

《心理学报》审稿意见与作者回应

题目：过程与结果线索信息对 4~7 岁儿童及成人对“聪明”和“努力”判断与评价的影响

作者：赵欣、张爱颖、张文宇

第一轮

审稿人 1 意见：

该研究聚焦儿童对聪明和努力的理解和评价，这一问题具有较好的理论和实践意义。研究者在前人的范式上进行创新，突破以往简化单一结果情境的局限，在范式中第二次考试的成绩根据儿童在第一次判断中的选择进行动态调整，通过比较两次考试情境中儿童的判断，来推断儿童对聪明与努力是基于过程导向还是结果导向。我的主要问题如下：

意见 1：首先，在问题提出中，作者提到“现有研究尚未探究儿童能否以及如何通过过程和结果信息对聪明和努力进行推测，以及其推测主要关注过程信息还是结果信息。”这一表述有失妥当，前人研究也有关于过程和结果信息对聪明和努力的理解。

回应：我们感谢审稿专家的宝贵意见。在修改稿中，我们已对这一表述进行修改，更加细致具体地阐述现有研究中尚未解决的问题，修改之后的段落为：“然而，现有研究尚未特别关注儿童在对聪明和努力进行推测的时候主要关注过程信息还是结果信息。现有的研究范式通常采用单一情境设定：儿童通常被要求比较两位故事人物，仅呈现故事人物一次任务中的过程与结果，其任务结果往往保持一致（如，都完成了相同数量的题目），只是其任务完成的过程（如，用时长短）有所不同。这种范式在一定程度上限制了我们深入理解儿童在评估“聪明”与“努力”时更依赖过程还是结果线索。一个关键的尚未被探究的问题是：当任务结果发生变化时，儿童对他人能力与努力的推测是否会受到影响？换言之，儿童在判断他人的能力与努力时，是更倾向于以结果为导向（即因结果变化而有所不同），还是以过程为导向（即更关注过程中的信息，不受结果的变化影响）？”

意见 2：有关被试的年龄段选择，虽然作者强调这个年龄段中儿童对聪明和努力的理解发生重要的变化。但是本研究结合中国文化背景下的学业成就情境，该情景更适合学龄儿童。而且用本研究的范式探讨学龄前到学龄期儿童对努力和聪明的理解的发展可能更有意义，比如儿童对努力者的偏好是否会随着学习经验的增長而发生变化，所以建议作者进一步扩展年龄。

回应：我们感谢审稿专家的这一宝贵意见。本研究选取了 4~7 岁儿童为研究对象，正是希望探讨从学龄前到学龄早期儿童在理解“聪明”和“努力”方面的发展变化。研究中所有 7 岁儿童及部分 6 岁儿童已进入小学阶段，我们的结果也确实揭示了从学龄前到学龄期在相关认知上的发展趋势，尤其是在对“努力”的理解方面，儿童在 5-7 岁之间有重要的发展变化。此外，既有研究也表明，5-7 岁是儿童在能力与努力认知上发生显著转变的重要时期（如 Yang et al., 2024）。我们在引言第 1.4 部分补充了关于年龄段选择的说明，具体修改见稿件第 4 页。

此外，我们已在讨论部分补充了未来研究的展望，指出后续工作可进一步拓展至更宽的年龄范围。修改之后的段落如下：

第 18 页：“小学阶段的儿童往往高度认同努力的价值，并倾向于给予努力者更高的评价；然而，随着年龄和经验的生长，他们逐渐意识到努力并非总能带来更好的结果，并开始感知到努力所伴随的成本(Jiang et al., 2018)。因此，他们对天赋与努力的看法也逐渐趋于更加平衡。未来研究可进一步扩大年龄范围，探究儿童在中学阶段对于能力者与努力者的评价模式，以及在何时出现对努力偏好下降的转折点。”

第 19 页：“未来研究有必要进一步拓展年龄范围，考察儿童中期及青少年阶段的发展趋势。随着儿童对“努力”及其所带来的可能的结果有了更多的体验和感受，他们对努力与能力的判断与评价是否会进一步发生变化，值得深入探究。”

意见 3：可能有所疏漏，结果分析中有关儿童对于能力者和努力者的评价，我只找到第一次考试，没有看到第二次考试。建议作者进一步确认或凸显，因为两次考试的变化是本研究的亮点。

回应：在原文稿中，我们未对第二次考试中的评价进行分析，原因是该部分分析未在预注册方案中注册，且当时并无明确假设。然而，受审稿专家意见的启发，我们认为有必要补充此部分分析。因此，我们在修订稿的结果部分中新增了儿童在第二次考试中对两位故事人物的评价结果。由于第二次考试中两位人物的表现结果（即名次）存在差异，本分析可以帮助我们考察儿童与成人在评价时是否受到结果（即长用时人物与短用时人物相对排名高低）的影响。结果显示，儿童与成人的评价均显著受到结果的影响：当长用时人物的排名较高时，参与者对其的评价也更高。此外，对于儿童而言，年龄与结果之间存在显著交互作用。随着年龄的生长，儿童在长用时人物排名较高时对其评价更高，而在长用时人物排名较低时则未发现显著的年龄效应。相关结果详见第 14-15 页。

.....

审稿人 2 意见：

意见 1：很荣幸受邀审阅本稿件。该研究聚焦于儿童对“聪明”和“努力”概念的理解与评价，通过精心设计的实验，对相关认知和行为进行了深入探讨和分析。但是，在理论背景的论述和分析方法的描述上，仍有提升空间。具体来说，理论部分对相关概念的界定和前人研究的梳理尚不充分，难以完全支撑研究假设；方法部分对混合效应模型的具体设定、变量处理及多重比较校正等细节描述不够明确，建议进一步补充完善，以增强研究的透明度和科学严谨性。

回应：非常感谢审稿专家对本研究的细致审阅和建设性意见。我们十分感激您对研究选题与实验设计的肯定，也认真采纳了您关于理论背景与方法描述的宝贵建议。根据您的意见，我们对论文的相关部分进行了较为全面的修改与补充，以提升研究的清晰度与科学严谨性。

具体而言，在理论部分，我们进一步梳理和扩展了有关儿童对“聪明”和“努力”概念理解的国内外研究，明确界定了相关核心概念，并阐述了本研究假设的理论依据，从而使研究的逻辑链条更加完整。

在方法部分，我们补充了对混合效应模型设定的详细说明，变量处理方式以及多重比较校正的具体信息，同时将原始数据和分析代码上传到了 OSF 网站，以增强研究方法的透明度与可复现性。

具体回复及修改见下文对具体意见的回复。

意见 2：方法部分引入了成人样本，但缺乏在引言和理论背景中对成人纳入的合理性与必要性的阐述，也未在结果与讨论中对成人数据进行充分分析和解释。建议作者补充成人样本的理论依据、研究目的及其在整体研究框架中的角色，明确成人组分析的具体方法和结果，提升研究的整体连贯性和说服力。

回应：非常感谢审稿专家对本研究提出的中肯意见。我们充分理解并认同您关于成人样本的理论依据与研究作用需要进一步阐明的建议。根据您的意见，我们对论文的相关部分进行了补充与修改，以增强研究的逻辑连贯性与说服力。

具体而言，在引言和理论背景部分，我们新增了对成人相关研究的综述（见第 4 页 1.3 部分第一段），以往有关研究主要关注的是成人对于能力者与努力者的偏好。此外，在 1.4 部分的第二段，我们指出，成人组的加入旨在作为儿童数据的发展性对照组，为儿童结果的解释提供一定的基准。

在方法部分，我们进一步说明了成人样本的招募方式及测量方式（见第 5 页 2.1 部分第一段，以及第 7 页 2.2 部分最后一段），成人样本为通过社交平台（如，微信朋友圈）招募，年龄范围为 18-52 岁。测量方式方面，成人组在问卷星上填写问卷，问卷流程去除了儿童实验中第一次考试前有关光荣榜的概念介绍环节，其余设置与儿童组保持一致。成人组阅读了与儿童组相同的文字描述和视觉材料，并回答与儿童相同的问题。

意见 3：为什么选择 4-8 岁年龄段，需要给出更多的理论考虑。

回应：本研究主要关于 4 至 7 岁(4.02~7.95)的儿童。之所以选择这一年龄段，是基于以往研究发现儿童在 4~5 岁左右可以对聪明与努力进行推测(Muradoglu & Cimpian, 2020; Zhao & Yang, 2023),并且其对聪明者与努力者的偏好在 5 至 7 岁之间发生重要的发展变化(Yang et al., 2023)。选择这一年龄段可以探讨从学龄前到学龄期儿童对努力和聪明的理解可能的发展变化。我们在引言第 1.4 部分补充了关于年龄段选择的说明，具体修改见稿件第 4 页。

此外，我们在讨论中也增加了有关未来研究可以进一步扩展年龄的展望。具体见第 18 页与第 19 页。

意见 4：在文献综述部分，建议作者调整论述顺序，先系统介绍儿童对“努力”和“能力”概念的认知理解，着重理论和实证研究的认知视角；随后再论述儿童的偏好表现及其相关研究。

回应：非常感谢审稿专家的宝贵建议。我们已经根据这一建议对文献综述部分进行了顺序的调整，使得论述逻辑更加清晰。具体而言，我们首先介绍了儿童对“能力”与“努力”的概念的认知理解，随后阐述了儿童在理解能力与努力的过程中可能形成的不同的认知导向，之后再论述成人及儿童对于能力者及努力者的偏好，并引出本研究的问题与假设。具体修改见文稿第 2-4 页。

