

第三轮修改说明

《5-8 岁儿童道德判断中的“黑羊效应”：重复性违规与代表性违规的影响》

尊敬的编辑老师与审稿专家：

您好！

非常感谢您再次抽出宝贵时间审阅我们的论文《5-8 岁儿童道德判断中的“黑羊效应”：重复性违规与代表性违规的影响》，并指出稿件在写作逻辑与表达严谨性上的完善方向。您的提醒让我们更清晰地意识到行文规范对论文质量的重要性，我们已针对全文的写作问题进行了系统梳理与逐处优化。为便于您审阅，本次修改稿的标注规则如下：

前两轮修改（对应初次和二次审稿意见）的内容仍保留红色和橙色字体标注；

本次新增修改的内容以绿色字体标注。

我们理解写作质量对论文录用至关重要，也会以此次修改为契机，进一步提升文稿的学术规范性。若您在审阅中仍发现有任何需要进一步调整之处，我们随时愿意配合完善。

再次感谢您耐心、专业的指导！

编辑部审核意见：论文修改后质量明显提升，同行一致认可。仍有写作问题需要大力完善。请根据同行评议修改写作，特别是逻辑整理和前后呼应一定要完善。需要特别提醒的是，近期部分论文在终审中由于写作问题最终被拒。因此，建议作者高度重视准确、严谨地表达。

回复：衷心感谢您对稿件的认可与细致指导！感谢您再次点明论文在写作逻辑与严谨性上的完善方向，您的提醒让我们更加重视行文的精准与连贯。我们非常认同您提出的意见，特别是关于全文写作逻辑的梳理与前后呼应的重要性，以及表达务必准确、严谨的要求。我们已对全文进行了系统性检查与修改。重点强化了以下方面：

（1）整体逻辑连贯性：重新审视并优化了引言、结果讨论等各部分之间的整体脉络，确保各部分之间衔接自然、论证层层递进；

（2）段落与句子间的呼应：强化了关键概念在不同部分的呈现一致性，避免前后脱节或重复表述；

（3）语言表达的准确性与严谨性：对冗长或晦涩的句子进行了拆分与重组，使论述更加流畅，易于理解，对可能引起歧义的表述进行了修正，力求用词精确、语句清晰，符合学术写作规范。

详见修改稿正文部分，请您审阅。

第二轮修改说明

《5-8 岁儿童道德判断中的“黑羊效应”：重复性违规与代表性违规的影响》

尊敬的编辑老师与审稿专家：

您好！

首先，衷心感谢您再次在百忙之中投入时间审阅我们的论文《5-8 岁儿童道德判断中的“黑羊效应”：重复性违规与代表性违规的影响》，并给予我们进一步完善稿件的宝贵机会。无论是初次审稿的细致指导，还是此次补充的修改意见，都为我们厘清研究逻辑、提升论文质量提供了关键方向。我们已逐字逐句研读所有意见，对提出的问题进行了深入核查与分析，针对每一条建议制定了具体修改方案，并全力以赴完成了相应调整。

为方便您高效审阅，我们在修改稿中做了如下标注：

针对首次审稿意见已完成的修改内容，仍以**红色字体**标注，保持与初次修改稿的连贯性；

针对二次审稿意见新增的修改内容，以**橙色字体**标注，便于快速识别本次完善的部分。

现将对本次二次审稿意见的逐条回复整理如下，其中“修改措施”以**蓝色字体**标注，“原稿件修改位置及修改内容”以对应颜色字体（**红色**为首次修改、**橙色**为本次新增修改）标注，敬请审阅：

编辑部审核意见：该文修改后明显提升了质量。同行尚有一些需要修改的意见，请作者务必仔细完成进一步修改。修改后将进一步评审。请作者注意本刊新近的要求，对照并对稿子予以完善，避免在终审环节因不符合本刊新近要求导致拒稿。

回复：衷心感谢您对稿件的认可与细致指导。我们已认真研读最新修改意见，将严格对照本刊新近要求，全力以赴完成进一步修改，确保完善所有细节，避免终审环节因格式或规范问题影响稿件进展。以下是我们针对审稿专家修改意见点对点的回复。希望我们的修改能够解决审稿人的意见和建议。

审稿专家一

审稿意见 1: 建议几个实验的流程可以综合成一张流程图, 有些环节可以删减, 比如指导语这一步就可以删减。目前的流程图, 只呈现了实验步骤, 并未给读者提供更多的信息, 通过看图, 不能了解作者的实验具体的操作。

回复: 感谢审稿专家的宝贵建议, 您的建议对提升流程图信息含量至关重要。为提升读者对实验流程的理解效率, 我们已将实验流程综合为一张图, 删减冗余环节并补充具体操作细节, 完成流程图的整合优化, 现已插入论文对应位置, 再次感谢您对提升论文可视化呈现质量的指导! 请您审阅。

详见修改稿 P6 “2.4 实验程序”, 摘录如下:

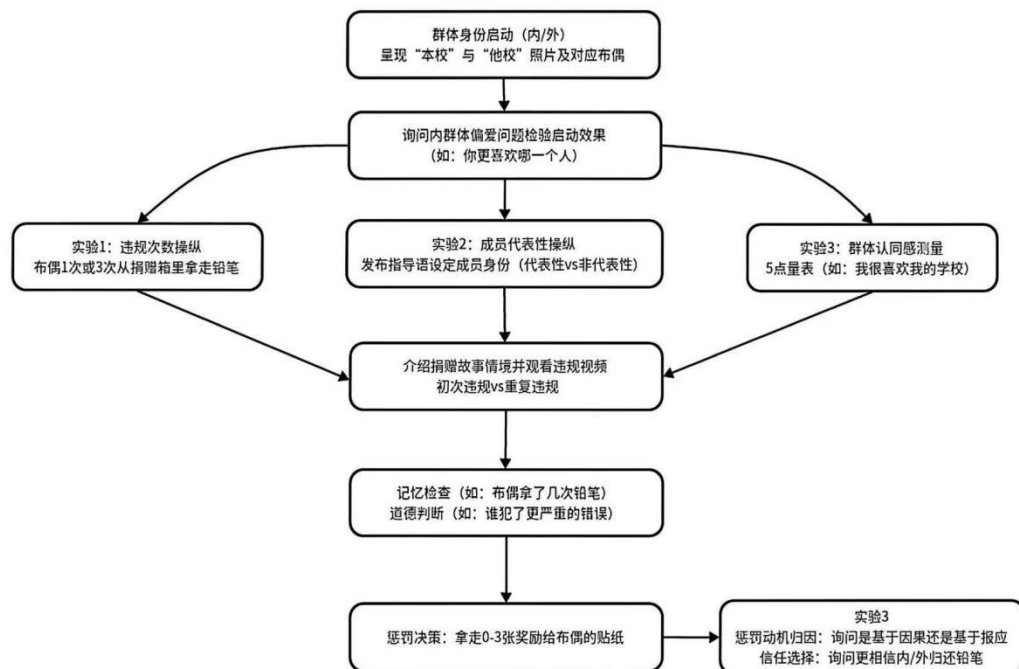


图 1 研究总体设计与实验流程

审稿意见 2: 结果的图需要把 SD 或者 SE 画出来, 而不能仅仅呈现均值。此外在图 2 中, 5-6 岁的儿童, 在外群体条件上, 似乎出现了重复违规的惩罚程度显著小于初次违规, 作者如何解释这一结果呢。

回复: 衷心感谢您对结果图表呈现细节及数据趋势解读的细致关注! 针对您指出的问题, 我们做如下说明和修改:

(1) 我们已在结果的折线图中补充了标准误 (SE), 以更完整地呈现数据变异情况,

提升结果的可解释性。详见修改稿 P8 图 2、P11 图 3，摘录如下：

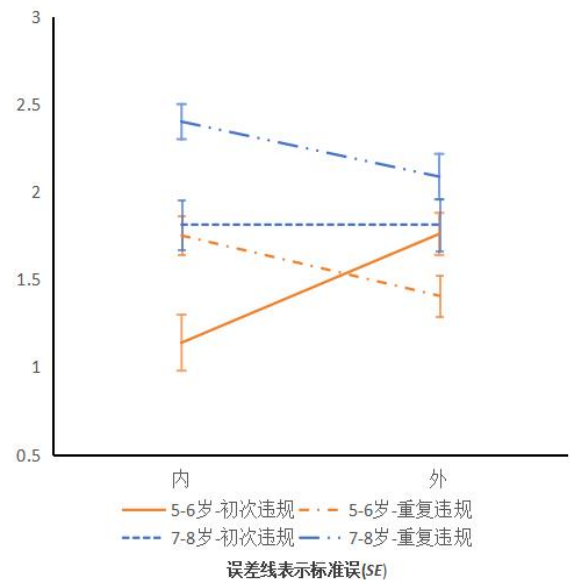


图 2 不同年龄儿童在不同违规情境下对内外群体的惩罚数量

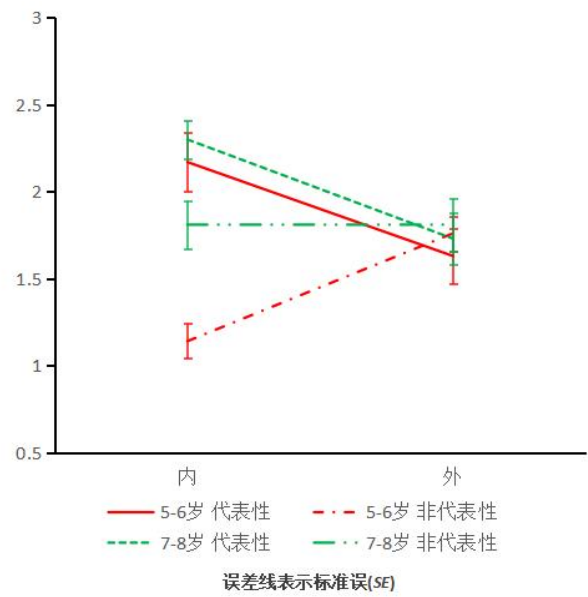


图 3 不同年龄儿童在不同代表性情境下对内外群体的惩罚数量

（2）关于“5-6 岁儿童外群体条件下重复违规惩罚程度小于初次违规”的趋势，统计检验显示其显著性为 $P=.059$ ，未达到统计学上的显著性水平。因此在统计学意义上，二者的惩罚差异并未达到显著程度。

另外我们再次梳理了实验流程，发现在实验一中，群体身份为被试内变量，年龄与违规次数为被试间变量，为了防止被试的疲劳，5-6 岁被试会随机分配至初次或重复违规中的任一条件，同一儿童不会同时参加两种违规次数情境，也就是说一组儿童仅经历内外群体成员

的初次违规情境，另一组儿童仅经历内外群体成员的重复违规情境。这意味着 5-6 岁外群体条件下，“初次违规组”与“重复违规组”的被试是完全独立的两拨样本，两组被试对惩罚强度的初始认知、贴纸惩罚的使用偏好等行为基准可能存在一些个体差异，但是这种差异在统计上未达到显著的水平。

综上，结合统计结果的非显著性与实验的设计，5-6 岁儿童外群体条件下的这一趋势不会影响本研究核心结论的可靠性。

审稿意见 3：作者应该当注意书面化表达，避免口语化表达。比如摘要中的“儿童惩罚内群体违规者主要是基于因果，希望其变好。”，“希望其变好”就是口语化表达，且表述的并不具体。文章中还有其他的口语化表达，希望作者全文通读并修改。此外，爱之深，责之切。在本文中使用也并不恰当，该句子通常用于描述父母对子女、老师对学生等关系中的深切关爱与严格要求，而本文主要的是关于研究内外群体的道德判断，更应该突出内群体社会规范维护。

回复：非常感谢您指出本文在书面化表达及特定表述适配性上的不足，您的意见为提升论文学术规范性、强化研究主题聚焦提供了关键指导。我们已对全文进行了多轮通读与系统修订。针对口语化表达及“爱之深，责之切”的使用问题开展系统修改，修改后的论文在学术规范性、表述严谨性与结构逻辑性方面均得到了有效提升。口语化语句加下划线标注，修改内容橙色标注。请您评阅。

详见修改稿，摘录如下：

第①处 摘要

（3）儿童惩罚内群体违规者主要是基于因果，希望其变好。研究表明，群体身份、违规情境影响儿童道德判断和惩罚行为，且儿童严惩内群体违规者体现出“爱之深，责之切”的心理。

