

社会情境下奖赏学习中的预测误差加工方式与应激反应的关联：基于计算建模视角

马意清^{1,2}

¹西南大学心理学部, 重庆

²西南大学教育部认知与人格重点实验室, 重庆

收稿日期: 2026年3月6日; 录用日期: 2026年4月13日; 发布日期: 2026年4月27日

摘要

心理社会应激常由社会评价威胁与不可控感驱动, 但个体在相似应激源下的反应强度与恢复轨迹存在显著差异。奖赏过程被视为潜在的保护机制, 然而既有研究多停留在总体性指标, 较少从学习更新环节定位差异来源。基于计算建模视角, 本研究在社会互动情境中考察奖赏学习的预测误差加工方式与应激反应的关联。80名大学生分别在两次实验中完成特里尔社会应激测验(TSST)与重复信任任务(rTG) (两环节间隔 ≥ 3 天)。TSST期间多时点评估主观压力感、不确定感与社会评价威胁, 并计算AUC_g与AUC_i表征急性应激动力学; 同时提取心电信号并计算心率变异性(HRV)指标。rTG行为数据采用分层贝叶斯强化学习模型拟合, 并以LOOIC与WAIC进行模型比较。结果显示, 效价特异性Rescorla-Wagner模型对rTG行为数据拟合最佳, 提示社会奖赏学习存在效价不对称更新。参数-应激对应关系上, 负性预测误差权重 α^- 与主观压力感AUC_g及社会评价威胁AUC_i呈显著正相关, 而 α^+ 与逆温度参数 τ 与主观应激指标未见稳定关联。生理层面, 模型参数与各HRV指标之间未观察到显著相关。补充分析显示, 社会逆温度参数 τ_{social} 与感知压力(PSS)总分及失控感维度呈正相关。上述结果提示, 社会奖赏学习中对负性预测误差的赋权更可能与急性社会应激的主观动力学特征耦合, 而更稳定的决策确定性特征可能在更长时间尺度上与慢性压力负荷相关, 为从计算成分层面理解社会情境下奖赏学习-应激关联提供了机制线索。

关键词

社会奖赏加工, 强化学习, 预测误差, 应激反应

Association between Prediction Error Encoding in Social Reward Learning and Stress Responses: A Computational Modeling Perspective

Yiqing Ma^{1,2}

文章引用: 马意清(2026). 社会情境下奖赏学习中的预测误差加工方式与应激反应的关联: 基于计算建模视角. *心理学进展*, 16(4), 379-391. DOI: 10.12677/ap.2026.164211

¹Faculty of Psychology, Southwest University, Chongqing²Key Laboratory of Cognition and Personality, Ministry of Education, Southwest University, Chongqing

Received: March 6, 2026; accepted: April 13, 2026; published: April 27, 2026

Abstract

Psychosocial stress is often driven by social-evaluative threat and perceived uncontrollability, yet individuals show marked heterogeneity in stress reactivity and recovery. Although reward-related processes have been proposed as protective factors, prior work has largely relied on global indices and has rarely localized individual differences to specific computational components of learning. From a computational modeling perspective, the present study examined whether prediction-error (PE) processing during social reward learning is associated with acute and chronic stress-related outcomes. Eighty Chinese college students completed the Trier Social Stress Test (TSST) and a repeated Trust Game (rTG) on separate days (≥ 3 -day interval). During the TSST, subjective stress, uncertainty, and social-evaluative threat were assessed at multiple time points, and area-under-the-curve indices (AUCg and AUCi) were derived to characterize stress dynamics. Electrocardiography was recorded to compute heart-rate variability (HRV) indices. Trial-by-trial rTG choices were fitted with hierarchical Bayesian reinforcement-learning models, and model comparison was conducted using LOOIC and WAIC. The valence-specific Rescorla-Wagner model provided the best account of rTG behavior, indicating asymmetric updating for positive versus negative social outcomes. At the individual-differences level, the negative PE weight (α^-) was positively associated with cumulative subjective stress (AUCg) and the increase in social-evaluative threat (AUCi), whereas the positive PE weight (α^+) and inverse temperature (τ) showed no robust associations with subjective stress indices. No significant associations were observed between computational parameters and HRV measures. In a supplementary analysis, the social inverse temperature (τ_{social}) was positively related to perceived stress (PSS), particularly the uncontrollability dimension. These findings suggest that stronger weighting of negative social PEs may preferentially couple with acute subjective stress dynamics, whereas more stable decisional certainty may relate to longer-term stress burden, providing mechanistic insights into the link between social reward learning and stress.

Keywords

Social Reward Processing, Reinforcement Learning, Prediction Error, Stress Response

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>

Open Access

1. 引言

在高竞争的现代社会中，心理社会应激源呈现出更高的暴露频率与更强的持续性，这类应激源通常与社会评价威胁、不可控感以及潜在排斥相关。大量实验研究表明，当情境同时包含被他人评价与失败风险/不可控成分时，应激反应更稳定且幅度更大，尤其体现在皮质醇等 HPA 轴指标的显著升高(Dickerson & Kemeny, 2004)。与此一致，纵向研究与综述同样指出，应激性生活事件，特别是涉及人际冲突与社会拒斥等社会性压力，是抑郁等内化问题的重要近端风险因素，并为抑郁发作提供了较为坚实的因果证据(Slavich & Irwin, 2014; Hammen, 2005; Kendler et al., 1999)。然而，应激暴露并不必然导向病理结局，许多个体能够在创伤或慢性逆境后维持或恢复相对稳定的功能状态，提示应激与健康之间的联系会受到保护

性因素的显著调节(Bonanno, 2004; Feder et al., 2009)。

近十余年的动物与人类研究将奖赏系统视为其中一个关键的保护机制。在动物研究中,自然奖赏能够减弱 HPA 轴、心血管与行为层面的应激反应(Ulrich-Lai et al., 2010);在人类研究中,通过实验操纵诱发的奖赏期待同样可缓冲急性心理社会应激引发的主观应激体验与总体内分泌反应(Hu et al., 2021)。但外部奖赏干预在现实应用中常受限于持续性与潜在副作用,因此研究关注逐渐从“奖赏是否有效”推进到更具解释力的问题:在相类似应激源下,个体为何会呈现差异化的反应强度与恢复速度?这一转向隐含的关键前提是,奖赏加工过程中存在相对稳定的个体差异,而这些差异可能构成塑造应激反应模式的重要来源。相关研究已开始的特质层面提供支持性证据,例如奖赏敏感性与应激暴露期间奖赏相关脑区反应以及皮质醇恢复等复原指标之间存在系统关联(Hu et al., 2022)。

