

研究假设的有效性及其评价

孙 健 敏

Abstract: One of the important characteristics of empirical research is hypothesis testing. This paper discusses the necessity and functions of research hypothesis in empirical research and proposes five criteria for the evaluation of research hypothesis and the operational steps.

一、引言

社会科学领域的研究方法问题,已经引起我国学者的广泛关注。特别是在社会学和管理学领域,采用问卷调查、统计分析、实地研究、实验研究等方法的论文数量不断上升,在20世纪90年代定量研究风起云涌的基础上,近年来采用焦点小组、深度访谈的定性研究也有上升的趋势,关于研究方法本身的讨论也日渐增多(风笑天、田凯,1998;风笑天,2000;林彬、王文韬,2000;侯龙龙,2001;陈向明,1996,1997)。在分析总结国内实证研究的基础上,学者们得出一个结论:我国的实证研究,或者没有充分利用数据,或者没有恰当利用数据,数据与理论脱节,导致研究失去了本来的意义和价值。于是,有人对实证研究的价值提出了质疑,好像实证研究不可能产生理论贡献。

笔者认为,我国的社会科学研究,应该大量提倡实证研究。导致数据与理论脱节的一个主要原因是忽视了研究过程中假设的作用。风笑天曾经呼吁研究者重视研究设计(风笑天,1999)。在研究设计中,形成假设是一个关键的环节。假设的好坏,对于研究质量和水平具有决定作用,对于数据的利用具有指导作用。一项研究的意义,在很大程度上取决于假设的潜在价值。

即使在社会学发源地之一的美国,实证研究也是在不断完善中,直到20世纪60年代,才有大量的运用假设检验的研究。韦尔斯和皮寇(Wells & Piou, 1981)曾经分析了1936年到1978年间发表在《美国社会学评论》(ASR)上的学术论文,真正运用理论框架进行研究设计并进行假设检验的论文数量,从1965年以后才开始达到了54%以上。当然,这种建立在演绎推理前提下的实证研究,并不是一花独放,进入20世纪80年代以后,大量研究者的兴趣又回到了定性研究的归纳推理上。

国内学者也曾对发表在《社会学研究》上的论文进行过统计分析。风笑天等(1998)统计了1986—1996年间《社会学研究》发表的经验研究报告,其中调查研究方式118篇,占63.4%。刘少杰(2000)的分析中涉及到论文有否理论前提,统计表明1987年到1998年11年间发表的统计调查报告中,只有29%的文章有明确的理论前提。笔者不禁有疑问:没有理论前提怎么能做实证研究(经验研究)呢?

实际上,科学方法的真正作用,不在于其具体的收集资料或分析资料的方法,而在于如何证明研究者自己的观点或命题,用什么逻辑来证明自己的观点或命题。著名社会学家理查德·鲁德纳(Richard Rudner, 1966: 5)指出:科学学科的方法论不在于它的手段或技术,而是它的证明逻辑(logic of justification)。确实,科学方法是基本根据,以此为基础来决定接受或拒绝假设或理论。相应的,认为社会科学在方法论上不同于非社会科学,并不是简单地认为社会科学使用不同的探究方法,而是一种全新的观点:社会科学需要一种不同于非社会科学的探究逻辑。要接受这种观点,就需要否定另一种观点:所有的科学在接受或拒绝假设时都以共同的证明逻辑为特征。显然,这个认识论范畴的命题并没有得到充分的证据。

二、假设的含义与作用

实证研究的一个重要特征是假设检验。一般的研究过程(包括定量研究和定性研究),包括提出问题、确定关键变量、形成假设、收集相关信息或数据、检验假设、形成(修正)理论、提出新问题(Creswell, 1992)等环节,其中形成假设是关键步骤。

关于假设的含义和作用,学者们有不同的看法。普遍的看法是当一个命题是用来做实证检验的,这个命题就变成了假设。假设是一个陈述,具有临时性和探索性。从理论上讲,凡是没有经过科学验证的观点或命题,都属于假设。例如:是否接受高等教育对一个人的收入水平有影响;提高公务员的报酬水平可以减少腐败现象;学历与工作满意度相关;工作中承受的压力越大,员工的工作满意度越低;中国和美国的员工不同等等。从科学研究的角度看,这些命题都属于假设。实证研究就是要去检验假设的正确性。所以,所有的实证研究,都应该首先有假设。

