

## 《心理学报》审稿意见与作者回应

题目：第三方惩罚中的规范错觉：基于公正世界信念的解释

作者：杨莎莎；陈思静

### 第一轮

首先，非常感谢编委老师及外审专家细致审慎的阅读和富有建设性的意见，这对提升本文的写作质量和学术水平有极大的帮助！基于专家们所提出的问题和修改意见，我们在行文逻辑、研究说明、文字表述和写作规范等方面对文章进行了认真的修改（正文中标记为蓝色的是修改后的文字）。专家意见及修改说明具体如下：

#### 编委初审意见：

作者从规范错觉的角度理解第三方惩罚，视角比较新颖，实验逻辑合理，操作基本规范，研究内容丰富。实验中还有较多的具体问题需要进一步澄清。

**意见 1：**请对实验程序进行更详细的报告，例如，实验一中独裁者任务每种分配方案的试次数？实验二和实验四为何只选择 8-2 的分配结果进行实验？

**修改说明：**非常感谢专家所指出的问题，我们重新阅读了实验程序部分，发现写的确实不够清楚，因此从以下几方面进行了补充说明。

①实验 1 采用的是策略方法（strategy method），即所有被试都需要对 6 种不同的分配方案做出估计与惩罚，它的好处是能够完整地揭示被试对所有分配方案的反应（Jordan et al., 2016）。在该方法中，每种分配方案按随机顺序仅呈现 1 次，由于我们表述上的疏忽，先前并未明确地指出试次数和呈现方式。非常感谢您的指正，我们已经在实验 1 的实验程序（位于正文 2.3 部分）中对此进行了补充。

②实验二和实验四之所以只选择 8-2 的分配结果，主要是出于以下考虑：

首先，如果在实验二中采取重复博弈，则前一轮博弈（例如 8-2 分配）的规范提示可能会影响后一轮博弈（例如 9-1 分配）提示前的估计与惩罚，因此实验二采用了单轮次博弈；而实验四选择了与实验二相同的分配结果，部分原因是希望能够与实验二的结果形成印证。

其次，更重要的是，我们在三种不公平分配方案中（10-0、9-1 和 8-2）做了不公平程度（编码方式同实验 1，即 10-0 记为 5，表示完全不公平，以此类推）对规范错觉的回归分析，不

论是在实验一中 ( $\beta = -0.03, t = -0.76, p = 0.450$ ) 还是实验三中 ( $\beta = -0.03, t = -0.70, p = 0.486$ ), 不公平程度对规范错觉的回归系数都是不显著的, 因此我们认为在实验二和实验四中选择任意一种不公平分配结果都是可以接受的, 而不会对分析结论有实质性的影响。

最后, 现有文献中有大量研究选择 8-2 作为一个典型的不公平分配方案(e.g., Chen et al., 2020; Csukly et al., 2011; Przepiorka & Liebe, 2016), 这也在一定程度上说明该选择的合理性, 较符合主流的研究范式, 因此, 在实验二和实验四中最终遵照前人的研究范式, 选择了 8-2 作为不公平分配方案的代表。

### 实验二正文 3.2 部分补充内容如下:

但实验 2 中被试只需针对 8-2 的分配方案进行估计和实施惩罚, 主要是出于以下两个方面的考虑: 1) 若此处选择重复博弈, 前一轮博弈的规范提示可能会影响后一轮博弈中提示前的惩罚和估计; 2) 实验 1 的结果显示, 对于三种不公平分配方案而言, 分配方案的不公平程度在很大程度上不会影响规范错觉的强弱: 我们在实验 1 三种不公平分配方案中(10-0、9-1 和 8-2)做了不公平程度(编码方式同实验 1, 即 10-0 记为 5, 表示完全不公平, 以此类推)对规范错觉的回归分析, 结果显示不公平程度对规范错觉没有显著影响( $\beta = -0.03, t = -0.76, p = 0.450$ ), 因此本文参考现有文献选取了 8-2 的分配结果(Chen et al., 2020; Przepiorka & Liebe, 2016)。

