应。

Pauls 和 Crost (2004)部分证实了 Paulhus 层级结构理论, 他们发现代理式的指导语(比如, 该测验用来选拔能力优秀的人才)能够激发个体在自我欺骗量表上故意作假, 公共性的指导语(比如, 该测验主要考察你的道德水平)只会使印象管理量表分数上升。Parmač, Galić 和 Jerneić(2014)利用两个研究检验了 Paulhus 层级结构理论的合理性。在研究一中, 他们设计了自我取向和道德取向的社会期望反应量表来测量在“诚实”条件下的自我欺骗和在“装好”条件下的印象管理。在研究二中, 研究者计算这些量表上的分数和外部自我欺骗和印象管理指标的相关。研究表明, 在“诚实”条件下, 社会期望反应量表和自我欺骗指标的相关最强, 在“装好”条件下, 社会期望反应量表的得分和印象管理指标的相关更强, 由此支持了 Paulhus 层级结构理论。

但是也存在一些研究不支持层级结构理论。Li 和 Reb(2008)发现, 虽然相比在个人主义文化中, 在集体主义文化中欺骗是一种更易被接受的行为, 且在个人主义文化中个体更倾向于夸大他们在自我欺骗上的得分, 但是来自集体主义文化的被试在印象管理量表和自我欺骗量表的得分都低于来自个人主义文化的被试。这说明两个量表与测试条件无关, 即并非自我欺骗测量了无意识, 而印象管理测量了有意识的虚假行为。Lönnqvist, Verkasalo 和 Bezmenova (2007)选取了 4 种具有不同文化或者亚文化的被试样本, 考察印象管理和自我欺骗量表与 Schwartz 价值量表的相关, 结果发现印象管理量表与公共性价值观相关, 自我欺骗则与公共性和代理价值观都没有相关。这说明, 印象管理和自我欺骗量表所测内容并非如 Paulhus 所假设的“自我欺骗量表只测量了  $\alpha$  因素, 印象管理量表则测量了  $\alpha$  和  $\gamma$  两种因素”, 所测内容其实更加复杂和多样化。此外, Kam (2013)发现, 自我欺骗得分能够比印象管理得分更有效地预测社会称许性反应, 这与 Paulhus 的假设印象管理量表比自我欺骗量表测量了更多的因素也不一致。

由于 Paulhus 层级结构理论至今仍缺乏相应的测量量表, 所以研究者对其兴趣并不高, 其影响力也相对较小。目前, 有一些研究引用了 Paulhus 层级结构理论中的观点(例如, Cvencek, Greenwald,

& Meltzoff, 2016; Elliot et al., 2016)。

### 3 讨论

前面综述了社会称许性反应的控制方法,以及事后控制方法的基础——社会称许性量表的发展脉络。由于对社会称许性反应的性质和结构的认识不同,研究者相继开发了一系列社会称许性量表,其中以 Edwards, Marlowe 和 Crowne 师生,以及 Paulhus 的贡献最大,它们分别代表了社会称许性反应测量的三个发展阶段。

Edwards 是第一个对社会称许性反应进行系统研究的测量学家,为社会称许性反应的理论研究打下了基本框架。比如,Edwards 开拓了使用因素分析法研究社会称许性反应的先河,随后几乎有 2/3 的研究者在做社会称许性反应的研究时都采用了因素分析法。Edwards 还发现项目特征(比如“社会称许”意涵)影响被试的夸大反应。这正好是项目反应理论(IRT)研究的关注点,于是上个世纪 90 年代以后 IRT 被迅速引入社会称许性反应领域,产生了很多有意义的结果。例如,在社会称许性反应研究中,常用的分数处理方法是:在一个 7 点量表上,被试如果对正向项目反应为“6”或“7”计 1 分;对负向项目反应为“1”和“2”计 1 分,其他反应都计“0”分,但是,Vispoel 和 Kim (2014)研究发现,这种分数处理方法信度不高,而利用原有计分方法,采用 IRT 等级反应模型(Graded Response Model, GRM)来处理这种信息,会得到更为准确的结果。

Marlowe 和 Crowne 师生提供了社会称许性量表的规范化编制过程,其选择项目的标准成为社会称许性量表开发中唯一得到广泛认可的参考规范。著名的 MCSD 量表是至今应用最为广泛的社会称许性量表,由于其使用方便,在人格研究领域当人们希望控制社会称许性反应对人格分数的污染时,一般都采用 MCSD 量表来测量被试的社会称许性反应分数。例如,Claes 等人(2014)使用大五人格量表来测量男性犯人的性格剖面时,利用 MCSD 对分数进行了控制,由此降低社会称许反应的影响;Aldridge 和 Gore (2016)在探索人格与幸福感的关系时,也对 MCSD 的得分进行了控制。而且,Lambert, Arbuckle 和 Holden (2016)研究表明,相比于 BIDR, MCSD 量表能够更准确地甄

