

运用混合模式提升调查数据质量的国际经验

孙 妍

内容摘要：为应对调查应答率下滑及调查执行成本上涨所带来的压力，混合模式已成为近些年来国际社会调查方法领域的热门研究专题。本文在回顾社会调查领域数据搜集模式发展史的基础上，梳理了混合模式产生的原因及背景，并对混合模式提升调查数据质量的机制及实施方式作了详细阐述。结合中国家庭追踪调查（CFPS）项目实施经验发现，设计良好的混合调查模式能够切实提升项目整体应答率，但同时也提升了调查项目设计及数据搜集的复杂性，给调查执行机构提出了新的挑战。在我国，社会调查应答率下降的趋势已逐渐显现，摸索运用混合模式提升调查数据质量在现阶段具有较强的理论及实践意义。

关键词：社会调查；混合模式调查；非抽样误差；数据质量；调查应合率；CFPS

中图分类号：C915 **文献标识码：**A **文章编号：**1004-7794(2020)05-0059-06

DOI: 10.13778/j.cnki.11-3705/c.2020.05.010

一、引言

社会调查是通过样本群体搜集所得的信息，是对更大总体属性进行定量描述的系统方法^[1]。社会调查数据质量由测量准确度及样本代表性共同决定，而调查项目执行中的诸多环节均会对最终产出的数据造成影响。问卷设计优劣决定着调查中测量指标的系统性误差；样本框的质量及抽样方法决定着代表性样本的可及性；调查执行情况则直接影响最终应答样本对研究总体的代表性及测量指标的随机误差。从执行的角度，社会调查项目组织实施过程中有一项重要但容易被忽视的决策：数据以何种方式搜集。早期的社会调查起始于纸笔模式，随着电话、计算机、网络等新技术的发展，社会调查数据搜集模式也在迭代更新。整体而言，数据搜集朝着成本更低、周期更短、质量更优的方向发展。在数据搜集模式不断突破的同时，为应对调查环境恶化、调查执行成本上涨所带来的压力，在同一项目中运用多种数据搜集模式（即混合模式）的现象在西方极为普

遍。“混合模式”成为近些年来国际社会调查方法领域的热门研究话题。在我国，自 2010 年起大型综合性社会调查大都选择采用计算机辅助面访的模式采集数据。

近些年随着样本流失率的增加，追踪调查项目已开始摸索运用混合调查模式，但目前尚未有成文的研究成果发布。整理和梳理国际混合模式调查的经验及研究发现，借鉴国际经验开发适合中国国情的混合调查模式，在现阶段具有较强的现实意义。本文首先对调查数据搜集模式发展史作简要回顾，在此基础上介绍社会调查中常见的 4 类混合调查模式，并对混合模式影响数据质量的机制及类型、混合模式问卷设计与模式效应评估方法进行阐述，最后以中国家庭追踪调查（CFPS）项目为例介绍混合模式调查组织及实施的实践经验。

二、数据搜集模式发展简介

邮寄问卷及访员面访作为最为古老的数据搜集方式，是早期社会调查项目常用的数据搜集模式。资料记载首个邮寄问卷调查启动于 1788 年，

邮寄调查执行成本较低,在很长一段时间内作为采集社会调查数据的主流模式,美国曾经运用邮寄问卷的方式开展人口普查。访员主导的面访模式,因其高应答率、低访问门槛的特点而广受社会调查执行机构欢迎,时至今日面访模式仍然被作为数据搜集的金标准。1960年代后期,随着西方国家电话覆盖率的提升,电话访问成本低、周期短的优势逐渐得以体现,这一时期曾一度出现面访调查向电访让渡的趋势,直至1990年代,达到电话调查鼎盛时期。由此,面访、电访、邮寄问卷形成数据搜集的3种主导模式。20世纪末,随着计算机技术的推广,传统数据搜集模式与计算机技术结合拓展出许多新模式。计算机在数据录入、访题呈现、信息传输等环节逐渐取代传统纸版模式,由此诞生了计算机辅助面访、计算机辅助电访、计算机辅助自助访问、互动式语音应答(IVR)等技术。21世纪互联网的普及掀起了网络调查的风潮,调查逐渐从线下向线上迁移。近些年,智能手机、社交媒体及各类自媒体的广泛运用,丰富了手机调查模式。

