

第一轮修改

审稿人 1 意见:

1. 模型选择与验证的严谨性。缺少 HDDM 对数据的绝对拟合指标，无法判断该模型是否适用于分析该数据，进而无法判断后续所有相关结果的正确性。

回应：感谢专家对模型验证严谨性的宝贵建议。为保证模型拟合的可靠性并选择最优模型，我们根据 Gelman-Rubin $\hat{R}$ 值和偏差信息量准则 (Deviance Information Criterion, DIC) 两个指标对上述 7 种模型进行比较。Gelman-Rubin $\hat{R}$ 值反映模型收敛程度 (Gelman & Rubin, 1992)，越接近于 1 表示同一模型下不同 MCMC 链内和链间的方差差异越小，模型越可靠；若大于 1.1 则表明该模型未收敛。DIC 是模型拟合比较的重要指标，其值越小说明模型越好。若两模型 DIC 差值超过 10，则认为模型间差异明显 (Spiegelhalter et al., 2002)。此外，我们还按照审稿专家 2 的建议，补充了对后验分布的差异检验以佐证基于均值的模型分析结果，增强模型选择的严谨性。相关内容已在正文对应部分进行补充和完善。

2. 未明确提供 HDDM 各参数的先验分布；由于研究样本量较小，先验分布（确认一下，查一下文献）的选择理论上会对后验分布产生较大影响。严谨角度，作者应探讨不同先验分布下的参数估计结果稳健性。

回应：感谢专家对模型参数估计稳健性的宝贵建议。关于 HDDM 先验分布的选择及其影响我们进行了以下关键检验：（1）通过四条独立马尔可夫链的收敛性诊断显示，所有核心参数 (v, a, t) 的 R -hat 值均接近 1 (正文已增加相应表格)，轨迹图（正文已增加相应图示）显示各链高度重叠，模型收敛良好；（2）最后 100 次迭代的自相关图表明参数估计的自相关性快速衰减至零，证实采样独立性；（3）阈值参数 t 的后验分布直方图显示其估计值集中分布在稳定的高概率区间，峰值显著且分布形态稳定。这些结果从数值指标和图形诊断两个层面共同验证了模型参数估计的稳健性，确保后续分析结论建立在可靠的统计基础之上。此外，HDDM 的分层特性使其参数估计主要依赖于试次级数据而非被试数量 (Ratcliff & Childers, 2015)，这有效降低了小样本带来的影响。相关内容已在正文对应部分进行补充和完善。

3. 未提供具体的 R_{hat} 数值或轨迹图等。

回应：感谢专家的建议。相关数值和轨迹图已在正文中添加。具体而言，在结果部分中，我们补充报告了所有关键参数的 R_{hat} 值（ R_{hat} 值均接近 1，表明马尔可夫链达到了极佳的收敛水平）。此外，还补充了收敛相关图，具体包括所有迭代的轨迹图、拟合的最后 100 次迭代自相关图以及阈值参数 t 的估计直方图，作为模型参数估计可靠性和可重复性的佐证。

4. 实验 1 中高控制条件的漂移率显著降低，作者将其解释为“自主选择干扰证据累积速度”；根据 DDM 假设， v 降低通常表示区分目标与干扰物的能力下降，这与“增强注意捕获”的结论似乎存在逻辑冲突。

回应：感谢专家对理论解释的提问，以帮助我们进一步澄清研究发现。DDM 模型中漂移率(v)反映了区分目标与干扰物的速度，值越大区分速度越快，表明区分任务越容易（或更容易在干扰物中找到目标）。实验 1 中漂移率对应的是目标识别决策，漂移率越低意味着目标识别越困难——即单例干扰物产生了更强的注意捕获效应，从而对目标识别造成了更大干扰。该结果与非决策时间 t 的增加一致，当目标识别过程受到干扰时，被试需要延长非决策时间以完成对目标的知觉编码。

5. 缺少公开数据与代码，研究结果的可重复性遭到质疑；建议在开发存储上公开数据与代码。

回应：感谢专家对研究可重复性的重要建议。我们完全认同开放科学实践的价值。我们目前正在整理所有研究材料（包括原始数据、分析代码和模型数据等），并将在文章正式接收后的通过开放科学框架(OSF)平台完整公开，同时申请 DOI 永久标识符。

