

社会化媒体调查与传统模式调查间的数据质量差异研究*

朱文龙 邵培基 方佳明

(电子科技大学经济与管理学院, 成都 610054)

摘要 论文从回复质量和社会期望作答两个视角研究了纸笔调查、网站调查、BBS 调查、微博调查以及即时通信调查间的数据质量差异。结果表明: 不同敏感程度的量表会对五组调查的数据质量产生不同影响; 同一量表下五组调查在回复质量方面存在显著性差异; 量表的敏感程度会对五组调查带来不同程度的社会期望偏差影响。本研究旨在指明五组调查间的模式差异对调查数据质量的影响, 以便调查者今后在整合混合模式数据时能够充分考虑可能存在的风险。

关键词 网络调查, 社会化媒体, 数据质量, 回复质量, 社会期望作答

中图分类号 C93

1 引言

社会化媒体(Social Media)是建立在 Web 2.0 理念和技术基础上的一系列能够促进个体共享信息、协作、创建并发展在线社区的软件工具集合^[1,2]。目前, 国内的社会化媒体主要包括社交网站(SNS)、微博、博客、在线论坛(BBS)、即时通信(IM)等。2012年6月, 中国社交网站用户规模达2.5亿, 即时通信用户规模已突破4.45亿^[3]; 2012年9月, 国内微博用户达3.27亿^[4], 这些现象都标志着社会化媒体作为一种新型信息服务模式已逐渐被网民所接受。

网络调查经过几十年的研究, 在现实中已得到广泛应用。运用混合调查模式进行数据收集已变得越来越普遍^[5,6]。然而, 许多研究却表明传统混合调查模式并不会产生较高的数据质量并减少总体调查误差^[7]。因此, 探寻一种能带来较好模式效果, 获得较高数据质量的调查模式已成为调查者关注的焦点。2008年, 社会化媒体的兴起和迅速成长为混合模式的设计提供了新的途径与平台。

由于社会化媒体具有社会性, 因此, 利用该平台实施调查有望获得更高的回复率。那么, 社会化媒体调查与传统模式调查所得结果在数据质量层面是否存在显著差异? 这是本研究所关注的问题。数据质量的优劣直接关系到数据的可用性及调查结果的可靠性, 如果调查者没有意识到不同模式下数据的适用性, 分析结果很可能会因调查模式的不同而产生偏差, 从而使我们得出错误的分析结论。

广义上讲, 调查模式间的数据质量差异主要由两类调查误差所引起: 非测量误差与测量误差。非测量误差是由样本涵盖率、调查回复率、选项回复率等方面的差异所带来的偏差; 而测量误差主要是指调查工具和数据收集过程所造成的偏差^[8]。本研究仅针对测量误差所造成的数据质量差异进行分析探讨, 旨在从作答结果层面来检验社会化媒体调查结果可否推广至一般总体。

* 基金项目: 国家自然科学基金(71101018), 中央高校基本科研业务费(ZYGX2012YB034, ZYGX2010J131)。
通信作者: 方佳明, 电子科技大学经济与管理学院, 副教授。E-mail: jmfang@uestc.edu.cn。

本研究基于创新精神量表(GI)、个人信息技术创新量表(PIIT)和称许性平衡量表(BIDR),从回复质量(Response Quality)和社会期望作答(Socially Desirable Responding)两个维度来对纸笔调查、网站调查、BBS 调查、微博调查以及即时通信调查间的数据质量差异进行探索性分析,明确各平台的调查模式效果,并对所得结果进行讨论。

本文研究成果能够为后续研究者在网络调查领域进行其他平台的数据质量分析提供参考与指导;现实调查中,若混合调查模式中存在基于社会化媒体平台的调查,则调查者可参考本文分析结果,充分了解运用不同类型量表实施调查时所产生的问题与偏差,并充分考虑整合混合模式数据时所面临的潜在风险;现阶段有越来越多的公司与企业基于社会化媒体进行内部或外部调查,本文研究结果能够为这些机构在实践调查时提供借鉴与参考。