假设是表明两个或多个变量之间关系的一个或一组陈述句(Mason & Bramble, 1989)。假设可以定义为从逻辑上推测的两个或多个变量之间的关系,以一种可检验的陈述句形式表达出来(Sekaran, 2000)。假设是把变量赋予个案的陈述,在这个意义上,个案可以定义为某个假设要谈论的一个整体或一件事情(Cooper & Schindler, 2001)。例如,总裁的成就动机比平均水平高,或者A公司的管理层成就动机比平均水平高。著名研究方法专家科林格(Kerlinger, 1979)认为:假设在科学研究中比结论本身重要,它是增进知识的有力工具。查威茨(Chavetz, 1978)在论述社会学理论的建构和检验时,专门阐述了假设在研究中的作用,以及假设检验的必要性。于是,这就形成了实证研究只能是检验假设这一普遍看法。但是,莫顿(Merton, R. K., 1948)并不完全同意实证研究的惟一功能是检验假设或验证假设,他认为实证研究至少有4个功能:指导研究方向;区分与问题有关或无关的事实和变量;选择最合适的研究设计或方式;形成结论的组织框架。笔者比较认同这个界定。

假设可以来自于已有的理论,或者别人的研究,或者系统观察。

三、假设必须满足的标准

什么样的假设算是有效的假设呢?如何评价一项研究的假设是好是坏呢?萨尔坎德(Salkind, 2000)提出衡量假设好坏的5个指标:以陈述句而不是疑问句的方式表达;说明变量之间的关系;以某个理论为基础或以一系列的研究文献为基础;简短精练;可检验。科林格和李(Kerlinger & Lee, 2000)提出了3个标准:第一,对目的的合适性:限定条件、范围、变量的分布等。具体表现在:假设揭示了最初的问题条件吗?假设清楚地区分并界定了相关的和无关的事实吗?假设是否严格地界定了对研究问题有意义的变量?假设是否清晰地从量上确定了变量的条件、范围、分布?假设是否为需要解释的事实提供了答案?假设是否隐含着某种研究设计最合适?第二,可检验性。具体表现在:假设提供了组织结论的框架吗?假设使用了现有的技术吗?是否需要解释与现有的物理和心理定律不一致的地方?是否揭示了能够从检验目的演绎出来的结果或变异?第三,比其他理论好。具体表现在:简单、限制条件或前提少;能解释更多的事实;适用的范围更广;人们都能认同。

上述解释都有其道理,但存在两个问题:一是没有完全解释研究假设的内在特性,二是对于评价好坏的条件的论述不够充分。笔者根据自己的研究经验,认为有效的研究假设应该具备5个条件。

1. 假设必须在概念上清晰明了,表述必须准确

假设的概念必须有明确的定义,可能的话应该给出操作定义。而且,这些概念应该是大众都能接受的,可以进行沟通的,而不是小圈子中的概念。

因为大多数的社会和行为现象都被看成是不可观察的概念之间的结构作用过程,研究者的任务就是揭示这些过程的规律和内在机制;而概念是抽象的东西,无法直接观察。因此,对这些概念的研究必须借助于某些可观察的指标。对指标的选择,在很大程度上决定了概念的内涵和外延。因此,概念的界

定必须十分明确。在可能的情况下,必须通过可观察的指标来进行定义,也就是操作定义。需要说明的是:观察的指标并不是概念本身,只不过是概念可能的表现形式而已。例如,研究者可能对员工积极性比较感兴趣,而员工积极性是一个抽象概念,无法直接观察,但可以找出一些能够体现员工积极性的指标进行观察,如按时上班、不缺勤、工作主动、工作质量高等。

把所有的概念和操作都明晰化,意味着不仅概念要清晰,而且对所采用的指标要进行描述和说明。例如,假设新员工的工作满意度比老员工要高,我们不仅需要定义工作满意度,而且要定义新员工和老员工;不仅要有明确的定义,而且要给出指标:什么是满意,什么是不满意(对薪酬、工作、同事、上司),什么是新员工,什么是老员工(工作多少年,多少月)。社会科学研究中,大量的指标需要研究者自己选择和确定,而指标的选择和确定在一定程度上也就决定了研究的质量和水平。

