

DOI: 10.16538/j.cnki.fem.20200204.301

消费者行为研究中的田野实验： 概念、操作介绍与开展建议

柳武妹, 黄河清, 叶富荣

(兰州大学管理学院, 甘肃 兰州 730000)

摘要: 田野实验是一种与实验室实验相对应的实验研究方法。该类实验尽管相比实验室实验内部效度较低(噪音控制弱), 但因为情境真实和外部效度高而近几年来受到学术界的持续关注。自2017年JMR、JCR和JM相继出版田野实验的专刊以来(Gneezy, 2017; Meyer, 2017; Morales等, 2017), 经笔者统计, 2018年和2019年发表在JCR、JMR、JM上接近1/4的消费者行为研究都宣称自己使用了田野实验。然而, 仔细阅读这些田野实验的文章后笔者发现, 有些关于田野实验的问题并没有被彻底回答清楚, 包括究竟什么是田野实验? 田野实验有哪些判别标准? 由于对田野实验的内涵界定存在不一致, 一些文章宣称自己使用了田野实验, 但实际上却不是严格意义上的田野实验。更重要的, 我国营销学术界对田野实验本身的关注度和重视度还非常不足, 发表的田野实验文章数量还非常有限。本文在JCR、JMR、JM、JCP这四份消费者行为顶级和权威期刊近5年(2014年8月—2019年11月)刊发的108篇使用了至少一项田野实验的文献的基础上, 将重点厘清田野实验的概念, 并总结归纳前人在田野实验中为了增强情境真实性和外部效度做了什么, 最后提出田野实验的开展建议。笔者希望, 这篇文章能为国内营销学者和对田野实验感兴趣的其他领域学者日后开展田野实验提供操作层面的参考。

关键词: 田野实验; 消费者行为研究; 概念; 操作介绍; 开展建议

中图分类号: F270 **文献标识码:** A **文章编号:** 1001-4950(2020)03-0035-22

一、引言

消费者行为是市场营销学科的重要研究领域, 这一领域使用的研究方法多数是实验法, 尤其是实验室实验。自1974年第一本消费者行为杂志《Journal of Consumer Research》(JCR)诞生至今, 发表在JCR、《Journal of Marketing Research》(JMR)、《Journal of Marketing》(JM)等顶级

收稿日期: 2019-10-15

基金项目: 国家自然科学基金面上项目(71972092); 中央高校基本科研业务费专项资金项目(人文社科类)(18LZUJBWZY007)

作者简介: 柳武妹(1985—), 女, 兰州大学管理学院青年教授, 博士生导师(通讯作者, wumeijiayou@163.com);

黄河清(1996—), 女, 兰州大学管理学院硕士研究生;

叶富荣(1994—), 女, 兰州大学管理学院硕士研究生。

期刊上的绝大多数消费者行为研究采取了实验室实验(lab experiment)来考察消费者的感知和认知、态度、偏好、决策和行为。比如,发表在JCR创刊号上的文章就用实验室实验探究了消费者行为自身的非理性(Katona, 1974)以及儿童消费者对广告的态度(Robertson和Rossiter, 1974)等主题。尽管实验室实验允许研究者对被试进行随机分配,对实验产生的噪音加以控制,这在一定程度上增强了研究结论的内部效度(马尔霍特拉,2009),但由于消费者的实际消费行为发生在田野环境而非实验室环境中,且田野环境本身充满各种干扰因素和噪音,因此实验室实验降低了研究结论的外部效度(马尔霍特拉,2009)。

鉴于实验室实验的上述缺点,早在20世纪70年代就有消费者行为领域的学者开展田野实验,来弥补实验室实验的不足。比如,Craig和McCann(1978)就用田野实验探究如何通过信息沟通影响消费者的节能行为。后来也有学者用田野实验探究在总统选举投票时个体先前的回答如何影响他们随后的问题回应(Simmons等,1993);探究便利店购物时消费者所处的购买目标阶段如何影响他们的购买行为及对产品的心理表征(Lee和Ariely,2006)。但整体而言,消费者行为领域使用田野实验的文章数量不多,绝大多数消费者行为研究是采用实验室实验完成的。直至2016年前后,心理学尤其是社会心理学自身出现了研究结论无法复制的危机争议(如Fiske,2016;刘佳等,2018)。学术界对无法复制危机的出现进行了多方面归因,如概念不精确(Schaller,2016)、理论缺乏说服力(Fiske,2016),并认为从研究方法层面来看无法复制危机与单一方法(尤其是实验室实验)的重度依赖有关(Maner,2016)。因此,学术界开始思考有没有其他方法可以和实验室实验搭配使用,以增强心理学研究结论的可复制性。Maner(2016)就明确指出,田野研究(尤其是田野实验)既可以增强研究结论自身的可靠性,又可以增加研究结论在真实世界中的影响力(real-world impact)。由于心理学尤其是社会心理学常常为消费者行为研究提供理论支撑,因此,心理学自身的可复制性危机争议也引发了营销学者对营销研究可复制性的思考,以及对田野实验的相应重视。这可以说明为什么在2017年,JMR和JCR等要出版田野实验专刊(Meyer,2017;Gneezy,2017;Morales等,2017),来鼓励消费者行为学者开展田野实验,以增加研究结论的可复制性和对实践的指导价值。

顶级期刊JMR、JCR等2017年对田野实验的鼓励,大大提升了2018年和2019年消费者行为领域田野实验文章的比例。经笔者统计,仅2018年这一年,发表在JCR、JMR、JM等消费者行为顶级期刊上标注自己至少开展了一项田野实验、田野研究或在田野中开展的实验的文章比例就分别为17/67(25.4%,JCR)、11/60(18.3%,JMR)、11/51(21.6%,JM)。2019年从1月到9月,发表在这三份期刊上包含田野实验、田野研究、在田野中开展的实验的文章的比例均有提升,分别为12/46(26.1%,JCR)、10/39(25.6%,JMR)、9/40(22.5%,JM)(参见图1)。这些数据说明,如果不开展田野实验,学者们未来在JMR、JCR、JM等消费者行为顶级期刊上发表文章的概率可能会降低。

笔者也在国内消费者行为领域权威期刊(如《管理世界》《南开管理评论》《营销科学学报》等)官网的2018年和2019年各期文章中进行了阅读和搜索。我们发现,使用至少一项田野实验来探究研究假设的文章数量相对较少。樊亚凤等(2019)和MBA学生联合会合作开展了一项线上捐赠研究。冯小亮和任巍(2018)和国内某虚拟平台运营方合作开展了田野实验,以探究不同类型虚拟产品和社会影响易感性之间的关系。我们还发现,李杨(2017)在北京和宁波这两座城市分别开展田野实验,探究了不同的干预策略对不同区域公众参与旧衣回收行为的影响。然而整体而言,尽管我国的消费者行为权威期刊对田野实验有鼓励,但国内学术界使用田野实验的文章数量明显少于国际学术界。

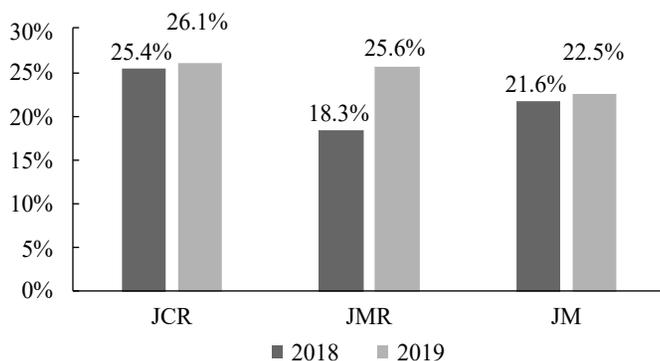


图1 2018年及2019年(1—9月)JCR、JMR、JM上发表的包含田野实验、田野研究或在田野中开展的实验的文章占比

仔细研读上述国际和国内权威期刊上至少包含一项田野实验的文章后,本文发现,学者们在田野实验的内涵、判别标准等关键问题上的看法存在不一致,这就使得一些宣称自己用田野实验来探究研究假设的文章实际上并不是真的使用了田野实验(后文会详细说明)。更重要的,究竟田野实验的自变量(IV)该如何操纵,因变量(DV)该如何观测?田野实验在开展过程中有没有一些一般性原理和方法可以遵循和借鉴?学者们对这些问题并没有详细回答。比如,Gneezy(2017)的文章仅简要介绍了营销研究中田野实验的概念、例子、类型以及潜在好处和成本,而对于田野实验究竟该如何开展,每类变量究竟该如何处理,不同类型的田野实验该如何开展,何时该开展田野实验何时又不该开展等问题,并没有进行探究。本文旨在厘清田野实验的内涵,并在研读了近5年(2014—2019年)发表在消费者行为领域权威期刊JCR、JMR、JM、《Journal of Consumer Psychology》(JCP)上的108篇田野实验文章的基础上,对田野实验中的变量操纵和观测做出归纳和总结,对田野实验的适用情境做出梳理并提出建议。

需要说明的是,本篇综述和国内外已有的田野实验综述有明显区别,并对这些综述有明显推进。首先,本文对Morales等(2017)的文章有推进。Morales等(2017)侧重于介绍如何提升消费者行为实验研究的真实性,文章的主题并不全部是田野实验,也涉及实验室实验。正是由于没有聚焦于田野实验,Morales等(2017)旨在高屋建瓴地提出怎样的消费者行为研究才是真实的研究,具有很强的理论导向性。但是Morales等(2017)并没有对田野实验中变量的处理方法进行归纳,也没有就田野实验的适用情境提供建议。本文对这一不足进行了弥补。其次,本文和Simester(2017)也有明显不同。Simester(2017)侧重于综述1995—2014年间发表在JM、JMR、《Marketing Science》、《Quantitative Marketing and Economics》以及《Management Science》上的因变量涉及行为的61篇包含至少一项田野实验的文章。Simester(2017)对这61篇文章进行了主题归纳和结论梳理。但在综述研究结论时,Simester(2017)没有在文中总结和归纳开展田野实验的一般性规律和方法。本文将站在Simester(2017)的肩膀之上,尝试总结出一些开展田野实验的一般性方法和建议。

二、田野实验及其概念介绍

(一)田野实验的定义

什么是田野实验/现场实验(field experiment)?表1详细列举了学者们对田野实验的定义及其侧重点。由表1可知,学者们对田野实验的界定有相似之处,也有差别之处。相似之处在于都

