论 操 作 化

张小天

鉴于现有的操作化理论的不足,本文提出了一种关于操作化的看法。首先对"概念"和"变量"加以澄清,指出变量是一种集合概念,是分类法。在日常生活中,与人们使用原始方法获知概念的出现相对应,人们使用原始测量方法去测量变量。然后,从科学研究要求实现可靠的测量出发,指出变量可分为两群:一群是社会研究者无法可靠地测出的变量,另一群是可以可靠地测量的变量。操作化就是为属于前一变量群的某个变量在后一变量群中找到对应的变量(即指标),以便能够用操作测量方法可靠地测量这个变量。操作化是通过定义和命题来建立变量间的联系。不同的联系,有效度不同。操作定义是在某些情形下社会研究者定义出新的变量的方式。

作者:张小天,男,1962年生,浙江大学哲学社会学系讲师。

测量是社会调查必不可少的内容。操作化是测量中的重要环节。不同的学者用"操作化"这一术语指代不同的过程:(1)形成指标、指数、或量表的过程,(2)形成测量的方法及步骤的过程。在本文里,"操作化"是指前一种过程。本文认为:有时候后一过程中包括有前一过程,这时前一过程是这些后一种过程的主要部分。同时总可以认为任何一个前一种过程都是某个后一种过程的一部分。因此可以说指代前一过程的"操作化"是狭义操作化,指代后一过程的"操作化"是广义操作化。本文所分析的是狭义操作化。

操作化的有效度直接影响测量的质量和社会调查的价值。因此为什么和如何进行操作化的社会调查具有重要意义。关于在测量中为什么要进行操作化,现有的理论认为:存在着两个层次,一个是理论的或概念的层次,一个是经验的层次。前者由概念构成,后者由经验现象所构成。概念是不能被人们直接观察到的。但是在概念层次的每一个概念在经验层次都有对应物。因此测量一个概念时就必须架通概念层次和经验层次的桥梁,找出一个经验地测量概念的方法,即在经验层次为这个概念找到一组指标或一个量表。操作化就是这种沟通概念层次和经验层次的过程。①

不难看出,这个理论其实暗含着两个推论:(1)对所有概念的测量都需要进行操作化。(2)与概念对应的指标由于处在经验层次,已不是概念。但是考察社会研究者实际进行的工作,就会发现情况并非如此。在社会研究中,对有些概念的测量并不需要进行操作化,比如对"性别"的测量;而且仔细分析与概念对应的指标就可以看到,每一个指标其实也是一个概念,只不过

① 参见肯尼思·D·贝利著,许真译:《现代社会研究方法》,上海人民出版社,1986 年版,第 71 -73 页;G·罗斯著,林彬、时宪民等译:《当代社会学研究解析》,宁夏人民出版社。第 37—38 页;衰方主编;《社会调查原理与方法》,高等教育出版社,1990 年版,第 121—122 页;G. S. Howard 《Basic Research Methods in the Social Sciences》(Glenview,Scott,Foresman and Company, 1985),第 18,101 页。

是不同的概念而已。比如做为"生活质量"的一个指标,"住房面积"也同样是概念。

更重要的是这个理论还不能扩展到较令人满意地告诉我们寻找指标的途径和不依靠其他指标体系对一个指标体系的有效度进行评估(即表面有效度评估)的方法或困难所在,而只能说指标与概念须有"认识性相关(epistemic correlation)"的关系、①表面有效度的评估"首先需要知道所测量的概念的定义,其次需要知道所搜集的信息是否和该概念密切有关。"②

事实上,有众多的社会学家早已不满意于这个操作化理论。有些社会学家试图重新构造一个操作化理论;^③ 而更多的社会学家在指出需要进行操作化并以事例加以说明的同时,对于为什么要操作化避而不谈。这两种做法都没能使我们获得更多的关于如何进行操作化和如何评估操作化的知识。本文拟从测量能否实现和测量的可靠性出发、就为什么要进行操作化提出一种初步的看法,并在理论上指出进行操作化和表面有效度评估的一些原则。由于"概念"和"变量"这两个概念与操作化有密切的关系,本文的分析将从澄清这两个概念的有关方面入手。

