

# 基于印象管理量表的反应模式筛查故意作假者\*

任 岩 潘逸沁 骆 方

(北京师范大学心理学部,应用实验心理北京市重点实验室,心理学国家级实验教学示范中心(北京师范大学),北京 100875)

**摘 要:**社会称许性反应由两个成分组成:自我欺骗(无意识地夸大反应)和印象管理(有意识地歪曲作答)。采用实验研究范式和 Mixed Rasch Model,探讨二者在 BIDR 印象管理量表上的得分及作答模式差异。结果发现,当发生印象管理时,量表分数高于仅发生自我欺骗时的分数,且被试倾向选择极端反应。进而采用 logistic 回归对量表分数划定分界线,将高于分数线的被试认定为故意作假者,低于分界线的被试认定为仅发生了自我欺骗,没有故意作假。

**关键词:**社会称许性反应;自我欺骗;印象管理;BIDR 量表

**中图分类号:**B841.2

**文献标识码:**A

**文章编号:**1003-5184(2018)03-0236-05

## 1 引言

社会称许性反应(Socially Desirable Responding,以下简称SDR),是指被试在填答人格测验时,倾向于朝向社会期望的方向回答(Edwards, 1961; Tracey, 2016)。研究者发现,SDR会污染人格测验,大大降低测验分数的真实性(Goffin & Christiansen, 2003; Dunning, Heath, & Suls, 2004; Krumpal, 2013; Perinelli & Gremigni, 2016)。1984年Paulhus提出了SDR有两个成份:自我欺骗和印象管理。自我欺骗是指无意识中发生的夸大反应,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。它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,与心理适应、乐观、自尊及一般的胜任感等人格概念有关(Farrow, Burgess, Wilkinson, & Hunter, 2015)。对任何测量自我欺骗的社会称许性量表分数加以控制,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。印象管理是指个体故意报告不真实的自我描述来传递较好的社会形象的反应趋势(Li & Bagger, 2006)。印象管理在概念上与人格特质无关,但仍对人格特质的自陈得分产生作用,是一种反应偏差,对人格测验的危害更大。

基于双成分模型,Paulhus开发出了社会称许性平衡量表(Balanced Inventory of Desirable Responding, BIDR, Paulhus, 1991),描述的都是人类不可能具有的美德(或不可能去掉的陋习),如果被试承认

(或否定)自己具有这类行为,则表明他(她)在无意或者故意地夸大反应。BIDR有两个子量表,其中自我欺骗量表的内容主要是有关性和侵犯方面的,这些念头在人的无意识中广泛存在,如果被试反应过度,则说明其具有自我欺骗的倾向。印象管理量表的项目则非常夸大和透明,个体对这些值得赞美的行为的过度宣称,一定涉及有意识的粉饰。Paulhus推荐对印象管理量表使用特别的计分方式:在一个7点量表上,对正向项目仅对“6”和“7”计分;对负向的项目仅对“1”和“2”计分。他认为自我欺骗在这些题目上只能引发“微弱”的反应,得分是较低的。

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,指导语中要求被试“尽可能诚实回答”;“作假”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”。Holden, Book, Edwards, Wasyliw 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.36,而在“作假”条件下,二者相关为0.80。很显然,自我欺骗量表在“作假”条件下测量了一些有意识作假的成分;而在“诚实”条件下,印象管理量表也测量到了无意识中发生的自我欺骗。因而,BIDR的两个量表并没有良好的区分效度,尤其是印象管理量表的得分受到了自我欺骗的污染,无法实现使用它来筛选故意作假的被试的目的(Holden et al., 2003)。值得探讨的是,当被试作答印象管理量表

\* 基金项目:北京教育科学规划重点项目(CADA17078)。

通讯作者:骆方, E-mail: luof@bnu.edu.cn。

时,发生自我欺骗与发生印象管理在印象管理量表上的得分是否有差别,以及二者的作答模式是否有区别。如果能够把二者有效区分开来,那么就有望将自我欺骗对印象管理量表的污染作用屏蔽掉,来筛选故意作假的被试,提高人格测验的效度了。