在调查模式不断丰富的时候,社会调查执行却越来越困难,调查项目应答率不断下滑^[2],调查执行成本却逐年提升。Gao(2013)表示剔除通货膨胀因素后,2010年美国单户家户普查成本是1970年的6倍多^[3]。高质量调查数据的需求与现实资金、执行周期的压力迫使调查方法不断革新。在单一调查模式已无法应对现实挑战的背景下,混合模式调查(Mixed-mode Survey)被广为运用。如今,混合模式调查已成为美国及欧洲社会调查领域搜集数据的常规或惯用模式^[4-5],其初衷是希望通过多种数据搜集模式相互组合、彼此互补,在可负担(较低)成本下搜集高质量调查数据。

学术界关于混合调查模式通常有两种定义方式:第一种是根据与受访对象沟通的方式进行判别,在调查项目执行不同阶段采用多种方式与受访对象进行接触即视为“混合模式”^[6];第二种定义方式则限定在数据搜集阶段,即混合模式指的是采用一种以上的方式搜集调查数据^[7]。前一种

定义过于宽泛,且调查方法学家已充分论证在访问不同阶段采用不同的方式与受访对象接触对最终调查数据质量具有正面提升效果;而在数据搜集阶段采用不同模式对整体数据质量的影响则较为复杂。因此,本文以下讨论将限定于数据搜集阶段的混合调查模式。

三、混合模式类型

(一)并行式混合模式

并行式混合模式是指调查执行机构在同一时间为受访对象提供多种调查模式供其选择。并行式混合模式设计的基本假设是:受访对象对不同数据搜集模式存在选择偏好,为其提供调查模式选择一方面有助于提升受访对象对调查机构的好感,另一方面也希望有助于提升项目整体应答率。但大量实证研究表明,同时给受访对象提供多种模式选择并不能显著提升项目的整体应答率^[8]。在Medway et al(2012)所作的一项网络与邮寄问卷的混合模式评估实验中,并行混合模式所得到的应答率还不及单一邮寄问卷模式^[9]。可能的解释是当受访对象同时面临多种模式选择时,增加了其决策过程的复杂度,造成响应行为延迟,并最终因遗忘而放弃参与调查。

(二)接续式混合模式

接续式混合模式指的是根据受访对象的不同受访状态采用不同的调查模式,受访对象在每个时段内仅接触到一种访问模式。降低项目整体运作成本是采用接续式混合模式的最直接动因,因此接续式混合调查项目通常选择从成本最低的模式开始,对于第一阶段无应答的受访对象转而采用成本较高的模式继续接触。美国社区调查(America Community Survey, ACS)是典型的接续式混合调查项目。2014年调查执行中,每个样本面板组(Sample Panel)访问执行周期为三个月。所有有详细邮寄地址的受访家户首先会收到一封参与网络访问的邀请信;随后,未完成网络调查的住户会收到一份纸版问卷;第二个月,针对仍未应答的住户启动电访调查;最后一个月,启动面访调查,由访员登门对尚未应答的住户进行劝说并完成问卷调查。

接续式混合模式的理论来源于霍克斯提姆 (Hochstim, 1967) 所做的一项数据搜集模式比较实验。通过分析实验数据, 霍克斯提姆认为邮寄问卷、电访和面访三种模式组合方式取得的结果十分相近, 由成本较低模式开始逐渐向高成本模式过渡能够获得与单一面访模式类似的效果^[10]。近些年对接续式混合调查模式的讨论主要集中于论证不同模式的启动顺序是否会对整体应答率产生影响。迄今为止, 这个领域的研究结论不尽相同。一些研究表明, 在接续式混合模式下存在次序效应, 从低应答率模式向高应答率模式递进的组合效果通常不如反向为之^[11]; 也有研究认为不存在次序效应^[12]。需要说明的是, 已有的实证研究通常基于某个特定调查项目, 针对特定的目标人群, 结论普适性较差。对于出现次序效应有两种可能的解释: 第一, 若受访群体存在模式选择偏好, 当首现模式与其选择偏好契合时会得到较高应答率, 反之亦然; 第二, 可能存在登门槛 (foot-in-door) 效应, 即受访对象若不接受首现模式, 也更倾向于拒绝参与后续模式的访问。