概念

概念就是一种文字的或有声的语词,用来代表一个事物或一类事物。值得一提的是:(1)特称概念只表示一个实体;(2)名词、动词、形容词和副词这些实词都是概念。概念的外延就是概念所代表的事物的全部。许多概念的外延边界是模糊的;有时无法确切地肯定某事物是否属于某概念。概念的内涵就是概念所指事物的共同特点。这些特点是在把这些事物与其他事物相比较时获知的;与不同的其他事物相比较,就会得出不同的特点,因而概念的内涵中包含了众多特点。与外延边界的模糊相对应,有些概念的内涵中特点的组成并不是完全确定的。

日常生活里,人们使用一个概念或创造一个新概念时,并不需要对概念的含义有完整的了解;只要能够大致了解概念外延的边界的可能方位或内涵中的某些特点,就可以利用概念完成通常所进行的思维和交流。但在科学研究中,尤其是在测量中,需要对一部分概念的外延边界有较精确的认识,或对内涵中的特点有较完整的了解。明确概念的含义有两种方法。(1)首先罗列众多的人们使用这个概念的事例,然后考察这些事例中这个概念所指代的事物以及这个概念与其他概念间的关系,由此明确其含义。但是有些概念被人们使用的事例较少,或者人们对其含义有不同的看法,用这个方法不能使我们获知足够明确的含义,这时就需要(2)对这些概念进行工作定义,即研究者为他所进行的研究对概念进行足够明确的定义,约定其外延或内涵,使研究工作可以在此含义下进行,也使其他研究者了解他使用这些概念时所指的含义。

概念所指的事物是可以被人类的感知(视、听、触、嗅和味)、内心体验和思维这些心理活动觉察或推断到的。比如,作为"吸烟"这个概念所指的事物,"丘吉尔在办公室里吸烟"这个现象是可以被他的助手看到的;做为"公正的"这个概念所指的事物,"王某认为这项分配制度是公正的"这个判断是王某这个人在比较之后推断出的。当然觉察推断到一个事物的方式有时不止一种,有时也包含了多种心理活动。比如做为"失望"这个概念所指的一个事物,"李某对他所处的社会感到失望"这个情感可以李某自己体验到,也可以其他人通过看和听他的言行来推断出。在后一方式中就使用了感知和思维两种心理活动。当我们通过感知、体验或思维这些心理

① 肯尼思·D·贝利著,许真译:《现代社会研究方法》,上海人民出版社 1986 年版,第74页。

② 同上,第94页。

③ 如贝利试图以概念层次、经验层次和操作化层次这三个层次来取代概念层次和经验层次这两个层次。但这样做"并不解决全部问题。参见贝利著,许真译:《现代社会研究方法》,同土,第91-92页。

活动觉察和推断到某一概念所指的事物出现于某一情景中时,我们就说我们获知这个概念出现于或存在于这个情景中。

显然,在获知概念的出现时,可能会出现错误:当某一概念确实存在于某一情景时,某人并不是必然能够获知到这个存在的;当某一概念不存在于某一情景时,某个人可能会错误地感知,体验或推断为此概念存在于这个情景中。

在日常生活里,人们所使用的获知概念出现的方法中有些方法包含了思维的成份,而且其中也有些方法对如何思维有明确的程序,但是所有这些方法都没有包含社会研究者有意识地、明确地规定出的关于如何进行思维的成份。这些数量众多的方法都是未经社会研究者有意识地加工的方法,不妨称之为原始方法。

变量

巴比($E \cdot R \cdot Babbie$) 认为"变量是属性的逻辑聚合(logical grouping of attributes)",而属性是用于描述客体的特点或性质。 布东指出,变量"是好些事物按一个或几个特定标准所作区分的结果",是"分类法","分类的标准"。 确实,变量是分类的方法,事实上,是做为变量的值的这组概念规定了分类的方法:当我们说到分类或分类的标准时,我们实际上是指这组概念;当我们说按某一变量进行分类时,我们也指按这组概念来分类。而且,论域也被这组概念所规定:论域就是这组概念的外延之和或具有这组概念所指属性的事物之和。所以变量被这组概念所规定;变量可以看成是由这组概念所构成的分类,是一种由其值组成的集合概念。

