

审稿意见及修改说明

非常感谢两位审稿专家对《不同社会情境对道德判断的影响》(编号: psysci2024-017)做出评审, 专家们提出的意见和建议对论文的修改非常有帮助。我们已经对评审意见逐一做出了回复, 并按照修改意见对文章进行了详细修改, 修改部分在文章中标记为蓝色。关于审稿意见的详尽回复如下:

审稿专家一意见:

意见 1: 文章标题中“不同社会情境”的表述过于模糊。

回复: 感谢审稿专家提出的宝贵意见。本研究中关注的实验情境无论是受他人观察还是双方共同参与, 都涉及道德决策的公开性, 只是公开程度和信息交流深度有所不同。情境公开会影响个体行为的匿名性, 在公开情境下, 个体会主动印象管理, 需要自我控制, 导致道德决策倾向出现变化(杜秀莲, 魏巍, 2021)。因此, 我们将“社会情境”修改为“公开情境”, 使其更加符合研究目的。同时, 综合考虑两位审稿专家的意见, 为了明确道德判断的具体指向并在标题中体现主要研究结果, 我们将文章标题修改为“不出手的道德? 公开情境对道德两难决策的影响”(见 p1、p9)。

此外, 我们将文章的关键词, 以及文中部分关于社会情境的表述内容修改为公开情境。

意见 2: 摘要部分缺少对研究问题的背景介绍, 也缺少对实验结果理论/实践意义的总结。因此, 很难从摘要部分了解本研究的意义和价值。

回复: 非常感谢审稿专家提出的意见。我们对摘要部分进行了全面的修订, 增加了关于研究背景的介绍, 并且修改了研究结果的表述。具体修改内容如下(见 p1):

“以往研究多集中于使用经典道德困境对单被试的决策进行探讨。本研究采用能够量化决策者结果敏感性、道德规范敏感性以及一般行动倾向的 CNI(consequence, norm, inaction)模型, 通过两个实验考察公开情境对道德两难决策的影响。实验一使用社会观察范式, 发现他人观察并不影响被试的道德决策。实验二使用双人平行判断范式和共同判断范式。结果显示, 相比于单独判断, 平行判断和共同判断导致被试出现明显的不行动倾向, 而且共同判断组对于道德规范的敏感性增加, 这两种情境下个体对道德困境中主体的行为做出更多的“不接受”反应。研究表明, 公开情境下个体看似道义的决策可能源于营造良好社会形象的需要, 以及避免承担行动责任的动机。”

意见 3: “本研究采用道德判断的 CNI 模型”摘要部分这句话, 英文没有给出全称, 也让不了解相关背景的读者无法理解。

回复: 感谢审稿专家提出的宝贵意见。CNI 模型概念中的三个字母分别代表个体在道德判断中三种不同的决策倾向: 结果敏感性(sensitivity to consequences, C), 道德规范敏感性(sensitivity to moral norms, N)和一般性的行动或不行动倾向(general preference for inaction versus action, I)(Gawronski et al., 2017), 参考云祥(2020; 2023)的研究, 将 CNI 模型补充书写为“CNI(consequence, norm, inaction)模型”。因此, 我们将摘要

中“本研究采用道德判断的 CNI 模型”改为“研究采用能够量化决策者结果敏感性、道德规范敏感性以及一般行动倾向的 CNI(consequence, norm, inaction)模型”(见 p1)。

意见 4:“基于多项式建模的道德判断的 CNI 模型可以解决上述问题”，CNI 在引言部分第一次出现时，也没有提供全称。

回复:感谢审稿专家提出这一问题。我们已将引言中首次出现 CNI 模型的部分进行修改，补充了全称。修改如下(见 p3):“基于多项式加工树模型(multinomial processing tree, MPT)建构的道德判断的 CNI(consequence, norm, inaction)模型可以弥补上述局限(Gawronski et al., 2017)。”

意见 5:因为作者对 CNI 模型的介绍不够充分，尤其是该模型与本研究的关系，因此作出以下假设时，很难理解这个假设的依据是源于怎样的发现或者理论推导：“本研究假设，相比于单独判断，在公开或多人判断情境下，个体会做出更多的道义决策，这一方面可能源于被试道德规范敏感性的增加(N 参数)，另一方面也可能是较强的不行动偏好(I 参数)导致的。”请作者结合自己的实验范式，详细论证 CNI 模型相较于双加工理论的优势，以及作者如何得到关于 CNI 中具体参数的假设。

回复:非常感谢审稿专家提出的问题，我们将对这两个问题进行详细阐述。首先需要说明的是，由于投稿文章篇幅相对有限，本文初稿并未对 CNI 模型进行充分介绍，可能会使读者对该模型产生一定的误解。事实上，CNI 模型并非与双加工理论对立，而是对使用经典道德困境范式的研究结果进行了更深入和细致的分析解释。一些研究将道义倾向和功利倾向视为非此即彼的对立关系，但双加工理论认为道义主义和功利主义是两个独立的加工过程，二者同时对道德判断起作用，占优势的一方产生决策结果，劣势的一方仅表现为神经水平上的激活(Greene et al., 2008)。CNI 模型在此基础上试图将道德敏感性和结果敏感性进行量化，同时分离出一般的不行动倾向。因此，接下来要讨论的是 CNI 模型相较于经典道德困境范式的优势。

意见 5-1, CNI 模型的优势:自双加工理论提出后，许多研究利用经典道德两难困境在该理论框架下开展研究。但却逐渐出现了一些不符合常理的结果，如精神病性被试的功利主义倾向并不能理解为这些被试具有权衡利弊的道德理性(Bartels & Pizarro, 2011)。研究者对其产生的质疑主要包括以下几个方面：

首先，基于经典道德困境的研究认为决策者接受功利性选择是受功利主义原则驱动，不接受功利性选择则是受道义主义原则驱动，将其视为对立的两极，其加工纯粹性是最大的局限所在(Conway & Gawronski, 2013)。个体在面对道德选择时，完全有可能综合考虑行为背后的社会规范以及行为带来的后果。神经研究证据表明，功利和情感的评估是独立并行进行的(Hutcherson et al., 2015)。所以功利主义和道义主义的反应倾向完全是可以相互独立，并行不悖的(刘传军, 廖江群, 2021)。

其次，经典道德困境无法对道义主义和功利主义的倾向程度进行量化。例如，当两种倾向程度差异较小时，被试可能需要更长的时间做出决策，但若两种倾向程度相当，在二元解释的框架下，决策结果不一定能够反映个体的真实想法。此外，经典两难困境中大多集中于对禁止性规范的考虑，涉及的行为都是道

德规范所不允许的，并未涉及道德规范允许的情况。所以，需要对行为结果的利弊大小，以及行为是否被道德规范所允许进行控制，才有可能分辨出决策者两种倾向的程度分别是怎样的。

最后，基于经典道德困境的研究在解释上存在模糊性，并且忽略了对一般反应倾向(行为惯性)的考虑。例如，在电车难题中如果同意“杀一救多”，既可以解释为功利主义增强的作用，也可以说是道义主义弱化的结果，甚至可能是不考虑结果如何，也不考虑道德规范，只是单纯地偏好行动造成的(比如精神病态被试的冒险倾向使其不断做出行动的决策)。所以，需要将一般的行动偏好纳入决策过程的考虑之中。

CNI 模型对决策结果的来源进行了分离，能够同时量化决策者的结果敏感性，道德规范敏感性以及一般的行动或不行动倾向。该模型设计了 12 个故事情境，每个故事情境又分为四种不同的版本，这四种版本在情境上对道德规范是否允许，以及行动结果利弊大小进行了操纵。通过多项式决策加工树建模，根据决策者的反应概率对不同的心理加工过程进行分离，从而实现对三种反应倾向程度的量化。此外，以往研究在使用经典道德困境对公开情境下的道德决策进行研究时，大多把个体的道义决策归结于声誉担忧和从众压力(Rom & Conway, 2018; Bostyn & Roets, 2016)，但是却缺少明确的数据支持，也无法从根本上说明反应倾向的变化。利用 CNI 模型对三种反应倾向进行量化，能够对个体决策时所关注的内容做出更加清晰的解释。

