

《心理科学进展》审稿意见与作者回应

题目：基于视频游戏的空间能力测评

作者：尚俊杰 石 祝 沈科杰

第一轮

审稿人意见：

意见 1：前言部分，空泛的议论较多，需要精炼；需要对每一主题，需要概括出几个要点。方法部分，简要说明研究思路，尤其重点说明需要观察的变量；

回应：

感谢专家的建议。原文在论述研究问题时的确有重点不清晰的问题。根据您的建议，我们在问题提出部分，在原稿基础上整体修改并将其重新按照“空间能力的重要意义；空间能力测评的现状和问题；基于游戏的测评的优势；本研究的方法、思路与创新点”这四个主要部分重新撰写。本研究选择“基于视频游戏的空间能力测评”为题，其主要思路是利用视频游戏的媒介特征支持空间能力测评：一方面基于视频游戏环境设置贴近真实情境的空间能力测评任务，另一方面利用游戏后台实现自动化、无感知的“隐形测评”，并且能利用视频游戏吸引参与者的特点，探索大规模测评的可能性。

对于关键测评变量和指标，在原稿问题提出部分没有清晰论述。根据专家的建议，我们首先在问题提出部分论述了变量指标的基本结构。在研究构想中对这些指标进行了详细的梳理，重点明确了空间能力测评工具的可观察变量，即被试者完成游戏化空间能力测评关卡的任务变量（反应时、正确率、运动距离、角度和方向误差、游戏内部得分等）（详见表 1 测评指标）。结合玩家玩游戏时的多模态生理指标（眼动注视时间、注视点分布、注视次数、扫视路径、脑电 N2、P2、P3 电位、皮电等）。由于在游戏化测评工具中，空间能力测评指标也会受游戏玩法和机制限制，要在游戏测评工具开发迭代过程中逐渐确定，难以在问题提出部分完全呈现观察变量，更细致的指标框架会随游戏开发逐渐完善细化。

意见 2：在研究现状部分，没有有效总结出基于游戏测评空间能力的创新之处；需要说明新发现的测量指标、空间能力的生理指标，以及信效度参数等；

回应：

感谢专家的建议。在研究现状部分，我们原稿总结的确不够系统到位，根据专家的建议对空间能力游戏化测评的发展脉络重新梳理为“基于游戏的测评方法、空间能力的游戏化测评研究、空间能力的多模态测评研究三个部分”。基于游戏的测评方法重点介绍游戏化测评的优势与特点，以及隐形测评的概念。基于游戏的测评方法能够设置贴近真实情境的测评任务，呈现丰富的空间信息，游戏后台可以有效收集学习者过程性数据，实现无感知的隐形测评，并且游戏化测评可以用于测量复杂问题解决和元认知等高阶思维过程，通过游戏发布平台能够吸引大量玩家参与，易于形成大规模数据。因此，基于游戏的测评在内容开发、信息呈现、数据收集、可解释性、测评体验、应用推广等方面都具有显著优势。

根据专家的建议，我们补充和梳理了当前与空间能力有关的计算机化测评和游戏化测评研究，并详细介绍了当前研究使用哪些测评指标、变量和算法来刻画空间能力水平。另外，由于研究者对空间能力的定义和分类并不完全一致，在测评指标选取上，要么是从经典量表改

编而来，只针对某一种空间能力子维度的指标，如心理旋转测试 (Xu et al., 2023)；要么是根据自身设计的游戏工具的特点来寻找指标，如 Ventura 等人 (2013) 开发的虚拟空间导航测试用游戏用时和游戏得分来计算空间能力。本研究综合以上两类研究的思路，根据空间能力分类框架，总结空间能力各个子维度对应的心理测评任务，再结合当前已有的数字化测评工具和游戏化测评的特征，提出空间能力测评指标。

意见 3: 基于多模态的测评与基于游戏的测评，是两个不同的领域，需要分开讨论；

回应:

感谢专家的建议，基于多模态的测评部分在初稿撰写的时候也斟酌权衡了很久，游戏化测评和多模态测评确实两个相对独立的部分，原稿存在混杂不清晰问题。根据专家的建议，我们重新撰写了 2.3 节空间能力的多模态测评部分，对当前针对空间能力的生理测评研究进行了简单的回顾，并探讨了与空间能力相关的生理信号指标。对于两种测评方法的关系，本研究以游戏化测评为主线，开展眼动和脑电研究的主要目标是取得更多学习者在游戏过程中的行为表现数据，结合游戏内数据来作解释和预测其空间能力水平，并通过多模态学习分析来解释其背后潜在的认知机制。

意见 4: 研究构想中，理论模型应该是在前面部分已经提出的，而不应该是研究构想中。目前的实验设计没有专注于变量的操控，需要说明观察数据的类型以及可能得到的结果；

回应:

感谢专家的建议。对于您提出的理论模型应在前面的建议，我们经过仔细考察和思考，选择在本章节呈现研究的整体框架。一方面在研究现状基础上，针对空间能力测评进行更细致的展开，同时介绍各个子研究的具体内容和之间关系，所以仍然选择在此处放置研究框架图。这项研究根据以证据为中心的设计 (ECD) 框架来组织和构建空间能力测评。研究一旨在通过文献调研分析空间能力的概念内涵和能力维度变量 (即熟练度变量)，根据专家的建议，我们在对空间能力的分类、任务和测评指标做了进一步的梳理 (表 1)，明确了测评空间能力的潜在指标变量。研究二中我们聚焦行为表现指标，包括玩家在虚拟空间中的行为表现指标 (如任务用时、行动轨迹、无效探索时间、完成任务数量、游戏内得分等指标) 和生理信号指标，包括眼动 (注视时间、注视点分布、扫视路径、和注视次数)、脑电 (N2、P2、P3 电位) 等，并呼应研究现状部分，对数据类型和可能观察到的结果进行了论述。

意见 5: 在理论建构与创新部分，讨论的问题杂乱无章，既有文献分析，也有研究方法介绍。建议聚焦在理论建构与创新上，例如是否与创新的行为测量有别，以及是否能够对比过去的基于游戏的测量有什么创新。目前的介绍、分析没有抓住重点。

回应:

感谢专家的建议。原稿中理论建构与创新部分的确存在论述缺乏重点，比较散乱的问题。根据专家的建议，我们对理论建构与创新部分整体重写，重点讨论了基于视频游戏的空间能力测评的创新点。根据专家的建议，本部分聚焦研究的创新点，删除了与理论建构不相关的内容，将部分内容整合到文献分析和研究方法之中。本项目的创新之处在于，进一步发展了基于视频游戏的空间能力测评方法，相比于先前的游戏化测评研究多聚焦于对游戏后台数据的事后分析，本项目将游戏化测评与多模态学习分析深度融合，实现了对空间能力的过程性评价和整体性理解。并且首次在空间能力测评中，将以证据为中心的设计框架 (ECD) 和贝叶斯网络 (BN) 相结合，用于识别汇总空间能力多模态行为数据，通过整合学习者行为数据、游戏后台数据、以及游戏过程中的生理信号数据，实现多源数据驱动的空间能力测评。

第二轮

审稿人意见：作者根据审稿意见，进行了基本的修改。目前只是字数还有 16000 多，不含文献，是否符合杂志要求。

编委意见：同意发表。