6. 被试群体的代表性问题，这似乎是很多“大学生心理学”研究的不足，作者应在适当地方说明研究结果/结论的跨群体(e.g., 非大学生)普适性。

回应：感谢专家关于研究群体代表性的建议。我们已在正文讨论部分添加对结果普适性的论述。

审稿人 2 意见:

当前研究结合行为、眼动和计算建模 (HDDM) 技术考察了两个维度的控制感对损失注意捕获的调节机制。研究发现: 基于选择的控制通过增强早期自动化定向和延长晚期注意维持增强捕获; 基于结果的控制通过优化晚期资源分配调节加工效率。当前研究的问题有趣, 方法也具有前沿性, 但在实验设计、数据分析和语言文字等方面还存在着一些问题有待解决。

主要问题

1. 引言部分对 HDDM 模型 4 个主要参数中的 3 个进行了解释, 但并未说明控制感与这些参数之间的可能关系, 也缺乏相关的明确假设, 建议在结尾段落补充对 HDDM 参数的具体预测。此外, 为何只考虑了 4 个参数中的 3 个, 而未考虑起始点 z ? 有无相关理论或者研究支持这一点?

回应: 感谢专家关于模型参数选择和理论假设的建议。1) 模型参数选择方面, 原文表述可能造成了误解。实际操作中, 所有拟合的模型都包括漂移率、决策阈值、决策偏向、非决策时间这四个基本参数, 并以正确(编码为 1)和错误(编码为 0)为上下边界。根据以往研究, 以正确和错误为边界且各条件均衡时决策起始点不会向某一边界偏倚(Zhang et al., 2023)。因此本研究将决策偏向(z)设置为默认值 0.5。相关内容已在 2.4 结果的建模方法说明中补充和完善。2) 控制感与模型参数的约束关系上, 一方面我们尚未发现通过 HDDM 建模来考察控制感如何影响注意捕获的文献参考, 另一方面本研究重点关注“控制程度 $\times$ 损失程度”对模型参数的约束。因此, 较难给出详细准确的模型预测。但我们根据模型参数在注意中的基本含义, 做出了相应预测。

2. 引言部分的内容偏重近年研究 (2020 年后), 但关键理论 (如 Rotter, Bandura) 的引用较老。可考虑补充近期元分析或理论综述 (如 2020 年后关于控制感的整合性论文), 以体现领域进展。此外, 引言部分虽然提到了“价值驱动注意捕获框架”, 但缺乏具体的解释。

回应: 感谢专家的宝贵建议。关键理论的依据已在正文相应部分增加了新近文献。同时, 正文也增加了对“价值驱动注意捕获”的简要说明。

3. 当前研究希望从选择和结果两个维度考察控制感对损失注意捕获的影响，但从操纵上来看，实验 1 中基于选择的控制感和实验 2 中基于结果的控制感在本质上似乎是一样的！在实验 1 中“高控制条件下，被试按空格键后，白色外框立即停止；低控制条件下被试按空格键后，电脑发出报错声，白色外框继续转动，由程序控制停止在预先设定的损失分值上”。低控制条件下，被试也进行了按键（做出了“选择”），只是结果由于电脑报错与预期不符。这与实验 2 对预期与结果匹配程度的操控看起来是一样的。因此，两个实验似乎考察的是同一个维度！

回应：感谢专家对实验设计的提问，使我们能对原文中未表达清晰之处进行说明和改进。实验 1 主要从是否具有施动感（即被试行为与结果的时空连续性）的角度来操纵基于选择的控制。施动感的研究表明，个体做出行为和行为结果出现的时间间隔会影响个体的施动体验（Moore & Obhi, 2012）。实验 1 高控制条件下，被试按键后光标立即停止，即被试会认为光标停止是由自己按键导致的（即基于选择的控制）。低控制条件下，被试按键后触发电脑报错声且光标延迟 1000ms 停止。根据施动感研究，这种物理上的间隔会破坏被试心理上行为与结果的时空连续性（施动感生成的时间窗阈值为小于 300ms），导致施动感受损（Moore & Obhi, 2012）。实验 2 主要从结果是否符合预期的角度操纵基于结果的控制感。实验 2 所有条件下，被试均具有施动感——无论光标是否停在被试期望的位置上（符合预期时，光标直接停在被试选择的数字上；不符合预期时，光标停在相邻的数字上），被试按键和光标停止的时间间隔均小于 300ms，行为和结果之间始终具有时空连续性。实验 2 在保持高施动感的前提下，通过改变结果匹配概率（50%/75%/100%），诱发结果控制感梯度变化。我们对正文相关内容进行了调整和改善。