(三) 嵌入式混合模式

前两种混合调查根据为受访对象提供模式选择的时机不同进行区分, 而嵌入式混合模式则是在测量工具层面引入多种模式, 即在一套问卷中部分问题采用区别于主体访问模式的其他模式完成。最常见嵌入式混合模式是在面访调查中, 采用自填的方式完成敏感题目的数据搜集。在覆盖健康主题的调查中, 嵌入式混合模式较为普遍。以全美药物使用与健康调查 (National Survey on Drug Use and Health, NSDUH) 为例, 当提问药物使用问题时, 访问模式由面访转变为受访对象自助辅助计算机访问模式。大量研究已证明, 针对敏感性问题采用更有助保护受访对象隐私的自填模式将降低社会期许效应、提高数据准确度。调查方法学界已就嵌入式混合调查模式有助于降低敏感问题测量误差达成共识。

(四) 追踪调查及跨国 (地区) 或跨文化调查项目中的混合模式

将时间和空间维度纳入后, 引申出两种不同

的混合模式。即同一受访对象在不同调查轮次接受不同模式的访问, 以及同一时期执行的一项调查项目在不同地区采用不同的数据搜集模式。

受限于样本框精确程度及受访对象联系信息完整程度, 为最大限度提高项目整体应答率, 大型追踪调查的基线调查通常采用最为昂贵但应答率最高的面访模式开展访问。在与受访对象取得联系并建立项目认同后, 则可转而采用成本较低的模式完成后续数据搜集。以美国收支动态长期追踪调查 (Panel Study of Income Dynamics, PSID) 为例, 1968 年基线调查采用面访的方式进行, 自 1972 年第四轮追踪调查起转为电访为主的数据搜集模式。追踪调查由成本较高模式向较低模式转换无疑会缓解资金方面的压力, 但如何从两轮数据差异中排除因模式转换导致的测量差异, 保持跨轮数据间的可比性是追踪调查混合模式执行的难点及重点。

在跨文化或跨国 (地区) 的调查项目中, 为保持数据可比性应最大限度保持问卷及执行模式一致, 但由于各国 (地区) 文化、地理、人口等方面的差异, 找到一种适用于所有调查区域的数据搜集模式难度巨大。因此, 跨文化或跨国 (地区) 调查中, 通常会允许各国 (地区) 执行机构根据区域特点选择最适合的调查模式。以世界社会情况调查 (International Social Survey Programme, ISPP) 为例, 最初项目设计是采用受访对象自填的模式搜集数据, 但在首轮 6 个参与的国家中, 就有两个国家为适应本国低文化人群较多的特点增加了面访数据搜集模式。

四、混合模式对数据质量的影响

从项目执行的视角分析, 影响调查数据质量的误差可分为抽样误差和非抽样误差两大类, 其中非抽样误差包含了覆盖性误差、无应答误差和测量误差三类, 混合模式的引入对三类非抽样误差均会产生影响。

第一, 混合调查模式的一项重要运用就是发挥不同模式覆盖人群不同的特点, 通过多种模式弥补单一模式目标人群覆盖不全的问题, 对于降低覆盖性误差有正面积的作用。20 世纪 90 年代, 电访盛行时期所采取的双框 (dual frame) 模

式便是混合模式降低覆盖误差的典型运用。Parackal (2003) 尝试将网络调查与纸版问卷调查模式相结合, 采用发送纸质调查邀请信的方式, 邀请受访对象参与网络访问, 并提供纸版问卷填答的选择, 有效降低互联网时代因“数字鸿沟”产生的覆盖误差^[13]。