### 实验四正文 5.2 部分补充内容如下:

但实验 4 中被试只需针对 8-2 的分配方案进行估计和实施惩罚, 一方面是因为实验 3 中不公平分配方案的不公平程度基本不会影响规范错觉的强弱( $\beta = -0.03, t = -0.70, p = 0.486$ ), 另一方面也是为了和实验 2 的结果相呼应。

**意见 2:** 实验一的目标之一在于检验规范错觉是否存在, 也并没有对规范错觉进行直接操纵, 在表述中直接把规范错觉作为自变量是否合适?

**修改说明:** 非常感谢专家的宝贵意见, 您的建议极大地提高了本文的严谨性! 我们根据您的建议修改了实验 1 中针对规范错觉的表述, 修改部分位于正文 2.2 和 2.4 部分(在 2.4 部分我们还删去了表 2 中“自变量”、“控制变量”的表述)。我们参照了 Fehr 和 Gächter (2002) 有关第三方惩罚的经典论文 *Altruistic punishment in humans* 的格式, 不再标明哪个变量是自变量或因变量等, 而是直接采用了变量的名称, 如规范错觉, 惩罚水平等。由于修改内容过于分散, 烦请专家参看正文中标蓝部分。

**意见 3:** 如何证明是规范错觉 (PO-M<sub>PS</sub>) 影响了惩罚行为, 而不是仅仅对他人惩罚水平的估计 (PO) 影响了惩罚行为?

**修改说明:** 感谢专家所指出的问题, 您的问题有助于我们对本文进行更深入的思考。我们主要考察  $PO - M_{PS}$  而不是 PO 本身对惩罚的影响是基于以下考虑:

首先, 由于对每个被试而言  $M_{PS}$  是个恒定值, 因此正如您指出的那样, 仅仅从统计分析的角度来说, 无论采用  $PO - M_{PS}$  还是 PO 都不会影响统计分析的结果。其次, 从定义上来看, 规范错觉指的是个体所感知到的规范与实际存在于群体中规范之间的差异 (Chung et al., 2020), 因此规范错觉包含两个部分: 个体对规范的感知 (在本文中即是个体对他人惩罚行为的估计) 和真实的规范 (在本文中即是他人实际的惩罚水平), 在理论上真实的规范在一定程度上独立于个体而存在, 所以我们确实也可以说个体对规范的估计决定了规范错觉的强弱, 进而影响了个体自身的行为。换言之, 规范错觉影响了自身行为或者个体的估计影响了自身行为, 这两种说法本质上是一体两面, 都是成立的。那为什么我们选择了  $PO - M_{PS}$  而不是 PO 作为影响自身惩罚行为的变量? 一个重要的原因是  $PO - M_{PS}$  可以表达出更丰富的内涵: 1)  $PO - M_{PS}$  可以直观地展示个体对他人惩罚水平的估计是否符合现实, 如果存在错估的话, 错估的方向如何? 我们只需通过展示  $PO - M_{PS}$  是否接近 0, 以及数值的正负即可以回答以上问题, 而如果采用 PO 则无法很好地达到上述目的; 2) 类似地, 在探讨规范错觉与自身惩罚行为的关系时,  $PO - M_{PS}$  同样可以直观地展示当规范错觉存在或不存在时自身惩罚行为的变化, 而采用 PO 则无法很好地做到这一点; 3) 在引言中, 我们提出根据两种不同的理论, 既可以得出规范错觉为正也可以得出规范错觉为负的结论, 而本文的一个重要目标就在于验证规范错觉的方向问题, 而如果仅仅考虑 PO 可能无法很好地回答上述问题。最后需要说明的是, 在现有文献中, 采用  $PO - M_{PS}$  作为规范错觉的操作定义是一种较为普遍的操作 (e.g., 陈思静等, 2021; Duong & Parker, 2018; Sandstrom et al., 2013), 从这个意义上来说, 我们的操作基本还是符合主流范式的。

