

• 研究方法(Research Method) •

管理心理研究中情景实验材料的使用现状与设计策略*

于璟婷 卫旭华 马艺恩

(兰州大学管理学院, 兰州 730000)

摘要 虽然情景实验是管理心理领域兼具内部效度和生态效度的一种研究设计,但仍缺少详细、规范的情景实验材料设计指南。基于此,本研究在介绍了情景实验内涵、类别、与情境启动法区别的基础上,编码分析了20本国内外管理心理领域期刊近5年的93项情景实验,从设计前、中、后三个阶段开发了情景实验材料的设计策略。其中,设计前阶段侧重能否采用情景实验;设计阶段侧重如何形成实验材料初稿,依次为确定情景数量、初步设计情景、选择情景媒介、规范情景材料和营造真实情景五个子阶段。设计后阶段侧重评价材料是否具有真实性和有效性。若未通过检验或存在扩展材料的需求,研究者应返回设计阶段不断修改或改写,直至使其符合标准。未来研究可以重视理论选择并借助关键事件技术、信息技术及可迭代的决策情景提升情景材料适切性、丰富性和真实性。

关键词 情景实验, 情景材料, 设计策略, 管理心理实验

分类号 B849: C93

1 引言

情景实验作为一种兼具内部效度和生态效度的前沿方法,近年来受到管理心理学者的关注(Ng et al., 2020)。其指研究者通过模拟现实的情景设计操纵自变量并控制干扰因素,以观测各情景对个体认知、情绪和行为的影响。虽然情景实验越来越多地应用于管理心理领域,但是传统的实验室实验法和调查法仍为主流(Freedman et al., 2018),究其原因可能是情景实验的材料设计提高了该方法的使用门槛。并且,即使采用情景实验的研究,其实验材料也大多存在不规范之处(Wason et al., 2002)。不规范的情景材料容易产生

框架效应、社会期许效应和虚假反应等效应威胁因素(Antonakis, 2017),这进一步印证了情景材料设计指南的必要性。

目前,已有研究从方法论层面定义情景实验的研究范式,然而其侧重情景实验的整体实施流程,鲜少讨论实验材料的设计细节(Aguinis & Bradley, 2014)。即使有研究试图填补上述文献盲点,也只局限于设计原则(Rungtusanatham et al., 2011; Wason et al., 2002),尚未全面系统地介绍情景材料的设计策略。因此,在管理心理领域,“如何设计情景实验的情景材料”仍是一个悬而未决的问题。由于缺少具体、详细、规范的情景设计指南,情景实验在管理心理领域面临着材料设计的难题、有效性的质疑和推广发展的困境。

鉴于此,本研究试图归纳和开发管理心理情景实验中的情景材料设计策略,为广大学者提供设计与规范实验材料的实用指南,从而提高情景实验的严谨性和有效性,促进其进一步推广与发展。本研究先阐释了情景实验的内涵、类型和与情境启动法的区别。其次,基于20本国内外管理

收稿日期: 2023-08-10

* 国家自然科学基金项目(71972093, 72372063); 甘肃省哲学社会科学规划项目(2023YB013); 兰州大学中央高校基本科研业务费专项资金项目(2023jbkyzx005); 教育部2023年产学合作协同育人项目(津发科技-工效学会“人因与工效学”项目)。

通信作者: 卫旭华, E-mail: weixuhuahao@163.com

心理权威期刊近 5 年发表的 93 项情景实验的编码数据,系统梳理了实验材料设计的关键环节和使用问题。再次,开发了情景材料设计策略。最后,展望了情景材料的设计前景。

2 情景实验

情景实验,又名情境实验、情景模拟实验或情境故事法(experimental vignette methodology、scenario experiment),是一种研究者通过模拟现实的情景操纵自变量和控制干扰因素,并将被试随机分配到不同实验组,以观测各情景对个体认知、情绪和行为影响的研究方法。作为核心要素之一,情景是系统整合人、事、物所形成的简短描述(Atzmüller & Steiner, 2010),其可以借助文本、视频等媒介生动地模拟现实,引发被试的真实反应(Hughes & Huby, 2002)。情景实验包括多种类别,详见表 1。按照研究目的分为内隐决策和外显决策情景实验(Aguinis & Bradley, 2014)。前者要求被试对包含不同属性水平的情景排序;后者要求被试做出明确的判断和选择。按照媒介形式分为单一感官媒介和多重感官媒介情景实验,前者只通过文本、图片等作用于被试的某一感官;后者借助视频、数字媒体等调动被试多重感官。按照实施地点分为传统实验室和线上情景实验,其分别借助实验室中实验设备或利用网络平台开展。按照被试视角分为第一视角和第三视角情景

实验,前者需要被试想象自己是情景中的主人公,后者需要被试以第三视角参与实验。

现有研究中存在情境启动法(contextual priming)与情景实验名词混用现象,这是因为两者均属于实验法,并且语境场景是应用方式之一。但是,基于定义,本文发现二者在以下三方面具有差异,详见表 2。一是方法特性,在情境启动法中,实验刺激先启动被试的某种心理过程,使其处于准备状态,此种准备状态会对后续无关任务产生无意识影响(张玥,辛自强,2016);情景实验中的被试无需经历上述准备状态,其通过代入直接对实验刺激反应。二是情景特点,情境启动法的材料无需模拟现实;相反,情景实验材料需以现实世界为摹本。三是操纵变量,情境启动法仅可以操纵被试的心理变量,如认知表征(Bargh & Chartrand, 2000);情景实验既可以操纵被试的心理变量,还能操纵环境、行为和关系类等变量,因此情景实验的应用范围更广。另外,方法异质性决定了技术的不同,情境启动法常用技术如词语搜索,而阅读材料、观看视频则多用于情景实验。

3 管理心理情景实验的情景材料设计现状

本部分以设计前、中、后三个阶段为基本框架,编码国内外管理心理领域的 20 本权威期刊

表 1 情景实验的类别

情景实验类别	含义	特点
研究目的(why)		
内隐决策情景实验	被试对包含不同属性水平的情景排序,目的是了解各属性的相对重要性	适用于了解隐式决策过程的影响
外显决策情景实验	被试需要做出明确的判断和选择	因变量明确,适用于评估明确的过程和结果
媒介形式(how)		
单一感官媒介情景实验	采用文本、图片和音频等媒介呈现情景材料	成本低、简单易行、单一感官
多重感官媒介情景实验	采用视频、数字和角色扮演等媒介呈现情景材料	成本高、设计流程复杂、多重感官
实施地点(where)		
传统实验室情景实验	在实验室中,借助专门的实验设备(如 VR)开展情景实验	媒介限制少、内部效度高、样本受限
线上情景实验	利用网络平台(如见数)招募被试并开展实验	成本低、外部效度被优化、媒介受限
被试视角(who)		
第一视角情景实验	被试想象自己是情景中的主人公	代入感强
第三视角情景实验	被试以第三视角参与实验	代入感低

表2 情景实验与情境启动的概念辨析

概念	定义	方法特性	情景特点	操纵变量	常用技术
情景实验法	研究者通过模拟现实的情景操纵自变量和控制干扰因素,并将被试随机分配到不同的实验组,以观测各情景对个体认知、情绪和行为影响	被试通过代入直接对实验刺激做出反应	模拟现实	被试的心理变量、环境、行为、关系等变量	阅读材料、观看视频等
情境启动法	研究者通过先前无关刺激激活了与自变量相关的心理过程,并由此对后续任务产生的无意识影响	实验刺激启动被试某种心理过程,使其处于准备状态,并由此对后续无关任务产生无意识影响	无需模拟现实	被试的心理变量	词语搜索、归纳任务等