第二, 混合模式设计的初衷是通过模式的转换给受访对象提供多种(次)选择, 从而达到提升项目的整体应答率的效果。如前文所言, 并行式混合模式在提高应答率方面所起的作用极为有限; 在接续式混合模式下, 后续访问模式是针对前一阶段无应答人群, 每新增一种模式必将增加联系次数, 从而提升项目整体应答率。Sala et al (2009) 的研究发现, 混合模式虽然无法彻底消除无应答误差, 但却能切实有效降低无应答误差^[14]。

第三, 不同搜集模式会对受访对象的应答过程产生不同影响, 由此导致不同模式间存在系统性偏差。排除测量工具的差异, 不同模式在信息传递方式、访问主导模式、访问压力及受访对象隐私保护方面的差异是导致出现模式效应的原因。Hochstim (1967) 在他奠定混合模式适用理论基础的实验研究中得到的结论是不同模式间的数据是可转换的, 即不存在模式效应^[10]。但后续聚焦模式效应的研究则得到了不同的结论, 大量研究表明信息传递渠道及访员是否在场均会对受访对象的回答策略产生影响^[4]。随着调查方法领域对混合模式研究的不断深入, 混合模式项目设计和执行水平也随之提升。Hox et al (2017) 研究表明, 在控制样本选择性问题后, 模式测量效应处在一个较低水平^[15]。

五、混合模式调查的设计与评估

(一) 混合模式测量工具设计

根据调查项目数据搜集策略的不同, 测量工具有两种可选设计模式: 统一设计和定制设计。

统一设计的目标是开发出一套适用于所有调查模式的测量工具。当调查项目中存在一个主导模式时, 在设计过程遵循主导模式测量最优化原则。例如, 面访主导、电访追踪的混合模式下,

问卷依然可以保留多选题, 并引入卡片等视觉辅助元素。当调查项目中存在多种并重的数据搜集模式时, 统一设计往往遵循受访对象负担最低的设计准则。例如, 在网络、电访、面访混合项目中, 所有的问题都采用简短的是否题提问方式, 避免使用多选题或长选项题^①。

定制设计的目标则是针对每种模式下受访者回答的特点设计最适合该模式的问卷。换言之, 定制设计追求的是语境而非形式相同。定制设计实施的难度在于设计者需要对受访对象在不同模式下的访问应答策略有非常全面、清晰的了解。遗憾的是, 目前学界尚缺乏足够的文献和资料对定制化设计的实施原则及实施效果进行论证。

(二) 测量及评估模式效应

混合调查中的“模式效应”差异形成的原因可拆分为两部分: 其一, 不同受访群体选择了不同的访问模式, 即“模式选择效应”; 其二, 不同数据搜集模式下受访对象采取不同的应答模式, 即“模式测量效应”。模式选择效应是混合模式存在的前提假设和理论基础, 在测量和评估模式效应时, 重点关注的是在控制模式选择效应后的模式测量效应。评估模式测量效应的基本思路是创建干预组和对照组, 在人群可比的前提下测量模式净效应。实验法和评估法是构造可比群体的两种方式。其中, 实验法指的是在数据搜集的过程中嵌入模式评估实验, 即在访问目标人群中随机选择受访对象运用不同数据搜集模式访问, 人为构造干预组和对照组, 为后期测量模式效应提供基础数据; 评估法则是在数据搜集完成后, 通过统计方法排除模式选择效应, 基本假定是客观人口学背景特征变量, 如性别、年龄、工作、教育程度等, 不容易受到模式测量效应的影响。通过匹配人口学背景特征, 相同特征人群在不同模式间的数据差异即为模式测量效应。