澄清概念的一种简单办法是把研究框架中所使用的概念列出一个清单。然后,从3个方面对这些概念给出定义:①从语义上;②用特定的操作术语(如观察次数,平均数的计算等);③参照以前研究的其他概念。还可与本领域的学者或同事讨论这些概念。通过讨论往往会发现,以前认为简单的概念实际上包含很多意义,不同的人在理解上是有分歧的。这样,就不难确定哪个概念是清楚了。

假设的表述必须严谨,意义必须明确。仅有清晰界定的概念不够,还必须保证概念之间的关系能明确地表达出来。

假设的表述方式有很多种,不同的表述,意味着研究问题的不同。著名教育学家和方法专家克瑞斯威尔(Creswell, 1992: 74)把研究假设划分为4种类型:一是语意虚无假设,以概念为基础,没有方向性。例如,受教育程度高低与一个人的收入没有关系。二是语意替代假设,以概念为基础,有方向性。例如,一个人受教育程度越高,收入越高。三是操作虚无假设,以操作性定义为基础,没有方向性。例如,本科大学生和硕士研究生的收入没有差异。四是操作替代假设,以操作定义为基础,有方向性。例如,硕士研究生的收入比本科生的收入高。毫无疑问,不同的表述方式形成了不同的假设,对研究者而言,需要用不同的方法来检验假设,最终形成的结论也就不一样。

2. 假设应该具有实证参考物

研究者早就指出过,科学概念必须有最终的实证参照物。假设应该是在实证基础上做出来的,是对客观事实的关系的陈述。这种陈述应该有客观的标准来衡量,而且这个标准本身是可以观察的,可以作为验证假设的效标。

有用的假设不应该包含道德判断的成分。诸如“罪犯不比商人更坏”、“女性应该有自己的事业追求”、“私有化才是出路”、“股份制是国有企业走出困境的良药”等陈述。不过“白天过去就是黑夜”之类众所周知的命题,根本算不上假设,言之无物,等于没说。换言之,虽然假设不可避免地包含价值判断的研究,但这不应该成为研究者的目的,而应该努力从研究假设中分离出去。

通过下面的例子,我们可以看到明确的假设表述中实证参照的重要性。有人要研究宗教信仰与离婚的关系。研究者把参加宗教礼拜作为宗教信仰的观测指标。那么,关于信仰宗教与离婚的关系就可以用下面的多种假设来表述,也就形成了不同的假设(Newman, 1997)。

- ①参加宗教礼拜是减少离婚的原因。
- ②参加宗教礼拜导致离婚减少。
- ③参加宗教礼拜与减少离婚有关。
- ④参加宗教礼拜影响到离婚的减少。
- ⑤参加宗教礼拜与降低离婚率有关联。
- ⑥参加宗教礼拜导致离婚的可能性下降。
- ⑦参加宗教礼拜次数越多,离婚的可能性越低。
- ⑧参加宗教礼拜会减少离婚的人数。
- ⑨如果人们参加宗教礼拜,那么离婚的可能性就会降低。
- ⑩参加宗教礼拜可以降低离婚率。

显然,所有的假设都有可衡量的标准,但并没有进行价值判断或评价。

在实际操作中,首先,对那些表达态度而不是描述或参考实证现象的概念进行分析,注意那些关键性限定词,如“应该”、“必须”、“只有”、“好坏”等等。然后把这些表述转化成更有用的概念。例如,“不称职的老师”是一个评价性概念,研究者应该在心中有个清楚的定义:什么样的老师是不称职的老师?如,

口齿不清楚、表达不准确？知识的深度和广度不够？备课不认真？对学生态度不好？不能贯彻教书育人的要求？衣冠不整？等等。只有定义清楚，才能进行研究。“应该”或“必须”也属于价值判断术语，在假设中应该避免。再如，研究者要回答的问题是女性为什么应该有自己的事业追求，女性追求事业可以避免对丈夫的完全依赖，或可以具有独立的经济地位或社会地位等。这是研究假设。

