

公众如何感知公立医院和私立医院的绩效差异？

——基于一项调查实验的比较研究

吴建南 刘 遥*

摘要：

【问题】人们如何看待公立和私立医院的绩效？欧洲语境下的调查实验结果在中国能否复现？

【方法】基于复制实验的原则，在中国情境下复制了 Hvidman 和 Andersen (2016) 的调查实验，以医院的“公私属性”和“是否实施绩效管理”将实验分为四组，采用分层随机抽样获取 141 份有效问卷。

【发现】与原实验一致的发现为：（1）中国公立医院的感知繁文缛节显著高于私立医院；（2）公立医院和私立医院的感知效果、感知仁爱的差异并不显著。不同点在于：（1）原实验发现公立医院的感知效率显著低于私立医院，在本次实验中不显著；（2）我们的实验发现实施绩效管理对中国公立医院的感知效率具有显著正向影响，在原实验中不显著。

【贡献】本文引入复制实验的方法，为复制实验的设计和实施提供了样板。研究表明，理论的可复制性和政策建议的可适用性需要得到更加审慎的讨论。

【关键词】复制实验 公私属性 绩效管理 感知绩效

【中图分类号】D63

【文献标识码】A

【文章编号】1674-2486 (2020) 06-0099-15

【政策之窗】

- ⊙ 相对于“私立”而言，人们会认为“公立”医院有更加繁琐的规章制度。
- ⊙ 实施绩效管理有助于改善中国公立医院的感知效率。
- ⊙ 借鉴不同国家的公共组织相关理论或经验时，应基于中国的本土化情境权变考虑。

* 吴建南，上海交通大学国际与公共事务学院教授，中国城市治理研究院常务副院长；刘遥，上海交通大学国际与公共事务学院博士研究生。

一、引言

“美国人一生中一直被教导：我们的政府是一片浪费的海洋，是无能的沼泽，是权力的大山，是平庸无止境的平原。我们的媒体和政客告诉我们，公共官僚机构规模膨胀，与商业机构相比效率低下，是令人窒息的工作场所，对普通公众漠不关心，是问题而不是解决方案。”（Goodsell, 2003: 3）公共组织效率低下、浪费的想法深植于美国人思维之中（Marvel, 2016）。西方社会对于公共组织绩效的普遍观点认为，公共组织绩效远低于私人组织（Poister & Henry, 1994）。这种认知可能会降低公共部门雇员的士气，影响人们对公共部门的信任，同时也会对类似征税的公权力使用形成阻碍（Marvel, 2015）。

Poister 和 Henry（1994）基于问卷调查比较了公众对消防、邮局、警察等公共服务部门和零售店、私人诊所、快餐店等私人服务机构的评价打分，结果发现受访者对两类服务机构的态度之间不存在系统差异。当然，组织本身所提供服务的不同，一定程度上会降低其可比性。Slyke 和 Roch（2004）研究发现，当未获得令人满意的服务时，公众更可能将原本是由非营利机构提供的服务误认为是由政府提供的，侧面反映出公众倾向于消极地评价政府服务。然而，Hodgkinson 等（2017）利用英国地方政府 2007 年和 2009 年公共休闲服务的二手数据，发现所有权类型对公众满意度没有直接影响。综上，实证研究的发现不一而足，其中可能存在的原因是遗漏变量、样本选择偏误等问题。

通过实验研究设计，可以控制其他因素的干扰，仅比较公共组织和私人组织间的感知绩效差异。Marvel（2015）利用一项调查实验比较了美国公共和私人邮政机构之间的感知绩效差异，结果表明人们对公共邮政服务的感知绩效低于私人邮政服务，存在“反公共部门偏见（anti-public sector bias）”，而且客观的绩效信息也无法完全扭转这一偏见。之后，Marvel（2016）面向美国邮政服务设计了三个调查实验，结果再次证实上述偏见的存在。以上两项研究以随机实验排除了其他诸如规模、职能等因素的影响，实验结果也都显示在不受其他因素干扰的情况下，公共组织的感知绩效显著低于私人组织。但是，这两项实验中感知绩效的测量均采用整体评分。Hvidman 和 Andersen（2016）将感知绩效分为四个维度，分别是感知效果、感知繁文缛节、感知效率和感知仁爱，并以丹麦医院为背景设计了调查实验，结果显示公立医院的感知繁文缛节显著多于私立医院，感知效率显著低于私立医院，而公立医院与私立医院在感知效果和感知仁爱维度上的差异不显著。那么在中国，人们如何看待公立医院和私立医院的绩效？Hvidman 和 Andersen（2016）的实验结果在中国是否成立？

复制 (replication) 是在相似和不同条件下评估理论的经验含义, 从而建立对理论的信心 (或证伪), 或检查研究发现的有效性, 并使之推广成为可能 (Walker et al., 2017)。Jilke 等 (2017) 指出, 复制能够提高外部效度、验证研究发现、探讨边界条件、减少发表性偏倚 (publication bias)。对实证结果的复制, 在科学研究中扮演着基础性作用 (Francis, 2012)。复制也越来越成为社会科学中备受关注的问题 (Tsang & Kwan, 1999; Walker et al., 2017)。