**意见 4:** 实验二中提示高估组和提示低估组在提示前的差别如何? 提示后的差别如何? 请进一步增加以上组间比较的结果。

**修改说明:** 感谢专家的提醒, 您的建议极大地提升了本文的严谨性和科学性。遵循您的建议, 我们分别比较了提示高估组和提示低估组在提示前和提示后的差别。结果显示, 提示前两组被试在惩罚、估计和规范错觉上并无显著差异; 而提示后, 提示高估组的惩罚、估计和规范错觉均显著低于提示低估组。我们已在正文中增加了对该结果的说明 (位于正文 3.4 部分第

2 段和第 3 段，由于修改内容相对分散，烦请专家参看正文中[标蓝部分](#)）。下表是完整的检验结果：

提示前两组被试惩罚、估计和规范错觉的平均数差异检验

变量	<i>t</i>	<i>p</i>	<i>BF</i> <sub>01</sub>	95%CI	
惩罚	-0.04	0.971	5.401	-0.26	0.25
估计	0.05	0.963	5.398	-0.25	0.26
规范错觉	0.09	0.932	5.386	-0.24	0.26

提示后两组被试惩罚、估计和规范错觉的平均数差异检验

变量	<i>t</i>	<i>p</i>	<i>d</i>	95%CI	
惩罚	-7.43	<0.001	-1.29	-1.32	-0.77
估计	-11.23	<0.001	-1.95	-1.91	-1.34
错觉	-4.01	<0.001	-0.70	-0.87	-0.29

**意见 5：**按照作者的假设，激活公正世界信念后规范错觉会上升，而实验四的结果却呈现与该假设相反的趋势，请对该结果给出进一步说明。

**修改说明：**感谢专家所指出的问题，这里我们确实存在表述不清的问题。本文规范错觉的操作定义为：规范错觉 =  $PO - M_{PS}$ 。在原文中，我们说“规范错觉的上升”本义指的是在数值的意义而非绝对值的意义上的上升，如在实验 4 中，操作自变量后，抑制组的规范错觉均值是-0.42，而激活组的规范错觉均值是 0.09，我们是在数值的意义上认为规范错觉上升了，因为  $0.09 > -0.42$ 。但正如您指出的那样，这样的表述确实容易给读者造成误会，因此我们将“规范错觉的上升”改成“公正世界信念的激活削弱了规范错觉的低估程度”。需要指出的是，尽管单纯从数字上来看，激活组的规范错觉为正，但实际上  $PO - M_{PS}$  间的差值并不显著，也就是说，激活组不存在规范错觉。类似的，我们将假设 3 从“公正世界信念正向预测了第三方惩罚中的规范错觉”重新表述为“公正世界信念越弱，规范错觉的低估程度越大，而公正世界信念的增强削弱了规范错觉的低估程度。”对文中其他容易引起误解的地方我们也做了相应的修改。

**意见 6：**作者提出了低估他人惩罚行为的三种解释，并指出本文重点考察第三种解释，但实

际上第一种解释也在前文中频繁提到，并且用于引出作者的 RQ1。建议明确此处，避免给读者造成误解。

**修改说明：**感谢专家的细心阅读与中肯的建议，我们遵循您的建议，删去了“本文重点考察第三种解释”这一表述。由于第三种解释是本文实验 3 的研究结论，而第二种解释是我们对未来研究提供的理论思路，因此也相应地调整了二、三两种解释的顺序。修改后的内容如下（见正文 6.1 第 2 段）：