意见 5：“但是，现有研究尚未探究儿童能否以及如何通过过程和结果信息对聪明和努力进行推测，以及其推测主要关注过程信息还是结果信息”，这个是非常重要的以往研究的不足，需要在前言的部分进行更详细的论述，以及本文打算如何去解决。不能只是在方法部分，才让读者去思考，这到底是什么。

回应：我们在前言中相应的部分更加细致具体地阐述现有研究中尚未解决的问题，并论述了本文打算如何解决（见文稿第 3 页）。修改之后的段落为：“然而，现有研究尚未特别关注

儿童在对聪明和努力进行推测的时候主要关注过程信息还是结果信息。现有的研究范式通常采用单一情境设定：儿童通常被要求比较两位故事人物，仅呈现故事人物一次任务中的过程与结果，其任务结果往往保持一致（如，都完成了相同数量的题目），只是其任务完成的过程（如，用时长短）有所不同。这种范式在一定程度上限制了我们深入理解儿童在评估“聪明”与“努力”时更依赖过程还是结果线索。一个关键的尚未被探究的问题是：当任务结果发生变化时，儿童对他人能力与努力的推测是否会受到影响？换言之，儿童在判断他人的能力与努力时，是更倾向于以结果为导向（即因结果变化而有所不同），还是以过程为导向（即更关注过程中的信息，不受结果的变化的影响）？本研究将设计一个创新性的包含两次任务的范式，儿童将观看两个故事人物在两次考试任务中的过程和结果。我们通过任务完成时间的长短来操控两位人物的过程信息，通过考试成绩排名操控其结果信息。具体而言，在第一次考试中，其中一位故事人物完成任务的用时更短，两位人物最终取得了相同的成绩结果。儿童需判断谁更聪明、谁更努力，并对两位人物进行比较评价。随后，儿童将被呈现这两位人物参与第二次考试的过程与结果。关键在于，第二次考试的成绩将根据儿童在第一次判断中的选择进行动态调整，即儿童先前认为更聪明（或更努力）的人物在第二次考试中成绩落后于另一位人物。随后，儿童需再次判断谁更聪明或更努力。简言之，通过比较两次考试情境中儿童的判断，我们可以推断其对聪明与努力的认知导向：若儿童主要依据任务过程信息（如，用时长短）进行判断，即便结果发生变化也不改变判断，我们则认为其对相应特质的认知为“过程导向”；相反，若儿童会根据考试排名的变化改变自己的判断，将第二次考试排名更高的人物判断为更聪明（或更努力），我们则认为其对相应特质的认知为“结果导向”。

意见 6：“关键在于，我们操控，在聪明者改变情境中，被试在第一次考试情景中认为“更不聪明”的人物进步到第二名，认为“更聪明”的人物维持在第三名。在努力者改变情境中，被试在第一次考试情景中认为“更不努力”的人物进步到第二名，认为“更努力”的人物则维持在第三名。” 1）这里不是很清晰。比如，“更聪明”与“更努力”的人物是如何确定的？是依据儿童第一次考试中的评价，还是研究者事先设定的？ 2）另外，根据图 1，似乎后续的操控是基于两种不同的“做题用时”与“名次”排列组合设计（如乐乐用时短但名次靠前或靠后），儿童需要根据这一设置判断谁更聪明、谁更努力。建议作者在方法部分对该设计逻辑进行更清晰的说明，包括时间用时、名次与聪明/努力评价之间的关系，以及儿童的判断任务。

回应：非常感谢审稿专家提出的宝贵意见。针对您的问题，我们已在方法部分进行了详细澄清与补充。具体回复如下：

针对第（1）点，“更聪明”与“更努力”的人物是根据儿童在第一次考试情境中对“谁更聪明”或“谁更努力”的判断确定的，而非由研究者事先设定。

针对第（2）点，我们已在文中进一步说明操控逻辑及其与儿童判断的关系。研究中的操控是基于儿童第一次判断结果而设定的。以“聪明者改变情境”为例，在第二次考试中，两位故事人物的考试过程（即做题用时）保持不变，但其考试结果（即光荣榜排名）由主试根据儿童第一次的迫选问题中的回答进行调整。如果儿童在第一次考试中认为用时短（或长）的人物更聪明，则该人物在第二次考试中的名次会被另一人物反超。由此，儿童最初认为“更不聪明”的人物在第二次考试中进步至第二名，而其认为“更聪明”的人物则保持在第三名。

“努力者改变情境”的操控逻辑与此相同：如果儿童在第一次考试中判断用时短（或长）的人物更努力，则该人物在第二次考试中的名次会被另一人物反超。这样，儿童最初认为“更不努力”的人物在第二次考试中进步至第二名，而其认为“更努力”的人物则维持在第三名。

我们采用这一设计旨在探究,当儿童最初认为更聪明或更努力的人物在第二次考试中被其认为较不聪明或较不努力的人物超越时,儿童是否会因结果变化而改变自己的判断。如果儿童在第二次考试中改变判断(即认为成绩更好的人物更聪明或更努力),我们将其视为结果导向型认知;若儿童判断保持一致,则视为过程导向型认知。

相关修改已在方法部分补充说明(见第6-7页)。

意见 7: 参与的成人和儿童,有什么关系?成人是大学生吗?年龄范围是多少?

回应: 感谢审稿专家的关注。本研究中参与的成人与儿童样本之间没有直接关系。成人样本通过社交媒体公开招募获得,参与者不限定为大学生。成人参与者的年龄范围为 18 至 52 岁。相关信息已在论文的“被试”部分进行了补充与说明。

意见 8: 分析方法部分,需要补充更多细节。比如使用了什么软件分析,比如如果是 R, 用了什么 package 等。

回应: 本研究的数据分析使用 R 软件进行,主要采用了 *lme4* 和 *lmerTest* 包(Bates et al., 2015)进行混合效应模型分析。相关信息已在论文结果的数据分析计划部分中补充说明(见文稿第7页2.3部分)。同时已将数据分析的代码上传到OSF网站。

意见 9: 建议在统计方法部分补充对样本量和检验效能(power)的说明。

回应: 我们已在方法部分补充了对样本量及检验效能(power)的说明(第5页)。使用 *G*Power 3.1* 进行的事后效能分析结果显示,对于儿童样本,在样本量为 120、显著性水平 $\alpha = 0.05$ (双尾检验)、最多包含 7 个预测变量的条件下,线性多元回归对检出小到中等效应量($f^2 = 0.15$)的检验效能为 87%。对于成人样本,在样本量为 100、显著性水平 $\alpha = 0.05$ (双尾检验)、最多包含 3 个预测变量的条件下,线性多元回归对检出小到中等效应量($f^2 = 0.15$)的检验效能为 90%,均达到较为理想的水平。

意见 10: 数据分析方法部分,过于简单。建议在结果前,单独加一个小结,介绍数据分析计划。

回应: 我们已在结果之前增加了数据分析计划这一部分:“数据分析采用 R 软件的 *lme4* 包和 *lmerTest* 包(Bates et al., 2015)。所有分析均按照预注册方案进行。首先,我们分析被试在第一次考试中对“聪明”与“努力”的推断与理解,其次结合两次考试中被试的回答分析其对“聪明”与“努力”推断的认知导向,最后分析其对故事人物的偏好。对于每一类变量,因研究主要关注 4~7 岁儿童的发展变化,儿童数据分析中将年龄作为连续变量处理;成人数据则单独分析。儿童与成人的比较分析结果见附录 2。按照预注册方案,我们在建模时均从完整模型出发(在适用时包含交互项),若交互项在似然比检验(likelihood ratio test, LRT, 使用 *drop1()* 函数)中不显著,则予以剔除。”见文稿第7页2.3部分。

意见 11: 关于儿童组和成人组的分析设计,发现儿童组分析中年龄被作为核心自变量纳入模型,而成人组分析却未将年龄作为核心变量,且两组数据被分别独立分析。如果成人组的主要目的是作为儿童组的对比样本,那么建议将儿童与成人样本合并到同一模型中,通过加入“组别(儿童 vs. 成人)”以及其与年龄等变量的交互项,来直接检验不同群体间的发展差异和年龄效应的异质性。这样不仅能提高统计效能,还能更系统地比较两个年龄群体的异同,避免分开分析带来的解释割裂与统计效率降低。建议作者在研究设计和分析方法部分补充这

一点的论述，并考虑相应调整分析策略。

回应：非常感谢审稿专家的细致建议。关于本研究的分析策略，我们选择将成人组作为基线，以帮助读者理解在同一任务中成人对相关问题的典型反应模式，而非用于直接比较儿童与成人之间的差异。本研究的核心目标在于考察儿童群体内部的发展变化，因此在儿童样本中将年龄作为连续变量纳入模型，以揭示随年龄增长而出现的认知与判断变化趋势。这一分析策略与我们在研究预注册中所设定的方法保持一致。

与此同时，为回应审稿专家的建议，我们在论文附录2（见文稿26-30页）中补充了儿童（整体样本）与成人样本的对比分析，以供参考，从而使读者能够更系统地了解两组样本在关键指标上的总体差异。

意见 12：以下分析存在冗余——

在3.1对第一次考试中聪明和努力的推断和理解部分，先做了一个“聪明/努力合并”的混合模型，然后又分开做聪明评分和努力评分的模型，相当于重复检验了同一批数据里的两类结果。如果最终目的是解释聪明评分和努力评分各自的模式，其实直接分开做两组LMM就够了（聪明评分一个模型、努力评分一个模型），但是这样操作，是对同一批被试做了多次假设检验，就有多重比较问题，需要校正p值。但是，如果作者关注努力和聪明评分之间的差异和区别，则需要放到一个大的混合模型上去做，然后发现显著的交互作用或者主效应后，进一步做简单效应分析。