别说谎者。

最近的 20 年内,关于社会称许性反应的大部分研究还是围绕着 Paulhus 的双因素理论及其 BIDR 量表展开的。主要有 4 种类型的研究:(1)探讨 BIDR 量表能否解决社会称许性反应是心理特质还是反应偏差之争。比如, Rosse, Stecher, Miller 和 Levin (1998)强烈推荐使用 BIDR 来测量反应偏差,认为仅有印象管理可以评估有意的歪曲反应偏差。(2)探讨 BIDR 量表的测量学属性是否理想。比如, Li 和 Bagger (2007)使用概化理论, Lanyon 和 Carle (2007)使用因素分析都发现, BIDR 量表的信度不好,两个量表的分数存在高度相关。此外, Kam (2013)发现,当 BIDR 分数被当作二分变量来计算时,其预测效度较低。(3)与传统的量表混在一起进行因素分析,考察量表的结构效度。比如, Leite 和 Beretvas (2005)发现这些社会称许性量表测量的既不是一个因素,也不是两个因素,因而需要更加谨慎的解释; Li 和 Li (2008)利用中国样本发现, BIDR 既不支持 Paulhus 从加拿大样本获得的两因子(自我欺骗,印象管理)模型,也不支持 Paulhus 和 Reid 提出的三因子模型(自我欺骗拔高,自我欺骗否定,印象管理),而是分析产生了四因子模型(自我欺骗拔高,自我欺骗否定,印象管理拔高,印象管理否定),说明自我欺骗和印象管理可以依据是拔高还是否定被进一步划分,以及 BIDR 的结构在不同国家或文化之间存在差异。(4)有一些研究并不依托社会称许性量表,而是直接采用实验研究的方式,直接对人格测验上的分数差(“装好条件”减去“诚实条件”,或者自评分数减去他评分数)进行因素分析,重新探讨社会称许性反应的结构。比如, Ziegler 和 Buehner (2009)认为社会称许性反应的结构比较复杂,并不是双因素结构。

可以看出,社会称许性反应领域的主流研究几乎都是围绕 Paulhus 的双因素理论和 BIDR 量表展开的,近些年更是以批判研究为多。但是,对 Paulhus 于 2002 年提出的层级结构关注很少,只有个别的实证研究来检验该理论的有效性,主要是由于没有开发出测量层级结构的社会称许性量表,使得其他研究者无法展开深入的探讨。

Paulhus 关于如何编制自我欺骗和印象管理量表的项目,曾有过详细的论述。他认为 Edwards SD 和 Wiggins Sd 量表,虽然都是由称许水平高的

项目构成,但是相关却很低,一个关键的区别是“同意率”,即被试承认自己符合项目所描述行为的比例。在日常情况下,SD量表有较高的“同意率”,Sd则“同意率”低,只有在特殊情况下(比如选拔测试中)Sd的“同意率”才会升高,因此Sd测量了一种有意的夸大,而SD量表则测量了自我欺骗。因此,Paulhus编制BIDR的印象管理项目时,特别强调“项目是如此夸大和透明,只有有意识的自我提升才能获得极端的高分”。这种操作很巧妙,无可厚非,但是实施的结果却暴露出一些问题:(1)BIDR的印象管理量表的项目倾向于测 $\gamma$ 因素。由于 $\gamma$ 因素需要激起作答者故意歪曲反应的欲望,因此项目大多与社会规范相关,又因为与社会规范相关的内容较少,所以印象管理量表测量的内容偏窄,具有局限性。而自我欺骗量表的项目测量的都是 $\alpha$ 因素,在诚实条件下已经有了较高的“同意率”。但是为了使项目分数提高,有一些作答者也会有意识的作假。所以,这些项目既受到作答者无意夸大的影响,也受到作答者故意作假的影响,即这些项目同时测量了自我欺骗和印象管理的成分,应该被删除。因而,BIDR量表需要增加大量的项目;(2)BIDR量表中大量项目含有极端的修饰词,例如,“从来不”,“总是”,“一直”,影响了被试的反应。杨中芳(1996)怀疑,推崇中庸之道的中国人对这些吹牛皮式的陈述不会肯定反应的,即中国人在进行自我报告时,有降低极端程度的倾向,把频率极高的事件弱化为频率较高的事件,把频率极低的事件弱化为频率较低的事件。刘萃侠(2001)的研究证明了这一点,中国大学生在这类项目上的社会称许性反应得分是较低的。