同时，需要说明的是，前人研究除了使用经典道德困境范式的局限外，在公开情境的设计上也较为有限。Lee 等人(2018)采用社会观察范式，两名观察者对被试的决策进行直接监控。Chen 等人(2020)让真被试与假被试共同进行道德判断，真被试在 fMRI 设备中，实验开始便无法看到另一人且不能获知对方的决策。Andersson 等(2020)将陌生被试随机组成 4 到 8 人的小组，让他们完成公共物品博弈、独裁者博弈和道德两难判断(仅判断 4 个困境)，处于公开情境下的被试做完决定后会在大屏幕显示他们的照片、姓名和决策。这些研究中无论决策人数如何变化，都仅仅是暴露决策结果而已，较少涉及被试间的信息传递，与真实社会生活中的情境仍有差距，所以在结论的推广性上存在不足。因此，本研究采用三种不同的实验情境设置，在不同程度的信息交流状态下探究个体道德规范敏感性、结果敏感性和不行动倾向三种决策反应倾向的变化。

综上，我们对 CNI 模型相较经典道德困境的具体优势在引言中作了修订和补充(见 p2-p3)。

意见 5-2，CNI 参数假设的得出：许多社会决策方面的研究表明，相比于单独或匿名情境，当有他人在场时个体的决策倾向会发生改变，这种变化体现在道德问题是复杂的，对 C、N、I 三参数的影响方向也是不同的，具体阐述如下：

以往研究发现，当处于第三方在场的公开情境时，人们参与志愿服务的热情有所提升(Bereczkei et al., 2010)、并且能够促进利他规范的执行(Kurzban et al., 2007)以及合作的经济决策(Egas & Riedl, 2008)。当个体在所处团体中暴露身份信息和捐赠行为时，被试的捐赠金额有着显著的增加(Andreoni & Petrie, 2004)。

这些结果都说明，公开或多人决策的情境下个体会产生更多符合社会规范的行为表现，即更遵守一些被普遍接受的、有关群体成员在给定情境下应当如何行为的规范(Fehr & Fischbacher, 2004)，并且更愿意表现出符合社会规范的形象(Anderson & Kilduff, 2009)。因此，本研究假设公开情境会提高个体对道德规范的关注，即表现出较高的 N 参数。

虽然多人情境下个体存在明显的利他倾向，却有相当一部分研究发现，在有其他人在场的情况下，个体的助人倾向显著下降(Fischer et al., 2011; Zoccola et al., 2011; Hortensius et al., 2016)，即存在“旁观者效应”。面对帮助情境时，个体会将自身行为产生后果所致的责任分散到其所在的群体中，从而减轻个体的责任感和压力(Bandura, 2002)。即便是面对道德规范所提倡的行为，由于责任的分散，愧疚感的减弱，决策者未必会出手相救。因此，我们假设公开情境会使个体在道德判断中倾向于不行动，即表现出较高的 I 参数。

另外，我们还补充了关于结果敏感性(C 参数)的假设。研究表明，个体会存在潜在的受他人评价的声誉担忧(Anderson & Kilduff, 2009)。人们有着不在公共场合展示功利主义道德决策的偏好，毕竟道义主义和功利主义的判断可能会导致截然不同的社会印象。道义的决定象征着温暖，知觉者通常认为拒绝伤害的决策者比接受伤害的更值得信任(Everett et al., 2016)、更有同理心(Uhlmann et al., 2013)；而功利主义决策由于关系到一个人追求有效目标最大化的动机，会引起人们对能力和自利倾向的看法(Lee et al, 2014; Rom et al, 2017)。个体知觉的大量研究揭示了社交中温暖和能力的作用并不对称，前者是印象形成和人际喜好的主要决定因素(Fiske et al., 2007; Lee et al., 2018)。因此，本研究假设公开情境会降低个体对行动结果利益的考虑，即表现出较低的 C 参数。

综上，我们对假设的具体提出过程在引言中做出了补充说明(见 p4)。

意见 6：“CNI 模型的研究中包括四种版本的故事：道德规范禁止，行动利大于弊(简称“禁止利大”)、道德规范禁止，行动弊大于利(简称“禁止弊大”)、道德规范提倡，行动利大于弊(简称“提倡利大”)、道德规范提倡，行动弊大于利(简称“提倡弊大”)。”请给予具体的说明。例如，在附录中，1a 这个例子，不是很好理解为什么这个是一个“道德规范禁止行动，行动的利益大于成本”的例子。

回复：非常感谢审稿专家提出的问题。CNI 模型需要对故事情境中的道德规范和利弊大小进行操纵。模型设计了 12 个基本的道德两难困境，每种困境有四个不同版本的变式，共形成 48 个道德两难困境。需要说明的是，对于道德规范的控制是指故事中主人公的行动(具体指故事最后一句话描述的行为决定)是否符合规范，以故事 1 绑架困境为例，在 1a 和 1b 中“否决支付赎金”是个体需要决策的行动，如果否决支付，那么游击组织就会将战地记者斩首，行为直接导致一人丧生，因此属于“道德规范禁止”的背景。在 1c 和 1d 中“批准支付赎金”是个体需要决策的行动，这时可以拯救战地记者的生命，属于“道德规范提倡”的背景。

关于行动结果利弊的大小则涉及故事后半部分，1a 中“否决支付赎金”带来的是战地记者一人的死亡，如果支付赎金就会协助该组织购买武器，这将“导致许多人死亡”，所以主人公不支付赎金的结果是丧失一人的生命，拯救更多人的生命，类似于“杀一救多”，是“行动利大于弊”的困境。事实上，在 CNI 模型的四类故事中，“禁止利大”的故事基本等同于经典道德两难困境，如：电车困境，都是规范不允许伤害但行动可以拯救更多人。其他三个版本故事中的利弊大小关系较好理解，就不再赘述。

我们在引言中对四种故事的区分进行了具体的说明，修改如下(见 p3):

“基于多项式加工树模型(multinomial processing tree, MPT)建构的道德判断的 CNI(consequence, norm, inaction)模型可以弥补上述局限(Gawronski et al., 2017)(见图 1)。该模型设计了 12 个故事情境，利用行动后果的功利化收益程度高低操纵结果(利大或弊大)，利用行为的直接后果操纵规范(不伤害生命或帮助弱者的规范)，每个情境生成四种版本：道德规范禁止，行动利大于弊(简称“禁止利大”)、道德规范禁止，行动弊大于利(简称“禁止弊大”)、道德规范提倡，行动利大于弊(简称“提倡利大”)、道德规范提倡，行动弊大于利(简称“提倡弊大”)。这样，无论结果利弊如何，在规范禁止时不行动，在规范允许时行动才能反映个体对道德规范的高关注。”

同时，为了方便读者阅读，我们在 2.1.2 材料部分列举了另一个较简单的例子便于理解，一并放在附录中(故事 2 免疫缺陷困境)，具体修改(见 p4-p5)。

意见 7：CNI 分析的具体流程、原理与目标需要给予更充分的说明。

回复：非常感谢审稿专家提出的宝贵意见。现对 CNI 模型的分析进行详细介绍。

CNI 模型为了区分决策者的道德规范敏感性、结果敏感性以及一般的行动倾向，对困境中的道德规范和行动结果利弊大小进行了操纵，使得每个故事背景下的困境都有四个变式。在 4 种困境中，接受和不接受的反应概率之和为 1。其中， C 参数表示对结果的敏感性， N 参数表示对道德规范的敏感性， I 参数表示在不考虑结果和规范时，个体的不行动偏好。模型用 4 条路径描述了由三种倾向驱动时行为选择的 4 种情况(见图 1)。当决策由结果驱动时，路径 C 的描述代表结果驱动行为选择的情况(图 1 第一行)；由道德规范驱动时，路径 $(1-C) \times N$ 代表结果无法驱动但规范驱动行为的情况(图 1 第二行)；当个体存在一般的不行动倾向时，路径 $(1-C) \times (1-N) \times I$ 代表了结果和道德规范都不驱动行为，个体存在不行动偏好的情况(图 1 第三行)；路径 $(1-C) \times (1-N) \times (1-I)$ 则描述了当个体存在一般的行动倾向时的情况(图 1 第四行)。