另外，分开分析儿童的聪明评分与努力评分中，在显著交互作用的基础上，作者进一步进行了简单斜率分析，以分别检验年龄在不同故事人物条件下对评分的预测作用。为控制多重比较引起的第一类错误率，需要简单斜率检验的p值进行校正，并报告了校正后的结果。在下面的几个维度的分析上，也存在类似的问题。

回应：感谢审稿专家的这一建议。在这一分析中，我们更加关注努力评分与聪明评分之间的差异与区别，因此保留了大的“聪明/努力合并”混合模型。这也与我们预注册中的分析策略一致。在发现显著交互作用后，我们进一步进行了简单效应分析。针对简单斜率分析中涉及的多次检验，为控制多重比较引起的第一类错误率，我们对p值进行了Holm校正，并在结果中报告了校正后的显著性水平，结论依然与之前的分析中的一致。具体修改见文稿第8页。

意见 13：在3.1对第一次考试中聪明和努力的推断和理解，分开分析儿童的聪明评分与努力评分中，对努力的评分，年龄和故事人物之间存在显著的交互作用。这里为什么不考虑用J-N法而用简单斜率？

回应：感谢审稿专家的这一建议，我们加入了J-N分析，发现从5.2岁起，儿童对于长用时人物与短用时人物的努力评分有显著差异

意见 14：在3.3对能力者与努力者的评价部分，“我们接下来分析儿童在第一次考试中对于评价问题的回答。我们对儿童的回答(0 = 短用时人物，1 = 长用时人物)进行二项线性混合效应模型分析，以问题(奖励、偏好、成功)、年龄及其交互作用作为自变量，加入情境(聪明/努力改变情境)作为控制变量。结果发现，交互作用不显著($ps > .265$)，因而删除交互项。最终的模型表明，问题的主效应不显著($ps > .296$)，年龄的主效应显著($B = 0.46$ SE = 0.10, 95% CI = [0.26, 0.67], $p < .001$)，这表明，随着年龄的增长，儿童对于用时长的人物评价更高。”当混合模型中年龄主效应显著且“问题”是被试内分类变量（奖励、偏好、成功）时，通常需

要进一步做简单效应分析。混合模型的主效应告诉你“总体上”年龄对儿童选择有影响，但“问题”有多个类别，可能年龄的效应只在某些问题类型上显著。简单效应分析就是分问题类型，分别检验年龄对儿童选择的影响，找出具体在哪种评价维度上年龄的影响最强或显著。如果作者不关注具体问题上的表现，那么一开始的时候，可以考虑把问题(奖励、偏好、成功)合成一个总均分（在计算信效度的基础上）。

回应：我们加入了分问题类型，分别检验年龄对儿童选择的影响（第13页）。年龄在所有问题类型中均显著正向预测儿童选择长用时人物的倾向(奖励: $B = 0.38, SE = 0.15, 95\% CI = [0.08, 0.67]$, Holm 校正后 $p = 0.014$; 偏好: $B = 0.60, SE = 0.16, 95\% CI = [0.29, 0.90]$, Holm 校正后 $p < 0.001$; 成功: $B = 0.42, SE = 0.15, 95\% CI = [0.13, 0.72]$, Holm 校正后 $p = 0.009$)。

意见 15：需要在文字描述结果部分，明确指出对应的图。图和内容，完全没有对应。

回应：感谢审稿专家的建议。我们已在文字结果描述中明确标注了对应的图。

意见 16：图 4， 右边的图，横轴要标注哪个是 0，哪个是 1.左边的图，看不懂是什么，建议给出详细的备注。

回应：感谢审稿专家的建议。对于图4 右侧的图，我们已在横轴加入了明确标注。关于左侧的图，如果审稿专家的疑问在于不同颜色线条代表的问题不明确，由于左图中颜色所代表的问题与右图一致，我们仅在右图保留了图注，以保持图示简洁。如果审稿专家认为有必要，我们可以在左图也添加图注以进一步说明。

意见 17：在前言部分，没有介绍成人。但是在讨论部分，把成人结果作为一部分进行讨论。没有前后一致。

回应：如前文所述，我们已在引言部分补充了对以往成人研究的综述，并对收集成人样本的理由进行了说明，以保持前后内容的一致性。

意见 18：研究不足可以讨论下被试的局限，年龄跨度等。

回应：如前文所述，我们在讨论部分补充了关于研究被试的局限，并展望了未来研究方向，包括拓展样本的年龄跨度，进一步研究儿童中期乃至青少年期的认知发展。

意见 19：建议在修改前言后，再系统修改讨论部分。现在前言和方法需要改动太多，不太好给出讨论的建议。

回应：在对前言、方法和结果进行了修改之后，我们已对讨论进行了相应的修改调整，以保持全文逻辑的一致性。我们非常感谢审稿专家的宝贵意见，并欢迎进一步的指导和建议。

审稿人 3 意见：

意见 1：研究较为系统地揭示了儿童对聪明和努力等成就相关概念的认知发展轨迹，但是缺乏相关机制的解释。为什么儿童会出现这样的认知变化？为什么成人在结果变化时保留原有判断？研究在讨论中提供了一些猜想，但是缺乏相关研究证据的支持。

回应：我们非常感谢审稿专家的宝贵建议。本文的核心目标在于揭示儿童及成人对聪明与努力等成就相关概念的认知发展轨迹，因此研究本身并未直接检验机制，我们承认这是本研究的一个局限。然而，在讨论部分，我们加入了针对潜在机制的系统性猜想，并提出完整验证

这些机制仍需未来研究进一步探索。

首先，关于儿童对于“努力”的推断发展稍晚于对“聪明”的推断，我们提出以下可能机制（文稿中修改见第16-17页）：

“一方面，这可能与家长和老师日常生活中对“聪明”和“努力”的讨论频率及语境差异有关。家长往往在儿童年幼时就频繁使用“你真聪明”这样的表扬，而在学校环境中，那些能快速回答问题或者高效完成任务的孩子也更容易被贴上“聪明”的标签。相比之下，关于“努力”的评价和讨论通常出现地较晚，往往在幼小衔接或小学阶段才开始被强调。与此一致，近期研究发现，成人对他人“聪明”的评价会影响儿童对努力价值的认知(Zhao et al., 2022)及儿童自身的努力行为(Ma et al., 2025)；同时，父母与孩子关于努力的讨论对儿童自身的毅力和坚持也有重要的影响(Haber et al., 2024; Lucca et al., 2019)。未来研究可以进一步探讨家长和老师关于聪明与努力的言语输入如何影响儿童对于这两个概念的理解与推断。

另一方面，“努力”的概念本身可能具有更高的认知复杂性。儿童较早即可通过直观线索（如完成速度、轻松程度）来判断他人是否“聪明”。而“努力”的线索则相对隐蔽，更多体现在持续思考、投入时间、反复练习、认真等过程性特征上。儿童需要一定时间的观察，以及一定程度的心理过程理解能力，才能识别这些行为背后的“努力”意图。此外，研究也表明，儿童对与“努力”紧密相关的“认知成本”这一概念的理解也同样需要发展时间(Radovanovic et al., 2023; Zhao & Kushnir, 2023)。未来研究还可采用更开放性的问题，探讨儿童在判断他人聪明或努力时依赖哪些信息与线索。”

第二，关于随年龄增长儿童对聪明和努力的判断呈现不同的认知导向，我们提出如下推测（文稿中修改见第17-18页）：

“首先，年幼儿童通常对结果信息更为敏感，这与在道德发展研究中的发现一致，年幼儿童在进行道德评价时往往更关注行为的结果，而非行为背后的意图(Cushman et al., 2013)。此外，如前文所述，年幼儿童对“努力”的概念理解尚不充分，因而更易依据外显结果进行判断。随着年龄的增长，儿童逐渐能够整合结果与过程信息，可能认为聪明的人能够取得好成绩，若某人表现不佳或进步较慢，则会被认为“不那么聪明”，只是“做得快”而非“做得好”。与此一致，近期研究发现，儿童认为一开始表现优异但后来逐渐退步的人其实并不聪明(Hu et al., under review)。相对而言，儿童逐渐意识到“努力”并不总能带来好结果。年长儿童与成人更倾向于将努力理解作为一种持续的精力投入，而非单纯由结果反映出来。随着他们在生活中逐渐体验到即使努力也未必成功的情境，他们对努力的判断也愈加过程化。近期的相关研究同样表明，4-7岁的儿童已经能够理解努力本身的过程价值，而非仅将其价值建立在结果之上(Hu et al., 2025)。”

最后，关于成人为何更多地呈现出过程导向，在结果变化时保持原有判断，我们提出以下解释（文稿中修改见第18页）：

“首先，这可能是因为成人的判断更关注过程，而非结果；这也与他们在道德判断中更加注重行为意图而非行为结果的倾向一致(Cushman, 2008)。其次，成人可能认为一次考试的名次变化不足以提供足够的信息来修正他们原有的判断，他们需要更持续或更显著的表现差异来调整之前的判断。与此相符，近期研究发现，与儿童不同，成人倾向于认为一开始表现出色但后来退步的个体比起一直保持中等表现的个体更聪明(Hu et al., under review)。最后，从更一般的认知机制来看，成人的判断可能更受“确认偏差”的影响，即既有信念一旦形成便较为稳固，不易被相反信息改变(Nickerson, 1998)。未来研究可通过引入多次考试结果的动态

