

CAPI不是万能的: 入户调查执行方式与系统性误差^①

邹宇春 张丹 张彬

(中国社会科学院社会学所 北京 100732)

摘要:现代信息技术的迅速发展掀起了诸多领域的革新,社会调查领域也不例外。相比传统的纸笔入户访问(PAPI),计算机辅助面访(CAPI)的使用极大地改善了数据质量并在一定程度上缩减了调查成本,正被越来越多的调查者和机构所采用。但是,CAPI不是万能的,它没能完全解决调查过程中误差的问题,同时也带来了新的挑战。两种方法在执行方式上存在异同,如果把社会调查分为前期准备、实地访问、数据回收与核查三个阶段,CAPI在入户调查的执行和质控方面能够消减系统性误差,但也面临由于技术和设备缺陷而仍会导致的系统性误差等新问题。

关键词:CAPI; PAPI; 入户调查; 质量控制; 系统性误差; 研究方法

中图分类号:C91 **文献标志码:**A **文章编号:**1002-462X(2019)06-0043-07

CAPI是一种使用便携设备如笔记本电脑、平板电脑等进行面对面访问的社会调查技术。访问结束后,通过互联网,访员可以及时把数据传输给研究者;同样地,在实地访问过程中,研究者也可以通过互联网给访员发出指令,及时调整访问进程。与传统纸笔访问(PAPI)相比,CAPI将社会调查的工作重心前移,在省去数据录入环节的同时,增加了问卷电子化等步骤。

当前,入户调查采用CAPI形式已成为不可阻挡的趋势,主要是因为它的两大优势:一是能够利用现代信息技术对实地调查的过程进行实时监控与及时干预,从而降低入户调查的系统性误差,提高了数据质量;二是省去PAPI中的一些环节,从而降低大规模或定期重复类社会调查的成本^[1]。但是,尽管CAPI为改善数据质量做出了

卓越贡献,但这并不意味着CAPI能完全解决传统纸版问卷入户调查中的系统性误差问题。换言之,CAPI并非是万能的,作为一项新兴技术,它只是在一定程度上消减了入户调查中的误差,与此同时也带来了新的问题。

随着近些年我国调查领域逐步推广CAPI的趋势日渐明晰,充分了解CAPI的优势、理性分析应对CAPI有可能面临的问题和挑战显得十分必要。笔者把社会调查分为前期准备、实地访问与数据回收三个阶段,分阶段阐述CAPI在入户调查的执行和质控方面如何消减误差以及面临的挑战,并在此基础上尝试提出解决思路。

一、入户调查执行方式的比较: PAPI vs CAPI

(一) 纸笔问卷面访(PAPI)

PAPI是传统的入户调查方式。在此种方法下,访员携带设计好的纸质问卷,依据抽样信息的指引,进入到受访者家中进行访问。期间,基于前期培训确定的行为准则,访员进行户内抽样并依问卷提问,尽可能获得受访者真实的回答,并由访员将受访者的回答登记在纸质问卷上。当所有问题回答完毕后,即视为完成一次访问。这种方法

^① 此文在与中国社科院李炜研究员、任莉颖副研究员,北京大学中国社会科学调查中心丁华副研究员等学者的讨论下形成。感谢他们提出的宝贵建议,文责自负。

作者简介:邹宇春,1979年生,中国社会科学院社会学研究所副研究员;张丹,1996年生,中国社会科学院社会学系研究生;张彬,1995年生,中国社会科学院—上海研究院研究生。

对设备和访员技能的要求较低。此外,由于研究者分身乏术,一般会将访问数据的质控工作分为两部分,分别由实地督导和后台督导予以完成。不合格的问卷将被作废,要求访员重访或补访。实地督导主要负责现场数据核查,与实地调查同步进行。后台督导主要负责数据的二次审核,多数是在实地访问已结束的情况下对问卷数据进行的

后期审核,发现问题后要求访员重访或补访的成本、难度都非常大。因此,现场实地审核是数据质控的重要环节,面临工作量大、流程复杂、相对耗时等压力。待访问和核查结束后,将进入数据录入阶段。同时,对于其中的缺失和异常值,调查者将会进行进一步清理,为数据挖掘和深入分析做准备。见图1。