改（3）儿童惩罚内群体违规者主要是基于因果，旨在使其将来的行为符合社会规范。研究表明，群体身份、违规情境影响儿童道德判断和惩罚行为，且儿童严惩内群体违规者体现出“严于律己，宽以待人”的心理。

第②处 2.5 结果与分析

这表明重复性违规会影响 5-8 岁儿童惩罚内外群体的严厉程度，且在群体身份的影响下，儿童会更严厉的去惩罚内群体重复违规者而非外群体重复违规者。

改 这表明重复性违规会影响 5-8 岁儿童惩罚内外群体的严厉程度,且在群体身份的影响下,儿童对内群体重复违规者施加的惩罚显著严于外群体重复违规者。

第③处 4.2 假设

假设 1: 儿童对待内群体的惩罚不仅仅是因为错误而惩罚,惩罚更重要的目的是为了**让犯错者改正错误,变得更好**,惩罚动机是**基于因果**;而对待外群体的错误,只是单纯为了惩罚而惩罚,惩罚动机是**基于报应**。

假设 2: 群体认同感的高低对影响儿童的信任选择,群体认同感越高的个体,越会选择相信内群体违规者经过严厉惩罚**未来变好**。

改 假设 1: 儿童对待内群体的惩罚不仅仅是因为错误而惩罚,惩罚更重要的目的是为了**让犯错者改正错误,使其行为符合社会规范**,惩罚动机是**基于因果**;而对待外群体的错误,只是单纯为了惩罚而惩罚,惩罚动机是**基于报应**。

假设 2: 群体认同感的高低对影响儿童的信任选择,群体认同感越高的个体,越会选择相信内群体违规者在经过严厉惩罚后,**未来的行为会更符合规范**。

第④处 4.5 结果与分析

这表明无论面对内群体还是外群体,儿童都**较少是为了惩罚而去惩罚他们**。

改 这表明无论面对内群体还是外群体,儿童**基于单纯的报应而进行惩罚的均较少**。

第⑤处 4.6 小结

研究三表明,7-8 岁儿童严惩内群体违规者的行为主要是**基于因果**(**希望违规者改正**),而非**基于报应**(**单纯的报复**)。

改 研究三表明,7-8 岁儿童严惩内群体违规者的行为主要是**基于因果**(**期望违规者纠正错误行为**),而非**基于报应**(**单纯的报复性惩罚**)。

第⑥处 5 总讨论

儿童严惩内群体违规者是**基于因果**(**希望他未来改正**),这表明,儿童对内群体的严厉惩罚实质上是一种“**爱之深,责之切**”的表现,旨在通过惩罚促使内群体成员**变得更好**。

改 研究三进一步发现,儿童严惩内群体违规者是**基于因果**(**期望其将来的行为符合社会规范**),这表明,儿童对内群体的严厉惩罚实质上是一种“**严于律己,宽以待人**”的表现,

旨在通过惩罚促使内群体成员遵守规范。

第⑦处 6 结论

（3）儿童对内群体和外群体的惩罚动机并不相同，具体而言，儿童严惩内群体违规者，其惩罚动机主要是基于因果，即希望通过本次惩罚，违规者能够吸取教训，以后能够变得更好。对于内外群体而言，儿童都较少地因为报应而产生惩罚行为。

改 （3）儿童对内群体和外群体的惩罚动机并不相同，具体而言，儿童严惩内群体违规者，其惩罚动机主要是基于因果，即期望通过惩罚使违规者认识到错误，并促使未来行为符合群体规范。对于内外群体而言，儿童都较少地因为报应而产生惩罚行为。

审稿专家二

审稿意见 1: 论文整体逻辑较为清晰。研究聚焦 5~8 岁儿童, 通过三项实验考察儿童对内外群体成员违反道德规范行为的道德判断惩罚特点及动机。实验设计合理, 但也存在需要改进的问题。

回复: 衷心感谢您对本文研究逻辑与实验设计所给予的充分肯定, 以及对论文修改提出的宝贵意见! 我们已仔细阅读并深入理解了您的每一条建议, 高度重视您所提出的问题, 并据此对文章进行了系统性的修改与完善, 旨在进一步提升研究的严谨性与表述的规范性。

审稿意见 2: 作者在引言部分对所研究内容的理论支撑及实证研究的呈现需要进一步梳理, 已呈现理论与实证研究需要相互佐证, 但修改后的综述, 在这一部分呈现较为生硬。

回复: 非常感谢您对本研究引言部分理论支撑与实证研究呈现逻辑的宝贵意见。针对“理论与实证需进一步梳理、呈现较为生硬”的问题, 我们高度重视并对引言部分重新梳理, 强化理论与对应实证研究相互佐证, 按照科学故事的逻辑对我们的引言部分进行重构, 新的结构将围绕“现象与矛盾→理论竞争解释→实证分歧→提出空缺→本研究方案”这一线索展开, 确保理论与实证相互佐证, 行文自然流畅。修改内容已插入引言对应位置, 橙色字体标注, 请您评阅。

详见修改稿 P1-3 “引言”部分, 摘录如下:

1 引言

群体身份认知会对个体的道德判断产生显著影响, 进而引发两种不同的现象。一种是内群体偏爱 (In-group Favoritism), 指个体在态度和行为等方面表现出来的对自身群体成员的偏爱。例如 1-2 岁幼儿当看到内群体成员受到伤害时, 会更期望旁观者去惩罚施害者, 而当内群体伤害他人时, 则会期待旁观者能给受害者提供一些帮助(Ting et al., 2019); 3-6 岁儿童会判定外群体成员踢人比内群体踢人影响更加恶劣, 更应受罚(Chapman et al., 2020); 与群体内违规者相比, 6 岁儿童更有可能惩罚自私的外群体违规者(Jordan et al., 2014)。与之对应的是“黑羊效应”(Black Sheep Effect), 即个体也可能会对内群体成员更加严厉和苛责。研究表明, 在第三方惩罚任务中, 5-6 岁女童更倾向于惩罚内群体违规者, 而男童会同等程度地惩罚内外群体违规者(Wu & Gao, 2018); 43%的儿童在目睹道德违规行为后选择惩罚内群体违规者, 仅有 17%的儿童惩罚外群体违规者(Yudkin et al., 2020)。

为何在面对内群体违规者时, 会出现从“偏爱”到“严惩”的转变? 群体身份如何作用于儿童的惩罚行为? 要解释这一看似矛盾的现象, 社会认同理论与主观群体动力理论提供了

两种竞争性的解释。社会认同理论(Tajfel & Turner, 2004)指出, 个体往往将内群体成员赋予积极的品质认知, 这种认知偏差使得人们对群体内部成员普遍表现出包容与偏爱倾向; 但主观群体动力理论(Abrams et al., 2003)表明, 人们存在限制和控制违规者以维持群体规范的共识, 当发现内群体成员出现违反群体规范的行为时, 个体可能会对内群体违规者更严厉, 即出现“黑羊效应”。这两种理论在儿童发展研究中均能找到支持证据, 但我们尚不清楚是哪些具体因素决定了儿童在特定情境下会表现出哪一种倾向。

随着研究的深入, 研究者开始探寻“黑羊效应”发生的条件。目前已有多个因素被证实能调节这一效应: 首先是违规行为所属的领域。道德在“黑羊效应”中扮演着核心角色(Tang et al., 2023), 社会领域理论将社会情境划分为道德领域、社会习俗领域和个人领域(Killen et al., 2013), 儿童在3岁时就优先表现出对道德规则的认同和更严格的惩罚倾向(Nucci, 1981; Yau & Smetana, 2003), 道德违规(相对于社会习俗违规)因其对群体核心价值的直接威胁, 更易引发强烈的“黑羊效应”(Tang et al., 2023)。其次是儿童的发展年龄。7岁前后被普遍视为群体认知与道德判断的转折点: 7岁前儿童更依赖群体身份进行判断, 表现出更强的内群体偏爱; 7岁后, 随着认知能力的发展, 儿童开始更多地考虑个体品质与情境因素(Aboud, 1988)。以往的研究表明, 5-11岁儿童随年龄增长更排斥内群体违规者(Abrams et al., 2003), 且“黑羊效应”仅在8岁及以上儿童中稳定出现, 6岁及以下儿童尚未表现出这一效应(Abrams et al., 2014)。第三是群体的人际互动模式。在竞争性互动中, 个体会对违背信任的内群体成员表现出更严厉的“黑羊效应”; 而在合作性情境中, 这一效应则消失(Cao et al., 2024)。这为主观群体动力理论提供了支持, 表明群体间的竞争会强化群体认同, 从而促使成员更严格地维护内部规范。

现有研究已揭示特定群体互动模式和儿童的发展阶段对“黑羊效应”的调节作用。然而, 根据主观群体动力理论, 对违规者的惩罚严厉程度, 取决于其行为对群体形象和规范威胁的严重性(Abrams et al., 2003)。其中, 违规行为的重复性与违规者的群体代表性, 是评估该威胁程度的关键维度, 但二者如何引发“黑羊效应”尚缺乏系统的实证探索。重复违规是否会放大群体身份的作用, 使儿童对内群体中的“害群之马”格外严厉? 当违规者是群体代表性的核心成员时, 儿童是会出于维护群体形象而从严处理, 还是会因内群体偏爱而从宽对待?

在道德违规情境中, 个体对群体违规者的评价与惩罚确实会受到群体身份和重复违规行为的影响(Gollwitzer & Keller, 2010; Ting & Baillargeon, 2021)。其理论意义在于, 单一偶发行为与重复性行为可能引发儿童对行为者意图的不同推断。例如, 在一项以隐藏公共图书为违规场景的研究中, 发现成人对内群体成员的重复违规惩罚更重(Gollwitzer & Keller,

2010)。让儿童观看布偶伤人视频，儿童也更倾向将重复多次伤人的布偶判定为坏人并支持对其进行惩罚(Langenhoff et al., 2022)。更有研究发现，相比单次伤害，三次连续伤害更易被幼儿判定为严重违规；同时该现象受群体身份影响，布偶伤害内群体成员时，仅需单次伤害行为幼儿即推测布偶后续会进行不公平分配，而布偶需多次伤害外群体成员，幼儿才会做出类似推测(Ting & Baillargeon, 2021)。然而这些研究未能直接回答：对于重复违规者，其内群体身份会导致更严厉的惩罚还是更宽容的对待？因此，研究一聚焦重复性违规情境，将“重复性违规”操作定义为同一主体在实验情境中连续三次违反同一道德规范，通过控制违规次数与违规者群体身份，系统考察儿童对内外群体成员违规行为的道德判断差异。

除重复违规外，违规者的群体代表性同样会影响个体对道德违规行为的判断(Bernard et al., 2016; Gollwitzer & Keller, 2010)。主观群体动力理论强调，对群体核心成员的监督与惩罚是维持群体价值的关键(Abrams et al., 2003)。研究表明，幼儿已能区分有影响力与无影响力的个体，并认为前者更有责任维系社会规范(Stavans & Baillargeon, 2019; Bernard et al., 2016)。作为“群体价值与规范体现者”的核心成员(Abrams et al., 2014; Fousiani et al., 2019)，其道德违规是会引发儿童的内群体偏爱还是“黑羊效应”？据此，研究二聚焦代表性违规情境，将群体核心成员界定为充分体现群体核心价值观、具有高影响力和示范效应，是群体内广受欢迎的榜样且被普遍模仿的成员(Abrams et al., 2003, 2014)，系统探讨儿童对内外群体代表性成员与非代表性成员道德违规行为的判断与惩罚特点，为儿童道德判断领域的研究提供实证补充。

同时，过去的惩罚动机研究表明，成人与幼儿惩罚道德违规兼具两种动机(Carlsmith et al., 2002; Marshall et al., 2020, 2022)，一种是基于因果，惩罚是为了让犯错者的行为符合社会规范；一种是基于报应，惩罚是让个体遭受痛苦，受到相应的报应。成人对内群体违规者惩罚更偏功利，主要是为了维护群体利益，而非单纯的报应(Fousiani et al., 2019)，但儿童对内外群体成员惩罚动机的差异，尚无研究涉及。因此，在研究一二的基础上，研究三将进一步探讨在特定的道德违规情境下，儿童惩罚内外群体违规者的惩罚动机。

