

《心理学报》审稿意见与作者回应

题目：共赢促进合作的认知计算机制：互惠中积极期望与社会奖赏的作用（亲社会行为专刊）

作者：吴小燕，付洪宇，张腾飞，鲍东琪，胡捷，朱睿达，封春亮，古若雷，刘超

第一轮

审稿人 1 意见：

该研究采用了心理学实验设计以及计算建模，并进行了两个实验，旨在揭示人们在实时互动过程中的心理计算机制，并展示了积极期望（即个体对他人往后合作行为的预期）以及社会奖赏在促进共同合作中的重要作用。研究中所使用的分析方法、模型验证以及实验设计都表现出高度的专业和严谨性。研究结果和配图呈现清晰。两个实验的结果一致地展示了相同的计算心理机制，以及积极期望和社会奖赏的重要作用。具体来说，合作对象的高概率合作行为会提高被试对其积极期望和社会奖赏的评估，从而促使被试更积极地进行合作，从而实现双方的合作。作者也充分考虑到实验 1 的结果可能会受到合作对象的合作行为变动性的混淆影响，因此进行了实验 2 以排除可能的混淆变量。总的来说，该研究的结论具有很高的可信度，反映了该学术领域的最新进展和水平。研究目的明确，数据可靠，条理清晰，与期刊要求相符，建议发表。我对本文的细节方面有以下一些建议：

意见 1：摘要中去掉本研究以及 $n=135$ ， $n=104$ 统计数字。

回应 1：非常感谢审稿专家的意见！我们在修改稿中已去掉“本研究”以及“ $n=135, n=104$ ”统计数字。

意见 2：统一用阿拉伯数字表示，各级号码之间加一小圆点，末尾一级不加小圆点。如一级用“1”、“2”、“3”，二级用“1.1”、“1.2”、“1.3”，三级用“1.1.1”、“1.1.2”、“1.1.3”。标题仅限三级。

回应 2：非常感谢审稿专家的意见！我们已在正文中将各级标题编号进行了修改。

意见 3：作者在自检报告中提到“本研究应用线性混合模型分析数据。线性混合模型的各变量在拟合模型之前均做为标准化（z 分数转换），这样我们所报告的 β 值为标准化系数（Standardized beta coefficients）”。请在中文的相应部分也对 z 分数标准化处理进行说明。

回应 3：感谢审稿专家的意见，我们已在 2.1.4 和 3.1.3 部分对 z 分数标准化处理的相应内容进行了补充说明。

意见 4：表 1 收益矩阵表的 partner 应居中，另外请考虑采用中文。图片全部统一为英文也可以，但对 subject 和 participant 请注意统一用词。

回应 4：感谢审稿专家的意见，我们已修整了表 1 收益矩阵。考虑到全文用词统一，方便读者理解对应表格和图片展示的结果，我们统一采用了英文表达，全文用 subject 表示被试。

意见 5：作者总共比较了 6 个模型，这 6 个模型看起来是有先后逻辑的关系。例如从基础的随机选择假设模型开始，逐次递进模型更加复杂，在计算建模部分建议增加对所有模型有一句话的介绍。

回应 5：感谢审稿专家的意见，我们已经在实验 1 的方法部分-计算建模对各个模型进行了简要介绍，阐明了我们构建模型的策略和原则。详见 2.1.5 部分的蓝色字体。

.....

审稿人 2 意见：

与同一个人进行重复互动，是我们在生活中经常面临的问题。本研究采用了计算建模的形式，同时探讨合作中的积极期望和社会奖赏。在对积极期望的模型构建上，还将一阶和二阶信念纳入到计算建模，其最优模型对动态合作过程进行了较为完整的刻画。该研究具有较强的创新性，为合作互动行为提供了更为完整的解释，是非常有意义的研究。

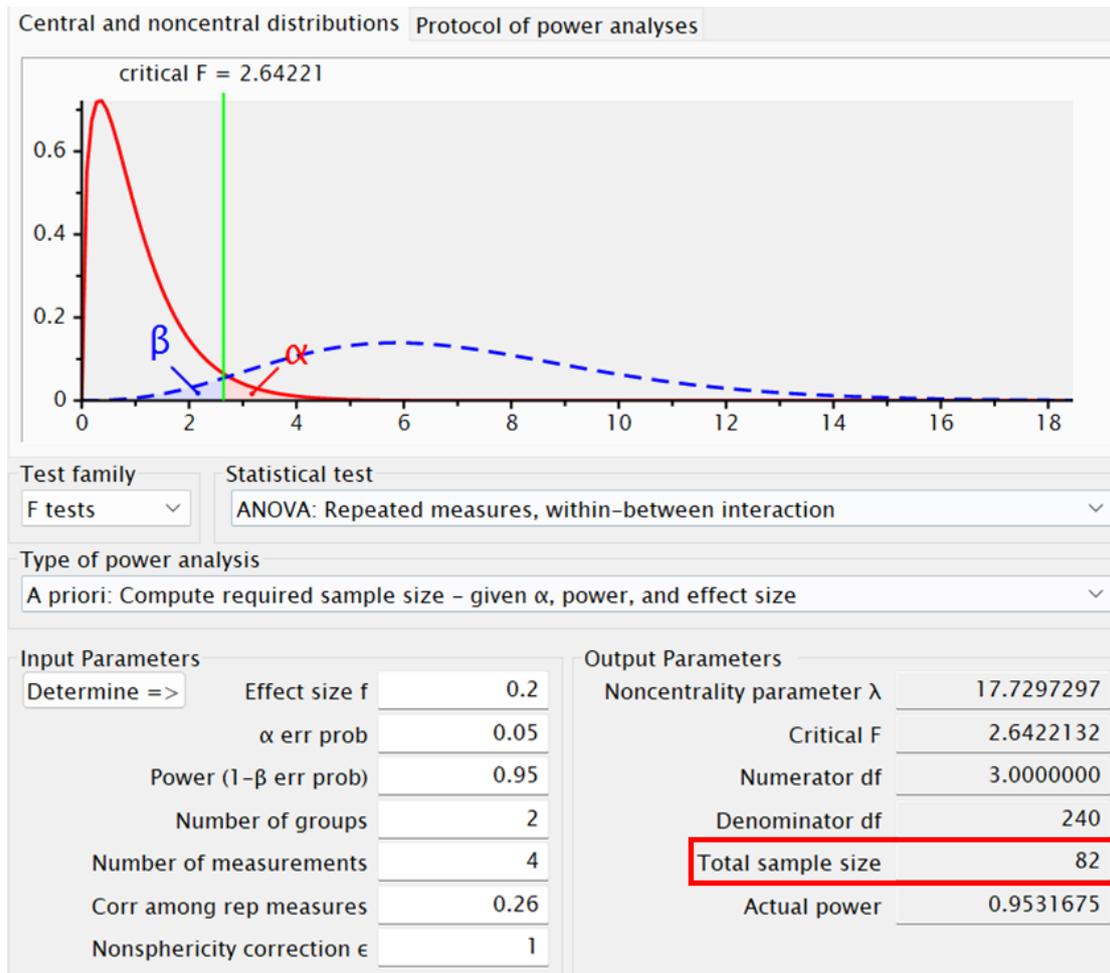
然而，研究中有诸多地方阐述不够详尽，对于数据和结论之间对应的关系我也存在疑问，使我对该研究结论的可靠性存在质疑。下面，我将一一列出我的问题，希望研究者能够进行详尽解答。

