

# 大学生自我报告的学习投入可靠吗\*

——大学生群体的社会称许性反应及对自陈式问卷调查的影响

郭菲 赵琳 连志鑫

(清华大学教育研究院,北京 100084)

**摘要:**以“学”为中心的高等教育质量评价理念,推动了国内大学生学情调查的广泛开展。此类调查使用自陈式问卷进行,其数据质量可能受填答者社会称许性反应的影响。使用清华大学2014-2016年“中国大学生学习与发展追踪研究”的数据研究发现:有关学习行为和主观感受的自陈性指标存在5-10%的社会称许性反应偏误,但自我汇报的客观情况(如学习成绩)偏误较小;学生的社会称许性水平越高,自陈性指标中的偏误越大;指标内容与社会称许性反应偏误的大小有一定关系。上述发现对大学生学情调查的科学性和工具开发的有效性具有方法论的意义。

**关键词:**社会称许性反应;自陈式问卷调查;大学生自我报告

## 一、问题提出

进入21世纪,高等教育质量评价与保障呈现出鲜明的转型特征:倾听学生声音,以“学”为中心的增值性过程评价,以“证据”为基础的可持续的评价方式,开始服务于高等院校的自我诊断与院校改进(赵琳,王传毅,2015)。上述转型既需要教育质量观、学生观、质量评价与保障观等理念上的转变与更新,也有赖于评价方法和测量工具的发展。近年来,各国纷纷开展学情调查,如全美大学生学习投入调查(NSSE)、澳大利亚大学生学习投入调查(AUSSE)、清华大学的“中国大学生学习与发展追踪研究”(CCSS)、厦门大学的“国家大学生学情调查”(NCSS)等,研究者尝试通过问卷调查来搜集学生的学习体验数据,并用以诊断和改进高等教育质量。这些研究需要以数据的真实性、有效性和可靠性为基本前提。

影响数据质量的一个重要因素是填答者的社会称许性反应(Socially Desirable Responding),它指的是人们出于获得社会赞许和接受的需要,按照社会赞许的方式而非个人实际情况来填答问卷,即做出过度正向自我描述的倾向(e.g. Zerbe & Paulhus, 1987)。由此造成的数据偏误被称为“社会称许性反应偏误(Social Desirability Response Bias)”(以下简称“社会称许性偏误”)。该偏误给自陈式问卷调查的数据质量带来较大威胁,因此对其进行识别、测量和控制一直受到社会科学研究者的关注。

本研究希望通过对清华大学“中国大学生学习与发展追踪研究”(以下简称CCSS调查)2014-2016年的数据分析,回答以下研究问题:在填答自陈性问卷时,我国大学生中不同群体的社会称许性反应如何?在学情调查中,基于自陈性数据构建的指标得分中存在多大的社会称许性偏误?以上两个问题关乎学情调查的科学性,因此本研究具有方法论上的意义;从工具开发、问卷与数据的精细化分析的

\* 基金项目:教育部哲学社会科学研究重大课题攻关项目“双一流建设背景下我国高校评价体系改革研究”(项目号:17JZD056)。

角度看,本研究也兼具重要的实践应用意义。

## 二、文献综述

上世纪50年代,心理学家在人格测验中提出了社会称许性问题(e.g. Cowen & Tongas, 1959; Edwards, 1953; Edwards & Horst, 1953; Fordyce, 1956; Marlow & Crowne, 1961),开发了相关量表(e.g. Crowne & Marlowe, 1960; Edwards, 1953; Wiggins, 1959),并对其理论维度进行了探索(e.g. Damarin & Messick, 1965; Paulhus, 1984, 1986; Sackeim & Gur, 1978)。在前人研究的基础上,Paulhus (2002)总结提出了社会称许性的双层理论构架(图1)。他认为,社会称许性反映了“自我中心趋向”和“道德趋向”两类人格特质,二者当中都有“自我欺骗”(即个体出于对自我信念的自我保护而下意识进行的心理调整)的因素和“形象管理”(即个体为了给他人留下良好印象而有意识地对他人进行欺骗)的因素。其中,“自我中心趋向”反映了人们对自身能力品质的过高认识,包括对自我认知中下意识的过高判断(自我欺骗提高)和在自我展示时(如应聘环境下)对自身能力的刻意过度描述(能动性形象管理);“道德趋向”则指人们出于对社会道德、规范、规则等的遵从,而否认自己有负面行为的倾向,包括下意识地否认自己反社会规则的冲动(自我欺骗否定)和在面对他人时有意识地掩饰自己的错误行为等(交流性形象管理)。