变化，进一步探究成人在“聪明”与“努力”判断中的认知过程与调整机制。”

意见 2：情境与结论的适用范围：本研究仅考察了成功情境下（两名角色均取得较好成绩）的过程/结果导向区别。未涉及努力者可能失败的情境可能限制结论的普遍性。例如，如果勤奋用时长的角色最终没有取得好成绩，儿童对其努力的评价和偏好是否还会增加？这一问题尚未在本文中得到检验。作者在讨论中提到了这一点，但建议进一步澄清该局限，对未涉及的失败情境加以讨论；如是否已有相关研究探讨过消极结果情境下儿童对努力的看法，可在文献中补充说明。如果考虑文章工作量的话，最好能补充一下这个情境的研究数据，从而使本文研究结论更为稳健。

回应：我们非常感谢审稿专家对情境适用范围的建议。正如您所指出的，本研究中的第一次考试中两位角色均取得较好成绩，因此研究结论在一定程度上受限于较为成功的情境，未涉及努力者可能失败的情境（如，排名很低）。

受到审稿专家这一意见的启发，我们注意到，实际上在第二次考试中，由于我们的实验操控用时长的角色有时排名较高（第二名），有时排名较低（第三名），一定程度上考察了当努力者表现不如他人时，儿童及成人如何对努力者进行评价。我们在修改稿中补充了在第二次考试中儿童及成人评价的结果，结果发现，无论成人还是儿童，其评价都受到结果信息的显著影响。当在第二次考试中两位人物的表现出现差异时，儿童(尤其是年龄稍大的儿童)与成人的评价明显倾向于排名更高的一方。我们认为这一结果可能反映了当下社会中以结果为导向的评价体系对于儿童及成人的评价的影响。具体数据分析及结果见文稿 14-15 页。

另外，我们也结合近期的研究进行讨论（文稿中的修改见第 19 页）：近期的研究直接探究了儿童对于努力的过程价值（如，努力是为了享受努力的过程）和结果价值（如，努力是为了取得好的成绩）的人物的评价，发现儿童对于既享受了努力过程也取得了好成绩的个体评价非常高，对于享受了努力过程但并未取得好成绩的个体评价次之，对于既没有享受努力过程也没有取得好成绩的个体的评价则相对较低，这体现了儿童在对于努力者的评价中对于过程与结果的综合考虑。未来研究可以进一步呈现多种结果变化情境，系统探究儿童在不同成功与失败情境下对努力和聪明的评价与偏好。

意见 3：幼龄儿童测量的有效性：参与者中包含 4 岁幼儿，其认知和表达能力有限。研究采用 6 点刻度的程度评定（“有点儿/比较/非常（不）聪明”等）来测量儿童对特质的判断，量表选项可能过于复杂。4 岁儿童是否真正理解这些定量差异值得关注。尽管作者采用了图片或措辞辅助（文中未详述），但仍建议在方法中说明如何确保幼童理解评分类别，并在结果中对年幼组的数据可靠性给予谨慎解释。如果有幼儿难以理解评分任务，可能导致他们的评分结果偏差或随机，应考虑对此做必要说明。例如，可以报告不同年龄组注意力检测通过率或对 4 岁组数据进行单独分析以证明其模式可信。确保测量工具对低龄儿童友好，对于验证结论非常重要。

回应：感谢审稿专家的建议！我们在方法中补充了对于这一问题的问法（文稿中修改见第 6 页），以聪明评分问题为例，主试首先询问儿童该故事人物是“聪明”还是“不聪明”。在儿童作出二分判断后，主试进一步追问其具体程度（“有点儿”、“比较”或“非常”）。例如，当儿童回答人物“聪明”时，主试会追问该人物是“有点儿聪明”、“比较聪明”还是“非常聪明”；若回答人物“不聪明”，则进一步追问是“有点儿不聪明”、“比较不聪明”还是“非常不聪明”。努力评分问题采用相同的流程。为帮助低龄儿童更好地理解不同程度之间的差异，主试在提

问过程中结合手势示范，以双手间距大小直观呈现程度差异。这种分步骤的6点计分方法在以往的对于4岁儿童的研究中也已被广泛使用(Zang et al., 2023; Shtulman & Phillips, 2018)。

此外，本研究设置了注意力检测问题，4岁儿童的通过率为96%，证明本研究的实验材料对于低龄儿童较为友好。

意见 4：除了上述主要问题外，论文在表达和格式上还有一些细节可以改进：

- 概念术语解释：文中提到“聪明者改变情境”和“努力者改变情境”等术语。建议在摘要或方法部分对这些术语加以简单解释，以方便读者快速理解实验设计。例如，可注明“聪明者改变情境”指第二次测试中原本较不聪明者成绩反超的情境。清晰的术语解释有助于读者把握情境区别。

回应：我们在方法部分对“聪明者改变情境”和“努力者改变情境”等术语进行了详细的解释（文稿中修改见第6页）。修改之后的段落如下：“在第二次考试中，两位故事人物的考试过程(即用时长短)与第一次考试保持一致，但其考试结果(即光荣榜排名)由主试根据儿童在第一次考试中迫选问题的判断进行操控。在“聪明者改变”情境中，如果儿童在第一次考试的聪明迫选问题中认为短(或长)用时人物更聪明，则该人物在第二次考试中的排名会被另一人物反超。也即，儿童最初判断“更不聪明”的人物在第二次考试中进步至第二名，而其判断“更聪明”的人物则维持在第三名。在“努力者改变”情境中，操控逻辑相同：如果儿童在第一次考试的努力迫选问题中判断短(或长)用时人物更努力，则该人物在第二次考试中的排名会被另一人物反超。也即，儿童最初认为“更不努力”的人物在第二次考试中进步至第二名，而其认为“更努力”的人物则维持在第三名。”

我们在摘要部分也对实验的操控进行了简单的解释“在第二次考试中，我们操控原本被儿童认为不那么聪明或不那么努力的人物取得了更好的考试结果，从而探究儿童的判断是否因结果变化而改变。”但由于摘要字数的限制，我们并未直接介绍“聪明者改变情境”和“努力者改变情境”等术语。

意见 5：方法细节描述：建议补充说明成人被试的操作流程。虽然文中提到成人通过问卷星填写问卷，但应明确成人是否观看了与儿童相同的故事情境描述和问题，以及成人组是否也是先后完成两个情境的判断。保证成人条件与儿童一致，才能使成人作为基线组的比较具有说服力。如果存在差异（例如成人只读文字描述，没有视觉材料等），应在方法中加以说明。

回应：我们已在方法中说明成人组在问卷星上填写问卷，问卷流程去除了儿童实验中第一次考试前有关光荣榜的概念介绍环节（因为对于成人来说过于简单），其余设置与儿童组保持一致。成人组阅读了与儿童组相同的文字描述和视觉材料，并回答与儿童相同的问题。具体修改见文稿第7页。

意见 6：在讨论末段，作者提到了儿童对聪明和努力认知与成长型思维、成就归因的关系值得进一步研究。若篇幅允许，建议稍充实对此的讨论。例如，可以结合 Dweck 的内隐智力观理论指出本研究发现的儿童早期聪明/努力观念的发展，可能对其日后形成固定型或成长型思维模式有何启示。这将使文章的理论意义更为丰满，供作者酌情考虑。

回应：我们感谢审稿专家这一宝贵意见。我们已在讨论中进行了补充：“本研究的发现对于理解能力与努力之间的关系，以及成长型思维和成就归因等理论，也具有一定的启示意义。儿童在对能力与努力的判断和评价中，可能已初步体现出与其成长型思维模式和成就归因相

关的个体差异。成长型思维强调个体相信能力可以通过努力得到提升(Yeager & Dweck, 2020)。那些较早理解“努力”的内涵及其意义的儿童,或许更倾向于认为努力能够促进能力发展,即具有更高水平的成长型思维。同时,这些儿童可能也更容易将成就与失败归因于努力,而非天赋,从而在面对挫折时能够持续投入、不轻言放弃,展现出更具适应性的行为。未来研究可进一步探究儿童对能力与努力概念的理解与评价,与成长型思维、成就归因等观念之间的关系,并探究这些信念与理解如何共同影响儿童的坚持、毅力以及挑战寻求等动机与行为表现。”文稿中修改见 19 页最后一段。

第二轮

审稿人 2 意见:

作者对稿件进行了显著修改和提升,体现出认真回应审稿意见的努力。文章在论述逻辑、文献引用、方法描述等方面都有改进。然而,仍存在若干问题需要进一步解决,包括部分段落的逻辑衔接、方法和统计分析的说明,以及结果呈现的细节。建议作者关注这些方面,以进一步提升稿件的清晰度和科学性。

意见 1: 在 1.2 儿童对能力和努力的认知导向部分,作者使用了“目前的研究结果表明……”“现有的研究范式通常采用单一情境设定……”等表述。凡是涉及对既有研究的总结、评价或概括性论断之处,均应提供相应的参考文献或文献出处以支持论点,增强论述的科学性与可检索性。

回应: 非常感谢审稿人的指正。我们已仔细检查 1.2 节中关于既有研究的总结性表述,并根据建议补充了相应的参考文献。见文稿第 3 页。

意见 2: 稿件在研究设计与结果讨论部分明确涉及成人样本,并以成人作为重要的比较或参考对象。然而,题目和摘要均未提及成人,仅呈现“儿童”。若成人在研究中承担关键角色(如核心比较对象、理论参照或重要的解释依据),则应在题目和摘要中明确呈现。