4. 当前研究关注的是控制感对损失注意捕获的调节，但从研究设计来看和结果来看（在实验 1 中损失大小主效应不显著），很难说明注意捕获是否是由损失引起的，还是仅仅是由颜色引起的！对于损失差异的不显著，作者认为是“实验 1 人为匹配了高、低控制条件的总损失分值，从而减弱了损失分值差异性（条件间的，和条件内的分值没有关系）”。但虽然总损失分值匹配，但不同损失程度的差异依然存在，且高、低损失程度的差异还不小（-50 vs. -5）。因此，这并不能解释为何损失程度对损失引起的注意捕获无显著影响！

回应：感谢专家对结果解释的提问。颜色凸显效应方面，实验 1 损失大小主效应不显著确实导致了区分损失效应和颜色效应的潜在困难。但控制程度的主效

应显著，表明实验 1 结果受到控制感的调节，而非简单的颜色凸显效应。损失程度影响方面，原文关于“实验 1 人为匹配了高、低控制条件的总损失分值，从而减弱了损失分值差异性”的表述可能造成了误解。如前所述（问题 3 回应），实验 1 的高、低控制条件分别对应施动感和无施动感两种条件。一方面，两种控制条件下损失分值均包括-50，-5 和 0 三种；另一方面两种控制条件被试损失的总分相同（人为控制）。即实验 1 中高低控制条件下被试的施动感虽然不同，但损失结果一致且总体上符合预期（得到较低损失分值）。我们考虑施动感带来的明显心理差异和全局损失程度的高度一致可能掩盖了局部的损失程度差异。因此，实验 2 平衡施动感程度，系统改变总体结果与预期的一致程度，进一步检验损失程度是否影响注意捕获。我们已修正原文相关表述。

5. 要说明存在损失注意捕获效应，两个实验是否也应该分析目标和干扰物注视指标之间是否存在差异？

回应：感谢专家的宝贵建议。我们已在正文部分补充了目标和干扰物首次注视百分比的差异检验结果。实验 1 结果发现，高控制条件下对干扰物的首次注视百分比高于目标，而低控制条件下对目标和干扰物的首次注视百分比无差异。即高控制条件下损失关联干扰物在注意定向阶段快速捕获了注意。实验 2 结果发现，仅高控制条件下对目标的首次注视百分比高于干扰物，中等控制和低控制条件下对目标和干扰物的首次注视百分比无差异。

6. 在实验 2 中更高控制始终与更低的损失对应，那么到底是控制程度起作用还是损失程度起作用，抑或是两者共同发挥作用？

回应：感谢专家的提问，以帮助我们更好地澄清研究结果。理论上，如果仅考虑单独因素的影响，控制感越高越容易捕获注意(Kumar et al., 2015)；损失程度越大越可能捕获注意(Le Pelley et al., 2016)。本研究中，控制感越高，损失分值越低。如果是损失分值单独起作用，应该是低控制感条件（对应高损失分值）注意捕获最强。但实验 2 结果显示，高控制组(总损失最低)表现最差（正确率最低）。因此，实验 2 结果更主要地反映了控制程度的影响。但由于实验 2 控制程度的操纵是通过改变损失程度实现的，因此也不能完全排除两者共同发挥作用的可能。相关内容已在正文讨论部分补充论述。

7. 实验 2 中虽然仅含 t 模型的 DIC 值最小，但也应该考虑理论合理性和当前的研究目的。若已有文献或任务设计支持“实验操纵仅影响非决策时间”，则该模型是合理的；反之，若理论预期 v 或 a 也应受影响，则需要谨慎（特别是考虑到方差分析显示控制程度主效应不显著）。建议检查其他指标（如 WAIC、后验预测检验）是否同样支持仅含 t 的模型。