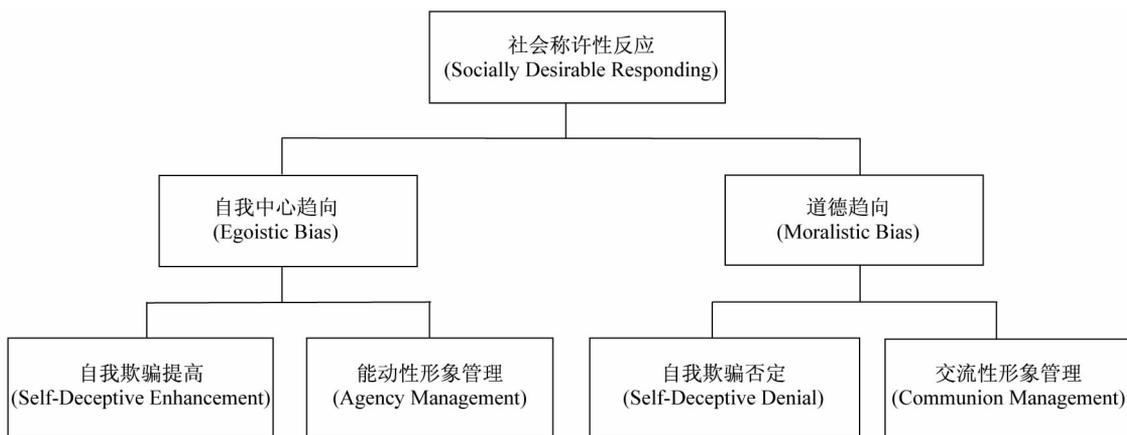


图1 Paulhus 的社会称许性反应双层结构 (Paulhus, 2002, p. 64)

根据以上概念构架,自陈性问卷调查中发生社会称许性反应有三个可能机制:第一,被调查者本身具有较高的社会称许性(Paulhus, 1984);第二,测验情境(如调查目的、引导语等)激发了被调查者的社会称许性反应,这对“形象管理”类社会称许性反应的影响尤其明显(Paulhus, 2002; Stark, Chernyshenko, Chan, Lee, & Drasgow, 2001);第三,受测验/调查题目内容的影响,当题目涉及个人的行为、情绪和主观感受且社会对这些行为、情感具有一定的评判取向时,社会称许性反应会表现得更加明显(Saunders, 1991)。

利用经典 Marlowe-Crowne 社会称许性量表(M-C量表),美国学者对大学生学情调查中的社会称许性反应偏误进行了检验。NSSE 调查机构分析了2010年6所高校的3169名学生样本数据,结果表明社会称许性与主要可比指标、自我报告的学习成绩和对院校的整体评价不存在显著相关(NSSE, 2012)。Miller(2012)基于美国6所院校的2352名学生进行了类似的研究,结果发现社会称许性偏误仅存在于少部分指标(支持性校园环境和自我汇报教育收获)中,且偏误很小。Bowman和Hill(2011)以美国某大学2009年一门课程的171位大学生为有效样本,使用NSSE调查的数据发现,社会称许性与大学生自我报告的教育收获间不存在显著相关。总体而言,这些研究都证明美国的大学生学情调查中的社会称许性偏误并不严重。

在我国,社会称许性相关研究也得到越来越多学者的重视。如韩振华和任剑锋(2002)等建议从研究设计的角度解决社会称许性偏见效应。一些学者对 Marlow-Crowne 社会称许性量表(M-C 量表)进行了中国情境下的实用性检验和修订(刘萃侠,2001;吴燕,2008;甄育玲,2011;赵丹,2011)。李强和徐晟团队基于 Paulhus 的双层模型,参考 M-C 量表和期待性回答平衡问卷(BIDR),结合我国社会学、心理学等领域专家的访谈,制定了《中国成年人多维社会赞许性量表》(李强,徐晟,李凌,黄艳,徐旻,2012),并在此基础上使用全国分层抽样样本建立了全国常模(徐晟等,2014)。蒋华林等(2014)利用2013年某“985”院校的“中国大学生学习与发展研究”数据,将自我报告的学习成绩与学校给出的真实 GPA 成绩做对比研究,发现社会称许性与学情投入调查结果存在一定的相关性。但是,国内现有研究均没有对社会称许性反应在自陈性问卷调查中的影响程度作出估计。

本研究借鉴 Miller 等人的研究方法,检验中国大学生自陈性问卷调查中的社会称许性偏误。研究提出以下假设:

假设1(存在性假设):学生填答问卷时存在社会称许性反应,即社会称许性水平越高,自陈性指标得分越高;

假设2(非线性假设):社会称许性偏误大小与学生社会称许性水平呈正相关关系,即社会称许性水平越高,自陈性指标中的偏误越大。

### 三、研究方法

#### (一)调查工具

本研究使用的调查工具为“CCSS 调查普通本科问卷”。该问卷重点关注大学生在校期间的学习性投入,通过其自我报告的课内外学习体验、学习态度、人际互动、教育收获和满意度等来诊断院校的教育教学过程。为控制调查中的社会称许性偏误,CCSS 问卷加入了测量社会称许性的专门量表。该量表以 M-C 量表为基础,结合中国大学生心理发展特点和生活语境修订而成,题项内容涵盖了 Paulhus 的双层模型中的各个维度,备选答案为“是”或“否”。计时符合社会称许的答案记为 100 分,不符合的记为 0 分,各题项的平均分为社会称许性指标得分。2014-2016 年量表信度及题项举例见表 1。

