

体育社会科学研究中问卷调查法使用的几个问题

李曼

(温州职业技术学院 公共教学部, 浙江 温州 325035)

摘 要: 问卷调查法是体育社会科学研究中的一种重要方法。但在问卷调查法的应用中, 其样本容量、效度和信度检验、问卷有效回收率统计等方面尚存在一定问题。样本容量不可随意, 应按照社会统计学的要求科学计算; 问卷的效度和信度检验不可或缺, 结构效度可以忽略, 内容效度必不可少, 重测(再测)信度应避免操作上的误区; 统计问卷的有效回收率, 在任何情况下都应该以发放的问卷份数作为分母。

关 键 词: 体育社会学; 问卷调查法; 社会调查

中图分类号: G80-05 **文献标志码:** A **文章编号:** 1006-7116(2013)05-0068-04

Several issues about the use of questionnaire survey in researches on sports social science

LI Man

(Department of Physical Education, Wenzhou Vocational & Technical College, Wenzhou 325035, China)

Abstract: Questionnaire survey is an important method used in researches on sports social science. However, there are certain issues in its application, such as sample size, validity and reliability test, valid questionnaire return rate etc. Sample size cannot be set at will, should be calculated according to social statistical requirements; questionnaire validity and reliability test is indispensable, structure validity is ignorable, content validity is indispensable, retest validity should be free of operational errors; the valid return rate of statistical questionnaires should be calculated by using the number of questionnaires sent out as the denominator under any circumstance.

Key words: sports sociology; questionnaire survey; social survey

20 世纪 80 年代出版的美国布莱洛克^[1]的《社会统计学》和北京大学卢淑华^[2]的《社会统计学》, 及之后出版的袁方^[3]《社会研究方法教程》和范伟达^[4]《现代社会研究方法》, 作为前置限制词的核心概念都是“社会”, 前者对资料收集技术和统计分析技术有所侧重, 兼顾了研究设计和研究方法等; 而后者则对研究原理、研究程序、研究设计和研究方法有所侧重, 兼顾了资料收集技术和统计分析技术。可以说, 这两类著作都是针对社会研究而言的指导性工具书。社会研究(Social research)是个比较宽泛的概念。它既包括社会科学各学科的研究, 也包括各个社会工作部门的研究。在一般情况下, 对这一概念并不作严格的定义^{[3]20}。国内出版的《中国大百科全书·社会学卷》对社会研究做了如下释义: “社会研究(Social Research)是运用科学方法对社

会生活现象加以了解、说明和解释的一切活动。它是人类社会为对象, 以科学方法为手段, 以解释和预测为目的, 以科学理论和方法论为指导的一个完整的过程。”^{[5]347} 社会研究有很多方法, 但问卷调查法是最常用的一种收集资料的方法, 特别是在社会调查中, 它的使用更为普遍。因此, 美国社会学家艾尔·巴比称“问卷是社会调查的支柱”, 英国社会学家莫泽说“十项社会调查中就有九项是采用问卷进行的”。可见社会研究者对问卷法的评价之高^{[3]231}。改革开放以来, 随着体育事业的迅速发展, 问卷调查法作为收集资料的一种重要方法在体育社会科学研究中得到了广泛的应用, 为描述、说明和解释体育这一社会现象做出了重要贡献。但在问卷调查法使用的过程中也出现了一些问题, 直接影响到研究结果的可信度。本研究试图通

收稿日期: 2013-01-31

作者简介: 李曼(1979-), 男, 讲师, 研究方向: 体育社会学、体育教育训练学。

过对样本容量、效度和信度检验、问卷有效回收率统计等问题的讨论,为规范问卷调查法的使用,促进研究结果可信度和体育社会科学研究质量的提高提供参考。

1 关于样本容量问题

问卷是要发放给被调查者的,发放给哪些被调查者,发放多少,这就涉及到抽样和样本容量的问题。为此,“抽样—问卷—定量分析三者的结合体,是现代社会科学定量研究中最常见,也是最重要的一种方式”^{[3]266}。也就是说,问卷调查必须与科学的抽样技术相结合才能得到可靠的研究结果。体育社会科学研究亦不例外。但是,这一点并没有引起研究者的足够重视,特别是对样本容量的确定随意性很大,许多研究都没有按照社会统计学的要求去确定样本容量,甚至120人或180人也要去代表具有相当规模的研究总体,例如某项关于广州市居民亚运会影响感知的研究中,以天河、越秀、海珠、番禺等4个区随机发放问卷180份(回收有效问卷151份)作为调查样本。实际上,广州市2010年第六次全国人口普查主要数据公报披露的天河区常住人口为1 432 431人、海珠区为1 558 663人、越秀区为1 157 277人、番禺区为1 764 869人,合计为5 913 240人。相对于590多万人,180个样本有代表性吗?例如某项关于某省成年女性身体素质现状及体育行为特征的研究,从24个地市(县)中抽取23 088人作为调查对象。那么是不是样本越多,其代表性就越好呢?