认为:田野实验需要在真实的田野环境中开展(如Simester, 2017; Morales等, 2017; 马尔霍特拉, 2009); 需要不做人为干预, 对实验过程不做过多控制(如Charness等, 2013; Gneezy, 2017; Morales等, 2017); 需要不让被试意识到自己正卷入一项研究(如Gneezy, 2017; Morales等, 2017)。存在争议的地方在于:田野实验中的因变量除了客观的实际行为之外, 是否还可以是自陈报告(问卷)测量的态度及偏好等。比如, Simester(2017)就认为田野实验的因变量是实际的行为, 而Morales等(2017)认为, 只要被试对先前的自变量操纵没有觉察, 田野实验的因变量就可以是用自陈报告测量的态度、偏好等。但Morales等(2017)也承认, 严格意义上的田野实验中的因变量应该涉及实际行为。结合上述共识性观点, 本文将田野实验界定为:“一种在真实田野环境下开展的探究自变量和因变量之间因果关系的实验类型。在该类实验中, 自变量是在真实田野环境下发生的、被试没有意识到的各类刺激; 因变量涉及实际行为; 同时, 由于研究人员对实验过程无法干预, 对被试无法随机分配, 该类实验设计的控制本身不严密, 但外部效率高。”

表 1 学者们对田野实验的定义

研究者(年份)	来源	观点	侧重点
Charness等(2013)	《Journal of Economic Behavior & Organization》	决定一项实验是否属于田野实验的一个主要因素是, 如果没有实验者的干预, 实验者感兴趣的活动中是否会发生, 如果会发生, 则是田野实验	实验者有无人为干预实验过程
Gneezy(2017)	《Journal of Marketing Research》	田野实验是一项旨在观察行为、充满噪音且对实验过程控制和干预很少的实验。在该类实验中, 被试没有意识到自己正卷入一项实验, 没有意识到自己的行为 and 结果正在被观测	旨在观察行为, 噪音大, 实验过程控制不严密, 被试没有意识到被观测
Simester(2017)	《Handbook of Economic Field Experiments》(第一版)	在真实场景(field)中开展的旨在观察行为反应的实验归类为田野实验。由于是营销研究, 所以该文关注的自变量是营销中的4P(product, price, promotion, place)刺激以及广告等刺激的有无、呈现方式等	在真实场景中开展, 因变量是行为反应
Morales等(2017)	《Journal of Consumer Research》	田野实验在严格意义上是指被试不知道自己正在一项研究中, 不知道一项实验操纵已经发生, 不知道自己的真实消费行为正在被观测。在该类实验中, 自变量是真实消费环境中的真实刺激, 因变量涉及行为	现场真实, 自变量真实, 因变量涉及行为, 研究者没有干预实验过程, 被试因此无觉察
Morales等(2017)	《Journal of Consumer Research》	如果在一个真实的消费场景下, 自变量是高度真实的, 而因变量以自陈报告的方式测量, 此时尽管因变量的真实程度比较低(因变量不是真实的行为), 但是这种研究仍然应该被视为田野实验	现场真实, 因变量用自陈报告的形式测量的话也可以是田野实验
马尔霍特拉(2009)	《市场营销研究: 应用导向》(第5版)	田野实验/现场实验是一种在现场真实环境中开展的、涉及研究组和对照组的实验。在该类实验中, 研究者对实验的干预和控制少, 因此实验本身内部效率低, 但外部效率高	在现场真实环境中开展, 有研究组和对照组, 干预和控制少, 内部效率低, 外部效率高

资料来源: 本文作者整理。

(二) 田野实验的判别标准

根据表1中学者们对田野实验的共识性理解, 本文认为符合下述所有5个条件的实验就是田野实验: (1) 在实验室之外的真实田野环境中开展; (2) 自变量是真实田野环境中的各类刺激, 涉及研究组和对照组; (3) 自变量涉及实际行为; (4) 实验者在整个实验过程中对被试的行

为表现不做任何干预;(5)实验设计不严密,允许干扰因素和噪音存在。相应的,我们认为宣称使用了田野实验,但具有下述4条特征的,则不是田野实验:(1)在非田野环境中开展;(2)自变量不是真实的刺激,而是一些问卷测项;(3)因变量不是行为,而是态度、偏好等主观报告;(4)实验者在实验过程中对被试的行为表现进行了人为干预。比如,尽管Lalwani和Wang(2019)标注自己使用了田野实验,但他们在研究消费者的文化背景和价值观如何影响他们对优惠券的使用倾向时,通过问卷测量了参与者的种族。这种问卷测量本身就可能引发被试对自己正在参与一项调研的警觉,从而可能导致行为反应失真。因此,这项研究不能算严格意义上的田野实验。又比如,Yang等(2019)在田野实验中让参与者思考和填写关于本地运动(vs.全球运动)的测项来操纵其本地身份(vs.全球身份)。由于在测量过程中,研究者需要介入,同时测量本身也引发了被试的警觉,所以这项研究也不能算严格意义上的田野实验。再比如,Wang等(2018)在咖啡馆开展田野实验,以检验解释水平和等待时间感知间的关系。但这项实验也不能算严格意义上的田野实验,因为在启动自变量时有研究者介入(通过不同的问题来启动高、低解释水平),同时这项实验的因变量属于主观报告,不涉及实际行为。

本文明晰上述田野实验判别标准的价值在于,通过明晰怎样的实验是田野实验,怎样的实验不是田野实验,(1)减少学者们对田野实验的误解;(2)促使学者们在进行田野实验设计时对如何降低被试的警觉有所考虑;(3)促使学者们开展严格意义上的田野实验,以及严谨地使用田野实验这一术语。

(三)田野实验及相关概念的区分

1.实验室实验(lab experiment)。实验室实验指一种在实验室情境下有计划地操纵自变量,控制噪音和干扰因素,以检验因变量变化的实验类型(麦克丹尼尔和盖茨,2019)。在这类实验中,被试会被随机分配,噪音会被尽可能地控制(郭秀艳,2004)。因此,实验室实验和田野实验是相对的,二者在实验发生的环境、分配被试的随机程度、实验控制的严密程度、被试对自己是否参与一项研究的觉察等方面具有明显不同。目前,消费者行为领域开展的绝大多数实验是实验室实验。在这类实验中,研究者常常在实验室中通过产品宣传册(如Noseworthy和Trudel,2011)、实验室中的环境因素如温度(如Huang等,2014)和拥挤程度(如Huang等,2018)等来操纵自变量,进而测量消费者的情感、认知和行为等的变化。

2.准实验(quasi experiment)。准实验是指根据现场的实际情况,灵活运用实验法的基本原理而进行的实验(时勘,1990)。这类实验对噪音的控制不及实验室实验那样严密,对被试不能做到完全随机分配,但是在解决实际问题上具有独特的功能,具有一定的外部效度(时勘,1990)。准实验和前实验(pre-experiment)以及真实验(true experiment)是相对的。在前实验中,研究者对噪音和干扰变量不会做任何控制,因此前实验是一种自然状态下的描述,内部效度和外部效度都很低(萨尔金德,2019)。在真实验中,干扰变量和噪音会被尽可能地控制,被试也会被完全随机分配到不同的水平下,该类实验常包含研究组和对照组,是一种控制严密的实验(萨尔金德,2019)。由此可知,田野实验和准实验比较接近,二者都在实际环境中开展,都对噪音和干扰因素控制不严密,都无法对被试随机分配。但二者也有明显不同,田野实验强调被试没有意识到自己正在参与一项研究以及因变量需要涉及实际行为等特征,而准实验并没有对这些进行强调。

3.田野研究(field study)。田野研究是指在实际的田野环境中开展的研究。这类研究可以是田野实验,也可以是二手数据研究,还可以是其他定性研究方法。田野研究区别于实验室实验的地方在于,研究情境是真实的而不是人为创造的。消费者行为研究中开展田野研究的文章有

一定数量,以二手数据研究为主。比如,Zwebner等(2013)在检验高温会提高产品估价的研究假设时,直接考察了2010年9月—2012年8月这24个月间的每日温度和顾客在一家网站八个品类上“点击购买”的点击流量间的关系。Gullo等(2019)通过分析超过25个月的一百多万户家庭每天不同时间的超市购物行为,发现消费者早上购买食物比其他时间段购买食物的多样化寻求程度更低。我国学者黄敏学和王薇(2019)也利用天猫平台收集了2016年8月26日—2016年11月1日某品牌41种白酒商品的16 410条后台销售订单数据,并检验了购物终端类型和产品价格水平对消费者延迟选择的影响。需要说明的是,除二手数据外,作为质性(定性)研究方法之一的扎根研究,也常常扎根于实际的田野环境来形成或检验理论(贾旭东和谭新辉,2010)。从这个意义上来说,我们认为,田野研究还包括扎根研究。

4. 在田野中开展的真实实验(realistic experiment in-the-field)。此类实验在实际的消费环境中开展,但消费者知道自己正在参与一项研究,且研究者会对于干扰因素和噪音进行设计控制(Morales等,2017)。举例来讲,在探究触摸的传染效应时,Argo等(2006,2008)让消费者来到大学书店并完成一项指定的试衣任务(衣服被其他消费者触摸过)。在他们的研究情境中,消费者知道自己正在被观察,但是也会进入研究任务,同时实验的因变量涉及产品评价。从这个意义上而言,在田野中开展的真实实验把田野当作实验室,但该类实验让被试对参与实验本身有觉察,因此并不是严格意义上的田野实验。

(四)田野实验的类型

目前,学术界对田野实验比较公认的划分方法是,根据是否涉及非学术合作方将田野实验划分为独立开展的田野实验(independent field experiment)和合作开展的田野实验(collaborative field experiment)(Gneezy,2017)。接下来介绍这两类实验各自的特征,并通过具体的文献例子加以说明。

1. 独立开展的田野实验。在独立开展的田野实验中,研究者可以自己设计并开展实验,以便让被试在毫无觉察的情况下做出真实反应。最常见的独立开展的田野实验是“pay-what-you-want”(为你想要的付款)范式的实验。举例来讲,Atasoy和Morewedge(2018)在检验实体产品是否比电子版产品更值钱时,让一位研究助理在曼彻斯特—波士顿地区Old North Church附近穿上Paul Revere(保罗·列维儿,教堂的知名历史人物)的衣服,邀请参观教堂的86名游客一一和他拍照。拍完照后询问游客愿意要哪一种照片及愿意为照片付多少钱。助理进一步告诉被试他们付的钱将用来维修教堂。但在整个实验流程介绍中,Atasoy和Morewedge(2018)并没有强调该实验和教堂合作,也没有强调该实验是为了解决教堂当前面临的某种担忧,因此我们把它归类为独立开展的田野实验。