表 1 CCSS 2014-2016 社会称许性指标构成及信度

	2014	2015	2016	例题
题项数	6	7	7	我从来没发过脾气(自我欺骗提高)
Cronbach's alpha	0.83	0.62	0.63	我从未失信于人(能动性管理)
观测值数	58,541	67,676	79,320	我从来没有讨厌过谁(自我欺骗否定)
均值(方差)	59.07(17.96)	67.08DG(25.40)	52.57(22.36)	只要是我的错,我都会承认(交流性管理)

为检验社会称许性指标的效度(即指标得分能否反映学生做出过度正向自我描述的倾向),本研究采用自我汇报高考成绩溢出值(即自我汇报成绩高于真实成绩的程度)作为校标进行检验。根据社会称许性指标得分分布的四分位点将样本分组,分别为低水平组(社会称许性得分处于整体分布的 1-25%(含)分位),中低水平组(25-50%(含)分位)和中高水平组(50-75%(含)分位),高水平组(75-100%(含)分位)。比较各组学生的真实高考成绩、自我汇报成绩和自我汇报溢出值,并检验各组中自我汇报成绩偏高(高于真实成绩 5 分)、偏低(低于真实成绩 5 分)、一致(与真实成绩相差 5 分及以内)的比例,结果如表 2。社会称许性水平越高的组,学生的平均高考成绩越低,但自我汇报溢出值却越大。同时,社会称许性水平越高的组,自我汇报成绩偏高的比例越大,而偏低的比例则没有明显变化。对自我汇报溢出值与社会称许性指标做相关分析,相关系数均为正,且在 0.001 的水平上显著。以上分析说明社会称许性指标得分高的学生更倾向于高报自己的高考成绩,即具有做出过度自我描述的倾向,社会称许性指标的效度得以证实。<sup>①</sup>

表2 社会称许性四分位组与自我汇报高考成绩差异

社会称许性四分位组	样本数	真实成绩	自我汇报成绩	自我汇报差值	自我汇报偏高比例	自我汇报偏低比例	自我汇报一致比例
低水平组	24,967	536.97(92.73)	540.59(90.17)	3.62(20.43)	14.26%	6.65%	79.09%
中低水平组	22,307	527.79(96.43)	531.98(93.49)	4.19(21.05)	15.59%	6.15%	78.27%
中高水平组	15,530	520.15(98.82)	525.15(95.65)	5.00(21.56)	18.65%	5.87%	75.47%
高水平组	17,106	510.52(101.63)	517.00(98.54)	6.47(23.90)	21.68%	6.06%	72.26%
总体	79,910	525.48(97.40)	530.14(94.41)	4.66(21.63)	17.07%	6.23%	76.70%

注:1. 括号中为变量标准差。

2. 自我汇报溢出值与社会称许性指标相关性检验: Pearson's  $r$ : 0.0472 ( $p < 0.001$ ), Spearman's  $r_s$ : 0.0653 ( $p < 0.001$ ), Kendall's Tau-a: 0.0391 ( $p < 0.001$ ), Kendall's Tau-b: 0.0498 ( $p < 0.001$ )。

## (二) 样本构成

2014-2016年参与CCSS调查的普通本科院校共计56所,其中“985”院校11所,非“985”的“211”院校13所,地方院校32所。院校分布于全国20个省,其中京津沪16所,其他东部地区(含东北)18所,中西部地区16所。院校样本在院校类型和地区分布上具有一定代表性。在参与院校中随机抽取每年级400-800名学生参与调查,2014-2016年样本总计221,591人,各年的问卷回收率分别为67.22%, 59.41%, 62.63%。剔除测谎题答案相差2分及以上或全部计分题项缺失超过2/3及以上的无效问卷后,共余有效问卷206,847份。表3展示了分析样本的人口学特征。

表3 CCSS 2014-2016 样本分布

	分布	样本数	比例(%)
学校类型	“985”高校	42,888	19.35
	“211”高校	42,890	19.36
	地方本科大学	99,360	44.84
	地方本科院校	36,453	16.45
性别	男	120,066	54.18
	女	101,494	45.8
民族	汉族	202,031	91.17
	少数民族	19,560	8.83
年级	大一年级	65,741	29.67
	大二年级	60,745	27.41
	大三年级	56,099	25.32
	大四年级	39,006	17.6
学科	人文学科	27,097	12.23
	社会科学	47,539	21.45
	自然科学	36,661	16.54
	工程类学科	107,095	48.33