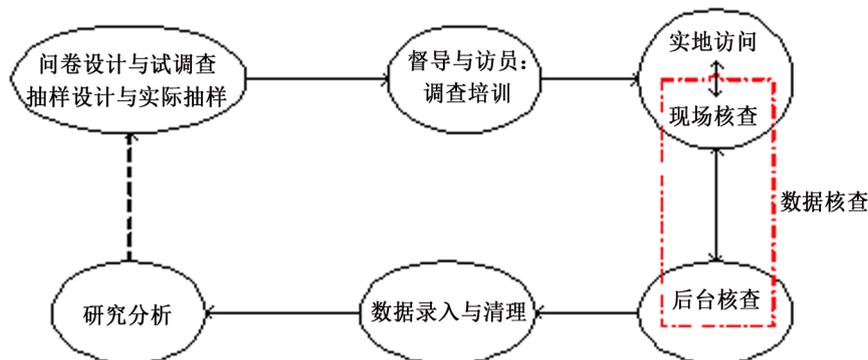


图1 PAPI流程图

(二) 计算机辅助面访(CAPI)

随着研究者对数据质量的追求和现代信息技术的发展,越来越多的调查者倾向选择使用CAPI技术进行入户调查。由于引入了现代信息技术和相关设备,CAPI引起了调查流程环节的变化。这种方法要求在入户前,需要将设计好的问卷进行电子化。经过一系列调试以确保问卷无误后,电子问卷被留存在系统之中。调查依据基础材料(如地址地图)和事先设计好的抽样方案进行,设计者将抽样信息与问卷结合,也存储在系统之中。目前,抽样的电子化正在逐步实现,尤其是入户后的现场抽样常被并入CAPI系统里,减少了抽样的耗时和因抽样误操作所产生的偏差,提升了访问的调查效率和质量。

携带受访户背景信息的电子问卷被分配给执行实地访问任务的各访员的电子设备。访员根据分配好任务的电子设备的指引,进入受访者的家中进行访问。期间,访员依据电子设备的提示进行询问,受访者按要求回答,同时,访员将受访者的回答记录在电子设备上。在电子问卷完成后,访员利用互联网把获得的调查数据连同系统记录

的访问时间、访问间隔、每题作答时长、受访者GIS信息等等并行数据传输给调查后台。

当数据传输回调查后台后,现场督导和后台督导能够同时看见调查数据,执行数据核查等质控工作。一般来说,数据传输被要求在访问结束后的较短时间段内完成,以确保这两个阶段大致是同步进行的。如督导发现问题,可以及时通过互联网将信息和要求传输给访员,并要求重访或补访。研究者能够根据计算机记录的并行数据大致还原出访问现场的情况,从而为数据核查提供依据。这大大压缩了后期数据清理的时间,但拉长了问卷准备和访员培训等前期环节的时间,并对访员培训内容提出巨大挑战。

与PAPI相比,CAPI在入户调查中运用了相对较多的现代信息技术。而这些技术的引入,在一定程度上加强了调查者对访问现场的调控,约束并规范了访员的行为,从而提高了数据质量。Caeyers等学者^[2]做了关于CAPI的引入对调查质量影响的量化研究,结果表明,CAPI大幅度消减了路径错误发生的可能性,每个家户调查路径错误的数量从10.4降低为0.0;同时,用CAPI进