回应：感谢专家对模型检验的宝贵建议。我们已按照建议补充了后验预测检验，结果与基于均值的差异检验一致。相关分析已在正文中补充。

8. 作者在引言中提到的自我效能理论和控制点理论，但在讨论部分并未进行相应的论述。此外，讨论中也缺乏与前人相关研究的比较分析。建议补充。

回应：感谢专家的宝贵意见。我们已在讨论部分补充了对控制点理论和自我效能理论的论述，并增加了与前人研究的比较分析。

9. 讨论中对部分结果的解释存在矛盾，需要再仔细考虑。例如：

a) 作者提到“另一方面，当个体感受到高度可控性时，其目标导向行为可能被过度自信削弱，导致对潜在威胁的注意维持不足(Hockey, 1997)。实验 2 高控制条件下被试对目标的注视时间最短且正确率最低，体现了这种由高度可控引发的“认知松弛”效应。”根据作者的表述，高控制感减少对潜在威胁的注意。那么，在当前实验背景下，相比于其他条件，高控制感条件是不是应该表现出更少的对损失相关干扰刺激注意、更多的对目标的注意和更好的任务表现，而这与实验 2 的结果截然相反。

回应：感谢专家的宝贵意见，以帮助我们进一步完善对结果的解释。我们同意原文关于“认知松弛”的解释与实验结果存在矛盾，现已在正文相应部分进行了修订。新的解释框架将高控制条件下的行为表现重新理解为多巴胺能系统介导的“效能优化策略”而非简单的注意资源减少。具体而言，当控制感达到峰值（100% 匹配）时，系统会触发以下关键变化：首先高控制条件下引起强直性多巴胺释放促使认知系统转向高效信息抽取模式，表现为注视时间缩短；其次，这种高效策略的代价是目标表征的过度简化和错误过滤机制的失效，从而导致正确率下降（Schultz, 2016）。该解释与实验 2 数据中注视时间与正确率的同步下降现象更一致。

b) “实验 2 中无论基于结果的控制程度如何，被试的按键后都轮盘都会马上停止。因此，被试始终具有很高的自主控制感。这可能削弱或掩盖了结果控制程度对任务表现的影响。”如果实验 2 中被试始终有很高的控制感，那么基于结果的控制的操纵是否有效？

回应：感谢专家的提问，以帮助我们澄清结果。原文此处表达不准确。实验 2 中结果控制程度不是没有影响。一方面，被试主观报告的控制感评分在各条件间呈现梯度变化且差异显著（低控 3.3，中控 4.6，高控 6.6）；另一方面，高控制条件下搜索任务正确率最低，注视持续时间最短。这里想强调的是内外控因素（即基于选择的控制）比自我效能感（即基于结果的控制）对注意选择的影响更大。我们已在正文相应部分调整了表述。

c) “只有当控制感超过临界水平时（本研究中为高控制条件），多巴胺能系统的神经递质释放模式才会从阶段性转变为强直性，进而促进认知系统采用更高效的加工策略(Schultz, 2016)。”这里提到的更高效的加工策略是指什么？如何与当前的结果对应？

回应：感谢专家的宝贵意见，以帮助我们进一步统一对结果的解释。这里提到的高效加工策略即是先前回应（第 9 问 a）提到的由多巴胺能系统介导的“效能优化策略”。这种策略在实验 2 中表现为对目标的注意持续时间缩短，反应正确率下降。反映了系统为追求效率而牺牲了部分表征精度。这种效率-精度的权衡正是“高效加工策略”的核心特征（Schultz, 2016）。

10. 作者最后提出的双通路模型中，基于结果的控制通过优化晚期资源分配来调节注意加工效率，调节注意加工效率与实验 2 的结果如何对应？此外，调节注意加工效率与损失注意捕获有何关系？作者需要加强相关论述，使其更容易被理解。

回应：感谢专家的宝贵意见，以帮助我们进一步统一对结果的解释。这里也存在表述不准确的问题。“基于结果的控制通过优化晚期资源分配来调节注意加工效率”对应的也是上述回应中提到“效能优化策略”。这部分内容已在正文相应部分更正为“基于结果的控制通过权衡加工效率和精度来调节注意选择”。

次要问题

11. 摘要部分缺乏明确的结论，建议补充上。

回应：感谢专家的宝贵建议。我们已在摘要最后添加明确结论："本研究表明控制感通过双路径调节损失注意捕获：基于选择的控制通过增强早期定向和晚期维持增强注意捕获；基于结果的控制通过权衡加工效率和精度来调节注意选择。"