## (三) 分析方法

本研究使用多元线性回归方法研究社会称许性反应对自陈性指标的影响。采用的回归模型为:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Year_i + \varepsilon_i \quad (1)$$

$$Y_i = \beta_0 + \beta_1 SD_i + \beta_2 X_i + \beta_3 Year_i + \varepsilon_i \quad (2)$$

$$Y_i = \beta_0 + \beta_1 SD\_quartile_i + \beta_2 SD\_i * SD\_quartile_i + \beta_4 X_i + \beta_5 Year_i + \varepsilon_i \quad (3)$$

各模型中,  $Y_i$  为CCSS调查的15个核心诊断指标,包括学业挑战度、师生互动、有效教学实践、学习态度、学习行为、教育收获和在校满意度等指标。此外,还选择了学生自我汇报的上一学年学分成绩(GPA)和百分制平均成绩作为因变量。 $X_i$  为人口学控制变量,包括性别、民族、学校类型、年级、学科。

$SD_i$  为各年标准化后的社会称许性指标得分。 $SD\_quartile_i$  为各年社会称许性指标得分的四分位分组。此外,为控制各调查年的特殊影响,模型中加入了年份固定效应,即  $Year_i$ 。所有模型均使用最小二乘法(OLS)进行估计。为纠正同院校样本的残差相关性,回归时以学校为单位对标准误进行了聚集。

模型 1 为基础模型,仅以人口学背景变量为解释变量。模型 2 在此基础上加入了社会称许性指标,其系数  $\beta_1$  估计值反映了控制人口学变量之后,社会称许性指标对自陈性指标的影响。模型 1 到模型 2 的  $R^2$  变化值,反映了在人口学背景之上,自陈性指标得分中的变异在多大程度上可被社会称许性解释,即社会称许性反应偏误的大小,以此验证假设 1。模型 3 将社会称许性标准化得分替换为得分的四分位分组,并加入了四分位分组的与标准化得分的交互项。该交互项的系数( $\beta_2$ )反映了在不同社会称许性水平上,社会称许性反应对自陈性指标的影响,通过分析各水平分组中  $\beta_2$  的变化来验证假设 2。

## 四、研究结果与发现

### (一) 我国大学生的社会称许性水平

根据社会称许性指标得分四分位分组检验我国大学生不同群体中的分布情况,结果如表 4 所示。

表 4 社会称许性四分位组在我国大学生不同群体中的分布

类别	社会称许性四分位组				总计	
	低水平组	中低水平组	中高水平组	高水平组		
性别	男	26.47%	27.80%	21.15%	24.59%	100.00%
	女	31.60%	28.98%	20.03%	19.40%	100.00%
年级	大一年级	31.83%	29.12%	19.46%	19.59%	100.00%
	大二年级	30.49%	28.68%	20.19%	20.64%	100.00%
	大三年级	27.71%	28.50%	21.10%	22.69%	100.00%
	大四年级	22.84%	26.27%	22.63%	28.26%	100.00%
学科	人文学科	28.57%	28.26%	20.31%	22.86%	100.00%
	社会科学	27.71%	28.45%	21.10%	22.74%	100.00%
	自然科学	31.57%	28.89%	20.06%	19.47%	100.00%
院校类型	工程类学科	27.97%	28.22%	20.84%	22.96%	100.00%
	985 院校	34.12%	28.78%	18.55%	18.56%	100.00%
	211 院校	28.76%	27.95%	21.07%	22.21%	100.00%
	大学本科	28.59%	28.48%	20.64%	22.29%	100.00%
	学院本科	23.48%	27.91%	22.51%	26.09%	100.00%

男生中社会称许性高水平及中高水平组占比约 46%,女生约 39.7%,说明我国大学生中,男生社会称许性水平高于女生。随着年级的升高,社会称许性高水平、中高水平组的占比不断提高,说明学生的社会称许性水平随年级逐步升高。各学科学生的社会称许性水平差异不大,相比之下自然科学学科学生社会称许性较低,高水平及中高水平组占比最低,为 39.5%,其他学科约为 43-44%。不同院校类型学生表现出一定的差异:“985”院校学生中社会称许性高水平、中高水平组占比 37%,非“985”的“211”院校和地方大学占比约 43%,地方学院本科占比约 49%。可见,我国地方学院本科学生的社会称许性水平相对较高。

### (二) 自陈性指标中的社会称许性偏误

对模型 2 中  $\beta_1$  系数的回归估计结果显示(见表 5),在控制了人口学变量和调查年固定效应的情况下,社会称许性指标在所有以自陈性指标为因变量的模型中系数均为正且统计学显著。标准化回归系数显示,社会称许性指标得分每增加 1 个标准差,自陈性指标得分随之增加 0.14-0.33 个标准差,且 15 个指标中的 12 个得分增加值在 0.2 个标准差以上。说明社会称许性对自陈性指标存在显著且有一定实际意义的影响。相比之下,对于客观性指标即学习成绩,社会称许性影响较小。