不过,若仅以单一指标概括奖赏加工,往往难以分辨不同心理过程的相对作用。现有理论与神经经济学框架普遍认为,奖赏加工包含可相对分离的预期阶段、价值判断与选择阶段、结果体验阶段,以及学习过程中由预测误差驱动的更新机制(Berridge & Robinson, 2003; Clithero & Rangel, 2014; Knutson & Greer, 2008; Schultz, 2016)。在解释应激差异时,尤需关注学习更新环节:预测误差不仅表示结果与预期之间的差距,更关键的问题在于偏差如何被赋权并进入更新过程,进而塑造后续期待与行为策略。若缺乏对学习更新机制的成分化刻画,奖赏-应激关系便容易停留在总体相关层面,难以进一步解释差异究竟发生在奖赏加工链条的哪个环节。

强化学习等认知计算建模为上述研究缺口提供了方法学支撑。与描述性指标不同,计算模型以生成方式刻画“期望-反馈-更新-选择”的动态过程,通过拟合试次序列推断潜在计算成分,从而将外显行为差异定位到可解释的参数层面(Daw et al., 2011; Sutton & Barto, 2018)。在社会情境下,这一优势更为突出:社会互动中的奖赏学习不仅涉及结果得失,还涉及对他人可信用度与互惠倾向的推断与更新。重复信任任务(repeated Trust Game, rTG)为此提供了自然范式:个体在多轮互动中依据对方回报不断调整投资决策,其学习过程可表述为预测误差驱动的价值更新。进一步地,模型可区分正性与负性预测误差,并以不同参数刻画对两类预测误差的赋权;同时,决策逆温度参数 τ 刻画价值差异向选择输出的转换强度,可反映决策一致性与噪声水平。由此,“预测误差加工方式”与“决策确定性”能够被分离估计,为解释社会评价威胁情境中的应激反应差异提供更精细的机制入口。

基于上述研究进展,本研究聚焦两个核心问题。其一,尽管奖赏过程被视为潜在的应激缓冲机制,但关于社会情境下奖赏学习的预测误差加工方式是否与急性社会应激的动态反应特征相联系,现有证据仍相对不足,尤其缺乏基于试次序列的计算分解来定位更新环节的差异来源。其二,应激效应具有时间尺度延展,除急性诱发反应外,社会学习中的稳定决策表型是否与更长期的压力负荷(如感知压力)相关,同样缺少直接检验。为弥补上述缺口,本研究采用 TSST 诱发急性社会应激,并以主观应激相关体验的时间进程刻画反应动态;同时纳入感知压力量表(PSS)表征慢性压力负荷,以检验社会学习参数在不同时间尺度上的对应模式。

基于上述理论与方法学框架,本研究关注社会奖赏学习的计算成分是否能够解释应激反应的个体差异。具体而言,首先检验学习更新环节的参数化特征(尤其是与预测误差加工相关的成分)是否与 TSST 诱发的主观应激反应动态相联系;其次,在更长时间尺度上,进一步考察社会情境下决策表型相关参数是否与感知压力负荷存在系统对应。

2. 方法

2.1. 被试

样本量在研究开始前通过 G*Power 3.1.97 进行估计(Faul et al., 2007)。在设定中等效应量(Cohen's $f =$

0.15)、显著性水平 $\alpha = 0.05$ 以及统计功效 $1 - \beta = 0.80$ 的条件下, 回归分析所需的最小样本量为 55 人。实际招募过程中, 共有 80 名来自西南大学的中国大学生自愿参与本研究(女性占 80.00%; 平均年龄 20.08 ± 1.75 岁)。参与者健康状况主要通过口头自陈进行初步筛查, 以尽量排除严重心理或躯体疾病史以及酗酒或药物滥用等情况; 此外, 所有参与者均为右利手, 无色盲或色弱, 视力或矫正视力正常。为降低短期生活方式因素对生理测量的潜在影响, 研究在实验前对参与者提出统一要求: 在应激任务开始前 1 小时内避免剧烈运动, 并避免进食或饮酒, 同时不饮用咖啡及刷牙。所有参与者均签署书面知情同意并获得相应报酬; 研究方案经当地机构伦理委员会批准, 研究过程遵循《赫尔辛基宣言》的相关原则。

2.2. 实验流程

本研究包含两个相对独立的实验环节, 分别用于诱发急性社会应激反应与测量社会情境下的奖赏学习个体差异。第一环节为急性应激诱发, 参与者在实验室完成特里尔社会应激测验(Trier Social Stress Test, TSST), 以社会评价威胁为核心线索, 在相对标准化条件下诱发急性应激反应(图 1)。为尽量降低短期生活方式因素对生理应激指标的潜在干扰, 实验开始前对参与者提出统一要求: 至少提前 2 小时避免剧烈运动, 并避免吸烟与饮酒。第二环节用于测量社会情境中的奖赏学习个体差异, 参与者完成重复信任任务(repeated Trust Game, rTG); 基于试次层面的决策与反馈序列进行计算建模, 从而估计与社会奖赏学习相关的个体差异参数(图 2)。考虑到两项任务可能在情绪状态与生理唤醒水平上产生交叉影响, 两环节安排在不同日期完成, 且间隔不少于 3 天, 以在设计层面尽可能降低任务间的潜在干扰。

2.3. 慢性应激量表

慢性应激水平通过感知压力量表(Perceived Stress Scale, PSS)进行测量。本研究采用 Cohen 等人编制的 14 项版本, 并使用由杨廷忠与黄汉腾修订的中文版量表实施测评(Cohen et al., 1983; 杨廷忠, 黄汉腾, 2003)。该量表旨在评估个体在过去一个月内的主观压力体验, 条目采用 5 点李克特评分, 从 0 (从不)到 4 (经常/非常频繁), 用于反映个体在多大程度上将日常生活情境知觉为具有压力的事件, 例如感到难以控制生活中重要的事情。量表得分越高代表过去一个月的感知压力水平越高; 在本研究样本中, 量表内部一致性良好(Cronbach's $\alpha = 0.90$)。其中第 4、5、6、7、9、10、13 题为正向表述条目, 按常规进行反向计分后, 与其余条目得分相加得到总分。除总分外, 本研究进一步计算两个分维度: 紧张感与失控感。其中, 紧张感主要反映个体在过去一个月中体验到的压力性情绪与困扰程度(如紧张、烦躁、被压力压迫等); 失控感主要反映个体对生活事件的可控性评估与无助体验, 即在面对日常要求时感到难以掌控、难以应对或缺乏控制感的程度。两项分维度得分越高均表示相应的慢性应激体验越强。