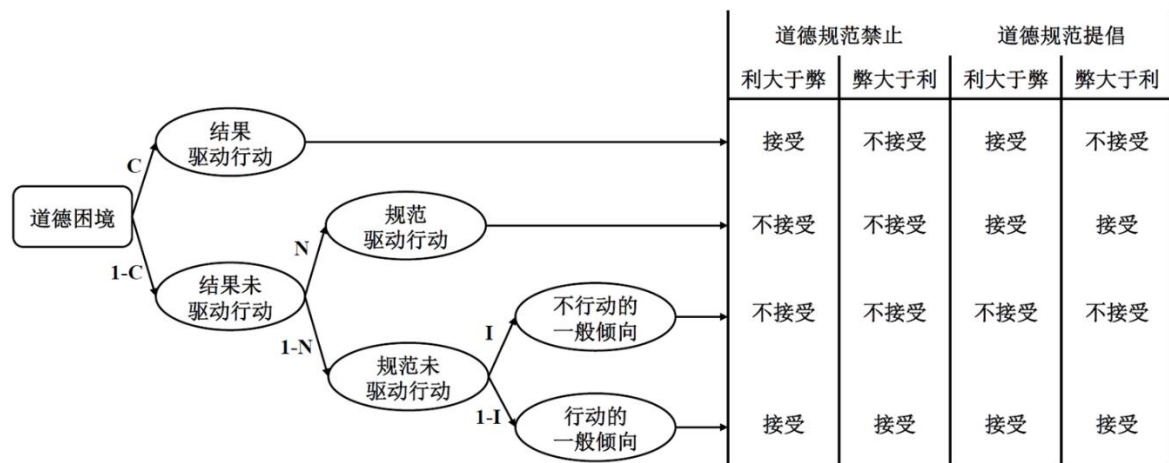


图 1 CNI 模型路径图解(资料来源: Gawronski et al., 2017)

通过这 4 条路径,可以推导出 4 种困境类型下选择接受或不接受概率的数学方程。因此,计算行为结果中 4 种故事类型下选择接受和不接受的反应率,代入方程可得三个未知数 C、N、I 的大小。使用最大似然估计法生成针对 3 个未知数的参数估计,利用拟合优度 G^2 评估模型在数据描述方面的充分性,观察到的概率和模型预测的概率之间的统计差异越小,模型拟合越好。所以,若 G^2 不显著($p>0.05$),表明数据与模型拟合良好;反之,模型拟合较差。Gawronski 等人(2017)提供了基于 multiTree 软件构建的多项式模型,在其中输入 8 个反应概率可得三参数的估计值(模型文件及材料见: <http://www.bertramgawronski.com/documents/CNI-ModelIndDiffMaterials.zip>)。对于 C 参数和 N 参数,估计值显著大于 0 时表明个体的决策受到结果敏感性或道德规范敏感性的驱动;对于 I 参数,估计值显著大于 0.5 时表明个体有一般的不行动的反应偏好,小于 0.5 时表明个体有一般的行动的反应偏好。

此外,可以通过假设两个变量的参数值相等,来比较两组参数之间的差异(如设置: $C_1=C_2$),若该操作使模型的拟合优度统计指标 G^2 显著性下降($p<0.05$),表明在该参数上两组存在显著差异;反之,表明两组之间在参数上不存在显著差异(Gawronski et al., 2017)。

综上,我们将 CNI 模型的基本逻辑补充在引言部分,同时对模型的适用性进行了更加详细的阐述,具体修改如下(见 p3):

“利用四种困境类型下的反应概率,模型可以推导出个体在所有困境中的决策情况,描述其对结果敏感的决策模式(图 1 第一行),对规范敏感的决策模式(图 1 第二行),以及不考虑结果和规范的一般的不行动或行动的决策模式(图 1 第三、四行)。以此对决策倾向进行了分离,实现对决策者结果敏感性(C 参数),道德规范敏感性(N 参数)以及一般的不行动或行动倾向(I 参数)的量化。由于社会化过程中道德规范的养成(Haidt, 2007),以及趋利避害的理性考虑,人们通常较少接受禁止弊大的行动,更容易接受提倡利大的行动,而对于规范和结果冲突的困境,个体在权衡之后做出的反应更能体现其决策倾向。目前,基于 CNI 模型的研究在情绪(Gawronski et al., 2018; 刘传军等, 2021)、权利感(Gawronski & Brannon, 2020; 云祥, 2020)、压力

(Zhang et al., 2018; Li et al., 2021)、人格特征(Luke & Gawronski, 2022)等因素对道德判断的影响中展现出良好的适用性,为理解道德决策的过程提供了新的视角和理论依据。”

此外,我们还在“2.1.4 CNI 模型分析”部分对 CNI 模型的具体分析方法进行了较为详细的阐述,具体修改如下(见 p5):

“基于多项式加工树模型原理,在构建好的 CNI 模型中根据被试的行为数据进行模型拟合,使用 mult iTree 软件(Moshagen, 2010)估计 3 个潜在心理过程(C 、 N 、 I)的概率。对于所有两难困境,将选择“不接受”赋值为 0,选择“接受”赋值为 1。用 G^2 表示数据与模型的拟合优度,若 G^2 不显著($p>0.05$),表明拟合良好。估计出的 C 和 N 参数若显著大于 0,表明被试有明显的结果和规范敏感性, I 参数若显著大于 0.5,表明被试存在一般的不行动倾向。在比较不同组别参数差异时,若 ΔG^2 显著($p<0.05$),说明两组被试参数存在显著差异(Gawronski et al., 2017)。”

意见 8: 实验一为什么与前人的发现不一致,需要更充分的说明。现在的解释说服力不够。

回复: 非常感谢审稿专家提出这一问题。我们再次搜集和梳理了有关社会观察范式下的相关文献,认为实验一未发现观察效应的原因除了观察员的观察作用和自然环境中社会观察的作用不同(Barnettler et al., 2012)之外,还有一个主要的原因可能是被试与观察者之间的社会联系较弱。Van Dijk 等人(2002)发现,共同参与公益实验的被试之间可以形成社会联系。个体对他人评价的关心程度与他们之间存在的社会联系密切相关,但是在实验室中的联系与现实生活中个体之间反复互动建立起的联系有着本质区别(Soetevent, 2005)。整个实验过程中被试和观察者之间不存在任何程度上的交流,且被试不会收到观察者对其决策的信息反馈,所以微弱的社会联系可能难以使被试对自我形象产生更多的关注。而且,人们在道德领域会对自己持有积极的看法(Epley & Dunning, 2000),并且能够以进行自我说服,使得任何一种决策合理化(Uhlmann et al., 2009),所以在没有反馈也无需考虑他人意见的情况下,被试的两难决策难以发生显著的变化。

一些相关研究还从文化差异和性别差异的角度解释类似的实验结果。Martinsson 等人(2013)发现决策暴露并不影响被试在公共物品实验中的贡献水平,认为可能是由于哥伦比亚和越南人集体主义更强,所以决策暴露的条件作用显得微乎其微。也许决策公开暴露对更具个人主义且更注重个人隐私的文化环境下的被试具有显著的影响作用,而中西方文化在这两方面存在一定差异。Sparks 和 Barclay(2015)采用眼睛图片作为监控线索,发现被试的道德判断并没有发生变化,认为一方面是由于法国和加拿大的文化环境差异,另一方面是由于女性被试过多。有研究指出眼睛图像的视觉线索对男性的慷慨程度有影响,但对女性没有影响(Rigdon, 2009)。视觉线索的作用在不同性别当中存在差异,本文实验一中女性被试略多于男性,这也可能是原因之一。

综上,我们对实验一结果可能的原因解释进行了补充,具体修改如下(见 p9):

