

《心理科学进展》审稿意见与作者回应

题目：联合行动中的自我-他人整合：基于层级任务表征的视角

作者：汪俊 李乾楷 赵海洋 赵梦飞 潘婷婷 郑峥

第一轮

审稿人意见：

该论文依托国家自然科学基金面上项目，聚焦联合行动领域人际如何协同这一核心科学问题，从层级任务表征的创新视角出发，系统构建了“现象-机制-调节”三层研究框架，融合行为实验、脑电超扫描与认知计算建模多技术手段，旨在揭示高低层级任务中自我-他人整合的分离、交互机制与边界条件。论文选题贴合社会认知与工程心理学领域的前沿热点，突破了以往单一层级研究的局限，创新性地提出了基于层级贝叶斯计算的自我-他人整合理论模型，兼具重要的理论创新价值与潜在的实践应用前景，整体符合《心理科学进展》“研究构想”栏目的核心定位与投稿要求。但仍存在以下重要问题需进一步深入思考并补充完善：

回应：非常感谢审稿专家的宝贵意见！以下内容是对修改意见逐条给出的修改说明。

意见 1：摘要部分虽包含核心要素，但对科学缺口的锚定不精准、创新点的核心增量不明确，建议进一步严格遵循“研究背景与核心科学缺口→研究整体框架→理论创新与学术价值”的逻辑重构；现有引言的逻辑链条也较松散，需按上述逻辑重构为四层递进结构；

回应：非常感谢审稿专家极具洞察力与建设性的指导意见。我们完全认同原稿在摘要与引言中存在逻辑递进不够紧凑、科学缺口聚焦不足以及理论增量凝练不力等问题。为此，我们严格遵循您提出的四层递进逻辑，并依据重新撰写的摘要版本，对回应细节做了更精准的匹配。具体修改如下：

1. 锚定研究背景与核心科学缺口

修改后的摘要开篇以“联合行动常呈具象至抽象的层级结构”切入，明确指出两大核心局限：“局限于单一层级的孤立探讨”与“将人际协同视为静态映射”。由此精准锚定核心科学缺口，即“层级间动态交互与信息累积机制的缺失”。

2. 重构研究整体框架

基于“层级任务表征”视角，我们将整体框架凝练为三个层层递进的核心科学问题：(1) 高低层级自我-他人整合在行为与神经层面的分离与交互；(2) 层级特异性是否源于先验信息积累模式的差异，及其交互如何通过先验与后验分布的动态耦合实现；(3) 不同人际共同基础情境如何调节该过程。由此，研究的问题导向更加清晰。

3. 明确研究方法路径

我们简明扼要地补充了本研究的创新性方法路径，即构建“层级嵌套任务范式”，融合行为实验、脑电超扫描与认知计算建模，从“现象—机制—调节”三层面探究层级特异性与交互规律。

4. 提炼核心创新点与学术价值

修改稿明确将核心理论增量凝练为“提出基于层级贝叶斯计算的自我-他人整合理论模型”，并概括其学术价值：实现从“静态单层表征”到“动态层级预测”的跨越，深化了对高低层级人际整合动态机制的理解，同时为协同多智能体系统开发提供了可计算、可验证的心理学依据。

基于相同的四层递进逻辑，我们亦对正文“1.问题提出”部分进行了系统性打磨与重组，使整体逻辑链条更为严密。鉴于修订量较大且篇幅有限，详细修改内容已在“摘要”及“1.问题提出”部分以蓝色字体标出。

意见 2：文献综述的核心要求是评述性而非罗列性，需系统梳理领域发展脉络，指出现有研究的争议与缺口，本研究现有综述存在结构松散、重罗列轻评述、理论推导不足等核心问题，需做进一步提升修改：首先需系统梳理联合行动中自我-他人整合的理论发展与研究现状，梳理从经典共同表征理论到最新预测加工框架的理论演进；然后是单一层级研究的局限与层级任务表征的研究进展；三是自我-他人整合的认知机制，评述贝叶斯模型在社会互动、联合行动领域的应用现状；最后是人际共同基础对自我-他人整合的调节作用，明确本研究引入“特质共同基础”的理论依据。