意见 1：被试数量。研究中实验一收集 135 名被试数据，实验二收集 104 名被试数据，是不小的被试数量。研究者在两个实验的方法部分都对被试数量进行了说明，进行了 G Power 测算。鉴于研究结果主要着眼于模型拟合的参数结果，需要具体汇报进行的何种敏感性功效分析以支持被试数量的合理性。

回应 1：感谢审稿专家的指正，我们已对稿件进行了相应修改。具体而言，我们依据所关心变量对被试合作概率的交互作用，进行了样本相关的分析。实验 1 的样本量参考了采用类似学习任务和计算模型探讨个体学习动态的现有研究（见以下表 1），同时兼顾了样本获取的便利性。未事先使用任何软件工具进行计算。对于实验 2，我们在事前采用 G power 确定了所需样本量，针对我们关注的交互作用——情境、波动性与概率，计算得出最小样本量应为 82 人（见以下图 2）。根据审稿专家的意见，我们已在文章的方法部分相应地进行了报告，具体见文章 2.1.1（实验 1）和 3.1.1（实验 2）部分。

参考文献	样本量
(Henco et al., 2020)	151
(Fang et al., 2024)	86
(Mukherjee et al., 2020)	128
(Bradshaw & McGettigan, 2021)	109
(Hackel et al., 2020)	169

表 1. 实验 1 样本参考文献。



回应说明图 1. 实验 2 的事前样本量计算。

意见 2: 计算建模构建。认知计算建模是研究的一大优势，需要得到更为详尽的解释才能使读者理解其合理性及其意义。

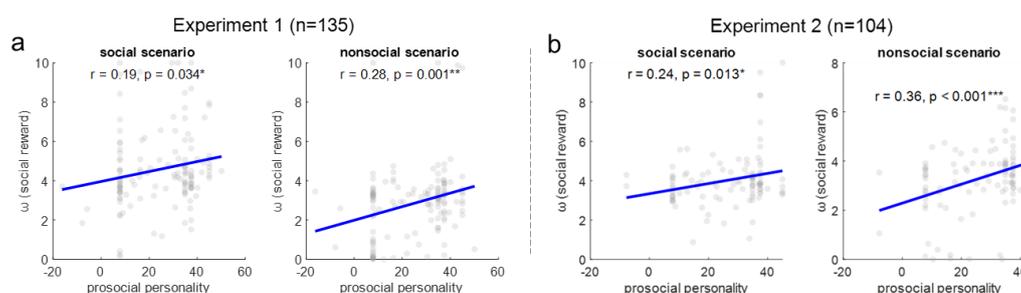
研究对于模型的构建，主要源于前沿中提出的假设以及相应的前人研究的模型形式。而在本研究所应用的多轮互动游戏中，研究者没有提供任何指标（被试主观汇报，或者回归分析结果？）指向当前关于预期和社会奖赏的认知模型。这是我对本研究的最大质疑。认知建模的自由之处在于，研究者可以进行各种可能模型假设，只要有一定的合理性，都能一定程度解释行为。而本研究使

用了六个模型，如果只从模型驱动的角度考虑，模型数量并不算多，应该还有挺多备择模型可以选择。

回应 2：感谢审稿专家意见。该问题的回应分为两个部分：

(1) 关于其他指标。主观汇报会为被试带来一定的暗示性和其他可能的影响。即使我们向被试提出一个问题去让被试主观报告其社会奖赏或愉悦程度，也容易收到社会赞许性的干扰。而通过计算模型，我们可以推测出被试的“社会奖赏”，这正是使用模型的优势。并将这一心理成分纳入模型，是有助于更好地解释被试的行为。此外，在文章中我们提供了相应的回归分析结果。在行为数据分析中，广义线性混合模型的分析显示对手的上一个回合能够显著地影响被试当前的行为，这表示被试会根据合作者的行为动态地调整自己的行为。在基于模型的分析中，以认知模型中所得到的动态预期以及社会奖赏为因变量构建的线性混合模型，能够帮助我们探索不同的实验条件如何影响个体的心理计算成分。

在审稿专家的建议下，我们还发现社会奖赏参数值和个体的亲社会人格（social value orientation）分数显著正相关（如下图）。这表明我们从认知建模所拟合得到的社会奖赏成分能够一定程度上反映个体的亲社会倾向，为认知模型所得到的社会奖赏部分的心理意义提供支持。根据审稿专家的宝贵建议，我们已将这部分分析结果新增至文中相应的结果部分，并将图片新增为附图 5。