回应: 根据建议,我们已对题目和摘要进行了相应调整,增加了对“成人”样本的明确呈现。见文稿第 1 页。

意见 3: 作者在题目里面用的是“评价”,但是在前言的 1.3 部分,用的是“偏好”。建议统一。

回应: 我们已将前言 1.3 部分中的“偏好”改为“评价”。

意见 4: 稿件中 1.1 “儿童对能力与努力的概念理解”与 1.2 “儿童对能力和努力的认知导向”均涉及儿童对能力与努力的认知判断。当前表述容易让人混淆,难以区分两部分的具体侧重点。其实 1.2 部分讨论儿童在评价他人时是否会受到结果信息的影响,即他们是否会根据最终结果调整对能力或努力的判断(结果依赖性/反向推断)。建议标题突出一下,和 1.1 区分开。在总讨论部分,也存在类似的问题。

回应: 感谢审稿专家的建议。我们已将 1.2 的标题改成“儿童对能力和努力推断的过程或结果导向”,与 1.1 明确区分,在总讨论部分,我们也将 4.2 部分的标题改为了“对能力和努力

推断的过程或结果导向”。

意见 5: 在 1.2 “儿童对能力和努力的认知导向”部分，第一段提出关注结果在儿童认知判断中的作用，但未提供参考文献。建议作者明确该观点的理论依据和已有研究发现：例如，结果信息在儿童的认知推断中通常起何种作用，已有研究发现儿童会根据结果调整判断，因此将这一机制引入对能力和努力的评价中是合理的。

回应: 我们已补充了有关关注结果在儿童认知判断中的作用的文献，并更加详细地阐述了结果信息在儿童的认知判断中的作用。目前该部分已修改为：“我们推测，由于以往研究表明年幼儿童在进行社会与道德评价时往往关注行为的结果而非过程或意图(Cushman et al., 2013)，因此他们在判断“聪明”和“努力”等概念时，也可能更倾向于依赖与结果相关的线索。具体而言，他们可能认为更聪明或更努力的个体理应取得更好的结果；如果原本较聪明或较努力的个体在结果上被另一人物超过，年幼儿童对“谁更聪明”或“谁更努力”的判断可能会随之改变。”见文稿第 4 页。

意见 6: 在 1.3 “儿童及成人对能力者与努力者的偏好”部分，目前的论述，还是有些不太通畅。第一段主要介绍成人研究，第二段讨论儿童随年龄变化的偏好。第二段描述儿童从“天赋者偏好”向“努力者偏好”的年龄发展趋势，但随后直接引入成人偏好，逻辑显得略为突兀。为了使叙述更顺畅，建议作者考虑通过扩大儿童样本的年龄范围（比如从儿童到青少年），完整呈现儿童偏好的年龄发展轨迹，而不是直接引入成人对比。

回应: 我们增加了有关于从儿童到青少年部分的介绍：“此外，已有研究表明，小学低年级儿童(7 至 10 岁)对努力者的偏好程度比成年人更高(Hu et al., under review; Zhao & Yang, 2023)。这启示我们，随着年龄增长，进入青少年期后，儿童可能逐渐意识到努力所伴随的成本(Jiang et al., 2018)，从而对于努力者的偏好可能有所下降。”见文稿第 4 页。

意见 7: 稿件在被试部分，使用 G*Power 3.1 进行了 power 分析，报告儿童和成人样本在当前样本量下对小到中等效应量 ($f^2 = 0.15$) 的线性多元回归检验效能分别为 87% 和 90%。然而，这种属于事后分析，不是事前。事前根据预期效应量、显著性水平 α 和预测变量数量，计算所需样本量，以确保研究在设计阶段就具备足够的统计效能。放在被试的部分，应该是事前分析。

回应: 感谢审稿人的提醒。我们完全认同事后 power 分析不应出现在“被试”部分。根据建议，我们已将该分析移至讨论部分，并明确说明是事后分析。见文稿第 22 页。

意见 8: 稿件中描述儿童对人物“聪明”与“努力”的评分流程时提到，每个问题有三个等级（如“有点儿聪明”“比较聪明”“非常聪明”），但随后又说明最终评分为 1~6 分。这一点容易引起混淆：如果每个维度本身只有三级评分，为什么最后呈现的是 1~6？建议作者澄清评分方式。

回应: 感谢审稿人的细致指正。我们已根据建议对评分流程进行了明确澄清。具体而言，儿童在评价角色的“聪明”或“努力”时，首先需要判断该角色是“聪明/努力”还是“不聪明/不努力”。若儿童判断为“聪明（或努力）”，则进一步在“有点儿聪明 / 比较聪明 / 非常聪明”三个等级中选择；若判断为“不聪明（或不努力）”，则同样在“有点儿不聪明 / 比较不聪明 / 非常不聪明”三个等级中细分。因此，两个方向（正向与反向）各包含三个等级，最终形成

1-6 分的评分体系 (1 =“非常不聪明/不努力”, 6 =“非常聪明/努力”)。我们已在方法部分对这一流程予以清晰说明, 以避免读者产生困惑。见文稿第 6 页。

意见 9: 稿件在 3.1 部分报告了年龄 × 故事人物 × 问题的三维交互作用显著。建议作者在三维交互显著的情况下, 进一步报告相关的二维交互作用, 并在此基础上分析简单效应, 以更全面地揭示各因素之间的交互模式。

回应: 我们补充了二维交互作用的简单效应: “我们接着解释二维交互作用。为了分析故事人物与问题类型之间交互作用, 分问题类型分析故事人物的简单效应, 在努力评分问题上, 儿童对长用时人物的评分($M = 5.15, SE = 1.10$)显著高于短用时人物($M = 4.53, SE = 1.51$), $B = 0.62, SE = 0.10, 95\% CI = [0.43, 0.81]$, Holm 校正后 $p < 0.001$; 但在聪明评分问题上, 儿童对两类人物的评分差异不显著(Holm 校正后 $p = 0.637$)。为了分析故事人物与年龄的交互作用, 分故事人物分析年龄的简单效应, 发现儿童对长用时人物的评分随年龄呈现正向变化趋势但并不显著($B = 0.14, p = 0.101$), 对于短用时人物的评分随年龄呈现负向变化趋势但并不显著($B = -0.06, p = 0.382$), 两类人物的年龄效应差异显著(Holm 校正后 $p = 0.001$)。为了分析问题类型与年龄的交互作用, 分问题类型分析年龄的效应。结果显示, 在控制了故事人物的效应之后, 儿童的努力评分($B = 0.07, p = 0.341$)与聪明评分均未随年龄显著变化($B = 0.01, p = 0.890$), 且两类问题的年龄效应差异亦不显著(Holm 校正后 $p = 0.359$)。”见文稿第 9 页。

意见 10: 在讨论部分, 作者提到“有趣的是, 成人在“聪明”与“努力”的判断上均展现出过程导向”。学术写作建议避免使用“有趣的是”这类主观性较强的词语。在学术写作中, 可以改为客观描述发现或结果的表达, 如“结果显示……”或“本研究发现……”, 以提升论文的学术规范性和客观性。

回应: 感谢审稿人的细致建议。我们已将讨论部分相关表述进行了修改, 将原文中“有趣的是”替换为“本研究发现”。

审稿人 3 意见:

本研究聚焦于儿童对“聪明”和“努力”概念的理解及评价, 探讨儿童在判断他人聪明或努力时是更倾向于关注任务过程信息 (如完成任务所花时间) 还是结果信息 (如考试排名)。作者采用了一个创新的双情境两次测试范式: 在第一次测试中, 两位故事人物成绩相同但完成用时不同; 在第二次测试中, 通过操控成绩排名使原先稍逊的一方后来居上。研究主要发现包括: 4 岁左右的儿童已经能够根据用时长短判断谁更聪明 (体现一定的结果导向), 而对谁更努力的判断则随年龄增长逐步从结果导向转向过程导向; 在评价偏好上, 年长儿童更倾向于喜欢并奖励“努力者”。成人组则始终更多地依赖过程线索判断聪明和努力, 对结果变化不敏感, 对勤奋者评价更高。总体而言, 研究结果清晰地揭示了儿童对于聪明与努力概念的认知随年龄发展的不同轨迹, 为成就归因领域提供了新的证据。

有以下细节建议进一步修改完善:

意见 1: 引言: 引言部分背景梳理充分, 文献综述全面, 新旧研究衔接顺畅地引出了研究问题和假设。但仍有细节可改进之处:

1. 突出研究意义：建议在引言结尾处进一步强调本研究的问题如何凸显创新之处。这将在一开始就让读者清晰理解研究的新颖性。

回应：感谢审稿人的宝贵建议。我们已在引言结尾部分进一步强调了本研究的创新点，具体说明了研究相对于既有文献的独特贡献：“本研究的主要创新在于系统考察儿童及成人在判断“能力”与“努力”时，是更倾向于结果导向（主要基于结果线索）还是过程导向（主要基于过程线索），并进一步探究他们对聪明者与努力者的评价是否受到结果变化的影响。本研究不仅丰富了儿童能力与努力认知的发展研究，也为理解不同年龄群体在社会推理中如何整合过程与结果提供了新的视角。”见文稿第 5 页。

意见 2：成年人研究的综述：作者在引言第 1.3 节引用了成人偏好天赋者或努力者的研究成果（如 Tsay, 2011; Zhao & Yang, 2023 等）。该部分可以适当扩充。尤其是如果成人组的数据在文章中占一定比重，那么在引言中应更明确地提出关于成人的研究问题或假设。目前引言主要聚焦于儿童，对成人相关理论着墨略少（稿件第 4 页 1.3 节）。建议增加对成人为什么可能表现出与儿童不同的偏好或认知导向的推测，以便读者了解引入成人组的理论动机。