“另一方面可能是被试与观察者之间的社会联系较弱,个体对他人评价的关心程度与他们之间的社会

联系密切相关，但是在实验室中的联系与现实生活中个体之间反复互动建立起的联系有着本质区别 (Soetevent, 2005)。Van Dijk 等人(2002)发现，共同参与公益实验的被试之间可以形成社会联系，然而实验一中的观察员和被试之间没有任何互动过程，因此难以产生社会联系，在没有反馈也无需考虑他人意见的情况下，被试的评价担忧较低，道德决策难以发生显著的变化。”

意见 9：本研究不同实验里，“不同社会情境”本质上操纵的是哪个潜在的变量，导致了不同的结果。请作者给出更明确的解释，不然本研究的发现也只是一些零散的发现，很难形成理论上的贡献。

回复：非常感谢审稿专家提出的宝贵意见。他人在场的道德判断实际上是一种社会决策，社会决策主要指在社会情境中个体行动之间相互影响进而产生决策的过程，在此过程中，个体的决策不仅要考虑问题本身，还要考虑参与决策的他人的行动，社会决策是一种不确定决策，包括决策问题带来的环境不确定性和其他决策者带来的社会不确定性(黄砾卉, 2010)。道德两难问题本身具有模糊性和不确定性，同时他人的存在又增加了一份不稳定性，衡量此种情况下的个体决策倾向的变化需要考虑更加复杂的因素。

人们通常会给自身所处情境赋予一定意义之后再做行为决策(林崇德等, 2003)。为了在人际交往时能够针对他人的情形和特定的社交场合做出灵活反应，大脑必须能够加工具有社会意义象征性或预测性的情境信息(Lee & Harris, 2013)，例如他人的身份、意图、行为和情绪等。那么他人在场的道德决策过程中，个体最直接最简单的预测信息来源就是决策内容的信息传递。

我们的研究首先探索决策暴露的影响，探究潜在的社会评价是否会改变被试的决策倾向，此时被试不会接收到他人明确给予的信息反馈，如果决策倾向发生改变，则基本来自于个体对他人行为的预测，以及自身的形象管理或声誉担忧。然而我们并未发现被试决策的显著变化，原因之一可能是被试和观察员之间不存在信息交流，单纯的决策暴露难以改变被试的决策倾向。因此我们设计了实验二，继续探索不同信息交流程度下的决策变化。

在平行判断的情境中，被试均能看到双方的道德判断结果，但不需要对决策结果进行评判和统一，考察当接收到社会信息反馈时，个体的道德判断是否会受到他人的影响。虽然实验过程中不涉及评价环节，但平行组高于单独组的确定性却反映着信息反馈对被试决策信心的增强作用。我们曾对平行组中每对被试决策的一致性进行计算，48 个故事中双方的一致性达到 62.8%(每个故事决策“一致”计为“1”，“不一致”计为“0”，计算每对被试一致性得分的总和后求平均)。与此同时，平行组 I 参数升高，展示出不行动倾向，进一步说明，两个决策者互相进行的信息暴露会使他们改变自身决策倾向。

共同判断的情境中，被试不仅能看到双方的道德判断结果，还要求双方根据决策信息反馈选择是否修改道德决策，使得最终达成一致决策。此时被试需要根据对方的判断结果，揣测对方意图，经由多轮次非言语的行为判断达成一致的道德判断。这种情况下信息交流程度进一步加深，被试需要对对方的意图有更加清晰和深入的揣测。最终双方同样表现出不行动倾向，但同时伴随着对道德规范敏感性的提升，然而，

共同组中最终轮的决策确定性相较于第一轮有所下降，说明信息交流使被试产生了妥协，此外，共同组较高的社会赞许性表明该组被试更加关注自身形象。

总的来说，“不同社会情境”本质上改变的是决策者社会信息交流的程度和方式。从单纯的个人决策公开，到双方决策相互公开，直至决策公开后需要反馈协调，被试间的信息交流程度逐层递进，在结果中也展现出不同的表现倾向。然而，我们的研究也存在一定局限性，并未探讨在言语交流下个体道德决策倾向的变化，虽然现实生活中部分道德场景不涉及旁观者之间的语言交流(如：偶然路过的欺凌现场)，但是，言语中展示着决策过程的思考和博弈，未来研究可以设计语言交流范式进一步进行探索。

综上，我们对实验操纵的本质内容以及局限性在文中讨论部分进行补充，具体修改如下(见 p10-p11)：

“本研究采用不同类型的社会情境范式对道德判断进行研究，分别考察了社会观察、平行判断和共同判断之间的决策倾向差异，在这三种实验条件下被试间的信息交流程度逐层递进，从单纯的个人决策公开，到双方决策相互公开，直至决策公开后需要反馈协调，被试展现出的道德决策倾向出现了明显的不同。可见，日常的道德判断与特定的社会情境密切相关，这对理解现实世界中个体的道德决策过程和道德行为，甚至公共道德事件都具有重要的参考意义。

本研究仍存在一定局限性，研究中被试的信息交流主要通过非言语过程实现，虽然现实生活中部分道德场景不涉及旁观者之间的语言交流(如：偶然路过的欺凌现场)，但是，言语中展示着决策过程的思考和博弈，未来研究可以设计言语交流范式进一步展开探索。”

审稿专家二意见：

该论文探究在他人观察或彼此参与等社会情境下个体道德决策的行为偏向。论文包括两项行为实验，并结合了 CNI 模型拟合。结果发现，平行判断和共同判断导致被试出现明显的不行动倾向，这种倾向可能源自个体对于道德规范的敏感性增加。该论文研究议题有趣，实验设计规范。研究采用了 CNI 模型拟合参数来揭示道义决策背后的潜在机制，具有较高的创新性。读下来，整个“故事”逻辑严密，观点具有较强的启发性。以下是一些修改意见，供参考。

意见 1：题目：是否可以更具体、更精准？社会情境、道德判断都太大，不够具体。如能进一步体现结果的主要观点则更好。

回复：非常感谢审稿专家提出的宝贵意见。我们对文章题目做了更加具体和精准的修改。本研究中关注的实验情境无论是受他人观察还是双方共同参与，都涉及道德决策的公开性，只是公开程度和信息交流深度有所不同。杜秀莲和魏巍(2021)以及杜秀芳等人(2022)在研究中对不同实验任务操纵了公开性和匿名性，且文中涉及关于公开情境的论述。因此，在借鉴前人研究基础上，我们将“社会情境”修改为“公开情境”，使其更加符合研究目的。

本研究所使用材料为道德两难困境，虽然已有研究中存在使用道德判断指代道德两难判断的写作方式(陈欣等, 2021; 丁凤琴, 马小红, 2023; 云祥, 2023)，但为了使题目表述更具针对性，我们将“道德判断”修改为“道德两难决策”。同时，为了能够在标题中体现主要研究结果，综合考虑后我们将文章标题修改为“不出手的道德？公开情境对道德两难决策的影响”(见 p1、p9)。

此外，我们还对文章的关键词“公开情境”和“道德两难”(见 p1、p9)，以及文中部分涉及社会情境的表述做了修订。

意见 2：摘要：建议先加一句背景介绍并凸显该研究的意义。建议首次提到 CNI 模型时能具体描述。最后一句“结果表明”太泛泛，是否可以更具体。

回复：非常感谢审稿专家提出的建议。我们对摘要部分进行了全面的修订，增加了关于研究背景的介绍，简要补充了 CNI 模型各参数的意义，并且修改了对研究结果的表述。具体修改内容如下(见 p1)：

“以往研究多集中于使用经典道德困境对单被试的决策进行探讨。本研究采用能够量化决策者结果敏感性、道德规范敏感性以及一般行动倾向的 CNI(consequence, norm, inaction)模型，通过两个实验考察公开情境对道德两难决策的影响。实验一使用社会观察范式，发现他人观察并不影响被试的道德决策。实验二使用双人平行判断范式和共同判断范式。结果显示，相比于单独判断，平行判断和共同判断导致被试出现明显的不行动倾向，而且共同判断组对于道德规范的敏感性增加，这两种情境下个体对道德困境中主体的行为做出更多的“不接受”反应。研究表明，公开情境下个体看似道义的决策可能源于营造良好社会形象的需要，以及避免承担行动责任的动机。”