本研究拟首先采用实验研究的范式来诱导被试发生自我欺骗和印象管理,考察他们在印象管理量表得分上的差异。然后,采用 Mixed Rasch Model (MRM, Fischer & Molenaar, 2012) 来考察这些被试在印象管理量表上的作答模式的差异。基于对这些差异的分析,采用 logistic 回归划定印象管理量表的分界线,将高于分数线的被试认定为有意的作假者,而低于分界线的被试认定为诚实者(仅在无意识中发生了自我欺骗)。

## 2 研究方法

### 2.1 研究工具

选取 BIDR 的印象管理量表中的 10 道题目作为社会称许性量表,嵌套在《大五人格测验简版》中使用。

### 2.2 被试

北京师范大学四年级 216 名学生,分属两个公共课班级,包含十多个专业,其中,76 名学生来自文科系,80 名来自理科系,60 名来自文理兼收系。平均年龄 22.36 岁( $SD = 1.04$ ),男生 64 人,女生 152 人。

### 2.3 施测

要求两个班级的学生按照指导语的要求完成人格测验。其中,一个班级接受诱使作假的指导语,一个班级接受诚实回答的指导语。

诚实指导语如下:“该测验是由若干描述性格特征的句子构成,请你判断自己的符合程度,并在相应的数字上打勾,数字越大表示越符合。测试结束后,你可获得测试结果,据此你可以更加了解自己。我们对你的测试结果严格保密,请真实回答!”

作假指导语如下:“该测验是由若干描述性格特征的句子构成,你不需要按照自己的真实情况回答,请把自己表现得尽量优秀,并在相应的数字上打勾,数字越大表示越符合。如果你的测验成绩足够优秀,排名在前 5%,我们两周后会公布录用名单,给你 20 元的奖励。”

诚实条件下被试在印象管理量表上的得分代表自我欺骗;作假条件下则代表印象管理。

### 2.4 统计分析

社会称许性量表同时测量了“诚实条件”组的

自我欺骗和“作假条件”组的印象管理,它们可视为一种潜在特质(夸大反应的动机)同时估计。采用 Mixed Rasch Model 来估计潜在特质  $\theta$ ,是因为(1)MRM 能够识别出具有不同反应模式的群体,可以揭示作假与 SDR 是否会促使不同类型的反应发生。(2)MRM 通过对项目功能差异的检验,探索出具有不同反应模式的群体,这与根据均值差异来判断类别的统计技术(比如,聚类分析,鉴别分析等)不同(Zickar, Gibby, & Robie, 2004)。(3)MRM 采用的是分部评分模型(Partial Credit Model, PCM)来分析个体的反应模式,PCM 是基于项目的相邻类级之间的关系来建构的,所以估计出的项目类级的难易参数  $d_i$  不一定是递增的。理论上, Likert 量表的各个类级(比如, 0~4 个级别分别代表很不符合-非常符合)的难易参数应该是递增的,也即类级越高,反应难度越大。如果个体不按照自己的真实水平反应,而更倾向于对某个类级反应(称许性高的),那么项目类级的难易参数的顺序将被打乱,作假者的反应模式会被清晰展现。

采用软件 WINMIRA 2001 进行 Mixed Rasch Model 分析, SPSS 16.0 进行 logistic 回归分析。

## 3 研究结果

### 3.1 自我欺骗和印象管理的区分

#### 3.1.1 自我欺骗和印象管理的社会称许性强度不同

使用 MRM 分析两个组作答社会称许性项目的反应数据,估计的潜在特质  $\theta$  值。结果显示,“诚实条件”组的  $\theta$  值均值为 0.115,“作假条件”组的  $\theta$  值均值为 1.272,二者有显著差异( $t = -9.407^{***}$ ),效应值  $Cohen's d$  为 1.304,说明印象管理的强度更大。需要注意的是,“诚实条件”组 27 人的  $\theta$  值大于 1,说明一些个体的自我欺骗水平较高。在  $\theta$  值的最高端(2~5),几乎全部是“作假条件”组,这说明采用社会称许性量表时划定较高的分界线识别严重作假者是可取的。

#### 3.1.2 印象管理与自我欺骗的反应模式不同

MRM 采用 CAIC 指标来判断分类结果,当把个体分为两类时 CAIC 最小(1 类: 5942.80; 2 类: 5940.40; 3 类: 6107.77),说明把个体分为两类最为合适。当将个体划分为两类时,所有项目的 Q 值在两个类别上均未超过  $\pm 1.96$  范围,说明所有项目的拟合良好。