回应：感谢审稿人的中肯建议。根据建议，我们在 1.3 节中扩充了有关成人研究的综述和文献：“此外，即便努力并未带来更好的结果，成人仍倾向于认为努力者更加道德，更值得被奖励(Gelniker et al., 2023)。研究还发现，环境因素如经济不平等和社会竞争程度也会影响人们对于天赋和努力的评价：经济不平等程度越大、竞争越激烈，人们越倾向于偏好天赋者(Bak & Yi, 2025)。”（见文稿第 4 页）。

此外，我们进一步明确提出了本研究中针对成人组的研究问题与预测：“此外，我们还将收集成人样本的数据，探究成人对聪明和努力的判断更偏向结果导向还是过程导向，以及他们对聪明者和努力者的评价。基于以往研究显示，当成人对普通人进行评价时，通常展现出对努力者的偏好(Yang et al., 2024)，我们推测在本研究中成人也可能会展现出对努力者相较于天赋者更积极的评价。对于成人对聪明和努力的判断是否会受到结果变化的影响，我们并未提出明确假设。一种可能性是，成人的模式与上文所述的年长的儿童相似：在判断聪明时更依赖结果信息，而在判断努力时更依赖过程信息；另一种可能性是，成人无论是在判断聪明还是努力时，都更加关注过程信息，而较少受结果影响。成人样本将作为儿童数据的重要发展性对照组，为儿童结果的解释提供一定的基准。”（见文稿第 5 页）。

意见 3：结果部分分为对第一次测试结果、认知导向分析、评价偏好等几个板块，层次清晰。统计报告详尽，但读者在众多数字中需要抓住要点。以下是针对结果部分的具体建议：
1. 强调关键发现：建议在每节结果的开头或结尾，加一句话总结该节最重要的发现。例如，在“3.1 对第一次考试中聪明和努力的推断”末尾点明：“综上，4 岁儿童已经能够区分更聪明者，但难以辨别更努力者；随着年龄增长，儿童对‘谁更努力’的判断逐渐倾向于用时长者。”这样的总结可以帮助读者迅速提炼信息。当前结果段落虽有阐述，但总结句不够突出，可适当强化（稿件第 9 页第 1 段已部分涵盖此意，建议措辞上再清晰直接一些）。

回应：非常感谢审稿人的建议。我们已在结果的每节最后加入了对于该段结果的总结（见文稿第 12-17 页）。

意见 4：成人组结果：虽然研究重点在儿童，但包含了 100 名成人的数据作为比较。建议

在结果部分适当提及成人组的结果要点。例如，可在每一小节最后一句概括成人的模式：“相比之下，成人在上述判断中几乎均未受结果变化影响，一直坚持认为短用时者更聪明/长用时者更努力”。目前成人的数据主要在讨论中描述，结果部分缺少对成人组的直接报告。这可能让读者误以为作者忽略了成人数据。精炼地报告成人组关键结果将使全文更完整。若篇幅不允，也可在正文提及“成人组结果详见附录 2”，但个人建议在主文中简短说明，以体现成人数据的价值。

回应：我们已在结果部分每节最后加入了对于成人组结果的总结（见文稿第 12-17 页）。

意见 5：讨论部分分主题深入剖析了研究发现的意义，与已有文献进行了充分对比，并提出了教育启示和未来展望。在完善讨论时，可考虑以下建议：

1. 加强结果与文献的呼应：讨论开篇已围绕三个研究问题重述主要发现。之后 4.1、4.2、4.3 小节分别探讨了能力与努力推断的起点差异、认知导向的发展分化以及偏好的年龄变化。在这些段落中，作者引用了大量相关文献解释结果。例如，用父母称赞和儿童理解“认知成本”等因素解释了努力概念出现较晚（稿件第 11 页第 2 段），用道德判断中结果导向的普遍现象解释幼儿更关注结果（稿件第 12 页第 1 段），这些解释合理且基于文献。不过，建议进一步强调本研究如何拓展或修正已有理论，使读者更清楚认识本研究成果在文献中的地位。

回应：我们已在讨论中进一步强调了本研究如何拓展或修正已有的理论：“这些结果共同对 20 世纪 70 年代提出的“学龄前和学龄早期儿童尚无法区分聪明与努力概念”的传统观点与理论框架(Nicholls, 1978)提出了新的挑战。我们的研究不仅表明，儿童在 5 岁半左右便能够在一定程度上区分两者，还进一步揭示了一个此前文献中未被充分讨论的发展差异：儿童对“努力”的推断出现得略晚于对“聪明”的推断。这一发现启示我们两者并非同步，而可能具有不同的发展轨迹与机制。”（见文稿 18 页）

“这些发现拓展了有关能力和努力认知发展的理论(Nicholls, 1978; Cimpian, 2017)，进一步揭示了儿童如何综合过程和结果的信息进行判断。本研究创新性地指出，儿童在判断“聪明”（能力）和“努力”时展现出不同的认知导向与发展轨迹：在幼儿期和学龄初期，儿童对“聪明”的推断主要依赖结果信息，而对“努力”的推断则随着年龄增长逐渐转向依赖过程线索，且不再受结果变化的影响。这些发现进一步印证了儿童对能力和努力的概念发展的差异性。”（见文稿第 20 页）

意见 6：局限性补充：作者在“4.4 局限与未来方向”中已诚恳地指出了样本年龄范围、任务情境（未涉及失败情境）等局限，并提出未来可在更广年龄段、不同学科领域、失败情境下继续研究。这些展望非常全面。除此之外，可考虑补充一项局限：文化背景的局限。本研究被试均来自中国文化环境，而作者在引言也讨论了中西方文化对努力价值的不同强调。有必要在局限中提醒读者，本研究结论主要适用于中国文化背景，其他文化下儿童的聪明/努力认知发展可能有所不同。虽然 Yang 等(2023)的跨文化研究表明 7 岁后中美儿童都偏好努力者，但关于更低龄儿童过程/结果导向的发展，尚缺乏跨文化比较。本研究为后续此类研究提供了范式，可在未来进一步检验文化因素的影响。

回应：我们已在局限部分加入了有关于文化背景的局限：“本研究的样本仅包括中国儿童与成人，因此研究结论主要适用于中国文化背景。在其他文化情境中，儿童对聪明与努力的认知发展可能呈现不同的模式。尽管 Yang 等(2024)的跨文化研究表明，中美儿童在 7 岁之后都更偏好努力者，但关于更低龄儿童在推断聪明和努力时是呈现过程导向还是结果导向，目

前尚缺乏系统的跨文化比较。本研究为后续此类研究提供了可行范式，可在未来进一步检验文化因素的影响。”（见文稿第 22 页）

意见 7：成人结果的讨论：作者在讨论中后段对成人与儿童差异进行了分析，提到成人可能因为更关注过程、一次名次变化不足以改变判断，以及确认偏见等原因，在两次测试中保持了稳定判断（稿件第 12 页第 2 段）。这些推测有文献支撑且富有见解。建议作者再明确一点：成人相比儿童在本研究中的对照意义。例如：“成人组的稳定表现为儿童的发展提供了基准线，说明结果导向倾向主要存在于儿童早期，随认知成熟而减弱消失。”这样点明成人数据的重要性。最后关于成人评价倾向的讨论（稿件第 13 页第 1 段），作者指出成人在奖励维度更偏好努力者，但在偏好和成功预测上更平衡。这一点可以和儿童的年龄趋势联系起来谈：似乎儿童在小学中高年级阶段可能也会逐渐走向这种平衡观，正如文中提到随着经验增长对努力成本的认识（Jiang et al., 2018）的作用。因此，可猜想儿童对努力偏好的上升曲线可能在青春期出现拐点。这些推论若能在讨论中稍作阐述，将使文章的发展脉络更完整。

回应：已根据建议对成人组的讨论进行了丰富：“小学阶段的儿童往往高度认同努力的价值，并倾向于给予努力者更高的评价；然而，随着年龄和经验的生长，他们逐渐意识到努力并非总能带来更好的结果，并开始感知到努力所伴随的成本(Jiang et al., 2018)。因此，他们在小学高年级或青少年阶段可能形成对天赋与努力更加平衡的看法。未来研究可进一步扩大年龄范围，探究儿童在中学阶段对于能力者与努力者的评价模式，我们猜想努力偏好的上升曲线可能在青春期出现拐点。”。具体见文稿第 21 页。

第三轮

审稿人 2 意见：

意见 1：核心概念“过程导向/结果导向”的建构效度仍不足：目前更像“对结果变化的敏感性”，未必等同“结果导向”

作者将“第二次考试改变迫选判断”定义为结果导向（2.2 节；3.2 节）。在统计学习/社会推断的标准框架下，如果第二次考试提供了新的、被试认为更诊断性的证据（排名变化），那么改变判断可能是合理的证据整合，而非“过度结果导向”。比如一次考试做的快不一定是聪明，而第二次考试还能做的那么快，真的是更聪明。关键在于：稿件没有充分证明参与者（尤其儿童）把第二次排名变化理解为“仅仅是结果变化而过程未变”，而不是“能力/努力真实变化”或“运气/任务难度变化”。

目前的范式叙事中，第二次考试被明确呈现为“这次考试他考了第二名/第三名”（附录指导语示例），儿童很可能把它解释为“这一次他真的更厉害/更努力”，从而更新特质判断。这种解释路径会削弱作者将其命名为“结果导向”的理论含义。