共性:均属于实验法;语境场景是应用方式之一

2019年1月至2023年5月发表的情景实验类文献,以总结情景材料的设计现状和方法。

3.1 研究方法

3.1.1 文献搜索与筛选

本研究根据 PRISMA 申明(Moher et al., 2010)检索与筛选期刊文献,流程见图1。

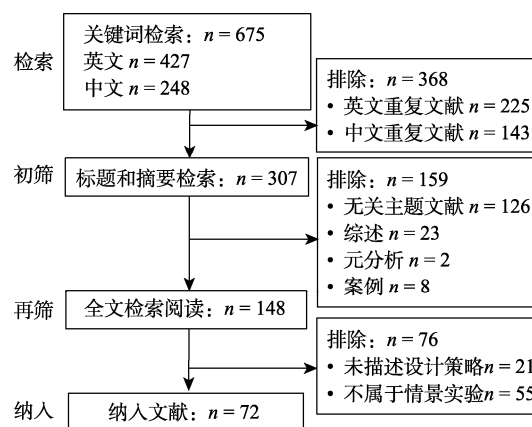


图1 PRISMA 流程图

检索。首先,中文检索基于中国知网、万方数据库和维普中文科技期刊数据库,英文检索基于 Web of Science 和 EBSCO 数据库。其次,中文检索词为“情景实验、情境实验、情景模拟、情境模拟、情境故事、角色扮演”;英文检索词为“vignette, paper people, scenario study or scenario experiment”。再次,鉴于 FT50 期刊上的情景实验材料设计规范、透明,因此本研究依据 FT50 学术期刊目录,选取管理心理领域的主流期刊,分别是 *Journal of Management*、*Journal of Applied Psychology*、*Academy of Management Journal*、*Journal of Business Ethics*、*Organizational Behavior*

and *Human Decision Processes*、*Organization Studies*、*Human Resource Management*、*Organization Science*、*Human Relations*、*Journal of International Business Studies*、*Journal of Management Studies*、*Administrative Science Quarterly*。另外, *Personnel Psychology*、*Journal of Organizational Behavior*、*Leadership Quarterly*、*Journal of Vocational Behavior* 也是该领域内有口皆碑的杂志,故一并纳入。中文期刊是《心理学报》《心理科学》《管理世界》和《南开管理评论》最后,鉴于 Aguinis 和 Bradley (2014)对情景实验的探讨是该领域经典的高被引方法论文。通过 Web of Science 分析其引文报告发现,其被引频次于 2019 年开始大幅上升。另外,研究目的决定了纳入分析的情景材料需具有规范性、完整性和透明性,而早期研究难以达到上述标准(Baguley et al., 2022),因此检索时间范围为 2019 年 1 月至 2023 年 5 月。

初筛。该阶段通过浏览标题和摘要进行初步筛选,所依标准为纳入管理心理领域文献,排除非实证研究。

再筛。该阶段通过阅读全文再次筛选初筛合格的文献,以排除未采用情景实验或未描述情景材料设计策略的文献。

纳入。最终纳入分析的文献为 72 篇(中文 7 篇,英文 65 篇),情景实验为 93 个。上述阶段由两名研究者独立完成,双方对有争议的文献进行再次阅读及分歧协商。

3.1.2 编码

我们从设计前阶段(pre-design stage)、设计阶段(design stage)、设计后阶段(post-design stage)分析了国内外管理心理情景实验中情景设计的现状。其中,设计前阶段侧重能否采用情景实验;设

设计阶段侧重如何形成情景材料初稿。基于其他领域研究(Matza et al., 2021)和情景实验的特征,本研究预设了5个子阶段:确定情景数量、初步设计情景、选择情景媒介、规范情景材料和营造真实情景;设计后阶段侧重情景材料初稿是否具有真实性和有效性。

3.1.3 编码的一致性信度

本研究由两位编码员共同编码,并对比二者编码的一致性来保证编码有效性。我们依据 Holsti (1969)的相互同意度及信度公式,求得编码一致性信度为0.978。对于不一致的编码,编码员解释编码原因,通过讨论达成一致。

3.2 情景材料的设计现状

3.2.1 情景材料设计前阶段的现状

本研究归类操纵变量发现,对于研究主题,30.55%、26.39%、19.44%文献的操纵变量分别涉及道德敏感、领导力、员工工作行为话题。另外,权力与地位、人际关系、工作家庭冲突等也均有出现。然而,团队管理相关操纵变量鲜少涉及(5.56%)。对于变量类型,52.78%、26.39%、12.5%分别涉及行为类、认知类和情境类操纵变量,对应示例如领导风格、组织认同和团队任务互依性。互动关系(如上下属关系)、个体特征(如残疾类型)、情绪状态(如真诚)等变量类型也有涉及。

3.2.2 情景材料设计阶段的现状

编码后发现,预设的5个子阶段基本涵盖了情景材料的必要设计工作。因此,探讨5个子阶段的现状有助于归纳情景材料的设计策略。

(1)情景数量确定的现状

基于样本实验发现,情景材料由背景材料和操纵材料构成。因此,情景数量取决于二者。结果显示,一是背景材料,在97.85%情景实验中,各版本情景材料的背景单一且相同。2.15%的情景实验具有两个及以上背景材料,其呈现形式为各实验组分别阅读不同背景的情景材料,即组间设计;或者各实验组情景材料的背景保持一致,但被试需依次阅读包含不同背景的材料,即组内设计。二是操纵材料,多因素情景实验常采用析因设计(factorial design),因此,操纵材料数量是操纵变量数量及水平的乘积函数。

(2)情景初步设计的现状

根据情景材料结构,研究者需进行相应的背景和操纵设计。自主设计、对现有背景材料沿用

或改编是常用的背景设计方法。操纵材料同样有三种设计方法:沿用现有操纵材料、改编现有操纵材料或相关测量量表、自主设计。在自主设计背景或操纵材料时,研究者可以通过访谈、实地研究、现实实践和官方报告等途径搜集素材,进而筛选、加工,形成初始情景材料。

(3)情景媒介选择的现状

媒介形式具有多样化,包括文本、图片、视频、角色扮演、音频、交互网站等,各占94.62%、9.68%、6.45%、4.30%、3.23%、1.08%。另外,20.43%的情景实验采用了混合媒介,其中,文本结合图片的比例最高。

(4)情景材料规范的现状

规范途径通常有6种:对称式设计、字数(时长)一致性、名称虚拟化、人物性别控制、措辞清晰准确和关键内容强调。其中,文本材料大多100~300字或英文单词。性别设定、中性化、随机分配和性别操纵为人物性别控制常用方式。然而,10.75%、27.96%、33.33%、34.41%、2.15%、58.06%的样本实验未利用上述途径标准化情景,这表明部分研究忽视了情景材料的规范工作。