一段时间以来, 国内外很多学科都在关注研究的可复制性和复制研究。国际上, King (1995) 提出了政治学研究的复制标准, 并据此对学者、管理者、评审人、编辑提出了关于建立复制研究正式规则的建议。Tsang 和 Kwan (1999) 以批判实在论的视角探讨了组织科学中的可复制性问题, 并从方法和数据两个维度将复制分为六大类别。Freese (2007) 认为社会学研究应当建立起复制标准, 并回应了一些可能的反对意见。Cesario (2014) 以启动效应为例, 探讨了复制的本质以及它与心理科学的关系^①。在国内, 廖飞等 (2010) 较早实施了一项关于管理者期权思维的复制实验, 结果发现, 对于中国本土的经理人员, 原实验揭示的行为偏差对实物期权主观估价的影响同样存在, 同时也发现个体折现率的结果与原实验存在显著差异。陈云松和吴晓刚 (2012) 提出要在社会科学定量分析中建立透明和开源的学术机制, 必须倡导“复制性研究”。吴小康 (2014)、黄小茹 (2017) 指出了期刊论文中存在部分不可重复验证问题。但国内对复制研究的关注仍亟待提升 (吴建南, 2018)。

实验研究具备很强的复制潜力, 因为研究人员会对实验处理方法和实施背景做出明确规定, 便于获取复制所需的信息 (Walker et al., 2017)。特别是在公共管理学科, 以 Richard Walker 教授为代表的一批学者选择实验研究作为复制的载体, 并在复制实验研究中取得了重要进展^②。本文拟复制 Hvidman 和 Andersen (2016) 的实验, 检验中国情境下公立和私立属性对医院感知绩效不同维度的影响, 以此探索西方理论在中国的适用性。

二、理论构建

(一) 公私属性对感知绩效的影响

Hvidman 和 Andersen (2016) 将感知绩效分为四个维度, 感知效果

^① King (1995)、Tsang 和 Kwan (1999)、Freese (2007) 和 Cesario (2014) 等文献被组织翻译, 并出版在《实证社会科学》(第五卷)。

^② *Public Management Review* 在 2017 年第 9 期上专门刊载了 8 篇复制实验研究。

◆公众如何感知公立医院和私立医院的绩效差异?

(perceived effectiveness) 即人们认为一个组织有能力且有效地完成其核心任务的程度;感知繁文缛节(perceived red tape)即人们认为一个组织具有繁琐的行政规则和程序,且对绩效产生负面影响的程度(Bozeman, 1993);感知效率(perceived efficiency)即人们认为一个组织能够控制其成本的程度;感知仁爱(perceived benevolence)即人们认为一个组织真正关心公众利益的程度。为与原文比较,本文采用这四个维度测量感知绩效。特别说明的是,原实验中设置感知仁爱维度,是为了研究公共组织在道德和规范方面是否具有积极的形象,这种关乎道德的观点经常出现在公共服务动机文献中(Hvidman & Andersen, 2016)。公共服务动机概念源于公私组织的对比研究,即公职人员与私营部门的同行不同,他们有着独特的动机(Perry et al., 2010; 王浦劬、孙响, 2018)。这里的动机即是指“由公共机构和使命产生的具体部署和价值观所激发的利他主义或亲社会动机”。(Perry et al., 2010: 682)

Hvidman 和 Andersen (2016) 认为公众对公共组织持有负面认知,并给出了三点产生对公共组织负面评价的原因:首先,媒体大量报道公共部门的负面新闻是公众对公共部门产生负面评价的一种解释;其次,政客或政治精英们为了获得民众支持而抨击政府的行为会助长公众对公共部门的负面成见;最后,新公共管理运动强调了公共部门效率低下,这种理念会影响公众对公共部门的评价。因此, Hvidman 和 Andersen (2016) 假设,“公共”一词能够让大多数被访者联想到低绩效,从而在评价组织绩效时会认为贴有“公共”标签的组织差于其他相同的组织。根据实验结果, Hvidman 和 Andersen (2016) 发现公立医院的感知效率显著低于私立医院、感知繁文缛节显著高于私立医院,而公私医院的感知效果、感知仁爱差异不显著。

在中国的医院体系中,公立医院无论是在专业设备还是人员配置方面都长期处于主导甚至垄断地位(庞瑞芝等, 2018),私立医院(即民营医院)^①更多的是作为医疗服务体系中的补充角色。数据显示,2017年公立医院入院人数达15594.7万人、住院病人手术人次达4479.7万次,民营医院则是3320.7万人和813.6万次^②,多数人会选择去公立医院就医。近年来出现的“假药假治疗”事件也引发舆论的普遍关注,人们对此深恶痛绝。在中国,公立医院和私立医院的感知绩效应当具有显著差异。