2.4. 社会急性应激诱发范式

特里尔社会应激测验(Trier Social Stress Test, TSST)是一种经典且应用广泛的急性心理社会应激诱发范式, 能够在以社会评价威胁为核心的情境下稳定诱发主观与生理层面的应激反应(Dickerson & Kemeny, 2004; Kirschbaum et al., 1993)。本研究采用经情境化改编的 TSST, 将任务设置为“模拟工作面试”场景, 以强化被评价与被比较的社会威胁体验(见图 1)。在整体流程上, 参与者到达实验室后首先佩戴生理信号采集设备(BIOPAC MP150), 随后进入 10 分钟静息阶段以获取基线水平; 基线结束后, 参与者在进入任务前完成一次主观应激评定。随后研究者告知参与者将进行面试任务, 并给予 5 分钟准备时间, 用于组织陈述内容与熟悉任务要求; 准备结束后, 参与者进入独立测试房间正式完成 TSST 面试环节, 在评委在场的情境下进行限时演讲(约 5 分钟), 要求围绕个人胜任力、岗位匹配与过往经历进行结构化陈述, 并被明确告知其表现将被评估与记录。任务结束后, 参与者返回原实验室并再次报告主观应激水平, 随后进入约 30 分钟的恢复阶段, 期间再进行两次主观应激评定, 以刻画应激反应的恢复轨迹。

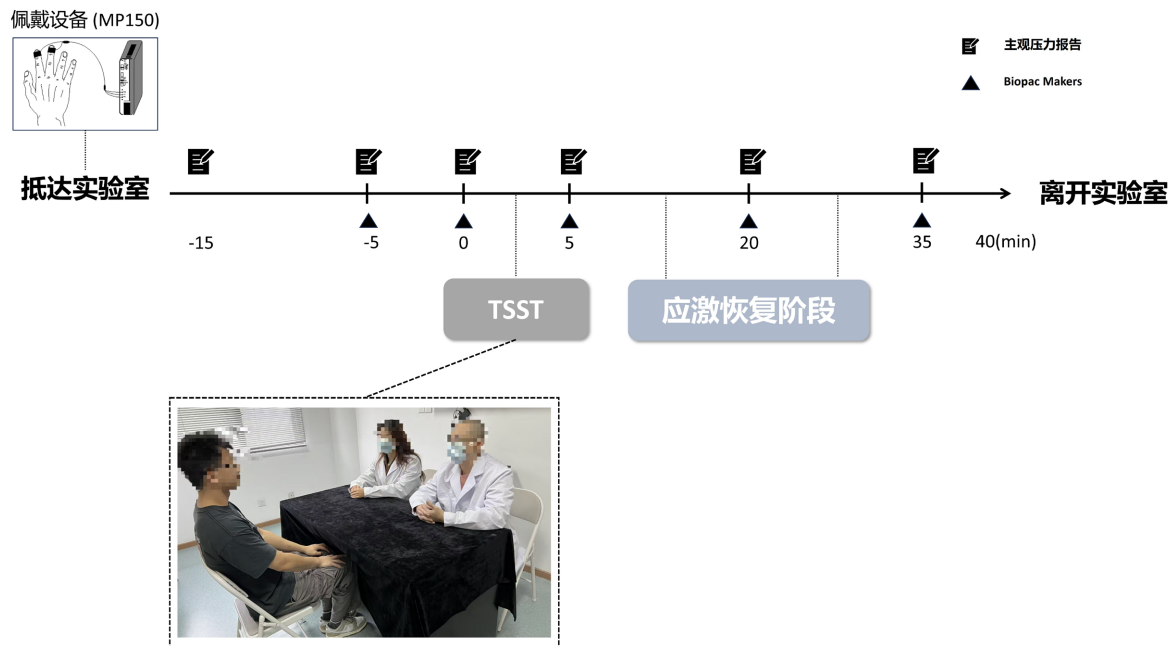


Figure 1. Trier Social Stress Test, TSST
图 1. 特里尔社会应激测验

2.5. 社会奖赏加工过程研究范式

本研究以重复信任任务(repeated Trust Game, rTG)作为核心行为测量范式,用于在社会互动情境中刻画奖赏加工的个体差异。研究重点放在这样一个过程上:个体如何在反复互动中依据他人的互惠或背叛反馈,逐步形成对他人互惠倾向的价值预期,并在试次层面持续修正这一预期。由于该范式同时具备重复互动、即时社会反馈与跨试次更新等结构特征,参与者的决策调整往往会自然伴随预测误差信号的产生与利用,从而为后续计算建模提供稳定的行为基础。基于试次层面的选择与反馈数据,研究可进一步提取与社会奖赏加工相关的个体差异指标,例如对社会结果的预测误差加工方式、对互惠概率的更新敏感性,以及更一般的选择策略和探索-利用倾向等。为维持参与者的投入与动机,任务开始前明确告知其决策表现将与最终获得的报酬金额挂钩。