综上所述，目前关于内群体偏爱的机制已较清晰，但关于“黑羊效应”在儿童道德发展领域上的研究仍存局限。随着年龄的增长，儿童的道德判断究竟是依赖于群体身份，还是越来越趋近于公平？为何个体在面对内群体违规者时存在内群体偏爱和“黑羊效应”两种现象？其中的影响因素（如重复违规、成员代表性）以及该现象的潜在动机需要实证研究去探明。目前对“黑羊效应”的研究大多聚焦于成人群体，而以儿童群体为调查对象的研究较少，“黑羊效应”在8岁以下儿童群体中究竟如何表现？年幼儿童是否也能够同时考虑到道德规

范和群体身份？基于上述疑问，本研究聚焦 5-8 岁儿童，其中 5-6 岁儿童的道德推理以单一道德规则为主，7-8 岁儿童处于认知转折关键期(Killen et al., 2013; Abrams et al., 2014; Aboud, 1988)。据此将儿童划分为 5-6 岁，7-8 岁两组，通过三项实验考察其在特定情境中对群体成员违反道德规范行为的道德判断、惩罚特点及内在动机，旨在通过细化年龄分组，系统考察两种现象在儿童不同发展阶段的表现特点，验证并拓展主观群体动力发展理论在儿童群体中的适用性。同时，深入剖析“黑羊效应”的影响因素及其背后的心理机制，以此揭示儿童视角下道德规范与群体身份认同的内在关联。研究结果可为道德教育提供理论支撑，也为家校教育实践提供科学参考，助力儿童道德认知与社会适应的健康发展。

审稿意见 3：作者在对已有研究梳理的过程当中提出的将儿童划分 5-6 岁及 7~8 岁两组来考察儿童道德判断的特点。但划分的依据并未明确说明，并且在结果的呈现过程中并未按已有划分的 5~6 岁，及 7~8 岁两个年龄段来分别呈现研究结果，在结果呈现中，会直接呈现 5~8 岁儿童的研究结果，这与文章实验设计研究设计不完全对应，建议进一步修改。

回复：非常感谢您的宝贵建议。针对您的建议，我们主要作出以下两点回复：

（1）非常感谢您指出本研究儿童年龄分组依据未明确的问题，我们已针对该问题补充完善了分组逻辑的说明，前人研究证实 8 岁及以上儿童“黑羊效应”稳定、6 岁以下无此效应，且 7 岁是儿童认知与道德判断的关键转折期，故本研究选取 5-8 岁并划分为 5-6 岁、7-8 岁两组。分组依据的详细说明已补充至引言，橙色字体标注，请您评阅。

详见修改稿 P2-3 “引言”部分第三段和第八段，摘录如下：

……其次是儿童的发展年龄。7 岁前后被普遍视为群体认知与道德判断的转折点：7 岁前儿童更依赖群体身份进行判断，表现出更强的内群体偏爱；7 岁后，随着认知能力的发展，儿童开始更多地考虑个体品质与情境因素(Aboud, 1988)。研究指出，5-11 岁儿童随年龄增长更排斥内群体违规者(Abrams et al., 2003)，且“黑羊效应”仅在 8 岁及以上儿童中稳定出现，6 岁及以下儿童尚未表现出这一效应(Abrams et al., 2014)。……

……目前对“黑羊效应”的研究大多聚焦于成人团体，而以儿童群体为调查对象的研究较少，“黑羊效应”在 8 岁以下儿童群体中究竟如何表现？年幼儿童是否也能够同时考虑到道德规范和群体身份？基于上述疑问，本研究聚焦 5-8 岁儿童，其中 5-6 岁儿童的道德推理以单一道德规则为主，7-8 岁儿童处于认知转折关键期(Killen et al., 2013; Abrams et al., 2014;

Aboud, 1988)。据此将儿童划分为 5-6 岁，7-8 岁两组，通过三项实验考察其在特定情境中对群体成员违反道德规范行为的道德判断、惩罚特点及内在动机，旨在通过细化年龄分组，系统考察两种现象在儿童不同发展阶段的表现特点，验证并拓展主观群体动力发展理论在儿童群体中的适用性。……

（2）针对“结果呈现需与 5-6 岁/7-8 岁分组设计对应”的建议，我们完全赞同。在此说明：本研究的研究一、二严格采用 5-8 岁儿童样本，并按 5-6 岁、7-8 岁分为两组开展实验，在对应“结果与分析”部分已分别报告了两个年龄段的具体数据（如不同违规情境下的惩罚强度、群体身份效应差异等），且明确分析了两组的年龄特异性特征，确保结果呈现与实验分组设计完全匹配。其次，研究三因核心目标是探究儿童惩罚行为背后的复杂心理动机，需儿童准确报告自身惩罚决策理由，我们在预实验中纳入 5-6 岁儿童，但发现该年龄段儿童因语言表达能力与逻辑推理水平尚未成熟，难以清晰、准确地描述这类抽象心理过程。因此，研究三仅选取认知能力更适配的 7-8 岁儿童开展实验并报告结果，该设计调整是基于研究结果的严谨和准确的考量，旨在保证动机数据的有效性与可靠性。

为使读者在第一时间就能把握核心发展性发现，我们已遵照您的建议，对“摘要”与“结论”部分进行了修改。严格按“5-6 岁儿童”和“7-8 岁儿童”两个组别，清晰地陈述了他们在不同违规情境下的行为特点，确保与研究设计的完全对应。请你评阅。

详见修改稿 P1 “摘要”部分、P16 “结论”部分，摘录如下：

摘 要：儿童的道德判断中存在内群体偏爱和“黑羊效应”两种相反的现象，揭示其不同的发生条件与动机至关重要。“黑羊效应”指对违反规范的内群体成员比外群体成员更严厉的现象。本研究聚焦 5-8 岁儿童，通过三项实验考察在重复性违规、代表性违规情境下，儿童对内外群体成员违反道德规范行为的道德判断、惩罚特点及动机。实验采用故事视频呈现违规行为，儿童通过拿走贴纸数量实施惩罚，并完成动机归因任务。结果发现：（1）违规次数调节惩罚行为，5-6 岁儿童在初次违规时显现内群体偏爱，而在重复违规时转为“黑羊效应”；7-8 岁儿童在初次违规中未表现偏爱，但在重复违规中呈现稳定的“黑羊效应”。（2）违规者代表性影响惩罚强度，5-6 岁与 7-8 岁儿童均对代表性内群体违规者惩罚更重。（3）儿童惩罚内群体违规者主要是基于因果，旨在使其将来的行为符合社会规范。研究表明，群体身份、违规情境影响儿童道德判断和惩罚行为，且儿童严惩内群体违规者体现出“严于律己，宽以待人”的心理。

6 结论

(1) 当群体成员初次违规时，5-6 岁的儿童给内群体成员较少的惩罚，外群体成员更多的惩罚，依旧保持内群体偏爱，7-8 岁儿童则未显现偏爱；而当群体成员重复出现违规行为时，5-8 岁的儿童会给内群体成员实施更加严厉的惩罚，而给外群体成员实施相对轻微的惩罚，出现了“黑羊效应”，且该效应在 7-8 岁组更稳定。

(2) 当违规者是群体中的代表性成员时，5-6 岁与 7-8 岁儿童均对内群体中的代表性成员实施更加严厉的惩罚，而给外群体成员实施相对轻微的惩罚，出现了“黑羊效应”。

(3) 儿童对内群体和外群体的惩罚动机并不相同，具体而言，儿童严惩内群体违规者，其惩罚动机主要是基于因果，即期望通过惩罚使违规者认识到错误，并促使未来行为符合群体规范。对于内外群体而言，儿童都较少地因为报应而产生惩罚行为。

在此，我们谨向编辑老师和各位审稿专家致以最诚挚的谢意。您们严谨而富有建设性的意见，是本研究得以完善的重要基石。这些宝贵的建议指明了稿件的完善方向，为我们系统理清研究逻辑、切实提升内容的科学性与严谨性提供了核心支撑，更促使我们对研究本身进行了更深层次的思考与提炼。我们已对全部意见逐一进行了认真回复与修改，相信修改后的稿件在逻辑严谨性、论述清晰度与结论说服力上均获得了实质性提升。再次感谢您们在百忙之中的指导与辛勤付出！

第一轮修改说明

《5-8 岁儿童道德判断中的“黑羊效应”：重复性违规与代表性违规的影响》

尊敬的编辑老师与审稿专家：

您好！

我们衷心感谢您在百忙之中审阅我们的论文《5-8 岁儿童道德判断中的“黑羊效应”：重复性违规与代表性违规的影响》，并给予我们修改稿件的机会。对于您提出的宝贵修改建议和意见，我们非常感谢并予以重视，全力以赴思考解决方案，进行了详细的修改和重新梳理。为便于审稿专家审阅，在修改稿中，对修改内容用红色字体进行标注。

以下是对您提出的宝贵审稿意见的逐条修改回复，包括我们的修改措施（蓝色字体标注）、原稿件修改位置及修改内容（红色字体标注）：

编辑部审核意见：同行肯定了研究的选题和基本构架，但也指出了论文存在的若干重要问题。请作者根据同行意见，逐条认真、细致和尽可能彻底地修改。修后再审。

回复：非常感谢您提出的宝贵意见和建议！我们已经对稿件进行了认真修改，使其更加清晰，便于读者理解。以下是我们针对审稿专家修改意见点对点的回复。希望我们的修改能够解决审稿人的意见和建议。

审稿专家一

审稿意见：该论文以“黑羊效应”为理论框架，通过三项实验系统探讨了 5-8 岁儿童在面对群体成员违规行为时的道德判断与惩罚动机，关注群体身份、违规情境（重复性与代表性）以及发展阶段在其中的作用机制。研究选题具有一定的理论创新性，将群体认同与道德发展相结合，尝试揭示儿童在复杂社会规范中的判断模式，具有积极意义。文章结构较为清晰，数据基本完整，结论具有启发性。但目前稿件在写作表达、方法描述与数据呈现等方面仍有较大提升空间，建议作者在以下方面进行进一步完善：

回复：感谢审稿专家对本研究选题价值、结构设计及结论意义的肯定与认可。您提出的关于写作表达、方法描述与数据呈现等方面的完善建议极具建设性，为我们进一步提升稿件质量指明了方向。我们将认真对照您的意见，逐点细致修改，力求在语言表达的精准性、研究方法的清晰度及数据呈现的规范性上得到显著提升，以期更好地呈现研究内容与价值。以下为具体修改说明：

审稿意见 1：建议在摘要首句中简要引入研究背景或选题意义，有助于读者更好地理解研究价值。同时，为便利非道德心理学背景的读者理解，建议在摘要中对“黑羊效应”做简要解释。

回复：十分感谢您提出的宝贵意见，根据您的建议，我们对摘要部分进行了相应的补充，并用红色字体标注，请您评阅。

详见修改稿 P1 “摘要” 部分，摘录如下：

儿童的道德判断中存在内群体偏爱和“黑羊效应”两种相反的现象，揭示其不同的发生条件与动机至关重要。“黑羊效应”指对违反规范的内群体成员比外群体成员更严厉的现象。本研究聚焦 5-8 岁儿童，通过三项实验考察在重复性违规、代表性违规情境下，儿童对内外群体成员违反道德规范行为的道德判断、惩罚特点及动机。实验采用故事视频呈现违规行为，儿童通过拿走贴纸数量实施惩罚，并完成动机归因任务。结果发现：（1）违规次数影响儿童惩罚行为，初次违规时，5-6 岁儿童表现出内群体偏爱；重复违规时，5-8 岁儿童出现“黑羊效应”。（2）违规者代表性影响儿童判断与惩罚，5-8 岁儿童对有代表性的内群体违规成员惩罚更重。（3）儿童惩罚内群体违规者主要是基于因果，希望其变好。研究表明，群体身份、违规情境影响儿童道德判断和惩罚行为，且儿童严惩内群体违规者体现出“爱之深，责之切”的心理。

审稿意见 2: 引言第二段中“代表性违规”“重复性违规”两个核心变量的定义尚不够明确, 建议作者在引用相关文献的同时, 结合本研究设计进一步阐明其操作含义和理论意义。

回复: 感谢审稿专家的细致建议。针对引言中“代表性违规”“重复性违规”定义不够明确的问题, 我们已结合相关文献和本研究设计进行了补充完善, 请您评阅。

详见修改稿 P2-3, 新增“引言”部分第 3、4 段, 摘录如下:

在道德违规情境中, 个体对群体违规者的评价与惩罚会受到群体身份和重复违规行为的共同影响(Gollwitzer & Keller, 2010; Ting & Baillargeon, 2021)。例如, 在一项以隐藏公共图书为违规场景的研究中, 对比内外群体成员的初次与重复违规行为, 发现内群体成员重复违规会引发被试更强的愤怒与更重的惩罚, 而外群体成员的重复违规无此现象(Gollwitzer & Keller, 2010)。让儿童观看布偶伤人视频, 结果显示儿童会更倾向将重复多次伤人的布偶判定为坏人, 并更支持对其惩罚(Langenhoff et al., 2022)。还有研究进一步发现, 儿童对违规行为的严重程度判断受违规频次的影响, 相比单次伤害, 三次连续伤害更易被幼儿判定为严重违规, 且会降低对违规者的道德评价, 同时该现象受群体身份影响, 布偶伤害内群体成员时, 仅需单次伤害行为幼儿即推测布偶后续会进行不公平分配, 而布偶需多次伤害外群体成员, 幼儿才会做出类似推测(Ting & Baillargeon, 2021)。然而现有研究仍存不足: 对于重复违规者, 其内群体身份会导致更严厉的惩罚还是更宽容的对待? 当面对外群体成员多次重复违规时, 儿童对他们道德判断和惩罚行为是否会一致? 因此, 研究一聚焦重复性违规情境, 将“重复性违规”操作定义为同一主体在实验情境中连续三次违反同一道德规范, 通过控制违规次数与违规者群体身份的差异, 系统考察儿童对内外群体成员违规行为的道德判断差异。

除重复违规外, 违规者的群体代表性同样会影响个体对道德违规行为的判断(Bernard et al., 2016; Gollwitzer & Keller, 2010)。已有研究初步揭示儿童对群体代表性的认知能力: 2 岁左右幼儿观察到布偶违反公平规范时, 会期望有代表性的领导者干预纠正, 而对非领导者干预无明确期待(Stavans & Baillargeon, 2019), 表明儿童此时已能初步区分有影响力者与无影响力者, 并认为前者有责任维系社会规范。3-5 岁幼儿在物品寻找与新异物品命名任务中, 会优先信任更具影响力的证言提供者(Bernard et al., 2016)。但现有研究尚未将群体代表性与道德违规情景结合, 儿童对群体成员道德违规行为的判断, 是否会因违规者的代表性身份不同而产生差异? 尤其对作为“群体价值与规范体现者”的核心成员(Abrams et al., 2014; Fousiani et al., 2019), 其道德违规是会引发儿童的内群体偏爱还是“黑羊效应”? 据此, 研究二聚焦代表性违规情境, 将群体核心成员界定为充分体现群体核心价值观、具有高影响力

和示范效应，是群体内广受欢迎的榜样且被普遍模仿的成员(Abrams et al., 2003, 2014)，系统探讨儿童对内外群体代表性成员与非代表性成员道德违规行为的判断与惩罚特点，为儿童道德判断领域的研究提供实证补充。

审稿意见 3: 建议作者补充儿童道德发展和动机归因的相关理论背景，例如社会域理论、道德动机的年龄差异研究等，以增强研究假设的理论支撑，也更有助于凸显本研究填补的文献空白。

回复: 感谢审稿专家的建议，您的这一建议对完善整个研究至关重要。充实社会域理论、道德动机年龄差异等相关理论背景，不仅能让研究假设的提出更具科学性和逻辑性，获得更坚实的理论支撑，也能更清晰地勾勒出当前研究在既有文献中的定位，有力凸显本研究所填补的文献空白。我们已根据意见进行了内容的补充，具体修改内容已用红色字体标记，请您评阅。

详见修改稿 P1-3 “引言” 部分，摘录如下：

1 引言

在个体的道德判断过程中，群体身份认知会产生显著影响，进而引发两种不同的现象。第一种是内群体偏爱（In-group Favoritism），指个体在态度和行为等方面表现出来的对自身群体成员的偏爱。例如 1-2 岁幼儿当看到内群体受到伤害时，会更期望旁观者去惩罚施害者，而当内群体伤害他人时，则会期待旁观者能给受害者提供一些帮助(Ting et al., 2019)；3-6 岁儿童会判定外群体成员踢人比内群体踢人影响更加恶劣，更应受罚(Chapman et al., 2020)；与群体内违规者相比，6 岁儿童更有可能惩罚自私的外群体违规者(Jordan et al., 2014)。与之对应的是“黑羊效应”（Black Sheep Effect），即个体也可能会对内群体成员更加严厉和苛责。研究表明，在第三方惩罚任务中，5-6 岁女童更倾向于惩罚内群体违规者，而男童会同等程度地惩罚内外群体违规者(Wu & Gao, 2018)；43%的儿童在目睹道德违规行为后选择惩罚群体内违规者，仅有 17%的儿童惩罚外群体成员(Yudkin et al., 2020)。内群体偏爱与“黑羊效应”的存在均具备合理性，社会认同理论指出，个体往往将内群体成员赋予积极的品质认知，这种认知偏差使得人们对群体内部成员普遍表现出包容与偏爱倾向(Tajfel & Turner, 2004)；但主观群体动力理论表明，人们存在限制和控制违规者以维持群体规范的共识，当发现内群体成员出现违反群体规范的行为时，个体可能会对内群体违规者更严厉，即出现“黑羊效应”(Abrams et al., 2003)。这种从“偏爱”到“严惩”的转变因何发生？群体身份如何作用于儿童的惩罚行为？仍需进一步探究。

社会领域理论将个体的道德发展置于具体的社会规则情境下理解,社会情境包括道德领域、社会习俗领域和个人领域(Killen et al., 2013),儿童在 3 岁时就表现出对道德规则以及其它规则的认同和偏好(Nucci, 1981),并且对道德违规惩罚更严格(Yau & Smetana, 2003)。同时,儿童道德判断受群体身份影响呈现年龄差异:5-11 岁儿童随年龄增长更排斥内群体违规者(Abrams et al., 2003);这一现象也符合儿童的认知发展轨迹,7 岁前的儿童以群体为中心,容易对内群体产生更积极的评价,7 岁后认知过程占据主导,更多的考虑个体品质,而不是其群体身份(Aboud, 1988)。值得注意的是, Abrams 等人的研究发现“黑羊效应”只出现在 8 岁及以上的儿童身上,未出现在 6 岁及以下的儿童身上(Abrams et al., 2014)。综上本研究选取 5-8 岁儿童,并将儿童划分为 5-6, 7-8 两组来考察儿童的道德判断特点,从而为理解该年龄段儿童道德判断的发展轨迹提供更细化的实证支持。

.....

同时,过去的惩罚动机研究表明,成人与幼儿惩罚道德违规兼具两种动机(Carlsmith et al., 2002; Marshall et al., 2020, 2022),一种是基于因果,惩罚是为了让犯错者改正错误,变得更好;一种是基于报应,惩罚是因为个体做错了,就应该遭受痛苦,受到相应的报应。成人对内群体违规者惩罚更偏功利,主要是为了维护群体利益,而非单纯的报应(Fousiani et al., 2019),但儿童对内外群体成员惩罚动机的差异,尚无研究涉及。因此,在研究一二的基础上,进一步开展研究三探讨在特定的道德违规情境下,儿童惩罚内外群体违规者的惩罚动机。

综上所述,目前关于内群体偏爱的机制已较清晰,但关于“黑羊效应”在儿童道德发展领域上的研究仍存局限。随着年龄的增长,儿童的道德判断究竟是依赖于群体身份,还是越来越趋近于公平?为何个体在面对内群体违规者时存在内群体偏爱和“黑羊效应”两种现象?其中的影响因素(如重复违规、成员代表性)以及该现象的潜在动机需要实证研究去证明,目前对“黑羊效应”的研究大多聚焦于成人群体,而以儿童群体为调查对象的研究较少。因此本研究聚焦 5-8 岁儿童,通过三项实验考察其在特定情境中对群体成员违反道德规范行为的道德判断、惩罚特点及内在动机,旨在验证并拓展主观群体动力发展理论在儿童群体中的适用性。同时,深入剖析“黑羊效应”的影响因素及其背后的心理机制,以此揭示儿童视角下道德规范与群体身份认同的内在关联。研究结果可为道德教育提供理论支撑,也为家校教育实践提供科学参考,助力儿童道德认知与社会适应的健康发展。

审稿意见 4：建议作者补充被剔除被试的具体人数及原因（如：年龄不达标（xx 名）等）。此外，由于作者分开比较了年长儿童和年幼儿童的结果，建议作者报告各年龄组的性别比例，以提高报告的透明性和严谨性。

回复：感谢审稿专家的细致指导，我们已仔细审核了数据，补充了实验一二三中被剔除被试的具体信息，我们也进一步在被试信息中报告了各年龄组的性别比例，确保数据报告的透明性与严谨性。修改内容如下，请您评阅。

详见修改稿 P4 “2.3.1 被试”、P8 “3.3.1 被试”、P12 “4.3.1 被试” 部分，摘录如下：

2.3.1 被试

从南昌市某幼儿园和小学随机招募 143 名 5-8 岁儿童，其中有 7 名年龄不达标，3 名月份不清，5 名没有展现出群体偏好，剔除 15 个无效数据，最后 128 名儿童参与研究，其中男性 67 名，女性 61 名；5-6 岁平均月龄 $M=72.18$ 月， $SD=6.49$ 月（男性 31 名，女性 30 名），7-8 岁平均月龄 $M=91.94$ 月， $SD=6.66$ 月（男性 36 名，女性 31 名）。这些儿童均自愿参与，分别实行个别实测。在本研究中，使用 G-Power3.1 软件计算事后统计效能，组间交互作用的效应量为.16，根据效应量 ($F=.16$)，alpha 设置为.05，总样本量设置为 $N=128$ ；结果检测到的统计检验效能为.95，符合标准。

3.3.1 被试

在南昌市某幼儿园和小学随机招募 129 名 5-8 岁儿童，其中有 3 名年龄不达标，1 名月份不清，1 名没有展现出群体偏好，剔除 5 个无效数据，最后 124 名儿童作为有效被试，其中男性 61 名，女性 63 名；5-6 岁平均月龄 $M=70.93$ 月， $SD=6.73$ 月（男性 28 名，女性 31 名），7-8 岁平均月龄 $M=91.78$ 月， $SD=6.01$ 月（男性 33 名，女性 32 名）。这些儿童均自愿参与，分别实行个别实测。在本研究中，使用 G-Power3.1 软件计算事后统计效能，组间交互作用的效应量为.24。根据效应量 ($F=.24$)，alpha 设置为.05，总样本量设置为 $N=124$ ；结果检测到的统计检验效能为.99，符合标准。

4.3.1 被试

在南昌市某小学随机招募 70 名 7-8 岁儿童，其中有 6 名月份不清，3 名无法正确理解实验要求，剔除 9 名无效数据，最后 61 名儿童参与研究，其中男性 33 名，女性 28 名，平均月龄 95.95 月。这些儿童均自愿参与，分别实行个别实测。在本研究中，使用 G-Power3.1 软件计算事后统计效能，成对样本 t 检验的效应量为.48。根据效应量 ($r=.48$)，alpha 设置为.05，总样本量设置为 $N=61$ ；结果检测到的统计检验效能为.95，符合标准。

审稿意见 5：建议加强实验流程描述，当前的实验流程描述略显笼统，建议作者补充具体的操作细节，例如儿童完成的“记忆类问题”“认知评估问题”的举例说明；群体内外的操作为被试内还是被试间的实验操作；此外，作者描述“因变量为儿童对内外群体成员违规行为严重程度的道德判断及对违规者的惩罚数量”，在实验流程中，仅看到了儿童对于违规行为的惩罚程度，请作者描述如何测量儿童的道德判断。

回复：感谢审稿专家的细致建议。针对您指出的问题，我们做如下说明和修改计划：

为完善实验流程描述，我们系统梳理了研究一二三各环节，从实验所需材料开始完善，在保证语句简洁的前提下详细说明了实验流程，确保操作目的、方式及注意事项明确可辨，以提升描述的透明性、严谨性与可重复性。专家提出来的问题也一一进行回答：儿童完成的问题已在文中列出：

实验一中群体身份为被试内变量，年龄为被试间变量，群体成员代表性为被试间变量，实验二中群体身份为被试内变量，年龄为被试间变量，群体成员代表性为被试间变量。已在实验设计处进行补充说明：

儿童的道德判断是通过询问认知评估问题（你认为布偶 1/2 犯了错误吗？谁犯了更严重的错误？），并将儿童的判断进行编码，统计两种违规情境下的判断比例，并进一步通过卡方检验分析差异，原文中进行了具体的说明。修改内容已用红色字体标记，请您评阅。