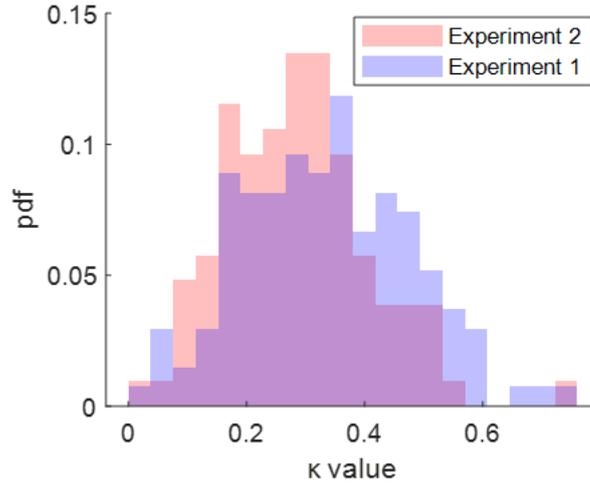
回应说明图 2. 社会奖赏与亲社会人格得分相关分析。

(2) 关于其他的备择模型。我们同意您的意见，从模型驱动的角度，模型的数量可以是无穷大的。然而，在没有充分证据和研究假设驱动下，尝试更多的模型未必能够为解释本研究的主要发现提供更多的帮助(Wilson & Collins, 2019)。同时考虑到现实因素（时间、精力成本等），我们采取的策略是结合我们感兴趣的问题和经典理论，来进行模型的构建。本研究从现有研究和假设驱动的角度出发而构建模型。我们的研究问题是回答条件合作的认知计算机制。根据已有的研究，影响个体做出条件合作的可能有两个主要的因素，分别是合作期望（即个体对他人合作的期望）以及双方合作共赢带来的额外奖赏（社会奖赏）(Pal & Hilbe, 2022)。我们通过系统地构建模型，用模型来构建我们关心的心理成分，通过模型比较和数据分析，最终根据结果给出解释和结论。我们比

较的 6 个模型能够带来很多确定的、有意义的信息。我们构建模型的思路先从最简到繁，先从简单的模型开始再到相对复杂的虚拟博弈(fictitious play)假设模型，我们从中择优再逐步优化模型。首先，从基线模型 1 开始，我们发现模型 2 基于奖励学习的假设以及模型 3 基于理性经济人的假设，并不优于基线模型 1，这暗示了模型 2 和模型 3 的方向可能是错的。模型 4 基于社会奖赏假设要显著优于没有社会奖赏假设的模型 3，这表明社会奖赏假设有一定的合理性。随后我们再模型 4 的基础上再增加不同的假设，比较了两种不同的学习模型，分别为一阶信念（模型 5）和二阶信念（模型 6）。至此，我们发现模型 6 在现有的模型中最优。同时，我们通过模型比较发现当前最优模型的预测能够很好地吻合观测的数据，并且能够很好地描述和刻画被试的行为模式。重要的是，当前模型所提供的证据已能够回答本文的研究问题。根据两位审稿专家的意见，我们文中 2.1.5 认知计算模型的这部分进行了修改，阐明了我们构建模型的策略和步骤。

意见 3: 关于最优模型的设置问题。（1）我对二阶信念的更新形式感到难以理解，为何被试实际的选择和被试推测合作者认为被试的合作概率之间的差越大，被试下一试次的合作概率就一定应该增加？换言之，为何 κ 只能大于 0,它可能小于吗？我观察到该参数在实际预测中有较多被试的估计值都在 0 附近。我能够理解研究将之解释为强化学习的更新形式，将 κ 限制在[0,1]之间，但是只能是这样的可能吗？

回应 3: 感谢审稿专家的提问。 κ 对应的二阶信念是指被试在模拟对方的心理过程。 Q_t 表示参与者在第 t 回合的实际选择，合作时 $Q_t=1$ ，背叛时 $Q_t=0$ 。 q'_t 代表参与者估计对手认为其合作的概率。因此， $\left(Q_t - q'_t\right)$ 反映了被试心理模拟的对方预测的合作概率与被试实际行为之间的预期差异。如果此值大于 0，表明对手低估了被试的合作意愿；反之，则表明对手高估了该意愿。 κ 用于指示参与者更新其二阶信念预期误差的程度，其理论值域为 0 到 1，0 表示完全不考虑预期误差，而 1 表示完全考虑，这与一阶信念的 α 参数处理方式相同。这个机制的逻辑基础是，如果被试认为自己合作了，那么对手也会提升对自己未来继续合作的预期（正如被试自己所应用的模型的一阶信念部分），进而提升对手的合作意愿；反之，如果自己背叛了则预计对手的合作意愿也会相应地降低。这一假设参考了最初提出影响模型（influence model）的研究（Hampton et al., 2008），符合条件合作理论(Keser & Van Winden, 2000)。超出这个值域不适用当前情景。在本研究中的值如果低于 0，则意味着参与者在认为自己的合作表现超出对手预期时，预测未来对手的合作概率反而会降低，我们认为这与常理以及与条件合作理论不符。 k 值接近 0 表示个体在学习过程较少地考虑到二阶信念去推测对方的合作概率。对于同一位被试，在不同的条件下 k 值均接近 0 的非常少（请见下图）。我们的实验中所有的被试的 k 值均大于 0。



回应说明图 3. κ 值的分布。纵坐标为概率密度函数（Probability Density Function）。

意见 4: 关于最优模型的设置问题。（2）此外，关于 q_t 的更新形式较为复杂，其中包含了社会奖励和被试的合作概率？它是如何推导出来的，其背后的心理意义是什么，请研究者给出详细的解释，不然，读者难以理解和判断其合理性。

回应 4: 感谢审稿专家提问，以下将详细展开对此项的推到和心理意义解释。 q_t 表示被试推测的对手（这里表示合作的另一方，即文中的合作伙伴）对自己合作概率的预期。例如“我觉得他应该会觉得我很愿意与他合作”，这是我从别人的视角去推测他们关于我的想法。这种复杂的心理功能，在社会决策中通常将其称之为“二阶信念” (Hill et al., 2017; Wu et al., 2020)。在模型中其形式由被试的合作概率方程 $q_t = \frac{1}{1 + e^{-\beta (U_d(t) - U_c(t))}}$ 决定。将 U_d 和 U_c 代入此公式得到：

$$q_t = \frac{1}{1 + e^{-\beta (4p_t + 2 - p_t (4 + \omega))}}$$

根据这一公式，我们先通过推公式，得到 p_t （预期对手合作概率）与 q_t （自己合作行为概率）的关系式。

步骤 1: 首先，将方程重写为：

$$1 + e^{-\beta (4p_t + 2 - p_t (4 + \omega))} = \frac{1}{q_t}$$

步骤 2: 然后，移项得到指数表达式：

$$e^{\beta (4p_t + 2 - p_t (4 + \omega))} = \frac{1}{q_t} - 1$$

步骤 3: 对两边取自然对数:

$$\beta (4p_t + 2 - p_t (4 + \omega)) = \ln\left(\frac{1}{q_t} - 1\right)$$

步骤 4: 解出 p_t :

$$\beta (4p_t + 2 - 4p_t - \omega p_t) = \ln\left(\frac{1 - q_t}{q_t}\right)$$

$$\beta (2 - \omega p_t) = \ln\left(\frac{1 - q_t}{q_t}\right)$$

$$2 - \omega p_t = \frac{\ln\left(\frac{1 - q_t}{q_t}\right)}{\beta}$$

$$p_t = \frac{2\beta - \ln\left(\frac{1 - q_t}{q_t}\right)}{\beta \omega}$$

由此，我们得到关系式。我们假定对手也应用同样的模型，即相同的主观价值函数，那么该公式也适用于对手用来推测被试自己的合作概率。而 q_t 在对手的角度应为对手自己的合作概率，在我们的模型中对手的合作概率表达为 p_t 。而被试从对手的角度预期被试的合作概率 p_t ，我们用 q_t' 来表示，由此我们得到：

$$q_t' = \frac{2}{\omega} - \frac{\ln\left(\frac{1}{p_t} - 1\right)}{\beta \omega}$$