行家户调查产生的异常值也显著低于 PAPI。此外,长远来看,由于电子设备和系统程序具备可反

复使用的特性,CAPI 能降低大型定期或重复调查的成本和耗时^[3]。

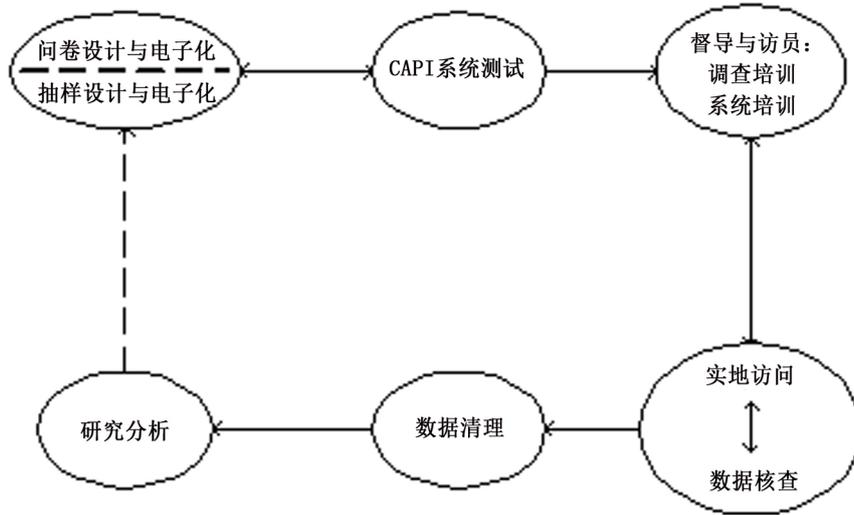


图 2 CAPI 流程图

二、计算机辅助面访的执行与质控: 摇摆的系统误差

误差大小是衡量入户调查数据质量的主要指标。所谓误差,是指调查结果与总体真实情况之间的偏差。误差分为两类: 系统性误差和随机误差。学界对后者的研究已较成熟,简单地说,随机误差的大小与样本量呈反比,当样本大小与总体相等时,随机误差为零。因此,在条件许可的情况下,研究者们偏向做大样本的社会调查。但随之而来的问题是样本量的扩大会导致系统性误差的增加。系统性误差即非抽样误差,由于受主观因素的影响明显,其原因较复杂,难以预测。系统性误差又可分为四类: 一是设计误差,由研究设计中的错误导致,如抽样框的偏差; 二是估计偏差,由估计方法有偏差导致; 三是调查误差,指收集样本数据中产生的系统性误差,受问卷设计、访员、受访者等多方因素的影响; 四是编辑误差,指在数据核查、清理、录入、计算中产生的系统性误差^[4]。

由上文可知,CAPI 改变了传统 PAPI 的调查流程和方式,通过技术约束来控制调查过程中主

观因素的影响。大量研究表明,这些技术的应用较好地控制了系统性误差,提高了数据质量,从而成为 CAPI 的主要优势之一。那么,这种改变如何影响作为衡量数据质量重要指标的系统性误差的大小,特别是调查误差和编辑误差有没有带来新的影响因素,就成为值得关注的问题。同时,需要注意的是,CAPI 仅仅是一种新的调查工具,本研究的讨论仅针对工具及其使用,而不涉及如测量指标选择等与工具本身无关的内容。接下来,笔者把入户调查的执行过程分为三个阶段,逐阶段分析 CAPI 对误差的影响。

(一) 前期准备阶段

问卷的质量即问卷获取受访者真实统计特征的能力是调查者普遍关心的问题。除指标体系之外,问题设计对调查误差的影响既体现在问题陈述上如陈述中的暗示等,也涉及问卷逻辑、形式等方面如询问的顺序。CAPI 的一大革新在于实现了问卷的电子化。在问题陈述方面,与纸质问卷相比,电子问卷可在大量认知测试的基础上根据受访者的年龄、受教育程度等特征,利用视频、音频、文字等方式实现不同风格、不同表达方式的问题陈述,从而使不同水平的受访者达到相近的理

解程度^[5]。这降低了对访员语言、沟通能力等技能的要求。研究表明,CAPI的使用显著降低了访员的调查经验、受教育程度等因素对访问过程中错误发生率等的影响^[2]。在问卷逻辑等方面,电子问卷通过利用现代信息技术进行硬检验和软检验,比如,设定跳转等询问路径、进行前后回答一致性检验、核实异常值答案的提醒、完成状况检查等内容,有助于对现场填答信息即时矫正,避免了实地访问中可能发生的一些逻辑错误和人为的填答错误。