其次，我们需进一步挖掘人们低估他人惩罚行为的原因。我们提出了三种解释。第一，由于存在归因偏差，人们总是倾向于低估他人的积极行为（Bursztyn et al., 2020; Dempsey et al., 2018），除非高估他人的积极行为能为自己带来收益。其次，公正世界信念在一定程度上解释了个体对他人惩罚行为的低估。胡金生等（2012）指出，公正判断更多地依赖于道德直觉而非社会线索，因此我们有理由认为，当被试无法获取关于惩罚的外部信息时，规范错觉的产生可能依赖于个体的某种直觉或信念。本文的研究结果显示：较弱的公正世界信念往往意味着较强的低估的规范错觉。具体来说，被试低估他人惩罚的重要原因可能是，他们对“善有善报，恶有恶报”这一信念的认同度较低，即对违规行为会受到应有惩罚这一观点持怀疑态度，因此人们低估了他人的惩罚水平，而这又进一步减少了被试的惩罚行为。第三，从社会发展的角度来考虑，几乎所有现代社会都处于惩罚权力高度垄断的阶段（Mathew, 2017），即公检法系统或类似的机构垄断了大部分对违规者的惩罚权力，因此尽管“路见不平拔刀相助”式的第三方惩罚行为依然存在，但对于庞大的惩罚系统而言只占到一个较小的比例，这可能在一定程度上加强了人们对他人惩罚行为的低估，因为在现代社会中人们已经习惯了奖惩行为由国家机构出面实施。上述解释提出了一种新的可能性，也就是说比起惩罚权力高度集中的现代社会来，在处于前国家时期的社群中（pre-state community），这种低估是否不存在或程度要小得多？这为未来的研究提供了一种新的理论思路。我们从行为性质、个人信念和社会发展三个角度提出了导致被试低估他人惩罚的原因，但这三个原因是独立的还是相互交织在一起，需要未来更多的研究才能给出可靠的回答。

**意见 7：**讨论的 6.1 部分作者在最后提到，“在这种情况下，外在参照点完全取代了内在参照点而成为了个体估计他人惩罚行为的依据。”文中哪项结果能够体现出外在参照点取代内在参照点成为个体的估计依据？感知社会距离增大时公正世界信念对规范错觉的预测作用消失，并不能够说明是内在参照点取代了外在参照点。

**修改说明：**感谢专家的细心阅读，您的建议极大地提升了本文在写作方面的规范性。诚如您

所言，本文主要关注内在参照点（公正世界信念）对规范错觉的影响，而外在参照点（社会距离）仅作为这一路径的调节变量，因此原文中的表述确实不够严谨。遵循您的建议，我们删去了“在这种情况下，外在参照点完全取代了内在参照点而成为了个体估计他人惩罚行为的依据”这一表述。

参考文献

陈思静, 濮雪丽, 朱玥, 汪昊, 刘建伟. (2021). 规范错觉对外出就餐中食物浪费的影响: 心理机制与应对策略. *心理学报*, 53(8), 904–918.

Chen, H., Zeng, Z., & Ma, J. (2020). The source of punishment matters: Third-party punishment restrains observers from selfish behaviors better than does second-party punishment by shaping norm perceptions. *PloS One*, 15(3), e0229510.

Chung, H. F., Shi, J. W., & Sun, K. J. (2020). Why employees contribute to pro-environmental behaviour: The role of pluralistic ignorance in Chinese society. *Sustainability*, 12(1), 239.

Csukly, G., Polgár, P., Tombor, L., Rádhelyi, J., & Kéri, S. (2011). Are patients with schizophrenia rational maximizers? Evidence from an ultimatum game study. *Psychiatry Research*, 187(1-2), 11–17.

Duong, H. T., & Parker, L. (2018). Going with the flow. *Journal of Social Marketing*, 8(3), 314–332.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.

Jordan, J. J., McAuliffe, K., & Rand, D. G. (2016). The effects of endowment size and strategy method on third party punishment. *Experimental Economics*, 19(4), 741–763.

Przepiorka, W., & Lieke, U. (2016). Generosity is a sign of trustworthiness—the punishment of selfishness is not. *Evolution and Human Behavior*, 37(4), 255–262.

Sandstrom, M., Makover, H., & Bartini, M. (2013). Social context of bullying: Do misperceptions of group norms influence children’s responses to witnessed episodes? *Social Influence*, 8(2-3), 196–215.

.....