意见 3：引言：首次出现 CNI 模型时，应给出三个单词英文全称。应适当介绍模型计算的基本原理及其适用性。

回复：感谢审稿专家提出的问题。CNI 模型概念中的三个字母分别代表个体在道德判断中三种不同的决策倾向：结果敏感性(sensitivity to consequences, C)，道德规范敏感性(sensitivity to moral norms, N)和一般性的行动或不行动倾向(general preference for inaction versus action, I)(Gawronski et al., 2017)，参考云祥(2020; 2023)的研究，我们将 CNI 模型补充书写为“CNI(consequence, norm, inaction)模型”(见 p2)。

由于考虑到投稿文章篇幅相对有限，初稿并未详细介绍 CNI 模型。CNI 模型为了区分决策者的道德规范敏感性、结果敏感性以及一般的行动倾向，对困境中的道德规范和行动结果利弊大小进行了操纵，使得每个故事背景下的困境都有四个变式。在 4 种困境中，接受和不接受的反应概率之和为 1。其中，C 参数表示对结果的敏感性，N 参数表示对道德规范的敏感性，I 参数表示在不考虑结果和规范时，个体的不行动偏好。模型用 4 条路径描述了由三种倾向驱动时行为选择的 4 种情况(见图 1)。当决策由结果驱动时，路径 C 的描述代表结果驱动行为选择的情况(图 1 第一行)；由道德规范驱动时，路径(1-C)×N 代表结果无法驱动但规范驱动行为的情况(图 1 第二行)；当个体存在一般的不行动倾向时，路径(1-C)×(1-N)×I 代表了结果和道德规范都不驱动行为，个体存在不行动偏好的情况(图 1 第三行)；路径(1-C)×(1-N)×(1-I)则描述了当个体存在一般的行动倾向时的情况(图 1 第四行)。

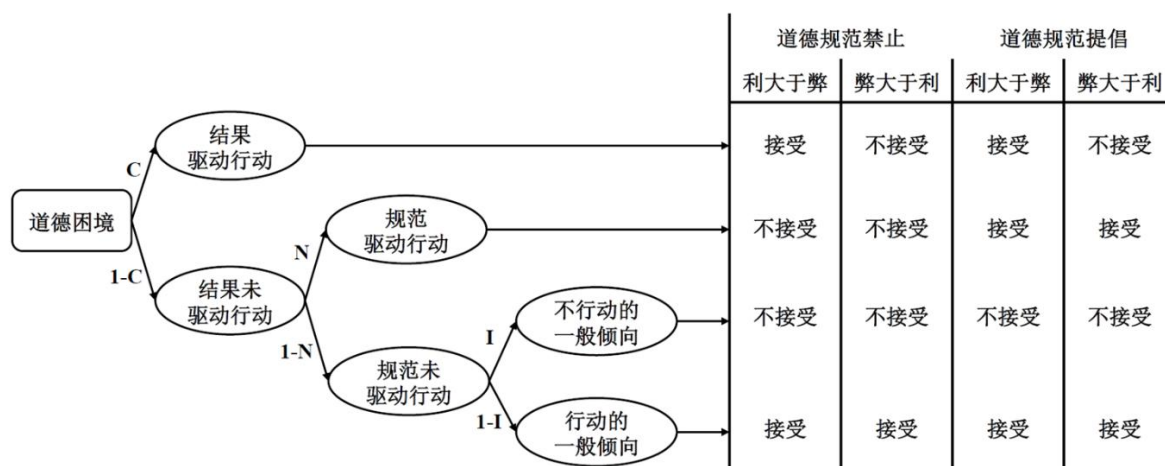


图 1 CNI 模型路径图解(资料来源：Gawronski et al., 2017)

通过这 4 条路径，可以推导出 4 种困境类型下选择接受或不接受概率的数学方程。因此，计算行为结果中 4 种故事类型下选择接受和不接受的反应率，代入方程可得三个未知数 C、N、I 的大小。使用最大似然估计法生成针对 3 个未知数的参数估计，利用拟合优度 G^2 评估模型在数据描述方面的充分性，观察到的概率和模型预测的概率之间的统计差异越小，模型拟合越好。所以，若 G^2 不显著($p>0.05$)，表明数据与模型拟合良好；反之，模型拟合较差。Gawronski 等人(2017)提供了基于 multiTree 软件构建的多项式模型，在其中输入 8 个反应概率可得三参数的估计值(模型文件及材料见：<http://www.bertramgawronski.com/docum>

ents/CNI-ModelIndDiffMaterials.zip)。对于 C 参数和 N 参数，估计值显著大于 0 时表明个体的决策受到结果敏感性或道德规范敏感性的驱动；对于 I 参数，估计值显著大于 0.5 时表明个体有一般的不行动的反应偏好，小于 0.5 时表明个体有一般的行动的反应偏好。

另外，可以通过假设两个变量的参数值相等来比较两组参数之间的差异(如设置： $C_1=C_2$)，若该操作使模型的拟合优度统计指标 G^2 显著性下降($p<0.05$)，表明在该参数上两组存在显著差异；反之，表明两组之间在参数上不存在显著差异(Gawronski et al., 2017)。

综上，我们将 CNI 模型的基本逻辑补充在引言部分，同时对模型的适用性进行了更加详细的阐述，具体修改如下(见 p3)：

“利用四种困境类型下的反应概率，模型可以推导出个体在所有困境中的决策情况，描述其对结果敏感的决策模式(图 1 第一行)，对规范敏感的决策模式(图 1 第二行)，以及不考虑结果和规范的一般的不行动或行动的决策模式(图 1 第三、四行)。以此对决策倾向进行了分离，实现对决策者结果敏感性(C 参数)，道德规范敏感性(N 参数)以及一般的不行动或行动倾向(I 参数)的量化。由于社会化过程中道德规范的养成(Haidt, 2007)，以及趋利避害的理性考虑，人们通常较少接受禁止弊大的行动，更容易接受提倡利大的行动，而对于规范和结果冲突的困境，个体在权衡之后做出的反应更能体现其决策倾向。目前，基于 CNI 模型的研究在情绪(Gawronski et al., 2018; 刘传军等, 2021)、权利感(Gawronski & Brannon, 2020; 云祥, 2020)、压力(Zhang et al., 2018; Li et al., 2021)、人格特征(Luke & Gawronski, 2022)等因素对道德判断的影响中展现出良好的适用性，为理解道德决策的过程提供了新的视角和理论依据。”

此外，我们还在“2.1.4 CNI 模型分析”部分对 CNI 模型的具体分析方法进行了较为详细的阐述，具体修改如下(见 p5)：

“基于多项式加工树模型原理，在构建好的 CNI 模型中根据被试的行为数据进行模型拟合，使用 mult iTree 软件(Moshagen, 2010)估计 3 个潜在心理过程(C 、 N 、 I)的概率。对于所有两难困境，将选择“不接受”赋值为 0，选择“接受”赋值为 1。用 G^2 表示数据与模型的拟合优度，若 G^2 不显著($p>0.05$)，表明拟合良好。估计出的 C 和 N 参数若显著大于 0，表明被试有明显的结果和规范敏感性， I 参数若显著大于 0.5，表明被试存在一般的不行动倾向。在比较不同组别参数差异时，若 ΔG^2 显著($p<0.05$)，说明两组被试参数存在显著差异(Gawronski et al., 2017)。”

意见 4：引言：对四种道德规范的个体倾向差异应适当做铺垫，描述其表现(例如，通常人们更容易做出行动利大决策，较少做出禁止弊大决策)并解释可能机制。以及，什么是“道义”的决策，需要澄清。

回复：非常感谢审稿专家指出的问题，我们将对这两个问题进行详细阐述。

意见 4-1，四种故事类型个体倾向差异的表现及可能的机制：人们的行动受到社会规范的制约，道德作为一系列引导社会行为的价值观念和传统的集合(Haidt, 2007)，是不同群体之间共同默认的行为准则和社

会规范。而不伤害他人被认为是人类道德中最基础的元素(Park et al., 2016)。与此同时, 追求更大福祉是人们道德理性的表现(Greene et al., 2008)。所以在通常情况下, 人们通常较少接受禁止弊大的行动, 更容易接受提倡利大的行动。而对于道德规范和结果产生冲突的困境, 个体在权衡之后做出的反应更能体现其决策的一般化倾向。

