

《心理科学进展》审稿意见与作者回应

题目：亲社会努力影响奖赏加工的双阶段模型

作者：郑亚 杨子阳

第一轮

审稿人意见：

该稿围绕“自我成本在亲社会行为中的作用机制”这一具有重要理论与现实意义的议题展开，选题新颖且具有明确的问题导向。文章在充分梳理既有亲社会行为与奖赏加工研究的基础上，敏锐指出当前研究中对“他人收益”的关注相对充分，而对“自我成本”维度关注不足的问题，进而提出以“亲社会努力”为切入点的研究构想，具有较强的创新性与理论补充价值。作者进一步整合认知神经科学视角，提出双阶段模型以解释努力前后奖赏主观价值的动态变化，整体逻辑较为清晰。稿件符合“研究构想”栏目对前沿性与理论创新性的要求，整体质量较高。

回应：

非常感谢审稿专家对本文的高度评价与建设性意见。这些建议帮助我们提升了稿件的严谨性与理论深度。在返修稿中，我们根据审稿专家的建议，对全文进行了系统性的修订。正文所有主要的修改之处均以“红色”字体高亮标出，接下来我们将针对审稿专家的意见逐一进行详细的回应与说明。

意见 1：在文献综述部分，建议对“自我成本”相关研究作更系统的梳理，可适当区分不同类型的成本（如时间、体力、机会成本等）及其可能存在的差异性影响，从而为后续理论推演提供更清晰的出发点。此外，文中提出“亲社会努力影响奖赏加工的领域一般性/特异性机制”这一思路较具启发性，但其理论依据可进一步展开。

回应：

感谢审稿专家的宝贵建议。我们也意识到有必要对自我成本的这种异质性进行更为充分的回顾和讨论，以使本文的研究思路与理论推演更加清晰、完整。据此，我们在文献综述部分补充梳理了不同类型自我成本的相关研究，并在成本-收益框架下进一步讨论了不同成本形式之间可能存在的差异。修订内容如下（对应内容详见“2.1 亲社会行为的成本-收益框架”小节第 2 段）：

值得注意的是，成本-收益框架中的个人成本并非单一维度，而是包含多种不同形式。研究中常见的个人成本包括金钱成本、时间成本、机会成本、道德成本、疼痛成本以及努力成本等多种形式。不同类型的成本在可量化程度、主观厌恶性、情境依赖性及其所涉及的心理与神经机制上存在差异。具体而言，金钱成本因其具有较高的可量化性、便于在实验中精确操控等特点而最早被引入相关领域的研究，但其影响也容易受到个体社会经济地位和社会文化等因素的调节(Vanags et al., 2025)；时间成本和机会成本则更多体现为个体在帮助他人时放弃其他活动或潜在收益的代价，往往具有较强的情境依赖性(Persson et al., 2024)；道德成本和疼痛成本则分别涉及规范违背和身体伤害所带来的负性体验(Lockwood et al., 2025; Qu et al., 2020)；而努力成本主要表现为个体完成特定任务时在体力或脑力上的消耗程度(Contreras-Huerta et al., 2020)。从研究范式来看，现有基于成本-收益框架的亲社会行为研究主要集中于金钱、疼痛和道德成本。关于金钱成本的研究多采用经济博弈范式，例如考察个体是否愿意向慈善机构捐赠金钱(Engel, 2011)或付出金钱让他人获益(Sellitto et al., 2021)，这些研究普遍发现，个体在此类情境中倾向于优先考虑自身经济利益，呈现出以利己为主导的行为模式。相比之下，采用道德成本的研究则揭示了明显的亲社会倾向。当个体需要在经济获利与施加痛苦（如电击）之间进行权衡时，相比于电击自身，个体对电击他人表现出更强的厌恶，需要更高的金钱补偿才愿意增加他人所承受的电击次数，表现出超利他倾向(Crockett et al., 2014; Crockett et al., 2017)，甚至愿意直接替他人分担身体疼痛(Hein et al., 2010)。

感谢专家对我们关于“亲社会努力影响奖赏加工的领域一般性/特异性机制”这一思路的认可。这一思路是基于成本-收益框架下亲社会行为研究中成本维度与收益维度均存在异质性这一角度而提出的。根据审稿专家的建议，我们在问题提出部分新增了一个段落，对领域一般性/特异性机制的推演补充了更多实证研究以及理论细节。修订内容如下（对应内容详见“3 问题提出”小节第2段）：

第二，现有研究尚未探讨亲社会努力影响奖赏加工的领域一般性/特异性机制。从成本-收益的视角来看，亲社会行为由自我努力成本和他人收益共同决定。其中，努力成本主要包括体力努力和认知努力两种基本形式，而他人收益既会出现在帮助别人获得奖赏或避免损失的经济博弈领域，也会发生在帮助别人避免伤害的道德决策领域。由此引发的一个基本问题是，亲社会努力影响奖赏加工的神经动态整合过程是遵循领域一般性原则还是领域特异性原则？现有亲社会努力研究在努力成本方面主要关注体力努力，忽视了认知努力；在收益方面

聚焦于为他人赢得奖赏，忽视了他收益领域。这种聚焦于单一努力成本和收益领域的研究无法回答亲社会努力影响奖赏加工的领域一般性/特异性问题。因此，有必要从努力类型和收益领域两个维度系统考察亲社会努力对奖赏加工的影响机制。来自非社会领域的研究表明，体力和认知努力存在既重叠又分离的行为和神经机制(Bustamante et al., 2023; Chong et al., 2017; Jiang & Zheng, 2023)。另外，个体的亲社会性会受到收益领域的影响：相较于赢钱领域，人们在损失或伤害领域表现出更多的亲社会行为，甚至愿意牺牲自我利益以避免对他人造成损失(Sellitto et al., 2021)或伤害(Liu et al., 2020)。这些研究表明亲社会努力和奖赏的动态关系在不同收益领域呈现不同模式。总之，亲社会努力对奖赏加工的影响机制如何受到成本类型（体力努力、认知努力）和收益领域（奖赏获得、损失回避、伤害回避）的调控仍不明确。