① 现有统计年鉴中,民营医院的英文翻译就是“Private Hospital”,也与本文复制的原实验中私立医院的写法相同,因此本文中将私立医院和民营医院等同。

② 数据来源于2018年《中国统计年鉴》。

（二）实施绩效管理对感知绩效的影响

绩效管理（performance management）是指通过明确的目标、标准、绩效指标、测量和控制系统，对组织和个人绩效的关键方面进行系统、定期和全面地采集、测量、监测和评估（Diefenbach, 2009）。2009年，伴随新一轮医改的开启，公立医院绩效管理改革成为重中之重（吴凌放, 2018），其目的在于提升公立医院行政效率、改善财政情况、提升服务水平。若公共组织实施了绩效管理，可能会让受访者认为其运行方式与私人组织接近，进而减小公共组织感知绩效与私人组织感知绩效间的差异（Hvidman & Andersen, 2016）。原实验将实施绩效管理对感知绩效的影响作为一个开放性问题，且原实验结果也并未发现绩效管理具有显著的直接或调节作用。因此，本次复制实验也将开放性地检验实施绩效管理与否对中国医院的感知绩效是否具有显著影响，或是否在公私属性对医院感知绩效的影响中能起到显著调节作用。

三、实验设计

（一）实验场景

本文在中国情境下复制了 Hvidman 和 Andersen（2016）的实验设计，以医院的公私属性以及是否采用了绩效管理将实验分为 4 组（见表 1）。调查问卷翻译自原实验问卷，包括个人信息、医院简介、感知绩效和检测题项 4 个部分。第 1 组和第 2 组将医院描述为私立医院，实验变量——公私属性赋值为 0；第 3 组和第 4 组将医院描述为公立医院，实验变量——公私属性赋值为 1。第 1 组和第 3 组的医院简介中不包括绩效管理信息，实验变量——绩效管理赋值为 0；第 2 组和第 4 组的医院简介中包括绩效管理信息，实验变量——绩效管理赋值为 1。除医院简介外，四个版本问卷的其他信息完全相同。

表 1 实验分组

		绩效管理	
		无	有
公私属性	私立	组 1	组 2
	公立	组 3	组 4

医院简介：

[某医院] 是一家有 450 名员工的 [公立/私立] 医院。该医院分为三个部分。该医院的

◆ 公众如何感知公立医院和私立医院的绩效差异？

最高管理职能由一名医疗院长、一名首席护士长和一名医院院长所组成的管理小组来执行。[某医院]的核心行政部门负责日常运转以便医院管理满足安全、有效治疗的要求。这项工作需要一个涉及医院所有领域政策和标准的综合系统。该医院的目标是提供最高质量的医疗服务，同时考虑到每位病人的特殊需求。[绩效管理信息：该医院引入了一套绩效管理流程，在这个流程中收集绩效信息，然后将其反馈给管理人员，并利用这些信息来调整目标。这一过程包括医院不断收集治疗质量和病人满意度的数据，然后用这些数据来评估医院的有效性。]该医院曾聘请了一家咨询公司对该组织进行全面评估。咨询公司得出结论说：

- “这家医院采用现代技术设备进行检查和治疗。”
- “这里的工作人员具有高度的专业精神，定期参加世界各地的培训活动。”

(二) 数据收集与均衡性检验

2017年12月，本实验以上海市某高校151名学生为调查对象，采用分层随机抽样的方法进行了问卷发放。首先，我们获取了参与实验的五个班级（编号：A至E）的学生名单；其次，利用Excel自带的随机数命令对五份名单各自进行排序，并分别分成四等份（编号：1至4）；再次，将五份名单按照随机分好的四等份汇总，即A1至E1汇总、A2至E2汇总、A3至E3汇总、A4至E4汇总；最后，根据四组汇总后的学生名单，依次发放对应于实验分组（组1至组4）的问卷。问卷为纸质版，三个班级是在课堂上发放，其余两个班级由调研员一对一发放。

问卷随机发放，避免了对照组和实验组之间出现系统误差。当受访者被要求比较公立医院和私立医院的绩效时，他们可能会按照他们认为应该做出的回应进行回答，也就是存在社会期望偏差问题。实验的方法将能大大降低社会期望偏差出现的风险（Hvidman & Andersen, 2016）。在此次实验过程中，每位受访者只填写一份问卷，且没有提前告知存在实验设计，从而避免受访者知晓要对公立和私立医院进行比较，以此降低出现社会期望偏差的风险。本次实验最终回收了141份有效问卷，样本信息见表2。

表2 样本信息分布

变量	Mean	SD	Min.	Max.	N
部门偏好					141
公共	0.34	0.476	0	1	48
私人	0.45	0.499	0	1	63
无偏好	0.21	0.411	0	1	30
公共服务动机	4.09	0.869	1	5	141
性别（0 = 男，1 = 女）	0.62	0.486	0	1	141
年龄	21.06	2.031	17	26	141