建议在引言/讨论中更清晰地区分并限定术语：建议把当前指标更准确地表述为“结果敏感性（outcome sensitivity）/对结果线索赋权更高”，或者解释“导向”即为此意思。

b) 增补（或在既有数据上补做）能够支持建构效度的证据：例如在任务后加入简短的原因解释（开放或选项式：你为什么觉得他更聪明/更努力），或至少报告附录中是否已有类似质性回答。若无法补测，应在讨论中把“结果导向”表述为推测，并系统呈现替代解释。

c) 在结果解释中避免将“改变判断”单向解释为“依赖结果而忽视过程”，而应写成“当过程线

索与结果线索冲突时，被试更倾向跟随结果线索（或更新判断）”。

回应：感谢审稿专家提出的宝贵建议。我们认同原文中使用的“结果导向”与“过程导向”等表述可能不够精确；根据审稿意见，我们已作出如下修改：

首先，我们在全文中将“结果导向”“过程导向”等术语进行修改调整，改用更贴合本研究设计所能反映含义的表述，如“结果敏感性”“因结果变化而改变判断”“受到结果信息的影响”等。此外，我们也修改了原本标题中的“结果导向”或“过程导向”的说法。

第二，由于本研究未设置开放式回答，因而我们在讨论部分补充了对多种替代性解释的分析。例如，部分儿童可能并未将情境理解为“结果变化而过程不变”，而是推断人物的能力和努力在第二次考试中也发生了真实改变；部分儿童可能受到需求特征的影响而倾向于在第二此判断中修改原有答案等。与此同时，我们也指出，即便在这种可能性存在的前提下，儿童在“聪明”和“努力”判断上修改答案的比例仍呈现出显著差异，表明儿童对于“聪明”和“努力”的判断的结果敏感性确实存在有意义的区别。具体见文稿第 20-21 页。

第三，我们对结果解释的语言进行了修改，避免将“改变判断”解释为“依赖结果而忽视过程”，采用更加准确、贴合数据的说法，如“对结果信息更敏感”，“倾向于根据结果的变化而调整其判断”。

意见 2：“动态调整（依据第一次回答来设定第二次结果）”的设计有强优点，但也引入了可观的需求特征与可比性问题，需要更透明报告。第二次考试的排名是根据被试第一次迫选判断进行操控，使其“先前认为更聪明/更努力者在第二次落后”（2.2 节）。这确保了每个被试都遭遇“反证信息”，但也带来两类风险：

1) 需求特征/叙事暗示：被试可能推断“故事的重点就是让排名反转，因此我应该改变答案”，尤其在远程、由主试主导叙事时更可能发生。

2) 刺激不等价：由于操控依赖第一次回答，不同年龄段/不同被试在第二次看到的“用时×排名”组合比例可能不同（尤其在努力维度上，低龄儿童第一次判断可能更随机），从而使所谓“发展变化”部分来自不同刺激暴露结构而非纯粹认知变化。

建议：

a) 在方法/附录中报告关键分布：在两种情境下，第二次考试中“短用时者排名更高/更低”的比例分别是多少？这些比例是否随年龄变化？（如果随年龄变化，就必须作为潜在混淆源讨论或在模型中控制。）

b) 在讨论中正面处理需求特征可能性，并说明你们如何降低（例如：是否有遮蔽目的的措辞；是否有与“改变答案”无关的填充问题；是否有主试标准化脚本训练等）。

c) 提供一个敏感性分析：例如仅在第一次判断“符合多数成人/规范推断”的子样本中（聪明：短用时更聪明；努力：长用时更努力）重复 3.2 分析，看“导向发展趋势”是否仍成立。这样可以降低“低龄随机性 + 动态操控”导致的偏差。

回应：感谢审稿人的这一建议。我们已根据审稿意见做出如下修改：

首先，我们在结果部分补充报告了在两种情境下第二次考试中“短用时者排名更高/更低”的

比例。在聪明者改变情境中，71.7%的儿童看到第二次考试中长用时人物排名更高，28.3%的儿童看到第二次考试短用时人物排名更高；在努力者改变情境中，64.2%的儿童看到第二次考试中短用时人物排名更高，35.8%的儿童看到长用时人物排名更高。由于该比例随年龄有所变化，我们进一步进行了补充分析，将第二次考试中的相对排名（即结果信息）作为控制变量纳入模型，结果显示，原有的特质的主效应和特质与年龄的交互作用依然显著。核心结论具有稳健性。

其次，我们增加了敏感性分析，对于在第一次判断中“符合多数成人/规范推断”的子样本中（即在聪明迫选问题中正确回答短用时人物，在努力问题中正确回答长用时人物的试次），我们重复了3.2节的分析，结果发现特质的主效应依然显著，即儿童对于聪明和努力的判断的敏感性仍存在显著差异；特质与年龄的交互作用仍存在原有的趋势，但未达到显著水平。这可能与子样本缩小导致统计效力下降有关。我们已在结果中汇报了这些补充分析，并在讨论中更加聚焦于较为稳定的效应（即儿童在“聪明”和“努力”判断上的结果敏感性差异），相应弱化了关于发展趋势差异的结论。具体见文稿第13页。

最后，我们在讨论部分补充了对于需求特征可能性的分析与讨论。在本实验的指导语与程序设计中，我们采取了多项策略以降低需求特征带来的潜在风险（即，儿童可能猜测研究的重点在于第二次考试中排名的变化，从而在第二次判断中倾向于改变答案）。首先，第二次考试被自然地嵌入故事流程中（“很快他们进行了第二次考试”），未突出强调其与第一次结果的矛盾性或“反证性质”。其次，在第一次考试与第二次考试的关键判断问题之间，儿童还完成了若干评价问题与注意力检测问题，以分散被试对回答一致性的关注。再次，实验过程均采用严格标准化脚本，并且未向被试提供任何正确性反馈。上述措施共同减少了被试基于情境线索形成策略性应答的可能。此外，我们也承认，部分儿童在第二次判断中的回答与第一次判断不同，可能一定程度上受到需求特征的影响。然而，即便在这种可能性存在的前提下，儿童在“聪明”和“努力”判断上修改答案的比例仍呈现出显著差异。这一模式表明，儿童对于“聪明”和“努力”的判断的结果敏感性确实存在有意义的区别。具体见文稿第20-21页。

.....

审稿人3意见：

总体而言，作者已根据前期意见对稿件进行了修改。然而，文章在发展框架的整体清晰性方面，建议稍微再改进下。

意见1：尽管作者已在题目中明确纳入成人样本，但在前言与讨论部分，对儿童与成人的整体梳理仍显不够清晰。目前，引言中仅在1.3和1.4部分较为集中地讨论了成人，而在前言开头及1.1、1.2部分，成人相关内容相对缺失。建议作者在引言的各个部分中，均适度纳入成人研究的视角，以确保儿童与成人的比较贯穿全文，而非局限于个别章节。同样地，在讨论部分，建议作者进一步强化以“儿童 vs. 成人”为核心的分析框架，系统比较两类群体在努力者偏好上的异同及其潜在机制，从而提升全文在理论结构上的一致性与聚焦性。

回应：非常感谢审稿人的指正。我们已在引言的各个部分纳入了成人研究视角，并在讨论部分系统讨论了儿童与成人的比较，尤其探讨儿童和成人在努力者偏好上的异同及其潜在机制。具体见文稿引言和讨论标红部分。

意见 2：“此外，已有研究表明，小学低年级儿童(7 至 10 岁)对努力者的偏好程度比成年人更高(Hu et al., under review; Zhao & Yang, 2023)。这启示我们，随着年龄增长，进入青少年期后，儿童可能逐渐意识到努力所伴随的成本(Jiang et al., 2018)，从而对于努力者的偏好可能有所下降。”——这个论述，逻辑不清楚，而且跳跃推导。已有研究没有青少年，但是其实青少年？”——关于努力者偏好的发展性推论，目前的论述在逻辑上尚不够清晰。已有研究仅比较了小学低年级儿童（7–10 岁）与成年人之间的差异，并未纳入青少年样本。在此基础上直接引入青少年期作为关键发展阶段，缺乏直接的实证依据。此外，Jiang et al. (2018) 主要表明个体随年龄增长对努力所伴随成本的理解逐渐增强，但该发现本身并不能支持“成本意识增强必然导致对努力者偏好下降”的推断。建议作者在相关表述中区分已有证据与推测性解释，弱化因果语气，或将该推论明确为有待进一步检验的假设。

回应：我们已根据审稿专家的建议对该段进行修改，注意区分已有证据与推测性解释，该段落修改之后如下：“此外，已有研究表明，小学低年级儿童(7 至 10 岁)对努力者的偏好程度比成年人更高(Hu et al., 2026; Zhao & Yang, 2023)。虽然尚未有研究明确探究青少年期对天赋者和努力者的偏好，但有研究表明，青少年对努力所伴随的成本非常敏感(Jiang et al., 2018)，因此我们推测，进入青少年期，儿童对于努力者的偏好可能逐步下降。”

第四轮

审稿人 2 意见：同意发表

审稿人 3 意见：

意见 1：预注册与当前稿件不一致，且稿件对这一点不够透明。

文稿在方法部分写道“所有分析均按照预注册方案进行”，且在自检报告中写明计划样本量与实际样本量一致；正文报告的儿童样本是 4.02–7.95 岁、N=120，成人 N=100。但作者提供并在文中引用的 AsPredicted 预注册链接(#94495)显示，原计划儿童样本为 4–6 岁、N=90，成人 N=100；同时，预注册对评价问题(RQ4)的分析明确写的是聚焦第一次考试的评价回答。但作者所链接的预注册页面写得很清楚：第二次考试后实际上还会再次收集 smartness/hardworking rating、forced-choice 和三项 evaluation questions。