12. 文中交替使用“控制感”和“可控性”，但二者在文献中可能存在细微差异（如前者更主观，后者更情境依赖）。应当明确区分或统一术语，避免概念混淆。类似的还有：控制程度、内部控制程度、内控程度。

回应：感谢专家的宝贵意见。全文已做如下修改：1.统一使用"控制感"（强调主观体验）替代"可控性"表述；2."控制水平"统一为"控制程度"。

13. 2.2 实验仪器与材料部分提到“屏幕分辨率为 1920×1080，刷新率为 500Hz”，屏幕的刷新率有达到 500HZ 这么高吗？请作者确认！

回应：感谢专家指出的问题。正文已更正为："屏幕分辨率为 1920×1080，刷新率为 120Hz。眼动数据通过 SR Research Eyelink 1000 Plus 系统采集，采样率为 500Hz。"

14. 实验 1 中“两次测试间隔两周以消除高控制条件颜色-分值联结的潜在影响”，是否在第二次测试前进行了前测，确认消除了影响？

回应：感谢专家的提问。实验 1 第二次测试前没有进行前测。一方面是为了避免被试在测验过程中有意回忆导致混淆；另一方面，前人关于联结学习效果持续性的研究表明两周间隔足以消除颜色-分值联结(Anderson, 2015)。已在正文相应部分补充说明。

15. 实验 1 和实验 2 学习阶段的转盘任务中白色外框的转动速度是如何确定的，为何两个实验不一致（实验 1：10.47 rad/s；实验 2：2.09 rad/s）？此外，在测试阶段的任务中，形状单例（搜索目标）与颜色单例（损失关联干扰物）位置平衡是如何实现的，是指二者位置相对，夹角为 180°吗？

回应：感谢专家的提问。转速差异方面，实验 1 采用高速(10.47 rad/s)是为了避免被试出现天花板效应。实验 2 降低转速(2.09 rad/s)是为了使被试在高控制条件下总能选中低损失分值。位置平衡方面，目标与干扰物夹角包括±60°、±

120° 和 180°，每个角度出现次数相等。已在正文相应部分补充说明。

16. 两个实验学习阶段的正确率在不同条件下是否存在显著差异？建议补充相应分析。

回应：感谢专家的提问。已在正文相应部分补充。

17. 2.4 中“剔除正确率低于 70% 的被试(实验 1 和 2 分别剔除 3 人和 2 人)，以及正负 3 个标准差外的极端值(实验 1 和 2 各阶段数据剔除量小于 4%)”，这里是什么的 3 个标准差？正确率还是反应时？数据剔除量是指试次还是被试？请作者描述清楚。

回应：感谢专家的提问。本研究数据剔除标准为：①被试正确率<70%；②单个试次反应时超出总体均值±3SD。剔除量指的是试次。已在正文相应部分修改。

18. 关于 HDDM 的参考文献建议添加最近的，并且说明为何不考虑含 z 的模型。此外，两个实验中关于 HDDM 模型比较的结果建议用一个表格呈现，内容包含模型参数、DIC 值以及模型收敛情况等，这样可能会清晰直观一些。

回应：感谢专家的宝贵建议。已补充相关 HDDM 参考文献，优化模型参数报告形式，说明参数选择依据。特别是不考虑含 z 的模型的原因已在引言部分补充相关说明。

19. 实验 2 中正确率上控制程度主效应显著时的事后检验是否对 p 值进行了多重比较校正？

回应：感谢专家的提问。本研究中涉及三水平及以上的事后检验均采用 Bonferroni 校正。相关说明已在正文相应部分补充。

20. 图 1： a) 建议将图 1 的两个部分合并在一幅图上，而非分开的两张图片； b) 注视点是否位于屏幕中央？如果是，请在流程图中正确呈现。如果不是，请说明为何不是？ c) 注视点呈现时长是 750ms 还是大于 750ms？如果是大于 750ms，那么具体时长如何确定？

