

行为公共管理研究中的实验方法： 类型与应用

代涛涛 陈志霞*

【摘要】 行为公共管理学作为一门正在孕育和形成中的分支学科，其旨在运用心理学的理论和方法从个体和群体的态度与行为层面来研究公共管理领域的问题，是对传统公共管理学的补充和完善。其最具特色的重要研究方法为实验法，该方法区别于其他方法的一个突出特点在于可进行因果推论。国内外公共管理学者已经在实验法上积累了初步经验，但是如何有效地选择和运用实验方法开展行为公共管理的研究仍有待系统梳理。论文介绍了实验法的基本原理，详细阐述了实验室实验、现场实验、问卷实验、自然实验和准实验等不同的实验类型，并且结合行为公共管理领域的前沿实证研究介绍了如何具体运用实验方法解决行为公共管理领域的研究问题。在此基础上分析了行为公共管理实验的优缺点，并且对如何选择具体的实验方法及其适用性进行了深入分析和讨论。不同的实验方法各有优劣，多样化的实验方法为开展行为公共管理领域的研究提供了更广阔的选择空间。

【关键词】 行为公共管理学 实验法 因果推论 心理学 态度与行为

【中图分类号】 D63

【文献标识码】 A

【文章编号】 1674-2486 (2019) 06-0166-20

* 代涛涛，华中科技大学公共管理学院，博士后；通讯作者：陈志霞（chpsy@126.com），华中科技大学公共管理学院，教授。感谢匿名评审人的意见。

基金项目：国家社科基金重点项目“领导排斥行为的影响因素的毒性三角模型及其作用机制”（17AGL014）、教育部社科基金后期资助项目“行为公共管理视角的城市居民幸福指数及其影响因素”（18JHQ080）、中国博士后科学基金面上资助项目“公共部门领导—下属匹配对主动行为的影响及心理机制”（0106407063）、国家财政专项华中科技大学人文社科发展基金项目“行为公共管理实验室平台项目”（HUST2019）。

一、引言

自20世纪四、五十年代开始，学者们一直试图将心理学中关于“人”的心理与行为的理论与方法知识引入到公共管理的研究中（Simon, 1947; Dahl, 1947; Truman, 1945）。随着行为经济学（Behavioral Economic）和政治心理学（Political Psychology）的巨大成功，这种呼声才开始在公共管理研究领域真正得到重视。2017年诺贝尔经济学奖得主理查德·塞勒（Richard H. Thaler）关于“助推”（Nudge）的思想和实证研究更是给公共管理学者们带来了极大的鼓舞。“助推”思想认为政府需要运用行为经济学的相关知识来优化政府政策设计，在不需要强迫的情况下巧妙地引导人们规避偏见和非理性判断，从而在教育、医疗、投资 and 环境保护等方面做出更好的选择（Thaler, 2008）。在此时代背景和学科背景的共同作用下，“行为公共管理学”（Behavioral Public Administration）作为一门正在孕育和形成中的分支学科也呼之欲出。

斯蒂芬·格里米利克惠森（Stephan Grimmelikhuijsen）等（Grimmelikhuijsen et al., 2017）认为，行为公共管理学是一门旨在运用心理学的理论和方法从个体行为和态度的微观层面研究公共管理领域问题的学科，是对传统公共管理学的补充和完善。张书维和李纾（2018：7）将行为公共管理研究的主要内容归纳为“政府行为－公民体验”双轮模型，主张从“政府公共服务与决策行为过程及其输出”和“公民认知与情绪体验过程及其反馈”两方面加以研究。在研究方法上，由于行为公共管理学聚焦于个体层面，因此尤其适合采用实验法。关于使用实验法的呼声不仅仅是行为公共管理领域独有的现象，而是整个公共管理学科对传统研究方法难以进行因果推论质疑后的一种普遍共识（曹堂哲，2018；李文钊，2018；朱春奎，2018）。国外学者已经运用实验法在公共管理领域开展了大量实证研究（Baekgaard et al., 2015; Battaglio et al., 2019; Blom－Hansen et al., 2015; James et al., 2017），国内学者也在公共政策（耿曙等，2016）、医疗保险（马超等，2015）等方面进行了初步尝试。国内也有学者从学科视角对实验法的整体内容和实验法在公共管理领域的重要性进行了阐述（李晓倩，2018；李晓倩等，2017；马亮，2015），对不同类型的实验方法进行了初步介绍（陈少威，2016；王思琦，2018）。可以说，国内学者在公共管理学的实验研究中积累了初步的经验，但是对于在行为公共管理领域中具体如何选择和运用实验方法开展研究却仍存疑虑。本文将介绍行为公共管理实验研究的一些典型类型，结合新近的实证研究解读如何运用实验法解决公共管理的实际问题，并试图比较

各种实验方法的优缺点，为实验方法的选择提供一些可行的参考。

二、行为公共管理研究中的实验

（一）行为公共管理学的实验法

行为公共管理学注重公共管理的微观基础，关注个体的态度与行为过程，极大地弥补了传统公共管理理论“致广大有余而尽精微不足”的弊端（张书维、李纾，2018：7）。斯蒂芬·格里米利克惠森等（Grimmelikhuijsen et al.，2017）阐明了实验法在“行为公共管理学”的重要地位。实验法的运用并不是为了取代其他研究方法，而是为了更好地补充和完善传统公共管理研究方法。研究方法的选择既取决于学科所聚焦的研究问题和惯用的研究范式，也应当遵从具有普遍性的符合科学规范的思维模式和研究逻辑。实验法是自然科学领域的基本研究范式，在社会科学研究中，为了寻求更加可靠的因果关系推论，实验法逐渐成为一种受重视的选择。目前，从事公共管理研究的学者们大多认可了实验法的优势和潜在价值。在新兴的行为公共管理领域更是如此，保罗·巴塔格里奥（R. Paul Battaglio）等（Battaglio et al.，2019）对行为公共管理领域中与认知偏向有关的84篇文献进行梳理后发现，43%的研究采用了问卷实验法，8%的研究采用了现场实验法，4%的研究采用了实验室实验法，这意味着在该领域中过半数的研究都采用了实验法。