2 研究现状与理论背景

2.1 文献回顾

数据质量是一个属性概念,反映了数据适合使用的程度,在调查领域表现为数据较高应用价值的特性^[9]。

目前,国外已有许多专家学者就该问题进行了研究探讨。Heerwegh 等人^[10,11]通过实证研究发现与网络调查相比,面对面调查的数据质量更高且更易受到社会期望偏差的影响;Fricker 等人的研究表明网络调查与电话调查在选项无差别数量、选项非响应数量以及作答完成时间三项指标上存在显著差异^[12];Hollbrook 等人^[13]分析了电话调查与面对面调查间的模式效果,结果表明面对面调查能够产生较好的回复质量;Newman 等人研究了敏感性问题对调查带来的影响,最终发现面对面调查较计算机辅助电话调查将会产生更频繁的社会期望回答^[14];Vannieuwenhuyze 等人通过实证对比研究发现网络调查、电话调查以及面对面调查的混合设计较单一模式调查会产生更好的选择效果与测量效果^[7];Dillman 等人对以邮件调查、电话调查、交互式音频调查和互联网调查构成的混合模式调查进行了深入研究,最终发现两阶段模式调查能够有效提升被试作答,但它并不能有效减小非响应误差^[15]。

国内研究方面,截至 2012 年 10 月 8 日,对 CNKI、万方、维普数据库检索主题和关键词,发现国内仅有 4 位学者在网络调查领域从事该方面研究。樊茗玥基于数据误差产生机理,对网络调查中的数据质量问题进行了相关分析与研究,包括指标体系的建立、误差修正技术的改进、控制方法的提出等^[16];王瑜对计算机辅助调查中敏感问题的数据质量进行了回归分析^[17];刘全等基于质量评价指标体系,运用层次分析法和模糊评价法对数据质量划分了评价等级^[18];杜婷等评估了网络抽样调查中数据的信度与效度^[19]。

2.2 理论背景

2.2.1 回复质量

依据 Krosnick 的满意理论,我们可知被试产生的回复质量与其在作答时所付出的认知努力程度有关^[20]。该理论认为,被试在作答时产生的满意倾向会受到题项难易程度、被试作答能力和被试作答动机三个方面的影响。在本研究涉及的两类调查模式中,我们认为只有被试作答动机水平会导致作答回复质量的差异。由于社会化媒体具有社会性,且其背后代表了圈群化的人际关系网络,因此基于该平台实施调查时,这种人际关系模式能够增加被试的作答动机水平,增强其认知努力并减少满意行

为,使被试们更倾向于花更多时间与精力来回答每一道问题,从而对回复质量产生影响。

依据社会交换理论,个人行为会受到其行为期望获得回报的影响。在调查领域,Dillman 认为有三个因素会影响到被试的作答行为:奖励、成本及信任^[5]。奖励是指期望获得的好处,成本指为获取奖励不得不放弃或花费的资源,而信任则是指从长远来看将会超过成本的奖励或好处。在本研究的实证调查中,只有信任这个因素会对被试调查结果产生影响。根据 Holbrook 的观点,“社会距离”有助于信任的发展,而信任的建立则有利于产生较好的调查效果^[13]。在具有一定社会特性的社会化媒体中,人与人之间的“社会距离”更近,被试对于调查者身份的认知会增强其对调查者的信任,从而会更用心地回答问题。此时,这种作答方式将会对回复质量带来影响。

2.2.2 社会期望作答

社会期望作答指被调查者总倾向于给出一个使其显示出良好形象的答案^[21],这种倾向会对调查效果产生偏差。在社会化媒体中,人与人之间具有不同程度的熟悉水平,因此大家彼此间的隐私就显得尤为重要。基于该平台实施调查时,调查者是以一种更为直接、私人的方式来联系被试,这意味着被试者匿名水平的降低。根据 Cho 等人的观点,匿名水平的下降会在潜在程度上导致调查效果的降低及作答结果的差异^[22]。如果匿名程度较低,那么被试在作答过程中可能会对其隐私更加敏感,且作答行为更易受到社会期望倾向的影响,尤其在面对敏感问题时,这种倾向将会更加明显^[23]。

Newman 等人的观点也表明,被试者在调查中披露的个人信息量与调查的匿名与隐私程度成正比^[14]。社会化媒体作为一个虚拟社交平台,人与人之间较高的熟悉水平会对被试产生一种压力,使其作答行为更易符合其感知到的社会行为标准,包括维持社会尊重、获得社会支持、保护隐私、维护社会期望形象等。此时,这种无意识的作答行为所带来的社会期望偏差会导致一定程度的测量误差,从而对调查效果产生影响。