回应：非常感谢审稿专家一针见血的指导意见。我们完全认同您对原稿文献综述“重罗列轻评述、理论推导不足”的批评。文献综述不仅是背景介绍，更是引出本研究核心假设的逻辑基石。为此，我们对“2.研究现状及发展动态”部分进行了全面的结构重组与理论深化，严格按照您建议的四个递进层次展开评述性撰写。具体修改如下：

1. 系统梳理理论演进：从“静态表征”到“动态预测”

在“2.1 联合行动中自我-他人整合的理论演进”中，我们不再停留于对联合行动实验现象的罗列，而是将经典的“共同表征理论”与“参照编码理论”加以对比评述，指出二者虽对整合的诱发机制存在争议，却共同受限于将自我-他人整合视为“静态的映射结果”。由此，我们引入最新的“预测加工框架”，强调人际协同在本质上是动态的证据积累与信念更新过程，从而为后续引入贝叶斯模型奠定了理论基础。

2. 梳理评述单一层级局限与层级表征进展：

在“2.2 单一层级的局限：联合行动中中层级任务表征视角的缺失”一节中，我们系统梳理了现有联合行动研究在低层级具象表征与高层级抽象表征两方面的进展，并回顾了个体层级任务表征的成熟理论与研究范式。以此为基础，论证了当前联合行动研究将高低层级孤立考察的局限，进而推导出构建“层级嵌套任务范式”的必要性。

3. 自我他人整合的认知机制与贝叶斯模型的应用现状：

在“2.3 贝叶斯模型视角下的自我-他人整合”一节中，我们首先概述了“社会贝叶斯大脑”假说，并系统评述了贝叶斯模型在显性社会推断（如合作博弈、计算精神病学）中的成熟应用。随后，我们指出现有研究在“自发性联合行动”领域尚存在应用缺口。在此基础上，我们着力强化了理论推导：结合行为层面的中心化偏差与神经层面的额顶叶 θ 频段功能连接证据，详细论证了高、低层级任务整合差异的本质，在于不同先验信息积累模式，即不确定性驱动的中心化效应，由此确立了本研究计算建模的理论合理性。

4. 明确引入“特质共同基础”的理论依据：

在“2.4 人际共同基础的不同维度如何特异性调节……”一节中，我们通过文献梳理，援引“人际协同多重表征模型”（宋晓蕾、董梅梅，2023）作为引入“特质共同基础”的理论依据。该模型指出，社会性表征（如同伴角色、性别、熟悉度等）对成功的联合行动同样不可或缺。在此基础上，鉴于特质作为稳定且显著的个人特征，能够显著影响我们对他人的社会判断（Lin et al., 2025）和刻板印象形成（Jia et al., 2022），理应成为社会性表征的重要组成，从而影响自我-他人整合。进一步，结合最新实证证据，即人际特质相似性（尤其是亲社会特质，如宜人性和外倾性的相似性）能够显著影响个体对同伴的行为表征与预测（Barnby et al., 2022; Arellano-Véliz et al., 2024），本研究推导出核心假设：相似的亲社会特质（外倾性/宜人性）将对高低层级的自我-他人整合产生不对称调节效应。

经由上述四个维度的重构，修改稿的文献综述部分强化了理论推导与评述，为本研究“现象

—机制—调节”的三层研究框架提供了有力支撑。鉴于修订内容较多且篇幅有限，详细修改已统一在正文“2.研究现状及发展动态”中以蓝色字体标出。

意见 3：行文规范性方面，核心表述注意避免不合规和错误，如研究三实验十一中出现“实验十一的结果发现”这类表述，严重违背“研究构想”栏目的定位（构想为尚未开展的系统性研究设计），需删除所有对未正式发表实验结果的确定性表述，仅保留基于前期研究的理论推导与预期；同时删除全文口语化、非学术化的表述，修正重复、冗余的段落，控制论文篇幅符合期刊要求（研究构想栏目一般不超过 15 页）。润色英文摘要与关键词，确保专业术语翻译准确、语句流畅，消除中式英语问题，符合英文学术写作规范。

回应：非常感谢审稿专家严谨细致的审阅。原稿误将尚未开展的实验表述为已获得结果，这一失误确实严重违背了“研究构想”栏目的定位，我们对此深表歉意。针对您的批评与建议，我们已对全文做出如下修改：

1. 纠正时态与不合规表述

已删除所有“实验十一的结果发现……”一类确定性结果表述。同时在“3 研究构想”部分（研究一至研究三各实验设计）中，所有结果相关描述均已一一核对，确保措辞严格基于前期研究的理论推导与预期。