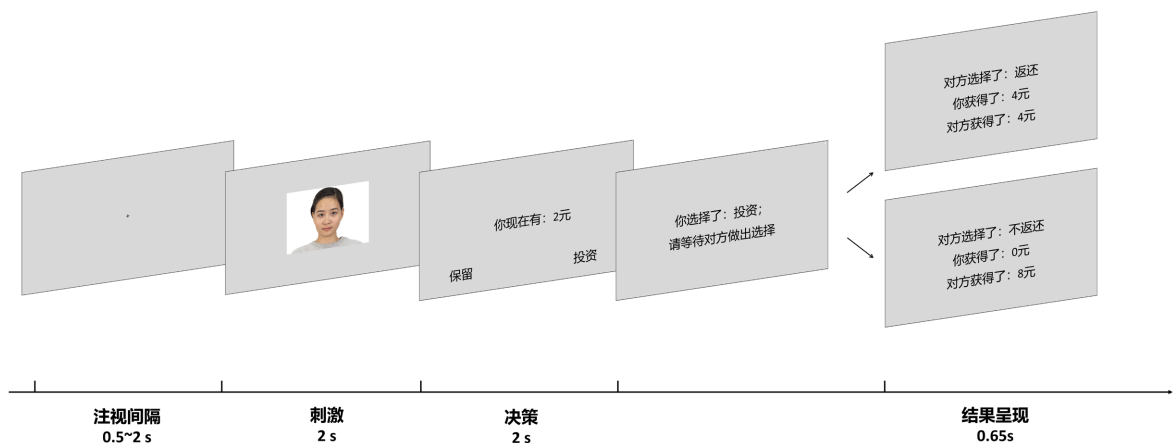


Figure 2. Repeated Trust Game, rTG
图 2. 重复信任任务

在任务设置上,参与者扮演投资者,与多名虚拟受托人进行重复互动,并在不确定条件下判断不同对象是否值得信任。每名受托人具有预设但对参与者不可见的互惠倾向,且互惠概率在不同受托人之间存在系统差异(90%, 70%, 50%, 30%, 10%)。每位受托人对应 20 轮交互,且多名受托人在任务过程中随机交替出现,这一安排使参与者需要在多目标的社会环境中并行学习不同对象的价值表征,并随受托人身份变化动态调整信任决策(见图 2)。

每轮试次的基本流程为:首先呈现抖动的注视间隔(0.5~2 s),随后呈现受托人刺激(2 s),并在选择阶段(2 s)要求参与者做出是否投资的决策。完成决策后进入短暂等待界面,继而呈现结果反馈(0.65 s),告知对方是否互惠以及由此带来的收益变化。通过这一设计,参与者在每次面对特定受托人时都需要形成对其行为结果的预期,而反馈阶段呈现的实际结果则与该预期形成偏差,从而在试次层面提供可观测且可重复的社会预测误差信息,为计算建模识别个体差异提供关键依据。

2.6. 模型建构与拟合

在本研究中构建了四种不同的强化学习模型,用以描述可能生成被试选择数据的奖赏加工过程。第一种为标准 Rescorla-Wagner (RW)模型,该模型认为抽象刺激的价值 V 会在试次层面基于对预测误差赋予的权重 α 与预测误差 PE 进行更新。其中, r_t 表示当前试次的奖赏结果,未被选择刺激 $V_t(B)$ 的价值保持不变。

$$PE_t = r_t - V_t(A)$$

$$V_{t+1}(A) = V_t(A) + \alpha * PE_t$$

第二种为效价特异性 Rescorla-Wagner 模型,该模型认为价值更新会受到预测误差符号的调节,即对正、负预测误差分别赋予不同的权重。当预测误差为正时使用 α^+ 更新价值;当预测误差为负时使用 α^- 更新价值。

$$PE_t = r_t - V_t(A)$$

$$\begin{cases} V_{t+1}(A) = V_t(A) + \alpha^+ * PE_t, PE_t > 0 \\ V_{t+1}(A) = V_t(A) + \alpha^- * PE_t, PE_t < 0 \end{cases}$$

第三种为 Pearce-Hall (PH)模型。与 RW 模型依赖固定权重不同,PH 模型通过一个随时间动态变化的“可联结性”参数来调节价值更新速度。这里,可联结性 $\alpha_t(A)$ 决定价值更新速率,其大小受预测误差幅度 $|PE_t|$ 影响,并由预测误差权重参数 γ 调节其变化幅度。

$$PE_t = r_t - V_t(A)$$

$$V_{t+1}(A) = V_t(A) + c * \alpha_t(A) * PE_t$$

$$\alpha_{t+1}(A) = \gamma * |PE_t| + (1 - \gamma) * \alpha_t(A)$$

第四种为效价特异性 Pearce-Hall 模型。在效价特异性 RW 模型的思路基础上,将类似的效价调节规则扩展到 PH 模型中,提出被试在调节可联结性时会对正、负预测误差赋予不同的权重,即分别使用 γ^+ 与 γ^- 来更新可联结性。

$$PE_t = r_t - V_t(A)$$

$$V_{t+1}(A) = V_t(A) + c * \alpha_t(A) * PE_t$$

$$\begin{cases} \alpha_{t+1}(A) = \gamma^+ * |PE_t| + (1 - \gamma^+) * \alpha_t(A), & PE_t > 0 \\ \alpha_{t+1}(A) = \gamma^- * |PE_t| + (1 - \gamma^-) * \alpha_t(A), & PE_t < 0 \end{cases}$$

对于所有模型,本研究均假设个体基于刺激价值通过 Softmax 函数进行选择。其中,逆温度参数 τ 调节探索与利用之间的权衡:

$$P_t(A) = \frac{e^{\tau * V_t(A)}}{e^{\tau * V_t(A)} + e^{\tau * V_t(B)}} = \frac{1}{1 + e^{\tau * (V_t(B) - V_t(A))}}$$

$$P_t(B) = 1 - P_t(A)$$

本研究在 R 环境中通过 RStan 接口调用 Stan 平台(Carpenter et al., 2017),对上述四个模型进行拟合与参数估计。所有模型均采用分层贝叶斯框架,即在群体水平分布的约束下估计个体水平参数,以在样本量与参数维度相对有限的情况下提升估计稳定性,并获得更精细的个体差异表征。后验抽样使用 6 条相互独立的 MCMC 链完成,每条链包含 4000 次预热迭代(Burn-In),随后进行 4000 次正式抽样迭代。为降低采样发散风险,将步长调节相关参数设置为较高水平(adapt = 0.99),从而在采样过程中采用更保守的步长策略。模型收敛性主要通过 \hat{R} 统计量评估,确保所有关键参数的 \hat{R} 接近 1 ($0.99 < \hat{R} < 1.01$)。

模型比较采用留一交叉验证(Leave-One-Out Cross-Validation, LOO)与广泛适用信息准则(Widely Applicable Information Criterion, WAIC)。两类指标均基于后验抽样下逐点对数似然的计算,用以估计模型的逐点预测准确性。在贝叶斯建模语境中,相较于 AIC、DIC 等依赖更多简化假设的预测误差指标,LOO 与 WAIC 通常具有更好的适用性与稳健性。完成行为数据的模型拟合后,进一步使用 R 的 loo 包计算各模型的 LOOIC 与 WAIC,并据此进行模型评估与比较。