抽样方案的设计和 implementation 是否严格遵守随机性原则是系统性误差产生的另一个来源。通过CAPI,抽样过程也同样可以实现电子化。依据调查研究人员提供的抽样基础信息,如调查点的村/居信息、受访户的地址清单列表、抽样规则等,CAPI技术人员设计相应程序指令,由访员现场点击抽样按钮并直接抽出受访者。而传统的PAPI中主要采取Kish选择法或生日法来抽取选择对象,以此保证家庭中的所有成年人都有同等的被选中的概率。Kish选择法需要事先准备若干套表格,并按年龄性别进行的排序、编号,最后根据一套程序选出受访者;生日法则需要统计家庭内全部成年人的生日,确定标准日期,然后确定最终的受访者^[6]。这两种方法都要访员提前准备和学习,正式访谈开始前的复杂程序也可能使受访者产生厌烦,影响后续访问质量。CAPI中的入户抽样简单快捷,录入家庭中符合访问要求的人员之后,计算机自动随机抽取受访者,无需提前准备表格或学习入户抽样方法,从而降低了访员抽样能力不足的负面影响,确保了随机性原则的实现。

不过,尽管CAPI系统的使用提高了减少系统性误差的可能性,但要从上述两个方面减少误差须以合适且准确的程序设计和访员的正确理解与操作为前提。这也极大增加了前期准备阶段的工作难度、访员培训的内容和对访员电子操作技能的要求,在一定程度上又可能导致系统性误差的增大。

一方面,CAPI系统的研发对调查执行团队的双重能力提出了巨大挑战。CAPI是一种调查工

具,这种工具要从“理想中的好用”变成“实际操作中的好用”,必须实现技术编程和调查操作的相互契合。目前,在我国社会调查领域,CAPI的软件技术开发和调查操作设计往往是由至少两个部门合作完成。多数情况下,调查研究设计人员须就调查流程的环节、工作内容和目标等反复与CAPI技术人员充分沟通,督促技术人员在CAPI系统中予以实现。尽管国内有些机构已研发出一些CAPI模板,但不同的调查目标会产生不同的调查需求,研究人员事先必须能意识到这种需求的特殊性并充分表达给技术人员,以便技术人员能够采用相应的编程技术在CAPI系统中予以实现。研究人员还需要高度关注整个调查流程的电子化设计,通过反复参与系统测试来体验是否实现了调查研究设计目标,以此避免出现因电子化系统缺陷而无法弥补的系统性误差。同时,由于编程技术的高专业性,调查研究人员对CAPI系统中的某些技术设置难以做出“是否最优”的评估,在CAPI的测试阶段需引入第三方进行技术测试和比对,以便更好地评估CAPI系统对访员的亲和性、现场的便利性以及调研目标的可达性。换句话说,如果软件技术人员不能很好地理解调查研究者的调查设计,而调查研究人员又未能在准备阶段及时发现这些问题,CAPI产生的系统性误差未必会小于纸版问卷调查。

另一方面,访员培训中需增加CAPI系统操作程序的培训和模拟测试。与传统调查培训不同,CAPI调查的访员培训要重点突出CAPI系统的操作化。一般来说,访员培训包括项目整体介绍、调查访问流程、调查访问技巧、访员安全须知、现场抽样操作、调查指标讲解、数据质量审核、现场调查模拟、试调查等内容。PAPI的培训以使用纸质问卷为背景进行以上内容的相关培训,而CAPI的培训则在电子化系统运行的背景下进行。与PAPI的培训相比,CAPI培训需要访员用立体思维去理解CAPI系统内各个模块、各个平台之间的多维链接关系,各个按钮的使用方法、注意事项、技术难点。缺乏电子产品使用经验的访员不太容易在短期培训内灵活掌握CAPI的使用方法,

尤其在互联网信号弱或技术支持人员缺位的情况下无法独立应对 CAPI 操作现场突发的基础技术问题。这不仅对访员在电子设备使用和电子程序知识方面的基本素养提出了更高的要求,也意味着前期的访员培训需要在此方面增加培训内容。^①