2. 合作开展的田野实验。在合作开展的田野实验中,研究者需要找到非学术合作方,以解决合作方目前面临的难题,并争取他们的赞助和支持。在已有的消费者行为研究中,需要合作开展的田野实验往往会和便利店、超市、餐厅、大学公共事务管理部门、酒店等合作。比如,Grinstein和Kronrod(2015)发表在JMR上的田野实验就是和一所美国中西部大学的管理部门合作的,以解决管理部门当前面临的大学生不愿意用洗手液洗手的担忧。再比如,Baca-Motes等(2013)的研究则是和酒店合作开展的,以解决酒店当前面临的环保节能需求。在他们的研究中,他们通过让酒店住客签署(vs.不签署)环保承诺书来检验环保承诺是否能够促进酒店住客的环保行为(如循环使用毛巾、离店关灯等),以及是否佩戴环保胸针的调节作用。Castro等(2013)的研究则和一家小型健康型食物商店合作,检验货架上鸡蛋的杂乱摆放是否会影响到消费者对产品的购买,从而帮助这家食物商店实现鸡蛋销量的突破。由此可见,在需

要合作开展的田野实验中,实现研究者的研究目标和非学术合作方的实践目标是关键。

(五)田野实验的作用

尽管田野实验近几年来开始受到国际期刊的青睐,但不可否认的是,田野实验自身有一些劣势导致学者们不愿意开展该类实验。Gneezy(2017)、Cialdini(2009)以及Charness等(2013)等文献对田野实验劣势的共识性看法包括:(1)内部效度和数据分析方面的挑战。在田野实验中,研究者很难随机分配被试,很难对田野中的噪音和干扰因素进行控制,这就使得通过田野实验只能得出相对较弱的因果推论,内部效度相对较低;同时,实验中存在大量干扰变量和噪音,给数据分析带来挑战。(2)无法探测中介机制。田野实验聚焦于观测能够观察到的行为,因此限制了研究者对行为背后的中介机制的探究,也限制了研究项目的理论贡献。(3)执行起来比较耗时、耗力。这一点在需要合作开展的田野实验中更为明显(Charness等,2013;Gneezy,2017)。研究者在实验开展前需要花时间找到非学术合作方,并且需要花时间洞察他们的需求,在准备实验和收集数据时也需要花费比实验室实验更长的时间。

然而,上述劣势并不能阻碍田野实验对一项研究的巨大作用,这种作用主要体现在能够提升研究的真实性和说服力。具体而言,田野实验可以帮助研究者检验一个已经建立的理论在实际生活中的适用性,能够帮助学者们提升研究结果的效应值,还能够帮助学者们捕捉研究结果的长期效应(VanEpps等,2016;Gneezy,2017)。同时,田野实验还能够解决自变量的内生性问题(Simester,2017),因此可以和问卷调查及二手数据法等自变量内生性问题相对明显的方法搭配使用。

三、消费者行为领域的田野实验:操作介绍

如前文所述,目前营销学术界发表的几篇介绍田野实验方法的文章要么是简短的编者案类文章(如Meyer,2017;Gneezy,2017),要么侧重点不全在如何开展田野实验上(如Morales等,2017),要么尽管综述了营销学中的田野实验,但没有额外总结出田野实验的开展建议(如Simester,2017)。那么,究竟消费者行为领域田野实验的自变量该操纵什么?如何操纵?因变量该观测什么?如何观测?研究者为了建立自变量和因变量间的因果关系可以在田野实验中做什么?这一部分将梳理和归纳近5年来(2014—2019年)发表在消费者行为期刊上的、至少包含一项田野实验的文章,以尝试回答上述问题。

有几点需要说明:(1)本文的聚焦点是消费者行为,因此我们关注四家消费者行为顶级和权威期刊(JM、JMR、JCR、JCP)上的田野实验文章。对于发表在社会心理学、认知心理学等期刊上的涉及消费者认知判断和行为的文章,我们没有搜索和纳入分析,如Zhou等(2012)在菜市场开展的金钱实验。(2)我们在浏览JM、JMR、JCR和JCP近5年的各期文章时,并没有简单地以关键词field experiment进行搜索,相反,我们下载了每一期文章进行泛读,并对其中含有至少一项田野实验(有些文章写的是field study,但实际上是field experiment,因此也计算在内)的文章进行精读。(3)基于上述工作,我们发现这四份期刊在2014年8月—2019年11月^①的5年间刊发的田野实验文章数量分别为26(JM)、46(JMR)、24(JCR)、12(JCP),共计108篇。(4)我们没有对这108篇文章像Simester(2017)那样按研究主题(关注产品、广告等)进行归纳、整理,而是基于所有文章对自变量的操纵以及因变量的观测方法进行归纳整理,并尝试得出一些共通的、一般性的规律。在这个过程中可能会漏掉和放弃个别几篇无法归类的文章,但尽可能做到

^①将近5年确定为2014年8月—2019年11月,是因为JM、JCR、JMR、JCP每年文章的期数不一致,而且JCR是偶数月刊发,JMR则是奇数月刊发。本文第二作者负责整理JCR和JMR上的2014年8月—2019年8月的田野实验文章,而本文第三作者负责整理JM和JCP上2014年11月—2019年11月的田野实验文章。三位作者讨论后进行汇总和该部分撰写。

全面和无偏。

(一)消费者行为领域的田野实验对不同类型自变量的处理方法

消费者行为领域的自变量根据研究需要往往五花八门,但大体上可以归结为以下几类:个体特征类、产品刺激类、环境情境类以及其他类型的自变量。接下来一一归纳这几类自变量的处理方法。

1. 个体特征类自变量的处理方法

(1)通过语言文字启动与自我相关的个体特征。研究者可以在实验开始时通过符合实际消费情境的语言文字来操纵个体特征。这些个体特征基本上都与自我相关,如自我建构(self-construal)、自我关注(self-focus)、自尊水平、社会身份等。具体而言,Simpson等(2018)和Allen等(2018)通过海报上的人称代词(我记得/我们记得)以及呼吁募捐时的措辞表述(你/我们可以向那些付出这么多的人致敬,并通过捐款来帮助他们)来操纵独立型自我建构/依赖型自我建构。Chang和Hung(2018)在真实的募捐呼吁书的顶部突出自我参照信息(例如“我捐赠”或者“他们捐赠”)以操纵自我关注。类似的,Fajardo等(2018)通过改变捐赠呼吁邮件和募捐标志中的信息内容来操纵捐赠者对自己(vs.组织)的关注。他们发现,强调捐赠者相关信息会提高捐赠率,而强调组织相关信息则会提高捐赠金额。此外,Grewal等(2019)则借助商店内的广告语(“你太棒了!”)来提高消费者的自尊水平。Garbinsky和Gladstone(2019)在探究消费者的社会身份是否会影响他们对价格的敏感性(用购买行为观测)时,在中国合肥开展了田野实验。他们让装扮成店员的研究助理向进入店铺的消费者发放店铺宣传手册,通过宣传手册上的语言来操纵消费者的社会身份(全球身份vs.本地身份)。本地身份组被试看到的宣传手册描述了一场支持当地社区和当地企业、重点关注当地新闻并维护当地传统的“地方思考运动”,而全球身份组被试看到的宣传手册描述了一场支持全球品牌和全球企业、关注全球新闻、强调世界其他地区文化的“全球思考运动”。

(2)通过语言文字启动个体与他人是否相似、是否和他人共用账户等变量。比如,Laporte和Briers(2019)通过在邮件中提及先前获奖者的名字,来操纵被试与先前获奖者的相似性感知。他们发现,当之前的获奖者和被试的姓名具有相似特征时,被试的抽奖参与率明显提升。又比如,Garbinsky和Gladstone(2019)在研究消费者的财务账户类型(单独拥有账户vs.和伴侣共用一个账户)对实用品购买的影响时,在研究1中就直接在校园篮球场入口处放置摆放了享乐品和实用品的桌子。当对桌子上的商品感兴趣的学生主动过来询问时,研究人员给学生一个装有1美元的信封,并通过信封上的语言文字操纵财务账户类型。单独账户条件下被试看到的信封封面语是“这钱是给你的(不是给你的伴侣的)”,而共用账户条件下被试看到的信封封面语是“这钱是给你和你的伴侣的”。在研究2中,他们继续在校园篮球场入口处做田野实验,他们先询问^①被试是单独使用银行卡还是和伴侣共同使用银行卡,根据被试的回答让他们从钱包中实际拿出自己单独使用的银行卡或和伴侣共同使用的银行卡。

2. 产品刺激类自变量的处理方法

(1)通过实物操纵产品特征或与产品相联系的他人特征。当自变量是具体的产品特征或与产品相联系的他人特征时,可以使用真实的产品实物来操纵。在这类田野实验中,产品实物可以直接让消费者看到,也可以隐蔽式出现在环境线索中。究竟该如何展示,取决于研究者的研究目的。如果研究目的是考察消费者对产品的价值估计或对产品属性信息的感知,就可以直接让消费者看到实物产品。反之,如果研究目的是探究和产品有联系的他人的存在如何影响消费

^①由于是由销售人员询问,这在日常消费情境中也常有,因此不会引发被试的觉察,属于田野实验。

者的决策,实物产品就可以隐藏式出现。

具体而言,为了检验消费者对实体产品和电子版产品的价值估计,Atasoy和Morewedge(2018)要求扮成Paul Revere的研究助手在与游客合影后,向游客提供实体照片或电子版照片,来观察游客对实体产品和电子版产品的价值估计(通过游客为照片支付的实际金额来反映)差异。类似的,Wu等(2017)通过美国西南部一家健身房浴室卫生纸的图案操纵了消费者对卫生纸的美学感知。低美学感知组被试看到的卫生纸是普通的、没有图案的白色纸,而高美学感知组被试看到的卫生纸则带有节日图案。无独有偶,Townsend(2017)通过学校社区年会邀请函的设计来操纵邀请函的美学感知。产品的展示(摆放)方式也可以通过实物来操纵。Guha等(2018)通过改变可口可乐实物饮料的摆放方式(对齐vs.不对齐)来操纵产品属性信息的展示方式。Deng等(2016)在美国一家大型连锁超市开展了田野实验,他们将该家超市的32种带有香味的洗手液水平摆放或垂直摆放,并记录了消费者选择的洗手液种类和数量。他们的研究发现,水平摆放(vs.垂直摆放)增加了消费者的多样化选择行为。在检验和产品有联系的他人的存在是否会影响消费者决策时,研究者则可以让实物产品隐藏式出现。比如,Otterbring等(2018)没有让店铺中的高价产品或大标志产品凸显,以便让消费者将注意力放在商店入口处的健壮他人是否存在上。他们发现,身体素质占优势的男性员工站在商店入口处欢迎顾客,会让身体素质缺乏优势的男性顾客感受到自我威胁,进而增加对高价格产品的消费。