图 1 描述了这两个类别在项目(横坐标)5 个类级上的反应概率。可见,第一类(左图)的个体在第

2 和第 3 类级上有较大的反应概率,而第二类(右图)的反应集中在第三和第四类级上,尤其对极端

的类级 4 有最大的反应概率。

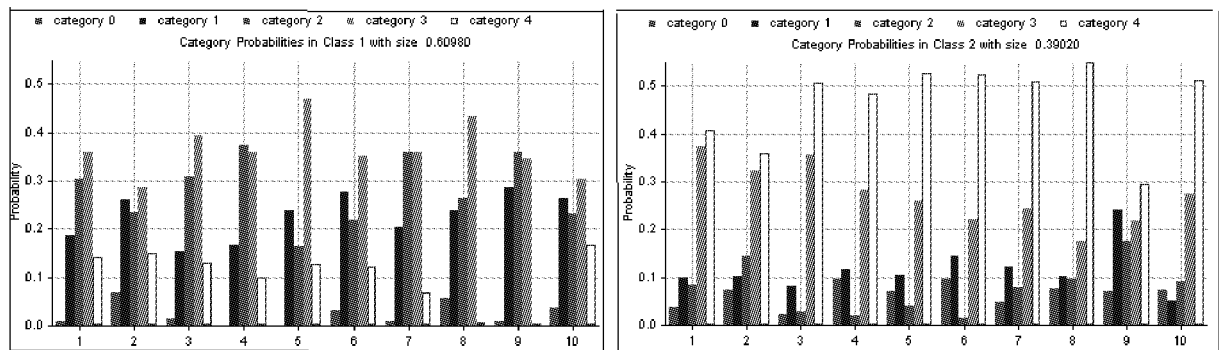


图 1 两个类别的反应概率图

图 2 是这两个类别在项目(横坐标)上的各个类级的难易参数图。第一类(左图)的各类级位置参数的顺序基本正确,第 2 和 3 类级的位置参数几乎重合,说明对个体而言两个类级(中间状态、比较符合)并没有区别,个体倾向于选择代表较好程度

的“比较符合”。第二类(右图)各类级位置参数的顺序是紊乱的,最低的是第 3 类级,4 类级和 1 类级的位置参数类似,说明个体对第 3 和 4 类级敏感,倾向于对极端好的选项反应。

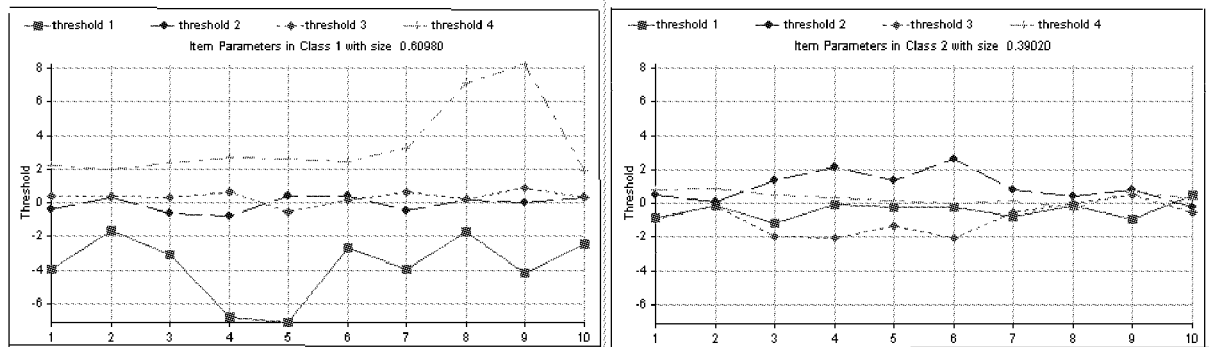


图 2 两个类别的难易参数图

可见,第一类的反应模式基本正确,而第二类的类级之间的难易参数完全被打乱,倾向于选择极端好的反应键。这说明第一类个体的反应正常,动机低只会在低水平类级上反应,动机高才会对高水平类级反应。第二类个体倾向于选择极端反应键,他们的反应更加剧烈。