2.3 评述

通过文献回顾,国外网络调查领域对数据质量问题的探索尚处于蓬勃发展阶段,研究成果具有较强的现实意义,研究视角主要基于经典的社会学理论与心理学理论;国内有关数据质量的研究主要集中在统计学、图书馆学、信息系统学等领域,管理学领域的研究成果还非常少。

目前,在网络调查领域内,国内外对数据质量问题的研究均集中在传统调查模式中。鉴于社会化媒体用户的逐年增多,基于该平台实施网络调查进行数据收集必将成为一种趋势。然而,不同调查模式会产生不同调查效果^[5],社会化媒体作为一种新型调查平台,其社会性、实名制等特性会导致该类调查的数据质量与传统调查的数据质量存在偏差。如果调查者今后需要将社会化媒体作为混合模式调查的平台之一,那么探查两类调查间的数据质量差异就显得尤为重要。因此,有必要对该问题进行研究。

3 方法与数据

3.1 研究指标

在实证调查模式的比较研究中,有很多测量数据质量差异的指标。De Leeuw 列出了测量数据质量的 5 个指标^[24]; Holbrook 通过社会期望偏差来探查其对数据质量产生的影响^[19]; 而 Heerwegh 则

运用回复质量和社会期望两个维度来分析各调查模式的数据质量^[10]。

本研究中,我们基于回复质量和社会期望作答两个视角来进行统计分析。其中,回复质量主要是从作答结果角度来分析调查模式间得分分布的差异,该维度能够反映被试者在作答过程中所付出的认知努力程度;而社会期望作答则主要是从作答方式层面来分析调查模式间均值得分的差异,该维度能够反映出被试者在调查过程中所处的作答心情与作答动机。上述两维度已广泛应用于调查领域研究,它们可以区分测量误差所带来的调查模式效果,能够有效反映各调查模式间存在的数据质量差异。

本研究中,回复质量效果通过四个指标来进行测量:调查项中“不知道(不清楚)”回复率^[20];选项无差异程度^[11];作答规范程度^[23];调查项中“默认”回复率^[10]。对于社会期望作答,我们采用社会期望得分率来检验两类调查模式在该维度上存在的差异。

3.2 量表

本研究依据调查问题是否敏感将调查量表划分为两类:敏感性量表与非敏感性量表。若调查问题致使被调查者不愿如实反映自己真实观点或调查问题本身涉及个人隐私,则该问题为敏感问题^[21],包含敏感问题的量表即为敏感性量表。本研究选用三个量表来进行数据质量差异分析:创新精神量表(GI)^[25]、个人信息技术创新量表(PIIT)^[26]以及称许性平衡量表(BIDR)^[27]。近年来,信息系统采纳的文献表明,个人创新精神能够显著地影响人们采用信息系统的倾向。GI作为一般创新量表,已在信息系统采纳研究中被广泛使用,其所测量的变量也与信息系统密切相关;PIIT源于信息管理系统领域,是GI在信息技术领域的具体扩展,属于具体领域的创新量表;BIDR来自心理学领域,是目前最为广泛使用的用于测量期望性回答的量表,该量表包括两个部分——自我欺骗(self-deception, SD)与印象管理(impression management, IM),前者衡量诚实但过于肯定的自我表现;后者反映迎合他人的自我表现。在本研究中,GI与PIIT为非敏感性量表,BIDR为敏感性量表。

研究选择上述三个量表,主要基于以下考虑:①三个量表已广泛应用于众多领域,都被证明具有良好的信度与效度;②GI与PIIT测量内容均不涉及被调查者隐私,所得结果不易受到社会期望作答的影响,能够较为准确地发现调查模式间的回复质量差异;③BIDR涉及较多敏感问题,且该量表可以从有意识及无意识两个维度充分考察被调查者对于情景反应的敏感程度,易于发现结果差异;④三个量表结构简单,填写方便且清晰易懂。

三量表均采用Likter 7点打分法,“1”表示强烈不同意,“7”表示强烈同意。GI与PIIT均由4个题项组成;为保证分析结果的效度,我们从BIDR 40个问题中选取10个题项,其中,自我欺骗与印象管理各包含5个题项。三量表中的调查题项如附录1所示。

3.3 数据收集

本研究选择现阶段三种主流的社会化媒体——BBS、微博和IM作为社会化媒体平台,并选择传统数据收集方法——纸笔调查以及网络领域的传统方法——网站调查作为传统模式调查。