(2)通过非常凸显的语言文字启动产品相关信息。当研究者关心的目标自变量是与产品相关的线索或信息(如产品的排名方式、产品广告的特征、产品的外观、产品是否可转换等)时,可以在开展田野实验时通过产品广告、海报、包装箱等上面的语言文字等来启动这类自变量。具体而言,Sevilla等(2018)通过文字操纵获奖奶酪广告的排名方式(数字排名vs.百分比排名),以观察不同的排名方式对消费者决策的影响。Wang等(2019)开发了两类星巴克广告(含促进导向说服性信息的广告vs.含预防导向说服性信息的广告),并将其和星巴克咖啡的故事结合,以检验动机不匹配策略对产品共同创新(cocreated innovation,即企业利用顾客的创意将顾客的想法做成产品成品)的影响。类似的,Grewal等(2019)通过农产品包装箱上的标识来操纵产品外观,吸引人的外观条件下的标识是“装着漂亮的苹果”,而不吸引人外观条件下的标识是“装着不漂亮的苹果”。由于该研究的目的是通过提升自尊水平让消费者购买外观丑陋的农产品,因此店铺里还张贴了横幅广告。高自尊广告的内容是“You are Fantastic! Pick Ugly Produce!”,而对照组广告的内容是“Pick Ugly Produce!”。此外,Winterich等(2019)通过网页上的牛仔裤回收广告内容来操纵产品可转换信息在产品回收广告中的凸显。他们发现产品可转换信息会提高消费者对产品的回收利用。从上述例子可知,通过语言文字启动产品相关信息的惯例在已有的消费者行为田野实验中运用得极其普遍。但在通过海报和广告中的语言文字启动产品相关信息前,还是需要开展预研究和预测试,以确保海报信息在消费者看来和当前消费情境是匹配的,不然消费者会觉察到自己正在被观察。

(3)通过环境中的隐藏线索启动产品相关信息。与上述第二类操纵方法不同,该类操纵方法常常会借助环境中隐藏的线索来启动产品相关信息(如产品广告的透明度、产品价格等)。例如,Kim等(2019a)通过改变网页侧边栏标题操纵广告透明度。具体而言,在广告透明度高的条件下,被试看到侧边栏标题为“根据您对我们网站的点击推荐”或者“根据您与我们共享的内容推荐”;而在广告透明度低的条件下,被试会看到标题仅仅为“推荐”。类似的,Atlas和Bartels(2018)和一家美国熟食配送公司合作,利用网页横幅来操纵产品价格的显示方式(每天16美元vs.每周99美元),发现相比每周99美元,每天16美元这种定价方式能够提升消费者对餐

食的购买量。线上田野实验具有数据收集方便客观(例如可以在后台记录顾客的点击次数,页面停留时间)、数据量大等优势,得到越来越多学者的使用(Simester, 2017)。

3. 环境情境类自变量的处理方法

(1)通过自然发生的环境情境因素操纵与空间感知相关的自变量。这类操纵常见于拥挤(如Machleit等, 2000; Puzakova和Kwak, 2017; Hock和Bagchi, 2018)、社会密度(如Knoeferle等, 2017)等方面研究。实验地点通常选择商场或者校园等人口集中的公共场所。这些研究的自变量各水平常自然发生,不需要人为创造,但在实验结束时需要对自变量的操纵进行主观核查检验。经过梳理,我们发现,在有关拥挤和社会密度的田野实验中,学者们对拥挤和社会密度的操纵核查分为两类:一是让被试或研究助理填写自变量的主观核查;二是让独立的研究助理对自变量进行客观数值计算。Hock、Consiglio等学者的研究就采用了第一种核查方法。例如, Hock和Bagchi(2018)在研究社会拥挤与卡路里消耗的关系时,在实验结束时让参与者对商场或者餐厅的拥挤程度进行主观评价。而Puzakova、Knoeferle等学者的田野实验则采取了第二种核查方法。例如, Puzakova和Kwak(2017)将单位时间内的人流量除以总面积作为拥挤的指标, Knoeferle等(2017)将一小时交易数量除以商店面积作为社会密度的指标。

我们认为,上述操纵核查方法尽管常用,但也存在一些弊端。如果让参与者主观评价拥挤程度等类似自变量的情况,很可能在最后降低实验的真实性,导致被试猜出实验目的。而如果让外部观察者评价或者通过客观数据衡量自变量的情况,则需要保证外部观察者的拥挤感知与参与者一致。因此,研究者自身需要在这两类操作核查方法间做出利弊权衡。

(2)通过人为创造的环境情境因素操纵偶发情绪。研究人员可以人为地让消费者处于不同的环境情境下来操纵偶发情绪,也可以在消费者所处的环境中提供某种刺激线索来操纵偶发情绪。目前,消费者行为领域的田野实验采用该类方法操纵的偶发情绪有控制感、鼓舞或灵感、内疚感等。例如, Jami(2019)通过改变海报在彩票销售柜台后面悬挂的位置高低来操纵视角的变化。他们发现,海报使用高海拔视角,会让被试产生一种虚幻的控制感,从而提高其风险承担能力。此外,研究者还通过店铺的促销展示来启动消费者灵感(inspiration)。比如,在Böttger等(2017)的研究中,在低灵感条件下,消费者看到的店内促销活动展示的是有机产品和相应促销材料;而在高灵感条件下,消费者看到店内除了展示有机产品外,还展示了三种使用特色产品作为配料的食谱(即汤、主菜和甜点的食谱),并提供了突出这三种食谱的宣传材料。消费者的内疚感也可以通过文字来启动。Thomas和Katherine(2015)的田野实验在英属哥伦比亚大学的校园里开展。高内疚感组被试看到的校园海报描绘了来自当地血液服务组织的信息,并且海报上的标语写着:“没有东西可以代替血液。您的犹豫可能会使他人丧命。”而对照组没有看到海报。他们发现,看到诱发内疚情绪的海报后,消费者会增加对自我提升类产品的选择。

(3)通过人为创造的环境情境因素操纵情境性感官体验。消费者的情境性感官体验,如嗅觉体验、触觉体验、听觉体验以及视觉体验等,也可以在田野实验中被操纵。比如,研究者可以通过在店铺中使用喷雾器散播放纵性食物的气味(披萨味)或健康食物的气味(苹果味)(Biswas和Szocs, 2019)来操纵嗅觉体验,通过用隐藏的扬声器发出频率不同的声音(Lowe等, 2019)以及在商店内播放节奏快慢不同的背景音乐(Knoeferle等, 2017)等来操纵听觉体验。Wang等(2016)在街头开展了一项真正的募捐活动,通过人为创造募捐材料的触感来操纵消费者的触觉体验(粗糙vs.光滑),并检验不同的触觉体验是否会影响到人们的捐赠行为。粗糙组看到的募捐材料的背面覆盖着一张60沙砾砂纸,而光滑组看到的募捐材料的背面覆盖着一张

光滑的投影纸。此外,Mehta等(2017)在探究屏幕色彩是否会影响儿童参与线上恶作剧的频率时,就通过在后台设置儿童等待时看到的网站屏幕上的色彩(红色vs.白色vs.蓝色)来操纵视觉体验。这项研究发现,相比于白色和蓝色,当等待屏幕是红色时,恶作剧对话的比例更高。

4.其他类型自变量的处理方法

除了上述可以明确分类的自变量,一些研究者还根据自己的兴趣关注了其他类型的自变量。这些自变量大体上可以分为以下几类:与金钱相关的自变量、与组织和消费者间信息沟通相关的自变量以及与机器和智能相关的自变量。

(1)与金钱相关的自变量可以实际记录或用文字操纵。比如,Chen等(2017)研究中的自变量为支付方式(银行卡vs.现金),他们直接记录了被试结账时使用的是银行卡还是现金。当无法直接记录时,可以通过语言文字对与支付相关的自变量进行操纵。比如,Mazar等(2018)研究中的自变量为信用卡公司的还款时间提示,他们通过按键提示来操纵这个自变量:按键“1”表示将在24小时内还款;“2”表示将在36小时内还款;“3”表示将在48小时内还款;“4”表示将在72小时内还款。在Mohan等(2018)的田野实验中,在“披露公司高管薪酬”条件下,研究者以文章照片的形式向消费者展示了零售商员工薪酬支付比率信息,Urban Outfitters首席执行官的薪酬是普通员工的三倍,Gap公司首席执行官的薪酬是普通员工的705倍;而在“未披露”条件下,消费者只在两家零售商的网站上看到了一般信息的截图,没有看到首席执行官薪酬的披露。

(2)与组织和消费者间信息沟通相关的自变量可以实际记录或实际操纵。组织和消费者之间的沟通媒介、沟通语调、沟通时表情符号的采用等会对消费者的行为产生影响。对于表情符号的采用这类可以直接记录和统计的自变量,研究人员选择了直接记录而非操纵。比如,Li等(2019)的研究就在真实的人工客服与消费者互动的情境中,直接记录了人工客服与消费者的在线交流是否使用了表情符号,以及使用表情符号的数量。而对于沟通媒介这类自变量,研究者可以直接采用不同媒介向消费者传输相同的信息。比如,Kaju和Maglio(2018)的研究就直接通过电子邮件(vs.手机短信)向消费者发送信息,两组被试看到的信息都是“我们最近增加了一些职位”,但是收到信息的渠道不同。该研究检验了不同的沟通媒介是否影响消费者的在线信息更新行为。对于沟通语调(肯定vs.非肯定,批评vs.表扬)这类自变量,Grinstein和Kronrod(2015)通过让大学教学楼公共洗手间张贴的告示以及发送给大学教工的邮件主题中含有must(vs.can)来凸显肯定(vs.非肯定)语调,含有not doing enough(vs.done a lot)来操纵批评(vs.表扬)性的沟通语调。

(3)与机器和智能相关的自变量可以通过文字和语言陈述来操纵。Roland等(2018)的研究检验了产品包装上的文字是手写还是机器书写是否影响消费者对产品的购买。他们在研究中直接让被试看含有手写体文字(vs.机器书写的文字)的包装。Bhargave等(2016)的研究通过酿酒厂代表的语言陈述来操纵“云”线索是否存在。“云”线索存在组听到的语言陈述是“我们有一个网站,在那里你可以找到关于我们所有葡萄酒的信息”,并且同时酿酒厂代表拿出平板电脑展示“云”信息;而“云”线索不存在组并没有听到相应的陈述。