表5 社会称许性指标对自陈性指标的影响

因变量	模型2 $\beta_1$ 估计结果			模型比较			
	回归系数	标准误	标准化回归系数	样本数	模型1 $R^2$	模型2 $R^2$	$R^2$ 增加值
学业挑战度(LAC)	3.107***	0.093	0.253***	202409	0.059	0.122	0.063
主动合作学习水平(ACL)	5.123***	0.165	0.275***	202,395	0.023	0.097	0.074
生师互动(SFI)	6.910***	0.209	0.324***	202,409	0.080	0.183	0.103
教育经验的丰富程度(EEE)	3.967***	0.144	0.228***	202,408	0.080	0.131	0.051
校园环境的支持度(SCE)	4.324***	0.118	0.263***	202,130	0.063	0.130	0.067
教学-课堂教学(ETP_CI)	3.020***	0.084	0.170***	202,409	0.015	0.043	0.028
教学-教师反馈(ETP_FB)	4.326***	0.088	0.239***	202,409	0.021	0.077	0.056
教学-激发学习志趣(ETP_MS)	3.758***	0.110	0.190***	143,892	0.017	0.052	0.035
教学-测评(多元评价)(ETP_AS)	5.028***	0.231	0.207***	143,802	0.056	0.098	0.042
向学/厌学(SWL)	3.793***	0.072	0.255***	202,144	0.017	0.081	0.064
高阶学习(HOC_HOL)	5.253***	0.148	0.244***	202,333	0.013	0.072	0.059
整合性学习(LS_IL)	5.530***	0.117	0.285***	202,408	0.027	0.106	0.079
反思性学习(LS_RL)	4.871***	0.097	0.254***	202,347	0.016	0.079	0.063
自我报告的教育收获(SSLO)	6.262***	0.117	0.333***	201,944	0.022	0.131	0.109
在校满意度(SSTF)	2.775***	0.137	0.137***	201,527	0.021	0.039	0.018
自我报告 GPA	0.021**	(0.008)	0.036**	134,396	0.073	0.075	0.002
自我报告平均成绩(百分制)	0.077	(0.046)	0.008	151,413	0.073	0.074	0.001

注: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ; 标准误为以学校为单位聚集后的稳健标准误; 模型控制了人口学变量和调查年固定效应; 缺失值采用 pair-wise deletion 的方法处理, 各方程因此被删除的样本为总样本的 1.5 - 2.7%。

对比加入社会称许性指标前后(即模型 1、2)的模型解释力即  $R^2$  发现, 以自陈性指标为因变量时, 模型 2 相比于模型 1 的  $R^2$  增加幅度为 0.018 到 0.11, 其中 11 个指标增幅大于 0.05, 2 个指标(生师互动、自我汇报的教育收获)大于 0.1。这说明在人口学背景之上, 社会称许性指标可以解释大部分自陈性指标的 5 - 10% 的变异性。但在以学习成绩为因变量的模型中, 不论是 GPA 还是百分制平均成绩, 社会称许性指标在人口学背景之上的贡献仅为 0.1 - 0.2%, 再次证明学生自我报告的学习成绩受社会称许性影响很小。

由上可知, 社会称许性指标对自陈性指标有显著且不小的正向影响, 本文假设 1 得以证实。

### (三) 社会称许性偏误的非线性特征

为检验社会称许性偏误是否存在非线性特点, 模型 3 加入了社会称许性得分四分位分组变量, 并将之与社会称许性指标得分进行交互。表 6 汇报了对模型 3 中  $\beta_2$  系数的回归估计结果。

首先, 在各社会称许性水平分组中, 社会称许性指标对自陈性指标的回归系数都显著为正, 说明在各水平上都存在一定的社会称许性偏误。其次, 效应量显示, 在社会称许性水平最低的组中, 社会称许性对大部分自陈性指标的影响在 0.2 个标准差以下, 不超过 0.3 个标准差, 说明影响几无实质意义; 在水平中等的两个组中, 达到 0.2 到 0.35 之间, 说明有一定实质影响但仍较小; 而在水平最高的组中, 效应量大多在 0.35 以上, 有 8 个变量达到 0.45 以上, 最高的达到了 0.72(生师互动指标), 说明有中等水平的影响, 同时有 7 个指标的效应量相比于前几组提高了一倍以上, 说明自陈性指标中的社会称许性偏误存在非线性特点, 即社会称许性水平越高, 偏误越大, 本研究假设 2 得到验证。相比之下, 在以学习成绩为因变量的模型中, 各水平分组中的社会称许性偏误都较小, 可忽略不计。