基于调查视角,社会化媒体可用做两种用途:问卷URL发布平台与问卷发布平台。前者主要用于被试招募;后者作为调查问卷放置平台。考虑到三种社会化媒体的适用性及特点,为保证数据质量,本研究将BBS与微博作为被试招募平台,IM作为问卷发放平台。我们依据调查结果是否存在明显作假或大范围重复作答的原则来筛选问卷,本次调查情况及数据统计状况如表1所示。

表 1 调查描述与统计状况

调查类型	调查方式	获取问卷	有效问卷	有效率/%	男性比例/%	18~28 岁比例/%
纸笔调查	纸质问卷	190	140	73.68	53.57	94.29
网站调查	邀请调查	196	137	69.90	41.61	81.75
BBS 调查	招募调查	177	128	72.32	85.16	98.44
微博调查	招募调查	166	128	77.11	50.78	86.72
IM 调查	发放问卷调查	132	130	98.48	63.85	92.31

纸笔调查中,被试来自中国西部地区三所综合性大学学生,为了保证问卷回收率并基于对被调查者现场解答疑惑的考虑,我们对问卷采取了当场发放、当场收回的方式,总共发放问卷 190 份,回收率为 100%。调查过程中调查者一直出现在作答现场,直至最后一位被调查者提交作答问卷为止。

网站调查选取问卷星网作为调查平台,通过邀请网站会员的方式来采集数据。该调查中被试与调查者以及被试间均互不熟悉。调查者可视参与会员的用户名,从而避免了网站会员的重复参与。

BBS 调查基于中国西部某高校的 BBS 论坛,采用招募被试的方式来实施调查。该 BBS 论坛实行实名注册制度,每名注册者只允许注册一个账号,确保了每个被试只参与一次。

研究选用新浪微博作为招募调查平台。在调查中我们鼓励“粉丝”们积极参与调查并转发该条调查信息,但规定每位“粉丝”只许参与一次。实际调查中,调查者利用微博留言来回答被调查者提出的各种疑问。

IM 调查中,我们选取腾讯 QQ 作为问卷发放平台。调查采取一对一发放方式,即只通过聊天窗口向好友发放问卷,而不利用群聊天或空间等其他平台。被调查者作答完毕后(不必当场作答)再通过聊天窗口将问卷传回调查者。

4 统计分析结果

为了确保各组调查间的统计相似性,我们分别对性别和年龄这两个分类变量的不同水平进行 K-S 检验,考察它们是否对问卷结果产生了影响。结果表明不同年龄段下三个量表的 Z 值均无显著性差异;而性别对研究结果的影响如表 2 所示。

表 2 不同调查组下性别对因子分数的影响

组别	纸笔组			网站组			BBS 组			微博组			IM 组		
	GI	PIIT	BIDR	GI	PIIT	BIDR	GI	PIIT	BIDR	GI	PIIT	BIDR	GI	PIIT	BIDR
Z 值	0.88	2.23*	0.80	1.29	1.59*	0.99	0.98	1.16	0.84	1.06	1.27	0.81	1.01	1.22	1.06
Sig. (双侧)	0.42	<0.01	0.54	0.07	0.01	0.28	0.29	0.13	0.49	0.21	0.08	0.52	0.26	0.10	0.21

注: * 表示 $P < 0.05$ 。

从表 2 中可以看出,性别只是对纸笔组和网站组中的 PIIT 产生了显著性影响,但影响未呈现出规律性。由此可知各组性别比例差异并不会对问卷结果产生影响。

本研究运用 SPSS 17.0 软件来进行统计分析,鉴于回复质量与社会期望作答中各指标统计值均为连续型变量,我们运用单因素 ANOVA 来分析五组调查间的指标差异状况,利用两独立样本 T 检验来分析任意两组调查中各测量指标的均值差异。在本调查中,我们对各组调查的有效问卷继续进行筛选,规定只有回答了所有题项且所得结果通过反向问题检验的问卷为最终有效问卷,具体情况如表 3 所示。本研究除作答规范程度外,其他指标的分析样本数我们均以两类量表各组中获取的最终