目前,人们使用人格测验的目的主要有两个,一是用来搜集数据,用于做研究,二是用于人才测评和选拔。其实,不只是人格测验存在社会称许性反应的问题,凡是基于自陈式题目(由被试判断自己是否符合某种行为特征,或者具有某种价值取向、态度和兴趣等)的调查和测试都存在同样的问题。大多数的心理学领域的研究还都是采用自陈式量表来搜集数据,其研究结果会不可避免的受到社会称许性反应的污染。因而,理解社会称许性反应的性质,对其进行准确的测量和控制,对于巩固和提升心理学领域的很多研究结论的合理性和可信性都是至关重要的。然而,由以上讨

论可见,尽管近一个世纪以来心理测量学家付出了大量的心血和努力,希望能够提出有效的控制社会称许性反应的方法,但是社会称许性反应的测量仍然不尽人意,这对于基于社会称许性量表的测试分数来进行矫正或者识别作假者的事后控制方法带来了很大的挑战。但是,这并不意味着当前心理学研究者不考虑自陈式量表的社会称许性反应的影响是可取的,他们往往直接采用自陈式量表分数进行研究,并不关心社会称许性反应对其研究结论可靠性的影响。因而,建议这些研究应该使用社会称许性量表来考察被试不诚实作答的程度,通过识别和剔除严重作假者,或者采用控制的方法把被试作假的分数从自陈量表分数中去除掉后,再进行后续的研究。通过对比控制和不控制社会称许性反应的研究结果是否有差异,来检验社会称许性反应的影响。如果研究结果差异很大,需要考虑在数据搜集阶段是否存在被试不诚实作答,导致数据被污染的可能性。对于采用人格测验来进行人才测评和选拔的研究者,采用社会称许性量表来识别作假者或者控制作假分数,更是非常必要,否则有可能被选拔的被试都是作假者,因为作假者往往集中在分数的顶端。但是,有必要增加其他的选拔手段来相互印证,比如面试,无领导小组等,以规避社会称许性量表测量不够精准的问题。

#### 4 小结

目前在心理学研究中用来收集数据的工具大多数都是自陈量表,很多研究者发现自陈量表的维度之间相关较高,由于共线性问题,之后进行的统计分析结果都不太好。维度之间的相关高多是由于社会称许性反应造成的,因而很多研究者希望能够控制社会称许性反应。本文对社会称许性反应的内部结构和测量量表进行了文献综述,希望能对研究者有参考价值。

#### 参考文献

- 刘萃侠. (2001). 马洛-克罗恩社会赞许性量表对中国被试适用性之初步验证. *社会学研究*, (2), 49-57.
- 王珊, 骆方, 刘红云. (2014). 迫选式人格测验的传统计分与IRT计分模型. *心理科学进展*, 22, 549-557.
- 杨中芳. (1996). *如何研究中国人: 心理学本土化论文集*. 台北: 桂冠图书公司.
- Aldridge, T. R., & Gore, J. S. (2016). Linking personality