注：公共服务动机有两个测量题项（Park & Rainey, 2008），采用的是五级里克特量表：（1）做有助于他人的工作；（2）做对社会有用的工作（Cronbach's alpha = 0.832）。

为了检验 4 组样本的分布是否均衡,我们利用 ANOVA 和 LSD 多重比较方法考察了问卷中所有协变量的均值差异(见表 3)。结果显示,仅第 3 组的性别在 0.05 水平下显著低于其他三组。其余变量的 F 检验均不显著,表明 4 组实验间具有良好的均衡性,能够较好地避免系统误差。

表 3 均衡性检验

	私立		公立	
	组 1	组 2	组 3	组 4
部门偏好_公共	0.34	0.34	0.34	0.33
公共服务动机	4.32	4.00	3.83	4.20
性别	0.61	0.74	0.37*	0.79
年龄	21.45	20.83	21.11	20.79

注:*表示 $p < 0.05$ 。

在问卷的最后一部分,我们设计了一道题来检验受访者是否真正接收到了实验信息。题目中询问受访者,前文描述的医院是“公立的”“私立的”或“不知道”。141 位受访者中,有 109 位正确填答了检测题项,占到 77.3%。如 Hvidman 和 Andersen (2016) 所述,鉴于医院的公私属性仅有一个微小提示,我们认为本次实验是有效的(原实验的正确回答比例为 76.2%)。

(三) 感知绩效测量与对等性检验

实验采用七级里克特量表测量受访者的感知绩效,包括九个题项(见表 4)。感知绩效的整体 Cronbach's α 系数为 0.639,通过了信度检验。KMO 值为 0.685,且 Bartlett 球形检验的 p 值小于 0.001,适合做因子分析。基于正交旋转的因子分析,九个题项被归到感知绩效的 4 个维度上,共解释了总方差的 74.458%。

当进行跨国调查研究时,需要考虑问卷的测量对等性(measurement equivalence),这一概念指的是测量工具的跨情境有效性(Jilke et al., 2017)。首先,本次复制实验是跨国直接复制(cross-national direct replication),我们采用了和原实验相同的变量与问卷。在问卷翻译过程中,我们遵从尽量少改动的原则,仅将原问卷中的医院名字“Billund”改成了“某医院”。我们对问卷翻译还进行了回译(back translation)检查,即先由一名研究者将英文问卷翻译成中文,再由另一名研究者将中文问卷翻回英文,并将回译的英文问卷交给第三位英文母语者检查。通过严格的翻译和回译检查,保证了此次复制实验与原问卷

共用相同的测量题项。

其次，我们根据 Jilke 等（2017）提出的测量对等性检验方法进行了构型对等（configural equivalence）检验。复制实验中的测量对等性问题取决于复制的类型和目的（Jilke et al., 2017）。此次复制实验是直接复制（采用相同的实验设计和量表），目的是比较回归系数的显著与否，不比较回归系数及截距的大小。因此，本次复制实验在潜变量度量方面需要达到构型对等，即潜变量具有相同结构的因子构造，这意味着潜变量在不同的情境下具有相同的含义。根据表 4 的因子分析结果，本次实验的潜变量——感知绩效的因子构造和原实验完全相同，达到了构型对等的标准。

表 4 感知绩效的测量与因子分析结果

题项	效果	繁文缛节	效率	仁爱
该医院是有疗效的	0.809	0.038	-0.016	0.072
该医院能有效完成它的核心任务	0.849	-0.049	0.146	0.174
该医院可以有效提供高质量服务	0.851	-0.037	0.058	0.172
该医院有非常繁琐的行政规章和程序	-0.061	0.786	0.132	0.191
高度繁琐的行政程序会抑制该医院的效果	0.041	0.850	-0.008	-0.126
该医院能有效降低其成本	0.060	0.020	0.866	0.023
该医院重点关注预算的合理性	0.055	0.099	0.876	-0.051
该医院会真正地关心病人的幸福	0.471	0.145	-0.138	0.653
该医院是为了病人的利益而作为	0.130	-0.029	0.045	0.919

注：数据分析中对繁文缛节的赋值进行了转换，以高分表示低水平的繁文缛节（即高绩效）。

四、实验结果

本文利用 SPSS 20.0 软件的 OLS 回归分析了公私属性和绩效管理对受访者感知绩效的影响及绩效管理的调节作用，结果见表 5。回归模型中，公私属性的对照组是“私立医院”，绩效管理的对照组是“无绩效管理”。数据分析时，先将公私属性和绩效管理两个实验变量纳入回归模型，再将二者的乘积项纳入模型，并将感知绩效的效果、繁文缛节、效率和仁爱维度依次作为模型的因变量。根据模型显著性可知，当以效果、仁爱作为因变量时，模型 1、模型 2、模型 7 和模型 8 均不显著；以繁文缛节和效率分别作为因变量时，模型 3、模型 4、模型 5 和模型 6 具有显著性。