事实上，在斯蒂芬·格里米利克惠森等（Grimmelikhuijsen et al.，2017）阐述“行为公共管理学”概念以前，学者们已经在陆续开展实验研究，并且成为公共管理学运用实验法的最主要的阵地。例如，陈少威（2016）介绍的自然实验法可以用于探讨扶贫政策对地方政府公共支出的影响，利用现场/田野实验的方法考察政府对互联网的规审行为。在李晓倩等（2017）以及岳经纶和张虎平（2018）对公共管理实验研究的梳理中可以发现，在排除自然实验、准实验和一些政策实验后，现有的公共管理领域的实验研究几乎都可以划分在行为公共管理实验范畴下。根据李晓倩（2018）对行为公共管理实验的梳理，运用实验法的研究内容涉及到公共管理的各个层面，既包括政府的主体和对象行为，也包括公民体验的各个方面。行为公共管理实验的本质在于通过操纵不同的刺激，然后观测“人”的心理与行为反应。

实验法之所以适用于行为公共管理研究，其原因主要在于以下方面：

首先，行为公共管理实验更关注个体层面的心理与行为，对个体进行实验

操控的可行性更高，成本也更为低廉。行为公共管理研究中的个体既包括公共部门中的个体，也包括公民个体（李晓倩，2018；张书维、李纾，2018）。相对于社会组织、社会文化和制度等社会背景因素，个体层面的心理与行为更为直观和可控，进行干预操控的空间更广，也更为现实。

其次，行为公共管理实验具有实验法的基本要素，可以进行因果推论。在社会科学研究中，为了获得可信的因果推论，需要解决变量的内生性问题（Wooldridge，2013）。此时，研究者有两种常用路径选择：要么通过统计控制来实现近似随机化，要么通过实验设计的外生干预来避免内生性问题（Blom - Hansen et al.，2015）。由于统计控制的方法需要大量严格的前提假定，并不能真正解决内生性的问题。因此，实验法成为进行因果研究时的最佳选择。

最后，行为公共管理实验更有助于解决公共管理的现实问题。相比心理学实验对基础理论的重视，行为公共管理实验更关注现实问题。公共管理的实际成效不仅取决于政府管理和决策层面，也取决于个体的感受和体验。在政府与公民的互动过程中，两者往往是相互影响的。相比于宏观层面的分析探究，微观的实验设计能够更快速准确地解析出政府与公民互动效果的因果路径和效应量，从而提高公共管理对策的针对性，及时优化和提升公共决策和服务质量。

（二）实验法的基本原理和分类

根据是否为实验设计，研究设计可以分为三种：实验设计（Experimental Design）、准实验设计（Quasi - Experimental Design）和非实验设计（Non - Experimental Design）（Reis & Judd，2000）。实验设计至少包含一个自变量的外生干预，以及被试在不同实验条件下的随机化处理；准实验设计也包含对自变量的操纵，但不要求被试的随机化处理；非实验设计则不包含自变量的操纵，也不需要被试的随机化处理。需要注意的是，对被试进行随机化可以是将被试随机分配到不同的组别，也可以是让被试按随机顺序接受多次实验处理。

在实验或准实验设计中，影响被试在因变量上的反应的来源包括被试因素和实验操纵。在严格控制的实验设计中，由于被试是随机化处理的，因此因变量上的反应差异不能归为被试因素，而只能解释为实验处理的差异（Reis & Judd，2000），这也是实验法的核心思想。例如，在被试间设计中，由于被试是随机分配到不同的组别中接受不同水平的处理，因此不同组别在因变量反应上的差异不能归结于被试本身的构成差异，只能解释为实验处理的差异；在被试内设计中，由于被试接受多次实验处理的顺序是随机的或者被试是被随机分配到不同的平衡组，那么因变量反应上的差异不能归结于被试接受实验处理顺序

或者自然成熟等其他因素，因此只能解释为实验处理的差异。在准实验设计中，由于被试并不是随机分配的，因此因变量的差异既包括实验处理的差异，也包括一定程度上被试本身的差异。因此，在对准实验的研究结果进行因果推论时需要格外谨慎。由于准实验包含外生的干预操纵，具备实验设计最基本的元素，在公共管理的实践中经常出现无法进行随机分组的情形，因此，准实验往往被视为广义的实验法的一种。

由于要解决和研究的实际问题不同，公共管理领域对实验方法的区分与其他学科存在一定的差异。詹斯·布洛姆汉森（Jens Blom - Hansen）等（Blom - Hansen et al. , 2015）通过 5 个问题来区分公共管理中的不同研究设计（见表 1），其中包括目前较为常见的 5 种实验设计形式，其对实验方法的区分得到了大量同行的认可。在随后对行为公共管理研究的梳理过程中，研究者们也大体上沿用了这种分类方法。值得注意的是，依据詹斯·布洛姆汉森等（Blom - Hansen et al. , 2015）的划分标准无法区分问卷实验和现场/田野实验，因此有必要进一步对这两种实验方法的特点进行区分。

表 1 公共管理实验设计的分类与比较

研究方法		实验组是否与基线组比较？	干预是否或者近似外生的？	被试是否或者近似随机化？	研究者是否控制干预？	研究者是否控制环境？
实验法	实验室实验（Lab Experiment）	√	√	√	√	√
	问卷实验/调查实验（Survey Experiment）	√	√	√	√	×
	现场实验/田野实验（Field Experiment）	√	√	√	√	×
	自然实验（Natural Experiment）	√	√	√	×	×
	准实验（Quasi - Experiment）	√	√	×	×	×
非实验法	传统大规模调查研究（Traditional Large - N Observational Study）	√	×	×	×	×
	传统个案研究（Traditional Single Case Study）	×	×	×	×	×

资料来源：Blom - Hansen et al. , 2015。

三、行为公共管理研究中的实验类型及示例

本文沿用了詹斯·布洛姆汉森等（Blom-Hansen et al., 2015）对公共管理实验的划分方式，同样将行为公共管理中的实验类型分为实验室实验、问卷实验、现场/田野实验、自然实验和准实验五类。虽然行为公共管理研究中对实验设计的区分非常细致，但本质内涵仍符合研究设计的基本原理，不同方法的划分仍然归根于对自变量和额外变量的控制方式的差异。下文将逐一介绍这些实验类型及其在行为公共管理领域的具体应用。在研究案例的选择上，本文遵循了以下原则：（1）每一种实验类型都至少包含1个实例。（2）研究案例需为新近10年内发表在具有一定国际影响力的公共管理类期刊，如 *Public Administration Review* 和 *Journal of Public Administration Research and Theory*。（3）研究话题尽可能涵盖行为公共管理不同领域的热点问题，既包括政府主体行为，也包括公民态度与行为。（4）涉及公共管理研究中常见的不同类型的被试群体，如政府或企事业单位工作人员、大学生、网民和其他普通公众。表2列举了本文所选取的实例，下文将详述之。