2.7. 数据分析

本研究的数据分析在 R 中完成,覆盖数据整理与描述性统计、急性应激诱发的有效性检验,以及围绕个体差异的后续推断分析。分析初期先对人口学变量与核心研究变量进行概括性描述,并同步完成缺失值模式与异常值的审查,以确保数据质量满足后续建模与推断的基本前提。随后,将急性应激诱发的有效性作为关键检验环节,针对不同测量时间点的主观应激指标实施重复测量方差分析,重点考察任务前后及恢复阶段的变化轨迹,从而验证主观应激水平是否呈现预期的动态波动。与此同时,为更集中地表征应激诱发过程中的总体变化幅度,本研究基于多时间点主观评分进一步计算曲线下面积指标,包括曲线下增加面积(Area under the Curve with Respect to Increase, AUCI)与曲线下总面积(Area under the Curve with Respect to Ground, AUCG),并将其作为后续分析的综合性表征指标。

生理数据处理在 Python 中完成。ECG 信号的预处理与 HRV 指标提取依托 NeuroKit2 包实现,主要步骤包括 R 峰识别、异常心搏校正以及 HRV 指标计算。HRV 分析聚焦于应激任务阶段,具体在 TSST 的 5 分钟任务窗口内计算各项 HRV 指标。

3. 结果

3.1. 效价特异性 Rescorla-Wagner 模型对社会奖赏加工数据拟合最佳

在使用强化学习模型提取概率性逆转学习任务中反映奖赏加工个体差异的潜在变量之前,首先对四个候选模型进行比较,以确定最适合刻画被试选择行为的模型。模型比较基于留一交叉验证信息准则(Leave-One-Out Information Criterion, LOOIC)与广泛适用信息准则(Widely Applicable Information Criterion, WAIC),并通过 R 的 loo 包计算获得。LOOIC 与 WAIC 数值越低表示模型对数据的预测拟合越好。结果

显示, 效价特异性 Rescorla-Wagner 模型(Model2)在 LOOIC 与 WAIC 指标上均优于其他模型, 因而对被试行为数据提供了最佳拟合(表 1)。

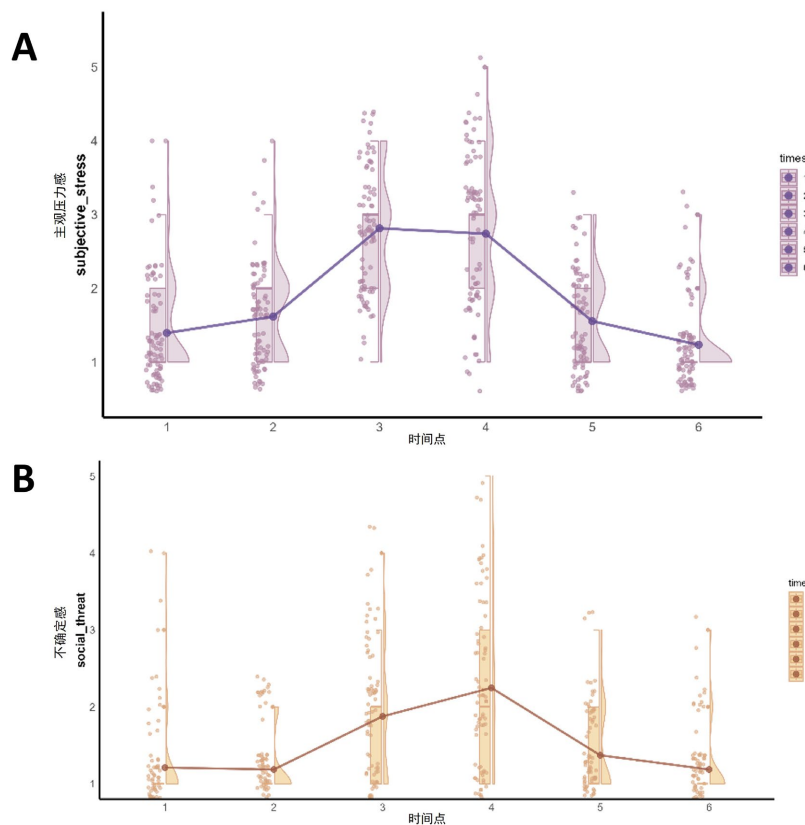
Table 1. Model comparison results

表 1. 模型比较结果

模型	LOOIC		WAIC	
	Values	SE	Values	SE
Rescorla-Wagner 模型	14159.8	98.3	14159.5	98.3
效价特异性 Rescorla-Wagner 模型	14031.6	99.4	14030.5	99.4
Pearce-Hall 模型	14218.8	97.3	14218.3	97.3
效价特异性 Pearce-Hall 模型	14218.9	97.3	14118.4	97.3

3.2. 社会急性应激诱发的操作性检验

急性社会应激诱发的操纵性检验以六个测量时间点的主观应激报告作为因变量, 分别针对主观压力感、不确定感与社会评价威胁实施重复测量方差分析。结果显示, 三个维度均存在显著的时间主效应: 主观压力感 $F(5,400) = 110.27, p < 0.001$; 不确定感 $F(5,400) = 52.02, p < 0.001$; 社会评价威胁 $F(5,400) = 46.30, p < 0.001$ 。总体而言, 从均值变化轨迹来看, 三项主观指标在任务前整体处于相对较低水平, 在 TSST 阶段显著升高, 并在随后的恢复阶段逐步回落(图 3)。上述模式支持所采用的 TSST 在实验室情境下能够有效诱发急性社会应激体验, 同时呈现随时间推移的恢复趋势, 从而为进一步分析急性应激反应及其与社会奖赏加工个体差异指标之间的关联奠定了基础。