有效样本数为准。

表 3 两类量表下各组调查的最终有效样本数

	纸笔组	网站组	BBS 组	微博组	IM 组
非敏感性量表最终有效样本量(有效率/%)	117(61.58)	116(59.18)	126(71.19)	117(70.48)	123(93.18)
敏感性量表最终有效样本量(有效率/%)	115(60.53)	123(62.76)	125(70.62)	118(71.08)	112(84.85)

4.1 回复质量

我们通过统计两类量表中得分为“4”的题项来计算各组回答中的“不知道”回复率^[10],两类量表下各组统计结果如表 4 所示。从结果来看,两类量表下各组调查中的“不知道”回复率均值都较为接近,单因素 ANOVA 结果如表 5 所示。分析结果表明无论量表的敏感程度如何,该指标在本研究选取的五组调查中均不存在显著性差异。

表 4 两类量表下各组调查回复质量指标的均值(标准差)状况

	非敏感性量表				敏感性量表			
	“不知道” 回复率	选项差异 程度	“默认” 回复率	社会期望 得分率	“不知道” 回复率	选项差异 程度	“默认” 回复率	社会期望 得分率
纸笔组	0.21(0.14)	0.63(0.14)	0.12(0.21)	0.42(0.27)	0.17(0.15)	0.71(0.09)	0.07(0.12)	0.25(0.19)
网站组	0.19(0.19)	0.60(0.17)	0.16(0.25)	0.38(0.29)	0.16(0.15)	0.71(0.11)	0.10(0.16)	0.29(0.20)
BBS 组	0.18(0.16)	0.64(0.11)	0.12(0.18)	0.34(0.28)	0.15(0.14)	0.72(0.08)	0.08(0.13)	0.23(0.19)
微博组	0.18(0.16)	0.64(0.11)	0.12(0.19)	0.34(0.27)	0.19(0.17)	0.71(0.09)	0.08(0.10)	0.24(0.17)
IM 组	0.21(0.15)	0.65(0.12)	0.08(0.17)	0.35(0.25)	0.18(0.15)	0.71(0.08)	0.06(0.09)	0.26(0.17)

表 5 两类量表下各指标的单因素 ANOVA 结果

	“不知道”回复率	选项差异程度	“默认”回复率	社会期望得分率
非敏感性量表下各指标 $F(4, 594)$ 值	0.90	2.90*	2.42*	2.39*
敏感性量表下各指标 $F(4, 594)$ 值	1.40	0.30	1.08	0.59

注: * 表示 $p < 0.05$

我们采用 McCarty 等人的回复差异指数 ρ_d ^[28] 来计算各组调查中的选项无差异程度。该指数与被调查者使用不同 Likter 量表得分的数量有关, ρ_d 值越高说明差异越大,反映了被调查者在作答时使用了更多的回复选项。虽然表 4 结果中两类量表下各组调查的 ρ_d 值差异不大,但单因素 ANOVA 发现非敏感性量表五组调查中存在显著的选项差异程度,表 6 中 T 检验发现网站调查与其余四种调查相比有显著低的平均差异率。该指标统计结果也显示出相对于非敏感性量表,敏感性量表会产生更高水平的选项差异程度。

作答规范程度分析中,我们比较各组调查中获取问卷的有效率大小。我们基于两原则来对问卷进行筛选:一是看结果是否存在明显作假或大范围重复作答;二是观察非响应作答状况及所得结果是否通过了反向问题的检测^[22]。两类量表下各组有效状况如表 3 所示,单因素 ANOVA 表明两类量表五组调查的作答规范程度均存在显著性差异。非敏感性量表下 $F(4, 856) = 13.41 (p < 0.01)$; 敏感性量表下 $F(4, 856) = 6.65 (p < 0.01)$ 。表 6 结果表明 IM 调查较其他四组调查具有显著高的作答规范程度。

表 6 两类量表下各组调查模式间测量指标的 T 检验结果

对比调查组	非敏感性量表				敏感性量表
	选项差异程度	作答规范程度	“默认”回复率	社会期望得分率	作答规范程度
纸笔组 与 网站组	2.02*	0.48	-1.46	1.02	-0.45
纸笔组 与 BBS 组	0.03	-1.95	0.15	2.39*	-2.04*
纸笔组 与 微博组	-0.18	-1.78	-0.17	2.30*	-2.11*
纸笔组 与 IM 组	-0.76	-7.58*	1.48	2.19*	-5.13*
网站组 与 BBS 组	-2.17*	-2.45*	1.69	1.26	-1.61
网站组 与 微博组	-2.33*	-2.26*	1.37	1.18	-1.69
网站组 与 IM 组	-2.81*	-8.19*	2.86*	1.02	-4.73*
BBS 组 与 微博组	-0.24	0.14	-0.34	-0.07	-0.09
BBS 组 与 IM 组	-0.89	-5.41*	1.48	-0.31	-3.06*
微博组 与 IM 组	-0.64	-5.43*	1.74	-0.23	-2.92*