详见修改稿 P4-5 “2.3.2 实验设计” “2.3.3 材料” “2.4 实验程序”；P8-9 “3.3.2 实验设计” “3.3.3 材料” “3.4 实验程序”；P12-13 “4.3.3 材料” “4.4 实验程序” 部分，摘录如下：

2.3.2 实验设计

采用 2（群体身份：内群体、外群体）×2（违规次数：1 次、多次）×2（年龄：5-6 岁、7-8 岁）的混合实验设计，其中群体身份为被试内变量，年龄为被试间变量，违规次数为被试间变量，因变量为儿童对内外群体成员违规行为严重程度的道德判断及对内外群体违规者各自的惩罚数量。

2.3.3 材料

（1）实验材料：布偶四个；彩色铅笔若干支；铅笔盒一个；自制捐款箱一个；学生所在学校照片若干张，与学生无关的其他学校的照片若干张；自创故事情境，并将故事情境拍摄成短视频，视频素材随机找数名适龄儿童进行施测，开展预实验，确保实验材料的有效性。

（2）惩罚材料：小红花贴纸若干；

（3）奖品：每位完成实验任务的被试均能获得一些可爱的小红花贴纸奖励， 或一些文

具用品。选用原则为不伤害儿童身心健康、儿童喜爱、家长可接受。

注：研究二 3.3.3 研究三 4.3.3 实验材料均与研究一相同，在此不多做赘述。

2.4 实验程序

在安静的房间内对儿童逐个进行测试，并且告知儿童参与游戏，完成后会有奖励，每个儿童大概需要 10 分钟时间完成实验。由于布偶与真人在引发儿童社会认知反应上具有等效性(Yu & Wellman, 2022)，儿童既清楚布偶的虚构属性，又会自然的将其视作可互动的社会主体，反映真实认知能力(Rakoczy, 2022)，因此实验采用布偶来呈现违规场景。正式开始前给被试介绍两位新成员，布偶 1 是和你一个学校的，而布偶 2 来自别的学校。整体实验流程为：（1）群体身份启动，沿用以往研究中对内、外群体的启动范式(Fehr et al., 2008；刘璐等, 2019)，向被试呈现两张含有学校标志性建筑的照片，分别为被试所在学校和其他一所学校。在两张照片前分别放置一个类似的布偶来代表小组成员和外群体成员（性别与参与者相匹配）。询问内群体偏爱问题检验启动效果，例如：“谁是你的同学？你会更喜欢哪一个人？你手中有一张贴纸，你想把它送给谁？”（2）介绍故事情境，两个布偶在上周末带着各自的彩色铅笔，一同前往参与为山区小朋友捐赠文具的公益活动，随后播放违规行为的视频，视频内容改编自过往实验中的反社会行为的操作(Misch & Dunham, 2021)，具体如下：内外群体布偶面前各有 3 支自己的铅笔，旁白介绍捐赠箱内已有 4 支铅笔，并说明可自主捐赠，箱子内所有铅笔将捐给山区没有彩色铅笔的儿童，随后呈现两名成员从捐赠箱拿铅笔的场景，视频中两名布偶有 1 次或 3 次从捐赠箱拿铅笔的行为，每人每次只拿一支铅笔。每个儿童先后观看内外群体成员的违规情境，视频中拿取顺序已进行了平衡。（3）故事呈现完毕后，研究人员向儿童被试依次提出三个记忆类问题来测试幼儿对故事的记忆和理解能力：①谁是你的同学？②布偶 1/布偶 2 捐铅笔了吗？③你还记得布偶 1/布偶 2 拿了多少次彩色铅笔吗？两个认知评估类问题考察儿童的道德判断：①你认为布偶 1/布偶 2 犯了错误吗？②谁犯了更严重的错误？并将儿童的判断进行编码，分别把判断内群体成员错误更严重、判断外群体成员错误更严重、一样严重编码为 1、2、3。所有儿童都正确地回答了这些记忆问题，并判断视频中的两位布偶都做错了。（4）儿童的惩罚决策，观看视频后，要求儿童决定是否要拿走原本奖励给布偶的三张小红花贴纸作为惩罚，可以选择不拿，如果想拿的话，你要拿几张贴纸？用拿走贴纸的数量来表示对内/外群体的惩罚力度。

3.3.2 实验设计

采用 2（群体身份：内群体、外群体）×2（群体成员代表性：代表性、非代表性）×2（年龄：5-6 岁、7-8 岁）的混合实验设计，其中群体身份为被试内变量，年龄为被试间变量，群体成员代表性为被试间变量；因变量与研究一相同。

3.3.3 材料

实验材料同实验 1。

3.4 实验程序

准备步骤及群体身份启动步骤同研究一。群体身份启动后，对部分儿童进行群体成员身份代表性启动，对群体成员代表性的界定为：代表性成员指充分体现群体核心价值观、具有高影响力和示范效应的成员(Abrams et al., 2003, 2014)，非代表性成员参考群体边缘成员特征，即参与度低、影响力弱(McGuire et al., 2020)。通过实验者发布指导语的方式进行启动，情境为描述布偶在学校的受欢迎程度和带头作用来确立其代表性或非代表性身份，共有四种身份形象：代表性情境-内群体代表/外群体代表、非代表性情境-内群体非代表/外群体非代表，例如代表性情境-内群体代表指导语为“在你的学校非常受欢迎，班上活动都是他/她带头做起的，是你们的榜样，大家都喜欢模仿他/她的行为”，非代表性情境-外群体非代表指导语为“在一个与你无关的学校不太受到关注，班上的活动他/她也不怎么参与，是班级里非常普通的一位同学，大家都不会去模仿他/她的行为”；身份形象确立后，询问儿童“最近学校要组织给贫困山区捐赠铅笔，你觉得他们会不会带头捐呢？”；接着介绍捐赠铅笔的故事情境，播放视频，视频内容同研究一违规情境中的片段；结束后询问儿童记忆问题、认知评估问题，并让儿童进行惩罚决策（同研究 1）。

4.3.3 材料

实验材料同实验 1。

4.4 实验程序

准备步骤及群体身份启动步骤与研究一相同。采用个别施测的方式，在相对独立的环境下对每个儿童进行操纵。群体身份启动成功后，测量儿童的群体认同感，采用标准化群体认同感量表(Tajfel & Turner, 2004)，翻译成中文后题目如下：“能在这里上学我很高兴”“能在这里上学我很光荣”“能在这里上学我很骄傲”“我很喜欢我的学校”“我很喜欢我的同学”；接着介绍捐赠铅笔的故事情境，播放研究一中的违规片段；让儿童实施惩罚：“布偶 1 和 2 都很喜欢这个贴纸，希望能留下来，现在你决定是否还要把贴纸拿走，如果拿走的话，要拿走多少张”，若儿童对内群体实施更严厉惩罚则继续研究。实验者口述惩罚动机，惩罚动机引用 Marshall 等人(Marshall et al., 2022)，其中包括基于因果动机（例如：“我希望我们

/其他学校的犯错者不会再去偷捐赠物品”）和基于报应动机（例如“我觉得我们/其他学校的犯错者就应该受到这样的处罚”），让儿童做出认可度判断，采用直接提问的方式，“惩罚她/他是想让他改正？（基于因果）还是因为他该受罚？（基于报应）”，最后测量儿童的信任问题，告知儿童捐款箱里的铅笔全部被还了回来，并且还多了三支，询问“你更愿意相信是谁做的这件事？”

审稿意见 6：为提高读者理解效率，建议增补一张简洁明了的实验流程图。

回复：感谢审稿专家的宝贵建议。为提升读者对实验流程的理解效率，我们分别对研究一二三各增补一张流程图，现已插入论文中，请您评阅。

详见修改稿 P5 “2.4 实验程序”、P9 “3.4 实验程序”、P13 “4.4 实验程序” 部分，摘录如下：



图 1 实验一流程图

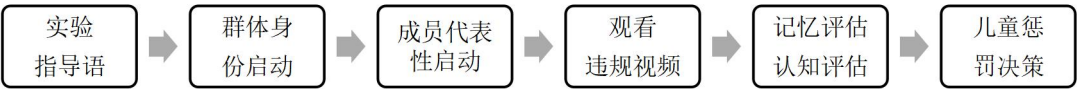


图 3 实验二流程图

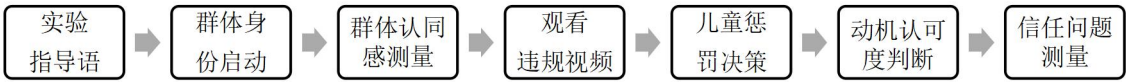


图 5 实验三流程图

审稿意见 7：实验中采用布偶呈现道德违规情景，建议作者简要说明为何选择该形式而非真人视频，是否已有文献支持其对该年龄段儿童的真实感或社交共鸣效果。同时，建议说明群体身份（“我们学校的布偶”）是否对儿童具有现实意义，是否经过群体身份的操纵

检验。

回复：感谢审稿专家的细致建议。本研究采用布偶呈现道德违规情景，主要基于以下考量：布偶在儿童研究中应用广泛，研究表明其与真人视频在引发儿童社会认知反应上具有等效性(Yu & Wellman, 2022)，布偶的使用符合儿童认知特点，儿童既清楚布偶的虚构属性，又会自然的将其视作可互动的社会主体，反映真实认知能力(Rakoczy, 2022)，并且在实验中，布偶形象更能吸引幼儿的注意力，且幼儿能够更好地将自己的认知与木偶的行为表现相联系，相比真人视频，木偶形式降低了幼儿理解过程中可能受到的无关因素干扰，使他们更专注于认知任务本身。关于群体身份的操纵检验，本研究在群体身份启动后，通过询问内群体偏爱问题来验证群体划分的有效性，问题如下：“谁是你的同学？你会更喜欢哪一个人？你手中有一张贴纸，你想把它送给谁？”核心逻辑是若群体划分有效，儿童应更倾向偏爱内群体，此类样本能更精准反映内外群体身份对道德判断的实际影响。未能展现出群体偏好的被试数据已被剔除，不进行下一步的研究（已在各实验“被试”部分注明剔除数量及原因），绝大多数的被试展现出了对内外群体的不同偏好，在研究一和研究二中有效被试出现内群体偏爱的儿童占比分别为 84.4%和 85.5%，并且具有统计显著性（卡方检验 $P<.001$ ），非随机误差导致。因此内外群体的划分在样本中是具有有效性的。具体修改和增加内容已用红色字体标记，请您评阅。

详见修改稿 P5-6 “2.4 实验程序” 部分，摘录如下：

2.4 实验程序

在安静的房间内对儿童逐个进行测试，并且告知儿童参与游戏，完成后会有奖励，每个儿童大概需要 10 分钟时间完成实验。由于布偶与真人在引发儿童社会认知反应上具有等效性(Yu & Wellman, 2022)，儿童既清楚布偶的虚构属性，又会自然的将其视作可互动的社会主体，反映真实认知能力(Rakoczy, 2022)，因此实验采用布偶来呈现违规场景。正式开始前给被试介绍两位新成员，布偶 1 是和你一个学校的，而布偶 2 来自别的学校。整体实验流程为：（1）群体身份启动，沿用以往研究中对内、外群体的启动范式(Fehr et al., 2008; 刘璐等, 2019)，向被试呈现两张含有学校标志性建筑的照片，分别为被试所在学校和其他一所学校。在两张照片前分别放置一个类似的布偶来代表小组成员和外群体成员（性别与参与者相匹配）。询问内群体偏爱问题检验启动效果，例如：“谁是你的同学？你会更喜欢哪一个人？你手中有一张贴纸，你想把它送给谁？”……

审稿意见 8：目前实验假设主要基于已有的研究结果，但理论推导略显薄弱。建议作者在提出假设时更明确地基于社会认同理论、主观群体动力模型或道德动机发展文献加以支持，或者前任研究证据，以增强研究逻辑的内在连贯性。

回复：感谢审稿专家的宝贵建议，您指出的实验假设理论推导薄弱问题十分关键，我们对此进行了深刻反思和认真修改。在修订过程中，我们重新梳理并强化了三个实验假设提出的理论逻辑：首先，我们更为明确地将社会认同理论(Tajfel & Turner, 2004)和主观群体动力理论(Abrams 等, 2003)作为核心理论框架，用以说明儿童对内外群体成员偏差行为做出不同反应的深层动机；其次，我们补充引用了近期儿童道德动机发展的相关文献，如(Marshall et al., 2020, 2022)，进一步从发展角度阐明儿童惩罚行为背后可能存在的基于因果动机与基于报应动机，并与社会认同建立联系。这些理论内容的增补，已体现在各实验假设提出前的论述中（原文已用红色字体标注），显著增强了从理论到假设的逻辑连贯性与推导合理性。再次感谢您的悉心指导。请您评阅。