回应：感谢专家的宝贵建议。相关图例已修改调整。注视点时间是 750ms。

21. 建议实验 1 和实验 2 中后验概率分布图的纵轴标题最好改成中文，并补

充上横轴标题。

回应：感谢专家的宝贵建议。相关图例已修改调整。

22. 语言文字： a) 2.4 结果部分“其次，分别对高、低控制条件下被试主观报告的的控制感程度和动机水平进行配对样本 t 检验。”中“主观报告的的控制感程度”多了一个“的”。 b) 2.4 结果部分“剔除首次眼跳起点偏离中央注视区 2° 以上以及扫视起始时间短于 80 ms 的试验(少于 8%)。”这里的“试验”应为“试次”。 c) 2.4 结果部分“高控制条件($M = 2001.792$)注视目标时的瞳孔直径显著高于低控制条件($M = 1762.535$)”、“高控制条件($M = 413.035$)注视目标的时间显著高于低控制条件($M = 370.716$)”。瞳孔直径和注视时间用“高于”是否不太合适？其他类似地方，请作者自行检查并修改表达。 d) η 应为斜体。 e) “50%、75%和 100% 匹配分别对应高、中和低控制。以低控制条件为例，在 50%的试次中，被试按键后外框停在最近的彩色正方形上，选中期望的低损失分值(-50 分)；剩余 50%的试次中，被试按键后外框停在最近的白色正方形上，选中不期望的高损失分值(-100 分)。”这部分的表达让人迷惑，50%匹配对应的到底是“高控制”还是“低控制”？ f) 讨论部分“种控制驱动的自动化加工在进化上可能具有适应性意义”，是否应为“这种控制.....”。 g) “后续研究可以采用认知神经可续技术进一步探究该机制的生物基础”，是否应为“认知神经科学技术”

回应：感谢专家细致的宝贵建议。已在正文部分按专家建议优化表达问题。

第二轮修改

审稿人 1 意见：

感谢作者对相关问题进行了细致地回应，不过有两个地方可能需要注意：

1. 作者对两个实验补充了后验概率差异分析（当然这很好），不过原本意见的第 7 条中的后验预测检验是指通过比较模型生成的模拟数据和真实数据检查模型是否合理，如果模型是合适的，用这些参数从模型中生成“新数据”应该和真实观测数据相似。该意见的原意是确认他指标或方法（如 WAIC、后验预测检验）是否同样支持仅含 t 的模型，希望作者通过计算所有模型的 WAIC(Widely Applicable Information Criterion)或者进行后验预测检验来说明实验 2 中使用仅含 t 的模型是合理的。

回应: 感谢专家的耐心说明和详细指导, 为我们确认模型合理性指明了方向。我们选取了后验预测检查(Posterior Predictive Check, PPC)直接检验实验 2 中使用仅含非决策时间(t)的模型是否合理。具体做法如下, 使用已拟合的 t 模型, 从后验分布中抽取 500 个参数样本, 为每个样本生成与观测数据相同大小的预测数据集。通过比较预测数据与观测数据分布特征的一致性, 进而验证模型解释数据的能力(Pan et al., 2025)。此外, 我们专门针对反应时(rt)分别在三种控制感条件下生成后验预测检查图(图 1A)。其中, 我们限制使用 100 个后验预测样本并设置固定的随机种子, 确保可视化结果的可重复性和计算效率(Pan 等, 2025)。结果显示, 观测数据与预测数据分布高度重合, 这说明 t 模型能够很好地解释数据。最后, 我们进行了 t 模型的统计推断, 计算了 95%最高后验密度区间(HDI), 并基于 $[-0.2, 0.2]$ 的 ROPE 区间进行实质性检验(Pan 等, 2025)。结果表明(图 1B), 三种条件下, 95%HDI 和 ROPE 之间没有重合, 表明该参数与 0 存在可信差异。因此, 从 PPC 及后验统计检验综合来看, 实验 2 中使用仅含非决策时间(t)的模型是合理的。但考虑到论文篇幅限制, 上述补充检验未加入正文。

Pan, W., Geng, H., Zhang, L., Fengler, A., Frank, M. J., Zhang, R.-Y., & Chuan-Peng, H. (2025). dockerHDDM: A user-friendly environment for Bayesian hierarchical drift-diffusion modeling. *Advances in Methods and Practices in Psychological Science*, 8(1), 25152459241298700.