注：* 表示 $p < 0.05$

根据 Billiet 等人的观点,“默认”回复率为被调查者选择完全同意(即得分为“7”)选项的比率^[29],该回复率反映了被调查者对题项的完全接受程度。表 4 展示了两类量表各组调查中的“默认”状况。单因素 ANOVA 显示非敏感性量表下各组调查的“默认”回复率存在显著性差异,表 6 中 T 检验表明只有网站组与 IM 组存在显著性差异;然而,该指标在敏感性量表中却不存在显著性差异。表 4 结果表明,相对于非敏感性量表,敏感性量表具有较低程度的“默认”回复率。

4.2 社会期望回答

本研究采用社会期望得分率来衡量被调查者的社会期望回答状况,即统计所得结果中得分为“6”或“7”的题项数量在各类量表中所占的比例^[26]。各模式调查中得分率越高表示该组调查越容易受到社会期望偏差的影响。

两类量表各组调查中社会期望得分率的统计结果如表 4 所示。表 5 分析结果发现非敏感性量表下各组调查的社会期望得分率存在显著性差异,表 6 中 T 检验表明纸笔调查与三种社会化媒体调查均存在显著性差异。表 4 统计结果也可发现,敏感性量表会产生较低的社会期望倾向。

5 结论与讨论

本文研究了两类量表下五组调查结果在数据质量层面的差异,通过实证分析最终得到以下三点结论。

第一,不同敏感程度的量表会对五组调查的数据质量产生不同影响。本研究中,与非敏感性量表相比,敏感性量表产生了更低水平的选项无差异程度、“默认”回复率及社会期望得分率。

该分析结果与 Newman 等人的观点一致。在面对敏感问题时,被试者会充分权衡自己行为的成本与收益,不但对隐私更为敏感而且其行为更易受到社会期望倾向的影响。本研究结果表明在面对敏感问题时,被调查者不太轻易完全表露自己的观点,且作答行为与作答动机也更为错综复杂。

第二,同一量表下五组调查在回复质量方面存在显著性差异。本文实证分析发现,在非敏感性量表下,网站调查与其他四组调查在选项无差异程度上表现出了显著性差异;IM 调查较其他四组调查具有显著高的作答规范程度,网站调查与三种社会化媒体调查在该指标上也存在显著性差异;此外,

与 IM 调查相比,网站调查具有显著高的“默认”回复率。采用敏感性量表时,IM 调查依旧具有显著高的作答规范程度,与三种社会化媒体调查相比纸笔调查表现出显著低的作答规范程度。

对于非敏感性量表的实证调查结果,我们认为主要原因可能是各组调查间潜在信任程度的差异。纸笔调查与网站调查中的被试与调查者均完全不熟悉,信任缺乏建立基础,与三种社会化媒体调查相比,两类传统调查中被试者的作答动机更低,不太倾向于花太多精力来回答每一道问题,有时是想办法应付了事。敏感性量表下,被试者谨慎的作答心理及混杂的作答动机缩小了不同调查模式对回复质量产生的模式效果。

第三,量表的敏感程度会对各组调查带来不同程度的社会期望偏差影响。非敏感性量表下,纸笔调查与三种社会化媒体调查产生了显著性差异;但在敏感性量表调查中,五组调查间均不存在显著性差异。

对于该结论,我们将主要原因归因于各组调查间作答情景的差异。非敏感性量表下,现实作答情景较虚拟作答情景可能更易使被试受到社会期望偏差的影响。在本研究中,纸笔调查具备现实感知性,在集体环境下被调查者极易受到公共环境的影响,且 GI 与 PIIT 均是对自身创新精神的评价量表,因此社会期望得分率较高也在情理之中。然而,在敏感性量表下,这种现实感知性并没有带来显著的社会期望倾向效果。