（一）实验室实验

典型的实验室实验往往包括：由研究者操控的外生性干预、对被试的随机化处理和对干扰因素的控制（Morton & Williams, 2010）。在公共管理情境中，不同部门和个体在文化价值观、工作绩效和其他方面均存在较大的差异。实验室实验正是要控制各种无关因素对因变量的影响，从而考察纯粹的自变量对因变量的影响。实验室除了要控制研究对象本身的差异带来的干扰外，也需要控制现场环境和实施程序给研究对象带来的干扰，例如他人在场可能带来的反应、实验实施时的温度和光照等自然环境的影响。这种干扰还包括研究者对研究对象的潜在影响，如“期望效应”（Rosenthal Effect）。为了尽可能减弱自然环境因素的干扰，实验室的自然环境在整个实验过程中需要保持恒定。为了控制研究者对研究对象的暗示，实验室实验经常采用“双盲实验”（Double Blind），即由不知晓实验目的的主试按照某套固定程序进行施测，并且向被试隐瞒真正的实验目的，以免由于被试的猜测产生“安慰剂效应”（Placebo Effect）。

雷内·卡伦斯（René Karens）等（Karens et al., 2016）采用实验室实验法考察了品牌化政策（Branding Policy）对政策信任的影响，该研究是招募大学生被试在实验室完成的。研究自变量是政策的品牌化，其中实验组的被试所阅读

的材料中阐述的政策信息加入欧盟旗帜符号并且声称这些政策出自欧盟委员会，而对照组的被试所阅读的材料仅仅阐述了政策信息，没有加入欧盟品牌元素。随后的因变量测量了被试对政策的信任程度。

彼得·马克（Peter K. Marks）和拉斯·格里茨（Lasse M. Gerrits）（Marks & Gerrits, 2017）同样采用实验室实验法考察了集体决策过程中前期决策对后续决策偏向的影响，研究被试为大学生，随机两人配对完成实验任务。研究自变量为第一次决策游戏的结果，即赢家和输家，通过采用哈希标记（Hash - Mark, HM）游戏范式来随机产生赢家和输家。研究因变量为第二次分配游戏的结果，采用两性战争（Battle - of - the - Sexes, BOS）游戏范式完成。该研究的最终目的是为了考察在决策过程中是否会出现所谓的“分配正义原则”（Lockean Principle of Distributive Justice）。

由于实验室实验对被试、实验操纵和环境变量的控制都较为严格，需要被试到专门的实验场所来完成实验任务，因此招募被试相对困难。在实际研究过程中，很多研究者都像雷内·卡伦斯等（Karens et al., 2016）以及彼得·马克和拉斯·格里茨（Marks & Gerrits, 2017）采用大学生作为被试群体。当以大学生为被试研究公共管理领域的问题时，需要尽量选择那些较少受到大学生这一身份所影响的议题，尽量突显大学生作为社会人时的心理和行为表现，从而使研究结果更具可推广性。例如，雷内·卡伦斯等（Karens et al., 2016）选择了政策信任这一议题，彼得·马克和拉斯·格里茨（Marks & Gerrits, 2017）选择了一般性决策行为作为结果变量，都尽量降低了学生身份对研究结果的影响。

（二）现场实验/田野实验

现场实验也称田野实验，它与实验室实验存在诸多共同之处，同样包括由研究者操控的外生性干预和对被试的随机化处理（Gerber, 2011）。现场实验与实验室实验唯一的差别在于实验实施的场景不一样，实验室实验是在研究者控制的环境下实施的，而现场实验是在研究问题所发生的真实场景中，例如，生活中求职者投放简历、劳动者进行生产、真实的救助行为、环境保护行为等等。

例如，尼古拉·贝莱（Nicola Bellé）（Bellé, 2015）的现场实验考察了薪酬设计对公共部门人员工作表现的影响，研究被试为意大利当地医疗管理局参加强制性培训项目的300名护士。该研究采用了三因素组间设计，三个自变量分别是报酬类型、报酬的可见性及与受惠者的联结。被试被随机分配到不同组别中。因变量是护士的工作表现，即在三小时的轮班期间正确组装手术套件的数量。

拉吉夫·达罗莉亚（Rajeev Darolia）等（Darolia et al., 2015）也采用现场实验法考察了就读私营大学（For-Profit College）经历对刚进入劳动力市场的年轻人的就业能力的影响。该实验通过编制不同类型的人员简历并且投递到不同的企业，考察雇主对这些简历的回复率。这些简历都至少包括一定程度的高等教育学历，要么是私营大学，要么是公共社区大学，即该研究的自变量操纵。将编制的8914份假简历投送出去后，等待雇主的信息反馈，即为该研究的因变量。

斯蒂芬·格里米利克惠森和阿尔伯特·克里金（Albert Klijn）（Grimmelikhuijsen & Klijn, 2015）同样采用现场实验法考察了司法透明对公共信任的影响，研究被试来自一个网络样本库。随后将被试随机分为实验组和控制组，实验组的被试会观看一个用于促进司法透明的系列剧集，在观看该系列剧集前、后分别对公共信任进行测量。控制组的被试先进行了公共信任的前测，但并没有让他们观看剧集，而是被告知一些关于司法机构希望能更多地了解关于审判和判决信任的信息。控制组也需要完成公共信任的后测。研究因变量为对司法信任的前后测中改变的程度。

实施现场实验的最大困难在于在实验控制过程中如何巧妙地还原心理和行为发生的场景。例如，上述三个研究都采用了实验设计的原理来进行操控，同时也从巧妙的角度切入，在真实的社会场景中考察与公共管理议题相关的人的心理与行为表现。另外，如果现场实验选择真实的行为指标作为因变量，其研究价值将更为凸显。