回归系数的显著性显示，公私属性显著影响感知繁文缛节，是否实施绩效管理显著影响感知效率，其余均不显著，具体而言：首先，公私属性对繁文缛

节维度具有显著负向影响,即受访者对公立医院的感知繁文缛节显著高于私立医院;其次,实施绩效管理对效率维度具有显著正向影响,即实行绩效管理制度能够显著提高受访者对医院的感知效率;最后,是否实施绩效管理的调节作用不显著,即没有证据表明绩效管理制度的实施能够减弱公共属性对受访者感知绩效的影响。

表 5 OLS 回归分析结果

变量	感知效果		感知繁文缛节		感知效率		感知仁爱	
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6	模型 7	模型 8
公私属性	-0.284 (0.168)	-0.056 (0.232)	-0.444** (0.165)	-0.308 (0.230)	0.108 (0.164)	0.121 (0.228)	-0.160 (0.169)	-0.132 (0.236)
绩效管理	0.131 (0.168)	0.359 (0.232)	0.060 (0.165)	0.196 (0.230)	0.516** (0.164)	0.529* (0.228)	0.117 (0.169)	0.145 (0.236)
公私属性× 绩效管理		-0.473 (0.334)		-0.282 (0.331)		-0.027 (0.329)		-0.058 (0.339)
常数项	0.074 (0.141)	-0.036 (0.161)	0.185 (0.140)	-0.120 (0.159)	-0.301* (0.138)	-0.308 (0.158)	0.020 (0.143)	0.007 (0.163)
N	141	141	141	141	141	141	141	141
F	1.729	1.829	3.654*	2.673*	5.207**	3.449*	0.685	0.463
R ²	0.024	0.039	0.050	0.055	0.070	0.070	0.010	0.010
调整 R ²	0.010	0.017	0.037	0.035	0.057	0.050	-0.005	-0.012

注:**表示 $p < 0.01$, *表示 $p < 0.05$, 括号内为标准误差, 高分表示低水平的繁文缛节(高绩效)。

从均衡性检验结果看(见表3),性别在四组样本之间存在显著差异。因此,我们借鉴 George 等(2017)的做法,将性别作为控制变量纳入到每个回归模型中进行稳健性检验(见表6)。结果显示,除模型4至模型12的模型显著性变为不显著之外,其余所有回归模型的显著性均未改变,所有回归系数的显著性及影响方向也都未改变。所以,我们认为本次实验结果具有稳健性。

表 6 稳健性检验

变量	感知效果		感知繁文缛节		感知效率		感知仁爱	
	模型 9	模型 10	模型 11	模型 12	模型 13	模型 14	模型 15	模型 16
公私属性	-0.292 (0.169)	-0.065 (0.237)	-0.426* (0.166)	-0.259 (0.234)	0.131 (0.164)	0.175 (0.232)	-0.166 (0.170)	-0.147 (0.240)
绩效管理	0.152 (0.175)	0.364 (0.234)	0.012 (0.172)	0.168 (0.231)	0.456** (0.170)	0.498* (0.229)	0.136 (0.177)	0.154 (0.238)

◆ 公众如何感知公立医院和私立医院的绩效差异？

(续上表)

变量	感知效果		感知繁文缛节		感知效率		感知仁爱	
	模型 9	模型 10	模型 11	模型 12	模型 13	模型 14	模型 15	模型 16
公私属性 ×		-0.462		-0.339		-0.091		-0.040
绩效管理		(0.339)		(0.335)		(0.332)		(0.344)
性别	-0.077	-0.040	0.178	0.205	0.223	0.231	-0.067	-0.064
	(0.182)	(0.183)	(0.179)	(0.181)	(0.177)	(0.179)	(0.183)	(0.186)
常数项	0.116	-0.012	0.089	-0.005	-0.422 *	-0.447	0.057	0.046
	(0.173)	(0.196)	(0.170)	(0.193)	(0.168)	(0.191)	(0.174)	(0.199)
N	141	141	141	141	141	141	141	141
F	1.206	1.374	2.765 *	2.331	4.019 **	3.013 *	0.499	0.375
R ²	0.026	0.039	0.057	0.064	0.081	0.081	0.011	0.011
调整 R ²	0.004	0.011	0.036	0.037	0.061	0.054	-0.011	-0.018

注：**表示 $p < 0.01$ ，*表示 $p < 0.05$ ，括号内为标准误差，高分表示低水平的繁文缛节（高绩效）。

数据分析结果显示，受访者认为公立医院和私立医院在感知繁文缛节上有显著差异（如图 1 所示）。并且据图 1 显示，感知繁文缛节方面的公立医院和私立医院差异在感知绩效四个维度中最大。受访者对公立医院和私立医院的感知效果、感知效率和感知仁爱差异不显著。但相较而言，对公立医院的感知效果和感知仁爱低于私立医院，对公立医院的感知效率高于私立医院。其中，感知效果、感知仁爱 and 感知效率三个维度的估计差异依次减小。