以上最终公式为文中所报告。该 Influence 模型由(Hampton et al., 2008)提出。随后被应用社会互动决策的研究中(Hill et al., 2017)。我们对该模型进行了调整以适应当前任务和价值函数。根据审稿专家的意见，我们已将这部分补充到附加第一部分，以便读者查阅。

意见 5: 关于最优模型的设置问题。(3) 在计算合作主观价值 U_c 时, 模型纳入了社会奖励这一参数。关于这个参数的估计结果, 让我感到非常意外, 被试在双方都合作的情况下实验设置能够得到的金钱奖励是 4, 但是合作本身带来的社会奖励 ω 平均值接近(图 2.d,图三.e)。在看每个被试的估计值时有相当多的被试都大于 4, 甚至到 10。也就是说, 社会奖励(按研究者所解释, 合作带来的愉悦度)要等同于甚至大于被试能够实际得到的金钱? 应该如何理解这样的现象? 同样是合作带来的愉悦度, 为什么低合作概率的时候合作带来的愉悦度显著降低了? 直觉上似乎在不常合作的情境中促成了双方合作, 带来的愉悦感更大? 既然是社会奖励, 为何在 non-social 情境下其估计均值依然在 3 左右? 这些都希望研究者给出合理的解释。

回应 5: 感谢审稿专家提问。对于第一个问题“社会奖励要等同于甚至大于被试能够实际得到的金钱? ”。现有的模型和结果不能支持这一结论。模型中合作的主观价值 $U_c = p(4 + \omega)$, 背叛的主观价值 $U_d = 4p + 2$ 。被试根据两个选项的主观价值差值做出决定, $U_c - U_d = p\omega - 2$ 。这里 $4p$ 在计算过程抵消, 影响被试合作决定的是 $p\omega$ 的乘积与 2 相比。当 $p\omega > 2$ 时, 被试选择合作的概率增加, 反则亦反。因而将 ω 与 U_c 函数中的绝对值 4 直接对比是没有意义的。举个例子, 假设我们改变收益矩阵如下:

		合作者	
		合作	背叛
被试	合作	10	0,12
	背叛	12,0	2,2

在这里收益矩阵下, 按照我们的模型 6, $U_c = p(10 + \omega)$; $U_d = 10p + 2$; $U_c - U_d = p\omega - 2$ 。这与当前实验的收益矩阵一致, 即被试会采用同样的计算过程而做出决策。那么在这个例子中, 如果我们将 ω 直接与 10 比较, 而得出“社会奖励一直低于/或不高于被试能够实际拿到的金钱”的结论是不可行的。

对于第二个问题“同样是合作带来的愉悦度, 为什么低合作概率的时候合作带来的愉悦度显著降低了? ”。按照条件合作理论, 当对方不合作时我便也不合作, 低概率会使得合作概率相应减少 (Keser & Van Winden, 2000; Suzuki et al., 2011)。这种减少有可能是由于预期降低, 也有可能是由于社会奖赏(即合作带来的主观奖赏价值)降低, 或两者同时产生影响。本研究对这一问题进行探索, 最终我们发现两者皆对影响合作行为起到作用。即当个体感知到对方合作意愿低时, 对其预期相应也降低, 双方合作共赢所带来的主观奖赏也较低, 进而个体也倾向于不与对方合作。审稿专家提到“直觉上似乎在不常合作的情境中促成了双方合作, 带来的愉悦感更大? ”。这里更多指向预期与正向预期误差 (positive prediction error)。当对方合作意愿非常低时 (p_t), 突来的合作 (P_t) 造成了非常大的正向预期误差, 即 $(P_t - p_t) > 0$ 。这会使得我们调整对下一次合作有更加积极的期待 (p_{t+1} 提升)。而 p_{t+1} 的提升进而增大了合作的主观价值, 进而促进被试

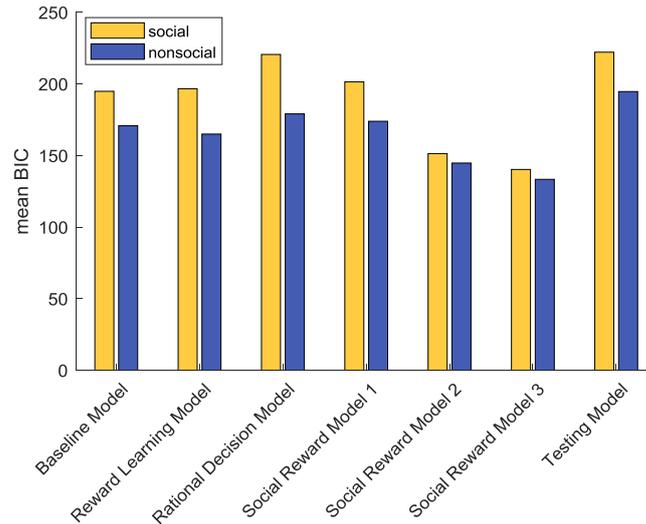
下一次选择合作。我们的结果显示，当明确地预期对方不可能合作时（相比于对方高概率合作的情形），双方合作成功所带来的社会奖赏是较低的（见正文图 2d 和 3e 的社会条件）。