详见修改稿 P3 “2.2 假设”、P8 “3.2 假设”、P11 “4.2 假设” 部分，摘录如下：

2.2 假设

根据社会认同理论，个体倾向于通过偏爱内群体成员来维持积极的社会认同(Tajfel & Turner, 2004)，而根据主观群体动力理论，儿童随着年龄的增长，会越来越反对内群体违规者，特别是违规行为威胁到群体规范的一致性，会引发更强烈的纠正动机(McGuire et al., 2020)，重复违规意味着对群体规范的持续挑战，此时儿童为维护内群体的积极形象，会对内群体违规者施加更严厉的惩罚(Gollwitzer & Keller, 2010)。据此提出以下假设：

假设 1：当内群体，外群体都是初次违规时，儿童对内群体的惩罚会小于外群体，即儿童仍旧存在内群体偏爱；当内群体，外群体重复违规时，儿童对内群体的惩罚会大于外群体。

假设 2：年长的儿童（7-8 岁）较年幼的儿童（5-6 岁）惩罚得更严厉，在初次违规情境下，年幼儿童较年长儿童更容易受到内群体偏爱的影响，对内群体违规者惩罚更少；在重复违规情境下，所有年龄段的儿童都会受到“黑羊效应”现象的影响，对内群体违规者惩罚更多。

3.2 假设

根据主观群体动力理论(Abrams et al., 2003)，群体代表性成员被视为内群体规范与价值额典型代表，其行为对内群体形象和认同具有更显著的影响，代表性成员的违规行为不仅是个体偏差，更可能被感知为对群体规范的严重威胁(Abrams et al., 2014; Fousiani et al., 2019)。

相比之下，非代表性成员的违规行为对群体认同的威胁较低，儿童更可能表现出一般性的内群体偏爱(Tajfel & Turner, 2004)。因此在研究二在研究一的基础上进一步提出以下假设：

假设 1：当内群体、外群体都是非代表性违规者时，儿童对内群体违规者的惩罚会小于外群体违规者，即保持内群体偏爱。

假设 2：当内群体、外群体都是代表性违规者时，儿童给内群体违规者比外群体违规者更加严厉的惩罚，出现了“黑羊效应”。

4.2 假设

道德动机发展研究表明，成人对道德违规行为的惩罚存在两种动机：基于因果动机与基于报应动机(Carlsmith et al., 2002)。新近针对幼儿的实验研究也得出类似结论，儿童作为第三方惩罚者时，其动机已表现出与成人相似的复杂性，但仍具有发展特异性：他们更倾向于支持可带来积极改变的教育性惩罚，而较为抗拒纯粹基于报复的惩罚(Marshall et al, 2022)。此外，基于主观群体动力理论(Abrams et al., 2003)，对内群体成员的惩罚更可能被视为一种规范的维护手段（基于因果），而对外群体的惩罚则更易被视为一种界限的划清与报复（基于报应），社会认同理论(Tajfel & Turner, 2004)认为，高群体认同感会强化个体对群体规范的维护承诺。因此，群体认同感高的儿童不仅更可能出于因果动机惩罚内群体违规者，也更可能相信惩罚的有效性，即预期违规者能够改过自新(Hogg et al., 2017)。基于此，本研究提出以下假设：

假设 1：儿童对待内群体的惩罚不仅仅是因为错误而惩罚，惩罚更重要的目的是为了犯错者改正错误，变得更好，惩罚动机是**基于因果**；而对待外群体的错误，只是单纯为了惩罚而惩罚，惩罚动机是**基于报应**。

假设 2：群体认同感的高低对影响儿童的信任选择，群体认同感越高的个体，越会选择相信内群体违规者经过严厉惩罚未来变好。

审稿意见 9：作者使用方差分析处理儿童惩罚行为数据（取值范围 1-3），由于该变量为离散分数，建议作者考虑是否使用非参数检验或秩次检验更为恰当。此外，对于显著交互效应（如实验一中的两因素、三因素交互），建议进一步报告简单效应分析。

回复：感谢审稿专家的宝贵建议，针对您提出的问题，我们进行了深入考量与补充分析，具体说明如下：

关于惩罚行为数据的统计方法：本研究的因变量是儿童惩罚行为数据，取值为 0-3（全距为 4）的离散分数（这里取值为 0 表示儿童可以选择不拿贴纸）。因此我们重新采用非参

数检验（Kruskal-Wallis 检验）对数据进行分析。

研究一探索年龄、违规次数和群体身份对儿童惩罚行为的影响，结果如下：①年龄对惩罚行为的影响，7-8 岁儿童的惩罚强度显著高于 5-6 岁儿童（ $H=26.565$ ， $df=1$ ， $P<.001$ ）②违规次数，重复违规会引发儿童更严厉的惩罚行为，违规严重程度越高，儿童惩罚强度越强（ $H=7.659$ ， $df=1$ ， $P=.006$ ）③群体身份单独对惩罚行为无显著影响（ $H=0.204$ ， $df=1$ ， $P=.651$ ）。

研究二探索年龄、代表性和群体身份对儿童惩罚行为的影响，结果如下：①年龄对惩罚行为的影响边缘显著，7-8 岁组惩罚强度有高于 5-6 岁组的趋势（ $H=3.714$ ， $df=1$ ， $P=.054$ ）②代表性对惩罚行为的影响，代表性情境下，儿童的惩罚强度显著更高（ $H=9.309$ ， $df=1$ ， $P=.002$ ）③违规次数，重复违规会引发儿童更严厉的惩罚行为，违规严重程度越高，儿童惩罚强度越强（ $H=7.659$ ， $df=1$ ， $P=.006$ ）④群体身份单独对惩罚行为无显著影响（ $H=2.074$ ， $df=1$ ， $P=0.150$ ）。

通过结果对比发现，非参数检验与原方差分析结果核心结论一致，但具体交互效应仍无法判断，我们又搜索到了一种非参数方差分析的方法，ART（Aligned Rank Transform，对齐秩变换），是一种用于分析非参数因子设计的方法，其核心逻辑是通过对齐过程分离效应（主效应和交互效应），再对对齐后的响应进行秩变换和方差分析。操作完之后收到提示说这种方法不太适合当前数据，具体原因不明。另外经查阅文献发现 2024 年《心理科学》的一篇题目为《5~9 岁儿童不公平厌恶的发展及其影响因素》的文章，该研究中因变量为不利不公平总分，取值范围为 0-4 分，同样采用多因素方差分析探究儿童发展差异(姜怡, 丁芳, 2024)，验证了方差分析在小范围整数数据、儿童发展研究中的适用性。综合考虑过后，我们原文中仍保留原方差分析结果，并且参照建议增加了显著交互效应（如实验一中的两因素、三因素交互）的简单效应分析。同时考虑是否可以在结果分析部分增加非参数检验结果来作为补充，用来支撑方差分析结果的准确性。审稿人若有进一步建议或想法，我们真诚期待再次沟通探讨，十分感谢！

详见修改稿 P7 “2.5 结果与分析”、P10 “3.5 结果与分析” 部分，摘录如下：

2.5 结果与分析

进一步进行简单-简单效应分析，结果显示，5-6 岁的儿童在初次违规情景下惩罚内群体的数量会低于外群体，7-8 岁儿童惩罚内群体和外群体的数量没有显著差异；此外，在重复违规情景下，两组儿童惩罚内群体的数量均高于外群体($P<.001$)。这表明重复性违规会影响 5-8 岁儿童惩罚内外群体的严厉程度，且在群体身份的影响下，儿童会更严厉的去惩罚内群

体重重复违规者而非外群体重复违规者。

3.5 结果与分析

进一步进行简单-简单效应分析,结果显示,在代表性条件下,各年龄组对内群体违规者的惩罚显著高于外群体违规者($P<.001$)。此外,在非代表性条件下,5-6岁儿童对内群体违规者的惩罚小于外群体违规者,而7-8岁儿童对内群体违规者与外群体违规者之间的惩罚不显著。此外,7-8岁的儿童比5-6岁的儿童的惩罚数量更多。

审稿意见 10: 建议统一按照 APA 第七版格式报告统计值。除非 $P < .001$, 其余应直接报告具体 P 值 (例如 $p = .034$), P 值保留三位有效数字, 其余统计指标保留两位。如当前文中出现 “ $P<0.05$ ” 等表达应修改为具体值。

回复: 非常感谢您指出关于统计值报告格式的问题, 这一意见对规范论文的学术表达至关重要。我们已认真对照 APA 第七版格式要求, 对全文的统计值报告进行了系统修正, 后续我们会进一步强化对 APA 格式的遵循意识, 确保论文细节的规范性。修改内容已用红色字体标注, 请您评阅。

详见修改稿 P6 “2.5 结果与分析”、P10 “3.5 结果与分析”、P13 “4.5 结果与分析” 部分, 摘录如下:

2.5 结果与分析

结果显示违规次数主效应显著, $F(1,124)=6.63, P=.011, \eta_p^2=.05$, 年龄主效应显著, $F(1,124)=22.35, P<.001, \eta_p^2=.15$, 群体身份和违规次数的交互作用显著, $F(1,124)=23.89, P<.001, \eta_p^2=.16$, 群体身份和年龄的交互作用显著, $F(1,124)=5.11, P=.026, \eta_p^2=.04$, 违规次数, 群体身份和年龄三者交互作用显著, $F(1,124)=6.18, P=.014, \eta_p^2=.05$ 。

3.5 结果与分析

其次, 通过方差分析不同代表性情境下各年龄组儿童惩罚内/外群体违规者的均值, 发现违规者代表性主效应显著, $F(1,120)=6.85, P=.010, \eta_p^2=.05$ 。此外, 违规者代表性与群体身份之间交互作用显著, $F(1,120)=39.45, P<.001, \eta_p^2=.25$ 。群体身份与年龄之间也存在显著的交互作用, $F(1,120)=5.80, P=.018, \eta_p^2=.05$ 。违规者的代表性、群体身份、年龄三者间存在交互作用, $F(1,120)=4.41, P=.038, \eta_p^2=.04$ 。

4.5 结果与分析

接着对内群体和外群体的惩罚动机认可程度进行配对样本 t 检验, 发现儿童对内群体和外群体成员的结果主义惩罚动机认可程度差异显著, $t(60)=4.24, P<.001$, 对报应主义惩罚动机认可程度无差异, $t(60)=1.81, P=.075$ 。这表明无论面对内群体还是外群体, 儿童都较

少是为了惩罚而去惩罚他们。在信任问题上，卡方检验结果显示两者差异显著， $\chi^2=17.85$ ， $P<.001$ 。这表明儿童更愿意相信内群体成员会改正错误。进一步将选择信任内群体违规者变好和选择信任外群体违规者变好的被试的群体认同感得分进行单样本 t 检验，结果发现差异显著， $t(59)=2.19$ ， $P=.032$ ，说明群体认同感越高的儿童，越倾向于信任内群体成员改正错误。

审稿意见 11：当前文章主要以表格形式呈现主要结果，建议采用图形（如条形图或折线图）呈现关键变量的交互结果，能更直观显示“黑羊效应”在不同群体与年龄间的差异，建议附图中标出显著性标记（如 *、**）与标准误。