审稿人 1 意见：

研究从规范错觉角度考察第三方惩罚，探讨二者之间的关系并进一步探究了其影响因素，有创新性，有一定的应用价值。整体行文完整，采用 4 个实验逐步论证和补充，基本验证了研究假设。但仍存在一些问题。

**意见 1：** 第三方惩罚本身是为了维护社会规范，而社会规范对个体行为的影响是一个复杂的认知加工过程，从对社会规范的认知到产生规范错觉，再到做出惩罚行为，这一过程中影响因素有很多但都未进行控制，影响研究的外部效度。

**修改说明：** 感谢专家的宝贵意见，您的意见给了我们很大的启发！首先，诚如专家所言，社会规范对个体行为的影响是一个复杂的认知加工过程，中间涉及到许多因素的影响，而这也是很多社会心理学课题所面临的问题。我们最初希望从两方面来尽量减弱无关因素的影响，



从而提高本研究的外部效度：1) 尽可能增加被试量；2) 将被试从学生群体扩展到社会群体。就第一方面而言，参与本研究的有效被试达到了 1031 名，这在社会心理学中属于被试量较大的研究，可以在一定程度上降低其他因素的影响，但第二点，由于目前学校管理层考虑到防疫问题，我们希望使用社会被试的研究申请未能得到批准，因此只能退而求其次在所有实验中都采用了学生被试，这显然导致本研究结果的应用范围在一定程度内受限，例如，我们不清楚个体的社会经历如求职、工作、创业、婚育等因素是否会影响个体的公正世界信念，从而对实验结果产生影响，这确实是本研究的一个局限之处，我们也希望未来研究能在这方面有所改进，从而提高本研究结论的说服力。

其次，在第三方惩罚中，目前鲜有文献讨论第三方惩罚中的规范错觉，本文首次证明了在不公平分配情境下确实存在对他人惩罚行为的规范错觉，并通过实验揭示了规范错觉对自身惩罚行为的影响，此外我们还探讨了公正世界信念的作用以及社会距离的调节效应。尽管上述结果都较为新颖，但总体而言，本文属于一个探索性的研究，还未能到达构建完整模型的研究阶段。换言之，我们将重点放在了引入一种新的视角，即通过公正世界信念来解释个体规范错觉的形成及其对自身惩罚行为的影响，但这并不意味着我们已经穷尽了所有的影响因素，恰恰相反，作为一种复杂的社会心理现象，我们认为肯定还存在其它潜在的影响因素，这些因素有哪些？如何识别它们？如何控制它们的影响？我们希望本研究能够抛砖引玉，引起更多学者的兴趣，从而在未来研究中更好地解决这些问题。概括而言，作为一个探索性研究，本文尝试着提出一种新的视角，而且至少目前的数据也在很大程度上支持了这一新视角，从这个意义上来说，我们认为尽管本研究尚有诸多可改进之处，但不失为一次有益的尝试。

第三，我们也广泛查阅了目前关于规范错觉影响行为的文献，发现这些研究主要集中在物质使用和亲环境行为这两个领域，而大多数研究也未能控制这一过程中的其他影响因素（e.g., Kenney et al., 2017; Dumas et al., 2019; Larimer et al., 2020）。当然，我们并不是说存在这样的例子就证明没必要控制其他影响因素，而是说考虑到社会心理现象的复杂性，研究者又受到手中资源的限制，很多时候我们只能通过一些间接的方法（如尽量扩大被试量）来减小其他因素的影响，从这个意义上来，我们的研究基本符合目前规范错觉领域的研究规范。

最后，我们也在正文中补充了关于这一问题的讨论和展望（位于 6.2 研究不足），补充的文字如下：

最后，需要说明的是，社会规范对个体行为的影响是一个复杂的认知加工过程。而本文作为探索性研究，主要考察了规范错觉的重要作用与影响因素，并未控制这一过程中的诸多影响因素，如何识别这些因素并加以控制是未来研究值得探讨的一个重要方向。