对于第三个问题，“为何在 nonsocial 情境下其估计均值依然在 3 左右？”。社会奖赏 ω 的绝对值大小对于理解心理意义作用不大，在研究中我们主要对比不同的实验条件下的相对大小。回到 $U_c - U_d = p\omega - 2$ 这一公式， $\omega=3$ 时表示，只有 $p>0.67$ 时（即被试预期对方合作的概率达 67% 及以上时）合作的主观价值才能大于背叛的主观价值，此时被试才会选择合作。当 ω 接近 2 时，只有极其少数的情况下（即预期对方的合作概率接近 100% 时）被试才会选择合作。在我们实验中，被试在 nonsocial 条件下的合作行为虽然低于 social 条件，却也不为 0。在实验中，合作者的合作概率上下浮动，被试的合作行为也跟随变化。我们的模型对每位被试的数据进行拟合，找到最优的 ω 值以解释被试的数据。当 ω 接近 3 时，表示被试预期对方合作的概率较高时，被试与对方合作的概率增加。对于 $\omega=10$ 的情况，即 p 只要 >0.2 ，合作的直观价值便已大于背叛的主观价值。 ω 接近 10 的情况在我们实验中非常少见，在实验 1 中只有 4 位被试（4/135），实验 2 中没有。

总得来说，我们的结果表明在 nonsocial 条件下，相比于 social 条件，被试选择合作的条件要求更高。这也是我们为何观察到，在两个实验中，被试在 nonsocial 条件中的合作行为远低于 social 条件。根据审稿专家的宝贵问题，我们已对文中相应的结果解释和讨论部分进行探讨和补充。

意见 6: 如果将 U_c 的形式稍微改变为 $U_c(t)=4*p_t+0*(1-p_t)+\omega$ ，模型拟合效果可能类似？但其意义就不再是合作带来的社会奖励，而是被试在特定情境中的基本合作偏好。

回应 6: 感谢审稿专家的建议。根据专家建议，我们尝试了这一模型。 $U_c(t) = 4 \times p_t + 0 \times (1 - p_t) + \omega$ 化简后为 $U_c(t) = 4p_t + \omega$ 。 $U_c(t)$ 与 $U_d(t) = 4p_t + 2$ 相比， p_t 相互抵消不起到影响被试决策的作用。因此，该模型的预测是当 $\omega>2$ 时被试倾向所有情形中都会选合作， $\omega<2$ 时，被试均选择背叛， $\omega=2$ 时，被试完全随机选择合作和背叛。在该模型的假设下，只存在这三类被试，这与我们实际观察到的数据不符。我们将这一模型（放在最后的 Testing Model）加入后模型比较结果如下图。根据以上，我们选用原先所报告的最优模型。



回应说明图 4. 加入测试模型后的模型结果比较。

意见 7: 研究汇报最优模型预测的被试合作率与被试实际的合作率在各个实验条件相关性都达到了 $r=0.98$ 以上。计算模型预测与被试实际性行为的相关其实并不是一个很好的选择, 更好方式是模型预测的合作率与被试实际合作率的接近程度, 甚至可以进行统计性检验其差异。同时, social 和 non-social 情境的表现一样好吗?

回应 7: 针对审稿专家的问题, 我们在重新评估了原始计算方法后认识到, 通过对数据进行平均化处理, 可能导致了数据的失真, 从而无法准确对比在不同条件下, 认知模型预测的合作行为(以下简称“预测行为”)与参与者实际选择的合作行为(以下简称“实际行为”)之间的差异。根据审稿专家的建议, 我们采用了线性混合模型对预测行为与实际行为之间的差异进行了系统性的统计检验。我们对文中所用的模型中, 加入一个名为“prediction”的二分类变量。该变量对预测行为标记为 1, 实际行为标记为 0。在这个模型中, “prediction”变量的主效应及其与其他变量的交互作用能够检验预测行为与实际行为之间可能存在的偏差。

实验 1 的结果显示, “prediction”变量的主效应 ($\beta = -0.03$, 95% CI = [-0.14, 0.07], $p = .536$) 以及所有交互效应均不显著 (cooperation probability \times prediction: $\beta = 0.06$, 95% CI = [-0.07, 0.20], $p = .378$; scenario \times prediction: $\beta = -0.05$, 95% CI = [-0.18, 0.08], $p = .487$; cooperation probability \times scenario \times prediction: $\beta = 0.01$, 95% CI = [-0.17, 0.20], $p = .875$)。

实验 2 的结果显示, “prediction”变量的主效应 ($\beta = 0.06$, 95% CI = [-0.06, 1.18], $p = .322$) 以及所有交互效应均不显著 (prediction \times probability: $\beta = 0.05$, 95% CI = [-0.12, 0.22], $p = .546$; volatility \times prediction: $\beta = -0.07$, 95% CI = [-0.22, 0.07], $p = .321$; probability \times volatility \times prediction: $\beta = 0.02$, 95% CI = [-0.19, 0.22], $p = .994$; probability \times scenario \times prediction: $\beta = 0.06$, 95% CI = [-

0.14,0.27] , $p=.551$; volatility \times scenario \times prediction : $\beta =0.03$, 95%CI=[-0.17,0.24] , $p=.748$; probability \times volatility \times scenario \times prediction : $\beta =-0.02$, 95%CI=[-0.31,0.26] , $p=.887$) , 这包括 prediction \times scenario 交互项的效应也不显著 ($\beta =-0.06$, 95% CI = [-0.20 , 0.08] , $p = .418$) 。

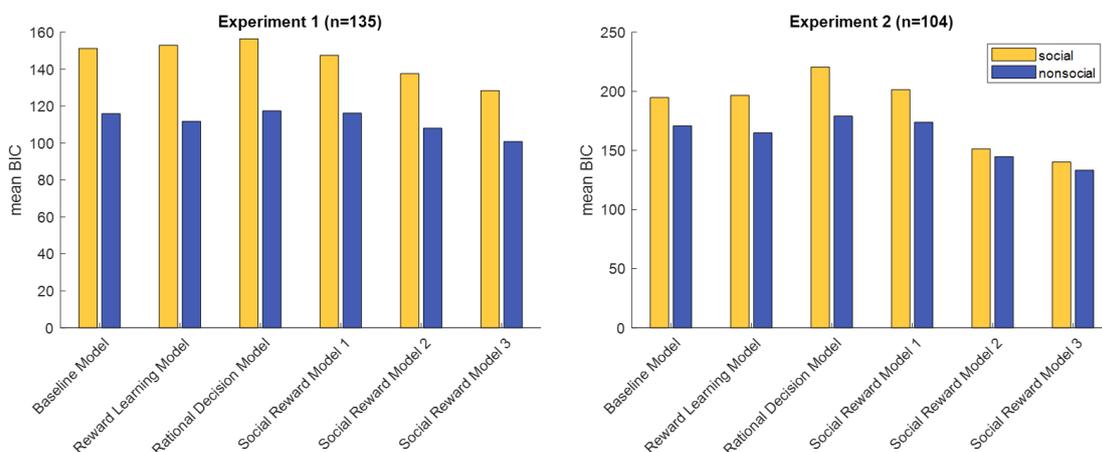
根据这些结果，我们得出结论：认知计算模型所预测的合作行为与我们实际观测到的合作行为在统计学上没有差异。在社会（social）和非社会（nonsocial）情境下，模型的预测效果没有统计学上显著的差异，即在这两种情境下，模型预测的效果具有相同的准确度。我们已将以上分析结果更新到附加的模型验证部分。