社会研究者对于"变量"这个概念有许多误解。(1)有的人把变量的值看成是变量的亚概念。③ 这是混淆了普遍概念与集合概念。其实值是变量这个集合体的分子;值或者用于代表具有值所指属性的事物的普遍概念是用于指代变量论域中事物普遍概念的亚概念。(2)有的人认为变量不是概念,这种看法往往以前述的操作化理论为基础。〕这显然与通常所认为的名词是概念的看法相违背。(3)有的人通过把变量与常量对立起来来理解变量。⑤ 其实常量与变量这一对概念是相对于研究情景而言的:如果在某研究情景中只有某变量的某一值所指的事物或者只有具有某一值所指属性的事物出现,则我们说这个变量在这个情景中是个常量;否则我们说这个变量在这里是个变量。(4)有的人把变量看成是属性。这个看法无法解释值为特称概念或普遍概念的变量。这个看法可能是来自"雇员是有收入的"这样一类语句使用的例子。其实,在这类语句中,本来表示变量的语词被做为表示变量所有的值的普遍概念在使用。

由变量所具有的特征可知,我们可以把任意若干个互斥的(即全异的)概念做为一个概念组,把这个概念组看成一个变量的值,从而构造出一个变量。视这个概念组是特称概念和普遍概念组还是属性概念组。虽然这个变量不能够用于对含有其论域以外事物的事物类进行分类,总是可以用于对属于其论域的任何一个事物类进行分类。

事实上,社会研究者常常以这种方式构造出新的变量。其中最为常见的是从一个属性概念

① E.R. Babbie, «The Practice of Social Research »(Belmont, Wadsworth Publishing Company, Inc., 1979),第38页。

② 雷蒙·布东著,黄建华译:《社会学方法》,上海人民出版社 1987 年版,第 33 页。

③ 参见肯尼思·D·贝利著,许真译:《现代社会究方法》,上海人民出版社 1986 年版,第53页。

④ 参见 G·罗斯著, 林彬、时宪民等译:《当代社会学研究解析》, 宁夏人民出版社, 第 37 - 38 页; 衰方主编:《社会调查原理与方法》, 高等教育出版社 1990 年版, 第 121 页。

⑤ 参见郭志刚编著:《社会调查研究的量化方法》,中国人民大学出版社 1989 年版,第 139 页。

出发,构造出两种变量:(1)构造二分变量。把不具有此属性概念所指属性做为另一个属性概念,视这两个属性概念为一个变量的两个值。往往给予这个变量的名称是"××状况";可有时候只用"××与否"或"××和非(或不)××"这种把值的名称联结起来的方式来称谓。比如变量"有没有对立情绪:有对立情绪/没有对立情绪"。(2)构造多值变量。在此属性概念上加上不同的副词限定语(因为属性概念一般是形容词或动词),形成几个互斥的属性概念,这些概念再加上指代不具有原属性概念所反映属性的属性概念,构成一个变量的几个值。这种变量往往以"××程度"来称谓。比如变量"认同程度:强烈认同/比较认同/没有认同"。另外常见的由一个概念构造出变量的方式还有以不同的形容词限定一个名词普遍概念,以不同的宾语限定一个动词属性概念,使之成为变量"××种类"的值。

瓦托夫斯基指出,测量就是"确认某事物属于一定的种类。" 从测量的含义可以推知,测量其实就是去获知当被测对象出现时,做为变量的值的一组概念中哪一个概念出现。因此测量变量就是试图去获知概念的出现;测量变量的方法就是获知概念存在的方法。这样,由上一节的结论可知:(1)测量是通过心理活动完成的。(2)并非每个人对每一环境下的每个变量都可以做出测量。(3)测量可能会出现错误,即错误地认定对象的取值。(4)在一个对象身上测量一个变量,有时候测量的方法不止一种。(5)有一部分测量方法中包含了被人们明确意识到的思维程序。虽然在日常生活中或在描述日常行为时我们并不使用"测量"和"变量"这些术语,但测量变量这种行为却是日常生活中的一部分。其实,日常生活中的任何一个对自己或对他人的提问都是在试图测量一个变量。因此,我们可以说,与获知概念出现的原始方法相对应,日常生活中人们广泛而且唯一地使用着原始测量方法进行测量。