这会进一步引出一个透明度问题：如果第二次考试后的 rating 也收集了，为什么主文完全没有报告？如果没有收集，那就与预注册链接不一致；如果收集了但未报告，那就需要说明是出于什么原因未分析。当前稿件的方法—结果—预注册链接三者没有对齐。请作者把第二次考试中的所有测量项目完整写清楚，并说明哪些进入主文、哪些进入附录、哪些没有分析以及原因。

回应：非常感谢审稿专家的意见。我们作如下澄清：

首先，关于样本量的问题。在最初的预注册中，我们计划招募儿童样本为 4–6 岁、N = 90，成人 N = 100。在数据收集过程中，我们发现 6 岁儿童的回答仍处于发展阶段，尚未呈现稳定的模式。基于这一观察，我们决定扩展年龄范围至 7 岁儿童，并进行了第二次预注册 (<https://aspredicted.org/et2rt7.pdf>)。因此，最终样本量为两次预注册合并后的儿童 N = 120 (4–7 岁)，这一点与稿件中报告的样本量和年龄范围一致。我们已在方法部分补充了相关

说明，以提高清晰度与透明度。

第二，关于第二次考试之后的评价问题。我们在此澄清：与预注册一致，我们确实收集了第二次考试后的评价数据。在预注册的分析计划中，我们原本仅计划对第一次考试后的评价结果进行分析。然而，在审稿过程中，审稿专家建议将第二次考试后的评价分析也纳入正文。我们认为这一建议非常有价值，因此在修订稿中已将该部分分析纳入正文。我们已在正文的方法与结果部分进行了明确说明。

编委意见：

论文依据审稿专家的意见做了较好的修改，但在语言表述上需要进一步修改。

意见 1：建议论文题目修改为：过程与结果线索信息对 4~7 岁儿童及成人对“聪明”和“努力”判断与评价的影响，全文依此题目予以修改。

回应：感谢您的建议。我们已采纳该修改意见，将论文题目调整为：“过程与结果线索信息对 4~7 岁儿童及成人对‘聪明’和‘努力’判断与评价的影响”。同时，我们已根据新的题目对全文相关表述进行了系统性修改，以确保题目与正文内容的一致性。

意见 2：摘要开头需要增加一两句研究背景的陈述。

回应：感谢您的建议。我们已在摘要开头补充了研究背景的相关表述，以增强研究问题的引入和整体逻辑的完整性。具体来说，我们增加了“个体对“聪明”和“努力”的判断与评价对其学业投入、动机及表现均可能产生重要影响，但既有研究尚不清楚 4~7 岁儿童及成人对“聪明”和“努力”的判断与评价如何受到过程与结果线索信息的影响。本文探究这一问题。”

意见 3：论文中许多地方对“聪明”和“努力”判断与评价表述不统一，不够严谨，有时候表述为对“聪明”和“努力”概念判断与评价，有时候表述为对“聪明”和“努力”理解与评价，也有时候表述为对“聪明”和“努力”概念的理解与判断的影响，需要确定题目后在全文中予以统一。

回应：感谢您指出这一重要问题。我们已根据修改后的论文题目，对全文中涉及“聪明”和“努力”的相关表述进行了全面梳理与统一。具体而言，我们统一采用“对‘聪明’和‘努力’的判断与评价”这一表述，并对个别不够严谨或前后不一致的表述进行了修订。

意见 4：论文问题提出 1.4 中最后段有关创新性的表述建议删除，研究问题的创新性要通过前述内容表述体现出来，而不是直接告诉读者。

回应：我们已删除第 1.4 节最后一段中关于研究创新性的直接表述。

意见 5：引言与讨论部分陈述的逻辑层次需要加强，包括核心概念的前后统一使用。

回应：我们已对引言与讨论部分进行了系统梳理和修改，重点加强了段落之间的逻辑衔接与层次结构。同时，对核心概念（如“聪明”和“努力”的判断与评价）在全文中的使用进行了统一和规范。

意见 6: 摘要中的首句有关研究背景信息建议修改为：儿童对“聪明”和“努力”的判断与评价是对过程线索还是结果线索更加敏感？儿童与成人对“聪明”和“努力”判断与评价有什么不同，本研究对这一问题进行了探讨。

回应：我们已将摘要首句改为：儿童对“聪明”和“努力”的判断与评价是对过程线索还是结果线索更加敏感？儿童与成人对“聪明”和“努力”判断与评价有什么不同？本研究对这些问题进行了探讨。

意见 7: 前次提出的“引言与讨论部分陈述的逻辑层次需要加强，包括核心概念的前后统一使用”的修改意见，还需作者依据该条建议进一步仔细推敲修改论文。不是简单修改核心概念的表述，做到核心概念表述前后一致即可，而是在核心概念前后表述一致的前提下，对论文的内在逻辑进行修改，使得论文对于研究的核心问题表述更加清晰，结果的讨论也要紧紧围绕核心问题的回答展开。

回应：根据您的建议，我们进行了如下修改。在引言部分，我们首先在开篇对核心概念进行了界定，并清晰提出本研究的核心问题；随后围绕这些核心问题，对相关文献进行了整合与评述，在此基础上提出相应研究假设。我们对整体论述逻辑进行了系统梳理与调整，使本研究所关注的核心问题得以更加清晰地呈现和论证，同时删减了与研究主线关联较弱的冗余内容。

在讨论部分，我们首先概括本研究的核心发现，随后围绕各核心研究问题，结合既有研究结果与相关理论进行了针对性讨论，并进一步分析了本研究的局限性及其理论与实践意义。

具体修改内容见文中标红部分。

主编意见：

与前几轮版本相比，论文在研究问题聚焦、儿童与成人比较框架、预注册透明性、方法细节说明、第二次考试评价数据补充以及讨论结构等方面均有明显提升。稿件围绕 4~7 岁儿童及成人对“聪明”和“努力”的判断与评价展开，采用两次考试、动态调整结果信息的实验范式，较系统地考察了过程线索与结果线索在不同年龄阶段中的作用，具有一定理论价值和教育启发意义。

几点建议：

1 请继续收束对核心结论的表述强度。

本研究更直接支持的是儿童和成人在判断聪明”与“努力”时对结果变化的敏感性差异，而不是对稳定“过程导向/结果导向”认知模式的直接测量。请统一检查摘要、讨论、结论和自检报告，尽量使用“对结果变化更敏感/较少敏感”、“更可能随结果变化调整判断”等表述，避免再出现过强的“结果导向认知模式”、“过程导向理解”等说法。

回应：已对自检报告、以及全文进行细致检查，避免出现“结果导向认知模式”、“过程导向理解”等说法。

2 请进一步统一样本量与预注册的说明口径。

正文目前一方面说明样本量是预注册计划确定的，另一方面又在局限部分写“未采用事前效能分析”。建议明确区分：样本量是基于既有研究和预注册计划预先设定，而文中报告的 power 属于事后补充说明，不是正式的样本量确定依据。正文、自检报告和附录中的相关表述请统一。

回应：我们在正文的方法、讨论部分以及自检报告中进行了澄清，说明本文的样本量是基于既有研究和预注册计划预先设定，并未进行事前效能分析（见文稿第5页被试部分）；在数据收集结束之后进行了事后效能分析（见文稿第22页）。

3 请对“努力判断的发展趋势”保持更审慎的表述。

当前数据较稳妥支持的结论，是儿童在“聪明”和“努力”判断上存在不同的结果敏感性模式；至于“努力判断的年龄变化趋势”，在敏感性分析中稳健性相对弱一些。请在摘要、结论和讨论中相应收束，避免写得过于确定。

回应：我们已对摘要、结论、讨论进行修改，避免强调努力的结果敏感性的年龄变化趋势，而更多强调儿童在“聪明”和“努力”判断上存在不同的结果敏感性模式。例如，摘要相应部分修改为：“相较于“聪明”判断，儿童对“努力”的判断较少受结果变化影响”。在讨论和结论部分也进行了相应的修改（见文稿第20-21页）。

4 请收束第二次考试评价结果的外推。

目前结果可以说明，在成就情境中，儿童与成人对人物的评价会结合结果信息进行调整；但尚不宜直接上升为对现实社会“结果导向评价体系”的较强判断。请使相关讨论与证据层级保持一致。

回应：已对相应段落进行了修改，使讨论更加符合证据：“这一结果表明，儿童与成人在评价中均表现出对结果的敏感性，这一发现也与既有研究中儿童在评价他人时重视结果的结论一致(Yang et al., 2024)。未来研究可进一步探索儿童如何整合过程与结果信息，对他人形成各位全面的评价。”见文稿第21页。

5 请做一次终稿级别的技术核查。

请重点检查并统一以下内容：

- (1) 自检报告中“不涉密证明”状态是否仍为“尚未”；
- (2) 文中对参考文献状态的表述是否一致，如“Hu et al., 2026”与“Hu et al., in prep”；
- (3) 统计术语是否统一规范，如“二项线性混合效应模型”的表述；
- (4) 个别拼写与排版错误，如2.3节中“ikelihood ratio test”的拼写问题。

这些虽然是技术性细节，但建议在终稿中一次性清理干净。

回应：已对全文进行技术核查，

- (1) 对“不涉密证明”状态改为：已发邮件。
- (2) Hu et al., in prep 统一改为 Hu et al., 2026
- (3) 对“二项线性混合效应模型”表述统一改为更规范的“广义线性混合效应模型”
- (4) 已改为“likelihood ratio test”，并对全文进行拼写问题的核查和修改。