表6 社会称许性指标对自陈性指标的影响,按社会称许性水平分组

因变量	低水平组(1-25%含)			中低水平组(25-50%含)			中高水平组(50-75%含)			高水平组(75-100%含)		
	回归系数 <sup>a</sup>	标准误	效应量 <sup>b</sup>									
学业挑战度(LAC)	2.199***	(0.135)	0.179	2.226***	(0.167)	0.181	2.385***	(0.383)	0.194	5.782***	(0.279)	0.471
主动合作学习水平(ACL)	3.460***	(0.202)	0.186	3.984***	(0.314)	0.214	3.929***	(0.636)	0.211	10.420***	(0.499)	0.560
生师互动(SFI)	3.248***	(0.226)	0.152	3.695***	(0.505)	0.173	2.095*	(0.842)	0.098	15.412***	(0.611)	0.722
教育经验的丰富程度(EEE)	2.149***	(0.126)	0.123	3.270***	(0.243)	0.188	3.739***	(0.424)	0.215	11.511***	(0.686)	0.661
校园环境的支持度(SCE)	3.736***	(0.217)	0.227	4.843***	(0.318)	0.294	3.498***	(0.591)	0.213	6.489***	(0.362)	0.395
教学-课堂教学(ETP_CI)	3.293***	(0.217)	0.186	3.952***	(0.355)	0.223	4.398***	(0.497)	0.248	4.051***	(0.403)	0.229
教学-教师反馈(ETP_FB)	3.154***	(0.179)	0.174	3.918***	(0.307)	0.216	3.075***	(0.479)	0.170	7.234***	(0.376)	0.399
教学-激发学习志趣(ETP_MS)	3.549***	(0.270)	0.179	3.221***	(0.477)	0.163	3.778***	(1.072)	0.191	3.899***	(0.479)	0.197
教学-测评(多元评价)(ETP_AS)	2.127***	(0.302)	0.088	2.687***	(0.658)	0.111	-2.509	(1.386)	-0.103	12.140***	(0.590)	0.501
向学/厌学(SWL)	3.682***	(0.125)	0.247	4.655***	(0.313)	0.313	7.091***	(0.389)	0.476	5.516***	(0.348)	0.371
高阶学习(HOC_HOL)	3.769***	(0.262)	0.175	4.937***	(0.328)	0.229	3.531***	(0.568)	0.164	10.042***	(0.563)	0.467
整合性学习(LS_IL)	3.963***	(0.193)	0.204	4.740***	(0.345)	0.244	5.760***	(0.512)	0.297	10.706***	(0.545)	0.551
反思性学习(LS_RL)	3.955***	(0.194)	0.206	4.739***	(0.294)	0.247	6.288***	(0.453)	0.328	8.258***	(0.501)	0.431
自我报告的教育收获(SSLO)	5.155***	(0.186)	0.274	6.003***	(0.316)	0.319	9.181***	(0.510)	0.488	10.143***	(0.540)	0.540
在校满意度(SSTF)	3.622***	(0.234)	0.178	3.820***	(0.314)	0.188	3.304***	(0.858)	0.163	0.621	(0.544)	0.031
自我报告GPA	0.017*	(0.007)	0.029	-0.018	(0.015)	-0.031	-0.007	(0.023)	-0.012	0.033**	(0.012)	0.056
自我报告平均成绩(百分制)	0.151	(0.093)	0.016	0.089	(0.249)	0.009	-0.203	(0.313)	-0.021	-0.123	(0.189)	-0.013

注：\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ; 标准误为以学校为单位聚集后的稳健标准误; 模型控制了人口学变量和调查年固定效应; 缺失值采用 pair-wise deletion 的方法处理, 各方程因此被删除的样本为总样本的 1.5-2.7%。

1. 表中汇报的是对模型 3 中  $\beta_2$  系数的回归估计。

2. 效应量为 Cohen's D 类型效应量, 计算方法为回归系数/因变量标准差, 由于自变量(社会称许性指标)已做标准化处理, 其意义为自变量变化 1 个标准差所引起的因变量标准差的变化, 与表 5 中标准化回归系数的意义相同。

## 五、结论与讨论

大学生学情调查为学生提供了发声的途径, 为院校诊断并改进教育教学提供了来自一线的证据。但由于以自陈性问卷为工具, 此类调查的结果不可避免地受到社会称许性的影响。我们使用 CCSS 调查了 2014-2016 年 56 所普通本科院校 20 多万学生样本, 研究发现: (1) 男生、大四学生、社会科学和工科专业的学生、非 985 院校的学生社会称许性水平较高。(2) 社会称许性反应对自陈性指标造成的偏误约为 5-10% 左右。(3) 对大部分指标来说, 社会称许性偏误主要存在于高社会称许性水平的学生当中, 水平中等或偏低的学生所汇报的指标虽然也受到一定影响, 但程度较轻。在本批样本中, 高社会称许性水平组(即 75% 百分位及以上)对应的得分分别为 2014 年 66.7 分, 2015 年 85.7 分, 2016 年 62.5 分。由此可推断, CCSS 调查中社会称许性指标得分高于 60 分的答题者, 其自陈性指标得分较实