本研究的局限性主要有以下三点:首先,研究选取的三个量表均用于测量被试的自我评价,属具体性量表。本文研究结论尚无法适用于综合性、全面性的量表,如领导力量表。其次,本次调查中各组参与群体大多较为年轻(18~28岁),相对一般群体来讲该部分被试更易于接受与尝试新技术,因此也会更习惯于社会化媒体调查。若调查被试为普通群体,研究结论是否还具有有一致性仍值得商榷。最后,本研究仅是基于测量误差视角来分析模式调查间的数据质量差异,并未探讨非测量误差对作答结果产生的模式效果。若实际调查中各模式调查的涵盖、抽样也对调查产生了实质性影响,则研究尚需进行更为深入详细的分析与探讨,这也是作者们今后未来的研究方向之一。

参 考 文 献

- [1] Kaplan A M, Haenlein M. Users of the world, unite! The challenges and opportunities of social media[J]. *Business Horizons*, 2010, 53(1): 59-68.
- [2] Chen Y B, Fay S, Wang Q. The role of marketing in social media: How online consumer reviews evolve[J]. *Journal of Interactive Marketing*, 2011, 25(2): 85-94.
- [3] 中国互联网络信息中心. 第 30 次中国互联网络发展状况统计报告[R]. 北京: 中国互联网络信息中心, 2012.
- [4] 中国互联网数据中心. 2012 中国微博蓝皮书[R]. 北京: 中国互联网数据中心, 2012.
- [5] Dillman D A, Smyth J D, Christian L M. *Internet, Mail, and Mixed-mode Surveys: The Tailored Design Method* [M]. Hoboken, NJ: Wiley, 2008.
- [6] Voogt R J J, Saris W E. Mixed mode designs: Finding the balance between nonresponse bias and mode effects [J]. *Journal of Official Statistics*, 2005, 21(3): 367-387.
- [7] Vannieuwenhuyze J, Loosveldt G, Molenberghs G. A method for evaluating mode effects in mixed-mode surveys [J]. *Public Opinion Quarterly*, 2010, 74(5): 1027-1045.
- [8] Bowling A. Mode of questionnaire administration can have serious effects on data quality[J]. *Journal of Public Health*, 2005, 27(3): 281-291.
- [9] Caro A, Calero C, Caballero I, et al. A proposal for a set of attributes relevant for web portal data quality[J]. *Software Quality Journal*, 2008, 16(4): 513-542.
- [10] Heerwegh D. Mode differences between face-to-face and web surveys: An experimental investigation of data quality and social desirability effects[J]. *International Journal of Public Opinion Research*, 2009, 21(1):

- 111-121.
- [11] Heerwegh D, Loosveldt G.. Face-to-face versus web surveying in a high-Internet-coverage population differences in response quality[J]. *Public Opinion Quarterly*, 2008, 72(5): 836-846.
- [12] Fricker S, Galesic M, Tourangeau R, et al. An experimental comparison of web and telephone surveys[J]. *Public Opinion Quarterly*, 2005, 69(3): 370-392.
- [13] Holbrook A L, Green M C, Krosnick J A. Telephone versus face-to-face interviewing of national probability samples with long questionnaires; Comparisons of respondent satisficing and social desirability response bias[J]. *Public Opinion Quarterly*, 2003, 67(1): 79-125.
- [14] Newman J C, Des Jarlais D C, Paone D, et al. The differential effects of face-to-face and computer interview modes[J]. *American Journal of Public Health*, 2002, 92(2): 294-297.
- [15] Dillman D A, Phelps G, Tortora R, et al. Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet[J]. *Social Science Research*, 2009, 38(1): 1-18.
- [16] 樊茗玥. 网络调查数据质量控制研究[D]. 镇江: 江苏大学, 2011.
- [17] 王瑜. 计算机辅助问卷调查中敏感问题数据的质量评价[J]. *统计与决策*, 2011, (13): 30-32.
- [18] 刘全, 朱胜, 何源. 网络调查数据质量的多级模糊综合评价方法[J]. *中国统计*, 2007, (5): 42-44.
- [19] 杜婷, 庞东. 网络抽样调查数据质量的评估与控制[J]. *统计与决策*, 2004, (4): 4-5.
- [20] Krosnick J A, Holbrook A L, Berent M K, et al. The impact of 'No Opinion' response options on data quality. Non-attitude reduction or an invitation to satisfice? [J]. *Public Opinion Quarterly*, 2002, 66(3): 371-403.
- [21] Tourangeau R, Yan T. Sensitive questions in surveys[J]. *Psychological Bulletin*, 2007, 133(5): 859-883.
- [22] Cho H, LaRose R. Privacy issues in Internet surveys[J]. *Social Science Computer Review*, 1999, 17(4): 421-434.
- [23] Heerwegh D, Vanhove T, Matthijs K, et al. The effect of personalization on response rates and data quality in web surveys[J]. *International Journal of Social Research Methodology: Theory and Practice*, 2005, 8(2): 85-99.
- [24] De Leeuw E D. *Data Quality in Mail, Telephone, and Face-to-face Surveys*[M]. Amsterdam: TT Publications, 1992.
- [25] Hurt H T, Joseph K, Cook C D. Scales for the measurement of innovativeness[J]. *Human Communication Research*, 1977, 4(1): 58-65.
- [26] Agarwal R, Parsad J. A conceptual and operational definition of personal innovativeness in the domain of information technology[J]. *Information Systems Research*, 1998, 9(2): 204-215.
- [27] Robinson J P, Shaver P R, Wrightsman L S. *Measures of Personality and Social Psychological Attitudes*[M]. San Diego, CA: Academic Press, 1991, 17-59.
- [28] McCarty J A, Shrum L J. The measurement of personal values in survey research: A test of alternative rating procedures[J]. *Public Opinion Quarterly*, 2000, 64(3): 271-298.
- [29] Billiet J B, McClendon M J. Modeling acquiescence in measurement models for two balanced sets of items[J]. *Structural Equation Modeling*, 2000, 7(4): 608-628.

