

《心理科学进展》审稿意见与作者回应

题目：管理心理研究中情景实验材料的使用现状与设计策略

作者：于璟婷，卫旭华(通信作者)，马艺恩

第一轮

审稿人 1 意见：

本文通过对管理心理研究中情景实验材料的使用现状与设计策略进行分析和总结，提供了详细的情景实验材料设计指南，并展望了未来情景实验材料设计的发展方向。这些优点使得本文具有实用性和指导性，对相关领域的学术研究具有积极的促进作用。本文还存在如下一些可以建议改进的地方：

意见 1：更加有效地呈现、区分和对比情境实验法和情境启动法的共性与差异。

回应：感谢审稿人对本研究的认可以及提出的宝贵意见！

根据老师的建议，我们在第二节对情景实验法和情境启动法进行了详细对比，并增设表 2 以更加有效地呈现二者之间的共性和差异，具体见第 34、35 页。

情景实验法是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内，以观测各情景对个体认知、情绪和行为影响的研究方法。情境启动法则被描述为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(张玥, 辛自强, 2016)。上述界定使得二者既存在共性，也具有差异，详见表 2。一方面，二者的共性之处在于均属于实验法，并且语境场景是应用方式之一。另一方面，二者在方法特性、情景特点和操纵变量三个方面存在区别。从方法特性来看，在情境启动法中，实验刺激先启动了被试的某种认知表征、思维过程或心理感受，这使其处于一种准备状态，这种准备状态会对后续无关任务产生无意识的影响(张玥, 辛自强, 2016)；在情景实验中，被试无需经历上述准备状态，这是因为其通过代入直接对实验刺激做出反应。从情景特点来看，情境启动法的材料无需模拟现实；相反，情景实验材料需要以现实世界为摹本进行设计。从操纵变量来看，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); 情景实验法除了可以操纵被试的心理变量外, 还能够操纵环境类变量、行为类变量、关系类变量或产品、品牌或广告等, 因此情景实验法的应用范围更广。另外, 方法的异质性决定了技术的不同, 情境启动法常用的技术是词语搜索、归纳/举例任务等(Bargh et al., 2001; Fujita & Carnevale, 2012), 而阅读材料、观看视频则多用于情景实验法中。

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内, 以观测各情景对个体认知、情绪和行为倾向影响	被试通过代入直接对实验刺激做出反应	实验材料需要模拟现实	除被试的心理变量外, 还可以操纵环境类变量、行为类变量、关系类变量或产品等	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	实验材料无需模拟现实	被试的心理变量, 如认知表征、思维过程和 心理感受等	词语搜索、归纳/举例任务等

共同之处: 1.二者均属于实验法; 2. 语境场景是应用方式之一

意见 2: 以附录的形式列出编码者对文献编码不一致的内容及最后的共识。

回应: 感谢审稿人提出的宝贵意见!

本研究由两位编码者独立编码, 并通过对比二者编码的一致性来保证编码的有效性。具体而言, 首先, 我们基于前人研究制定编码规则, 并邀请专家进行论证; 其次, 尽可能获取研究的实验材料; 最后, 编码尊重原文。编码完成后, 我们依据 Holsti (1969)提出的相互同意度及信度公式, 求得两位编码者的编码一致性信度为 0.944。对于不一致的编码, 先由编码者指出原文依据, 并解释说明编码原因; 再对照编码规则, 充分研读原文和实验材料; 最终, 小组讨论直到达成一致。

根据老师的建议，我们以附录的形式列出了与研究内容相关的不一致编码和共识，并通过举例，阐述了如何对不一致的编码达成共识，具体可见修改说明的附录。

意见 3: 目前是否存在符合研究者完整情景实验设计思路的研究？如果有，宜给出推荐的范文列表及相关信息，从而增强证据力。

回应: 感谢审稿人提出的宝贵意见！

本研究虽然对设计前、中、后三个阶段的每一条策略都通过表格的方式进行了举例，但缺少对完整情景实验设计思路的印证。目前，编码文章中中的确存在符合所提出设计思路的研究，因此，根据老师的建议，我们在 4.3.2 后增设 4.4 实例分析一节，列出了相应的范文和阐述了相关信息，以增强证据力，详见第 51 页。具体如下：

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(Hofenk et al., 2019)，因此，为更好地说明和印证所提出的材料设计思路，本研究接下来以 Hofenk 的研究为实际例子进行演示。该研究主要探讨的是零售商的可持续性努力如何以及何时转化为积极的消费者反应，其自变量为分配可持续性和分销可持续性。

在设计前阶段，一是鉴于该研究的目的是探究零售商的可持续性努力对消费者反应的因果效应，二是自变量为客观变量，属于情景实验法可操纵变量的范畴，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法来开展研究。