进一步考察“诚实条件”组和“作假条件”组在这两个类别上的人数。结果显示,“作假条件”组有 55 人,“诚实条件”组有 25 人属于第二类,“作假条件”组有 53 人,“诚实条件”组有 83 人属于第一类。也即,“诚实条件”组更倾向于第一类别,“作假条件”组更倾向于第二类别。由此可见,作假者和诚实者的差异不仅大小强度不同( $\theta$  值),而且反应模式也不太一样(作假者更容易选择极端的反应键,得分更高)。

### 3.2 Logistic 回归划定分界线

MRM 分析的结果显示,研究者能够通过 SDR

分数的高低区分自我欺骗和印象管理。采用 Logistic 回归划定分界线识别作假者:以个体在社会称许性量表的得分(SDR 分数)为自变量,两种反应 01 编码的识别变量(0 为诚实,1 为作假)为因变量,对 216 名被试的反应进行 Logistic 回归。方程的解释率 Cox & Snell  $R^2$  和 Nagelkerke  $R^2$  分别为 0.336 和 0.448,说明该社会称许性量表能够较好的识别作假反应。

Logistic 回归可以估计个体的作假概率,如果大于临界概率,可将个体判为作假者。下面考察三种临界概率下误中率和误判率的大小,以及相应的社会称许性量表的得分(即分界线)。结果显示,当临界概率为.3 时,分界线较低(3.35),容易将个体判定为作假者,误中率很小(4.6%),但是 20.8%的误判率较高,可能会将一些诚实者误认为作假者。临界概率为 0.5 和 0.7 时,分界线升高,更容易将个体判定为诚实者,误中率增大(12.5% 和 21.3%),但

是被误判的诚实者减少(13.4%和4.2%)。可见,分界线低时,误中率低,但误判率高;分界线高时,误判率低,但误中率高。因而,误中率和误判率不能同时减小,显然划定分界线是一个两难的决策。

本研究在作假条件下的测验结束后询问被试:你对大约多少的项目作了夸大反应?请在以下5个反应项进行选择(10%、20%、30%、40%、50%)。结果显示,无论在何种临界概率下,被错误判为诚实反应的被试基本上都是作假比例为10%的轻微作弊者。比如0.5的临界概率下被误中的27人中,有22人都是这种轻微作弊者。在0.7的临界概率下误中的46人中,有37人是轻微作弊者。这说明,当选择较高的临界概率(分界线为4)时,虽然误中率较高,但在误中的人群中大多数是轻微作弊者,严重作弊者(约20%)是能够较好识别的。而选择较低的临界概率(分界线为3.5)时,更多的人(约60%)被判为作弊者,如此大比例的剔除被试,显然会大大降低使用人格测验的意义。

## 4 讨论

### 4.1 SDR 的实验研究

SDR的研究范式通常是:随机分配两组被试完成相同的人格测验,对一组要求“夸大反应,尽可能表现出最好的形象”(作假条件组),对另一组要求“尽可能诚实回答,结果匿名”(诚实条件组),通过两个组的反应对比了解SDR的性质。研究采用这种设计,使用不同的指导语创建诚实条件组和作假条件组,诱使作假条件组发生印象管理。本研究对作假条件组的指导语做了精心安排,鼓励被试作假并给优秀者20元的奖励,期望能够最大可能的模拟作弊者心境。Mueller-Hanson, Heggstad和Thornton(2003)认为奖励的方式会诱发出更加真实的作假反应。

研究表明,“作假条件”组的夸大反应( $\theta$ 值)比“诚实条件”组大很多,而且在反应模式上也体现出更容易选择极端反应的特征。对于MRM分出的两个类别而言,“作假条件”组有55人,“诚实条件”组有25人属于第二类,“作假条件”组的人数显然更多,说明选拔情景中个体更倾向于极端反应。个体在诚实条件下有各种反应定势(Response Set)包括极端反应、趋中反应、顺从效应等,这表明一部分个体(25人)在回答测验时,可能倾向于选择极端反应键,他们视5点量尺与2点量尺相同,并没有程度上的差别。“作假条件”组有53人,“诚实条件”