（三）问卷/调查实验

问卷实验也称调查实验，主要是在实验设计中采用问卷调查的形式来呈现实验刺激材料，由于这种方法简便易行，因而成为行为公共管理研究中使用较多的一种实验方法。问卷实验同样包含实验设计的基本要素，即由研究者操控的外生性干预和对被试的随机化处理（Sniderman, 2011）。问卷实验往往通过呈现不同的问卷版本来进行实验操纵，被试随机得到一种版本的问卷并作答。问卷实验往往采用大规模问卷发放的形式，其突出特点是便于操作，但对实验实施的环境的控制性较差。

例如，西蒙·安徒生（Simon C. Andersen）（Andersen, 2017）采用问卷实验法考察了不同种族来源的校长对以不同形式表达的多元文化政策的态度。研究者通过邮件邀请了美国某州的学校校长自愿参与问卷调查，这些校长来源于不同的种族和学校背景，被随机分配到实验组和控制组。该研究设置了三种不

同版本的问卷：控制组的问卷要求被试想象该州教育部门决定所有的学校都必须开展多元文化课程；实验组 1 的问卷同样要求被试想象该政策，并且告知他们实施该政策的目的是为了重视种族和文化差异；实验组 2 的问卷也要求被试想象该政策，并且告知他们实施该政策是因为有研究证据表明这样做可以减少儿童的偏见。最后，所有的被试都被询问在多大程度上支持该政策。

金承贤（Seung H. Kim）和金三木（Sangmook Kim）（Kim & Kim, 2015）采用问卷实验法考察了不同种族被试在报告社会服务动机时的社会称许性差异。被试为不同种族背景美国人，被随机分为实验组和控制组。该研究主要测试三个敏感性问题，分别是公共服务、自我奉献和工作满意感。研究者给控制组的被试一系列中性的非敏感性问题，与社会称许性没有关联。被试只需要给他们回答问题时同意的条目数，而不需要报告针对每个问题的意见，通过这种匿名的方式最大程度让被试表达自己的真实态度。而实验组的被试会回答相同的系列问题，并且在这些问题中增加了一个敏感性问题。实验组和控制组的中性问题条目数是一样的，因此两个组别的平均差异反映了对敏感性问题的态度差异。

问卷实验虽然对环境的控制较弱，但是简便易行，便于施测，因此在行为公共管理实验中使用率最高。在采用问卷实验法时，有两个问题值得注意：首先，问卷操纵的有效性，即实验操纵是否成功。例如，在西蒙·安徒生（Andersen, 2017）的研究中，让被试想象某类政策，如果被试没有积极参与实验，该操纵可能是无效的，因此往往需要增加其他测量来检验实验操纵的有效性。其次，不同版本的问卷是否包含自变量以外的其他因素。例如，金承贤和金三木（Kim & Kim, 2015）的研究中，不同版本的问卷中的题目条数是不一致的，问卷本身的内容长度也可能对研究结果产生影响，因此在设计问卷时需要进行平衡控制，使得不同版本的问卷在除了操纵的变量以外是等价的。

（四）自然实验

自然实验，顾名思义是指实验的操纵是在自然状态下发生的，并不是由研究者操控的。自然实验也包括外生性的干预和对被试的随机化处理（Dunning, 2012）。与现场实验和问卷实验不同的是，自然实验中的实验操纵并不是由研究者操纵的，而是自然发生的，或者由政治或行政程序引发的。虽然自然实验的外生干预不是由研究者掌控，但实验组和控制组却是随机形成的，或者“近似”随机形成的，因此可以较好地控制内生性问题。

例如，马丁·乔克利（Martin Chalkley）等（Chalkley et al., 2010）运用自

然实验考察了不同的合同制对激励牙医提供公共服务的影响。英国国家卫生署针对牙医服务制定过不同的政策，牙医有三种不同的合同类型：第一种是收取费用合同（A Fee - for - Service Contract）；第二种是准聘用合同（Quasi - Employment）；第三种是基于活动的合同（Activity - Based Contracting）。在 2006 年，英国英格兰和威尔士地区的政策发生了变化，而北爱尔兰地区的政策仍然维持不变，因此在不同地区出现了“不同的外生干预”。研究者搜集了政策前后 96 位牙医的公共服务表现，并比较了不同的合同政策变化前后牙医们的公共服务表现。

由于自然实验法中的实验操纵不是由实验者来控制的，这些操纵往往是公共管理中极其关注的政策变革、公共事件或者自然灾害等变量。例如，中国在深化改革过程中多采用“先试点后推广”的模式，自然形成的实验组和控制组为自然实验的实施提供了大量可能性。以往公共管理的研究多从宏观视角分析这类变量对社会经济发展的影响，或者从理论层面分析对社会群体的影响，较少采用定量方法分析个体心理与行为的变化。随着行为公共管理学的兴起，公共管理中“人”的心理与行为更受关注，自然实验法将成为一种有前景的能解决公共管理理论和实践问题的科学方法。

（五）准实验

准实验仅包含实验法最基本的要素，即对实验组和控制组分别施以不同的外生干预，进而测量这种干预所引发的效应（Cook & Campbell, 1979）。由于准实验无法进行随机分组，它较难区分实验结果到底是由实验操纵引发的，还是由组别本身的内在差异导致的，所以准实验面临的最大挑战就是如何解释组别差异对实验结果的影响。准实验常常采用回归间断设计、间断时间序列设计或者不均等控制组设计，采用前、后测通过统计控制来提高内部效度（Reis & Judd, 2000）。

例如，约瑟夫·巴蒂（Yosef Bhatti）等（Bhatti et al., 2015）的研究运用准实验法考察了丹麦在 2007 年的改革前后政府组织变革对政府雇员健康行为的影响。在这项政府改革中，有些地方政府进行了重组，而另外一些地方政府则依然保持原样。研究者将地方政府间的重组视为实验组，而没有进行重组的地方政府视为控制组。研究者通过政府官方部门渠道获取了此次数据，采集了个体层面的健康数据，也采集了一些政府层面的数据作为控制变量。

理查德·格兰特（Richard W. Grant）等（Grant et al., 2014）运用准实验考察了医务助理在初级保健随访中收集锻炼信息对人们参加运动锻炼的行为的

影响。该研究是利用在四家美国旧金山的地方医疗中心开展的 EVS 项目，该项目改变了医务助理的工作流程，增加了电子医疗记录，以便医务助理能够快速结构性地询问患者的运动锻炼信息。该研究比较了是否参加 EVS 项目对人们参加运动锻炼的影响。该研究的因变量包括患者主观报告的运动锻炼信息，也包括体重、体质指数和糖化血红蛋白等医疗中心采集的客观指标。