在设计阶段，一是确定情景数量。对于背景而言，为了增强研究结论的外部效度，研究者选择多背景设计，即两种背景：办公用品和个人护理。对于操纵而言，该研究具有两个存在高低水平的自变量。因此，该实验采用 2(分配可持续性：高/低)*2(分销可持续性：高/低)*2(商店类型：办公用品/个人护理)的因子设计，共 8 个情景材料。二是初步设计情景，研究者主要利用自下而上法，通过现实素材的搜集与筛选，形成包含地点(荷兰、OfficeSupplies)、现实环境(零售商逐渐走向绿色发展之路)、有助于理解操纵的背景材料。操纵材料的设计同样基于自下而上法。三是选择情景媒介，报纸报道与该研究场景适配，因此研究者选用了文本和图片相混合的媒介形式。四是规范情景材料，8 则材料的信息量相似，措辞清晰准确，字数均在 400 个英文单词左右，且使用了虚构的商店名称(如 OfficeSupplies)，并利用字号、斜体和加粗的形式强调关键内容。五是营造真实情景，初步设计时，情景材料来源于现实生活中的绿色商店；媒介选择时，材料呈现了与商店类型相关的产品图片；材料规范时，研究者模仿报纸的写作文风和排版版式。

在设计后阶段，该研究在预测试中运用可信与否对情景材料进行真实性评价，并运用操纵检验进行构念效度评价，二者均通过检验。随后，根据被试的反馈与建议，材料的措辞被进一步地调整和修改。

意见 4: 关于指导情景实验材料设计的理论选择及适切性判断需要引起重视，建议在未来展望中点一下。

回应: 感谢审稿人对本研究提出的宝贵意见！

根据老师的建议，我们在“5 未来展望”中进行补充，详见第 53 页。具体如下：

5.4 重视情景实验材料设计的理论选择及适切性判断

已有研究在设计情景材料时理论意识尚浅，因此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，恰当选择情景实验法来开展研究。二是设计阶段，对于操纵材料来说，理论驱动确定操纵变量及水平。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），进而在概念内涵的指导下，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，在该阶段，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、有效性和完整性，并对相关内容予以修改和完善。

另外，本研究对此部分内容涉及的中英文摘要及其他章节也进行了相应的修改，具体见第 32、52、59 页。

.....

审稿人 2 意见:

作者基于个人创建的情景实验材料设计规范框架，分析了近五年国内外管理心理领域期刊中的 185 项情景实验。在分析现状的基础上，作者针对情景材料设计的前、中、后三个阶段，全方位地提出了具体的设计建议和注意事项。总之，这是国内外如此全面、系统地分析情景材料设计的为数不多的研究之一，具有非常高的创新价值。为了更好地反映该研究成果，特提出以下的修改意见和建议，仅供参考：

意见 1: 在“表 1 情景实验的类别”中,除了按照媒介形式和研究目的对情景实验进行分类外,是否还可以考虑从其他维度进行分类。

回应: 感谢审稿人对本研究的认可以及提出的宝贵意见!

根据老师的建议,本研究在第二节基于研究目的(why)、媒介形式(how)、实施地点(where)和被试视角(who)等维度,补充了情景实验的分类,见 33、34 页。具体如下:

情景实验包括多种类别,鉴于此,本研究通过研究目的(why)、媒介形式(how)、实施地点(where)、被试视角(who)等维度,介绍情景实验的分类,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)。内隐决策情景实验要求被试在包含不同属性水平的情景之间进行排序,目的是了解各个属性的相对重要性;外显决策情景实验则要求被试做出明确的选择,适用于评估明确的过程和结果。按照媒介形式分为单一感官媒介情景实验和多重感官媒介情景实验,二者的区别在于情景材料的媒介形式,前者只作用于被试的某一种感官,如文本、图片和音频等;后者则借助视频、数字和角色扮演等媒介,调动被试多重感官。按照实施地点分为传统实验室情景实验和线上情景实验,其分别在实验室中借助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情景实验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,后者则需要被试以第三视角参与实验。

表 1 情景实验的类别

分类依据	情景实验类别	含义	特点
研究目的 (why)	内隐决策情景实验	被试需要在包含不同属性水平的情景之间进行排序,适用于了解隐性决策过程的影响	通常呈现某产品或服务不同属性及水平的组合,关注消费者的内在价值体系
	外显决策情景实验	被试需要做出明确的决定、判断和选择,适用于评估明确的过程和结果	通常因变量明确,关注被试对假设情景的明确回应
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的情景实验	成本低、简单易行、真实性低、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的情景实验	成本高、设计流程复杂、真实性高、多重感官
实施地点 (where)	传统实验室情景实验	在实验室情境下,借助专门的实验设备(如 VR),通过有计划的控制实验条件以检验因变量变化情况	媒介限制少、内部效度更高、样本受限
	线上情景实验	利用见数、Amazon Mechanical Turk 等网络平台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	被试需要想象自己是情景中的主人公	常用于工业与组织心理学
	第三视角情景实验	被试需要以第三视角参与实验	常用于消费心理学

意见 2: 建议对背景材料和操纵材料分别加以具体介绍和举例说明。

回应: 感谢审稿人提出的宝贵意见!