六、中国家庭追踪调查混合模式的实践经验

中国家庭追踪调查(China Family Panel Studies,

①迪尔曼等人(Dillman et al, 2014)详细阐述了混合模式统一设计的原则, 有兴趣的读者可进一步阅读。

CFPS)是北京大学中国社会科学调查中心组织实施的固定样本追踪调查。该项目于2010年启动全国首轮基线调查,调查样本覆盖全国25个省(区、市),649个区县的17960户家庭。基线调查后,项目保持隔年追踪的模式,回访基线调查界定出的所有家庭成员及其所在的家庭。在前三轮调查中,CFPS均采用面访主导的模式完成数据采集。随着调查追踪轮次的增加,人口流动、家庭分裂等因素共同作用导致每轮需要访问的样本规模不断攀升,且样本地域分布逐渐扩散。在2016年第四轮全国调查中,CFPS采用接续式混合调查模式尝试在可控成本下尽力提高项目整体应答率。项目组尝试运用更为便捷的电访替代面访,电访模式主要适用于两类样本:一是分布在面访访员覆盖区域外的零散样本;二是面访模式无法成功访问的样本。此外,为更好地实现跨轮调查模式的转化,CFPS项目组在减少测量误差、评估受访对象和调查数据质量等方面进行了有效尝试。

第一,为了尽可能减少因引进电访模式所带入的测量误差,项目组采用统一设计的思路,对原有的面访问卷进行调整,重新设计出一套同时适用电、面访的问卷,从源头上最大限度降低模式测量效应。结合电访主要依赖语言传递、无视觉辅助的特点,在保障追踪数据稳定可比的前提下,项目组遵循受访者负担最低的原则对问卷进行了两方面调整。(1)简化、拆分问题,减轻受访对象应答负担,例如,从多选题中提取关键选项,转化为是否题进行提问;回答难度大或答案生成规则较为复杂的问题则将测量的综合指标进行分解,拆分为若干道简单问题进行提问。(2)改变提问方式,取消受访者卡片,并引入访员现场编码,部分客观问题,如大学学科、雇主性质等,在访员提问后由受访者直接描述自己的情况,访员根据受访者提供的信息现场编码后选择最恰当的选项。

第二,鉴于CFPS问卷内容较为丰富、访问时长较长(个人问卷平均访问时长约50分钟),项目组有针对性地组织了两轮电访测试,评估受

访对象对采用电访模式完成长问卷调查的适应性。首轮测试选择了28名各年龄段、各学历水平的便利受访者,受访者年龄最小17岁、最大74岁,进行完整问卷调查。测试结束后,通过受访者回访、访员小组座谈、行为编码、数据分析等方式对测试中暴露出的问题进行总结归纳,并根据测试结果对问卷作出调整。第二轮测试安排在CFPS2016全国追踪调查启动前,项目组挑选出557个地理集中度相对零散的家庭样本,开展电访预调查。预调查旨在对电访问卷适应性、受访者酬金设计、访员薪酬设计、混合模式样本管理系统运转状况等方面进行全流程演练,确保正式访问顺利开展。两轮适应性测试后,CFPS正式启动2016年第四轮追踪调查,所有样本首先由面访访员尝试接触,面访访员无法触及或面访失访的样本转而采用电访模式继续追踪。

第三,调查结束后项目组从应答率、数据缺失率、社会期许倾向、跨轮数据差异等维度对电访及面访采集的数据质量进行评估。从应答率角度看,电访的引入切实提高了项目整体应答率。首批电访发放的557个家庭样本中,成功完访398条,完访率为71%;面访失访且有联系电话的8604个家庭转入电访后,最终电访完访3714条,逆转率为43.2%。数据质量方面,由于电、面两种模式都属于有访员介入的依靠听觉传播的数据采集模式,加之CFPS在不同模式中均使用相同的问卷,因此在控制受访者个人特征及访问经历后,两种模式采集所得数据间并未发现显著差异。为方便数据用户使用,CFPS2016数据发布时,在数据库中增加了数据采集模式的分类变量标注每条样本数据采集的模式。

七、小结

近年来随着样本应答率的下降,我国的调查执行机构也已开始尝试和摸索运用混合调查模式,欧美国家混合模式调查项目实施经验及相关研究成果具有较强的借鉴意义。

首先,调查项目设计人员应在思想上重视模式效应的存在,在项目设计及实施过程中采取适当预防或矫正措施;其次,不同数据搜集模式以

不同方式组合会产生不同的效果,制定符合受访群体特点的样本流转规则是保证混合模式调查成功的关键;再次,需要根据不同模式组合策略对测量工具进行适应性调整,在对受访对象个人属性及应答模式没有充分了解的情况下,可将统一设计作为问卷设计的起点;最后,在调查项目执行过程中需要重视搜集混合模式效应的评估数据,为后续数据合并及校验提供素材。条件允许的情况下,调查机构可以组织随机实验评估模式测量净效应。