际情况偏高的可能性很大,在分析时需要格外注意。(4)学生汇报的学习成绩中存在的社会赞许性偏误较小,数据基本可信。

本研究进一步发现,问卷题目内容与社会赞许性偏误的大小存在一定关系,这与 Saunders (1991) 等的发现一致。涉及对自身、老师、学校的评价且社会赞许倾向明显的题目中存在的社会赞许性偏误较大。如偏误最大的“自我汇报的教育收获”和“生师互动”指标( $R^2$  增加值分别为 0.11 和 0.10),一为对自身能力提高程度的评价,一为对社会赞许行为的自我报告。各学习行为指标和学习态度指标中的偏误均在 6% 以上;涉及学校整体情况的题目,如“学业挑战度”和“校园环境支持度”等,社会赞许性偏误也均高于 6%。而社会赞许倾向不太容易被识别的题目,如反映课程考评方式的“教学-测验”指标,以及实际情况普遍较好的指标(如反映教师是否清晰合理进行教学的“教学-课堂教学”和反映教师是否鼓励学生主动学习的“教学-激发学习志趣”指标),受社会赞许性的污染则较小。此外,社会赞许倾向清晰可辨的指标中,社会赞许性偏误在全体学生中普遍存在,如表 6 中,社会赞许性对“自我汇报的教育收获”、“向学/厌学”等指标得分影响的效应量在各水平分组中均大于 0.2。而社会赞许倾向不明显的指标,尽管总体偏误较小,但在高社会赞许性水平组中仍存在较大的偏误,如社会赞许性对“教学-测验”影响的效应量在社会赞许性中等及偏低的三组中均小于 0.15,没有实际意义的影响,但在高水平组中达到了 0.5,属于中等水平的影响。这说明,即使对于社会赞许倾向不明显的题目,仍然要警惕高社会赞许性水平学生做出的反应偏误。

与美国学者的研究发现相比,本研究发现的社会赞许性偏误存在范围更广、程度也更高。这可能是由于中美文化存在一定差异。有比较研究发现,中国被试相比于美国被试,尽管自吹自擂的倾向较低,但更重视集体主义,以及社群对个人的期望;且中国学生尽管在公开场合表现得更为谦逊,但在匿名情况下对自己能力的评价并不低于美国学生(赵志裕,邹智敏,林升栋,2010)。一项基于四大洲 26 国消费者市场调查的研究发现,中国消费者在社会赞许性指标上,不论是“自我中心趋向”还是“道德趋向”,得分都高于美国和大多数国家,其中“道德趋向”高于所有国家(Steenkamp, de Jong, & Baumgartner, 2010)。这些发现可以在一定程度上解释本研究所发现的大学生对自身和学校评价时容易根据社会赞许的方式作出回答的情况。

综上,我国大学生学情调查中存在的社会赞许性偏误不容忽视。因此,在设计学情调查问卷时,应尽量减少带有价值判断色彩的题项,以减少学生的社会赞许性反应。问卷中应考虑加入社会赞许性相关量表,以方便在事后识别高社会赞许性水平的学生、控制社会赞许性的影响。在使用学情调查数据进行分析,尤其是进行群体间比较时,应考虑不同群体学生社会赞许性水平的差异,以免因未控制社会赞许性偏误而做出错误的判断。

## 参考文献

- 韩振华,任剑峰.(2002). 社会调查研究中的社会赞许性偏见效应. *华中科技大学学报(社会科学版)*, 16(3), 47 - 50.
- 蒋华林,邢全超,吴芳,朱晓华.(2014). “社会赞许性”影响大学生学情调查的实证研究. *大学教育科学*, 6(6), 105 - 111.
- 李强,徐晟,李凌,黄艳,徐旻.(2012). 心理健康素质测评系统·中国成年人多维社会赞许性量表的编制. *心理与行为研究*, 10(4), 255 - 261.
- 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.
- 吴燕.(2008). 人格测验中社会赞许性反应的测定与控制. 陕西师范大学.
- 徐晟,李强,李凌,尹艺璇,梁栋.(2014). 心理健康素质测评系统·中国成年人社会赞许性全国常模的制定. *心理与行为研究*, 12(6), 748 - 755.
- 赵丹.(2011). 马洛-克罗恩社会赞许性量表的修订及相关研究. 浙江师范大学.
- 赵琳,王传毅.(2015). 以“学”为中心:研究生教育质量评价与保障的新趋势. *学位与研究生教育*, (3), 11 - 14.
- 赵志裕,邹智敏,林升栋.(2010). 文化与社会赞许反应:社会个人互动的观点. *心理学报*, 42(1), 48 - 55.
- 甄育玲.(2011). 社会赞许性反应量表的修订. 湖南师范大学.
- Bowman, N. A., & Hill, P. L. (2011). Measuring how college affects students: Social desirability and other potential biases in college student