操作化

科学知识的一个重要特点就是这种知识具有一致性(consensus)。^②是不具有个人性的。有些学者把这个特点称为客观性,有的则称之为可复制或可检验性,或称之为"主观共享性"。^③

做为科学知识的一致性的一个方面,科学研究中所进行的测量必须具有一致性,或者说必须具有可靠性。即通俗表达中称之的测量要正确和准确,在自然科学中被称为系统误差和偶然误差要小,而在社会科学中则被称为要有有效度和可信度。

应该注意到,一名社会研究者在测量变量时,可以利用以下三种途径:(1)只利用研究者自己的知觉或思维完成测量。(2)询问不做为研究人员的被调查者,去获知被调查者使用他自己的知觉、体验和思维对变量加以测量的结果"。④(3)询问被调查者的某些知觉、体验或思维结果,在此之上研究者进行思维推断来完成测量。

社会研究者在使用原始测量方法测量变量时,利用上一节的结论可推知,他能否获得测量的结果及测量的可靠性如何,会因下列因素的不同而不同:被测变量,被测对象,所利用的途径,所使用的方法,所选择的被调查者,被调查者是否了解被测变量的含义或研究者提问的含

① M·W·瓦托夫斯基著,范岱年等译:《科学思想的概念基础—— 科学哲学导论》,求实出版社,1982 年版,第 214 页。

② 参见M·W·瓦托夫斯基著,范侨年等译:《科学思想的概念基础》,同上,第28页-第29页。

③ Babbie 《The Practice of Social Research》,第52页。

④ 在使用这个途径时,如果被调查者测量的结果是"变量 A 取值 a",那么在有些情况下,研究者所获得的测量结果并不是"变量 A 取值 a",而是"他认为变量 A 的值为 a",所实现的是对变量"他所认为 A"的测量。这里的问题涉及到了社会科学的经验基础和社会科学所使用的概念的特点,本文不拟展开讨论。

义,测量所处的环境,以及被调查者是否愿意回答和是否愿意真实回答研究者的询问。对于某个变量和某个对象而言,研究者可以通过选择做各种努力,而其结果可能是这变量被测出,也可能是仍无法测出;在前一种情况下,研究者还可以通过选择上述因素(如果有必要且还有选择余地的话)去试图进行可靠的测量。这个试图也有两种结果:可以实现可靠的测量和无法使测量结果可靠。

因此,研究者使用原始测量方法测量变量时,总有一部分变量无法被测出,并总有一部分变量只能被不可靠地测量出来,换言之,变量总可以分为两群,一群变量无法用原测量方法可靠地测出,另一群变量可以被这种方法可靠地测量出来。

这样,在科学研究中我们就不能使用原始测量方法去直接测量前一群变量。但是,对于某个我们想去加以测量却又属于 前一群的变量,如果我们能够为客观存在在后一群变量中找到一个或一组对应的变量,并能够根据对象在这个或这组对应变量上的值可靠地推断出对象在此欲测变量上的值,那么我们就可以可靠地测量出这个无法用原始测量方法可靠地测出的变量。欲测变量在后一群变量中所对应的这个或这组变量就是欲测量的指标或量表,寻找指标并指出如何由指标的值来推断欲测变量的值就是对欲测变量的操作化。这种测量首先用原始测量方法测量指标,然后由指标的值来推断欲测变量的值,而把哪个或哪些变量做为指标以及如何进行推断都是研究者在操作化过程中明确地界定的,所以这种测量方法已不是原始测量方法,不妨称之为操作测量方法。

显然,我们总能够对前一群变量中的一部分变量实行操作化,而且对其中有些变量的操作化的方式不止一种,但也总有一部分变量无法或难以操作化;而操作化也可能是可靠的,甚至是完全可靠的操作化,也可能是不可靠的操作化。

操作化就是在一个变量与一个或一组其他变量间建立起联系,使我们可以由后者推知前者。操作化中可以使用的变量间的联系有多种多样,但首先可区分为单步联系和多步联系:多步联系使用了中间变量,它是由二个以上的单步联系把变量联系起来。而单步联系又有由定义建立起来的联系和命题联系。中于操作化不仅是指找出对应的指标,而且还包括明确指出如何由指标的值推断出欲测变量的值,所以操作化、或简单地说指标,总是具有完全的可信度的,操作化或指标的可靠性就是指有效度。这样,操作测量的可信度就是指对以原始测量方法测量时的可信度,操作测量的有效度包括原始测量的有效度和操作化的效度。