(5)真实情景营造的现状

62.37%的情景实验未采取措施营造真实情景,这使其陷入“是否真实地反映现实世界”的质疑。本研究归纳了6种真实性增加的方式:取材自现实世界(54.29%)、采用沉浸式媒介(31.43%)、巧妙安排实验程序(20.00%)、呈现相关图片(8.57%)、使用真实名称(2.86%)和模仿媒体风格(2.86%)。另外,在努力营造真实情景的实验中,25.71%综合运用了以上多种方式。

3.2.3 情景材料设计后阶段的现状

(1)情景材料评价的现状

情景材料的评价内容涵盖真实性、内容效度和构念效度,评价方法分为定性法和定量法。其中,定性法或定量法评价真实性,各占总实验的3.23%、11.83%;定性法或定量法评价内容效度各占3.23%、2.15%;构念效度评价主要通过定量法,即操纵检验,占比91.40%,这既说明定量法的普遍应用,也说明研究者有待提升对真实性和内容效度的关注。另外,样本文献还存在两个评价问题。一是在正式实验开始前评价情景材料的实验仅17.20%,这可能导致由于设计不周,盲目开展实验造成的人财物浪费。二是8.60%的情景材料

未操纵检验。而在经过操纵检验的研究中,直接操纵检验、辨别式操纵检验和指导式操纵检验(Ejelöv & Luke, 2020)各占 85.88%、7.06%、47.06%,可以看出研究者只能确保激活了预期变量水平,无法消除产生无关变量的疑虑。

(2)情景材料扩展的现状

当评价结果较好,设计工作完成。不过,我们发现 16 篇文章开展多个情景实验,旨在通过差异化情景提高结论的外部效度。其操作化路径主要有三个角度:更换背景信息(58.82%)、更改操纵方式(35.29%)或更改操纵水平(11.76%)。35.29% 的情景实验对其综合运用,其中背景信息更换和操纵方式更改相结合最受研究者欢迎。

4 管理心理情景实验的情景材料设计策略

情景设计分为设计前阶段、设计阶段和设计后阶段,流程见图 2。情景材料设计的第一步是确定采用情景实验,即设计前阶段。其次,在设计阶段,研究者需要通过确定情景数量、初步设计情景、选择情景媒介、规范情景材料和营造真实情景五个子阶段,形成情景材料初稿。再次,在设计后阶段,初稿需要经过真实性、内容效度和构念

效度评价。倘若未通过检验或存在扩展情景材料的需求,研究者需要返回设计阶段,直至使其满足 5 项标准:符合研究主题、准确激发相应的变量水平、囊括必要信息、控制混淆因素、真实地反映现实情况。

4.1 设计前阶段

在设计情景实验的实验材料前,研究者应明确能否采用情景实验来开展研究。为回答该问题,本研究通过阐述两种情形以帮助研究者判断,即情景实验的适用情形和不适用情形。详情及实例,见表 3。

4.1.1 情景实验的适用情形

研究对象存在样本获取困难。当调查法或二手数据法所需的大量目标样本难以获取时,情景实验是不错选择。相比传统实验法,情景实验具有触及广泛主题、省时省力成本低等优点,加上其内在逻辑是设计模拟现实的情景以获得被试在类似现实情况下的反应(Aguinis & Bradley, 2014),这使得研究者无需寻找目标样本,只需通过媒介模拟与研究对象相关的现实场景,即可获得研究数据。

研究主题敏感。在研究敏感话题(如辱虐管理)时,传统实验法或调查法会使参与者面临潜在的

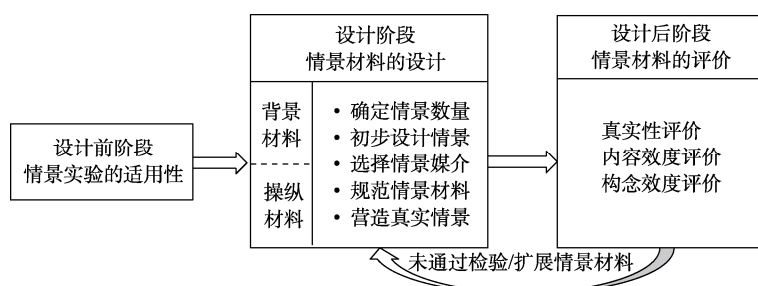


图2 情景材料设计流程

表3 情景材料设计前阶段

设计要点	设计策略及内容	研究举例
情景实验的适用情形	研究对象存在样本获取困难; 研究主题敏感; 操纵变量常为社会互动各方的相关变量或社会互动关系。前者包括被试的认知、权力、地位和行为等; 与被试社会互动的个体或群体, 其权力、地位、特质、认知、行为和人口统计学特征等。组织、团队和工作特征等互动情境变量也有涉及	Coyne 等(2019); Zhong 等(2023); Pieper 等(2019); 胡琼晶等(2021)
情景实验的不适用情形	核心变量均无法被有效操纵 虚拟情景无法充分反映现实世界	稳定的人格特质, 人口统计学特征, 认知风格, 价值观等; 并购, 系统动力学等

资料来源: 作者根据文献整理。

心理伤害,从而不仅参与者容易出现回避性作答,还常导致研究者陷入伦理指责。相比之下,情景实验借助虚拟情景引导被试想象的方式建立了被试与研究对象间的心理距离(Bradbury-Jones et al., 2014),这既不会伤害其心理,也使其无需基于个人经历作出反应,从而有效保护被试,为探讨敏感主题提供可行之法。

适合情景实验的操纵变量类型。社会互动各方的相关变量或社会互动关系是研究重点。前者包括两类,一是被试的认知、权力、地位和行为等;二是与被试进行社会互动的个体(同事、领导)或群体,其权力、地位、特质、认知、行为和人口统计学特征等。另外,组织、团队和工作特征等互动情境变量也有涉及。在操纵被试心理时,研究者通常操纵第三方虚拟人物的行为或外部环境,以激发被试的认知和情绪。在操纵被试行为时,研究者既可以通过主人公想象的方式直接操纵,也可以构建场景促使被试做出相关行为。其他类型的变量则可直接操纵。以员工建言变量为例,当探讨领导对员工建言的反应时,被试被赋予领导角色,操纵员工(被试社会互动方)的建言行为。当探讨员工建言对自身的影响时,员工建言为被试的行为变量,操纵方式为建言行为想象或促使实施。

需要说明,研究者在选择情景实验时,应持谨慎态度,这是因为该方法常陷入三点质疑:一是外部效度较低。尽管情景实验存在研究情境更换的可操作化路径,但其毕竟无法穷尽场景并且与现实世界存在鸿沟(Klebe et al., 2021)。在实际操作中,研究者可以结合调查等方法缓解上述不足。二是行为与倾向的差异。对于行为类因变量,情景实验测量的往往是倾向而非实际行为,这将导致研究结论的过度推断(Klebe et al., 2021)。恰当的媒介形式或许对此有所缓解。三是释义问题。因果推断会受到被试信息加工模式差异、个体经历、情景材料信息不足、特定信息元素的干扰(Matza et al., 2021)。

4.1.2 情景实验的不适用情形

鉴于情景实验既属于实验法,又具有真实性,因此本研究以此为出发点,探讨情景实验的不适用情形,以避免该方法的误用、滥用。一是核心变量均无法被有效操纵。例如稳定的被试特征变量(稳定的人格特质,人口统计学特征,认知风格,