(二) 实地访问阶段

实地访问阶段是入户调查实现收集数据的重要阶段,也是系统性误差的主要来源。由于这一阶段涉及多个主体,故误差的来源也较为复杂。访员是研究设计者在这一阶段的代言人,他们直接接触受访者并记录访问结果,在一定程度上影响受访者的回答。因此,他们成为调查误差的一大来源。影响方式有两个方面:其一,由于访问是一个社会互动的过程,所以访员的人口学和社会经济特征可能影响受访者的回答;^②其二,访员的整个访问行为可能影响受访者的回答,如解释或改述问题造成的理解差异、记录错误造成的误差等^[7]。此外,作为研究对象的受访者是有反应性的个体,他们会根据自己的知识库存和心理、调查时的情境、访员的表现等选择自己的答案,从而影响了回答的真实性。在 PAPI 中,这些难以控制的因素往往产生了大量无法计算的系统性误差。而 CAPI 通过智能化的提问方式、提供选项图片、增加录音、私密问题的自我填答等方式,在一定程度上解决了这些问题,与此同时,却也带来了新的影响因素。具体来说,有以下两个方面:

1. 电子设备的影响

就对电子设备的要求而言,要想成功降低系统性误差需要满足两个条件:一是电子设备本身不会引起受访者的负面情绪,二是设备和网络的稳定。一方面,民众对电子设备的信任水平部分地决定了他们对调查的态度,进而影响数据质量。Baker 等人^[8]的研究表明,在欧美国家,控制其他条件不变时,47%的受访者认为 CAPI 比 PAPI 更机密,CAPI 有更高的应答率和应答意愿,且获得的数据在分布上无明显差异。这说明,欧美国家的绝大部分受访者能够且倾向于接受电子问卷,且基本不会影响受访者的回答。但这一结论在中国是否成立尚有待验证。比如,CAPI 的自动录音功能为数据核查提供最即时和真实的现场访问信

息,但随着我国居民对个人隐私保护意识的增强以及当前社会里非法使用个人隐私数据的现象激增,入户访谈时采用录音的方式很可能会增加某些类型受访者拒访或隐瞒真实信息的可能性,进而降低了数据的信度。

另一方面,设备和网络的稳定性也是需要关注的问题。在实地访问过程中,特别是对大范围调查而言,由于各地基础设施条件的差异,网络稳定性常受影响。这导致现场的调查数据无法及时传输给调查后台,进而影响调查者在数据核查上的及时性和有效性。另外,一旦在实地访问过程中发生设备故障,受访者的心理极有可能受到影响,并导致拒访;更有甚者,系统故障还可能致未传回的数据丢失,影响访问进度。

2. 不规范的访问

不规范的访问行为有多种形式,既包括有意识的行为,如为了减少工作量而群访、臆测结果,也包括无意识的行为,如登记错误、声调的不同等^[7]。PAPI 主要通过现场督导的监督来解决这一问题,但效果有限。而 CAPI 则利用现代信息技术设置限制或分析并行数据使这一问题得到了较好的解决。一方面,通过录音和 GPS 定位等技术记录的并行数据,研究者可以较清楚地掌握访员的整个访问过程,包括访问的位置、访问过程中的语音和时间信息。这些信息可以在一定程度上还原访问过程,并通过录音回放和访问时间分析

^① 本研究此处强调的是调查培训中为使用 CAPI 而需要增加的相应培训内容,但并没有否认,CAPI 相比 PAPI 在调查培训上更具有其他优势。

^② 当访问的题目与访员的某一或某些特征相关时,为避免不必要的潜在冲突,受访者可能选择隐藏自己的真实想法,根据访员的特征做出判断和回答。同时,访员在场本身可能会增强受访者自我保护的意识或带来遵守社会规范的压力,进而影响受访者对敏感问题的回答,如收入、同性恋等问题,出现虚假回答或社会赞许行为。需要注意的是,CAPI 无法消除或减少访员本身的人口学或社会经济特征所带来的影响,但可以减少访员在场所带来的影响。与 PAPI 相比,CAPI 最大的不同在于引入了现代信息技术和相关设备。然而,现代信息技术和设备虽然可以利用音频、视频等技术替代访员的询问,但无法改变访员的特征,也无法解决访员在场的问题。这是入户调查方法一直存在的不足,并非 CAPI 引起的新问题。

来判断访员的沟通技巧、态度等是否影响受访者的回答。另一方面,电子问卷只能在完成上一份问卷的前提下才能进行下一个访问,这避免了群访的可能性。这些技术的应用在一定程度上规范了访员的访问行为。