- self – reported gains. *New Directions for Institutional Research*, 2011(150), 73 – 85.
- Cowen, E. L., & Tongas, P. N. (1959). The social desirability of trait descriptive terms: Applications to a self – concept inventory. *Journal of Consulting Psychology*, 23(4), 361 – 365.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349 – 354.
- Damarin, F., & Messick, S. (1965). *Response styles as personality variables; A theoretical integration of multivariate research*. ETS Research Bulletin Series, 1965(1), 1 – 116.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37(2), 90 – 93.
- Edwards, A. L., & Horst, P. (1953). Social Desirability as a Variable in 2 Technique Studies. *Educational and Psychological Measurement*, 13(4), 620 – 625.
- Fordyce, W. E. (1956). Social desirability in the MMPI. *Journal of Consulting Psychology*, 20(3), 171 – 175.
- Marlow, D., & Crowne, D. P. (1961). Social desirability and response to perceived situational demands. *Journal of Consulting Psychology*, 25(2), 109 – 115.
- Miller, A. L. (2012). Investigating social desirability bias in student self – report surveys. *Association for Institutional Research*, 36(1), 30 – 47.
- NSSE. (2012). Are NSSE scores influenced by a desire to respond in a socially desirable manner? Retrieved August 23, 2017, from [http://nsse.indiana.edu/pdf/psychometric\\_portfolio/2010\\_social\\_desirability\\_formatted.pdf](http://nsse.indiana.edu/pdf/psychometric_portfolio/2010_social_desirability_formatted.pdf).
- Paulhus, D. L. (1984). Two – component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 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: Springer Berlin Heidelberg.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In I. Brown, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and education measurement* (pp. 49 – 69). Mahwah, NJ: Erlbaum.
- Sackeim, H. A., & Gur, R. C. (1978). *Self – Deception, Self – Confrontation, and Consciousness*. Springer US.
- Saunders, D. G. (1991). Procedures for adjusting self – reports of violence for social desirability bias. *Journal of Interpersonal Violence*, 6(3), 336 – 344.
- Stark, S., Chernyshenko, O. – S., Chan, K. – Y., Lee, W. – C., & Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology*, 86(5), 943 – 953.
- Steenkamp, J. – B. E. M., de Jong, M. G., & Baumgartner, H. (2010). Socially Desirable Response Tendencies in Survey Research. *Journal of Marketing Research*, 47(2), 199 – 214.
- Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instruction. *Journal of Consulting Psychology*, 23(5), 419 – 427.
- Zerbe, W. J., & Paulhus, D. L. (1987). Socially Desirable Responding in Organizational Behavior: A Reconciliation. *The Academy of Management Review*, 12(2), 250 – 264.

## 注 释：

①本研究样本中同时具有真实高考成绩和自我报告高考成绩样本共 86,764 个,为了消除奇异值的影响,将非零高考成绩差分布前后各 5% 的观测值删除,共删除了 2.5% 的样本。再删除社会称许性变量缺失的个案,最终用于表 2 效度分析的样本数为 79,910 个。

(责任编辑 胡 岩)

the influence of family background, employment expectation and experience of senior high school as explanatory variables on teenagers' international competency. Moreover, this paper attempts to explore whether highly competitive middle schools play a mediating role between family background and international competency. According to the results, family background, employment expectation and experience of senior high school significantly affect teenagers' international competency.

**Keywords:** global competence; international competency; family background; employment expectation; influencing factor

## **Reliability of Self-reported Data in College Student Engagement Surveys: Social Desirability Bias in Self-reported Survey**

GUO Fei ZHAO Lin LIAN Zhixin

(Tsinghua University, Institute of Education, Beijing 100084, China)

**Abstract:** Inspired by learner-centered and evidence-based evaluation in higher education quality assurance, investigations on college student experience and engagement become increasingly popular in China. However, with self-reported survey as a major instrument, the quality of data collected in such investigations is vulnerable to bias caused by socially desirable responses. Using data from the China College Student Survey (CCSS) in 2014-2016, the authors found that indicators constructed from student self-reported data on learning activities and perceptions are subject to a social desirability response bias of 5-10%, while the bias in objective information reported in the survey (such as GPA) is negligible. What's more, the higher the students' social desirability level, the larger the bias. It's also found that the size of social desirability bias shows its correlation with the content of the question items and indicators. These findings have methodological significance for quantitative investigations in social sciences. Further, implications for survey instrument development are discussed.

**Keywords:** self-reported survey; social desirability response bias; college students

## **Education, Background and Policy: The Allocation Mechanism of Minority Students' Access to Higher Education**

HUANG Yuheng SHI Jinghuan

(Institute of Education, Tsinghua University, Beijing 100084, China)

**Abstract:** Surveys on minority students' access to higher education tend to involve little pre-college experience and policy data. Based on existing theory and data from the China College Student Survey (CCSS, 2011-2016), this study developed a new framework to analyze minority students' access to higher education. Related factors are grouped into three dimensions: family background, pre-college educational experience and educational policy. Using the higher education stratification mobility rate, logistic regression model and the Sheaf Coefficients, the authors found that, with equal opportunity policy in place, pre-college educational experiences are the main factors influencing minority students' access to higher education, especially in elite u-