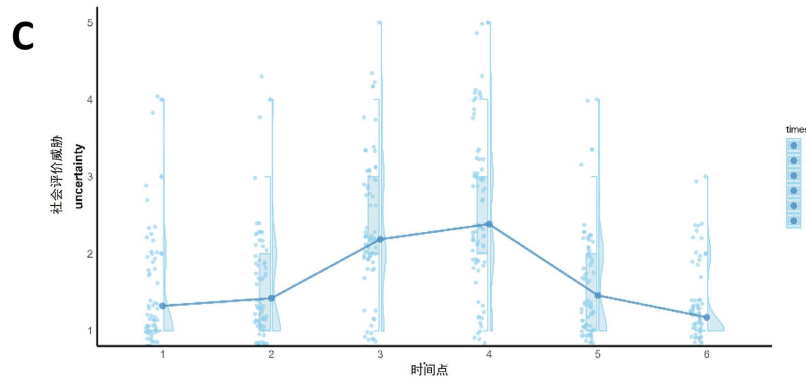


Figure 3. Subjective stress ratings during the TSST
图 3. 主观应激报告(TSST)

3.3. 社会奖赏加工过程中个体差异与急性应激反应的联系

随后，进一步检验模型参数所表征的社会奖赏加工个体差异是否与急性社会应激反应期间主观应激的总体水平相对应。主观应激的总体水平分别以曲线下总面积 AUC_g 与曲线下增长面积 AUC_i 加以表征。结果如图 4 所示，在三个模型参数中，负性预测误差权重参数 α^- 与主观压力感的 AUC_g 呈显著正相关 ($r = 0.26, p < 0.05, 95\% CI [0.04, 0.45]$)，并且与社会评价威胁的 AUC_i 亦呈显著正相关 ($r = 0.22, p < 0.05, 95\% CI [0.01, 0.42]$)。该关联模式表明，在社会互动情境中更倾向于对负性预测误差赋予较高权重的个体，应激诱发过程中累积的主观压力水平更高，同时其社会评价威胁体验的增长幅度也相对更为明显。相比之下，正性预测误差权重参数 α^+ 以及逆温度参数 τ 与各主观应激指标的 AUC_g 或 AUC_i 之间未观察到显著关联(图 4)。

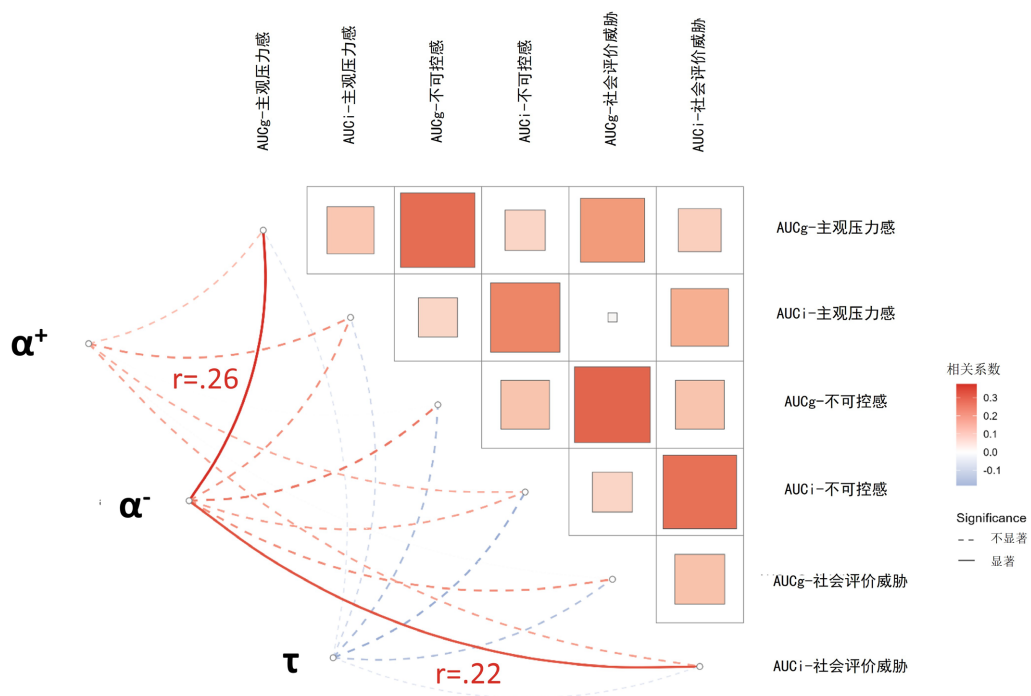


Figure 4. Correlations between model parameters and subjective social stress
图 4. 社会奖赏加工模型参数与社会主观应激之间的相关性

生理指标层面,进一步考察了模型参数与应激反应期间多项 HRV 指标之间的相关模式。结果显示,负性预测误差权重参数 α^- 、正性预测误差权重参数 α^+ 以及逆温度参数 τ 与各 HRV 指标之间均未达到显著相关。这一结果在一定程度上表明,本研究所捕捉的社会奖赏加工相关计算差异更倾向于体现在主观应激体验的个体分化上,而在心率变异性指标所反映的自主神经活动层面,尚未观察到相对稳定的一致对应关系。

3.4. 社会奖赏加工过程中个体差异与慢性应激的联系

为检验本研究的第二个核心问题,即奖赏加工中的个体差异是否会在更长时间尺度上与慢性应激水平形成系统关联,本研究以社会奖赏学习任务中计算模型提取的关键参数作为奖赏加工差异的操作化指标,并将其与感知压力量表(PSS)总分及分维度进行相关分析,从而评估社会奖赏加工的潜在计算成分在解释慢性应激个体差异方面的作用。

总体而言,PSS 总分均值为 32.36 (SD = 8.82)。PSS 两个维度与总分之间呈高度正相关(不可控感: $r=0.90, p < 0.01, 95\% CI [0.86, 0.93]$; 紧张感: $r = 0.88, p < 0.01, 95\% CI [0.83, 0.92]$), 两维度之间也呈中等相关($r = 0.59, p < 0.01, 95\% CI [0.44, 0.71]$), 说明二者在反映共同压力体验的同时仍保留一定的区分。

在社会奖赏学习参数方面,社会逆温度参数 τ 与 PSS 总分呈显著正相关($r = 0.23, p < 0.05, 95\% CI [0.03, 0.41]$), 并且与不可控感维度的正相关更为稳定($r = 0.27, p < 0.01, 95\% CI [0.07, 0.44]$); 其与紧张感维度的相关未达到显著水平($r = 0.13, 95\% CI [-0.07, 0.32]$)。同时,社会正性与负性预测误差权重参数 α^+ 与 α^- 与 PSS 总分均未呈现显著相关(α^+ : $r = -0.05, 95\% CI [-0.25, 0.15]$; α^- : $r = -0.12, 95\% CI [-0.31, 0.08]$)。