上述介绍的几种实验方法是行为公共管理研究中较为常见的几种类型，但不能涵盖所有的实验类型。例如，保拉·肯特莱利（Paola Cantarelli）等（Cantarelli et al., 2018）的研究采用了现场问卷实验法（A Survey – in – Field Experiment）考察了公共部门雇员在进行管理和政策决策时的认知偏差。该研究通过不同的问卷版本来实现实验操纵，但是所操纵的内容仍是雇员在工作中真实发生的情境，所以该研究既有现场实验的特点，同时也具有问卷实验的特点。因此，本文介绍的实验方法是较为典型的实验研究方法，在实际操作过程中应该根据实际情况进行灵活选择和应用。

四、行为公共管理实验方法的选择

虽然实验法的应用尚未成为公共管理研究的主流研究方法，但是顶级公共管理学术期刊都越来越青睐采用了实验法的论文（李晓倩等，2017）。对实验法的呼吁源自实验法可以进行因果推断这一重要的优越性，也源自发展和完善公共管理学科的内在需求。实验法无法取代传统的个案研究和社会调查研究，但它对提升公共管理的学科严谨性和科学性具有重要的价值。以往对实验法的忽视，主要是因为公共管理学科取向强调应用性，重视研究结论的可推广性而一定程度上牺牲了研究的科学严谨性（朱春奎，2018）。同时，在研究内容上以往公共管理研究更重视整体层面的剖析，较少关注个体体验与感受，对整体目标进行操控的难度太大（马亮，2015），将其他社会科学惯用的实验方法迁移到公共管理领域的难度较大。此外，公共管理研究没有运用实验法的传统，相关研究者也缺乏系统的实验方法的训练，在进行研究设计时缺乏相应的技术支持。行为公共管理学的兴起为实验法的大规模运用提供了可能性。根据张书维和李纾（2018）总结的“政府行为 – 公民体验”双轮模型，行为公共管理的研究内容聚焦于政府行为过程和公民体验两大领域，而这两大领域都极其适合采用实验法进行研究，因此实验法也成为行为公共管理的方法特色。

表2 不同的实验方法在行为公共管理中的应用举例

实验方法	代表性论文	研究问题	被试	自变量	刺激信息	因变量	控制变量
实验室实验	Karens et al., 2016	品牌化政策对政策信任的影响	大学生	欧盟政策品牌化：有欧盟品牌元素 VS. 无欧盟品牌元素	政策信息	对政策的信任程度	对政府机构的信任倾向
实验室实验	Marks & Gerrits, 2017	多次决策情境下前期决策对后期决策的影响	大学生	第一次 Hash - Mark 游戏的结果：成功 VS. 失败	Hash - Mark 游戏	分配游戏中的决策倾向	对他人的信任倾向 性别、文化背景
现场/田野实验	Bellé, 2015	报酬类型对公共部门人员工作表现的影响	护士	报酬类型：固定报酬 VS. 固定报酬外加绩效报酬 VS. 固定报酬外加象征性报酬 报酬可见度：可见 VS. 不可见 与受惠者的联结：有 VS. 无	告知报酬方式	工作表现	性别、年龄、工作经验、公共服务动机、自我效能感、尽责性、自我调节、对工作的尊重和公平敏感性等
现场/田野实验	Darolia et al., 2015	就读私营大学对刚进入劳动力市场的年轻人的就业能力的影响	企业雇主	简历中展示的高等教育的类型：私营大学 VS. 公共大学	简历信息	雇主回应的比例 雇主要求面试的比例	简历中的个人基本信息、受教育经历、工作经验、个人技能及其他的信
现场/田野实验	Grimmelikhuijsen & Klijn, 2015	司法透明对公共信任的影响	网络样本库	是否观看了促进司法透明的系列剧集：是 VS. 否	系列剧集	观看剧集前后司法信任的改变程度	
问卷/调查实验	Andersen, 2017	不同种族的校长对多元文化政策的态度	校长	政策表达方式：控制组 VS. 规范价值组 VS. 基于研究证据组	情境想象	对政策的支持程度	校长个人特征和所在学校的特征

(续上表)

实验方法	代表性论文	研究问题	被试	自变量	刺激信息	因变量	控制变量
问卷/ 调查实验	Kim & Kim, 2015	不同种族被试的在 报告社会服务动机 时的社会称许性的 差异	不同种族的美国 人	问卷中是否包含敏 感性问 题: 是 VS. 否	社会服务动 机问卷材料	被试对敏感性问题的 态度差异	
自然实验	Chalkley et al., 2010	不同的合同制刺激 励牙医提供公共服 务的影响	牙医	合同制政策是否变化: 变化 VS. 维持不变	—	政策前后 96 位牙医的 公共服务表现	
准实验	Bhatti et al., 2015	改革前后政府组织 变革对政府雇员健 康行为的影响	政府雇员	地方政府改革: 政府重组 VS. 维持不变	—	长假缺席时间、访问 心理学家的次数、访问 精神病学家的次数 访问一般执业医师的 次数、服用与压力相 关的药物情况、服 用抗抑郁药的情况和 整体药品使用情况	政府层面的变量
准实验	Grant et al., 2014	医务助理在初级保 健随访中系统收集 锻炼信息对人们参 加运动锻炼的行为 的影响	医院患者	是否参加 EVS 项目: 是 VS. 否	—	患者主观报告的运动 锻炼信息 体重 体质指数 糖化血红蛋白	

资料来源: 作者整理。

（一）行为公共管理实验的优缺点

行为公共管理实验是利用心理学的实验方法来研究公共管理领域的问题，因此它既具有心理学实验的一般特征，又兼具公共管理的研究特色。相比于传统的社会调查研究方法和个案研究法，实验法能够确立因果关系，进行因果推论，这是实验法最基本的特点和优势。不同的实验法均对变量进行了一定程度的控制，使得其他研究者可以对研究结果进行重复检验，这一点也不同于以往公共管理中常用的案例研究，而研究结果的可重复性恰是检验研究结论是否可靠的最重要的标准。在特定时间、特定情境、针对特殊对象开展的研究虽然也具备其科学价值，但如果研究结果不能得到重复验证或者推广到其他研究主体，其研究价值就会大打折扣。另外，实验法所需要的样本量往往不大，研究设计也普遍简便易行，耗费的人力和物力相对较少，经济成本较低。研究被试的选择既要基于科学的抽样方法，也要考虑研究问题本身的性质。在进行行为公共管理实验时，如果采用科学合理的研究设计方案，所需要的被试量往往不大，在一定程度上可以降低研究成本。