然而,这也体现出 CAPI 本身的限制。其一,虽然录音可以部分还原访问现场的情境,但却无法反映访员和受访者在访问之前的交流以及访员的肢体语言等,故而无法杜绝访员和受访者联合作弊的情况和访员行为的影响。其二,利用录音回放进行数据监测具有一定的滞后性,有时会出现发现问题却来不及干预的情况。其三,访员的经验是把“双刃剑”。标准化的询问过程虽然避免了逻辑错误,但也限制了访员自由发挥的空间,如根据访问现场的实际情况及时适当地调整询问顺序能够获得更好的访问效果^[3],但固定的访问顺序致使访员无法根据现场情况做出合理可行的访问顺序的调整。其四,当受访者的回答前后不一致或提供了对前面题目的补充回答时,CAPI 访员必须返回修改,比 PAPI 更加耗时且欠缺灵活性。在 PAPI 的访问现场,当访员发现受访者的回答存在逻辑不符或前后冲突时,可以在问卷空白处加以备注,留待访问结束后再返回问卷相应题目进行修改,不用中断正在进行的访问。相比之下,CAPI 系统缺乏这种能够随时补写信息、随时定位单个题目的灵活性,电子界面的翻页模式或模块选择模式都相对耗时,打断了正常的访问过程,访问的流畅性受损,并可能引起受访者的厌烦而导致中途拒访的情况发生,进而导致应答率下降。

(三) 数据回收与核查阶段

这一阶段的主要任务是收集和核查访员获得的数据,它是连接实地访问和数据分析研究的关键步骤。在传统的纸版问卷调查中,由于无法进入每个访问现场,调查研究者只能依据现场核查、问卷回收后的数据信息来判断数据的质量即核查误差大小,并进行下一步的分析研究。相对而言,CAPI 的数据核查与数据回收可以同步进行,并且省去了数据录入的过程,研究分析者可以更加快捷地得到调查数据并进行误差分析。

1. 数据回收与计算

对于纸版问卷调查,将收集的数据录入计算机并进行必要的计算是进行数据分析的必要步骤。在这一过程中,无论人的工作精确度有多高,都会存在编辑错误,即编辑误差。编辑误差的存在影响着数据分析结果的准确性。对 PAPI 收集的数据在录入环节的误差,调查者用校对和非法值检验等方法来部分地消除此类误差^[9]。

与 PAPI 不同的是,CAPI 的数据回收与核查阶段在时间上与实地访问几乎同时进行。基于对互联网技术的应用,它实现了即时数据回传,省去了人工的数据录入环节;同时,计算机还可以自动进行检验,因此,也避免了计算环节中编辑误差的产生。Bhatia 等人^[10]发现 CAPI 中发生的错误数量平均比 PAPI 少 60%。但是,如前文所述,系统和网络的稳定性可能导致数据丢失或格式错误,从而影响数据回收的效率,徒增访问过程的工作量。此外,CAPI 系统的后台极易产生冗余数据,需要花费一定的时间对其进行清理以便得到更为清洁的调查数据和平行数据,这是 PAPI 所未曾遇到的问题。

2. 数据核查

如前文所述,CAPI 产生了大量的并行数据。所谓并行数据,就是在访问过程中产生的记录访问过程的数据,如访问时间和时长、位置、录音等。并行数据是对访问过程进行质量控制的关键材料,因而受到了广泛的关注,成为 CAPI 的一大优势。

然而,现有技术还难以对并行数据进行深度发掘,特别是将音频资料转换为可以进行统计分析的量化资料。囿于技术的限制,目前并行数据的使用需要依靠大量的人力进行,费时、费力,成本高。由于并行数据的时效性强,因而其发掘的高成本低效率极大地限制了它的使用^[11]。因此,将来能否提高并行数据的开发使用,成为减少 CAPI 的系统性误差的关键。