回复：感谢您提出的宝贵建议，意见对提升论文结果的直观性和可读性极具启发。我们已按要求进行修改，补充折线图来更直观显示“黑羊效应”在不同群体与年龄间的差异。

详见修改稿 P7 “2.5 结果与分析”、P11 “3.5 结果与分析” 部分，摘录如下：

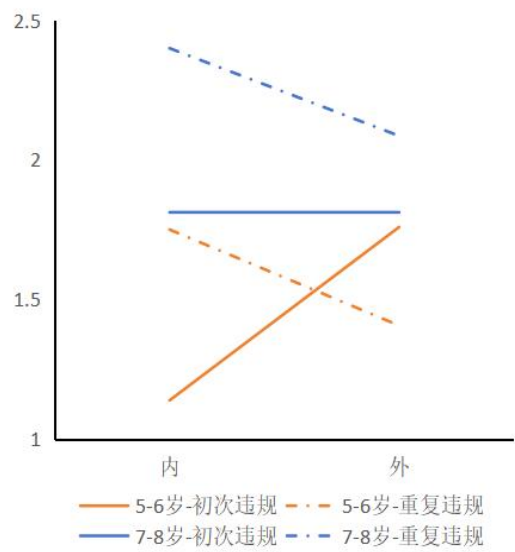


图 2 不同年龄儿童在不同违规情境下对内外群体的惩罚数量

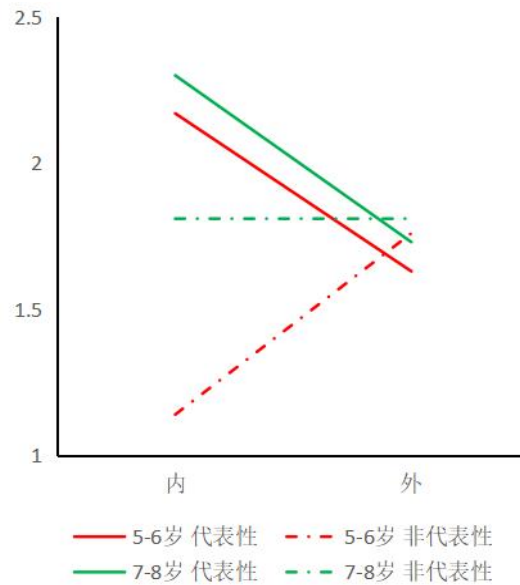


图 4 不同年龄儿童在不同代表性情境下对内外群体的惩罚数量

审稿意见 12：实验三的结果部分，建议作者更清晰地说明所使用的群体认同感量表是否为标准化工具，若为标准量表应当引用文献，若为自编量表应详细问卷题目。动机判断采用口头询问方式是否存在实验者效应，比如可能促使儿童选择社会期望性的“结果主义”选项。实验三的讨论中提到“群体认同感在其中起调节作用”，但未见相应的调节或中介分析，建议补充相应的统计验证。

回复：感谢审稿专家的细致建议。为我们完善研究提供了关键指引，针对实验三的相关问题，我们已认真修改，具体说明如下：

（1）本研究中群体认同感测量的量表是基于 Tajfel 与 Turner 在其著作中提出的群体认同感量表(Tajfel & Turner, 2004)，通过翻译与文化适配来确保题目表述符合儿童语言理解能力及本研究场景，形成中文版量表，用于评估儿童的群体认同感。具体题目包括：“能在这里上学我很高兴”“能在这里上学我很光荣”“能在这里上学我很骄傲”“我很喜欢我的学校”“我很喜欢我的同学”，已在研究三“实验程序”部分补充文献引用以及完整题目内容。

（2）关于动机判断采用口头询问方式可能引发实验者效应这一问题，我们进行了深入的思考，认为在本研究中，这种方式的潜在影响是可控的，具体理由如下：①我们采用标准化语句，分维度询问两种动机，含控制变量题项，实验者仅朗读题干，无额外引导，限制主观干扰；另外我们并非单纯依赖儿童的口头回答来判断动机，而是结合儿童的惩罚行为强度、道德判断等多方面信息进行综合考量，最大程度减少因口头询问可能带来的偏差；②本研究

的被试是 5-8 岁的儿童，这一年龄阶段的儿童虽然已经开始有一定的社会期望意识，但他们在回答问题时往往更倾向于表达自己内心的真实想法，这在一定程度上降低了选择社会期望性选项的可能性。我们的实验是在较为轻松、自然的环境中进行的，通过使用布偶等道具营造出一个贴近儿童日常生活的故事场景，让儿童能够放松地表达自己的观点；③从收集到的数据结果来看，儿童给出的惩罚动机是多样的，不仅有基于因果动机，还有基于报应动机以及其他类型的动机。且不同群体的动机占比存在显著差异（外群体基于因果的动机显著低于内群体），具体差异已在结果部分呈现，这表明儿童的回答并非受到实验者效应的干扰，而是基于自身真实动机做出判断，非单纯迎合社会期望。综上，尽管口头询问方式确实存在引发实验者效应的潜在风险，但通过研究设计的优化、对被试特点的考虑以及数据结果的验证，我们认为在本研究中这种方式是可行的，能够较为准确地获取儿童的惩罚动机。

（3）关于群体认同感这一变量，我们研究设计的初衷是想以群体认同感作为额外变量，来进一步探究儿童对待内群体惩罚严厉、外群体惩罚轻微的动机。在研究三中，首先运用群体认同感量表获取儿童群体认同感数据，借助 SPSS 软件，对群体认同感得分进行描述统计，得出其平均值为 22.23、标准偏差为 2.89，以此呈现该变量在样本中的整体分布；后续为探究其作用，将群体认同感得分与儿童信任内/外群体违规者变好的选择进行分析，对选择信任不同群体违规者变好的被试的群体认同感得分进行单样本 t 检验，发现群体认同感越高的儿童，越倾向于信任内群体成员改正错误，以此明确群体认同感对儿童信任决策的影响，现已针对文中不准确表述（如错误提及“群体认同感在其中起调节作用”）进行全面修正，重新梳理群体认同感与惩罚动机、行为的关联分析，确保表述贴合研究设计本意。后续我们也会强化写作规范意识，对论文语言表达进行多轮校验，避免类似疏漏。再次感谢您的专业指导，若修正内容仍有不足，恳请您继续指正，我们会全力完善！

详见修改稿 P11 “4.1 目的”、P12 “4.4 实验程序”、P13 “4.6 小结” 部分，摘录如下：

4.1 目的

通过实验法检验动机类型，探究 7-8 岁儿童对待内群体严厉、外群体轻微惩罚的动机，并且引入群体认同感作为额外变量，考察群体认同感的影响。

4.4 实验程序

准备步骤及群体身份启动步骤与研究一相同。采用个别施测的方式，在相对独立的环境下对每个儿童进行操纵。群体身份启动成功后，测量儿童的群体认同感，采用标准化群体认同感量表(Tajfel & Turner, 2004)，翻译成中文后题目如下：“能在这里上学我很高兴” “能

在这里上学我很光荣”“能在这里上学我很骄傲”“我很喜欢我的学校”“我很喜欢我的同学”；

4.6 小结

研究三表明，7-8 岁儿童严惩内群体违规者的行为主要是基于因果（希望违规者改正），而非基于报应（单纯的报复）。同时，高群体认同感儿童更倾向于通过惩罚内群体维护群体规范，而低认同感儿童则无此倾向。

审稿专家二

审稿意见 1: 作者论文题目为 5~8 岁儿童道德判断中的黑羊效应。在研究中实验设计也围绕 5~8 岁儿童进行年龄的划分,但是在研究的综述当中,已有研究证据主要聚焦为 3~6 岁及 6~8 岁儿童的相关研究,对 5~8 岁儿童的道德判断缺乏系统性的综述。且在实验设计当中,实验 1 及实验 2 对于自变量年龄进行划分,然而在研究的综述中并未就 5~6 岁及 7~8 岁年龄段道德判断差异进行系统的梳理,对年龄的划分已经论证不够充分。因此作者对于所研究内容的相关综述不够全面,三个实验的逻辑关系需要进一步梳理。

回复: 十分感谢您提出的宝贵意见和建议。针对综述中 5-8 岁儿童道德判断相关研究梳理不足及年龄划分论证不充分的问题,我们将补充 5-6 岁与 7-8 岁儿童道德判断差异的系统性综述,以充分论证年龄划分的合理性。另外我们也在引言部分进一步梳理三个实验的逻辑关系,使论文结构更加清晰,具体增加和修改内容如下,请您评阅。

详见修改稿 P1-3 “引言”部分,摘录如下:

1 引言

在个体的道德判断过程中,群体身份认知会产生显著影响,进而引发两种不同的现象。第一种是内群体偏爱(In-group Favoritism),指个体在态度和行为等方面表现出来的对自身群体成员的偏爱。例如 1-2 岁幼儿当看到内群体受到伤害时,会更期望旁观者去惩罚施害者,而当内群体伤害他人时,则会期待旁观者能给受害者提供一些帮助(Ting et al., 2019); 3-6 岁儿童会判定外群体成员踢人比内群体踢人影响更加恶劣,更应受罚(Chapman et al., 2020);与群体内违规者相比,6 岁儿童更有可能惩罚自私的外群体违规者(Jordan et al., 2014)。与之对应的是“黑羊效应”(Black Sheep Effect),即个体也可能会对内群体成员更加严厉和苛责。研究表明,在第三方惩罚任务中,5-6 岁女童更倾向于惩罚内群体违规者,而男童会同等程度地惩罚内外群体违规者(Wu & Gao, 2018); 43%的儿童在目睹道德违规行为后选择惩罚群体内违规者,仅有 17%的儿童惩罚外群体成员(Yudkin et al., 2020)。内群体偏爱与“黑羊效应”的存在均具备合理性,社会认同理论指出,个体往往将内群体成员赋予积极的品质认知,这种认知偏差使得人们对群体内部成员普遍表现出包容与偏爱倾向(Tajfel & Turner, 2004);但主观群体动力理论表明,人们存在限制和控制违规者以维持群体规范的共识,当发现内群体成员出现违反群体规范的行为时,个体可能会对内群体违规者更严厉,即出现“黑羊效应”(Abrams et al., 2003)。这种从“偏爱”到“严惩”的转变因何发生?群体身份如何作用于儿童的惩罚行为?仍需进一步探究。

社会领域理论将个体的道德发展置于具体的社会规则情境下理解,社会情境包括道德领域、社会习俗领域和个人领域(Killen et al., 2013),儿童在 3 岁时就表现出对道德规则以及其它规则的认同和偏好(Nucci, 1981),并且对道德违规判断更严格(Yau & Smetana, 2003)。同时,儿童道德判断受群体身份影响呈现年龄差异:5-11 岁儿童随年龄增长更排斥内群体违规者(Abrams et al., 2003);这一现象也符合儿童的认知发展轨迹,7 岁前的儿童以群体为中心,容易对内群体产生更积极的评价,7 岁后认知过程占据主导,更多的考虑个体品质,而不是其群体身份(Aboud, 1988)。值得注意的是, Abrams 等人的研究发现“黑羊效应”只出现在 8 岁及以上的儿童身上,未出现在 6 岁及以下的儿童身上(Abrams et al., 2014)。综上本研究选取 5-8 岁儿童,并将儿童划分为 5-6, 7-8 两组来考察儿童的道德判断特点,从而为理解该年龄段儿童道德判断的发展轨迹提供更细化的实证支持。

在道德违规情境中,个体对群体违规者的评价与惩罚会受到群体身份和重复违规行为的共同影响(Gollwitzer & Keller, 2010; Ting & Baillargeon, 2021)。例如,在一项以隐藏公共图书为违规场景的研究中,对比内外群体成员的初次与重复违规行为,发现内群体成员重复违规会引发被试更强的愤怒与更重的惩罚,而外群体成员的重复违规无此现象(Gollwitzer & Keller, 2010)。让儿童观看布偶伤人视频,结果显示儿童会更倾向将重复多次伤人的布偶判定为坏人,并更支持对其惩罚(Langenhoff et al., 2022)。还有研究进一步发现,儿童对违规行为的严重程度判断受违规频次的影响,相比单次伤害,三次连续伤害更易被幼儿判定为严重违规,且会降低对违规者的道德评价,同时该现象受群体身份影响,布偶伤害内群体成员时,仅需单次伤害行为幼儿即推测布偶后续会进行不公平分配,而布偶需多次伤害外群体成员,幼儿才会做出类似推测(Ting & Baillargeon, 2021)。然而现有研究仍存不足:对于重复违规者,其内群体身份会导致更严厉的惩罚还是更宽容的对待?面对外群体成员多次重复违规时,儿童对他们道德判断和惩罚行为是否会一致?因此,研究一聚焦重复性违规情境,将“重复性违规”操作化为同一主体在实验情境中连续三次违反同一道德规范,通过控制违规次数与违规者群体身份的差异,系统考察儿童对内外群体成员违规行为的道德判断差异。

除重复违规外,违规者的群体代表性同样会影响个体对道德违规行为的判断(Bernard et al., 2016; Gollwitzer & Keller, 2010)。已有研究初步揭示儿童对群体代表性的认知能力:2 岁左右幼儿观察到布偶违反公平规范时,会期望有代表性的领导者干预纠正,而对非领导者干预无明确期待(Stavans & Baillargeon, 2019),表明儿童此时已能初步区分有影响力者与无影响力者,并认为前者有责任维系社会规范。3-5 岁幼儿在物品寻找与新异物品命名任务中,会优先信任更具影响力的证言提供者(Bernard et al., 2016)。但现有研究尚未将群体代表性与