从以上对田野实验中自变量操纵方法的总结归纳中可以看出,消费者行为领域的田野实验对自变量的操纵方法整体上分为三类:自然展示、实际记录和语言文字操纵。其中语言文字操纵用得最多,学者们常常用促销海报、店铺广告、手机信息等上面的语言文字来操纵感兴趣的自变量。同时也可以看出,对于同一类型或同一主题的自变量,学者们可以将三类方法交叉使用,但需要确保操纵没有引发被试的警觉。

(二) 消费者行为领域的田野实验对不同类型因变量的观测方法

尽管田野实验中的因变量涉及实际行为,但是消费者的行为纷繁复杂。近5年的108篇田野实验文章究竟关注了哪些类型的行为?对这些行为进行了怎样的操纵?通过梳理,我们发现,这108篇田野实验文章关注的行为整体上可以分为以下几类:(1)与食物选择相关的行为;(2)与捐赠相关的行为;(3)与网站点击和网站访问相关的行为;(4)与产品购买和再购相关的行为;(5)与信息分享和信息更新相关的行为;以及(6)销量、销售额、购买金额等行为结果。对于上述行为以及行为结果,学者们一般采取下述方法进行观测:

1. 采取“pay-what-you-want”范式衡量消费者的产品估值和捐赠意愿。这类实验往往涉及实际的支付行为或购买行为(如Atasoy和Morewedge, 2018),会通过“pay-what-you-want”范式完成。其内在原理是,消费者越愿意付款表明其对产品的价值估计越高。Atasoy和Morewedge(2018)发表在JCR上的文章就采取了这一范式。他们让参与者为实体照片或电子版照片付出任何数量的金额来衡量其对实体产品或电子版产品的价值感知。尽管“pay-what-you-want”范式一般用来衡量消费者对产品的价值估计,但新近有学者开始将这一范式运用到捐赠研究中。其潜在假设是,越愿意为慈善机构付款表明对慈善机构和需要帮助者的捐赠意愿越高。比如,Simpson等(2018)、Fajardo等(2018)以及Chang和Hung(2018)等在捐赠相关研究中就采用了这一范式。

2. 实际记录消费者的捐赠行为。比如,Hsee等(2015)在记录消费者的捐赠行为时,通过个体捐赠的硬币数量衡量消费者捐赠行为的强弱。Townsend(2017)通过募捐公益组织中不同美学设计和成本的门票的实际销售情况间接衡量消费者的实际捐赠行为。Simpson等(2018)则用是否捐赠及捐赠金额来衡量参与者的捐赠行为。Anik等(2014)通过某非营利组织网站上选择是否从一次性捐赠转变为每月捐赠的捐赠者的比例,来衡量消费者的捐赠承诺。

3. 对消费者的是否型二元选择行为进行记录。当研究的结果变量不涉及金钱付出时,最真实的因变量观测方法就是记录消费者的真实选择。这类真实选择是在实验人员营造的真实情境下诱发的消费者反应,常用来测量消费者的多样化寻求行为、产品选择行为、优惠券选择行为以及垃圾回收行为等,目前被田野实验的研究者频繁使用。具体而言,在研究消费者的多样化寻求行为时,Yoon和Kim(2018)在实验结束时记录消费者会选择常见产品还是不常见产品。在研究消费者的产品选择行为时,研究者会记录消费者对不同产品的真实选择,例如消费者究竟是选择了高热量饮料还是低热量饮料(Guha等, 2018),是选择了外观丑陋的苹果还是外观吸引人的苹果(Grewal等, 2019)等。在研究消费者的优惠券选择行为时,研究者会记录消费者是选择了优惠时间信息水平显示还是垂直显示的优惠券(Romero等, 2019)。在研究消费者的垃圾回收行为时,研究者会记录消费者在看到产品可转换信息后是否选择回收废弃物(Winterich等, 2019)。从上述例子可以看出,不同形式的产品选择其实在测量不同的抽象概念。用对漂亮外观苹果的选择来反映消费者的美观偏好(Grewal等, 2019);用对水平或纵向显示时间的优惠券的选择来测量消费者的时空概念。因此,研究人员应当基于特定的抽象概念,结合实验设计情境,将抽象概念用产品选择来操作化定义,并对产品选择进行记录。

4. 对消费者的真实选择数量或食物摄入数量等连续类因变量进行统计。如果说上一种方法针对的是“是否型选择”这类类别型因变量,该类方法则适用于购买金额、食物摄入数量、选择数量等连续型因变量。在实际的田野实验中,研究者常对这类因变量进行客观计数或计算。首先,针对购买金额类因变量,研究者常会统计消费小票中的金额。具体而言,在研究他人的身体优势是否会增加男性消费者的消费支出时,Otterbring等(2018)就利用男性消费者的消费小

票计算出了他们的店内消费总额。而在研究彩票海报视角的海拔高度是否会影响消费者的风险寻求时,Jami(2019)同样记录了彩票小票中的消费总额。其次,针对食物摄入总量和热量这类因变量,研究者常会计算出食物订单中的总热量。比如,Hock和Bagchi(2018)在研究拥挤环境对消费者放纵性消费的影响时,就通过消费者在菜单上勾选的食物类别和食物数量计算出了每份食物订单的总热量。同样,Biswas和Szocs(2019)借助教育局和学校行政部门统计的自助餐厅中食物的出售数据来统计消费者对不健康食物的选择数量。他们发现,长时间暴露在放纵性食物的气味(如披萨味)中会降低学生对不健康食物的选择数量。此外,还有研究统计了暴露在繁忙线索下学生的食物购买量(Kim等,2019b),暴露于声音威胁条件下消费者为了规避风险所消耗的酸奶样品数量(Lowe等,2019)。

整体而言,采用数量类连续型因变量的田野实验由于借助小票、菜单、订单等二手客观资料上的数据来衡量消费者的放纵性消费、风险规避消费等,得出的证据更客观。同时,这类田野实验研究也常常在实际消费环境下开展,涉及非学术合作方,因此研究结论的说服力和实践参考价值更强。相比之下,尽管“是否型选择”这类类别型因变量的研究结论也是客观的,但由于缺乏消费小票、订单等客观资料的支撑,以及缺乏非学术合作方的支持,研究结论的说服力和实践参考价值相对弱一些。从这个意义上而言,开展以购买金额、选择数量、摄入数量等为因变量的田野实验更值得鼓励。

5. 对网站访问量和点击率等行为数据进行记录。除上述测量指标外,尤其是在线上消费情境下,网站访问量和点击率也为相关田野实验研究提供了丰富的因变量数据来源。很多研究采用广告点击率测量消费者的真实选择(Winterich等,2019;Li等,2019)。Kim等(2019a)用推荐项目的点击率、花在推荐项目页面上的秒数,以及推荐项目产生的收入等真实表现来测量消费者的产品购买意愿(即广告有效性)。Gai和Klesse(2019)也采用推荐点击率来衡量消费者的真实选择偏好。Paharia和Swaminathan(2019)以及Kamleitner等(2019)的研究均通过记录Facebook上的点击率来测量用户的实际偏好。Putnam-Farr和Riis(2016)与一家大型健康服务公司合作,向这家公司的客户发送电子邮件,通过链接点击率和实际后续活动参与率来测量客户的参与行为。Grinstein和Kronrod(2015)在其第二项研究中通过记录不同操纵条件下视频链接的点击率来测量消费者的期望行为。

除点击率数据之外,网站访问量也是相关田野实验研究常用的因变量数据来源。Lobschat等(2017)通过一家德国零售商公司网站的访问量来测量不同操纵条件下的广告效果。与此同时,也有研究通过个人cookies追踪来测量消费者的广告浏览行为(Hoban和Bucklin,2015)。相比“pay-what-you-want”范式,以及消费者是否型选择(类别变量)和数量型选择(连续变量),网站访问量和点击率数据的测量相对间接,且较为客观,更适用于在线消费情境下的田野实验研究。

6. 对消费者个人信息共享和采集相关行为进行记录。在在线消费情境下,除网站访问量和点击率数据之外,消费者个人信息的共享和采集本身亦可作为相关田野实验研究因变量的数据来源。Bone等(2014)通过采集银行人员所提供的信息和消费者申请贷款所需要的信息测量了对少数族裔的限制性选择这一因变量。Huang和Broniarczyk(2015)通过记录参与者信息共享的长度(单词数)测量了其自主信息共享行为。而在Kaju和Maglio(2018)的研究中,应聘者被要求点击一个链接查看新的工作,研究者通过其是否点击另一个链接更新他们的在线资料来测量消费者的信息更新行为。

(三)通过田野实验检验自变量和因变量间因果关系的一些方法和技巧

田野实验由于在真实的消费环境下开展,对噪音不加干预和控制,因此成功率会很低。为了增强田野实验的成功率,研究者可以做什么?这一部分将对此问题进行简要回答。整体而言,通过文献梳理我们发现,在消费者行为顶级期刊上汇报的田野实验(这些田野实验都是结果显著和成功的)具有下述特点:

1. 理论自身强健。理论自身的强健和可复制是一项研究的开展基础,因为强健的理论才能增强实验和研究自身的可复制性(Fiske, 2016)。上文介绍的田野实验文章都存在一个共有的特点:理论和假设的推演极为严密。比如,Hock和Bagchi(2018)在开展拥挤影响高热量消费的田野实验之前,详细分析了社会拥挤影响分心的文献,以及分心对高热量消费的影响,从而得出了自己的理论假设。类似的,Atasoy和Morewedge(2018)在教堂前开展产品形式(电子版vs.纸质版)是否会影响产品估价的田野实验之前,也详细梳理了产品形式和拥有者身份感知领域的文献,以及产品拥有者身份感知和产品估价领域的文献。从这些例子可以看出,开展田野实验的基础之一是理论自身强健,而理论自身强健的基础则是对过往文献深入全面的梳理。

2. 通常在实验室实验成功之后开展田野实验。尽管少数消费者行为领域的田野实验文章在理论推演完成之后就先开展田野实验(如Barasch等,2018;Kim等,2019b;Shen等,2019;Wang等,2019),但绝大多数文章还是在开展完几项实验室实验后才开展田野实验(如Chang和Hung 2018;Jami, 2019;Kim等,2019a)。这样做可以利用实验室实验的结果更好地设计田野实验,进而提高田野实验的成功率。