意见 2：实验 1 结果显示，被试仅在面对不公平分配方案时才产生规范错觉，表明分配方案的公平程度是影响第三方惩罚中规范错觉的一个重要变量，但这一点并未在接下来的几个实验中进行进一步论证，并且未在文章摘要和讨论部分提及。建议对这部分内容进行补充和解释。

修改说明：感谢专家的细心阅读与中肯的建议，您的建议有助于我们深入挖掘数据结果。我们试着从以下几个方面来回答您提到的问题。

**①总体而言，分配是否公平影响了规范错觉的存在性，具体表现为：仅在分配方案 10-0、9-1 和 8-2 中存在规范错觉。**我们在实验 1 中提到，目前有许多研究认为分配给对方的金额约小于 30% 是一种违规行为 (Csukly et al., 2011; Fehr & Fischbacher, 2003)。如果我们将 10-0、9-1 和 8-2 的分配方案视为不公平分配，将 7-3、6-4 和 5-5 的分配方案视为公平分配，上述结果就可以总结为：规范错觉主要存在于不公平的分配方案（即违规情境）中。但诚如专家所言，我们并未在摘要和结论中体现这一结果。遵循您的建议，我们在摘要部分和结论部分（正文 6.1 第一段）强调了违规情境这一重要前提。

**②在三种不公平的分配方案中，分配方案的不公平程度对规范错觉并无显著影响。**首先，我们合并了实验 1 中组 1（先惩罚后估计）和组 2（先估计后惩罚）的样本，针对 10-0、9-1 和 8-2 这三种分配方案，进行不公平程度（编码方式同实验 1，即 10-0 记为 5，表示完全不公平，以此类推）对规范错觉的回归分析，结果显示：不公平程度对规范错觉的回归系数是不显著的（ $\beta = -0.03, t = -0.76, p = 0.450$ ）。进一步，我们使用实验 3 的数据进行了上述回归，也发现了同样的结果（ $\beta = -0.03, t = -0.70, p = 0.486$ ）。因此我们有理由认为，在不公平的分配方案中，分配方案的不公平程度在很大程度上不会影响规范错觉的强弱。基于上述结果，我们在实验 1 的结果和讨论部分也进行了相应的补充说明（位于正文 2.4 表 2 下方第 1 段）：作为补充说明，我们还通过不公平程度对规范错觉的回归分析探究了违规情境下不公平程度对规范错觉的影响（考虑到在公平分配方案中不存在显著的规范错觉，因此仅选取 10-0、9-1 和 8-2 这三个不公平分配方案），结果显示：违规情境下不公平程度对规范错觉的回归系数是不显著的（ $\beta = -0.03, t = -0.76, p = 0.450$ ）。

**③本文仅关注不公平的分配方案，理由包括以下两个方面。**

1) 第三方惩罚主要是基于违规情境，在本文中表现为不公平的分配（10-0、9-1 和 8-2）。具体而言，违规情境下的第三方惩罚可以被视为一种牺牲个人利益、维护社会规范的积极行为 (Enge et al., 2017)。分配给对方的金额约小于 30% 是一种违规行为，恰好对应本文中 10-0、



9-1 和 8-2 这三种分配方案；而针对 7-3、6-4 和 5-5 这三种分配方案的惩罚，缺少了违规情境这一重要前提，并不在本文的重点讨论范围之内。

2) 实验 1 的分析表明，在 7-3、6-4 和 5-5 分配中，不存在显著的规范错觉，这也是我们未在后续分析中讨论这三个分配方案的原因之一。

对于上述 2 点理由，我们也在实验 1 中（位于正文 2.4 表 2 下方第 1 段的括号内）进行了补充说明：

考虑到在公平分配方案中不存在显著的规范错觉，因此仅选取 10-0、9-1 和 8-2 这三个不公平分配方案。

#### ④在实验 3 中，我们并没有将分配方案的不公平程度纳入分析框架，主要是出于以下三方面的原因。

1) 仅在违规情境下被试才会产生规范错觉，而基于实验 1 和实验 3 的分析都表明，在违规情境下不公平程度不影响第三方惩罚中的规范错觉，因此我们没有考虑不公平程度对规范错觉的影响。