道德违规情景结合，儿童对群体成员道德违规行为的判断，是否会因违规者的代表性身份不同而产生差异？尤其对作为“群体价值与规范体现者”的核心成员(Abrams et al., 2014; Fousiani et al., 2019)，其道德违规是会引发儿童的内群体偏爱还是“黑羊效应”？据此，研究二聚焦代表性违规情境，将群体核心成员界定为充分体现群体核心价值观、具有高影响力和示范效应，是群体内广受欢迎的榜样且被普遍模仿的成员(Abrams et al., 2003, 2014)，系统探讨儿童对内外群体代表性成员与非代表性成员道德违规行为的判断与惩罚特点，为儿童道德判断领域的研究提供实证补充。

同时，过去的惩罚动机研究表明，成人与幼儿惩罚道德违规兼具两种动机(Carlsmith et al., 2002; Marshall et al., 2020, 2022)，一种是基于因果，惩罚是为了让犯错者改正错误，变得更好；一种是基于报应，惩罚是因为个体做错了，就应该遭受痛苦，受到相应的报应，成人对内群体违规者惩罚更偏功利，主要是为了维护群体利益，而非单纯的报应(Fousiani et al., 2019)，但儿童对内外群体成员惩罚动机的差异，尚无研究涉及。因此，在研究一二的基础上，进一步开展研究三探讨在特定的道德违规情境下，儿童惩罚内外群体违规者的惩罚动机。

综上所述，目前关于内群体偏爱的机制已较清晰，但关于“黑羊效应”在儿童道德发展领域上的研究仍存局限。随着年龄的增长，儿童的道德判断究竟是依赖于群体规范，还是越来越趋近于公平？为何个体在面对内群体违规者时存在内群体偏爱和“黑羊效应”两种现象？其中的影响因素（如重复违规、成员代表性）以及该现象的潜在动机需要实证研究去证明，目前对“黑羊效应”的研究大多聚焦于成人群体，而以儿童群体为调查对象的研究较少。因此本研究聚焦 5-8 岁儿童，通过三项实验考察其在特定情境中对群体成员违反道德规范行为的道德判断、惩罚特点及内在动机。旨在验证并拓展主观群体动力发展理论在儿童群体中的适用性。同时，深入剖析“黑羊效应”的影响因素及其背后的心理机制，以此揭示儿童视角下道德规范与群体身份认同的内在关联。研究结果可为道德教育提供理论支撑，也为家校教育实践提供科学参考，助力儿童道德认知与社会适应的健康发展。

审稿意见 2: 研究 2 探究 5~8 岁儿童对具有代表性和非代表性特征的内外群体成员违规行为的道德行为及惩罚行为在实验设计当中，对群体成员代表性及非代表性成员的操作定义缺乏理论依据，代表性成员及非代表性成员是否需要通过的预实验确定其有效性。

回复：非常感谢您的宝贵建议。针对您的建议，我们主要作出以下两点回复

（1）研究 2 中对代表性成员和非代表性成员的操作定义基于主观群体动力理论和社

会认同理论的经典研究，具体理论依据如下：

①代表性成员的理论基础：根据 Abrams 等人（2003, 2014）的主观群体动力理论，群体中的代表性成员是指“充分整合了群体的核心价值观和信念，以及理解群体规范重要性的人”，其行为对群体规范具有更强的示范效应，其违规行为也被认为是对群体认同的更大威胁。

Fousiani 等人（2019）进一步指出，代表性成员的核心特征包括“高群体影响力”“被广泛模仿”“与群体规范高度一致”，这与本研究中“带头参与活动、成为榜样、被模仿”的操作定义直接呼应。

②非代表性成员的理论基础：非代表性成员的定义参考了 McGuire 等人（2020）对“群体边缘成员”的描述，即“较少参与群体活动、对群体规范影响力弱、不被其他成员关注或模仿的个体”，其行为对群体认同的影响较小。这与本研究中“不太受关注、不参与活动、不被模仿”的操作定义一致，符合“群体边缘性”的理论内涵。

这一部分内容已整合到原文中，使读者对于群体成员代表性及非代表性成员有更清楚的了解。

（2）另外您指出的关于群体成员代表性及非代表性操作性定义的理论依据和有效性验证问题，对于提升研究严谨性具有重要意义。本研究在实验程序中采取了相应措施以保障操纵的有效性：一方面，实验中设置了严格的记忆核查问题（如询问儿童“谁是你的同学”“你觉得他们会不会带头捐铅笔呢”等），所有儿童均能正确回答，表明他们已准确接收并理解了代表性与非代表性成员的特征描述；另一方面，从结果来看，儿童对代表性与非代表性成员的道德判断和惩罚行为存在显著差异（如在代表性条件下，各年龄组对内群体违规者的惩罚显著高于外群体，非代表性条件下则呈现不同趋势），这些结果共同支持了本研究对群体成员代表性及非代表性操纵的有效性。具体增加和修改内容如下，请您评阅。

详见修改稿 P8-9 “3.4 实验程序”，摘录如下：

3.4 实验程序

准备步骤及群体身份启动步骤同研究一。群体身份启动后，对部分儿童进行群体成员身份代表性启动，对群体成员代表性的界定为：代表性成员指充分体现群体核心价值观、具有高影响力和示范效应的成员(Abrams et al., 2003, 2014)，非代表性成员参考群体边缘成员特征，即参与度低、影响力弱(McGuire et al., 2020)。通过实验者发布指导语的方式进行启动，情境为描述布偶在学校的受欢迎程度和带头作用来确立其代表性或非代表性身份，共有四种身份形象：代表性情境-内群体代表/外群体代表、非代表性情境-内群体非代表/外群体非代表，例如代表性情境-内群体代表指导语为“在你的学校非常受欢迎，班上活动都是他/她带头做起的，是你们的榜样，大家都喜欢模仿他/她的行为”，非代表性情境-外群体非代表指导语

为“在一个与你无关的学校不太受到关注，班上的活动他/她也不怎么参与，是班级里非常普通的一位同学，大家都不会去模仿他/她的行为”；身份形象确立后，询问儿童“最近学校要组织给贫困山区捐赠铅笔，你觉得他们会不会带头捐呢？”；接着介绍捐赠铅笔的故事情境，播放视频，视频内容同研究一违规情境中的片段；结束后询问儿童记忆问题、认知评估问题，并让儿童进行惩罚决策（同研究1）。

审稿意见3：研究在实验一、实验二及实验三都涉及内外群体的划分，但对内外群体划分，样本的有效性并未做进一步检验，且研究者在研究的过程中运用了过多的相关研究概念如因变量结果主义惩罚、报应主义惩罚动机等，概念虽然来自于西方研究但翻译较为生硬，缺乏与本土化结合后的基于中国文化的概念命名及解释，因此论文需要进一步修改。

回复：感谢审稿人提出的重要意见，这些建议对于提升研究的严谨性和文化适应性具有重要意义。针对您指出的问题，我们做如下说明和修改：

（1）关于“内外群体划分的样本有效性检验”，本研究在群体身份启动后，通过询问内群体偏爱问题来检验启动效果并筛选有效样本，验证群体划分的有效性。问题如下：“谁是你的同学？你会更喜欢哪一个人？你手中有一张贴纸，你想把它送给谁？”核心逻辑是若群体划分有效，儿童应更倾向偏爱内群体，此类样本能更精准反映内外群体身份对道德判断的实际影响。未能展现出群体偏好的被试数据已被剔除，不进行下一步的研究（已在各实验“被试”部分注明剔除数量及原因），绝大多数的被试展现出了对内外群体的不同偏好，在研究一和研究二中有效被试出现内群体偏爱的儿童占比分别为84.4%和85.5%，并且具有统计显著性（卡方检验 $P<.001$ ），非随机误差导致。因此内外群体的划分在样本中是具有有效性的。

（2）感谢审稿人细致指出概念本土化的问题！您提到“西方研究概念翻译生硬、缺乏本土化结合”的建议非常关键，我们深刻认同，由于国内儿童惩罚动机的本土化研究较少，相关概念的本土适配表述与系统阐释较少，所以翻译稍显生硬，这确实是本研究的不足，我们已调整表述，将结果主义改为“基于因果”，报应主义改为“基于报应”，使概念更适配中国本土化语境。“基于因果”指儿童在道德判断中，认为惩罚不只是对错误的回应，更要让犯错者“因罚改错”——通过惩罚让对方明白“错误行为会带来负面后果”，进而纠正行为、变得更好，这一逻辑贴合本土“有错能改，善莫大焉”的教育观。“基于报应”指儿童认为做错事就要付出代价，认为惩罚是对错误行为的直接回应，这一逻辑呼应本土“自作自受，因果报应”的朴素公平观，反映儿童道德判断中“维护基本公平秩序”的本能需求，

是早期以罚惩恶正义感的直观体现，使概念更适配中国教育文化语境与儿童道德认知逻辑。
原文中出现全部的“结果主义”已调整为“基于因果”，“报应主义”已调整为“基于报应”

详见修改稿 P3 “引言”第五段、P11 “4.2 假设”、P12 “实验程序”、P13 “4.5 结果与分析” “4.6 小结”、P16 “6 结论”，摘录如下：

1 引言

同时,过去的惩罚动机研究表明,成人与幼儿惩罚道德违规兼具两种动机(Carlsmith et al., 2002; Marshall et al., 2020, 2022),一种是基于因果,惩罚是为了让犯错者改正错误,变得更好;一种是基于报应,惩罚是因为个体做错了,就应该遭受痛苦,受到相应的报应,成人对内群体违规者惩罚更偏功利,主要是为了维护群体利益,而非单纯的报应(Fousiani et al., 2019),但儿童对内外群体成员惩罚动机的差异,尚无研究涉及。因此,在研究一二的基础上,进一步开展研究三探讨在特定的道德违规情境下,儿童惩罚内外群体违规者的惩罚动机。

4.2 假设

假设 1: 儿童对待内群体的惩罚不仅仅是因为错误而惩罚,惩罚更重要的目的是为了
让犯错者改正错误,变得更好,惩罚动机是**基于因果**;而对待外群体的错误,只是单纯为了惩
罚而惩罚,惩罚动机是**基于报应**。

4.4 实验程序

……实验者口诉惩罚动机,惩罚动机引用 Marshall 等人(Marshall et al., 2022),其中包括
基于因果动机(例如:“我希望我们/其他学校的犯错者不会再去偷捐赠物品”)和**基于报
应动机**(例如“我觉得我们/其他学校的犯错者就应该受到这样的处罚”),让儿童做出认
可度判断,采用直接提问的方式,“惩罚她/他是想让他改正? (**基于因果**)还是因为他该
受罚? (**基于报应**)”……

4.5 结果与分析

表 5 群体认同感与惩罚动机的平均数和标准差 ($M\pm SD$)

	基于因果(内)	基于因果(外)	基于报应(内)	基于报应(外)	群体认同感
$M\pm SD$	5.46 (1.78)	4.56 (1.78)	2.00 (1.47)	1.67 (1.54)	22.23 (2.89)

4.6 小结

研究三表明,7-8 岁儿童严惩内群体违规者的行为主要是**基于因果**(希望违规者改正),
而非**基于报应**(单纯的报复)。同时,高群体认同感儿童更倾向于通过惩罚内群体维护群体
规范,而低认同感儿童则无此倾向。

6 结论

（3）儿童对内群体和外群体的惩罚动机并不相同，具体而言，儿童严惩内群体违规者，其惩罚动机主要是**基于因果**，即希望通过本次惩罚，违规者能够吸取教训，以后能够变得更好。对于内外群体而言，儿童都较少地因为**报应**而产生惩罚行为。

再次向编辑老师与各位审稿专家致以最诚挚的感谢！您提出的每一条宝贵指导意见，都精准地指出了稿件的完善方向，为我们理清研究逻辑、提升内容质量提供了关键支撑。在逐字逐句修改的过程中，我们不仅打磨了稿件的细节，更深化了对研究问题的思考，正是这些细致且专业的建议，让本研究的科学性与严谨性得到了切实提升。目前我们已根据意见完成全部修改，衷心希望修改后的稿件能符合期刊要求，也期待未来能继续得到您的指导与指正。