意见 2: 在关于努力折扣（如内部成本理论、机会成本理论）与努力回溯性提升奖赏价值（如努力合理化理论、心理对比理论）的解释路径，现有论述主要借鉴于非社会决策领域的研究。建议作者进一步考虑在亲社会决策情境中对这些不同理论框架加以区分与整合，明确其适用边界及可能的差异性预测。

回应:

感谢审稿专家的宝贵建议！关于社会特异性的问题，一直是社会认知神经科学和心理学领域关注的核心议题之一，亲社会努力研究同样面临这一问题。为更好地区分并整合源自非社会决策领域的相关理论，我们在修订稿中进一步论述了非决策领域研究的预测，内部成本理论和机会成本理论更适合解释努力执行之前的前瞻性价值决策，而努力合理化理论和心理对比理论则更适合解释努力完成之后的结果评估过程；同时，这些理论在亲社会情境下的作用还可能受到受助对象特征、社会距离、共情水平及道德意义等因素的调节。修订内容如下（对应内容详见“2.2 努力与奖赏之间存在悖论关系”小节第3段）：

然而，上述关于努力折扣与努力增值的理论解释及实证研究主要来自非社会决策领域，其能否推广至亲社会情境尚不清楚。在非社会决策中，个体权衡的是自身所获奖赏与所付努力之间的关系；而在亲社会情景中，努力所换取的结果由他人获得，个体的价值计算不仅需要整合努力成本，还需要对他人收益赋予主观社会价值(Ruff & Fehr, 2014)，即从追求自我收益最大化转向为他人付出努力以帮助其获益的社会目标(Lockwood et al., 2020)。基于这些考虑，利己和利他努力在努力与奖赏的关系上可能存在以下两方面的差异。第一，从理论预测

来看，不同理论框架对利己努力与利他努力的预测存在分离。就前瞻性努力贬值而言，内部成本假设对两类努力的预测相似，因为无论受益对象是自己还是他人，努力本身产生的内部成本是类似的。然而，机会成本假说的预测可能不同：为他人付出努力时，个体放弃了将同等努力用于自身收益的机会，因此利他努力会伴随更大的机会成本。就回溯性努力增值而言，两种理论的预测同样存在差异。努力合理化理论强调个体对自身付出的意义建构与认知合理化，是一种晚期的、自上而下的加工过程；当受益对象为他人时，个体缺乏直接的自我相关动机来合理化付出，因此利他努力在回溯性阶段可能不会产生增值效应，甚至可能出现努力贬值。相反，心理对比理论描述的是一种自动的、自下而上的加工过程，这种努力与奖赏之间的对比可能在价值评估的早期阶段以不区分受益对象的方式进行，从而利己与利他努力产生相似的增值效应。第二，从边界条件来看，亲社会努力涉及到对他人收益赋予主观社会价值，因此可能表现出更大的情境敏感性，更容易受到社会因素的调节。首先，社会距离是影响亲社会行为的重要因素，人们通常对与自己的群际关系（内外群体）或心理距离（亲疏关系）更近的他人表现出更多的亲社会行为(Jones, 2021)。其次，个体层面的共情能力和互惠影响也可能调节亲社会努力的价值评价。共情-利他假说认为，对他人困境的共情会诱发减轻或消除其困境的动机，从而促进利他行为(Batson et al., 1991)；而互惠理论则指出，善意行为会通过社会规范等机制诱发受助者的回馈行为(Gouldner, 1960)。共情水平较高或互惠预期较强的个体，更可能倾向于为他人努力，并由此将获得的奖赏赋予更高的主观价值。综上，亲社会情境下的努力-奖赏之间的关系可能受到社会距离、共情能力和互惠信念等多种社会因素的调节，表现出比非社会情境更大的可塑性。

意见 3：决策前的努力折扣具有较为明确的行为学含义，即可能降低个体做出亲社会选择的概率。相比之下，决策后由努力引发的奖赏价值回溯性提升，其行为后果仍有待进一步澄清。例如，这一过程是否会增强后续类似情境中的亲社会倾向。建议作者在后续研究中结合具体实验设计或数据分析，对上述过程加以更系统的刻画与检验。

回应：

感谢审稿专家的宝贵建议！我们完全认同，相较于决策前努力折扣对亲社会选择概率的直接影响，努力执行后所引发的奖赏价值回溯性变化，其行为学后果确实尚缺乏充分检验。审稿专家提出从后续类似情境中的亲社会倾向变化来考察这一过程，具有重要的启发作用和现实意义。我们将在后续实证研究中结合具体实验设计和数据分析，系统考察努力执行后的奖赏价值变化是否会对后续亲社会决策产生持续性影响，从而更全面地揭示该过程的行为学

意义。例如，未来研究可以采用序列决策范式，考察个体在为他人付出努力并获得反馈后，是否会在后续相似任务中表现出更高的亲社会选择概率或更大的努力投入；也可以结合试次间分析，检验前一试次中奖赏评估的神经反应是否能够预测下一试次的亲社会决策或努力程度。鉴于本文的研究构想定位，我们未能在正文中对这一问题展开论述，但我们会在后续实证研究中进行上述分析。再次感谢审稿专家的建设性意见。

第二轮

审稿人意见：作者已充分、细致地回应了我提出的所有问题，无其他建议。

编委意见：同意发表，无其他建议。