2. 提升学术性与精炼冗余段落

我们已对全文进行系统性文字润色，剔除了口语化、非学术化表述，并重点对“1 问题提出”与“2 研究现状及发展动态”进行缩减与逻辑合并，删除了冗余表达，使论述更简洁、逻辑更紧凑。

3. 控制论文篇幅

通过上述去重与精炼（尤其是压缩过度发散的文献罗列与重复的假设推演），我们在完整保留“现象—机制—调节”核心框架的前提下，将正文篇幅进行了适当缩减使整体行文更加干练。目前全文（不含参考文献和摘要）共 12 页。

4. 润色英文摘要与关键词

我们邀请具有心理学专业背景的英语母语研究者对英文摘要和关键词进行了重新编译。重点对“自我-他人整合（Self-other integration）”“层级任务表征（Hierarchical task representation）”等核心术语的英译做了标准化统校，并重写了摘要中的复杂长句，彻底消除了“中式英语”痕迹，确保表达完全符合国际英文学术写作规范。

第二轮

审稿人意见：此轮修改论文质量得到较大提升，但还有一些问题需进一步补充修改完善：

回应：非常感谢审稿专家的宝贵意见！以下内容是对修改意见逐条给出的修改说明。

意见 1：在 2.1 节“联合行动中自我 - 他人整合的理论演进”中，虽对比了共同表征理论与参照编码理论的核心分歧，但未进一步讨论两种理论的适用边界与整合趋势。例如：两种理论是否在不同任务层级或不同人际情境下分别发挥作用？预测加工框架如何实现对两种理论的整合？

回应：非常感谢专家这一极具建设性的意见，这促使我们对理论部分的探讨更加深入。针对专家提出的两个问题，我们已在 2.1 节进行了重点补充：

第一，明确了两种理论的适用边界。我们补充论述了在低级空间协调任务中，两种理论

均可解释联合行动效应；但当任务涉及意图理解或人际关系表征时，共同表征理论更具解释力，并补充了联合统计学习任务及浪漫伴侣情境下的实证依据予以佐证。

第二，阐明了预测加工框架对两种理论的整合。我们提出将两者视为同一动态加工过程的两个信息流：共同表征提供自上而下的“社会先验预期”，而参照编码强调的知觉显著性提供自下而上的“感觉输入”，个体通过计算预测误差实现信念更新与自我-他人整合。以上补充与修改已在正文中以蓝色字体标出，敬请专家查阅。

意见 2：2.4 节中“相似亲社会特质促进低层级整合、削弱高层级整合”的假设推导不够充分。现有证据仅说明特质相似性提升预测准确性，但未解释为何会对高低层级产生相反影响。

回应：感谢专家的深刻提问，这确实切中了此前假设推导中的逻辑薄弱环节。我们对推导逻辑进行了补全：

已有证据表明，人际特质相似性，尤其是亲社会特质（宜人性和外倾性）的相似性，能够显著提升个体对同伴行为的预测准确性，使参与者更容易准确预测与自己社会偏好相似的伙伴的行为（Barnby et al., 2022; Arellano-Véliz et al., 2024）。进一步，人际特质相似性还会显著拉近人际心理距离，促使个体采用“低水平建构”（Trope & Liberman, 2010），从而降低心理表征的抽象性，将注意力更多聚焦于同伴具体的行动细节（Arellano-Véliz et al., 2024）。另一方面，Sened 等人（2024）发现，感知到的相似性会促使个体在社会推断中更多依赖自我知识（self-knowledge），即倾向于“以己度人”。这种基于自我知识的推断策略容易引发社会投射偏差，使个体高估他人与自己的共识。当任务要求个体表征他人的高层级任务信息（如统计学习规则）时，自我知识的锚定可能导致个体用自身任务信息替代同伴的任务信息，从而削弱对他人任务信息的表征。综上，人际相似性虽然提升了对具体行为细节的预测，却可能降低对同伴高层级抽象任务信息的表征。

意见 3：研究构想部分，实验一将联合西蒙任务与联合统计学习任务嵌套，需同时测量联合西蒙效应（低层级）与联合统计学习效应（高层级），但未说明两种效应的测量方式如何避免相互干扰。例如：刺激序列中内嵌的统计规则是否会影响西蒙效应的稳定性？