意见 8: 关于模型比较方式研究者需要阐述清楚具体的计算形式和合理性。研究比较的模型使用参数数量并不相同，似乎更合适方式的是使用包含参数数量和试次数量惩罚的 BIC 而不是 ΔAIC ，相应地，PEP 的计算也应该是 BIC 惩罚后的结果。

回应 8: 感谢审稿专家的建议。我们使用的是考虑了参数数量样本大小的矫正过后的 AIC，简称为 AICc (Hurvich & Tsai, 1989)，其公式为：

$$A I C c = -2 \log(L) + 2K + \frac{2K(K+1)}{n-K-1}$$

与 AIC 相比，AICc 考虑了样本大小和参数数量的比例，减少了在小数据样本情况下选择过于复杂模型的倾向，使得 AICc 在数据样本量较小时比 AIC 更加可靠，这是我们选择 AICc 的原因。根据审稿专家的意见，我们现将改用 BIC，其模型比较结果不变（如下图）。正文相对应的内容和图片已修改。根据审稿专家的意见，模型恢复的分析也改用为 BIC 指标，文中相应的内容和图片已更新。



回应说明图 5. 基于 BIC 的模型比较结果。

意见 9: 最优模型估计的对合作者的合作预期 p_t 与实验设计的真实预期之间接近程度如何？我观察到在 social 情境拟合度较好，但在 non-social 情境下，模型估计的 p_t 与实验设计的差异较大(附图 5)。考虑到这是一个用强化学习形式模型估计的结果，对于一个学习序列来说，似乎被试根本不在乎电脑到底是选择合作还是背叛，不对其进行学习？这难以理解，特别是在两个实验中都有 60 或 80 个稳定概率的合作试次。说明这个模型并不能较好的解释被试在 non-social 下的行为？

从实际被试结果来看，被试确实在 non-social 下的合作行为更少，是因为被试预测电脑的合作概率更低？

回应 9: 感谢审稿专家的问题。我们的结果显示，无论是 social 还是 nonsocial 条件，被试均会进行学习。从模型比较和模型预测的结果，我们发现我们的模型能够很好地解释 social 和 nonsocial 情况下被试的行为决策。其次，正如审稿专家所说，被试在 nonsocial 条件下的预期要比 social 条件低。请见图片 2.e (实验 1) 和图片 3.f (实验 2)。这表明尽管合作者的行为完全一致，但当被试认为对方是电脑时，对对方的预期要低一些（而非不学）。同时也表明被试是在乎电脑的选择的，他们会根据电脑的反应动态地调整行为策略。对于这一结论，除了模型比较的结果（即有学习假设的模型优于无学习假设的模型），线性混合模型也显示合作者的前 1 个回合的选择显著地影响被试当前的行为。

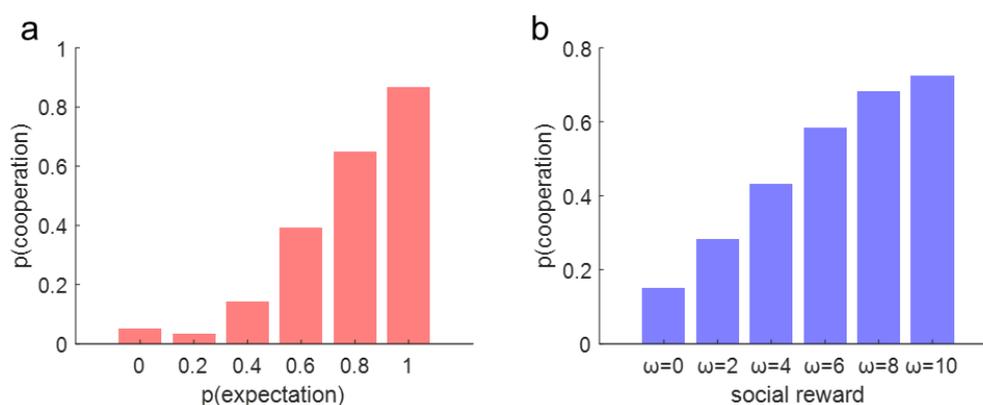
最后，审稿专家提问到的“从实际被试结果来看，被试确实在 nonsocial 下的合作行为更少，是因为被试预测电脑的合作概率更低？”，正如审稿专家所说，这是一方面的原因。另一方面的原因是社会奖赏影响的。从我们的结果图 2 和图 3，结合 social reward 和 expectation 的结果来看，被试在 nonsocial 条件下的合作行为更少，是由较低的社会奖赏（即与计算机合作所带来的主观奖赏感受少一些）以及对计算机较低的预期两者共同决定的。感谢审稿专家的问题，这对我们解释结果有非常大的启发。我们根据审稿专家的提问，完善了文章的讨论。

意见 10: 结果解释与结论合理性。研究者在摘要中提到，“个体合作行为的增加由积极期望（即合作成功使得个体对合作者建立了积极的期望）与社会奖赏（由合作本身带来的额外奖赏）所共同驱动”，这个结论的数据证据是什么？研究者只是对模型估计的社会奖励和期望进行了回归模型分析，即对不同条件下的参数值进行了统计检验，这与合作行为增加的关系是什么？我设想一个可行但较为粗略的方法是使用回归检验模型估计参数比如社会奖励和期望对被试合作比例是否都存在独立且正向的影响。

回应 10: 感谢审稿专家提问。支持摘要中的结论的证据来源于图 2 和图 3 的结果，对于结果的解释来源认知模型的计算机制。模型 6 的框架里，较高的预期或较高的社会奖赏，都能够提高被试对合作的主观价值计算，从而提升合作行为。从认知模型的主观价值函数（见公式 1 和公式 2），我们可以看到预期 (p) 和社会奖赏 (ω) 的增加都能够增加合作的主观价值 (U_c)，进而提高被试