价值观等)。被试代入虚拟人物的方式看似可行,但是,其可能具备与虚拟人物相反的特征,这将导致被试难以代入角色,预期变量水平无法被准确激活。二是虚拟情景无法充分反映现实世界。例如虚拟情景难以充分捕捉和模拟并购决策背后的动态复杂环境。此时,被试对虚拟情景和真实情景的反应存在较大差异,研究者何谈精确探究决策背后的心理机制。类似研究还如系统动力学。

4.2 设计阶段

情景材料由背景材料和操纵材料构成。背景是承载情景产生、发展的舞台,传递着人物特征、人物关系、时间地点、工作特点、操纵铺垫等必要信息,并控制混淆因素(Rungtusanatham et al., 2011)。操纵材料描绘了操纵变量不同水平对应的现实情况。例如,凌文铨等(2019)的研究依托于组织场景,故背景中交代了地点(公司)、岗位(管理人员)、人物名称及人物关系(小王是直接下属)。鉴于自变量上下属关系分为亲近和疏远两水平,操纵材料则分别对应描述了小王与“我”的关系情况。

设计阶段主要是通过5个子阶段,设计背景材料和操纵材料,形成情景材料初稿,研究实例见表4。

4.2.1 确定情景数量

“设计哪些情景”是设计阶段的首要问题。虽然没有唯一答案,但却可以抽象化为背景材料数量和操纵材料数量的乘积函数。

在设计背景材料时,研究者需要选择单背景或多背景。单背景是各版本情景共享同一背景,两个及以上背景材料则是多背景的特征。相比单背景,多背景的外部效度较高。因此,为增强结论的外部效度,研究者可以设计多个背景材料。多背景采用组间设计或组内设计,鉴于在前者中,除了操纵材料外,各则材料的背景信息也不一致,无法排除替代性解释,因此后者更优。

“设计哪些操纵材料”可以引导研究者明确操纵变量及水平,原因是操纵材料数量往往是操纵变量数量及水平的乘积函数(Atzmüller & Steiner, 2010)。需要注意,操纵变量及水平的确定应由理论驱动(Aguinis & Bradley, 2014),必要时可操纵控制变量以排除混淆因素的干扰。

4.2.2 初步设计情景

初步设计情景是设计阶段的中心环节。研究者通常可以遵循两种思路:沿用和重新设计。后者包括自上而下法和自下而上法,方法选择流程见图3。

表 4 情景材料设计阶段

设计步骤	设计策略及内容	研究举例
确定情景数量	背景材料数量：多背景(组间、组内设计)和单背景	Cook 和 Kuhn (2021); Coyne 等(2019)
	选择依据：为提高外部效度选择多背景	
	操纵材料数量为操纵变量数量及水平的乘积函数	Coyne 等(2019)
初步设计情景	沿用：沿用现有材料，使用应思考文化、概念、翻译方面是否存在局限	Ciampa 等(2021)
	自上而下：基于现有文献进行设计	Zhong 等(2023); Hassan 等(2021); 凌文铨等(2019); Hurst 等(2019); 蒋旭婷等(2023)
	背景：改编现有材料；基于操纵变量或因变量的概念背景、影响因素及理论；	
	操纵：改编现有材料；改写量表；基于操纵变量的定义、理论及研究	Pieper 等(2019); Cook 和 Kuhn (2021) 胡琼晶等(2021)
	自下而上：现有文献无法支撑设计	
	操纵：广泛搜集原始典型素材(访谈、报告、其他研究)；分类与筛选(基于操纵变量水平)；形成操纵材料	
	背景：广泛搜集原始典型素材(操纵材料素材搜集渠道，实践、被试信息、网络等)；筛选(标准：必要性、可信度、相关性等)；形成背景材料	
	自上而下法和自下而上法可以相互结合	
选择情景媒介	文本；图片；音频；视频；交互网站；角色扮演；混合媒介等 选择依据：研究主题、权衡逼真度和成本	蒋旭婷等(2023); Amarnani 等(2022); Smallfield 等 (2020); Greenwood 等 (2022); Liao 等(2023); Pieper 等(2019);
规范情景材料	设计：对称	凌文铨等(2019)
	单背景：操纵材料应句式、措辞、情节、结构和形式相似	
	多背景：背景和操纵材料应句式、措辞、情节、结构和形式相似	Zhong 等(2023)
	信息量：一致，且不超载	
	名称：虚拟	Cheng 等(2023)
	人物性别控制：性别设定、中性化、随机分配、性别操纵	Levine 和 Wald (2020); 胡琼晶等 (2021); Ritzenhöfer 等(2019)
	措辞：清晰准确	Liao 等(2023)
	关键背景内容和操纵重点：强调 文本：斜体、加粗、字号、字体颜色 音频：语调、语气	Pieper 等(2019)
	除上述通用原则外，还需根据理论确定混淆因素并予以干预	
营造真实情景	情景初步设计时，	Cook 和 Kuhn (2021); Paustian-Underdahl 等(2023)
	背景：取材自现实世界、巧妙安排实验程序	
	操纵：取材自访谈	胡琼晶等(2021)
	情景媒介选择时，视频、交互网站等媒介；背景中呈现相关场景图片	Smallfield 等(2020); Pieper 等(2019)
	倘若情景材料以网站、APP 界面等形式呈现，模仿相应写作文风、排版版式	Greenwood 等(2022)

资料来源：作者根据文献整理。

研究者首先应该判断可否沿用与操纵变量相关的现有情景材料。经学者论证的情景材料往往真实性和有效性较高。因此，沿用成熟的情景材料有利于确立实验的可信度。但是，在实际应用时，研究者需要考虑三个问题，一是文化适用性，国外的情景材料能否为本国样本理解；二是概念适用性，选用的情景材料能否全面、准确地激发相应的变量水平(Telwatte et al., 2017)。三是翻译准确性，由外语译成中文的情景材料是否掺杂混淆因素。另外，现有情景材料并非都通过了有效性和真实性检验。在沿用时，研究者应对未经充分论证的材料持谨慎态度。