- traits with well-being: The influence of primary social roles. *Psychological Studies*, *61*, 233–244.
- Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, *81*, 261–272.
- Block, J. (1965). *The challenge of response sets: Unconfounding meaning, acquiescence, and social desirability in the MMPI*. New York: Appleton-Century-Crofts.
- Brown, A. (2016). Item response models for forced-choice questionnaires: A common framework. *Psychometrika*, *81*, 135–160.
- Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, *71*, 460–502.
- Brown, A., & Maydeu-Olivares, A. (2013). How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychological Methods*, *18*, 36–52.
- Claes, L., Tavernier, G., Roose, A., Bijttebier, P., Smith, S. F., & Lilienfeld, S. O. (2014). Identifying personality subtypes based on the five-factor model dimensions in male prisoners: Implications for psychopathy and criminal offending. *International Journal of Offender Therapy and Comparative Criminology*, *58*, 41–58.
- Cvencek, D., Greenwald, A. G., & Meltzoff, A. N. (2016). Implicit measures for preschool children confirm self-esteem's role in maintaining a balanced identity. *Journal of Experimental Social Psychology*, *62*, 50–57.
- Edwards, A. L., & Walker, J. N. (1961). A short form of the MMPI: The SD scale. *Psychological Reports*, *8*, 485–486.
- Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology*, *84*, 155–166.
- Elliot, A. J., Aldhobaiban, N., Murayama, K., Kobeisy, A., Gocłowska, M. A., & Khyat, A. (2016). Impression management and achievement motivation: Investigating substantive links. *International Journal of Psychology*, doi: 10.1002/ijop.12252. (in Press)
- Furnham, A., Petrides, K. V., & Spencer-Bowdage, S. (2002). The effects of different types of social desirability on the identification of repressors. *Personality and Individual Differences*, *33*, 119–130.
- Gómez, F. J., Crespo, G. S., & Tobón, C. (2009). A social desirability scale for the MMPI-2: Which of the two: Wiggins (WSD) or Edwards (ESD)? *The European Journal of Psychology Applied to Legal Context*, *1*, 147–163.
- Graham, J. R. (2000). Use with special groups. In J. R. Graham (Ed.), *MMPI-2: Assessing Personality and Psychopathology* (pp. 215–244). New York: Oxford University Press.
- Greenberg, B., & Weiss, P. (2012). Validation of a short form of the Marlowe-Crowne for use with law enforcement personnel. *Journal of Police and Criminal Psychology*, *27*, 123–128.
- Guenole, N., Brown, A. A., & Cooper, A. J. (2016). Forced-choice assessment of work-related maladaptive personality traits: Preliminary evidence from an application of thurstonian item response modeling. *Assessment*, doi: 10.1177/1073191116641181. (in Press)
- Holden, R. R., Book, A. S., Edwards, M. J., Wasylkiw, L., & Starzyk, K. B. (2003). Experimental faking in self-reported psychopathology: Unidimensional or multidimensional? *Personality and Individual Differences*, *35*, 1107–1117.
- Hontangas, P. M., de la Torre, J., Ponsoda, V., Leenen, I., Morillo, D., & Abad, F. J. (2015). Comparing traditional and IRT scoring of forced-choice tests. *Applied Psychological Measurement*, *39*, 598–612.
- Jiménez-Gómez, F., Sánchez-Crespo, G., & Ampudia-Rueda, A. (2013). Is there a social desirability scale in the MMPI-2-RF? *Clínica y Salud*, *24*, 161–168.
- Kam, C. (2013). Probing item social desirability by correlating personality items with Balanced Inventory of Desirable Responding (BIDR): A validity examination. *Personality and Individual Differences*, *54*, 513–518.
- Kelly, E. L., Miles, C. C., & Terman, L. M. (1936). Ability to influence one's score on a typical pencil-and-paper test of personality. *Journal of Personality*, *4*, 206–215.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, *47*, 2025–2047.
- Kurz, A. S., Drescher, C. F., Chin, E. G., & Johnson, L. R. (2016). Measuring social desirability across language and sex: A comparison of Marlowe-Crowne Social Desirability Scale factor structures in English and Mandarin Chinese in Malaysia. *PsyCh Journal*, *5*, 92–100.
- Lambert, C. E., Arbuckle, S. A., & Holden, R. R. (2016). The Marlowe-Crowne social desirability scale outperforms the BIDR impression management scale for identifying fakers. *Journal of Research in Personality*, *61*, 80–86.
- Lanyon, R. I., & Carle, A. C. (2007). Internal and external validity of scores on the balanced inventory of desirable responding and the Paulhus deception scales. *Educational and Psychological Measurement*, *67*, 859–876.
- Leite, W. L., & Beretvas, S. N. (2005). Validation of scores on the Marlowe-Crowne social desirability scale and the balanced inventory of desirable responding. *Educational and Psychological Measurement*, *65*, 140–154.
- Leite, W. L., & Cooper, L. A. (2010). Detecting social desirability bias using factor mixture models. *Multivariate*