4. 讨论

本研究聚焦社会互动情境下的奖赏学习个体差异,并检验其是否会在以社会评价威胁为关键信号的急性社会应激中表现为更强的主观反应与更大的时间维度累积,从而呈现更突出的“应激惯量”。结果整体呈现出较为清晰且集中的结构:其一,重复信任任务的行为序列更符合效价特异性 Rescorla-Wagner 模型,提示社会奖赏学习中对正、负结果信息的整合同样具有非对称性;其二,参数-应激的对应并未广泛泛化到多项计算成分,而是更集中地指向负性社会预测误差赋权(α^-)与主观应激动力学指标之间的耦合;其三,模型参数与 HRV 指标之间未观察到稳定相关;其四,在更长时间尺度上,社会情境下的逆温度参数(τ)与慢性压力负荷(PSS),尤其是失控感维度,呈现系统对应。总体而言,这些发现为“社会情境下奖赏学习-应激反应”关系提供了更具机制指向的证据,并提示不同计算成分可能在不同时间尺度上对应不同类型的应激结局。但需要强调的是,本研究采用横断面设计,所有变量之间的关系均基于同一阶段内获得的相关证据,因此当前结果仅支持社会奖赏学习中的计算成分与应激指标之间存在系统对应,而预测误差加工方式或决策确定性对急性/慢性应激的单向因果作用还需要未来研究进一步探究。

首先,模型比较结果显示,允许正、负社会预测误差以不同权重进入价值更新的效价特异性 Rescorla-Wagner 模型对行为数据的拟合最佳。这一发现提示,在包含他人反馈、不确定性与反复互动的社会学习过程中,个体并非以对称方式整合不同价向的反馈信息,而更可能呈现效价不对称的更新机制。该模式与强化学习文献中“效价不对称更新更具解释力”的观点相一致(Frank et al., 2007),也与以信息准则进行模型比较与泛化评估的实践路径相契合(Vehtari et al., 2017)。从方法学意义上看,这一结果为后续参数解释提供了更稳健的基础:与其依赖总体性行为指标,不如在生成模型层面对学习更新过程进行成分化分解,从而更精确地定位预测误差加工方式在社会学习个体差异中的作用位置。

更关键的是,参数与应激指标的对应关系呈现出较为集中的耦合路径。与其说社会学习参数与应激

反应存在广泛关联, 不如说更清晰地指向负性社会预测误差权重 α^- : α^- 越高, TSST 诱发过程中主观压力感的累积总体水平(AUC_g)越高, 同时社会评价威胁体验的增长幅度(AUC_i)也更为明显; 相比之下, α^+ 与逆温度参数 τ 并未显示出同等稳定的对应。该模式更适合被理解为一种“负性社会预测误差赋权—应激惯量”之间的衔接机制。社会互动中的负性反馈往往同时携带损失与关系性威胁含义, 当个体在学习更新中更倾向于放大并整合“比预期更差”的社会结果信息时, 更容易形成偏负性的社会预期并提高对不确定性的评估, 从而使得以被评价、被比较与失败风险为核心的情境更快被编码为威胁, 并在应激推进过程中表现为更高的累积负荷以及更陡峭的威胁上升轨迹(Dickerson & Kemeny, 2004; Kirschbaum et al., 1993)。从“奖赏加工差异如何嵌入应激反应”这一主线来看, 上述结果为机制定位提供了更明确的情境化解释: 社会情境未必要求引入全新的计算成分, 但同一类差异——尤其是对负性预测误差的赋权方式——在社会线索占主导的环境中更容易与社会应激的动力学过程发生耦合, 从而更直接塑造主观应激反应的强度与持续性。

与此同时, 模型参数与 HRV 指标之间未观察到显著相关。这一结果一方面可能意味着, 本研究所捕捉到的社会奖赏加工差异更优先体现在对社会威胁与压力体验的主观分化上; 另一方面也提示, 在 TSST 这类多成分、强情境依赖的社会评价任务中, 自主神经活动可能同时受到基线状态、情绪调节策略、任务投入程度与个体生理敏感性等多重因素的共同影响, 因此较难与单一计算参数形成稳定一致的线性对应。后续研究可进一步通过更精细的时间窗口划分(反应性与恢复的分段指标)、更丰富的生理系统测量(如皮质醇或炎症相关指标), 以及与任务过程更匹配的动态建模方式, 检验社会学习计算成分与生理应激反应耦合的边界条件。

此外, 社会奖赏学习中 τ 与慢性压力负荷, 尤其是失控感之间的关联, 提示慢性压力在社会情境下更可能对应于决策确定性的差异, 而不一定体现为对社会负性反馈的加速更新。逆温度参数刻画的是价值差异向选择偏好的映射强度, τ 越高, 决策越依赖当前价值差异, 行为越趋于确定而非探索。在长期压力水平较高的个体中, 更强的决策确定性可能体现为一种降低社会不确定成本的策略倾向; 然而社会环境本身具有更高的外源波动性与不可控性, 这类策略未必提升实际可控性, 反而可能在反复遭遇不一致反馈时强化“自身影响力有限”的主观判断, 从而更紧密地对应于失控感体验。相关理论与证据同样指出, 应激状态下个体的行为控制更容易从灵活的目标导向模式转向相对固化的习惯化模式, 表现为策略层面的稳定性增强与灵活性下降(Schwabe & Wolf, 2011)。结合社会评价威胁在社会应激中的核心地位, 该发现进一步提示: 与社会情境相关的奖赏加工差异, 在长期层面可能更多体现为控制与确定性环节的变化, 从而与慢性压力负荷形成更稳定的对应。

需要指出的是, 本研究采用横断面设计, 主要基于同一阶段内的相关证据识别社会奖赏学习参数与应激指标之间的对应关系, 因此尚不足以支持明确的因果推断。尤其是社会学习参数与慢性压力负荷之间的关联, 既可能反映特定计算成分对应激易感性的影响, 也可能反映长期压力暴露反过来塑造了社会决策表型。与此同时, 本研究样本主要由中国大学生构成, 年龄范围较为集中, 文化背景相对单一, 且女性比例较高(80.00%), 因此样本整体具有较强的同质性。这种样本特征虽然有助于降低部分无关变异, 但也可能限制研究结论向其他年龄群体、不同文化背景人群以及更均衡性别构成样本的推广。未来研究可结合纵向追踪、实验操纵以及跨年龄、跨文化和更具代表性的样本设计, 进一步检验社会奖赏学习计算成分与应激反应之间关系的方向性与普适性。