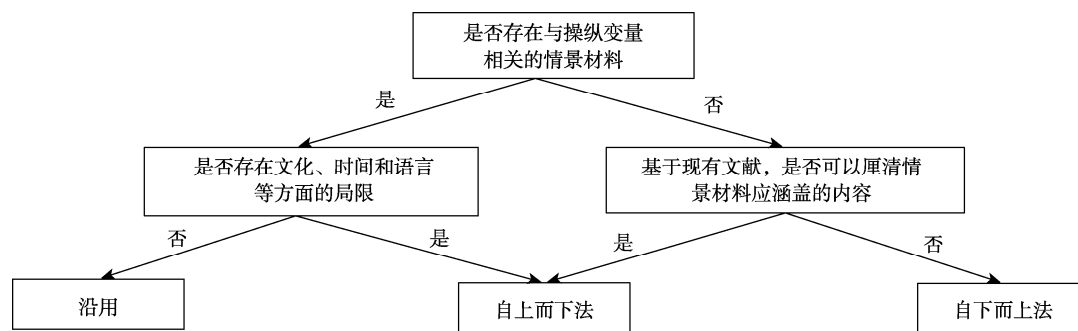


图3 初步设计情景材料的流程图

自上而下法是基于现有文献设计情景材料。对于背景而言,在梳理文献后,研究者能够明确有助于被试理解操纵材料并呈现完整故事所需的必要信息。在该前提下,研究者可以改编与操纵变量相关的现有背景材料,或基于操纵变量、因变量的概念背景、影响因素,发展背景材料(Hassan et al., 2021; Zhong et al., 2023)。Hassan 等(2021)在探究道德准则的沟通是否会降低不道德行为时,之所以选择公立医院的采购服务作为背景,是因为已有研究证实此时会出现不道德行为,如腐败。在实验操纵设计中,根据前人,研究者可以确认操纵材料应涵盖的内容。其主要有三种设计模式:改编现有操纵材料;改写测量量表;基于操纵变量的定义、理论及研究,发展操纵材料(凌文铨等, 2019; Hurst et al., 2019)。其中,模式三的操作化路径为,研究者需对某一操纵变量进行清晰的构念说明,确立其理论边界,厘清其与相关概念(如服务型领导和道德型领导)的共性和区别,必要时可利用理论和研究为操纵提供理论依据,从而撰写与操纵水平匹配的操纵材料。之后,根据析因设计,组合、整合多个变量不同水平的操纵材料,以使材料完整。Hurst 等(2019)在操纵辱虐管理时,既改写了相关量表,也借鉴了该概念所囊括特质(如自恋)的研究。

自下而上法应用于当现有文献无法支撑开展情景实验之时。对于操纵材料来说,一般分三步:(1)广泛搜集原始素材。研究者可以借助三种途径尽可能地搜集符合操纵变量概念的典型素材,一是访谈素材,如焦点小组访谈、个人访谈、已有访谈资料等(胡琼晶等, 2021);二是官方报告(Rice et al., 2023);三是其他子研究所提供的经验素材(Pieper et al., 2019)。(2)分类与筛选。研究者

既可以利用专家小组量化评定或定性讨论,也可通过问卷调查,将原始素材根据操纵变量水平分类,从而挑选出符合实验操纵要求的素材。(3)形成。整合合格素材,形成初始的操纵材料。背景材料设计同样分三步:(1)广泛搜集原始素材。背景材料的原始素材来源与操纵材料基本一致,还包括现实实践、网络素材、被试信息等(Becker et al., 2023; Cook & Kuhn, 2021; Liu et al., 2021)。(2)筛选。通过专家评定、调查等方式,以必要性、可信度、相关性为判断标准,信息全、代入强、相关高的背景可以被择出。(3)形成。在整合背景材料和操纵材料时,可以适度修改,以使语句通顺。例如 Cook 和 Kuhn (2021)首先在网络平台搜集因下班后的社交媒体行为被解雇的案例;其次通过问卷调查,对案例的道德违反情况(相关性)予以评分;最后,当案例被 60%以上的参与者一致认为涉及道德违反行为时,即被用于初始背景材料。另外,自上而下法和自下而上法并非对立关系,综合运用反而有助于提高材料的有效性和完整性。胡琼晶等(2021)综合两种方法设计情景材料。

4.2.3 选择情景媒介

媒介是可视化情景材料的工具。若想搭建起被试和情景材料间的桥梁,研究者需要从以下 6 种媒介形式中选择一种或多种媒介:文本、图片、音频、视频、交互网站、角色扮演。

上述 6 种媒介所营造的情景逼真度依次递增,逼真度越高,被试不仅沉浸感增强(Lonati et al., 2018),更大程度地参与实验,而且研究的生态效度也有所提升,但同时这也面临着引入混淆因素和提高成本的风险(Eckerd et al., 2021)。因此,如何选择恰当的情景媒介是困扰研究者的难题。本

研究认为可以从两个角度思考: 研究主题、逼真度与成本的平衡。对于某研究主题而言, 不同媒介的有效性存在差别。比如在道德研究中, 文本材料会引起道德判断, 虚拟现实技术则会导致道德行为(Terbeck et al., 2021)。因此, 研究者应根据研究主题, 选择恰当媒介, 从而减少对真实反应与情景实验结果存在差异的质疑。此外, 情景逼真度的获得是以高成本为代价的, 如雇佣演员、购买设备。研究者虽然可以通过故事板、租用设备等方式降低成本或直接选用使用率最高、成本最低的文本, 但也应根据实际情况在逼真度和成本之间评估和取舍。

4.2.4 规范情景材料

规范情景材料有助于控制混淆因素, 更准确地估计核心研究变量间的关系, 是设计阶段的重要环节。研究者需要综合运用通用原则与特殊情况下的有效准则来指导规范工作。

虽然情景材料各异, 但有6条通用规范原则。一是对称式设计。对于单背景情景材料来说, 各组材料的背景保持一致, 操纵材料存在差异。此时, 操纵材料应满足句式、措辞、情节、结构和形式相似等要求。对于多背景情景材料来说, 变化的背景和操纵材料均需满足上述要求。然而存在研究未满足标准。Coyne等(2019)的情景材料无论是句式结构、关键词强调(“urgent”大写与否)还是情节设定(仅有一则材料包含同事要求延长任务截至日期的情节)均有较大出入。非对称设计易使被试表现出框架效应(Wason et al., 2002)。二是信息量一致性。各版本的情景材料应信息量相似, 体现在视频时长、文本字数和内容难度等方面, 否则会引入阅读时长、被试理解差异等干扰因素。然而信息量一致并不意味着信息量超载。虽然无关信息有助于缓解需求效应, 但是应在提供必要信息的前提下(Aiman-Smith et al., 2002)。多数研究将单个情景的文本篇幅限制在100~300字或英文单词, 其中背景材料或操纵材料字数分别在50~150或50~200字(英文单词)不等。同样, 信息量一致并不代表信息趋同, 不同水平变量的操纵材料应差异化。三是名称虚拟化。虚构的人名、地名、公司名有助于防止真实名称的干扰。四是人物性别控制。研究者可以通过性别设定、中性化、随机分配、性别操纵排除性别影响。Ritzenhöfer等(2019)发现男性和女性的情绪表达

存在差异, 因此通过姓氏和人称代词将虚拟人物的性别平分并随机分配。五是措辞清晰准确。情景材料需要避免行话、修辞、歧义表述、不确定性用词和带有感情色彩的俗语, 力求文字通畅、清晰易懂。六是关键内容强调。为使被试注意到关键背景内容或操纵重点, 研究者可以通过斜体、加粗、字号和字体颜色等强调文本材料, 通过语气、语调等强调音频材料。但是过多强调也可能造成注意分散(Matza et al., 2021)。另外, 多数情景材料内含情景人物与被试同龄的假设, 但却往往被被试忽略(Grol-Prokopczyk, 2014)。因此, 人物年龄明确有助于获得更好的反应一致性。

除通用原则外, 特殊情况还需被考虑。当媒介为图片时, 图片大小、像素和位置应一致。当媒介为视频时, 演员相貌吸引力和背景复杂等因素都可能带来替代性解释。当代入第三方视角时, 客观的评价方式或实验任务或许可以降低社会期许效应。总之, 研究者需要根据理论确定混淆因素并予以干预(Bernerth & Aguinis, 2016)。