- Behavioral Research*, 45, 271–293.
- Li, A., & Bagger, J. (2007). The Balanced Inventory of Desirable Responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement*, 67, 525–544.
- Li, F., & Li, Y. J. (2008). The balanced inventory of desirable responding (BIDR): A factor analysis. *Psychological Reports*, 103, 727–731.
- Li, A., & Reb, J. (2008). A cross-nations, cross-cultures, and cross-conditions analysis on the equivalence of the Balanced Inventory of Desirable Responding (BIDR). *Journal of Cross-Cultural Psychology*, 2, 214–233.
- Lönnqvist, J. E., Verkasalo, M., & Bezmenova, I. (2007). Agentic and communal bias in socially desirable responding. *European Journal of Personality*, 21, 853–868.
- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51, 882–888.
- Miller, B. K., & Gallagher, D. G. (2016). Examining trait entitlement using the self-other knowledge asymmetry model. *Personality and Individual Differences*, 92, 113–117.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15, 263–280.
- Parmač Kovačić, M., Galić, Z., & Jerneić, Ž. (2014). Social desirability scales as indicators of self-enhancement and impression management. *Journal of Personality Assessment*, 96, 532–543.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598–609.
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires* (pp. 143–165). Berlin Heidelberg, Germany: Springer.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17–59). New York: Academic Press.
- Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology*, 74, 1197–1208.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 67–88). Mahwah, NJ: Erlbaum.
- Pauls, C. A., & Crost, N. W. (2004). Effects of faking on self-deception and impression management scales. *Personality and Individual Differences*, 37, 1137–1151.
- Perinelli, E., & Gremigni, P. (2016). Use of social desirability scales in clinical psychology: A systematic review. *Journal of Clinical Psychology*, 72, 534–551.
- Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634–644.
- Sackeim, H. A., & Gur, R. C. (1978). Self-deception, self-confrontation, and consciousness. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 139–197). New York: Springer.
- Sărbescu, P., Costea, I., & Rusu, S. (2012). Psychometric properties of the Marlowe-Crowne social desirability scale in a Romanian sample. *Procedia-Social and Behavioral Sciences*, 33, 707–711.
- Schinka, J. A., Kinder, B. N., & Kremer, T. (1997). Research validity scales for the NEO-PI-R: Development and initial validation. *Journal of Personality Assessment*, 68, 127–138.
- Sjöberg, L. (2015). Correction for faking in self-report personality tests. *Scandinavian Journal of Psychology*, 56, 582–591.
- Vésteinsdóttir, V., Reips, U. D., Joinson, A., & Thorsdóttir, F. (2015). Psychometric properties of measurements obtained with the Marlowe-Crowne Social Desirability Scale in an Icelandic probability based Internet sample. *Computers in Human Behavior*, 49, 608–614.
- Vésteinsdóttir, V., Reips, U. D., Joinson, A., & Thorsdóttir, F. (2017). An item level evaluation of the Marlowe-Crowne Social Desirability Scale using item response theory on Icelandic Internet panel data and cognitive interviews. *Personality and Individual Differences*, 107, 164–173.
- Vispoel, W. P., & Kim, H. Y. (2014). Psychometric properties for the Balanced Inventory of Desirable Responding: Dichotomous versus polytomous conventional and IRT scoring. *Psychological Assessment*, 26, 878–891.
- Vu, A., Tran, N., Pham, K., & Ahmed, S. (2011). Reliability of the Marlowe-Crowne social desirability scale in Ethiopia, Kenya, Mozambique, and Uganda. *BMC Medical Research Methodology*, 11, 162.
- Wang, W. C. (2015, October). *New item response theory models for ipsative data*. Paper presented at the 2015 Taiwanese Psychological Association Annual Convention and The International Convention of Learning, Teaching, and Assessment, Taiwan Normal University, Taipei, Taiwan.
- Wiggins, J. S. (1964). Convergences among stylistic response measures from objective personality tests. *Educational*

- and Psychological Measurement*, 24, 551–562.
- Whitaker, L., Long, J., Petróczi, A., & Backhouse, S. H. (2014). Using the prototype willingness model to predict doping in sport. *Scandinavian Journal of Medicine & Science in Sports*, 24, e398–e405.
- Ziegler, M., & Buehner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement*, 69, 548–565.
- Ziegler, M., & Kemper, C. (2013). Extreme response style and faking: Two sides of the same coin. In P. Winker, N. Menold, & R. Porst. *Interviewers deviations in surveys—impact, reasons, detection and prevention* (pp. 217–233). Frankfurt, Germany: Peter Lang.
- Ziegler, M., MacCann, C., & Roberts, R. D. (Eds.). (2012). *New perspectives on faking in personality assessment*. Oxford, UK: Oxford University Press.

## Measurement and control of Socially Desirable Responding

PAN Yiqin<sup>1</sup>; LUO Fang<sup>1,2</sup>

(<sup>1</sup> Faculty of Psychology, Beijing Normal University)

(<sup>2</sup> Beijing Advanced Innovation Center for Future Education, Beijing 100875, China)

**Abstract:** Socially Desirable Responding (SDR) is one of the most common sources of bias that affect the validity of personality tests. Considerable attempts have been made to investigate the concept and explore its underlying structure, resulting in a number of SDR measurement scales. The aim of the article is to review the literature on the development of SDR measures and identify unsolved issues. From both preventive and retrospective perspectives, the authors compared current approaches to coping with SDR and revealed the influence of retrospective control on the validity of SDR scales. Lastly, additional guidance is provided on potential methods of SDR reduction.

**Key words:** Socially Desirable Responding (SDR); self-deception; impression management