3. 人为地准备田野实验的条件以得出自变量和因变量间的因果关系。首先,需要人为地准备实验场地。涉及非学术合作方时,研究人员需要提前和非学术合作方沟通,在消费者正在消费的餐厅(Hock和Bagchi, 2018)、奶酪店(Sevilla等, 2018)、大学教学楼的公共卫生间(Grinstein和Kronrod, 2015)等场所开展实验,以检验所提理论。其次,需要选择合适的实验材料、实验时间和地点,并对一些意外因素进行些许控制,以得出自变量和因变量间的因果关系。Sevilla等(2018)在研究产品不同的排名格式对消费者判断的影响时,选择刚刚获奖的Cirrus奶酪作为实验材料,并选择与位于旅游景点的一家奶酪店合作。研究人员进行如上准备正是因为Cirrus奶酪在2017年美国奶酪协会比赛中的真实获奖满足了实验对产品排名格式的要求,并且合作奶酪店的位置保证了每日客流量。而Grinstein和Kronrod(2015)的田野实验就选择在大学教学楼的公共卫生间使用频繁的周一到周五来开展,而不是在周六和周日学生没课的时间开展。

4. 分析数据时对噪音和干扰变量加以控制。田野实验的干扰因素之一是自然环境中的不确定性所产生的噪音,这是田野实验特有的噪音。这种噪音不可能被完全消除,但是研究人员在分析数据时也应该尽可能地进行控制。Gneezy(2017)认为,在控制田野实验的噪音时,研究人员应该开发和维护一个系统,使他们能够实时捕捉和解决意外情境因素(如时间、温度的影响)。Gneezy(2017)进一步认为,在分析数据时研究人员应该将这些意外因素控制起来。一些研究者采纳了这一建议。Hock和Bagchi(2018)在研究拥挤对卡路里消耗的影响时,选择在周二(餐厅最忙)和周五(餐厅最不忙)的午餐时间进行调查,从而证明在最拥挤和最不拥挤的日子里,繁忙对卡路里消耗这一变量没有影响。又比如,由于天气已经被证明会影响消费,因此Knoeferle等(2017)在探究商店音乐节奏与社会密度对消费者支出的影响时,就从国家天气报告数据库中获取了实验期间的气象数据(每小时室外温度、每小时降雨量)作为额外的协变量,以排除气象数据对消费者支出的影响。

四、消费者行为领域两类田野实验的开展建议

上文提到,消费者行为领域的田野实验大体上可以分为独立开展的田野实验以及需要和非学术合作方合作开展的田野实验。整理JCR、JMR、JM以及JCP上近5年的108篇田野实验文章后我们发现,这四份期刊上独立开展和合作开展的田野实验文章的数量分别如下:JCR含10篇独立开展、14篇合作开展的田野实验文章;JMR含6篇独立开展、40篇合作开展的田野实验文章;JM含6篇独立开展、20篇合作开展的田野实验文章;JCP含7篇独立开展、5篇合作开展的田野实验文章^①。通过上述数据可以看出,JMR和JM上的文章多侧重于合作开展的田野实验,因为这两份期刊关注的研究主题相对宽广,对研究本身的营销实践意义有明显要求;而JCR和JCP上的文章在独立开展和合作开展这两类田野实验上的分布相对均匀。接下来分别给出这两类田野实验的开展建议。

(一)需要独立开展的田野实验的开展建议

1.灵活地选择田野场地。研究人员在规划独立开展的田野实验时,需要考虑到实验实施的便利性,并灵活地选择田野场地。研究人员不一定要选择校园外的消费环境开展实验,很多时候在校园内甚至课堂上也可以开展田野实验。例如,Grewal和Townsend(2018)利用课堂情境操纵课程内容的持续时间,研究不同暴露时长下被试对刺激强度不同的产品(颜色强度不同的姓名牌)的选择。此外,线上环境也可以充当田野。Goldstein等(2014)开展的广告研究就通过被试完成任务时网页面显示的文字操纵了广告类型(没有广告vs.好广告vs.坏广告),他们通过MTurk网站开展了田野实验。

2.通过人员和产品等凸显实验情境的真实性。研究者需要通过让研究助手扮演成真实购物环境下的角色(如导购员、教堂人员等)来增强实验的真实性(Mogilner和Aaker,2009;Atasoy和Morewedge,2018)。同时,还需要结合节日背景(如在加拿大将士阵亡纪念日开展募捐实验)(Simpson等,2018)和产品背景(如实验材料选择刚刚获奖的奶酪)(Sevilla等,2018)来布置实验场地。此外,在这类独立开展的田野实验中,研究人员需要向消费者提供真实在售的产品,以及相应的收款设备。比如,Garbinsky和Gladstone(2019)在大学篮球场入口处开展的田野实验就在桌子上提供了两类产品(享乐特征明显的大学足球日历vs.实用特征明显的大学校历),让消费者用1美元购买。

3.被试数量、收集数据的时间等可以灵活考虑。已有的独立开展的田野实验,被试数量基本上都没有合作开展的田野实验大,所花的时间也没有后者多。比如,Huang和Susan(2015)开展的分享行为田野实验就采取重复测量设计,观察和记录了49名消费者在7天内的步行分享行为。而Garbinsky和Gladstone(2019)在大学篮球场开展的田野实验也仅采集了68名消费者在当天8小时之内的享乐品和实用品选择数据。当然,一些独立开展的田野实验的被试数量也可能很大,如Goldstein等(2014)就记录了1233名消费者在看到恼人广告(坏广告)时的任务表现。上述数据表明,在规划独立开展的田野实验时,研究者在规划被试数量和数据收集期限时不需要有严格的人数或时间限制,可以根据实地情况灵活处理。

4.考虑通过研究助理记录被试的人口学信息,而不是直接询问。田野实验要求被试觉察不到自己正卷入一项研究,而研究人员或研究助理通过问卷测项询问消费者的人口学信息会让消费者觉得自己正卷入一项研究,从而损失田野实验的自然性。因此,一些田野实验(包括独立开展的田野实验)会让研究助理记录被试的性别,并估计他们的年龄、种族等人口学特征。比如,Savary等(2015)和Madzharov等(2015)就采用了这种方法。

^①划分独立开展还是合作开展的依据是,研究者是否在实验介绍中提及自己在和一家非学术机构合作。

(二)需要合作开展的田野实验的开展建议

1.充分洞察非学术合作方对研究的需求,并把他们的需求考虑进实验设计。比如,Wadhwa等(2019)选择将防晒喷雾广告作为实验材料也是因为防晒喷雾是合作零售商有兴趣紧急销售的产品,并且已经计划过夏季特卖。如果非学术合作方是公共部门,研究人员就需要在设计实验方案时考虑公共部门在制定政策时的关注点。例如,Grinstein和Kronrod(2015)和美国一家管理部门合作,通过改变沟通方式更有效地呼吁大学生使用洗手液,因为鼓励学生使用洗手液会降低疾病发生率,这也是公共部门的关注点。

2.对一些非学术合作方做匿名处理。一些非学术合作方出于自己以及消费者隐私的考虑,会要求研究者对其做匿名处理。比如,Fong等(2015)在文章中就明确提到和他们合作开展田野实验的移动服务商要求对其做匿名处理。类似的,Hoban和Bucklin(2015)的田野实验也明确提出他们和合作公司之间签署了研究协议,根据这一协议,他们无法共享公司名称以及确切的业务细节。Lara等(2017)的田野实验也写到,出于保密原因他们没有提供零售商名称。其他一些田野实验尽管没有明确提到非学术合作方是否存在匿名要求,但都仅仅很笼统地介绍了非学术合作方所在的位置(如南美一家通信公司)(Eva等,2016)。

3.需要在撰写研究结论时明确写出实验给非学术合作方带来的实际效果。比如,Eva等(2016)的文章就明确写出,积极主动地鼓励客户改变计划会令人惊讶地提高而不是降低客户流失率,因为在干预后的三个月内,对照组只有6%的客户流失率,而实验组却有10%的客户流失率。Sterling等(2017)在文章中注明,在调查结束一年后,那些接受积极的开放式反馈调查的顾客比那些没有接受的顾客多消费8.25%。Guha等(2018)的研究也注明,与“现在降低Y%”的折扣框架相比,在财务上等价的“原来高X%”的折扣框架提高了产品销量。Sahni等(2016)在文章中提到,通过实验操纵,14.6%的用户在4周内再次访问了网站。而Atlas和Bartels(2018)发现,定期定价的价格框架增加了购买量。在这项研究之后,其合作伙伴——美国一家熟食配送公司采取了定期定价策略。综上,撰写实际田野实验效果,一方面能够说明研究理论自身的强健性,另一方面也能为非学术合作方日后推广实验中采用的操纵方案提供科学依据。

4.大样本量更被普遍采用。合作开展的田野实验往往由非学术合作方提供真实的田野环境,研究者无须自己准备实验场地。然而,正是由于这类实验在真实的消费环境下开展,研究者无法像独立开展田野实验那样准备实验场地,因此干扰因素和噪音会更大,自变量和因变量间的因果关系更不易得出。因此,这类实验的开展者通常会选取大样本量来让研究结果变得稳定。这类实验的样本量普遍较大,如Sahni等(2016)的田野实验样本量为3 092,Ascarza(2016)的田野实验样本量为12 137,Fong等(2015)的田野实验样本量为86 294。

基于上述总结和归纳,我们认为独立开展的田野实验和合作开展的田野实验各有利弊。独立开展的田野实验如果情境设置不真实就很有可能使消费者觉察到自己正卷入一项调查,从而使实验结论的说服力降低。而合作开展的田野实验需要研究者有很好的数据统计基础以及前期较为丰富的田野实验开展经验,不然会对大量的数据无法分析。因此,我们认为,独立开展的田野实验适合新手研究者,而合作开展的田野实验更适合资深研究者。此外,从这一部分前面提及的JCR、JCP、JMR、JM这四份期刊上两类田野实验的文章数量来看,JCR、JCP这两份期刊上刊登的独立开展田野实验的文章数量显著高于JMR和JM,而JMR和JM这两份期刊上刊登的合作开展田野实验的文章数量显著高于JCR和JCP。这就意味着,如果研究者独立开展田野实验,JCR和JCP相对更适合;相反,如果研究者和非学术合作方合作开展田野实验,则会更受JMR和JM青睐。

五、消费者行为领域的田野实验:并非必需但值得鼓励

尽管本文的目的是通过文献梳理归纳总结田野实验的开展原理,并呼吁更多的消费者行为学者开展这类实验,但是这并不意味着本文对田野实验持有100%的赞成态度。实质上,本文认为,学者们在决定是否要开展田野实验时需要充分考虑研究自身的特征以及大量的约束和限制因素。同时,从方法论的角度来看,田野实验也并不是万能、完美和唯一的,应交叉使用多种方法,做到研究方法为理论服务。