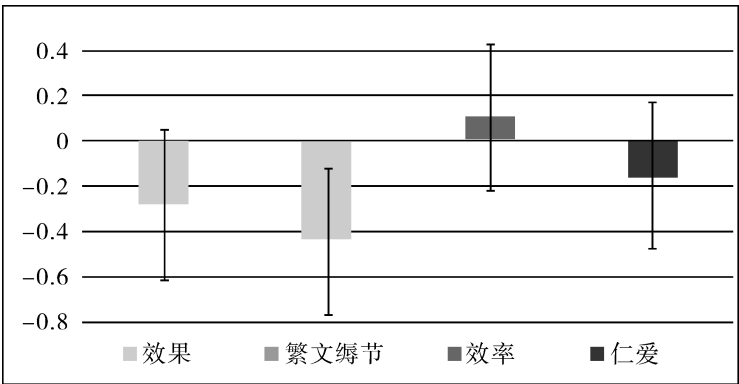


图 1 公立医院和私立医院感知绩效的估计差异

注：对照组是私立医院。条形图表示公立医院与私立医院在感知绩效（因子得分）上的估计差异，误差线表示回归系数的置信区间（95%），高分表示低水平的繁文缛节（高绩效）。

为了进一步验证绩效管理对公立医院感知绩效各个维度的作用，我们将样本按公私属性取出公立医院样本，再利用回归分析计算实施绩效管理对医院感知绩效的影响（见表7）。结果显示，实施绩效管理对公立医院的感知效率具有显著正向影响，对其他维度感知绩效影响不显著，即实施绩效管理能显著提升受访者对公立医院的感知效率。除回归分析之外，我们还对公立医院样本进行了四组 T 检验，结果与回归分析结果一致。

表7 是否实施绩效管理对公立医院感知绩效的影响

变量	感知效果	感知繁文缛节	感知效率	感知仁爱
	模型 17	模型 18	模型 19	模型 20
绩效管理	-0.114 (0.250)	0.086 (0.250)	0.502 (0.231) *	0.088 (0.258)
常数项	-0.91 (0.174)	0.188 (0.174)	-0.186 (0.161)	-0.125 (0.180)
N	68	68	68	68
F	0.209	0.117	4.736 *	0.115
R ²	0.003	0.002	0.067	0.002
调整 R ²	-0.012	-0.013	0.053	-0.013

注：**表示 $p < 0.01$ ，* 表示 $p < 0.05$ ，括号内为标准误差，高分表示低水平的繁文缛节（高绩效）。

五、讨论与启示

本次复制实验以医院为场景，检验了公私属性和绩效管理对医院感知绩效的影响。对比 Hvidman 和 Andersen（2016）的实验结果，两次实验既有相同的发现，也有不同的发现。两次实验中，公立医院的感知繁文缛节均显著多于私立医院，两次实验均表明公私属性对医院感知效果和感知仁爱的影响不显著；是否实施绩效管理对感知效果、感知繁文缛节和感知仁爱的影响以及是否实施绩效管理的调节作用在两次试验中都未表现出显著性。

复制实验证实了原实验关于公私医院感知繁文缛节差异的结论，即公立医院具有显著高于私立医院的繁文缛节评价。由此可知，无论是中国受访者还是西方受访者均认为公立医院有更多的繁文缛节。换言之，公众对于公立医院与私立医院繁文缛节的认知差异已经超越不同国家、不同城市。故而，公立医院管理者应减少不必要的规则和程序。同时，对就医事项和流程的不了解可能会

◆ 公众如何感知公立医院和私立医院的绩效差异?

导致对公立医院复杂繁琐的认知,公立医院管理者应让公众更加清晰地了解医院的运行模式。

两次实验均表明公立医院和私立医院在感知效果和感知仁爱上不具备显著差异。可能的解释是,公立医院往往人满为患,医生每天接诊人数往往处于超负荷状态,病患数量和医院职员数量的结构失衡导致公立医院病人难以感知到医生的治疗和关照;部分媒体对“医患矛盾”“医闹事件”的持续报道也可能使人们形成对公立医院医生的不佳印象,以至于公私差别未能显现。作为公立医院管理者,可强调疗效和仁爱的多重目标追求,进一步展现出公立医院“公益性”的优势。

与原实验比较,本次实验结果存在两点不同。由于本次实验是跨国直接复制,关于核心变量即感知绩效的测量也经过了对等性检验,因此我们判断不同的实验结果主要是情境差异所导致的。以下将结合中国情境,对两处不同进行解释。