评估操作化或指标的表面有效度可以依据上面对变量间联系的讨论进行。对于一个已形成的操作化结果,首先要考察欲测变量与指标之间的联系是多步联系还是单步联系;把多步联系分解为单步联系。其次判定单步联系是哪一种类的联系。最后考察不同种类的单步联系,看它们是否与定义相吻合,是否存在着函数关系所要求的条件,是否依据了可以可靠地推断欲测变量的值的因果关系或相关的强度如何。除了下列三种情况外,我们总可以对一个操作化或指标体系做出表面有效评估:(1)某一单步联系不能被划归为上述的任何一个单步联系种类。(2)联系中使用了不能给出完全可靠的推导的因果关系,而且不知道有关的相关关系。(3)联系中包含了相关关系或回归关系的联系,而我们并不知道有关的相关关系或回归关系的数值。

58

① 参见肯尼思·D·贝利著,许真译:《现代社会研究方法》,上海人民出版社,1986年版。

操作定义

许多社会①家把指标或量表看成是概念的操作定义,却又提及指标或量表的有效度有问题。② 严格地讲,这是相矛盾的。

事实上,逻辑学家区分了对概念的三种定义:唯名定义(nominal definition),实在定义(realdefinition)和操作定义(operational definition)。^② 当我们谈到变量的定义时,我们实际是在说变量的值的定义;如果一个变量的值是依某种定义方式定义出的,则可以说这个变量是以这种方式定义出的。社会科学研究中所使用的大部分概念和变量是以唯名定义或实在定义加以定义的,虽然定义往往是通过语言环境而不是通过明确的定义语句完成的。本文上一节讨论操作化时所涉及的变量,实际上是指以这两种方式定义出的变量,因此不能把操作化形成的指标称为操作定义。目前,社会研究者给出的定义,一般都是工作定义。

但是在下列情形下,社会研究者自觉或不自觉地以操作定义方式定义出新的概念和新的变量:(1)原本想对某个无法用原始测量方法可靠地测量的变量 X 进行操作化,以便对 X 实行操作测量。但在找到一组指标和一套推导方法之后,如果这个操作化的有效度很低,或者难以评估它的有效度,那么这组指标和推导方法就不应该被看成是对变量 X 的测量,而应被看成是对一个新的变量 X' 及其取值的操作定义。虽然在社会研究中变量 X 和变量 X' 以及它们的取值还使用相同的语词标识,但已是不同的变量和概念。

(2)对于变量 X 与其维度变量 \mathbf{Y}_1 \mathbf{Y}_2 … \mathbf{Y}_m 之间的联系(见上节单步联系),社会研究者不仅可以由 \mathbf{Y}_1 \mathbf{Y}_2 … \mathbf{Y}_m 的取值来推断 X 的取值,而且还可以给每个 \mathbf{Y}_n 的不同值赋予不同的数值,然后把各个 X_i 的数值取值直接相加或加权相加,认为其和是一个新的变量 X 的取值,从而给出了变量 X' 及其值的操作定义。在社会研究中变量 X' 和变量 X 一般使用相同的语词标识。这两个变量也具有同构性:X'的每个值与 X'中的一组值对应,当 X 取某值时,X'的取值就在相对应的值组中,反之亦然;而且对应于 X_i 的每个 X'的值都小于对应于 X_i +1 的每一个 X'的值。但是这两个变量的值的名称一般不同,而且变量 X'是比变量 X 更加精细的分类法。

(3)有些概念的含义本身很含糊,对这些概念的唯名定义或实在定义没有精确到科学研究所需的程度,或者没有普遍接受的唯名定义或实在定义。有时候与其对这种概念 $X_1, X_2, \cdots X_n$ 的含义进行工作定义,尚不如直接以某些可以用原始测量方法可靠的测出的量为基础对它们进行操作定义。但是当针对 X_1 约定这些变量的一个或一组值并指出这个或这组值出现 X_1 就出现时,可以认为所建立的其实已是对一个新的概念 X_1 的操作定义。由这些 X_1 所组成的变量相应地也就是由操作定义方式定义出的新变量。