4.3.5 营造真实情景

真实性是情景实验的关键特征和方法优势(Jared, 2019)。因此, 真实情景的营造应该嵌入各个子阶段。

在初步设计中, 对于背景材料而言, 取材自现实世界、巧妙安排实验程序有助于构建被试熟悉的场景。例如, 真实感受的表达促使被试与情景产生更多交流(Paustian-Underdahl et al., 2023)。操纵材料真实感增强的常用设计手段则是从访谈中积累素材。在媒介选择中, 视频、交互网站等媒介可以为被试提供多重感官体验。并且在背景中增设相关场景的图片也会进一步增加可信度。在材料规范中, 倘若情景材料以网站、APP界面等方式呈现, 那么模仿相应写作文风、排版版式将有助于营造临场般的浏览氛围。另外, 真实名称、第一人称具有同样效果。

研究者需要在情景材料规范和真实性营造间进行权衡, 这是因为逼真度增强是以混淆因素引入为代价的(Eckerd et al., 2021)。为解决该难题, 研究者应基于理论, 控制真实情景营造所带来的混淆因素。对于难控制的因素, 研究者可根据预测测试真实性评价的结果进行取舍, 也可设计不同版本的情景材料开展多个情景实验。例如真实公司名可以提高情景的逼真度, 但同时也使被试遭

受先验知识的干扰。针对是否使用真实公司名的疑惑,研究者可以评估未使用真实公司名的情景材料是否已然具备可信性,还可以运用真实公司名和虚拟公司名的情景材料分别开展情景实验。

4.3 设计后阶段

设计阶段形成的情景材料无法直接应用于正式实验,这是因为该材料是否真实、清晰和完整,以及能否准确激活预期变量水平,我们仍不得而知。因此,设计后阶段的主要任务是评价情景材料,并根据评价结果修改完善,直至满足标准。另外,研究者还需考虑如何扩展实验的情景材料来提高结论的外部效度。相关研究见表 5。

4.3.1 评价情景材料

研究者应进行真实性、内容效度和构念效度评价,以确保情景材料初稿不仅具备真实性、清晰性和完整性,也可以成功实现实验操纵。

真实性评价指采用定性法或定量法审查情景材料的真实性。在定性方法中,研究者需要招募不一定熟悉研究问题的审查小组进行评价。评价问题如“是否熟悉所描述的情景?”(Rungtusanatham et al., 2011)。研究者还可以通过预测试从定量角度进行评价。例如, Pilling 等(1994)开发了 4 题项量表。

内容效度评价指通过定性法或定量法评估情景材料的清晰性、相关性和完整性。定性法的操作化路径是,专家小组判断情景材料是否存在混淆或矛盾的术语、短语、句子或段落;操纵是否符合特定构念;是否拥有形成判断、做出决定或推荐解决方案的必需信息等。此外,学者们也讨

论了定量方法。一是内容效度指数(content validity index, CVI, Marie et al., 2021)。首先,专家需要分别评估情景的清晰性、相关性和重要性,其中“1”代表不清晰、不相关和不重要,“3”则相反。其次,计算 CVI。将“每个维度得分为 3 分的专家人数”除以“专家总人数”得出每个维度的 CVI,进而通过三个维度 CVI 的均数得到总 CVI。最后,判断 CVI。当 CVI 大于 0.8 时,情景材料通过内容效度检验。二是在预测试中运用量表评估用词的清晰性和适当性。需要注意,多数研究将真实性、清晰性和完整性作为所设计情景的暗含假设,忽视了审查与评价,这将会导致肤浅反应和认知偏差(Harrits & Møller, 2021),因此有必要将真实性评价和内容效度评价作为独立步骤,予以重视。

为检验实验操纵的有效性,研究者可以在预测试中通过操纵检验评价情景材料的构念效度。研究者应综合运用三种操纵检验形式,首先,直接操纵检验判断操纵是否激活预期构念。其次,辨别式操纵检验排除竞争解释。最后,使用具有客观答案的指导式操纵检验排除不认真填写样本。

在评价结束后,研究者需要根据反馈意见或测试结果修改存在问题的情景材料,直至情景材料具备清晰性、真实性、完整性和有效性。另外,对于三种评价来说,定性法和定量法可以互为补充。

4.3.2 扩展情景材料

当评价结果较好时,情景材料即可用于研究。虽然真实情景营造有助于提高生态效度,但是其仅仅依托于某一情形,因此研究者仍需思考“是否有必要开展其它情景实验来提高结论的外

表 5 情景材料设计后阶段

设计步骤	设计策略及内容	研究举例
评价情景材料	真实性评价: 定性: 招募小组评价材料的熟悉度和代入度; 定量: 在预测试中运用量表	Becker 等(2023)
	内容效度评价: 定性: 专家小组判断材料的清晰性、相关性和完整性; 定量: CVI、在预测试中运用量表	Coyne 等(2019);
	构念效度评价: 在预测试中操纵检验(定量)	Zhong 等(2023)
扩展情景材料	设计新情景材料: 更改操纵变量水平或操纵方式	蒋旭婷等(2023) Levine 和 Wald (2020)
	改写情景材料: 故事情节更换、虚拟人物的工作类型或所属行业改换, 人物视角转换, 无关信息删除和相关变量进一步控制等	凌文铨等(2019)

资料来源: 作者根据文献整理。

部效度?”。如果答案为“是”,研究者需要扩展本实验的情景材料。

改写本研究材料是省时省力的常见选择。通常,研究者将视角聚焦于背景,通过故事情节更换、虚拟人物的工作类型或所属行业改换,人物视角转换,无关信息删除和相关变量进一步控制等途径扩展材料。比如凌文铨等(2019)既明确了虚拟人物的行业,也把管理人员重新设定为研发人员。鉴于此时操纵材料并未被修改,且已被证明有效,因此,研究者只需规范背景、增强和评价背景的真实性。若是涉及更改操纵变量水平或实验操纵方式,研究者就需重新遵循设计流程,发展新材料。

4.4 实例分析

编码文章中存在契合上述情景材料设计策略的研究(Amarnani et al., 2022; Bode et al., 2022),因此,为更好地说明和印证所提出的材料设计思路,本研究以 Amarnani 的研究为实例进行演示。该研究探讨顾客欺凌如何产生和缓解对员工顾客导向组织公民行为的负面影响。

在设计前阶段,一是研究目的为探究因果效应,二是自变量为他人行为类变量,属于情景实验可直接操纵变量的范畴,三是研究主题不涉及难以捕捉的动态复杂环境,故可以考虑采用情景实验。

在设计阶段,一是确定情景数量。对于背景而言,研究者选择常用的单背景。对于操纵而言,自变量为顾客欺凌,包括有、无两个水平。因此该实验采用单因素设计,共2则情景材料。二是初步设计情景。该研究以服务行业的员工为研究对象,故研究者利用自下而上法,通过访谈相关行业员工,形成包含地点(电信公司的客服中心)、工作内容、有助于理解操纵的背景材料。操纵材料的设计结合使用了自下而上法和自上而下法,即基于自变量的定义和访谈数据。三是选择情景媒介,鉴于研究场景是客服中心,因此研究者选用了混合媒介:文字、音频。四是规范情景材料,2则材料设计对称、时长相似(50 s左右),措辞清晰准确,同一配音演员,且使用了虚构的公司名(XYA Services Ltd),并利用语调、语气强调关键信息。五是营造真实情景,初步设计时,情景材料来源于现实访谈;媒介选择时,材料以顾客语音邮件形式呈现。