当然，采用实验法研究公共管理问题也存在一些不足之处。首先，由于需要对被试进行外生的干预，因此其适用范围具有一定的局限性。例如，考察某类政策（如二胎政策）对公众的影响，如果政策出台（外生干预）已经完成，此时就无法再去测量政策出台之前公众的心理和行为反应。在真实的公共管理中，公共管理行为往往会给被试带来一定的影响，这也面临着一些法律和伦理道德上的风险。因此，部分干预无法真实实施，特别是那些可能对被试带来消极影响的实验操纵需要进行格外严格的控制，例如实施惩罚或者启动反社会倾向。另外，公共管理过程中影响人的心理与行为的因素非常多，包括社会背景因素、组织因素和个体因素等各个方面，在实际实验操纵过程中很难对这些变量进行完全的控制，只能根据研究的实际情况进行合理的取舍。

（二）如何选择行为公共管理实验方法

研究设计的选择一方面着眼于所考察的研究问题，另一方面也面临很多客观条件的限制，因此需要灵活选择。通过对行为公共管理中各种研究方法的分析和比较，本文认为可以根据以下六个问题来逐步对不同的研究设计进行区分和选择（见图1）：第一，是否需要比较不同目标？如果不需要比较，则只需要对单个个案进行分析即可。例如，艾瑞丝·柯瑟根（Iris Korthagen）（Korthagen, 2015）采用案例法考察了政策报道过程中的媒体偏见。第二，是否考察外生的

干预操作？如果不考虑进行外生的干预操作，则采用大规模社会调查即可。例如，许多研究者从行为公共管理学的视角出发采用问卷调查法在环境保护（Bell et al., 2009）、电子政务（Piehler et al., 2016）和公民－政府关系（Del Pino et al., 2016）等领域进行了相应研究。当然，如果试图追求更可靠的因果关系，则需要施以外生干预操作，采用实验法。在拟采用实验法时，需要考虑：第三，是否可以对被试进行随机化？如果遇到无法进行随机化的情形，则采用准实验的方法。例如，虽然约瑟夫·巴蒂等（Bhatti et al., 2015）研究的政府变革过程中遇到不能随机分组的情况，采用准实验设计同样有效地解决了研究问题。第四，外生干预是否由研究者实施的？公共管理中的干预多来自于政府政策或自然事件，在研究者自身无法实施干预操纵的情况下，可以采用自然实验法。第五，能否控制环境干扰？由于环境变量非常庞杂，对环境的控制是一个复杂的问题。一般而言，保持研究场所的恒定，尽可能降低其对研究对象的干扰，并且尽可能保证环境刺激对所有研究对象而言是一样的。如果有必要对环境因素进行有效的控制，则采用实验室实验。第六，是否采用问题发生时的真实场景？当无法严格控制环境变量时，仍然需要对研究场景加以考虑。如果能够在

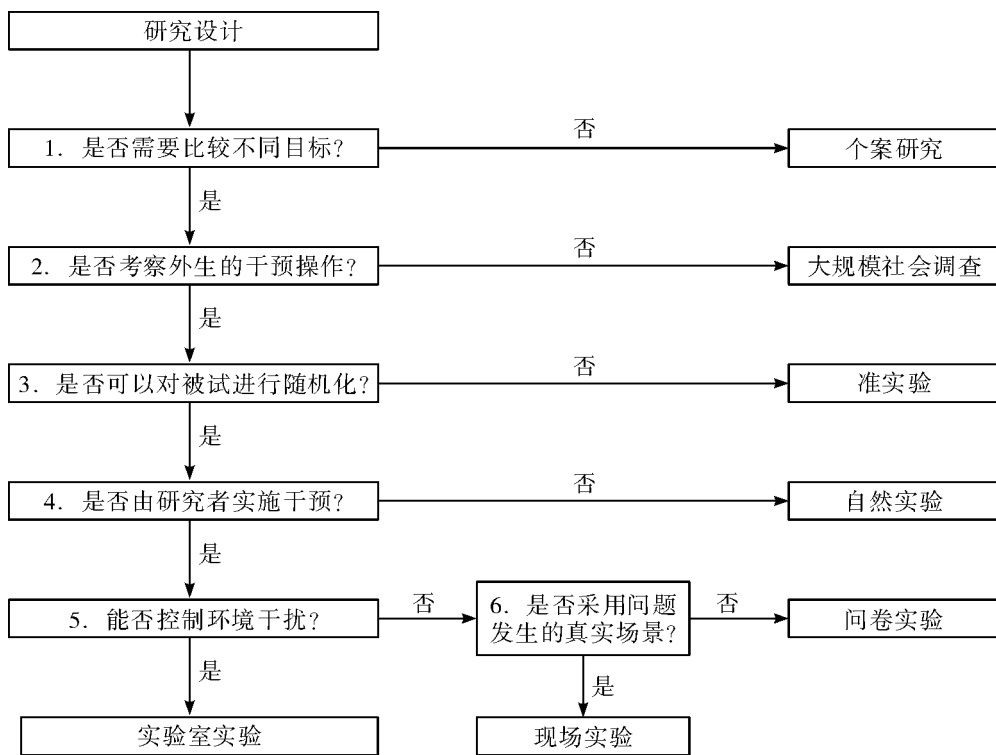


图1 行为公共管理学中的研究设计及实验方法的选择

资料来源：作者自制。

实验操控中还原问题发生的场景，则采用现场实验法，现场实验法在生产行为、助人行为、环境保护和公共部门人力资源管理等领域都有很好的应用。如果不考虑还原问题发生的场景，则采用问卷实验法。采用问卷实验法时，仍需要尽可能降低环境因素对研究结果的影响。

（三）不同实验方法的适用性分析

实验设计的选择既需要一定的客观条件，更需要研究者主动地创造性地营造实验环境。不同的实验方法都有其相对的优势和局限性（见表3）。研究者只有了解各种实验设计的适用性条件，才能更快速和灵活地在各种实验方法中做出选择。