第一,原实验中公立医院的感知效率显著差于私立医院,本次实验中结果不显著且影响方向相反。丹麦公立医院的资金绝大部分来自于国家和市政府财政(劳动保障部丹麦考察团,2007)。丹麦公立医院的所有治疗和药品均免费向患者提供(The Ministry of Health, 2017: 43)。中国的公立医院是指国有或集体资产举办的非营利性医疗机构,为社会公众利益服务,不以营利为目的,其收入用于弥补医疗服务成本,实际运营中的收支结余只能用于自身的发展^①。改革开放以来,中国公立医院的财政投入一直不足,政府只能允许“以药补医”机制的存在,导致了公立医院出现逐利趋势(岳经纶、王春晓,2016)。在2009年开始“新医改”后,国家陆续出台的政策文件已经在强调取消“以药补医”机制,增加对公立医院的财政投入,确保公立医院回归公益性(陈竺、张茅,2012)。即使“新医改”后对公立医院的财政投入逐年稳步增长,2016年财政补助也才占到公立医院总收入的9%,2018年占比9.5%^②。很大程度上,中国的公立医院仍然需要以药品收入和医疗服务收入来支撑医院运转,员工绩效工资、奖金也与之挂钩。这些情况都会不同程度地影响受访者,即中国公立医院同样会努力控制成本,其感知效率并不低于私立医院。

第二,原实验中是否实施绩效管理对医院感知绩效影响均不显著,本次实验发现实施绩效管理对医院感知效率具有显著正向影响。自改革开放以来,从

① 参考2001年《关于城镇医疗机构分类管理的实施意见》和2018年《中国卫生健康统计年鉴》。

② 数据来源于2019年《中国卫生健康统计年鉴》。

目标责任考核制到绩效评估、绩效管理,从政府到学校、医院都在强调绩效管理,人们已经形成了绩效管理会对绩效产生影响的普遍认知。而2009年发布的《中共中央国务院关于深化医药卫生体制改革的意见》明确提出“公共卫生机构收支全部纳入预算管理”,“实行核定任务、核定收支、绩效考核补助的财务管理方法”等。这将人们关于公立医院绩效管理的认知聚焦在了财务管理方面。本次实验也是以“成本控制和预算合理性的评价”作为感知效率的测量。因此,实施绩效管理对受访者的公立医院感知效率带来了明显提升。

总体而言,西方国家研究中普遍得到验证的“反公共部门偏见”(Marvel, 2015),即无关乎客观绩效水平而认为公共部门绩效比私人部门更低,在中国公立医院和私立医院场景实验中部分存在。效果、繁文缛节、效率和仁爱四个绩效维度之中,调查实验的证据表明人们对中国公立医院的感知繁文缛节显著高于私立医院,这一点与原丹麦实验一致,也符合已有西方研究中的理论观点。不同之处在于中国公立医院和私立医院的感知效果、感知效率和感知仁爱不存显著差异,这三个维度未显现出“反公共部门偏见”。特别是原实验发现丹麦公立医院的感知效率显著差于私立医院,而在本次实验中结果不显著且影响方向相反。对此,上文结合中国公立医院在资金来源、运营收支等方面的特点做了进一步解释。通过与原实验及西方理论观点的比较,我们发现在不同国情背景的约束下,有的理论可以跨越国界,也有的理论存在很强的情境依赖性。

毫无疑问,复制概念的引入对实证研究提出了更高要求,意味着研究者必须更加注重研究设计,从研究设计之初就需要考虑研究的可复制性、所提出理论的适用性;而研究完成之后,更要反复考察研究的可复制性。对于公共管理的研究者而言,若期望特定的对策建议为决策者采纳,更是必须审慎地从探究现实世界规律出发,考虑特定现实情境下的理论构建,而这将是公共管理科学化进程的必经之路(吴建南, 2018; Walker et al., 2019)。出于复制的考量,本次实验尽可能严格采纳了原实验的变量,以至于中国情境未能充分展现。未来实验中可以增加更多贴合中国实际的变量,也可考虑将实验场景扩展至中小学教育、社会福利、城市保洁、交通服务等领域。如果要继续探究中国公立医院和私立医院感知绩效的差异,未来可直接在公立医院和私立医院现场调查。

参考文献

- 陈云松、吴晓刚(2012). 走向开源的社会学——定量分析中的复制性研究. 社会, 32(3): 1-23.
- 陈竺、张茅(2012). 取消“以药补医”机制, 深化公立医院改革. 求是, 9: 33-35.
- 黄小茹(2017). 科研成果不可验证性问题——发现机制的失效及可能的对策. 科学学研究,

◆ 公众如何感知公立医院和私立医院的绩效差异?

35(7): 961-966.

劳动保障部丹麦考察团(2007). 丹麦医疗保障体制探析. 中国劳动保障, 12: 10-11.

廖飞、茅宁、刘文娟(2010). 不确定下的管理者期权思维: 一个基于复制实验的探讨. 现代管理科学, 1: 90-93.