样本容量又称样本大小、样本规模,指的是样本内所含个体数量的多少。样本容量的确定是抽样设计的最重要的内容之一^{[3]224}。确定样本容量,首先,要确定样本的精确度(置信区间)和置信水平(置信度)。样本精确度(置信区间)表示允许误差的范围,社会学研究允许误差的范围为1%~7%;置信水平(置信度)则表示对研究结论正确性的把握程度,社会学研究设置了两个置信水平,即95%和99%。其次,要清楚样本容量和总体的关系。样本容量和总体二者不是直接关系,而是缓慢的曲线关系。范伟达^{[4]210}给出了不同总体规模所需的样本数量,可供确定样本容量时参考。

根据范伟达确定样本量的方法,在置信水平(置信度)为95%,置信区间(精确度)为 $\pm 3%$,总体异质性较大的条件下,1 000人的总体需要350人的样本,10 000人的总体需要600人的样本,而100 000人的总体只需要680人的样本,当总体为500 000人的时候,样本量只增至700人。可见,样本量并非随着总体的增加而相应增加,也不是样本越大越能代表总体,如美国民意测验的样本数一般在1 600~2 000人,最多不超

过3 000人^{[3]225-228},中国十大社会阶层调查样本为6 240人^{[6]363}。社会学家风笑天^[7]指出,在95%的置信度下按抽样绝对误差不超过3%的要求进行计算,需要抽取样本量: $n = u_{\alpha}^2 p(1-p)/d^2 = 1.96^2 \times 0.25 \div 0.03^2 = 1 067$ (d 为抽样绝对误差取0.03, u_{α} 在置信度为0.95时为1.96, $p(1-p)$ 最大取0.25)。在采用多阶段的复杂抽样时,设计效应(deff)一般会在2和2.5之间。假如我们要将某一大城市接近1 000万人的户籍人口作为研究总体,而且需要采用多阶段的复杂抽样,则可将设计效应(deff)确定为2,设计效应和按公式计算出的样本量是倍数关系,即 $1 067 \times 2 = 2 134$ 。这样,即可初步确定对该大城市市民调查的样本容量是2 134人。考虑到可能会有20%的被调查者或拒访、或无回答等的情况,故实际调查所用的样本量应为 $2 134 \div 80\% = 2 668$ 人。唐盛明^{[8]154-155}指出:当总体为1 000万或1 000万以上的时候,样本比例的增加实际上不产生作用。除非有特殊的要求,样本的数目一般无需超过2 500人。

2 关于问卷的效度和信度检验问题

问卷是由围绕某一主题的相互关联的测量指标组成的测量社会现象的测量工具。任何一项将要使用问卷调查法的社会研究,在遵循社会研究的规范(原则)完成问卷设计之后,都面临着该问卷的效度和信度问题,即有效性和可靠性的问题。使用未经效度和信度检验或未能正确进行效度和信度检验的问卷,将会使研究结果的有效性和可靠性大打折扣。范伟达^{[4]130}指出:“效度是指用测量工具测出变量的准确程度,即准确性。”也就是说:“在用指标去度量体育这一社会现象时,有一个该指标能否反映测量目标与测量结果的一致性问题。”^{[9]36}一般效度检验的方法有内容效度、准则效度(亦称效标效度与统计效度)、建构效度(亦称理论效度)和因素分析效度等,最常用的是内容效度。信度是指作为测量工具的问卷的稳定性与可靠性问题。“所谓信度是指测量数据(资料)与结论的可靠性程度,即测量工具能否稳定地测量到它要测量的事项的程度”^{[3]187}。“测量的信度通常以相关系数来表示。由于测量分数的误差变异之来源有所不同,各种信度系数分别说明信度的不同层面而具有不同的意义”^{[4]134}。

2.1 被忽略的效度与信度检验

在体育社会科学研究中,只要使用了问卷调查法,都应该对其效度和信度检验做出必要的说明。然而,也不乏有被研究者忽略的。例一:某作者关于潜优势射箭项目青少年训练问题的研究,使用了问卷调查法,并将回收的问卷数据利用SPSS软件进行了统计处理,但却没有交代问卷的效度与信度检验。为此,其数据