表3 行为公共管理实验方法优缺点比较

实验方法	优点	缺点
实验法	能够确立因果关系；可重复性强；可控制性强；经济成本较低	适用范围有限；伦理道德和法律限制；无关变量的控制难度较大
实验室实验	可操控性强；对无关变量的控制更好；样本量往往不大	生态效度低；被试选择时会受到一定的限制
现场/田野实验	生态效度高	需要考虑的可操控性因素更多；研究成本更高；易产生伦理问题
问卷/调查实验	经济实惠，简便易行；可以大规模施测	操纵干预强度较弱；存在问卷回收率的问题；可能影响生态效度
自然实验	可以解决无法进行实验操控的研究问题	研究实施取决于自然或者政治事件；变量的选择受限于“随机性”因素
准实验	能解决公共管理中的部分现实问题	内部效度不高，只能提供较弱的解释力；需要采用多样化的方法来提高内部效度

资料来源：作者自制。

实验室实验操控性很强，对无关变量的控制更为精准，其需要的被试量也较少。但严格的实验控制也会带来潜在的负面效应（Morton & Williams, 2010）。例如，实验室环境与真实场景的差异使得研究对象在实验室的反应与真实场景中的反应存在一定的差异。这种反应上的差异会削弱实验结果的生态效度，研究结果的适用性和可推广性有所局限。另外，为了能够开展实验室实验，对被

试的挑选会受到一定限制，能够参与实验室实验的往往是特定人群，且大部分是学生被试，容易导致在样本选择上出现偏差。所选择的样本能否代表公共管理过程中所关注的对象的心理和行为倾向是一个需要考量的问题。

现场实验能够真实还原行为发生的场景，因此生态效度极高，这是现场实验相对于其他实验法最突出的优势。但是，实验现场的场景往往复杂多变，很多额外因素难以控制。例如，考察标语内容对是否践踏草坪行为的影响时，对每一个被试而言，实验现场的环境是不一样的（如周围同伴数量、温度等）。研究者为了使得研究场景更为真实，往往也需要耗费更多的成本投入去营造相应的环境。另外，在真实情境中对不同的被试施以不同的操纵时，在公共管理实践中可能会产生一定的伦理风险。

问卷实验的施测过程与以往的问卷调查没有差别，可以大规模施测，相对于其他实验方法而言，这种方法简便易行。在问卷发放过程中，同样存在回收率的问题，当回收率过低时，填写问卷的样本容易出现偏差，影响研究的生态效度。另外一个值得注意的问题是，通过问卷的方法进行实验操纵时，很难确定实验操纵的效果强弱。例如，让被试想象某种政策实施的情境，如果缺乏相应的操纵检查，很难保证被试真正按照要求完成该操纵任务。

与其他实验法相比，自然实验可以解决一些其他实验法无法进行操控的问题，例如警察巡逻对犯罪率的影响（Blom - Hansen et al. , 2015）。由于自然实验的外生干预不是由研究者掌控，因此自然实验的实施很大程度上取决于自然的或者政治的因素，如自然灾害或者政治改革等。这导致研究者在选择研究问题时受限于这些“随机性”的因素，并且在操纵变量的选择上只能由外部因素决定，它不一定能完全满足研究者进行假设检验的需要。在研究者无法掌握外生干预的情况下，可以考虑提前谋划研究设计或者通过多种途径采集外生干预前后的被试数据，例如在政策正式出台的前后分别采集数据。

相比于个案法和传统的大规模社会调查法，准实验在解决内生性问题上有所进步。鉴于行为公共管理中的许多实际问题并不是都能进行随机分组，在没有其他选择的情况下，准实验仍不失为一种较好的解决策略。例如，在考察政府变革过程中，无法对是否进行变革进行随机分组，因此只能尽可能控制其他因素的影响，为政府变革的后续影响提供更多的因果解释力。

根据本文的梳理可以看出，行为公共管理学中的实验设计多种多样各有优劣，多样化的实验方法给予研究者极大的选择空间，研究者可以根据研究问题选择恰当的研究方法。个案法、社会调查法和实验法等都有其独特的价值和丰富的方法学体系，无法相互替代。但是在追寻更可靠的因果关系的研究中，实

验法具有明显优势。随着社会治理中对人的因素日益重视，以及公共管理学对学科严谨性的追求，行为公共管理实验必将大有用武之地。

参考文献

- 曹堂哲(2018). 基于结果链的影响评估及其实验方法. 公共行政评论, 1: 108-123.
- 陈少威(2016). 公共管理研究中的实验设计: 自然实验与田野实验. 国外理论动态, 5: 76-84.
- 耿曙、余莎、孔晏(2016). 实验方法在公共政策研究中的应用——以纳税遵从为例. 公共管理与政策评论, 5(3): 45-52.
- 李文钊(2018). 因果推理中的潜在结果模型: 起源、逻辑与意蕴. 公共行政评论, 1: 124-149.
- 李晓倩(2018). 行为公共管理学实验: 基于 SSCI 期刊(1978—2016)相关文献的分析. 公共行政评论, 1: 37-60.
- 李晓倩、吕孝礼、徐浩、刘求实(2017). 国外公共管理实验研究的进展与启示. 公共管理评论, 2: 16-34.
- 马超、顾海、孙徐辉(2015). 参合更高档次的医疗保险能促进健康吗?——来自城乡医保统筹自然实验的证据. 公共管理学报, 12(2): 106-118.
- 马亮(2015). 公共管理实验研究何以可能——一项方法学回顾. 甘肃行政学院学报, 4: 13-23.
- 王思琦(2018). 公共管理与政策研究中的实地实验: 因果推断与影响评估的视角. 公共行政评论, 1: 87-105.
- 岳经纶、张虎平(2018). 实验方法在公共管理研究中的应用: 基于 PAR 和 JPART 两种期刊(2010—2017)的文献分析. 中国公共政策评论, 15(2): 39-59.
- 张书维、李纾(2018). 行为公共管理学探新: 内容、方法与趋势. 公共行政评论, 1: 7-36.
- 朱春奎(2018). 公共管理研究需要强化因果推理与实地实验. 公共行政评论, 1: 83-86.
- Andersen, S. C. (2017). From Passive to Active Representation—Experimental Evidence on the Role of Normative Values in Shaping White and Minority Bureaucrats' Policy Attitudes. *Journal of Public Administration Research and Theory*, 27(3): 400-414.
- Baekgaard, M., Baethge, C., Blom-Hansen, J., Dunlop, C. A., Esteve, M., Jakobsen, M., Kisida, B., Marvel, J., Moseley, A., Serritzlew, S., Stewart, P., Thomsen, M. K. & Wolf, P. J. (2015). Conducting Experiments in Public Management Research: A Practical Guide. *International Public Management Journal*, 18(2): 323-342.
- Battaglio, R. P., Belardinelli, P., Bellé, N. & Cantarelli, P. (2019). Behavioral Public Administration Ad Fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations. *Public Administration Review*, 79(3): 304-320.