回应：非常感谢审稿专家严谨细致的审阅与准确指正。我们完全同意审稿专家关于嵌套范式中两种效应测量方式可能相互干扰的问题，因为两者都用的是反应时指标。为解决这一可能问题，本研究做了以下补充：为进一步避免该层级嵌套任务在行为指标（反应时）测量中可能存在的相互干扰效应，实验一将结合脑电指标进行区分。针对低层级任务表征（刺激-反应映射），基于联合西蒙任务的冲突控制特性，本研究拟采用额中央 N2 成分作为特异性神经指标。已有研究表明，该成分对刺激-反应对应关系高度敏感，能特异性地反映冲突监控与抑制过程（Hoppe et al., 2017; Miti et al., 2026）。针对高层级任务表征（统计规则），基于联合统计学习的预测特性，本研究拟采用顶枕区 P3 幅值作为指标。这些指标随统计规则的习得呈现规律性变化（例如 P3 幅值降低反映预测性增强），反映了规则表征的积累与自上而下控制的减少（Zheng & Wang, 2024b）。通过结合不同时间窗口的 ERP 成分（早期 N2 反映冲突，晚期 P3 反映预测），本研究可在神经层面有效区分高低层级任务表征，从而避免行为测量指标中潜在的信号重叠与干扰。上述内容已用蓝色字体标识在原文中。

意见 4：3.2 节中采用朴素贝叶斯框架构建模型，但层级任务表征本身具有嵌套结构，朴素贝叶斯的“特征独立性假设”可能不完全适用；同时，参数 η 的取值范围（ $0 < \eta < 1$ vs $\eta > 1$ ）的理论依据未充分说明。4.1 节提出的层级贝叶斯自我-他人整合模型未明确其适用边界。例如：该模型是否适用于竞争性联合行动？是否适用于多人（ ≥ 3 人）联合行动？

回应：感谢审稿专家的重要意见。我们从三个方面回答本问题。

1. 非常感谢审稿专家的明确指正, 朴素贝叶斯模型所要求的 "特征独立性假设" 可能不完全适用于本研究使用的联合嵌套任务。正如原文图 1 所示, 高低层级间存在自上而下和自下而上的交互作用, 而非完全独立。因此, 本研究是分两步构建贝叶斯模型 (在此再次为原文表述不够清晰给审稿人造成的误解致歉)。首先, 在每个层级内部 (即高层级任务和低层级任务分别考虑), 任务表征的信息积累过程相对独立, 因此我们采用朴素贝叶斯框架对单层级数据进行拟合。其次, 针对层级间的交互作用, 我们通过引入高低层级间的相互调节系数 α 与 β , 分别量化层级间的自上而下和自下而上影响。具体建模细节请参见正文 3.2 章节。
2. 我们充分认同参数 η 取值范围需要坚实的理论支撑。在修改稿中, 我们结合朴素贝叶斯框架中“先验预期”与“观测数据”的权衡逻辑, 为参数 η 的取值范围提供了坚实的理论依据。具体而言, η 取值本质上反映了认知系统在“先验预期”与“观测数据”之间的权重分配。在朴素贝叶斯推断框架中, 后验由先验与似然共同决定; 标准推断对二者并无固有偏重, 其相对权重客观上取决于各自的可靠性 (Vilares & Kording, 2011)。本研究假设, 面对不同层级的任务结构, 认知系统会动态调节这一权重分配, 而参数 η 正是用于刻画对观测数据权重的调节程度。当 $\eta = 1$ 时, 模型与标准朴素贝叶斯模型完全一致。在涉及抽象规则提取的高层级任务中, 环境不确定性较高, 个体缺乏稳固的先验预期。此时, 认知系统需要对新线索保持高度敏感, 赋予观测数据更大权重, 以快速捕捉环境变化 (Nassar et al., 2010)。因此, $\eta > 1$ 意味着个体采用“数据驱动的快速更新”策略, 即放大新数据的影响力, 加速信念更新, 从而更快地习得新规则。相比之下, 在低层级任务 (如联合西蒙任务) 中, 个体已形成稳定且确信的预期。根据 Mathys 等人 (2011) 的观点, 在先验确定性较高的情况下, 认知系统倾向于维持既有预期。因此, 个体倾向于采用“维持先验稳定性”的策略 ($0 < \eta < 1$): 削弱观测数据的权重, 避免新输入信息动摇稳固的先验预期, 从而保持行为的稳定与连贯。
3. 感谢审稿专家对本模型适用边界的指正。我们已在原文中补充了明确的适用边界说明: “需要说明的是, 当前构建的贝叶斯自我-他人整合模型具有特定的适用边界, 并非适用于所有联合行动场景。在竞争情境中, 自我与他人的目标往往相互对立, 个体不再倾向于整合双方信息, 反而需要抑制对方或突出自身优势, 因此本模型无法直接推广至竞争性任务。在群体规模方面, 本模型基于双人互动框架构建。当参与者人数达到三人及以上时, 群体动力学会发生质变, 例如出现责任分散、社会从众、多方协调等复杂过程, 此时原先一对一的自我-他人整合需扩展为多节点的网络交互结构, 贝叶斯更新的计算复杂度也将呈指数级上升, 因此当前模型仅适用于双人联合行动场景。”以上补充有助于明确模型的适用范围, 避免未来研究的误用或过度推广。再次感谢审稿专家的严谨建议。