操作定义就是把被定义变量或概念与一个或一组可以用原始测量方法可靠地测量的变量 X,或变量的值联系起来,通常相联系的这个或这组变量或变量的值也被称作指标。在操作定 义中,寻找指标并指出如何由指标的值或指标的出现与否来推断被定义变量的值或被定义概

② 对此的有关讨论可参见.C.G. Hempel, 《Fundamentals of Concept Formation in Empirical Science》; 载于《International Encyclopedia of Unified Science》』, No. 7(1952),第6页。这里转引自Babbie, 《The Practice of Social Research》,第124页。Babbie, 《The Practice of Social Research》,第124页。

念的出现与否,是与对被定义变量或概念的定义同时完成的;这一过程并不是为了测量已被定义的变量所进行的操作化。而且对于由操作定义定义出的变量,在测量时已不需要进行操作化。但对它们的测量仍是先以原始测量方法测量其指标,再由指标的值推断出它们的值;所使用的仍然是操作测量方法。不过,测量的可靠性就是对其指标进行测量的可靠性。

可见,指标或量表可能是进行操作定义的结果,这时可称之为操作定义;也可能是进行操作化的结果,这时不能称之为操作定义。前一种情况是对概念或变量的一种界定方式,不存在有效度问题;后一种情况是对变量的一种测量方式,有效度是其重要的方面。当然,操作定义方式定义出的概念或变量总与一些以另两种方式定义出的概念或变量有相似之处,但它们之间的不同并不是以有效度来衡量的。

结论

综上所述,变量可分为四类:(I)可以用原始测量方法可靠地测量的变量;(I)可以进行可靠的操作化,从而可以操作测量方法可靠地测量的变量;(I)由操作定义定义的,可使用操作测量方法可靠地测出的变量;(N)无法可靠地测出的变量。在对变量进行测量时,有些变量(第(I)类)的测量需要进行操作化,有些变量(第(I)类)的测量需要进行操作化,有些变量(第(I)类)的测量需要进行操作化,有些变量(第(I)类)无法操作化从而无法实现测量;操作化是在不同的变量间建立起联系,是在第(I)类和第(I)类变量间架起桥梁,从而使能够可靠地加以测量的变量范围增大;操作定义也是在不同的变量间建立起联系,是以第(I)类变量为基础定义出第(I)类变量,是以可实现可靠的测量目标来定义变量的;操作化中的指标定义中的指标都是第(I)类变量,对第(I)类和第(I)类变量的测量都包含了对第(I)类的测量,因而第(I)类变量是所有测量的基础;测量变量有两种方法,测量第(I)类变量用原始测量方法,测量第(I)类和第(I)类用操作测量方法,后一种方法中包含了对原始测量方法的使用,因此原始测量方法是基本方法;虽然测量方法相同,但测量第(I)类变量的可靠性包括其中原始测量的可靠性和操作化的有效度,测量第(II)类变量的可靠性就是其中原始测量的可靠性;虽然都有指标,但第(I)类变量和第(II)类变量与指标的关系不同;最后,第(II)类变量与其他三类变量的定义方式不同。

需要指出,每一个类型中包括哪些变量是随着下列因素的改变而变化的:被测量的对象,选择的被调查者,以及测量所处的环境。这些因素改变时,第(I)类变量中所包含的变量就会改变,从而使那些变量可以可靠地操作化随之变化,也使能以操作定义定义出的变量的范围发生变化。因此在某种情况下不须操作化的变量在另一情况下可能就要操作化;在某次测量中能够可靠地操作化的变量在另一次测量中就可能无法操作化;在某个社会中能够使用的操作定义在另一个社会中可能就不能使用。

在本文的最后,简略地提一下"概念"和"变量"这两个术语的使用问题。变量是概念的一种,同时变量是由一组概念所组成,尤其是有些变量是由一个概念构造出来的。由于两者的这种关系,许多文献中对这两个术语的使用没有严格的区分。本文认为,在清晰的表述中,应该划分这两个术语的使用范围:对变量则称之为"变量",而不泛称为"概念";对做为变量的值的概念和其他不是变量的概念则称之为"概念",而不误称为"变量"。这样,测量是"对变量进行测量";操作化是为了测量变量而进行的,是"对变量的操作化"。

责任编辑:张宛丽