为了方便读者更好地理解 CNI 模型在故事中做出的设计, 我们对四种故事的区别以及可能的表现进行了详细的说明, 具体修改如下(见 p3):

“基于多项式加工树模型(multinomial processing tree, MPT)建构的道德判断的 CNI(consequence, norm, inaction)模型可以弥补上述局限(Gawronski et al., 2017)(见图 1)。该模型设计了 12 个故事情境, 利用行动后果的功利化收益程度高低操纵结果(利大或弊大), 利用行为的直接后果操纵规范(不伤害生命或帮助弱者的规范), 每个情境生成四种版本: 道德规范禁止, 行动利大于弊(简称“禁止利大”)、道德规范禁止, 行动弊大于利(简称“禁止弊大”)、道德规范提倡, 行动利大于弊(简称“提倡利大”)、道德规范提倡, 行动弊大于利(简称“提倡弊大”)。这样, 无论结果利弊如何, 在规范禁止时不行动, 在规范允许时行动才能反映个体对道德规范的高关注。

利用四种困境类型下的反应概率……, 由于社会化过程中道德规范的养成(Haidt, 2007), 以及趋利避害的理性考虑, 人们通常较少接受禁止弊大的行动, 更容易接受提倡利大的行动, 而对于规范和结果冲突的困境, 个体在权衡之后做出的反应更能体现其决策倾向。……”

另外, 我们在引言的最后对假设的具体提出过程进行了补充, 对于可能出现的结果以及原因做出了详细阐释, 具体如下(见 p4):

“根据以往研究, 公开情境可能对 CNI 模型中的三个参数产生不同方向的影响, 从而改变道德决策倾向。研究发现第三方的存在能够增加亲社会行为(Bereczkei et al., 2010; Hallgeir, 2019), 并促进利他规范的执行(Kurzban et al., 2007), 个体有展现符合社会规范形象的需求(Anderson & Kilduff, 2009), 因此公开情境可能会提高个体对道德规范的关注。而且, 道义主义和功利主义决策带来的社会印象截然不同, 知觉者认为拒绝伤害的决策者更值得信任(Everett et al., 2016)、更温暖并富有同理心(Uhlmann et al., 2013), 而功利主义决策则代表着能力和自利倾向(Rom et al, 2017), 并且温暖是印象形成和人际喜好的主要决定因素(Lee et al., 2018)。所以, 公开情境可能会降低个体对行动结果利益的考虑。然而, 也有相当一部分研究指出, 他人的存在会降低个体的助人倾向(Zoccola et al., 2011; Hortensius et al., 2016), 此时面对困境的责任由群体共同承担, 个人的责任感和压力都会有所减轻(Bandura, 2002), 使被试表现出一般化的不行动倾向。”

意见 4-2, 对“道义”决策的澄清: 道德判断的过程中存在理性和直觉的共同作用, 据此, Greene(2001)等人提出了道德判断的双加工理论, 认为在实际的道德认知过程中, 理性推理和情绪反应共同驱动道德判断。个体的道德判断过程中存在两个不同的加工系统: 一个是外显的对道德原则的认知推理

过程；另一个则是内隐的与社会适应相关的情绪动机过程。个人在对道德两难问题进行决策时会产生两种相互冲突的动机倾向，一种是避免对他人造成伤害的道义主义倾向，一种是为了利益最大化而不惜做出伤害行为的功利主义倾向(Greene, 2007; 2014)。根据道义主义者的观点，行为的适当性完全取决于行为的类型，无论在特定情境下行为的后果如何，都不能违背道德原则。相反，根据功利主义者的观点，符合道德的行动是在特定情境下，最大限度地使自己或他人受益的行动(Greene et al., 2009)。

综上，我们在引言部分对“道义决策”进行了补充说明，修改如下(见 p2):

“根据道德判断的双加工理论，情绪和认知加工共同影响决策结果，情绪系统驱动个体做出避免任何伤害的道义主义决策，认知加工则使个体产生无视伤害追求利益最大化的功利主义决策(Greene, 2009)。”

意见 5：方法：2.1.2 实验材料部分，是否对一些情境可以适当举例，以帮助读者在不参看附录时也能明白其含义。

回复：非常感谢审稿专家提出的建议。CNI 模型提倡需要对故事情境中的道德规范和利弊大小进行操纵。因此，设计了 12 个基本的道德两难困境，每种困境有四个不同版本的变式，共形成 48 个道德两难困境。需要说明的是，对于道德规范的控制在于故事中主人公的行动(具体指故事最后一句话描述的行为)是否符合规范。我们在方法部分补充了一个新的便于读者理解的例子(见附录 1，故事 2 免疫困境)，以此为例，在 2a 和 2b 中，“给学生用药”是被试需要决策接受与否的行动，若接受，虽然可以阻止病毒传播，但是却会“让学生死于副作用”，因此属于“道德规范禁止”的故事，在 2c 和 2d 中，“送学生回国治疗”是具体行动，这可以挽救学生的生命，所以属于“道德规范提倡”的故事。

关于行动结果利弊的大小则涉及故事后半部分，2a 中“给学生用药”会使她一人死亡，但是可以拯救众多老人和儿童，类似于“杀一救多”，所以是“行动利大于弊”的困境。事实上，在 CNI 模型的四类故事中，“禁止利大”的故事基本等同于经典道德两难困境，如：电车困境，都是规范不允许伤害但行动可以拯救更多人。2b 中感染病毒的后果只是胃痉挛，并未危及生命，若给学生用药会伤害一条生命，所以属于“行动弊大于利”的困境。另外两种故事以此类推较好理解，在此不再赘述。

综上，我们对 2.1.2 材料部分的具体修改如下(见 p5):

“道德两难困境来自对 Körner 等(2020)研究材料的翻译和适当修改(云祥, 2023)。在 12 个基本故事背景下，两难困境以(道德规范：禁止行动、提倡行动)×(结果：行动利大于弊、行动弊大于利)的方式组合，共形成 48 个两难困境(见附录一)，被试需要决定是否接受故事中的行动。示例见表 1，故事背景为：“你是欠发达国家中一家医院的院长。一名在该国做志愿者的外国学生感染了一种罕见的病毒。” ”

表 1 四种道德判断材料示例

禁止利大	禁止弊大	提倡利大	提倡弊大
这种病毒传染性极强，对老人和儿童都是致命的。唯一能有效阻止病毒传播的药物有严重的副作用。尽管病毒不会杀死她，但这名学生患有慢性免疫缺陷，用药将使她死于严重的副作用。在这种情况下给学生用药是可以接受的吗？	这种病毒传染性极强，被感染者会出现严重的胃痉挛。唯一能有效阻止病毒传播的药物有严重的副作用。尽管病毒不会杀死她，但这名学生患有慢性免疫缺陷，用药将使她死于严重的副作用。在这种情况下给学生用药是可以接受的吗？	这种病毒传染性极强，被感染者会出现严重的胃痉挛。这名学生患有慢性免疫缺陷，如果不回国接受特殊治疗，她将死于病毒。将她带出隔离区会带来相当大的病毒传播风险。在这种情况下，把学生从隔离区送回她的祖国接受治疗是可以接受的吗？	这种病毒传染性极强，对老人和儿童都是致命的。这名学生患有慢性免疫缺陷，如果不回国接受特殊治疗，她将死于病毒。将她带出隔离区会带来相当大的病毒传播风险。在这种情况下，把学生从隔离区送回她的祖国接受治疗是可以接受的吗？

意见 6：方法：2.2.1 似乎是方法部分，放成 2.1.4 似乎更合适。

回复：感谢审稿专家提出的建议。我们已将“2.2.1 数据分析”修订为“2.1.4 CNI 模型分析”，并且对 CNI 模型的具体分析方法进行了较为详细的阐述，具体修改内容详见意见 3 的回复。