选择合作的概率（见公式 5）。在认知模型的假设下，我们通过被试的真实选择来拟合模型。通过对比不同的模型来验证不同的结果，最终给出结论。

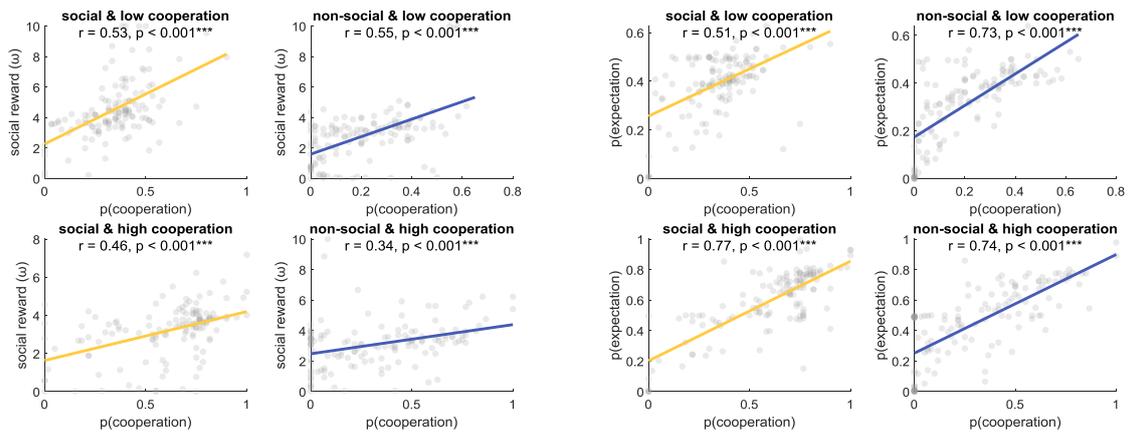
根据审稿人的建议，我们对预期和社会奖赏这两者对个体合作行为的影响做了模拟分析（simulation）。在模拟分析中，我们通过变化虚拟被试的预期或社会奖赏的大小，观察他们对预测个体的合作行为的影响。在模拟预期时，我们将虚拟被试的其他参数值设置为组平均值并保持不变，每次模拟只变化所关心的变量。结果如下图。从图中可以看到预期（左图）或社会奖赏（右图）的提高都能够提高个体的合作概率。我们已将这部分加入到附加的模型预测与验证，并将图片添加为附图 4。



回应说明图 6. 预期（a）与社会奖赏（b）影响合作行为的模拟分析。

关于独立，认知模型中的预期和社会奖赏对于提高合作概率的作用并非完全独立，而是正向的非线性的交互作用。在模型中，两者的乘积起到影响主观价值的作用。参数恢复的结果验证这两部分本身的可甄别性和可区分性。因此，预期和社会奖赏这两部分是可相互分离而又互相作用，从而起到促进合作行为的作用。

关于正向相关，我们通过认知模型所获得的预期和社会奖赏，均通过被试行为进行拟合得到。因此必然是和被试的行为存在相关的（如下图所示）。我们通过行为获得的心理计算成分，由于我们的模型较为简单，如若再与行为做相关，这样在科学上存在循环验证的嫌疑，因此我们不推崇这样做。



回应说明图 7. 左侧图的横坐标为被试合作行为, 纵坐标为社会奖赏参数值。右侧图的横坐标为被试的合作行为, 纵坐标为对合作者合作行为的预期。各个子图对应不同的实验条件。以上数据来自实验 1 ($n=135$), 每个图的点代表一名被试。各个字体上方的统计值为皮尔逊相关分析。

意见 11: 文本写作。(1) 前沿第一段最后一句(Lines 124-126), “根据已有的研究, 影响个体做出条件合作的可能有两个主要的因素, 分别是合作期望(即个体对他人合作的期望)以及合作成功带来的喜悦, 也称为社会奖赏(social reward)”, 这似乎是作者的归纳, 加入对应的文献引用更为合适。

回应 11: 感谢审稿专家意见。在先前诸多实证研究及综述中曾提到社会奖赏(social reward)和预期(expectation)均对积极的社会互动(例如, 信任、合作)非常重要(Bhanji & Delgado, 2014; Ng et al., 2016; Pal & Hilbe, 2022; Rand et al., 2009; Tamir & Hughes, 2018)。我们已在相应的地方加入文献。

意见 12: 文本写作。(2) 前言中还存在诸多写作语法上的问题, 希望作者仔细写作后修改。比如, Line140, “但不足在是”; Line144, “人们进行社会互动较为基础心智化功能”; Line 150, “影响对手的信念从响对手接下来的行为”, Line171, “已经有足够的信能够帮助个体做出决策”。这里不再进行一一列举了。

回应 12: 感谢审稿专家指正, 我们已对全文文字进行细致的检查和修正。

意见 13: 文本写作。(3) 图 3, 图注的表达“e & f & h & i 展示社会奖赏和预期的统计分析结果”, 并没有 h,i 图。并且, panel e,f 呈现的 non-social 在 high/low 的结果是否存在问题, 图示与正文的描述的结论相反。

回应 13: 感谢审稿专家指正。我们已对相应的地方进行细致的检查和改正。

意见 14: 文本写作。(4) 附图 1.最优模型的参数, s, v 分别是什么? 比如 α_s 和 α_v ? 正文方法部分提到对 social 和 non-social 分开进行了参数估计, 对应为 α_h 和 α_l , 见 Line 273 行 likelihood 公式。我很困惑研究到底进行了怎么样的模型构建, 请保持一致性, 并阐述清楚。

回应 14: 感谢审稿专家的指正, 我们已对原文进行细致的检查和改正。原文的 s, v 是对 stable 和 volatile 的简写。我们对每位被试在不同的实验条件下进行参数估计。在实验 1 的模型构建, 我们对每位被试的 social 和 nonsocial 两种条件分别估计高概率 (high) 和低概率 (low) 两种条件。在实验 2 中, 我们对每位被试的 social 和 nonsocial 两种条件分别估计稳定 (stable) 和变动 (volatile) 两种条件。实验 2 的高概率和低概率为组间设计, 同一被试只有一种可能性, 我们对每位被试分开稳定和变动的条件拟合得到参数值后, 再按照组别进行统计分析。我们已在文中相应之处作出调整和改正, 并全文保持一致。