在设计后阶段,该研究在预测试中运用可信与否评价材料真实性,并运用操纵检验评价观念效度。随后,根据被试的反馈,材料措辞被进一步修改。

5 未来展望

5.1 推动关键事件技术与情景材料设计的有效结合

未来研究可以运用关键事件技术设计情景材料。关键事件技术指从个体处获取关键事件,从而探究关键事件对其情感、认知和行为的影响(刘周颖等, 2020)。因此,研究者若想要操纵被试的某种情绪、认知或行为,不妨在初步设计情景时,先要求其他样本详细回忆相关的关键事件,从而基于这些事件素材设计实验材料。

5.2 利用信息技术提高情景材料的真实性

未来研究可以借助信息技术营造真实情景。VR技术为情景材料的真实性增强和被试的沉浸式体验提供工具支撑。具体地,当变量与静态的物理空间有关,如办公场所私密性,研究者可以制作VR全景,呈现立体空间;当情景涉及对话时,研究者可以利用VR虚拟数字人技术模拟交流场景;当实验依托的背景不常见且需要交互时,如基于水害逃生场景探讨矿工工作时的压力,研究者可以借助VR穿戴设备。另外,当情景涉及线上交流时,研究者可以运用软件模拟互动过程。

5.3 设计可迭代决策的情景材料获取真实行为数据

在操纵被试行为时,已有情景材料通常仅涉及简单的文本段落,被试无法影响情景的发展进程,这将导致参与度低。具有迭代决策特征的情景材料可以优化上述问题,原因在于被试并非被随机分配到各实验组,而是基于自主选择。以道德许可研究为例,在传统情景实验中,被试只能被随机分配到有、无道德行为的情景材料中,而具有迭代决策特征的情景材料则允许被试自主选择是否实施道德行为。因此,在未来,研究者可以尝试设计可迭代决策的情景材料,以获取真实、丰富的行为过程数据。目前,已有交互式平台Twine,帮助开发此类材料(Freedman et al., 2018)。

5.4 重视情景实验材料设计的理论选择及适切性判断

未来研究应重视理论指导并判断适切性。一

是设计前阶段,研究者应基于研究目的、研究主题、研究资源和操纵变量性质,恰当选择情景实验。不过,情景实验常深陷外部效度低,情景材料不完整、不真实、结构化过度等质疑,因此应与其他研究方法搭配使用。二是设计阶段,操纵变量及水平应基于理论。在初步撰写操纵时,研究者应率先明晰所操纵构念的核心特征,厘清其与其他相似构念的理论边界,界定操纵层次(个体或团队),进而沿用、改编或自主设计操纵材料。另外,操纵材料的规范同样需要研究者根据理论确定混淆因素并予以干预。对于背景材料来说,外部效度提高与否决定了是否设置多背景。研究主题、操纵场景和实际情形则是初步设计背景材料的理论依据。在规范背景材料时,研究者也需要基于理论纳入人物性别、年龄等信息。三是设计后阶段,研究者应运用专家小组、内容效度指数(Marie et al., 2021)、操纵检验等方式判断所设计材料的真实性、清晰性和有效性。

参考文献

- 胡琼晶, 魏俊杰, 王露, 谢小云. (2021). 犯错者地位如何影响同事容错?——任务目标偏离度和团队互依性的作用. *管理世界*, 37(6), 113–1857.
- 蒋旭婷, 吴小玥, 范雪灵, 贺伟. (2023). 员工愤怒表达对领导力涌现的影响: 温暖和感知能力的中介作用以及愤怒道歉的弥补作用. *心理学报*, 55(5), 812–830.
- 凌文轻, 李锐, 聂婧, 李爱梅. (2019). 中国组织情境下上司-下属社会交换的互惠机制研究——基于对价理论的视角. *管理世界*, 35(5), 134–148.
- 刘周颖, 赵宇翔, 朱庆华. (2020). 付费知识问答平台中提问者转移行为的影响因素分析. *图书情报工作*, 64(12), 75–86.
- 张玥, 辛自强. (2016). 社会心理学中的启动研究: 范式与挑战. *心理科学进展*, 24(5), 844–854.
- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, 17(4), 351–371.
- Aiman-Smith, L., Scullen, S. E., & Barr, S. H. (2002). Conducting studies of decision making in organizational contexts: A tutorial for policy-capturing and other regression-based techniques. *Organizational Research Methods*, 5(4), 388–414.
- Amarnani, R. K., Restubog, S. L. D., Shao, R., Cheng, D. C., & Bordia, P. (2022). A self-verification perspective on customer mistreatment and customer-directed organizational citizenship behaviors. *Journal of Organizational Behavior*, 43(5), 912–931.
- Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, 28(1), 5–21.
- Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology*, 6(3), 128–138.
- Baguley, T., Dunham, G., & Steer, O. (2022). Statistical modelling of vignette data in psychology. *British Journal of Psychology*, 113(4), 1143–1163.
- Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). Cambridge: Cambridge University Press.
- Becker, L., Ertz, E., & Büttgen, M. (2023). A relational perspective on supervisor-initiated turnover: Implications for human resource management based on a multi-method investigation of leader-member exchange relationships. *Human Resource Management*, 62(4), 547–564.
- Bernerth, J. B., & Aguinis, H. (2016). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, 69(1), 229–283.
- Bode, C., Rogan, M., & Singh, J. (2022). Up to no good? Gender, social impact work, and employee promotions. *Administrative Science Quarterly*, 67(1), 82–130.
- Bradbury-Jones, C., Taylor, J., & Herber, O. R. (2014). Vignette development and administration: A framework for protecting research participants. *International Journal of Social Research Methodology*, 17(4), 427–440.
- Cheng, B., Guo, G., Tian, J., & Kong, Y. (2023). 'I disdain the company of flatterers!': How and when observed ingratiation predicts employees' ostracism toward their ingratiating colleagues. *Human Relations*. Advance online publication. <https://doi.org/10.1177/00187267231170175>
- Ciampa, V., Sirowatka, M., Schuh, S. C., Fraccaroli, F., & van Dick, R. (2021). Ambivalent identification as a moderator of the link between organizational identification and counterproductive work behaviors. *Journal of Business Ethics*, 169(1), 119–134.
- Cook, W., & Kuhn, K. M. (2021). Off-duty deviance in the eye of the beholder: Implications of moral foundations theory in the age of social media. *Journal of Business Ethics*, 172(3), 605–620.
- Coyne, I., Gopaul, A.-M., Campbell, M., Pankász, A., Garland, R., & Cousins, F. (2019). Bystander responses to bullying at work: The role of mode, type and relationship to target. *Journal of Business Ethics*, 157(3), 813–827.
- Eckerdt, S., DuHadway, S., Bendoly, E., Carter, C. R., & Kaufmann, L. (2021). On making experimental design choices: Discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes. *Journal of Operations Management*, 67(2), 261–275.
- Ejelöv, E., & Luke, T. (2020). "Rarely safe to assume": Evaluating the use and interpretation of manipulation checks in experimental social psychology. *Journal of Experimental Social Psychology*, 87, 103937.
- Freedman, G., Seidman, M., Flanagan, M., Green, M. C., & Kaufman, G. (2018). Updating a classic: A new generation