Differences in Data Quality between Social Media and Traditional Mode Surveys

ZHU Wenlong, SHAO Peiji, FANG Jiaming

(School of Management and Economics, University of Electronic Science and
Technology of China, Chengdu 610054, China)

Abstract This paper studied the differences among data qualities of pen-and-pencil survey, web survey, BBS survey, microblog survey and Instant Messenger survey from two perspectives of response quality and socially desirable

responding. The results show that scales with different sensitivities produce different effects on data qualities of five survey methods. There are significant differences in response quality among five survey methods when using the same scale. Furthermore, sensitivities of scales have different impact on social desirability bias of five survey methods. This research aims to indicate the impact of mode differences on data quality among five survey methods so that researchers can consider the potential risks when integrating the similar mixed-mode data in the future.

Key words Web surveys, Social media, Data quality, Response quality, Socially desirable responding

作者简介

朱文龙(1985—),男,电子科技大学经济与管理学院博士研究生。研究方向:社会调查研究方法、社会化媒体调查。E-mail: zhuwenlong1985@163.com。

邵培基(1946—),男,电子科技大学经济与管理学院教授,博士生导师。研究方向:信息管理与电子商务。E-mail: shaopj@uestc.edu.cn。

方佳明(1982—),男,电子科技大学经济与管理学院副教授,博士。研究方向:电子服务、信息技术投资对企业的商业价值测量、网络消费者行为决策过程、社会化媒体营销策略。E-mail: jmfang@uestc.edu.cn。

附录

本附录为研究中三个量表的调查题项。

量表名称	维度	调查题项
创新精神量表(GI)	—	G1: 我乐意接受新思想(观念)
		G2: 我乐意尝试新事物
		G3: 我认为自己是一个具有创新力的人
		G4: 我不断尝试新的生活与做事方式
个人信息技术创新量表(PIIT)	—	P1: 当我听说一种新 IT 技术,我总是想方设法去体验它
		P2: 我对使用新的 IT 技术缺乏热情
		P3: 和周围人相比较,我总是第一个尝试新 IT 技术
		P4: 我乐意体验新的 IT 技术
称许性平衡量表(BIDR)	印象管理(IM) 自我欺骗(SD)	I1: 我对他人的第一印象通常都是正确的
		I2: 对我而言改掉任何一个坏习惯都很难
		I3: 我一点也不想知道别人对我怎么看
		I4: 我并不总是对自己诚实
		I5: 我总是清楚知道自己为什么会喜欢某些东西
		S1: 在不得已时,我有时也会说谎
		S2: 我从不掩饰自己的错误
		S3: 我有时也会欺负他人
		S4: 我从不说脏话
S5: 有时,我宁可对别人进行报复也不愿宽恕或遗忘		