有时我们会发现，我们所依赖的参照物不过是一种朦胧的感受，不能给出操作定义来进行观察，这时候就需要进一步研究问题，找出我们真正想要发现的东西。如近几年流行的“以人为本”的管理，究竟什么是以人为本？要回答这个问题，首先要界定“人”的内涵，然后再来界定什么是“以人为本”。

3. 假设必须具体化

由假设所表明的所有操作和预测必须明确表达出来，这样，才可能评价假设检验的有效性。从实际情况看，假设的表述常常是比较笼统的术语，包括的范围太大，或过于抽象，使得检验假设成为不可能的事。之所以会出现这种情况，往往是因为宏大的观念容易产生诱惑力，让人们觉得很有份量，很重要，容易产生轰动效应。实际上，科学研究不可能一下子就解决一个大问题。无论是学者还是学生，应该尽量避免这种倾向，养成严谨的习惯，培养扎实的技能，通过更切实的概念，来提出自己的观点或主张。

这样的具体格式化要求可以保证研究的可行性、实用性，可以提高研究结果的效度，减少犯错误或无效劳动的几率。因为，所提出的定义或概念越宽泛，越容易使用选择性证据来证明自己的假设，也更容易陷入模棱两可的误区，这样的研究结论显然是不充分的。大多数的算命先生就是利用这种技巧来蒙骗人。他们给出的命题一般都比较宽泛，可以从不同的角度去理解，这样，无论你提供什么证据，都会落入他的“预测范围”。毫无疑问，这不是科学研究的态度和作风。我们可以用统计的观念来表述：预测越具体，靠偶然的运气实现的概率越小；反之，预测越概括，靠运气获得证据的概率越大。科学预测或科学假设当然要通过清晰具体的界定来避免使用选择性证据。

如何做？决不要满足于普遍性的预测或假设，如果可能分解成更精细的假设，一定要进行分解。例如，“中国的未来一定会更美好”，这样的假设或预测没有任何科学性，因为它过于一般化了。我们需要具体的假设，需要具体到时间、地点、方式、程度等。只有把这些概念界定清楚，才可以进行研究，或者说研究的结论才有价值，才可能建立理论框架。一般来说，做到这一点并不难，可以对命题进行概念分析，或建立相关假设。如“未来”是什么意思？是指今后的3年、5年、10年？美好是什么意思？是指健康？美丽？富有？长寿？中国是什么意思？是指作为一个国家？或者是她的人民？再进一步，健康或美丽的指标是什么？可能的相关假设有：在未来5年内，中国人民的平均寿命将增加1岁，或国民生产总值将增加100%。这样的假设，第一，可以澄清数据和结论之间的关系，使数据更有意义；第二，使研究任务更加可操作。

4. 假设应该与可利用的技术相关

学术界早就强调理论与方法是相关的，也就是说，方法的水平在一定程度上决定了理论的水平。不知道什么方法可以用来检验假设的研究者很难构建出有用的理论来。

这并不是一条绝对的指令或规定，因为目前，很多复杂的假设还不能用现有的技术进行检验，但这并不意味着这些假设就没有价值了。我们这里所说的方法与理论的相关只是一种评价要求，在早期阶段为了判断是否能够进行研究，可以应用于所有的问题。

有时，虽然是不可能的假设，但其某些方面的特点却使得形成这样的假设是值得的。例如，如果一个问题非常重要，可以把它作为参考框架，那么就值得建立这样的假设，而不必考虑现在是否能够进行检验。最典型的例子是马克思关于政治经济学的许多假设，他的假设并没有得到数据的支持，无论是当时还是现在都不具备必要的技术。然而，马克思的理论框架是许多更精细的、可证明的命题的重要资源。这一点对涂尔干的自杀理论同样成立，对涂尔干社会内聚力的假设非常有用。上述两位理论家的研究对社会学非常重要，虽然他们的理论还不可能用现有的技术进行验证。而且，提出不可能的问题会刺激技术的发展。现代研究技术发展的推动力之一来自于对一些重要研究的批评，这些研究因为技术

方法的局限被认为是没有意义的或不够严谨。在实证传统占优势的社会科学界,这样的思潮更加盛行:如果你不能用科学的方法证明你所提出的问题,那么,尽管这个问题可能很重要,从研究的角度来看它也会被认为是没有意义的。虽然这样的观点有些偏激,但作为规范社会科学研究的途径,尤其是对于学生和年轻的研究人员,从技术手段上保证研究课题的相关性是进行科学研究的基本功。