意见 7：结果：3.2.2 给出了 CNI 参数及其在三个条件下的方差分析。这些参数是否可以进一步在四种道德规范类型下分别计算得到？如可以，是否有有趣发现？

回复：非常感谢审稿专家提出的建议。由于我们在引言中未对 CNI 模型进行充分介绍，无法帮助读者更好地理解模型的具体分析方法。关于 CNI 模型的分析原理详见意见 3 的回复，CNI 参数的计算需要四种故事类型下所有的反应概率，因此无法在每种道德规范类型下分别计算得到。

事实上，3.2.2 中对 CNI 参数的比较并非由方差分析得来，而是在多项式模型的基础上对参数进行的两两比较。对于 CNI 参数的分析有两种方式，一种是计算每种条件下被试在所有故事类型下反应的概率(即选择“接受”和“不接受”的次数)，将其放入模型进行拟合得出每种条件下的参数，即文章中表 3 和表 5 展示的结果；另一种是计算每个被试在所有故事类型下反应的概率，放入模型拟合得到每个被试的参数，之后计算每种条件下所有被试参数平均值，这时得出的参数可以进行方差分析。以往许多研究采用第一种方法进行数据分析(Hennig & Hütter, 2021; Kroneisen & Steghaus, 2021; Li et al., 2021; Gawronski, 2022)，本实验中使用第二种方式计算得到的结果与第一种趋势相同，为了节约文章篇幅，同时使结果分析更具针对性，因此在本文中报告第一种计算方法所得结果。第二种方法的分析结果如下：

首先将所有被试对每种故事类型的道德决策分别代入模型进行拟合，得出每个被试的参数水平(见辅表 1)。以组别为自变量，分别以 C、N、I 三个参数的均值为因变量进行单因素方差分析，结果显示，对于 C 参数，组别主效应不显著， $F(2,172)=1.278$ ， $p=0.281$ ；对于 N 参数，组别主效应不显著， $F(2,172)=2.113$ ， $p=0.124$ ；对于 I 参数，组别主效应显著， $F(2,172)=5.778$ ， $p=0.004$ ， $\eta_p^2=0.063$ ，共同组和平行组的 I 参数显著大于单独组， $ps \leq 0.005$ 。对单独组和共同组的 N 参数进行独立样本 t 检验，两组差异边缘显著， $t(119)=1.865$ ， $p=0.065$ ，共同组的 N 参数有高于单独组的趋势。

辅表 1 个体层面道德决策的 CNI 参数($M \pm SD$)

	C 参数	N 参数	I 参数
单独组	0.199 \pm 0.134	0.449 \pm 0.277	0.418 \pm 0.235
平行组	0.238 \pm 0.141	0.461 \pm 0.273	0.566 \pm 0.222
共同组	0.227 \pm 0.130	0.543 \pm 0.280	0.548 \pm 0.292

意见 8：结果：压力知觉和社会赞许性这些量表的结果是否可以与接受率以及 CNI 参数如道德规范敏感性等做相关分析？如可以，是否有重要发现？

回复：非常感谢审稿专家提出的建议。在意见 7 中的回复中提到，我们在文中保留的是 CNI 的第一种分析方法，此时每种条件下得出的参数源于所有被试的反应集合，即各参数只有一个估计值，这种情况下无法进行相关分析。但使用第二种方法时可以对压力知觉和社会赞许性与 CNI 三参数做相关分析。需要说明的是，CNI 参数的估计源于被试在各故事类型下的反应概率，所以在数据的本质表现上不会有太大差异。

对个人 C、N、I 参数及确定性分数与量表得分均值进行 pearson 相关分析(见辅表 2)，整体上 N 参数和社会赞许性之间存在显著正相关， $r=0.161$ ， $p=0.033$ ，决策确定性和压力知觉呈显著负相关， $r=-0.269$ ， $p<0.001$ ，和社会赞许性呈显著正相关， $r=0.211$ ， $p=0.005$ 。

辅表 2 相关分析结果

	压力知觉	社会赞许性
C 参数	-0.081	-0.014
N 参数	-0.050	0.161*
I 参数	0.033	-0.043
道德决策确定性	-0.269***	0.211**

文中 3.2.3 量表的差异检验结果表明，共同组的社会赞许性显著高于单独组，从整体的相关结果上看，社会赞许性可以显著正向预测 N 参数的变化，可见，人们由于在意社会评价而导致决策过程中更加关注道德规范(Fehr & Fischbacher, 2004; Rom & Conway, 2018)。此外，高压压力知觉会降低被试的决策信心，压力状态下的决策可能更多处于妥协和让步，同时高社会赞许性对高确定性的正向预测，进一步说明，被试想要展现良好的道德形象，不仅会改变道德倾向，还会展示出对自己决策的高度认可。

意见 9：讨论：该研究的重要贡献，即创新性发现所带来的理论拓展等，需要更突显并深入阐述。

回复：感谢审稿专家提出的宝贵意见。本文采用不同类型的社会情境范式对道德判断进行研究，分别考察社会观察、平行判断和共同判断下的决策倾向差异。首先，现实世界中许多道德难题本质上都是社会性的，因为它们要求公开披露自己的道德立场(Lee et al., 2018)。这可能会影响我们将道德信念转化为行动的过程，而这在实验室实验中经常被忽略。本研究结果表明，我们日常的道德判断与特定的社会情境密切

相关，这对理解现实世界中的道德现象具有重要的参考意义。

其次，先前研究表明，人们在道德困境中的决定并不完全是由道德原则的严格实施驱动的。相反，会受到一些看似无关的变量的影响，比如性别(Friesdorf et al., 2015)、个性(Bartels & Pizarro, 2013)、文化背景(Han et al., 2014)等。同样地，公开情境中的决策判断与单独或匿名决策不同，除了从声誉担忧和自我呈现的角度解释之外(Rom & Conway, 2018; Lee et al., 2018)，还需要考虑群体从众压力、责任分散等社会因素，才能使研究更具生态效度。有必要重新审视前人的诸多实验发现，很可能是多种动机因素导致的公开情境下的决策变化。本文在一定程度上揭示了个体在公开情境中做出道义决策的心理过程，并对潜在的心理机制提供新的见解，丰富了公开情境下的道德判断研究。

最后，不同程度公开情境对行为的影响研究其实并不局限于道德困境，这也提示我们在面对其他具有社会属性的研究议题时，需要将社会信息交流的作用纳入研究之中。

综上，我们在讨论中对研究的理论意义进行了深入的补充，具体修改如下(见 p10):

“本研究采用不同类型的社会情境范式对道德判断进行研究，分别考察了社会观察、平行判断和共同判断之间的决策倾向差异，在这三种实验条件下被试间的信息交流程度逐层递进，从单纯的个人决策公开，到双方决策相互公开，直至决策公开后需要反馈协调，被试展现出的道德决策倾向出现了明显的不同。可见，日常的道德判断与特定的社会情境密切相关，这对理解现实世界中个体的道德决策过程和道德行为，甚至公共道德事件都具有重要的参考意义。公开情境中的决策判断与单独决策不同，除了从声誉担忧和自我呈现的角度解释之外(Rom & Conway, 2018; Lee et al., 2018)，还需要考虑群体从众压力、责任分散等社会因素，才能使研究更具生态效度。因此，有必要重新审视前人的诸多实验发现，很可能是多种动机因素导致了个人在公开情境下的决策变化。本文在一定程度上揭示了个体在公开情境中做出更加“道义主义”决策的心理过程，并对潜在的心理机制提供了新的见解，丰富了公开情境下的道德判断研究。此外，不同公开程度对社会行为的影响并不局限于道德领域，其他具有社会属性的研究议题也应当重视社会情境中人与人之间信息反馈的作用。”

编委会意见：请作者根据审稿人意见进行修改，并对文章语言和写作规范性和流畅性进行进一步修改和提升，邀请同行进行挑剔性阅读。

回复：已根据审稿专家的意见对文章进行逐一修改，并进一步检查了文章的语言表述和写作规范。再次感谢各位专家及编辑部对本文的审阅！

参考文献(修改说明)