综上, 研究结果支持如下讨论要点: 社会互动中的奖赏学习具有明显的效价不对称更新特征; 负性社会预测误差赋权(α^-)与急性社会应激的主观动力学指标呈现更集中的耦合关系, 符合应激视角下对累积负荷与威胁上升轨迹的解释; 生理层面 HRV 未见稳定对应提示主观—生理映射可能存在边界条件; 而

在更长时间尺度上, 社会情境下的决策确定性特征(τ)与慢性压力负荷, 尤其失控感体验, 呈现更稳定的对应, 为从计算成分面理解社会奖赏学习-应激关系提供了进一步的机制线索。

5. 结论

总而言之, 本研究结果表明, 社会互动中的奖赏学习可能具有明显的效价不对称更新特征, 且负性社会预测误差赋权与急性社会应激的主观动力学指标之间存在更集中的对应关系; 在更长时间尺度上, 社会情境下的决策确定性特征也与慢性压力负荷, 尤其是失控感体验, 呈现一定关联。需要强调的是, 由于本研究采用横断面设计且样本主要来自年龄、文化背景较为一致的中国大学生群体, 同时女性比例较高, 因此上述发现应被理解为相关性证据和机制线索, 而非明确的因果结论, 其普适性仍有待在更具代表性的样本中进一步检验。

参考文献

- 杨廷忠, 黄汉腾(2003). 社会转型中城市居民心理压力的流行病学研究. *中华流行病学杂志*, 24(9), 760-764.
- Berridge, K. C., & Robinson, T. E. (2003). Parsing Reward. *Trends in Neurosciences*, 26, 507-513. [https://doi.org/10.1016/s0166-2236\(03\)00233-9](https://doi.org/10.1016/s0166-2236(03)00233-9)
- Bonanno, G. A. (2004). Loss, Trauma, and Human Resilience: Have We Underestimated the Human Capacity to Thrive after Extremely Aversive Events? *American Psychologist*, 59, 20-28. <https://doi.org/10.1037/0003-066x.59.1.20>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M. et al. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76, 1-32. <https://doi.org/10.18637/jss.v076.i01>
- Clithero, J. A., & Rangel, A. (2014). Informatic Parcellation of the Network Involved in the Computation of Subjective Value. *Social Cognitive and Affective Neuroscience*, 9, 1289-1302. <https://doi.org/10.1093/scan/nst106>
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, 24, 385-396. <https://doi.org/10.2307/2136404>
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, 69, 1204-1215. <https://doi.org/10.1016/j.neuron.2011.02.027>
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin*, 130, 355-391. <https://doi.org/10.1037/0033-2909.130.3.355>
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. *Behavior Research Methods*, 39, 175-191. <https://doi.org/10.3758/bf03193146>
- Feder, A., Nestler, E. J., & Charney, D. S. (2009). Psychobiology and Molecular Genetics of Resilience. *Nature Reviews Neuroscience*, 10, 446-457. <https://doi.org/10.1038/nrn2649>
- Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic Triple Dissociation Reveals Multiple Roles for Dopamine in Reinforcement Learning. *Proceedings of the National Academy of Sciences*, 104, 16311-16316. <https://doi.org/10.1073/pnas.0706111104>
- Hammen, C. (2005). Stress and Depression. *Annual Review of Clinical Psychology*, 1, 293-319. <https://doi.org/10.1146/annurev.clinpsy.1.102803.143938>
- Hu, W., Liu, Y., Li, J., Zhao, X., & Yang, J. (2021). Early Life Stress Moderated the Influence of Reward Anticipation on Acute Psychosocial Stress Responses. *Psychophysiology*, 58, e13892. <https://doi.org/10.1111/psyp.13892>
- Hu, W., Zhao, X., Liu, Y., Ren, Y., Wei, Z., Tang, Z. et al. (2022). Reward Sensitivity Modulates the Brain Reward Pathway in Stress Resilience via the Inherent Neuroendocrine System. *Neurobiology of Stress*, 20, Article ID: 100485. <https://doi.org/10.1016/j.ynstr.2022.100485>
- Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). Causal Relationship between Stressful Life Events and the Onset of Major Depression. *American Journal of Psychiatry*, 156, 837-841. <https://doi.org/10.1176/ajp.156.6.837>
- Kirschbaum, C., Pirke, K., & Hellhammer, D. H. (1993). The "Trier Social Stress Test"—A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, 28, 76-81. <https://doi.org/10.1159/000119004>
- Knutson, B., & Greer, S. M. (2008). Anticipatory Affect: Neural Correlates and Consequences for Choice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 3771-3786. <https://doi.org/10.1098/rstb.2008.0155>
- Schultz, W. (2016). Dopamine Reward Prediction Error Coding. *Dialogues in Clinical Neuroscience*, 18, 23-32.

<https://doi.org/10.31887/dens.2016.18.1/wschultz>

- Schwabe, L., & Wolf, O. T. (2011). Stress-Induced Modulation of Instrumental Behavior: From Goal-Directed to Habitual Control of Action. *Behavioural Brain Research*, 219, 321-328. <https://doi.org/10.1016/j.bbr.2010.12.038>
- Slavich, G. M., & Irwin, M. R. (2014). From Stress to Inflammation and Major Depressive Disorder: A Social Signal Transduction Theory of Depression. *Psychological Bulletin*, 140, 774-815. <https://doi.org/10.1037/a0035302>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.
- Ulrich-Lai, Y. M., Christiansen, A. M., Ostrander, M. M., Jones, A. A., Jones, K. R., Choi, D. C. et al. (2010). Pleasurable Behaviors Reduce Stress via Brain Reward Pathways. *Proceedings of the National Academy of Sciences*, 107, 20529-20534. <https://doi.org/10.1073/pnas.1007740107>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC. *Statistics and Computing*, 27, 1413-1432. <https://doi.org/10.1007/s11222-016-9696-4>