结果的有效性和可靠性令人生疑。例二：某作者关于体育教师绩效考核指标体系的研究，使用了问卷调查法，说明了回收率和有效回收率，但无问卷效度与信度检验的说明。例三：某作者关于高师体育教育专业某门课程学习评价方法的研究，使用了问卷调查法，也没有说明是如何进行效度与信度检验的。例四，某作者关于某省乡镇社区体育的调查研究，在139个区、县、市中获得调查问卷1594份，但也没有说明问卷的效度与信度检验。

2.2 结构效度与内容效度孰是孰非

在一次旁听某大学体育学硕士论文答辩时，有专家在答辩人(体育社会科学的论文，非心理学研究)已经说明对问卷进行了内容效度检验的情况下，提出了还应该进行问卷结构效度检验的质疑。我当时的理解是，问卷的内容效度是指问卷的有效性，结构效度无非就是指问卷设计的结构是否合理。后来，看到多数体育社会科学(非心理学研究)的文章在运用问卷调查法时，大多使用内容效度评价，但也有采用结构效度评价的。例如，某作者关于某省学校体育现状调查与发展对策的研究，没有对问卷进行内容效度检验，而是请专家对问卷进行了结构效度检验，但文中并未说明该“结构效度”指的是什么。正是带着这样的问题本人开始学习了解问卷的结构效度和内容效度，并有了一个初步的认识。百度百科对结构效度的定义是：“结构效度是指一个测验实际测到所要测量的理论结构和特质的程度，或者说它是指测验分数能够说明心理学理论的某种结构或特质的程度；是指实验与理论之间的一致性，即实验是否真正测量到假设(构造)的理论。”^[10]结构效度的概念最初是由美国心理学会(APA)、美国教育研究协会(AERA)和美国国家教育测量协会(NCME)(1954)联合委员会提出，发表于《心理测验和诊断技术的技术建议》。此后，在半个世纪的发展中，结构效度的定义经历了许多微妙的变化。结构效度(Construct Validity)亦称(构念效度)。结构是个心理学概念(或称构念)，是对某种不能被测量而且不能直接观察到的人类行为所做的理论定义。例如，智力、成就动机、紧张、成绩、态度和阅读理解等等。换言之，结构是一个人潜在的心理属性，是不能直接被观察和测量的一种或一组能力，但是能够通过测试成绩做出推断^[6]。在实践中，结构效度多用于心理学的研究或是考试(教育测量)。传统的结构效度检验是实验或考试结束后的实证研究。到目前为止，尽管对试前和实验实施阶段的理论验证已经引起了学者们的重视，但多数的结构效度验证仍然是在实验或考试后的实证研究。这一点与社会研究对内容效度检验的要求有很

大的不同，内容效度检验要求在前，而不是事后的测试。要检查问卷内容是否能达到测量目的与测量结果的一致性，即是否有效，只能依靠多数专家的共同意见。因此，至少目前在体育社会科学研究中的问卷调查法使用中(心理学研究除外)是可以忽略结构效度的。

2.3 重测(再测)信度是指全部还是局部

社会科学研究中常用的信度检验方法有重测信度(亦称再测信度)、复本信度、折半信度等。通过查阅期刊论文和博士硕士论文可知，重测信度在体育社会科学研究中的问卷调查法使用中得到了最广泛的应用。重测(再测)信度是指用同一份问卷，对同一被测群体前后调查两次，再根据调查结果计算(第1次 X 和第2次 Y)的相关系数。但在现实研究中，常常遇到的问题是，利用重测(再测)法进行信度检验的第2次调查(重测)是针对全部被调查对象，还是部分被调查对象的问题。而许多研究恰恰回避了这一问题，即在使用重测(再测)法进行信度检验时，没有说明进行第2次问卷调查时，是针对全体被调查对象，还是只针对部分被调查对象。此类事例较多，这里简举3例：例一，某作者关于某省城市老年人体育生活方式的社会学分析一文，使用了问卷调查法，发放问卷600份，其信度检验用的是再测法，信度系数 $R=0.82(P<0.05)$ ，但没有交代两次调查的时间间隔以及第2次调查的发放对象是全部还是部分。例二，某作者关于从某城市体育人口的性别结构透视其妇女体育开展的研究，发放问卷3000份，采用再测法，间隔1个月，用同样问卷分别进行2次测试，相关系数 $R=0.85$ 。但没有具体说明第2次测试是针对3000人，还是针对部分人。