三、总结与建议

综上所述,CAPI 与 PAPI 相比具有两大主要特点:一是缩减了大规模或定期重复类社会调查的成本;二是在一定程度上消减了系统性误差,提

高了调查质量。但要实现这两个优势需要满足一系列条件,就像“旅行中飞机更高效”的前提是没有设备、技术故障和交通管制以及天气条件允许等。一方面,成本缩减是对某些类型的社会调查而言的,比如,需要开展大规模或定期重复的调查,这些调查会因为电子设备和系统可持续使用性使其成本随时间推移而降低,但若只是开展小型且快速的探索性调查,使用纸版问卷调查也能得到不错的数据结果。另一方面,数据质量提高的条件则更复杂,包括四方面:一是操作程序设计具有恰当性、准确性和稳定性,二是受访者对电子设备具有一定接受度和信任度,三是访员正确理解和操作,四是互联网的稳定性。

因此,在缩减系统性误差方面,CAPI 主要存在两方面的问题使其成效折损:一是由于无法满足上述条件而引起的,二是 CAPI 本身的限制导致的。对于前者,根据条件的类型,可以将面临的问题归结为三类:一是技术和设备的缺陷导致的问题,如系统稳定性导致数据等资料的缺失和中途拒访、访问过程的标准化与僵化、并行数据难以有效利用等;二是受访者的反应性问题,如电子设备可能导致不信任感上升、设备损毁带来的高风险等;三是技术和设备的使用对首次使用 CAPI 的调查而言带来了新的问题,如增加了问卷设计等环节的压力、访员培训成本高等。对于后者,作为入户调查的基本特性使得 CAPI 本身还无法消解所有的系统性误差,例如,CAPI 无法完全消减访员特征对调查误差的影响,尽管 CAPI 系统的使用使得这个时间差在缩短,也无法完全解决数据回传和实际访问间的时间差。

其中,对于因无法满足前提条件而引起的三类问题,第一类问题是亟待解决的关键问题。现有研究表明,CAPI 的成功在很大程度上依赖于编程和测试应用程序上的努力以及对底层数据管理和传输系统的建设^[2]。为此,可以积极开展学科间合作以及调查机构和相关企业的合作,以针对性地解决技术上的限制。而第二、三类问题,可以随时间的推移而得到缓解。一方面,随着互联网的普及,民众对电子设备的了解程度逐渐加深,这使受访者更容易信任电子设备,同时对访员计算

机技能的培训也大有助益;另一方面,大量研究表明,由于软件和设备可以反复利用,准备阶段的压力和成本会随时间的延长而降低^[10]。也就是说,随着时间的推移,现代信息技术的普及会提高民众对电子设备的接受度和信任度,而调查和技术经验的积累会缩减调查的物质和时间成本,从而使第二、三类问题得到缓解或解决。

参考文献:

- [1] Banks R, Laurie H, “From Papi to Capi The Case of the British Household Panel Survey”, *Social Science Computer Review*, Vol.18, No.4, 2000, pp.397-406.
- [2] Caeyers B, Chalmers N, Weerd J D, “Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment”, *Journal of Development Economics*, Vol. 98, No. 1, 2012, pp.0-33.
- [3] 丁华《计算机辅助调查与访员引起的调查误差控制》,《统计与决策》2017年第19期,第36-39页。
- [4] 杨清《统计数据质量研究新思路——误差研究》,《统计研究》2000年第8期,第33-37页。
- [5] Nicholas Bateson, Paul Hunter, “The Use of CAPI For Official British Surveys”, *Bulletin of Sociological Methodology*, No.30, 1991, pp.16-26.
- [6] 风笑天《浅谈当前抽样调查中的若干失误》,《天津社会科学》1987年第3期,第47-51页。
- [7] Groves R M, “Survey Errors and Survey Costs”, *Journal of the Royal Statistical Society*, 1991.
- [8] Baker R P, Bradburn N M, Johnson R A, “Computer-assisted personal interviewing: An experimental evaluation of data quality and cost”, *Journal of Official Statistics*, Vol.11, No.4, 1995, pp.413-431.
- [9] 尹海洁、刘耳《利用 SPSS 消除统计数据录入误差的一种方法》,《西安交通大学学报》(社会科学版) 2003年第3期。
- [10] Bhatia M S, Jaiswal A, 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence) - “Empirical analysis of data acquisition techniques: PAPI vs CAPI” 2016, pp.326-330.
- [11] 任莉颖、严洁《并行数据与社会调查质量探讨》,《统计与决策》2014年第6期,第27-32页。

[责任编辑:朱磊,张斐男]