在运用混合模式提升调查数据这个研究领域中国目前仍处在探索阶段,有许多基础性问题值得学者深入研究。欧美国家在混合调查模式研究方面的积累都将为中国混合模式调查项目设计提供理论基础及实践指导。中国各大调查项目可发挥后发优势,设计并执行适合我国国情的混合模式调查,提升社会调查数据质量、降低数据搜集成本。CFPS项目实施经验表明,面电接续混合模式在追踪调查中实施效果良好,但引入混合模式影响调查数据质量的机制将是下一阶段研究的重点及方向。

参考文献

- [1] 邱泽奇. 调查方法[M]. 重庆: 重庆大学出版社, 2017.
- [2] Tourangeau R, Conrad F G, Couper M P. The Science of Web Surveys[M]. Oxford: Oxford University Press, 2013.
- [3] GAO. 2020 Census: Progress Report on the Census Bureau's Efforts to Contain Enumeration Costs. GAO-13-857T. Washington, DC: United States Government Accountability Office, 2013.
- [4] Dillman D A, Smyth J D, Christian L M. Internet, Phone, Mail, and Mixedmode Surveys[M]. NY: Wiley, 2014.
- [5] Brick J M, Williams D. Explaining Rising Nonresponse Rates in Cross-sectional Surveys[J]. Annals of the American Academy of Political and Social Science, 2013, 645(1): 36-59.
- [6] Leeuw E D D. To Mix or not to Mix Data Collection Modes in Surveys[J]. Journal of Official Statistics, 2005, 21(2): 233-255.
- [7] Schouten B, van den Brakel J, Buelens B, et al. Disentangling Mode-specific Selection and Measurement Bias in Social Surveys[J]. Social Science Research, 2013, 42(6): 1555-1570.
- [8] Griffin D H, Fischer D P, Morgan M T. Testing an Internet Response Option for the American Community Survey. Paper Presented at the 56th Annual Conference of the American Association for Public Opinion Research, Montreal, Quebec, May 2001.
- [9] Medway R L, Fulton J. When More Gets You Less: A Meta-analysis of the Effect of Concurrent Web Options on Mail Survey Response Rates[J]. Public Opinion Quarterly, 2012, 76(4): 733-746.
- [10] Hochstim J R. A Critical Comparison of Three Strategies of Collecting Data from Households[J]. Journal of the American Statistical Association, 1967(62): 976-989.
- [11] Lynn P. Alternative Sequential Mixed-Mode Designs: Effects on Attrition Rates, Attrition Bias, and Costs[J]. Journal of Survey Statistics and Methodology, 2013(1): 183-205.
- [12] Olson K, Smyth J D, Wood H M. Does Giving People Their Preferred Survey Mode Actually Increase Participation Rates?[J]. Public Opinion Quarterly, 2012, 76(4): 611-635.
- [13] Parackal M. Internet-based and Mail Survey: A Hybrid Probabilistic Survey Approach[J/OL]. Proceedings of the 2003 Australian Web Conference at <http://ausweb.scu.edu.au/aw03/papers/parackal.html> (accessed April 15, 2004).
- [14] Sala E, Lynn P. The Potential of a Multi-Mode Data Collection Design to Reduce Non Response Bias—The Case of a Survey of Employers[J]. Quality & Quantity, 2009, 43(1): 123-136.
- [15] Hox J J, Leeuw E D D, Klausch T. Mixed-mode Research: Issues in Design and Analysis[M]//P P Biemer, E D de Leeuw, S Eckman, et al. Total Survey Error in Practice. Wiley Series in Survey Methodology, 2017: 511-531.

作者简介:

孙妍, 女, 1981年生, 辽宁人, 2009年毕业于北京林业大学经济管理学院, 获管理学博士学位, 现为北京大学中国社会科学调查中心助理研究员, 研究方向为社会调查方法。

(责任编辑: 孙娜娜)