组有83人属于第一类,这些“作假条件”组的个体也许能够理解反应键的不同强度,谨慎的传递理想个体形象,他们的夸大反应是更加精明的。

### 4.2 使用 Logistic 回归划定分界线

使用 Logistic 回归可以帮助研究者寻找影响因变量分类的关键自变量,确定该怎样对自变量划定分界线,能够更好的预测因变量的类别。本文把既有的实验分组定为因变量(1为作假,0为诚实),社会称许性量表分数作为自变量,结果显示,该分数能够较好的预测因变量。但是,基于临界率划定分界线是一个两难问题,如果临界概率较低,分界线会较低,个体容易被判定为作弊者,误中率很小,但是误判率升高,可能把一些诚实者误认为作弊者。相反,临界概率较高时,分界线升高,个体更容易被判定为诚实者,误中率增大,但是被误判的诚实者减少。可见,分界线低时,误中率低,但误判率高;分界线高时,误判率低,但误中率高,误中率和误判率不能同时减小。

MRM的结果部分显示:划定较高的分界线更为合适,因而建议划定较高的分界线(4分)识别严重的作弊者,虽然误中率较高,但是误判率能够大幅度的降低,即不会把诚实的回答者误认为是作弊者。

## 5 结论

研究采用实验研究范式,探讨了当被试发生自我欺骗或印象管理时,其在BIDR印象管理量表上的得分差别有多大及二者的作答模式是否有区别。MRM的分析结果显示,印象管理和自我欺骗的分数强度不同,而且它们会促使不同类型的反应发生,有更多的作弊者倾向于选择极端的反应键,项目得分更高。基于MRM的分析结果,继续采用 Logistic 回归技术对量表分数划定了分界线,虽然分界线的划定是一个两难问题,但是划定较高的分界线识别严重的作弊者更为合适。

## 参考文献

- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106.
- Edwards, A. L., & Walker, J. N. (1961). *A short form of the MMPI: The SD scale*. Psychological Reports.
- Farrow, T. F., Burgess, J., Wilkinson, I. D., & Hunter, M. D. (2015). Neural correlates of self-deception and impression management. *Neuropsychologia*, 67, 159-174.

- Fischer, G. H. , & Molenaar, I. W. (2012). *Rasch models; Foundations, recent developments, and applications*. Springer Science & Business Media.
- Goffin, R. D. , & Christiansen, N. D. (2003). Correcting personality tests for faking: A review of popular personality tests and an initial survey of researchers. *International Journal of Selection and Assessment*, 11(4), 340 – 344.
- 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.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47(4), 2025 – 2047.
- Li, A. , & Bagger, J. (2006). Using the BIDR to distinguish the effects of impression management and self – deception on the criterion validity of personality measures: A meta – analysis. *International Journal of Selection and Assessment*, 14(2), 131 – 141.
- Mueller – Hanson, R. , Heggstad, E. D. , & Thornton III, G. C. (2003). Faking and selection: Considering the use of personality from select – in and select – out perspectives. *Journal of Applied Psychology*, 88(2), 348.
- Paulhus, D. L. (1984). Two – component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598 – 609.
- 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 psychology attitudes*. New York: Academic Press.
- Perinelli, E. , & Gremigni, P. (2016). Use of social desirability scales in clinical psychology: A systematic review. *Journal of clinical psychology*, 72(6), 534 – 551.
- Tracey, T. J. (2016). A note on socially desirable responding. *Journal of Counseling Psychology*, 63(2), 224.
- Zickar, M. J. , Gibby, R. E. , & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed – model item response theory. *Organizational Research Methods*, 7(2), 168 – 190.

## Using the Responses in Impression Management Scale to Identify Fakers

Ren Yan Pan Yiqin Luo Fang

( Faculty of Psychology, Beijing Normal University, Beijing Key Lab of Applied Experimental Psychology,  
National Experimental Teaching Demonstrating Center of Psychology( Beijing Normal University) , Beijing 100875 )

**Abstract:** Social desirability consists of two components, self – deception and impression management. Self – deception refers to exaggerating responses unintendedly and impression management means distorting responses intendedly. The result from Mixed Rasch Model showed; the degree of social desirability response, of the subjects in the fake condition are higher than ones in the honest condition; the subjects in the fake condition tend to choose the extreme options. The result of logistic regression analysis showed the probability of 0.7 is a suitable cutting – off point for identifying severe fakers.

**Key words:** social desirability; self – deception; impression management; BIDR scale