庞瑞芝、高贤泽、邓忠奇(2018). 公立医院“三重垄断”与医疗行业效率——基于我国省际面板数据的研究. 当代经济科学, 40(1): 1-12.

王浦劬、孙响(2018). 公务员公共服务动机与社会联系偏好的关联性研究——基于我国四地级市的实证调查. 中共中央党校学报, 22(5): 90-100.

吴建南(2018). 社会科学研究: 实验、复制与中国学者的使命. 载钟杨主编《实证社会科学》(第五卷). 上海: 上海交通大学出版社.

吴凌放(2018). 改革中的我国医院绩效管理: 现况、困境与出路. 同济大学学报(社会科学版), 29(2): 123-129.

吴小康(2014). 可复制性与经验研究的透明化. 经济学动态, 10: 121-129.

岳经纶、王春晓(2016). 堵还是疏: 公立医院逐利机制之破除——基于广东省县级公立医院实施药品零差率效果分析. 武汉大学学报(哲学社会科学版), 2: 29-38.

Bozeman, B. (1993). A Theory of Government “Red Tape”. *Journal of Public Administration Research and Theory*, 3(3): 273-304.

Cesario, J. (2014). Priming, Replication, and the Hardest Science. *Perspectives on Psychological Science*, 9(1): 40-48.

Diefenbach, T. (2009). New Public Management in Public Sector Organizations: The Dark Sides of Managerialistic ‘Enlightenment’. *Public Administration*, 87(4): 892-909.

Francis, G. (2012). Publication Bias and the Failure of Replication in Experimental Psychology. *Psychonomic Bulletin & Review*, 19(6): 975-991.

Freese, J. (2007). Replication Standards for Quantitative Social Science: Why Not Sociology? *Sociological Methods & Research*, 36(2): 153-172.

George, B., Desmidt, S., Nielsen, P. A. & Baekgaard, M. (2017). Rational Planning and Politicians’ Preferences for Spending and Reform: Replication and Extension of a Survey Experiment. *Public Management Review*, 19(9): 1251-1271.

Goodsell, C. T. (2003). *The Case for Bureaucracy: A Public Administration Polemic*. Washington, DC: CQ Press.

Hodgkinson, I. R., Hughes, P., Hughes, M. & Glennon, R. (2017). Does Ownership Matter for Service Delivery Value? An Examination of Citizens’ Service Satisfaction. *Public Management Review*, 19(8): 1206-1220.

Hvidman, U. & Andersen, S. C. (2016). Perceptions of Public and Private Performance: Evidence from a Survey Experiment. *Public Administration Review*, 76(1): 111-120.

Jilke, S., Petrovsky, N., Meuleman, B. & James, O. (2017). Measurement Equivalence in Replications of Experiments: When and Why It Matters and Guidance on How to Determine Equivalence. *Public Management Review*, 19(9): 1293-1310.

King, G. (1995). Replication, Replication. *Political Science and Politics*, 28(3): 541-559.

- Marvel, J. D. (2015). Public Opinion and Public Sector Performance: Are Individuals' Beliefs about Performance Evidence-Based or the Product of Anti-Public Sector Bias? *International Public Management Journal*, 18(2): 209 – 227.
- Marvel, J. D. (2016). Unconscious Bias in Citizens' Evaluations of Public Sector Performance. *Journal of Public Administration Research and Theory*, 26(1): 143 – 158.
- Park, S. M. & Rainey, H. G. (2008). Leadership and Public Service Motivation in U. S. Federal Agencies. *International Public Management Journal*, 11(1): 109 – 142.
- Perry, J. L. , Hondeghem, A. & Wise, L. R. (2010). Revisiting the Motivational Bases of Public Service: Twenty Years of Research and an Agenda for the Future. *Public Administration Review*, 70(5): 681 – 690.
- Poister, T. H. & Henry, G. T. (1994). Citizen Ratings of Public and Private Service Quality: A Comparative Perspective. *Public Administration Review*, 54(2): 155 – 160.
- Slyke, D. M. V. & Roch, C. H. (2004). What Do They Know, and Whom Do They Hold Accountable? Citizens in the Government-Nonprofit Contracting Relationship. *Journal of Public Administration Research and Theory*, 14(2): 191 – 209.
- The Ministry of Health (2017). *Healthcare in Denmark*. <http://www.sum.dk/English.aspx>.
- Tsang, E. W. K. & Kwan, K. M. (1999). Replication and Theory Development in Organizational Science: A Critical Realist Perspective. *Academy of Management Review*, 24(4): 759 – 780.
- Walker, R. M. , James, O. & Brewer, G. A. (2017). Replication, Experiments and Knowledge in Public Management Research. *Public Management Review*, 3:1 – 14.
- Walker, R. M. , Brewer, G. A. , Lee, M. J. & Witteloostuijn, A. V. (2019). Best Practice Recommendations for Replicating Experiments in Public Administration. *Journal of Public Administration Research and Theory*, 29(4): 609 – 626.

责任编辑:张书维