因此,需要查阅对同一个问题进行过研究的报告或论文,找出别人研究中测量重要因素或变量所使用的各种方法,并熟悉这些方法的应用情况。如果还有不清楚的地方,那么在进行正式的研究之前,最好先把必要的研究技术弄明白。也许,这种研究技术上的欠缺,可能意味着你要研究的问题过于庞杂,以你现有的能力和资源不足以应对。在查阅别人使用的研究方法时可能会发现,有些问题是用不同方法进行研究的。这时候需要分析:为什么一项研究采用这种方法而没有采用那种方法?不同方法有什么优劣?研究者对方法做过什么改进?哪种方法可能对你的研究更有用?在查阅别人的研究报告时要注意有关的批评意见,发现别人研究的弱点或不足,以便在自己的研究中克服或避免。

有时可能会发现,有些问题的研究不是采用严格的测量方法,研究设计也不规范。这时候可以用两种办法:一是考察这些文献本身,看为什么会是这种情况,是否这个问题过于宽泛,可以分解成几个更小的、容易操作的题目?二是看文献的出处,是什么性质的刊物发表的论文?是严肃的学术期刊还是流行的大众读物?需要说明的一点是国内社会科学(广义)领域的研究中,总的来说缺乏科学严格的规范,如从问题的提出、方法的说明、数据的收集与分析、结论的形成,一直到参考文献的罗列,都不够规范。

5. 假设应该与现有的理论体系有关

这个准则往往被刚进入研究领域的学生所忽视,尤其是硕士和博士生们。在做论文时,他们更喜欢选择那些自己“感兴趣的”或“重要的”题目,而不太考虑类似的题目别人做过什么研究,得出过什么结论,我的研究是否真的有助于反驳、证明、或支持已有的理论或观点。科学是一种积累,是在继承中的创新。这种积累只有每个研究者把自己的研究建立在已有研究结论、理论观点、或事实证据的基础上才能实现,如果每个研究都是独立的实验或论述,那任何科学都不会进步。

一个典型例子是关于智力的社会特征理论。19世纪末期,文化人类学家发现了不同社会中社会习俗的巨大差异,但同时也得出结论:人类社会生活中存在着大量的共同成分,如家庭制度,宗教模式,社会化过程等。社会学的法国学派,包括莱维布卢尔、涂尔干、马赛尔·茅斯(Marcel Mauss)、亨利·胡伯(Henri Hubert)等人在19和20世纪转折时期提出了一系列命题,主张人类心灵的智力结构是由社会结构决定的。也就是说,知觉和思想是由社会决定的,而不仅是由我们的眼睛、耳朵和其他感受器官的生理解剖结构单独决定的。思想的形态随着社会的不同而不同。这表明玻利尼西亚人的智力不能用欧洲人的标准去衡量,他们的思维方式不仅在量上,而且在质上不同。

几乎同时,测量智力的技术也发展起来,从而标准化为各种不同的智力测验分数,也叫智商。当把这些测验应用于不同的人群时,一个明显的结论是群体间的智力差异很大:美籍意大利人的后裔和黑人的后裔在这些测验上得分比较低,北部的黑人比许多南部的白人得分要高,华裔美国儿童和日裔美国儿童得分更高。因为一般的假设是这些测验测量的是“潜在智力”,因此这些数据有时可以得出这样的结论:某些民族比其他民族优越。

但是,这样的结论显然是冒犯理性原则的,是不符合人类追求的平等原则的,因此需要进行检验。于是,出现了两派主要观点:一派认为智力主要是由社会经验决定的;另一派认为智力是先天决定的。要检验这些互相矛盾的观点,就需要一种研究设计,在更具体的情景下检验其逻辑期望结果。例如,如果一个来自于下等群体的人的智力真的是由生物因素决定的,那么环境的改变不应该改变这个人的智商。另一方面,如果社会环境是决定因素,我们可以预期社会经验的变化会导致某种特定的智商变化。我们还可以作出更进一步的推论,如果把双胞胎在出生早期就分开,分别在两种截然不同的社会环境下抚养,他们的智商应该有显著的差异。或者,如果把一群黑人儿童从比较贫困和落后的南部学校和社区转移到比较富裕的北部学校,那么群体的平均智商应该有变化。后者正是奥托·克林伯格在20世纪30