意见 15: 文本写作。(5) Line303, 对不同条件的被试选择数据进行统计检验而不是回归模型参数进行进一步的统计检验是否可称为“简单效应分析”。

回应 15: 感谢审稿专家提问。传统的处理方式是对被试的合作行为分条件计算平均合作概率, 然后用方差分析进行统计检验。因本研究中自变量的数量较多, 且考虑到平均处理被试的合作行使得数据一定程度上失真, 因此我们的分析策略是先构建一个逻辑回归模型, 将我们考虑到的变量都纳入后, 在逻辑回归的结果上再进一步对关心的效应进一步解读。我们将这一策略详细地在分析方法部分阐述。考虑到简单效应分析一词多用于方差分析。我们在文中对应的地方修改为事后检验 (post-hoc analysis), 这一词参考了(van den Berg, 2021)。

意见 16: 文本写作。(6) 请写清楚参数恢复和 model recovery 的具体计算方法。

回应 16: 感谢审稿专家建议。根据审稿专家的建议, 我们已对文中对应的部分改进, 阐述其具体的计算方式。并更新了附图 1, 增加了实验 2 的参数恢复。

(1) 参数恢复。请见附加的参数恢复部分: 我们执行参数恢复分析评估了模型 6 的过度拟合情况及参数可靠性 (见图 S1)。参数恢复旨在评估认知模型参数估计的准确性和可靠性。在本研究中, 我们首先基于被试的行为数据拟合模型, 视得到的参数值为真值。使用这些真值, 我们在模型和实验设计条件下预测被试行为, 生成虚拟数据。通过对虚拟数据的重新拟合, 我们得到参数的恢复值, 并通过计算真值与恢复值的相关系数来评估参数恢复效果。高度的相关性表明所用模型和参数估计方法能可靠地反映被试认知过程, 确保模型预测的准确性。结果显示, 所有参数真值与恢复值的皮尔逊相关系数均超 0.7, 证明模型有效区分参数, 无冗余或过度拟合。

(2) 模型恢复。请见附加的模型恢复部分: 模型恢复分析旨在评估最优模型与其他候选模型在区分能力和预测准确性方面的表现。分析步骤如下: 首先, 利用各模型参数生成模拟数据。随后,

所有模型被拟合至其生成的模拟数据上，意味着每次拟合过程中六个模型同时竞争，以此来模拟各模型对数据的解释能力。接着，采用 BIC 评估每个模型的拟合优度，并通过此准则分析各模型生成数据的拟合情况，以验证是否能准确识别出“真实”模型。该过程重复 100 次，每次准确识别“真实”模型的情况记为 1，否则记为 0。通过将识别准确的次数除以 100，我们计算出模型恢复的概率。分析结果支持模型 6 的优越性，显示其为数据提供最佳解释的概率为 0.73，如图 2b 所示。

参考文献

- Bhanji, J., & Delgado, M. (2014). The social brain and reward: social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1), 61-73. doi:http://dx.doi.org/10.1002/wcs.1266
- Bradshaw, A. R., & McGettigan, C. (2021). Instrumental learning in social interactions: Trait learning from faces and voices. *Quarterly Journal of Experimental Psychology*, 74(8), 1344-1359. doi:https://doi.org/10.1177/1747021821999663
- Fang, Z., Zhao, M., Xu, T., Li, Y., Xie, H., Quan, P., . . . Zhang, R.-Y. (2024). Individuals with anxiety and depression use atypical decision strategies in an uncertain world. *Elife*, 13. doi:https://doi.org/10.7554/eLife.93887.1
- Hackel, L. M., Mende-Siedlecki, P., & Amodio, D. M. (2020). Reinforcement learning in social interaction: The distinguishing role of trait inference. *J Exp Soc Psychol*, 88. doi:https://doi.org/10.1016/j.jesp.2019.103948
- Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences*, 105(18), 6741-6746. doi:http://dx.doi.org/10.1073/pnas.0711099105
- Henco, L., Brandi, M.-L., Lahnakoski, J. M., Diaconescu, A. O., Mathys, C., & Schilbach, L. (2020). Bayesian modelling captures inter-individual differences in social belief computations in the putamen and insula. *cortex*, 131, 221-236. doi:https://doi.org/10.1016/j.cortex.2020.02.024
- Hill, Suzuki, Polania, Moisa, O'Doherty, & Ruff. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nat Neurosci*, 20(8), 1142-1149. doi:http://dx.doi.org/10.1038/nn.4602
- Mukherjee, D., Filipowicz, A. L., Vo, K., Satterthwaite, T. D., & Kable, J. W. (2020). Reward and punishment reversal-learning in major depressive disorder. *Journal of Abnormal Psychology*, 129(8), 810. doi:https://doi.org/10.1037/abn0000641
- Ng, G. T. T., Au, W. T., & Review. (2016). Expectation and cooperation in prisoner's dilemmas: The moderating role of game riskiness. *Psychonomic Bulletin*, 23, 353-360. doi:https://doi.org/10.3758/s13423-015-0911-7
- Pal, S., & Hilbe, C. (2022). Reputation effects drive the joint evolution of cooperation and social rewarding. *Nat Commun*, 13(1), 5928. doi:https://doi.org/10.1038/s41467-022-33551-y
- Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D., & Nowak, M. A. (2009). Positive interactions promote public cooperation. *Science*, 325(5945), 1272-1275. doi:https://doi.org/10.1126/science.1177418
- Tamir, D. I., & Hughes, B. L. (2018). Social rewards: from basic social building blocks to complex social behavior. *Perspectives on Psychological Science*, 13(6), 700-717. doi:https://doi.org/10.1177/1745691618776263
- van den Berg, S. M. (2021). Analysing data using linear models. In: University of Twente.
- Wilson, R., & Collins, A. (2019). Ten simple rules for the computational modeling of behavioral data. *Elife*, 8, e49547. doi:http://dx.doi.org/10.7554/eLife.49547

第二轮

审稿人 2 意见：

作者详细回答了我的问题，并对文章进行了相应修改，论文质量得到较大提升。我没有其它问题了，建议发表！

编委意见：该文章结合实验设计及计算建模，探究合作的机制，兼具理论和现实意义。建议发表。

主编意见：同意审稿人和编委意见，建议发表。