- 陈欣, 陶欣蕾, 李梦鸽. (2021). 清洁启动和主体对道德判断的影响: 道德自我意象的中介作用. *心理研究*, 14(06), 505 – 511.
- 丁凤琴, 马小红. (2023). 净脏启动对道德判断的影响: 有调节的中介效应. *心理与行为研究*, 21(02), 193 – 199.
- 杜秀芳, 胡卫红, 王静, 李方. (2022). 情境公开性与印象管理动机对建议采纳的影响. *心理科学*, 45(05), 1198 – 1205.
- 杜秀莲, 魏巍. (2021). 情境公开性对大学生道德心理许可效应的影响. *中国特殊教育*, (07), 82 – 88.
- 黄砾卉. (2010). *社会决策中意见分歧的信息交流效应*. (博士学位论文, 西南大学).
- 林崇德, 杨治良, 黄希庭. (2003). *心理学大辞典: 上册*. 上海: 上海教育出版社.
- 刘传军, 廖江群. (2021). 道德困境研究的范式沿革及其理论价值. *心理科学进展*, 29(08), 1508–1520.
- 云祥. (2023). 精神病态特质对大学生道德判断的影响: 基于 CNI 模型的分析. *心理技术与应用*, 11(08), 471–483.
- 云祥. (2020). 权力感对道德困境判断的影响: 基于 CNI 模型的分析. *心理与行为研究*, 18(04), 544 – 551.
- Anderson, C. & Kilduff, G. J. (2009). The Pursuit of Status in Social Groups. *Current Directions in Psychological Science*, 18(5), 295–298.
- Andersson, P. A., Erlandsson, A., Västfjäll, D., & Tinghög, G. (2020). Prosocial and moral behavior under decision reveal in a public environment. *Journal of Behavioral and Experimental Economics*, 87, 101561.
- Andreoni, J., & Petrie, R. (2004). Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of Public Economics*, 88(7–8), 1605–1623.
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, 31(2), 101–119.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121, 154 –161.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121(1), 154–161.
- Bereczkei, T., Birkas, B., & Kerekes, Z. (2010). Altruism towards strangers in need: Costly signaling in an industrial society. *Evolution and Human Behavior*, 31(2), 95–103.
- Bostyn, D. H., & Roets, A. (2017). An asymmetric moral conformity effect: Subjects conform to deontological but not consequentialist majorities. *Social Psychological and Personality Science*, 8(3), 323–330.
- Chen, C., Martínez, R. M., & Cheng, Y. (2020). The key to group fitness: The presence of another synchronizes moral attitudes and neural responses during moral decision-making. *NeuroImage*, 213, 116732.
- Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. *Journal of Personality and Social Psychology*, 104(2), 216–235.
- Egas, M., & Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 275(1637), 871–878.
- Epley, N., & Dunning, D. (2000). Feeling “Holier than thou”: Are self-serving assessments produced by errors in self or social prediction? *Journal of Personality and Social Psychology*, 79, 861–875.

- Everett, J. A., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, 145(6), 772–787.
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2), 63–87.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., . . . Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137, 517–537.
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83.
- Friesdorf, R., Conway, P., & Gawronski, B. (2015). Gender differences in responses to moral dilemmas: A process dissociation analysis. *Personality and Social Psychology Bulletin*, 41(5), 696–713.
- Gawronski, B. (2022). Moral impressions and presumed moral choices: Perceptions of how moral exemplars resolve moral dilemmas. *Journal of Experimental Social Psychology*, 99, 104265.
- Gawronski, B., & Brannon, S. M. (2020). Power and moral dilemma judgments: Distinct effects of memory recall versus social roles. *Journal of Experimental Social Psychology*, 86, 103908.
- Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., & Hütter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: The CNI model of moral decision-making. *Journal of Personality and Social Psychology*, 113(3), 343–376.
- Greene, J. D. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology*, 45, 581–584.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998–1002.
- Han, H., Glover, G. H., & Jeong, C. (2014). Cultural influences on the neural correlate of moral decision making processes. *Behavioural Brain Research*, 259, 215–228.
- Hennig, M., & Hütter, M. (2021). Consequences, Norms, or Willingness to Interfere: A proCNI Model Analysis of the Foreign Language Effect in Moral Dilemma Judgment. *Journal of Experimental Social Psychology*, 95, 104148.
- Hortensius, R., Schutter, D. J. L. G., & de Gelder, B. (2016). Personal distress and the influence of bystanders on responding to an emergency. *Cognitive, Affective and Behavioral Neuroscience*, 16, 672–688.
- Hutcherson, C. A., Montaser-Kouhsari, L., Woodward, J., & Rangel, A. (2015). Emotional and utilitarian appraisals of moral dilemmas are encoded in separate areas and integrated in ventromedial prefrontal cortex. *Journal of Neuroscience*, 35(36), 12593–12605.
- Kroneisen, M., & Steghaus, S. (2021). The influence of decision time on sensitivity for consequences, moral norms, and preferences for inaction: Time, moral judgments, and the CNI model. *Journal of Behavioral Decision Making*, 34(1), 140–153.
- Kurzban, R., DeScioli, P., & O'Brien, E. (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior*, 28(2), 75–84.
- Lee, M., Sul, S., & Kim, H. (2014). The impact of moral decision style on impression formation. *Korean Journal of Social and*

- Personality Psychology*, 28(2), 201–223.
- Lee, M., Sul, S., & Kim, H. (2018). Social observation increases deontological judgments in moral dilemmas. *Evolution and Human Behavior*, 39(6), 611–621.
- Lee, V. K., & Harris, L. T. (2013). How social cognition can inform social decision making. *Frontiers in neuroscience*, 7, 259.
- Li, Z., Gao, L., Zhao, X., & Li, B. (2021). Deconfounding the effects of acute stress on abstract moral dilemma judgment. *Current Psychology*, 40(10), 5005–5018.
- Luke, D. M., & Gawronski, B. (2022). Temporal stability of moral dilemma judgments: A longitudinal analysis using the CNI model. *Personality and Social Psychology Bulletin*, 48(8), 1191–1203.
- Martinsson, P., Pham-Khanh, N., & Villegas-Palacio, C. (2013). Conditional cooperation and disclosure in developing countries. *Journal of Economic Psychology*, 34, 148–155.
- Park, G., Kappes, A., Rho, Y., & Van Bavel, J. J. (2016). At the heart of morality lies neuro-visceral integration: Lower cardiac vagal tone predicts utilitarian moral judgment. *Social Cognitive and Affective Neuroscience*, 11(10), 1588–1596.
- Rom, S. C., & Conway, P. (2018). The strategic moral self: Self-presentation shapes moral dilemma judgments. *Journal of Experimental Social Psychology*, 74, 24–37.
- Rom, S. C., Weiss, A., & Conway, P. (2017). Judging those who judge: Perceivers infer the roles of affect and cognition underpinning others' moral dilemma responses. *Journal of Experimental Social Psychology*, 69, 44–58.
- Soetevent, A. R. (2005). Anonymity in giving in a natural context—a field experiment in 30 churches. *Journal of public Economics*, 89(11–12), 2301–2323.
- Sparks, A., & Barclay, P. (2015). No effect on condemnation of short or long exposure to eye images. *Letters on Evolutionary Behavioral Science*, 6(2), 13–16.
- Uhlmann, E. L., Pizarro, D. A., Tannenbaum, D., & Ditto, P. H. (2009). The motivated use of moral principles. *Judgment and Decision making*, 4(6), 479–491.
- Uhlmann, E. L., Zhu, L. L., & Tannenbaum, D. (2013). When it takes a bad person to do the right thing. *Cognition*, 126, 326–334.
- Van Dijk, F., Sonnemans, J., van Winden, F., 2002. Social ties in a public good experiment. *Journal of Public Economics*, 85, 275 – 299.
- Zhang, L., Kong, M., Li, Z., Zhao, X., & Gao, L. (2018). Chronic stress and moral decision-making: An exploration with the CNI model. *Frontiers in Psychology*, 9, 1702.
- Zoccola, P. M., Green, M. C., Karoutsos, E., Katona, S. M., & Sabini, J. (2011). The embarrassed bystander: Embarrassability and the inhibition of helping. *Personality and Individual Differences*, 51, 925–929.