(一)并不是所有的消费者行为研究都需要开展田野实验

1. 田野实验适于可爱的研究而非漂亮的研究。根据Park(2012)发表在JCP上的编者案文章,有吸引力的研究可以分为两类:可爱的研究(cute study)和漂亮的研究(beautiful study)。在Park(2012)看来,可爱的研究自身关注有趣的现象,但理论不一定成熟和完备,而漂亮的研究在理论的完备性和创新性上都略胜一筹。由于田野实验在探测中介机制上有自身的缺陷(Cialdini, 2009; Gneezy, 2017),而完备的理论往往对中介机制的证明有要求,因此如果学者们做的是漂亮的研究,田野实验就不一定需要。但是,如果学者们做的是可爱的研究,田野实验则有助于复制实验室里得出的现象,从而增强现象自身的说服力。

2. 田野实验适于检验已经通过其他方法验证成功的理论而非提出一个全新的理论。如果一项研究的目标是要通过田野实验构建一个新理论,那么研究者就需要慎重考虑。由于田野实验在开展过程中无法随机分配被试、无法控制田野中的干扰因素和噪音、很难对实验过程进行人为干预(不然会增加被试的警觉)(Cialdini, 2009; Charness等, 2013; Gneezy, 2017),因此构建新理论的研究即便使用了田野实验,也不能让读者信服新理论中自变量对因变量的贡献度。

3. 田野实验适于行为指标作为因变量的研究而非认知、态度等作为因变量的研究。通过表1对田野实验定义的归纳可以看出,学者们对田野实验的因变量应该涉及实际行为这一点存在共识。因此,如果一项研究关注的是消费者的信息加工、态度、偏好、信念等非行为指标,那么田野实验并不适合,因为信息加工、态度、偏好、信念等变量常常需要借助一些自陈报告来测量,而这种测量方法本身会引发被试对自己正在参与一项研究的警觉,从而降低田野实验自身的自然性。如果研究者需要把实验室中得出的结论在田野中复制,那么在田野中开展的真实实验更加适于测量非行为类因变量。就行为本身而言,从第三部分的梳理可以看出,线上点击行为和网站访问行为、信息分享与采集行为、食物摄入和选择行为、支付行为、捐赠行为、风险寻求行为等非常适合田野实验。

4. 田野实验适于自变量是相对外显、容易操纵的营销刺激的研究情境。从第三部分的文献梳理可以看出,和产品、促销、价格、沟通、支付等外显的营销刺激相关的自变量非常适合田野实验。同时,空间感知、偶发情绪、感官体验等容易在自然环境下操纵的自变量也非常适合田野实验。相比之下,消费者的信息加工方式、个体特质、思维方式、认知负荷、内在需求和价值观等需要借助量表和选项来测量的自变量则不适合田野实验。

5. 田野实验适于发生在消费情境中且容易探究的研究主题。从第三部分和第四部分的梳理可以看出,本身发生在消费情境中且容易探究的研究主题(如店铺中的食物摄入与选择、捐赠等亲社会行为、支付、购买、点击与访问、广告、包装、价格、促销、传播沟通等)非常适于在田野实验中探究。而在消费情境之外也会发生的研究主题(如思维方式、睡眠、社会排斥、灾难暴露、权力、价值观等)则更适合用实验室实验或问卷调查、访谈等方法探究。在实际消费情境中通过田野实验探究这些研究主题,可能会让消费者产生被研究和被观测的警觉(因为这些主题

的研究需要借助量表和工具)。

(二)并不是在所有情形下都需要开展田野实验

1. 当研究者缺乏田野实验的开展经验时,不一定需要立即开展复杂的田野实验。田野实验尤其是需要和非学术合作方合作开展的田野实验需要研究者充分准备研究文案,并巧妙设计实验和分析数据,因此对研究者的研究能力要求更高。所以,当研究者自身的研究专长是实验室实验时,建议先做独立开展的田野实验,之后再做和非学术合作方合作开展的田野实验,而不是直接从实验室实验进入合作开展的田野实验。这样一方面可以降低失败率,另一方面也可以积累田野实验的开展经验。

2. 当研究者面临时间、财力等方面的约束时,不一定需要开展田野实验。尽管一些田野实验开展起来花费的时间较少,如Deng等(2016)的洗手液实验仅花了1天时间,但是通过前文第三部分的田野实验文章梳理我们发现,多数田野实验尤其是合作开展的田野实验相对比较耗时。比如,Tu和Soman(2015)和印度的一个农村社区合作开展的田野实验光收集数据就花费了2个月时间,Petersen和Kumar(2015)和一家公司合作开展的田野实验花费了6个月时间收集数据,Mazar等(2018)和一家金融公司合作的田野实验花费了9个月收集数据,Peter等(2015)和一家大型商场以及商场内的38家零售商合作开展的优惠券兑换田野实验则花了2年时间收集数据。除耗时外,田野实验需要准备实验场地、购买真实的实验材料,这一过程会耗费一些财力,如果研究者存在经费约束,也不一定需要开展田野实验。

3. 当用其他研究方法也可以复制实验室实验的结论时,也不一定需要开展田野实验。研究方法作为工具服务于所提理论,在实验室实验检验了主效应和中介机制后,研究者也可以通过其他定性研究方法(如深度访谈、问卷调查等)和定量研究方法(如二手数据等)为主效应提供相关性证据。从这个角度而言,田野实验也并非一定要开展。

(三)田野实验并不是万能、唯一和完美的,但值得鼓励

作为一种具体的研究方法,田野实验在增强外部生态效度的同时却由于噪音的存在而降低了内部效度,因此它并不是唯一、万能和完美的。正如McGrath(1981)所说,研究者对研究方法的选择本身就是一种三角困境,因为每类方法都有自身的优缺点。McGrath(1981)明确提出,研究者需要像投票者一样从缺点少的方法中进行选择。由于田野实验相比实验室实验缺点更少一些(被试对研究的觉察程度更低、研究自身的外部效度更高),因此这种方法值得鼓励。但是在梳理本文第三部分涉及的田野实验文章时我们发现,这些文章基本上都只把田野实验当作其中的一项小研究,同时还搭配了其他研究方法,尤其是实验室实验。绝大多数文章会先开展几项实验室实验,再开展一项田野实验。学者们用田野实验复制主效应,用实验室实验检验中介效应。除实验室实验外,田野实验也可以和二手数据法、专家组座谈法、民族志法等方法搭配使用,以增强研究结论在方法层面的可复制性。针对田野实验无法探测中介机制这一缺点,一些研究者在开展田野实验时尝试通过调节实验来间接反映中介,如Baca-Motes等(2013)发表在JCR上的文章为了探究酒店住客的环保承诺与其节能省电等期望行为间的关系是否通过身份彰显的中介作用而实现,就直接通过是否佩戴环保胸针来操纵身份彰显,对这种方法感兴趣的学者也可以借鉴。

六、结束语

本文针对前人文献中存在的研究空缺,对田野实验的内涵和判别标准进行了界定,对田野实验中的变量操纵进行了介绍,并就不同类型的田野实验该如何开展以及田野实验自身的适

用情形提出了建议。尽管如此,本文依旧存在不足:(1)本文作者不是方法研究者,因此没有对田野实验这一方法进行模型和技术上的推进;(2)本文第三和第四部分侧重于介绍和梳理前人的田野实验是如何开展的,有一定的罗列性,尽管其他部分有我们自己的观点和洞察,但是文章整体理论性相对不高;(3)本文仅梳理了JCR、JMR、JM、JCP这四份期刊近5年的田野实验文章,而没有对这四份期刊近40年的田野实验文章进行梳理,同时也没有关注其他消费者行为期刊。感兴趣的学者可以在本文的基础上对本文的这三点不足进行解决和推进。整体而言,本文对于学者们构思和开展田野实验具有操作层面的借鉴价值。

主要参考文献

- [1]樊亚凤,蒋晶,崔稳权.网络公益平台默认选项设置对个人捐赠意愿的影响及作用机制[J].心理学报,2019,(4):415-427.
- [2]冯小亮,任巍.社会影响易感性和虚拟产品采用:用户等级和产品类型的调节作用[J].南开管理评论,2018,(5):73-82.
- [3]郭秀艳.实验心理学[M].北京:人民教育出版社,2004.
- [4]黄敏学,王薇.移动购物更快吗?决策场景与思维模式的相容性[J].心理学报,2019,(5):612-624.
- [5]贾旭东,谭新辉.经典扎根理论及其精神对中国管理研究的现实价值[J].管理学报,2010,(5):656-665.
- [6]李杨.干预策略对公众旧衣回收行为的启动机制研究[J].营销科学学报,2017,(3):113-127.
- [7]尼尔·J·萨尔金德(著),童定(译).心理学研究方法/心理学译丛[M].第9版.北京:中国人民大学出版社,2019.
- [8]时勘.现场研究中准实验设计的几个问题[J].心理科学进展,1990,(2):20.
- [9]小卡尔·麦克丹尼尔,罗杰·盖茨(著),李桂华(译).当代市场调研[M].第10版.北京:机械工业出版社,2019.
- [10]Allen A M, Eilert M, Pelozo J. How deviations from performance norms impact charitable donations[J]. *Journal of Marketing Research*, 2018, 55(2): 277-290.
- [11]Argo J J, Dahl D W, Morales A C. Consumer contamination: How consumers react to products touched by others[J]. *Journal of Marketing*, 2006, 70(2): 81-94.
- [12]Argo J J, Dahl D W, Morales A C. Positive consumer contagion: Responses to attractive others in a retail context[J]. *Journal of Marketing Research*, 2008, 45(6): 690-701.
- [13]Atasoy O, Morewedge C K. Digital goods are valued less than physical goods[J]. *Journal of Consumer Research*, 2018, 44(6): 1343-1357.
- [14]Atlas S A, Bartels D M. Periodic pricing and perceived contract benefits[J]. *Journal of Consumer Research*, 2018, 45(2): 350-364.
- [15]Baca-Motes K, Brown A, Gneezy A, et al. Commitment and behavior change: Evidence from the field[J]. *Journal of Consumer Research*, 2013, 39(5): 1070-1084.
- [16]Barasch A, Zauberan G, Diehl K. How the intention to share can undermine enjoyment: Photo-taking goals and evaluation of experiences[J]. *Journal of Consumer Research*, 2018, 44(6): 1220-1237.
- [17]Biswas D, Szocs C. The smell of healthy choices: Cross-modal sensory compensation effects of ambient scent on food purchases[J]. *Journal of Marketing Research*, 2019, 56(1): 123-141.
- [18]Castro I A, Morales A C, Nowlis S M. The influence of disorganized shelf displays and limited product quantity on consumer purchase[J]. *Journal of Marketing*, 2013, 77(4): 118-133.
- [19]Chang H H, Hung I W. Mirror, mirror on the retail wall: Self-focused attention promotes reliance on feelings in consumer decisions[J]. *Journal of Marketing Research*, 2018, 55(4): 586-599.
- [20]Charness G, Gneezy U, Kuhn M A. Experimental methods: Extra-laboratory experiments-extending the reach of experimental economics[J]. *Journal of Economic Behavior & Organization*, 2013, 91: 93-100.
- [21]Fajardo T M, Townsend C, Bolander W. Toward an optimal donation solicitation: Evidence from the field of the differential influence of donor-related and organization-related information on donation choice and amount[J]. *Journal of Marketing*, 2018, 82(2): 142-152.
- [22]Fiske S T. How to publish rigorous experiments in the 21st century[J]. *Journal of Experimental Social Psychology*, 2016, 66:

145-147.