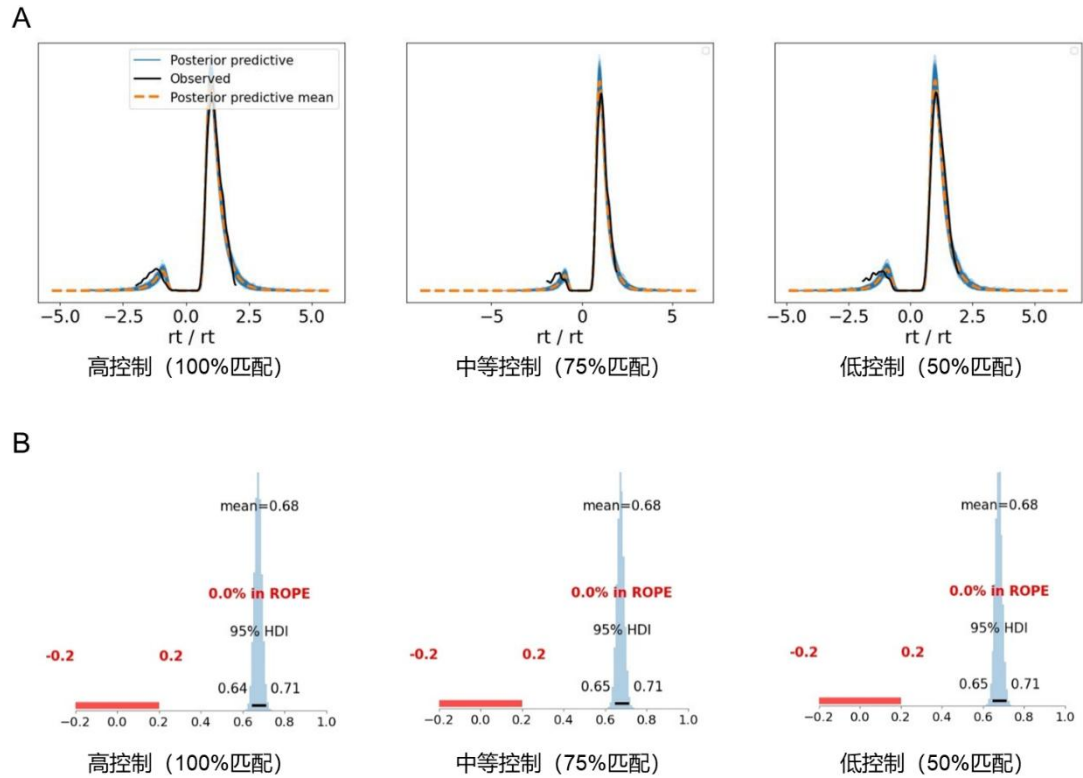


图 1 实验 2 高、中、低控制条件下 A) PPC 观测与预测数据结果和 B) 统计推断结果

2. 参考文献部分引文 “Müller, P., & Wentura, D. (2023). Undeserved reward but not inevitable loss biases attention: Personal control moderates evaluative attentional biases in the additional-singleton paradigm. *Quarterly Journal of Experimental Psychology*, 76(4), 888 – 904.” 与正文中的 “Müller 和 Wentura(2022)” 的年份不同，建议作者对参考文献全面进行检查，查看是否存在类似或者其他错误

回应：感谢专家的宝贵意见。我们已经对全文正文及参考文献列表进行了检查，以排除类似问题。

主编意见:

退稿。做重大修改后，可以寄回，直接终审。注意，作者只有最后的一次机会。总体而言，这是一个有意思的研究。但文章存在一些比较重大的缺陷。

1, 论文的文字、写作水平比较差。比如摘要部分，不清不楚，没有遵循基本的规范。

回应：感谢主编的宝贵意见，使我们的论文在写作方面得到了重要的提升机会。我们已根据意见对全文进行了全面修订，具体包括：强化论述逻辑，优化语言表达以使其更为准确流畅，并完善写作格式以符合学术规范。特别是针对摘要部分，我们按照学术规范进行了重写：明确了研究背景、目的、方法、结果和结论的完整结构。

2, 实验设计中，在一些可以进行被试间平衡的地方，似乎没有做平衡。比如高、低控制条件的实验顺序，似乎总是先做高控制，后做低控制；高、低、无损失条件与颜色的对应关系，似乎也总是固定的一套对应关系。这削弱研究结论的可靠性。如此，则作者必须明确承认，并讨论可能存在的影响。