年代进行的经典研究(Klineberg, 1935)。他追踪了从南部转学到北部的黑人孩子,发现转到北部越早,平均智商增加越多;转学越晚,增加越少。这个研究设计中有一个明确的假设,而这一假设来自于一个更广泛的命题的推论;而且,这个假设得到了有效的验证。

通过这个理论继承和发展的例子不难看出,当一项研究系统地建立在已有理论体系的基础上时,才可能产生有价值的贡献。换言之,要想使研究有价值,假设不一定是非常谨慎和精细的,但应该与已有的理论相关。

综上所述,我们在实施一项课题研究时,首先要收集所有与要研究的题目有关的文献。如果这样做的难度太大,那么其假设很可能过于宽泛了。其次,努力把不同研究的各种结论和命题联系起来,形成比较抽象的结论或预测。第三,设想我们的这项研究是否可以对这些命题进行推论,包括自己的假设。第四,用某些理论模型来检验自己的推论,看是否成立,是否漏掉了某些重要变量或命题。第五,特别要对自己的一系列命题与经典研究的命题进行比较,如果发现这样的抽象比较难度太大,也可以找一些验证经典研究的文献进行比较。无论自己的假设来自哪里,重要的是它在逻辑上必须是建立在已有研究的基础上,或从已有的理论中推论出来的。

参考文献:

- 陈向明, 1996《社会科学中定性研究方法》,《中国社会科学》第6期。
——, 1997《质的研究中的“局内人”与“局外人”》,《社会学研究》第6期。
风笑天, 1999《社会学者的方法意识和方法素养》,《社会学研究》第2期。
——, 2000《社会学方法20年:应用与研究》,《社会学研究》第1期。
风笑天、田凯, 1998《近10年我国社会学实地研究评析》,《社会学研究》第2期。
侯龙龙, 2001《质的研究还是新闻采访》,《社会学研究》第1期。
林彬、王文韬, 2000《对当代中国社会学经验研究及研究方法的分析与反思》,《社会学研究》第6期。
刘少杰, 2000《中国社会调查的理论前提》,《社会学研究》第2期。
张小天, 1997《显著性水平的意义》,《社会学研究》第2期。
Chavetz, J. 1978, *A Primer on the Construction and Testing of Theories in Sociology*, Itasca, IL: Peacock.
Cooper, D. R. & Schindler, P. S. 2001, *Business Research Methods*, McGraw-Hill International Edition.
Creswell, J. W. 1992, *Research Design: Qualitative and Quantitative Approaches*, Thousand Oaks, California: Sage Publications.
Hoover, K. K. 1991, *The Elements of Social Scientific Thinking*, 5th Edition, St. Martin's Press.
Howard G. S. 1985, *Methods in Social Sciences*, Glenview, IL: Salt, Foresman.
Kerlinger, F. N. 1979, *Behavioral Research: A Conceptual Approach*, New York: Holt, Rinhart & Winston.
Kerlinger, F. N. & Lee, H. B. 2000, *Foundations of Behavioral Research* 4.
Klineberg, O. 1935, *Negro Intelligence and Selective Migration*, New York: Columbia University Press.
Mason, E. J. & Bramble, W. J. 1989, *Understanding and Conducting Research: Applications in Education and the Behavioral Sciences*, New York: McGraw-Hill.
Merton, R. K. 1948, *The Bearing of Empirical Research upon the Development of Social Theory.*, *The American Sociological Review*. Vol. 5.
Newman, W. L. 1997, *Social Research Methods*. Allyn & Bacon.
Rudner, Richard 1966, *Philosophy of Social Science*, Englewood Cliffs: Prentice-Hall, Inc.
Salkind, N. J. 2000, *Exploring Research*, 4th edition, New Jersey: Prentice-Hall, Inc.
Sekaran, U. 2000, *Research Methods for Business*, John Wiley & Sons Inc.
Wells & Picou 1981, *American Sociology: Theoretical and Methodological Structure*, Washington D. C.: University Press of America.

作者系中国人民大学劳动人事学院教授,博士生导师
责任编辑:张宛丽