- [23]Gneezy A. Field experimentation in marketing research[J]. *Journal of Marketing Research*, 2017, 54(1): 140-143.
- [24]Grewal L, Hmurovic J, Lamberton C, et al. The self-perception connection: Why consumers devalue unattractive produce[J]. *Journal of Marketing*, 2019, 83(1): 89-107.
- [25]Grinstein A, Kronrod A. Does sparing the rod spoil the child? How praising, scolding, and an assertive tone can encourage desired behaviors[J]. *Journal of Marketing Research*, 2015, 53(3): 433-441.
- [26]Guha A, Biswas A, Grewal D, et al. An empirical analysis of the joint effects of shoppers' goals and attribute display on shoppers' evaluations[J]. *Journal of Marketing*, 2018, 82(3): 142-156.
- [27]Hock S J, Bagchi R. The impact of crowding on calorie consumption[J]. *Journal of Consumer Research*, 2018, 44(5): 1123-1140.
- [28]Huang X, Huang Z Q, Wyer J R Jr. The influence of social crowding on brand attachment[J]. *Journal of Consumer Research*, 2018, 44(5): 1068-1084.
- [29]Huang X, Zhang M, Hui M K, et al. Warmth and conformity: The effects of ambient temperature on product preferences and financial decisions[J]. *Journal of Consumer Psychology*, 2014, 24(2): 241-250.
- [30]Jami A. Having control over and above situations: The influence of elevated viewpoints on risk taking[J]. *Journal of Marketing Research*, 2019, 56(2): 230-244.
- [31]Kim T, Barasz K, John L K. Why am I seeing this ad? The effect of ad transparency on ad effectiveness[J]. *Journal of Consumer Research*, 2019a, 45(5): 906-932.
- [32]Kim J C, Wadhwa M, Chattopadhyay A. When busy is less indulging: Impact of busy mindset on self-control behaviors[J]. *Journal of Consumer Research*, 2019b, 45(5): 933-952.
- [33]Knoeflerle K M, Paus V C, Vossen A. An upbeat crowd: Fast in-store music alleviates negative effects of high social density on customers' spending[J]. *Journal of Retailing*, 2017, 93(4): 541-549.
- [34]Lalwani A K, Wang J J. How do consumers' cultural backgrounds and values influence their coupon proneness? A multimethod investigation[J]. *Journal of Consumer Research*, 2019, 45(5): 1037-1050.
- [35]Laporte S, Briers B. Similarity as a double-edged sword: The positive and negative effects of showcasing similar previous winners on perceived likelihood of winning in sweepstakes[J]. *Journal of Consumer Research*, 2019, 45(6): 1331-1349.
- [36]Li X S, Chan K W, Kim S. Service with emoticons: How customers interpret employee use of emoticons in online service encounters[J]. *Journal of Consumer Research*, 2019, 45(5): 973-987.
- [37]Lowe M L, Loveland K E, Krishna A. Quiet disquiet: Anxiety and risk avoidance due to nonconscious auditory priming[J]. *Journal of Consumer Research*, 2019, 46(1): 159-179.
- [38]McGrath J E. Dilemmatics: The study of research choices and dilemmas[J]. *American Behavioral Scientist*, 1981, 25(2): 179-210.
- [39]Meyer R. Introduction to the journal of marketing research special section on field experiments[J]. *Journal of Marketing Research*, 2017, 54(1): 138-139.
- [40]Morales A C, Amir O, Lee L. Keeping it real in experimental research—Understanding when, where, and how to enhance realism and measure consumer behavior[J]. *Journal of Consumer Research*, 2017, 44(2): 465-476.
- [41]Noseworthy T J, Trudel R. Looks interesting, but what does it do? Evaluation of incongruent product form depends on positioning[J]. *Journal of Marketing Research*, 2011, 48(6): 1008-1019.
- [42]Otterbring T, Ringler C, Sirianni N J, et al. The abercrombie & fitch effect: The impact of physical dominance on male customers' status-signaling consumption[J]. *Journal of Marketing Research*, 2018, 55(1): 69-79.
- [43]Paharia N, Swaminathan V. Who is wary of user design? The role of power-distance beliefs in preference for user-designed products[J]. *Journal of Marketing*, 2019, 83(3): 91-107.
- [44]Puzakova M, Kwak H. Should anthropomorphized brands engage customers? The impact of social crowding on brand preferences[J]. *Journal of Marketing*, 2017, 81(6): 99-115.
- [45]Romero M, Craig A W, Kumar A. Mapping time: How the spatial representation of time influences intertemporal choices[J].

- Journal of Marketing Research*, 2019, 56(4): 620-636.
- [46]Sevilla J, Isaac M S, Bagchi R. Format neglect: How the use of numerical versus percentage rank claims influences consumer judgments[J]. *Journal of Marketing*, 2018, 82(6): 150-164.
- [47]Shen L X, Hsee C K, Talloen J H. The fun and function of uncertainty: Uncertain incentives reinforce repetition decisions[J]. *Journal of Consumer Research*, 2019, 46(1): 69-81.
- [48]Simester D. Field experiments in marketing[A]. Banerjee A V, Duflo E. Handbook of field experiments[C]. North Holland: Elsevier, 2017: 465-497.
- [49]Simpson B, White K, Laran J. When public recognition for charitable giving backfires: The role of independent self-construal[J]. *Journal of Consumer Research*, 2018, 44(6): 1257-1273.
- [50]VanEpps E M, Downs J S, Loewenstein G. Advance ordering for healthier eating? Field experiments on the relationship between the meal order-consumption time delay and meal content[J]. *Journal of Marketing Research*, 2016, 53(3): 369-380.
- [51]Wadhwa M, Kim J C, Chattopadhyay A, et al. Unexpected-framing effect: Impact of framing a product benefit as unexpected on product desire[J]. *Journal of Consumer Research*, 2019, 46(2): 223-245.
- [52]Wang H S, Noble C H, Dahl D W, et al. Successfully communicating a cocreated innovation[J]. *Journal of Marketing*, 2019, 83(4): 38-57.
- [53]Wang J, Hong J W, Zhou R R. How long did I wait? The effect of construal levels on consumers' wait duration judgments[J]. *Journal of Consumer Research*, 2018, 45(1): 169-184.
- [54]Winterich K P, Nenkov G Y, Gonzales G E. Knowing what it makes: How product transformation salience increases recycling[J]. *Journal of Marketing*, 2019, 83(4): 21-37.
- [55]Yang Z Y, Sun S J, Lalwani A K, et al. How does consumers' local or global identity influence price-perceived quality associations? The role of perceived quality variance[J]. *Journal of Marketing*, 2019, 83(3): 145-162.
- [56]Zwebner Y, Lee L, Goldenberg J. The temperature premium: Warm temperatures increase product valuation[J]. *Journal of Consumer Psychology*, 2013, 24(2): 251-259.

Field Experiment in Consumer Behavior Research: Definitions, Manipulation Methods, and Suggestions

Liu Wumei, Huang Heqing, Ye Furong

(School of Management, Lanzhou University, Lanzhou 730000, China)

Summary: Field experiment is an experimental research method corresponding to lab experiment. It has received continuous attention in recent years because of its high realism and external validity. However, past research did not reach consensus on the definition of field experiment, which puzzled the readers to understand the meaning of field experiment and to carry out a real field experiment. This research gap motivated the current paper.

First, we collect previous scholars' definitions of field experiment. Based on these past definitions, we put forward a relatively clear definition of and a criterion for field experiment. In the same part, we further distinguish the definitions of field experiment between those of some other related terms such as lab experiment, quasi experiment, field study and realistic experiment in-the-field. This part is followed by introducing two subcategories of field experiment, namely, independent field experiment and collaborative field experiment, and discussing the advantages and disadvantages of field experiment. We argue that although field experiment has disadvantages in uncovering the underlying mechanism, controlling noise, analyzing data, etc, it can improve the realism and external validity of a scholarly

work. Importantly, it can increase the publication likelihood of the scholarly work to some extent.

Next, we analyze a total of 108 articles which include at least one field experiment in four leading marketing journals, i.e. JCR, JMR, JM and JCP. The time span of these articles ranges from August 2014 to November 2019. Based on the summary of these 108 field experiment articles, we elaborate on the general methods of IV manipulation and DV observation. Specifically, we divide IVs of these 108 articles into several categories: individual characteristics, product stimuli, environmental situations, and other categories, etc. These IVs are mainly manipulated through natural display, actual recording and language highlighting. Regarding DVs, we find that the DVs of these 108 articles are generally observed by adopting the “pay-what-you-want” paradigm, recording participants’ actual choices, recording participants’ food intake quantity and product purchase quantity, counting website visits and clicks, and recording relevant behaviors such as information sharing and information collection, etc.

In the following part, we provide a few suggestions on how to conduct two different types of field experiment (independent field experiment and collaborative field experiment). For independent field experiment, researchers can consider designing experiment flexibly, placing extra emphasis on increasing the authenticity of experiment, and avoiding participants’ awareness. While for collaborative field experiment, researchers can consider using a large sample size, understanding the needs of non-academic partners, and bringing practical benefits to non-academic partners, etc. Sometimes researchers also need to meet the anonymity requirements of non-academic partners.

In the final part, we discuss when researchers should carry out field experiment and when they should not. We argue that field experiment is not necessary, perfect and omnipotent, so researchers should choose appropriate research methods according to their own research objectives, research characteristics and resource constraints. We hope this paper could provide reference value for domestic researchers who plan to carry out field experiment in the future.

Key words: field experiment; consumer behavior research; definitions; manipulation methods; suggestions

(责任编辑:王舒宁)