根据老师建议,我们对 4.2 部分进行了修改和补充,见第 42 页,具体如下:

情景材料由背景材料和操纵材料两部分构成,分开来看,背景是承载情景产生、发展的舞台,主要传递人物信息、人物关系、时间地点、现实环境、操纵铺垫等必要信息,并控制混淆因素(Rungtusanatham et al., 2011)。操纵材料则描绘了操纵变量不同水平所对应的现实情况,以观测其对个体认知、情感和行为产生的因果效应。例如,凌文铨等(2019)对中国组织情境下上司与下属之间社会交换的互惠机制进行了探讨,自变量之一为上下属关系。由于其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)。基于上下属关系分为亲近和疏远两个水平,操纵材料则分别对应描述了小王与“我”的关系情况。

意见 3: 为什么只选择近五年的相关研究文献?可以简单说明一下理由和最初的想法。

回应: 感谢审稿人提出的宝贵意见!

根据老师建议,本研究在“3.1.1 文献搜索与筛选中的(1)”检索时间处阐述了为什么选择近 5 年相关研究文献,见第 36 页。具体如下:

最后,检索时间范围为 2019 年 1 月至 2023 年 5 月。本研究之所以选择上述检索时间是因为,一方面,Aguinis 和 Bradley 对情景实验的探讨是该领域经典的高被引方法论文章(Aguinis & Bradley, 2014)。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。另一方面,本研究的目的在于通过编码情景实验文章,归纳总结情景实验材料设计策略,所以这要求实证研究中实验材料设计的规范性、完整性和透明性,而早期研究难以达到上述标准(Baguley et al., 2022)。

意见 4: 由于在情景材料由背景材料和操纵材料两部分组成,因此建议在“4 管理心理情景实验的情景材料设计策略”部分中,分别针对背景材料和操纵材料谈一下设计策略的问题。

回应: 感谢审稿人提出的宝贵意见!

根据老师建议,本研究在“4 管理心理情景实验的情景材料设计策略”部分,针对相关步骤分开阐述了背景材料和操纵材料的设计策略。

一是设计阶段的确定情景数量,见第 44 页。目前已分段论述了如何确定背景材料和操纵材料的数量。

二是设计阶段的初步设计情景，见第 45、46 页。在介绍自上而下法时，先通过“对于背景而言”的句式，阐述了背景材料的设计策略，再通过“在自上而下法的实验操纵设计中，”论述操纵材料的设计策略。自下而上法的介绍并未遵循上述思路，而是杂糅在一起，故修改。

三是设计阶段的规范情景材料。我们通过重新编码或措辞明确化方式对信息量一致、对称式设计和关键内容强调进行了补充，以更明确背景材料和操纵材料的对应情况，具体见第 47、48 页。

四是设计阶段的营造真实情景。我们也对这部分进行了修改，详见第 48 页。上述修改在表 4 中同样进行了相应的调整和明确，可见第 43、44 页。

第二轮

编委 1 意见：同意发表

编委 2 意见：

意见 1：杂志的选择问题，作者探讨的是管理心理学中情景实验方法的应用，为什么有两个营销学杂志也纳入进来？如果是这样，management science 上也有很多类似的方法应用，为什么不纳入其他管理领域的杂志？同时，与管理相关的这 6 个杂志是如何选择的，诸如，personnel psychology, journal of organizational behavior 也是很好的杂志，为什么没有纳入进来？是否存在选择性的偏差？

回应：感谢审稿人提出的宝贵意见！

本研究的出发点是想探讨管理学中与心理有关主题的情景实验。但是，的确如审稿专家所说管理心理学是一门研究组织管理活动中人的行为规律及其潜在心理机制的学科。因此，本研究充分汲取编委专家的意见，在期刊选择上做了两方面的更改：(1) 删除期刊，将 2 本营销学杂志删除，且未纳入其他管理领域杂志，如 Management Science。(2) 增补期刊，在原来 6 本杂志的基础上，将 FT50 学术期刊目录中有关管理心理领域的主流期刊，以及 Personnel Psychology, Journal of Organizational Behavior 等杂志一并纳入，详见第 28 页，具体期刊选择流程如下：

本研究选择管理心理领域 20 本国内外权威期刊的文献。具体地，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》。