- of vignette experiments involving iterative decision making. *Advances in Methods and Practices in Psychological Science*, 1(1), 43–59.
- Greenwood, B., Adjerid, I., Angst, C. M., & Meikle, N. L. (2022). How unbecoming of you: Online experiments uncovering gender biases in perceptions of ridesharing performance. *Journal of Business Ethics*, 175(3), 499–518.
- Grol-Prokopczyk, H. (2014). Age and sex effects in anchoring vignette studies: Methodological and empirical contributions. *Survey Research Methods*, 8(1), 1–17.
- Harrits, G. S., & Møller, M. Ø. (2021). Qualitative vignette experiments: A mixed methods design. *Journal of Mixed Methods Research*, 15(4), 526–545.
- Hassan, S., Pandey, S., & Pandey, S. K. (2021). Should managers provide general or specific ethical guidelines to employees: Insights from a mixed methods study. *Journal of Business Ethics*, 172(3), 563–580.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Massachusetts Menlo Park, California: Addison-Westly Publishing Company.
- Hughes, R., & Huby, M. (2002). The application of vignettes in social and nursing research. *Journal of Advanced Nursing*, 37(4), 382–386.
- Hurst, C., Simon, L., Jung, Y., & Pirouz, D. (2019). Are “bad” employees happier under bad bosses? Differing effects of abusive supervision on low and high primary psychopathy employees. *Journal of Business Ethics*, 158(4), 1149–1164.
- Jared, M. D. (2019). Avoiding the hypothetical: Why “mirror experiments” are an essential part of survey research. *International Journal of Public Opinion Research*, 32(2), 266–283.
- Klebe, L., Felfe, J., & Klug, K. (2021). Healthy leadership in turbulent times: The effectiveness of health - oriented leadership in crisis. *British Journal of Management*, 32(4), 1203–1218.
- Levine, E. E., & Wald, K. A. (2020). Fibbing about your feelings: How feigning happiness in the face of personal hardship affects trust. *Organizational Behavior and Human Decision Processes*, 156, 135–154.
- Liao, Z., Yam, K. C., Lee, H. W., Johnson, R. E., & Tang, P. M. (2023). Cleansing or licensing? Corporate social responsibility reconciles the competing effects of unethical pro-organizational behavior on moral self-regulation. *Journal of Management*. Advance online publication. <https://doi.org/10.1177/01492063231154845>
- Liu, X. L., Lu, J. G., Zhang, H., & Cai, Y. (2021). Helping the organization but hurting yourself: How employees’ unethical pro-organizational behavior predicts work-to-life conflict. *Organizational Behavior and Human Decision Processes*, 167, 88–100.
- Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, 64(1), 19–40.
- Marie, B. S., Jimmerson, A., Perkhounkova, Y., & Herr, K. (2021). Developing and establishing content validity of vignettes for health care education and research. *Western Journal of Nursing Research*, 43(7), 677–685.
- Matza, L. S., Stewart, K. D., Lloyd, A. J., Rowen, D., & Brazier, J. E. (2021). Vignette-based utilities: Usefulness, limitations, and methodological recommendations. *Value in Health*, 24(6), 812–821.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D., & The Prisma Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Survey*, 8(5), 336–341.
- Ng, T., Wang, M., Hsu, D., & Su, C. (2020). Voice quality and ostracism. *Journal of Management*, 48(2), 281–318.
- Paustian-Underdahl, S. C., Little, L. M., Mandeville, A. M., Hinojosa, A. S., & Keyes, A. (2023). Examining the role of maternity benefit comparisons and pregnancy discrimination in women’s turnover decisions. *Personnel Psychology*. Advance online publication. <https://doi.org/10.1111/peps.12577>
- Pieper, J. R., Trevor, C. O., Weller, I., & Duchon, D. (2019). Referral hire presence implications for referrer turnover and job performance. *Journal of Management*, 45(5), 1858–1888.
- Pilling, B. K., Crosby, L. A., & Jackson Jr, D. W. (1994). Relational bonds in industrial exchange: An experimental test of the transaction cost economic framework. *Journal of Business Research*, 30(3), 237–251.
- Rice, D. B., Taylor, R. M., Wang, Y., Wei, S., & Ge, V. (2023). My company cares about my success horizontal ellipsis I think: Clarifying why and when a firm’s ethical reputation impacts employees’ subjective career success. *Journal of Business Ethics*, 186(1), 159–177.
- Ritzenhöfer, L., Brosi, P., Spörrle, M., & Welpe, I. M. (2019). Satisfied with the job, but not with the boss: Leaders’ expressions of gratitude and pride differentially signal leader selfishness, resulting in differing levels of followers’ satisfaction. *Journal of Business Ethics*, 158(4), 1185–1202.
- Rungtusanatham, M., Wallin, C., & Eckerd, S. (2011). The vignette in a scenario-based role-playing experiment: Vignette in an SBRP experiment. *Journal of Supply Chain Management*, 47(3), 9–16.
- Smallfield, J., Hoobler, J. M., & Kluemper, D. H. (2020). How team helping influences abusive and empowering leadership: The roles of team affective tone and performance. *Journal of Organizational Behavior*, 41(8), 757–781.
- Telwate, A., Anglim, J., Wynton, S. K., & Moulding, R. (2017). Workplace accommodations for employees with disabilities: A multilevel model of employer decision-making. *Rehabilitation Psychology*, 62(1), 7–19.
- Terbeck, S., Charlesford, J., Clemans, H., Pope, E., Lee, A., Turner, J., ... Bussmann, B. (2021). Physical presence

- during moral action in immersive virtual reality. *International Journal of Environmental Research and Public Health*, 18(15), 8039.
- Wason, K. D., Polonsky, M. J., & Hyman, M. R. (2002). Designing vignette studies in marketing. *Australasian Marketing Journal*, 10(3), 41–58.
- Zhong, M., Wayne, S. J., & Michel, E. J. (2023). When the past and the present collide: Contrast effect of sequential psychological contract breaches on employee outcomes. *Journal of Management*, 49(3), 913–943.

The vignette in experimental vignette methodology: Current status and design strategies in managerial psychology research

YU Jingting, WEI Xuhua, MA Yi'en

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

Abstract: Although the experimental vignette methodology (EVM) enhances both internal and ecological validity, there is still a lack of detailed and standardized guidance for the design of vignettes in the field of management. Based on this, this study first introduced the definition and types of EVM. Additionally, we clarified the differences between EVM and contextual priming methodology. Second, we coded and analyzed 93 scenario experiments published in 20 major domestic and foreign management-related journals over the past five years. Utilizing the coded data, we presented an in-depth examination of how to design the vignettes in EVM across three stages: pre-design, design, and post-design. During the pre-design phase, the focus was on determining when to utilize the EVM; The design stage centered on how to draft the vignettes, which was divided into five sub-stages: determining the number of vignettes, drafting the vignettes, choosing the media, standardizing vignettes and enhancing vignettes' realism. The post-design stage focused on evaluating whether the vignettes were clear, realistic, complete and effective. If the vignettes failed to meet the desired standards or another scenario experiment was required, researchers should return to the design stage to iteratively modify or rewrite the vignettes until they met the desired criteria. Future research could utilize incident technique and information technology to develop vignettes. Furthermore, vignettes with iterative decision are encouraged to be created. It is also important to note that the design process of vignettes should be guided by theoretical principles. By doing so, it will enhance the applicability, richness and authenticity of vignettes.

Keywords: experimental vignette methodology, vignette, design strategy, managerial psychology experiment