意见 5: 格式规范性上还得加强。删除 "助力形成以心理学为基础的 '强人类智能'" 这类过于宽泛且不严谨的表述。统一 "自我他人整合" 的表述, 全文均改为 "自我-他人整合" (带连字符)。图 1 "研究内容的具体框架图" 可进一步优化: 用不同颜色区分三个研究, 标注每个实验的核心操纵变量与预期结果; 补充图注: "图 1 联合行动中高低层级任务表征自我-他人整合的研究框架 (作者绘制)"。检查参考文献中作者名字的缩写一致性, 例如: "Freek van Ede, & Anna C. Nobre." 应改为 "van Ede, F., & Nobre, A. C." 等。

回应: 非常感谢审稿专家对格式规范性的严谨把关, 这些细节指正对提升文章的严谨度非常关键。我们已按照您的建议逐一进行了修改:

1. 删除了“助力形成以心理学为基础的‘强人类智能’”这类宽泛且不严谨的表述。
2. 全文已统一改为“自我-他人整合” (带连字符), 并仔细核对确保无遗漏。
3. 图 1 已重新绘制优化, 采用不同颜色区分了三个研究, 清晰标注了各项实验的核心操纵变量与预期结果, 并补充了图注“图 1 联合行动中高低层级任务表征自我-他人整合的研究框

架（作者绘制）”。

参考文献的作者姓名缩写已全部核对并统一，如专家所指出的，已将原错误格式修正为“van Ede, F., & Nobre, A. C.”等规范形式。以上所有修改之处，均已在正文中以蓝色字体标出，敬请专家查阅。

第三轮

审稿人意见：较好完成前二轮修改工作，文字格式进一步按期刊要求规范后可考虑发表。

回应：非常感谢您对本文前两轮修改工作的认可，以及提出的“文字格式进一步按期刊要求规范”的宝贵意见。为此，我们对照《心理科学进展》的撰稿要求，对全稿进行了逐项、细致的校核与修改，修改内容已用蓝色字体在文中进行了标识，具体修改工作如下：

1. 参考文献及文中引用格式规范：

严格按 4.8 条要求采用了“著者-出版年制”。正文中外国文献仅保留姓氏，中文文献保留全姓名，出版年置于括号内；文后文献列表已严格按中文在前、英文在后排列，分别按汉语拼音和姓氏字母排序，并仔细核对了同一著者同年出版文献的 a、b、c 标注。

2. 字体字号与排版格式规范：

严格按 4.9 条要求调整了全文的字体字号及段落格式：中文题目改为小二号黑体；作者姓名改为小四号仿宋；作者单位、城市及邮编改为小五号宋体；摘要、关键词及分类号改为五号宋体；正文改为五号宋体（其中一级标题改为四号黑体，二级标题改为五号黑体）；英文部分各层级字体字号均按要求逐一调整。全文数字及英文字体已统一设置为 Times New Roman，行距调整为 1.5 倍。

3. 标点符号规范：

依据语文出版标准及期刊要求，对全文的中英文标点符号进行了全面核查，确保中文语境下使用全角标点，英文及数字语境下使用半角标点，修正了此前存在的混用与格式瑕疵。

再次感谢编辑老师和审稿专家的辛勤付出！

编委意见：同意发表