鉴于期刊选择发生了变化，故本研究重新进行了文献检索、筛选与编码，并对编码一致性重新计算，对现状再次梳理。虽然样本文献有所变化，但本研究归纳的设计策略较为稳健，仅需对文中涉及的营销学例子予以替换，具体改动见第 20、27-35、37-41 页。

为进一步缓解选择性偏差的影响，本研究另在 Web of Science 核心合集 SSCI 数据库进行关键词检索，随机挑选 10 篇描述情景材料设计方式的管理心理领域文章进行编码。通过编码发现，本研究总结的情景材料设计策略基本能够囊括现有情景材料的设计过程。

意见 2: 文章的定位需要更加清晰，文章的重点是评价这一研究方法的优劣，还是告诉读者如何进行情景实验设计？如果是前者，建议对这一方法的优劣进行深入的分析，如果是后者，请给出更加详细的设计思路。

回应: 感谢审稿人提出的宝贵意见！

本文的重点是“如何进行情景实验材料的设计”。自查后发现，可能是由于引言的写法使得文章的定位模糊，故对引言进行了修改，详见第 25 页。具体如下：

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)。其是指研究者需要通过模拟现实的情景设计来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。虽然情景实验法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，即使采用情景实验法的研究，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度威胁因素(Antonakis & John, 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，然而其侧重点在于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，情景实验法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。首先，本研究阐释了情景实验法的内涵、类型和与情境启动法的区别。其次，基于 20 本国内外管理心理权威期刊近五年发表的 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，归纳和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。

为提供更加详细的设计思路，我们在第 4 章开始部分整体叙述设计流程并绘图、另在“4.2.2 初步设计情景”、“4.2.4 规范情景材料”、“4.3.5 营造真实情景”和表 4 中增补细节信息。详见第 31、34、35、37-39 页。

情景设计分为设计前阶段、设计阶段和设计后阶段，流程见**错误！未找到引用源。**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，直至使其满足五项标准：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。

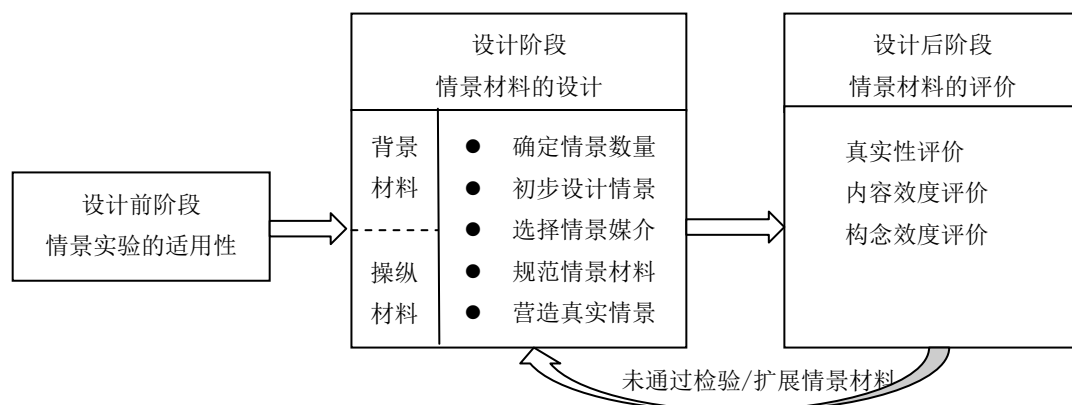


图 2 情景材料设计流程

意见 3: 情景实验的缺点本文讨论的较少，实际上，现在很多 top tier journal 对于该方法的使用是很谨慎的，因为它的外在效度不高，paper person 很难应用到现实中，建议作者在这一方面深入探讨。

回应: 感谢审稿人提出的宝贵意见！

根据老师的建议，我们不仅在“4.1.1 情景实验的适用情形”处论述了情景实验的缺点，也在未来展望 5.4 节处再次强调，从而帮助研究者在选用该方法时进行谨慎思考。详见第 33、42 页。具体如下：

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)。

5.4 重视情景实验材料设计的理论选择及適切性判断

未来研究应重视理论指导并判断適切性。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。

意见 4: 文章内容太长，很多地方没有必要过于详细的介绍，如 3.2 一节太琐碎，建议控制在 15000 以内。

回应: 感谢审稿人提出的宝贵意见！

根据老师的建议，我们通过文字精炼、参考文献整合以及章节详略安排(如简述 3.2 节)，

缩减字数 5000 多字，见全文绿色字体。

意见 5：文字的精炼与准确需要增加。个别地方的语言需要统一，如“五”还是“5”。

回应：感谢审稿人提出的宝贵意见！

根据老师的建议，我们对全文进行了检查并对语言进行了统一，如“近五年”。

第三轮

编委 2 意见：文章基本达到了发表水平，建议发表！

主编意见：该文经过多位评审专家的评审，作者进行了认真的修改，达到发表水平，同意发表。