- Bell, J. , Huber, J. & Viscusi, K. W. (2009). Voter – Weighted Environmental Preferences. *Journal of Policy Analysis and Management*, 28(4) : 655 – 671.
- Bellé, N. (2015). Performance – Related Pay and the Crowding Out of Motivation in the Public Sector: A Randomized Field Experiment. *Public Administration Review*, 75(2) : 230 – 241.
- Bhatti, Y. , Gørtz, M. & Pedersen, L. H. (2015). The Causal Effect of Profound Organizational Change When Job Insecurity Is Low—A Quasi – Experiment Analyzing Municipal Mergers. *Journal of Public Administration Research and Theory*, 25(4) : 1185 – 1220.
- Blom – Hansen, J. , Morton, R. & Serritzlew, S. (2015). Experiments in Public Management Research. *International Public Management Journal*, 18(2) : 151 – 170.
- Cantarelli, P. , Belle, N. & Belardinelli, P. (2018). Behavioral Public HR: Experimental Evidence on Cognitive Biases and Debiasing Interventions. *Review of Public Personnel Administration*, DOI: 10.1177/0734371X18778090.
- Chalkley, M. , Tilley, C. , Young, L. , Bonetti, D. & Clarkson, J. (2010). Incentives for Dentists in Public Service: Evidence from a Natural Experiment. *Journal of Public Administration Research and Theory*, 20(Supplement 2) : i207 – i223.
- Cook, T. D. & Campbell, D. T. (1979). *Quasi – Experimentation: Design & Analysis Issues for Field Settings*. Boston: Houghton Mifflin.
- Dahl, R. A. (1947). The Science of Public Administration: Three Problems. *Public Administration Review*, 7(1) : 1 – 11.
- Darolia, R. , Koedel, C. , Martorell, P. , Wilson, K. & Perez – Arce, F. (2015). Do Employers Prefer Workers Who Attend For – Profit Colleges? Evidence from a Field Experiment. *Journal of Policy Analysis and Management*, 34(4) : 881 – 903.
- Del Pino, E. , Calzada, I. & Díaz – Pulido, J. M. (2016). Conceptualizing and Explaining Bureauphobia: Contours, Scope, and Determinants. *Public Administration Review*, 76(5) : 725 – 736.
- Dunning, T. (2012). *Natural Experiments in the Social Sciences: A Design – Based Approach*. Cambridge: Cambridge University Press.
- Gerber, A. S. (2011). Field Experiments in Political Science. In Druckman, J. N. , Green, D. P. , Kuklinski, J. H. & Lupia, A. Eds. *Cambridge Handbook of Experimental Political Science*. Cambridge: Cambridge University Press.
- Grant, R. W. , Schmittdiel, J. A. , Neugebauer, R. S. , Uratsu, C. S. & Sternfeld, B. (2014). Exercise as a Vital Sign: A Quasi – Experimental Analysis of a Health System Intervention to Collect Patient – Reported Exercise Levels. *Journal of General Internal Medicine*, 29(2) : 341 – 348.
- Grimmelikhuijsen, S. & Klijn, A. (2015). The Effects of Judicial Transparency on Public Trust: Evidence from a Field Experiment. *Public Administration*, 93(4) : 995 – 1011.

- Grimmelikhuijsen, S. ,Jilke, S. R. , Olsen, A. L. & Tummers, L. (2017). Behavioral Public Administration: Combining Insights from Public Administration and Psychology. *Public Administration Review*, 77(1): 45 – 56.
- James, O. , Jilke, S. R. & Van Ryzin, G. G. (2017). *Experiments in Public Management Research: Challenges and Contributions*. Cambridge: Cambridge University Press.
- Karens, R. , Eshuis, J. , Klijn, E. & Voets, J. (2016). The Impact of Public Branding: An Experimental Study on the Effects of Branding Policy on Citizen Trust. *Public Administration Review*, 76(3): 486 – 494.
- Kim, S. H. & Kim, S. (2015). Ethnic Differences in Social Desirability Bias: Effects on the Analysis of Public Service Motivation. *Review of Public Personnel Administration*, 37(4): 472 – 491.
- Korthagen, I. (2015). Who Gets on the News? The Relation between Media Biases and Different Actors in News Reporting on Complex Policy Processes. *Public Management Review*, 17(5): 617 – 642.
- Marks, P. K. & Gerrits, L. M. (2017). Association between Decisions: Experiments with Coupled Two – Person Games. *Public Management Review*, 20(7): 960 – 979.
- Morton, R. & K. C. Williams. (2000). *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.
- Piehler, R. , Wirtz, B. W. & Daiser, P. (2016). An Analysis of Continuity Intentions of eGovernment Portal Users. *Public Management Review*, 18(2): 163 – 198.
- Reis, T. H. & Judd, M. C. Eds. (2000). *Handbook of Research Methods in Personality Psychology*. Cambridge: Cambridge University Press.
- Simon, H. A. (1947). A Comment on “The Science of Public Administration.” *Public Administration Review*, 7(3): 200 – 203.
- Sniderman, P. M. (2011). The Logic and Design of the Survey Experiment: An Autobiography of a Methodological Innovation. In Druckman, J. N. ,Green, D. P. , Kuklinski, J. H. & Lupia, A. Eds. *Cambridge Handbook of Experimental Political Science*. Cambridge: Cambridge University Press.
- Thaler, R. H. & Cass, R. S. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Truman, D. B. (1945). Public Opinion Research as a Tool of Public Administration. *Public Administration Review*, 5(1): 62 – 72.

责任编辑:张书维