重测(再测)信度一般要求对被测群体全体成员前后进行两次调查。虽然《社会研究方法》或《社会统计学》的书中并没有给出使用重测(再测)法进行第2次问卷调查时可以选取部分被调查对象，但有时因为样本量较大或很大，一些研究者在第2次调查时，只随机抽取部分被调查对象。这种做法虽然亦可接受，但必须注意避免操作上的误区。例如，第2次调查随机抽取了50名被测群体成员，然后再从第1次调查被测群体全体成员中随机抽取50名成员，根据其两次问卷填写的结果计算其相关系数，就陷入了误区。因为无法保证这50名成员中的每个成员都是自己前后填写的两份问卷，其中有不少是“张冠李戴”的，无法测量出真实的问卷信度。正确做法的关键是第1次向被测群体全体成员发放的问卷一定要编号。第2次随机抽取的50名被测群体成员的问卷也要编号，假如第2次随机抽取的50名成员为51~100号，则应是51 Y 、52 Y 、53 Y 、54 Y ……100 Y 。问卷回收以后，将第1

次回收的问卷为 X , 第 2 次回收的问卷为 Y , 从而形成 51 X -51 Y 、52 X -52 Y 、53 X -53 Y ……100 X -100 Y 两组相对应的问卷。由于多数问卷都是由围绕主题的相互关联的问题组成的, 还必须将其转换成数量关系方能计算其相关系数。一般可采用对封闭式答案赋值的办法, 以被测成员 51 号为例, 在其第 1 份问卷中可以得到若干 X 的分值, 在其第 2 份问卷中可以得到若干 Y 的分值, 总加之后则得到 $\sum X$ 与 $\sum Y$ 两个分值。

3 关于问卷有效回收率统计问题

问卷的有效回收率统计是使用问卷调查法的重要一环。但有的文章在统计问卷的有效回收率时存在错误。例如, 某作者关于优秀运动员退役原因的研究, 该文在其调查过程中写到: “问卷共发放 560 份, 回收 541 份, 回收率 96.6%, 有效回收率 100%”。这里必须强调的是, 有效率可以和回收率相等, 但有效率高于回收率则有悖常识。卢淑华^{[2]562} 强调指出: “发放的问卷数应不少于样本的容量, 但在任何情况下都应该以发放的问卷份数作为分母, 而不是样本容量。回收率(笔者代注: 指有效回收率)不得少于 70%, 否则将严重破坏随机抽样的原则。”据此, “在任何情况下都应该以发放的问卷份数作为分母”的情况下, 该问卷的有效回收率也应该是 96.6%, 而不是 100%。也就是说, 100%有效只是相对回收的 541 份问卷全部有效而言, 但整个问卷的有效回收率是以发放的问卷份数作为分母, 以回收的有效问卷数作为分子来计算的(541/560), 因此该问卷的有效回收率也应该是 96.6%。

再如, 某作者关于体育赛事影响的论文在问卷发放与回收中写到: “共计发放问卷 1 084 份, 回收问卷

921 份, 回收率为 84.96%, 其中有效问卷 784 份, 有效回收率为 85.12%”。有效回收率高于回收率。此例中统计的错误在于, 统计有效回收率时没有将发放问卷数作为分母, 而是将回收问卷数作为分母, 从而导致有效回收率(85.12%)比回收率(84.96%)还高。此例中统计的正确结果应该是有效回收率 72.32%。

参考文献:

- [1] 布莱洛克. 社会统计学[M]. 傅正元等, 译. 北京: 中国社会科学出版社, 1988.
- [2] 卢淑华. 社会统计学[M]. 北京: 北京大学出版社, 1989.
- [3] 袁方, 王汉生. 社会研究方法教程[M]. 北京: 北京大学出版社, 1997.
- [4] 范伟达. 现代社会研究方法[M]. 上海: 复旦大学出版社, 2001.
- [5] 中国大百科全书总编辑委员会. 中国大百科全书·社会学卷[M]. 北京: 中国大百科全书出版社, 1991.
- [6] 陆学艺. 当代中国社会流动[M]. 北京: 社会科学文献出版社, 2004.
- [7] 风笑天. 再谈样本规模和调查回收率[J]. 社会学研究, 2007(6): 195-212.
- [8] 唐盛明. 社会科学研究方法新解[M]. 上海: 上海社会科学院出版社, 2003.
- [9] 吕树庭, 杨霆, 车建平. 体育社会测量方法[M]. 广州: 广东高等教育出版社, 1998.
- [10] 百度百科 结构效度. [EB/OL]. <http://baike.baidu.com/view/1350460.htm>.