回应：感谢主编的宝贵意见和建议，使我们可以对实验设计中的关键细节进行更深入的说明和澄清。

首先，实验 1 始终按先高控制、后低控制条件的顺序进行施测。这样做的目的有两个。其一，排除损失总分对结果的干扰，保证实验 1 高、低控制感的差异源于是否可以自主选择，而非总体损失程度的不同。具体而言，在高控制条件完成后，电脑可根据被试的实际总分，为随后的低控制条件随机分配试次得分，从而确保两种条件下的损失总量一致。其二，最大限度保证高控制条件操纵的纯粹性。若顺序相反，先进行低控制条件，则电脑预设的总分可能无法匹配被试在高控制条件下的实际表现，从而影响高控制感操纵的纯粹性。该安排可能会带来练习或疲劳效应。为控制施测顺序的潜在干扰，实验 1 在两次测试间设置了两周的间隔。一方面，练习或疲劳效应通常随测试间隔的增加而减弱。我们比较了实验 1 高控制条件后 50%试次和低控制条件前 50%试次的任务表现。结果发现，反应时上，低控制条件前 50%试次显著慢于高控制条件后 50%试次， $t(27) = 2.428, p$

= 0.022, Cohen's $d = 0.459$ 。正确率上, 两者差异不显著, $t < 1$ 。这表明高控制测试中可能存在的练习效应并未延续至两周后的低控制测试中。另一方面, 对彩色单例干扰物的抑制和脱离依赖于注意控制能力。研究表明, 注意控制本身具有一定跨时间稳定性, 单纯的重复测试难以使其得到提升(Unsworth et al., 2024; Unsworth et al., 2025)。如, Zhang 等人(2024)在一项注意控制训练研究中发现, 未接受任何训练的控制组被试, 其注意定向能力在间隔一周的前、后测中并未发生显著变化。上述实验顺序的操纵已在实验 1 方法部分明确说明, 并在实验 1 讨论部分对其可能产生的影响进行了论述。

其次, 关于高、低、无损失条件与颜色的对应关系, 尽管实验设计已进行被试间的平衡处理, 但原文表述存在歧义, 我们已对正文相关内容进行了改进。具体设计如下: 转盘刺激包含白色与彩色两种方块。彩色方块的可选颜色有两套: 1) 红色(RGB:255,144,144)、蓝色(RGB:15,170,225)和绿色(RGB:114,201,101); 2) 橙色(RGB:255,180,85)、紫色(RGB:255,144,225)和青色(RGB:85,255,255)。高、低控制条件各使用其中一套颜色, 该对应关系在被试间平衡。同时, 每个方块内都有一个数字, 显示选中该方块将被扣除的分数。白色方块始终显示“-100 分”, 作为背景分数; 彩色方块则显示“-50 分”(高损失)、“-5 分”(低损失)或“0 分”(无损失), 作为目标分数。每种控制条件下, 三种目标分数分别与一种颜色唯一绑定, 该绑定关系在不同被试间进行随机平衡, 但对每位被试保持一致。原文流程图仅为一种可能的绑定示例。此外, 实验 2 中的颜色-损失的联结也进行了相同的被试间平衡程序。

3, 更重要的, 作者基于实验 1 和实验 2 的结果模式差异进行了很多推论, 但并没有在统计上直接比较实验 1 和实验 2 的数据。这是绝对不允许的。如果这些比较重要, 就必须提供统计检验结果;如果这些不重要, 就不能提供基于你眼睛看到的所谓差异的见解。

回应: 感谢主编提出的宝贵意见与建议, 使我们认识到原文在此问题分析上的局限性, 并为提升论文科学性指明了方向。我们完全同意, 原文在分析实验 1 与实验 2 的结果差异时存在过度推论的缺陷。本研究的核心发现应聚焦于“基于选择与基于结果的控制所表现出的差异性调节路径”, 即: 前者通过加速早期定

向与延长晚期维持来增强注意捕获；后者则通过权衡效率与精度来调节注意选择。而基于实验间差异对两种控制的权重进行比较，并非论文重点。同时，鉴于两个实验在控制感操纵上并不具备直接可比性，此类比较亦缺乏充分依据。因此，在修改稿中，我们删除了有关实验间差异的过度推论，并将此问题留待未来研究系统探讨。