2) 我们在实验 1 中进行的逐步回归表明：不公平程度会在一定程度上影响惩罚行为，表现为不公平程度越大，惩罚行为越高；但即便我们控制了不公平程度的影响，规范错觉还是能较好地解释惩罚行为。这也是我们没有进一步考虑不公平程度对惩罚行为影响的原因。

3) 我们进一步使用实验 3 的数据，在条件过程模型中加入了不公平程度这一控制变量，使用 PROCESS3.3 (Bootstrap  $N = 5000$ , Model = 7) 进行了检验。结果显示：(1) 在以公正世界信念为自变量、规范错觉为因变量、社会距离为调节变量时，公正世界信念 $\times$ 社会距离的回归系数仍然显著为负 ( $\beta = -0.21, t = -5.93, p < 0.001, 95\%CI = [-0.28, -0.14]$ )，这也意味着不公平程度在很大程度上不会影响社会距离对公正世界信念-规范错觉这一路径的调节作用；(2) 在以公正世界信念为自变量、惩罚行为为因变量、规范错觉为中介变量、社会距离为调节变量时，社会距离对中介效应的调节作用仍然显著为负 (Index =  $-0.13$ , BootSE =  $0.02$ , BootLLCI =  $-0.18$ , BootULCI =  $-0.09$ )。基于上述结果，我们没有将不公平程度纳入到条件过程模型的分析。

对于上述 3 点理由，我们也在实验 3 中进行了补充说明（位于正文 4.4.2 第 3 段），具体文字如下所示：

需要说明的是，在上述检验中，我们并未将不公平程度纳入分析框架，主要是出于以下原因。

第一，实验 1 的分析结果表明，在违规情境下，不公平程度对规范错觉无显著影响 ( $\beta = -0.03, t = -0.76, p = 0.450$ )；我们在实验 3 中进行了类似的回归分析，得到了一致的结果 ( $\beta = -0.03,$

$t = -0.70, p = 0.486$ )。第二, 实验 1 的结果表明, 虽然不公平程度在一定程度上影响了惩罚行为, 但即便我们控制了它的影响, 规范错觉仍然能在较大程度上解释惩罚行为。第三, 进一步使用实验 3 的数据, 在如图 7 所示的条件过程模型中加入了不公平程度这一控制变量, 使用 PROCESS3.3 (Bootstrap  $N = 5000$ , Model = 7) 进行检验后, 结果显示: 1) 在以公正世界信念为自变量、规范错觉为因变量、社会距离为调节变量时, 公正世界信念 $\times$ 社会距离的回归系数仍然显著为负 ( $\beta = -0.21, t = -5.93, p < 0.001, 95\%CI = [-0.28, -0.14]$ ); 2) 在以公正世界信念为自变量、惩罚行为为因变量、规范错觉为中介变量、社会距离为调节变量时, 社会距离对中介效应的调节作用仍然显著为负 (Index =  $-0.13$ , BootSE =  $0.02$ , BootLLCI =  $-0.18$ , BootULCI =  $-0.09$ )。基于上述结果, 我们并未在条件过程模型中考虑不公平程度的影响。

**意见 3:** 另外, 有些细节性问题建议补充: 1) 实验 1 共进行几轮博弈, 6 种分配方案的呈现顺序是否得到控制; 2) 实验 2 及实验 4 中选择“8-2”这一分配方案的依据。

**修改说明:** 非常感谢专家所指出的问题, 我们重新阅读了实验程序部分, 发现写的确实不够清楚, 因此从以下几方面进行了补充说明。

①实验 1 采用的是策略方法 (strategy method), 即所有被试都需要对 6 种不同的分配方案做出估计与惩罚, 它的好处是能够完整地揭示被试对所有分配方案的反应 (Jordan et al., 2016)。在该方法中, 每种分配方案按随机顺序仅呈现 1 次, 由于我们表述上的疏忽, 先前并未明确地指出试次数和呈现方式。非常感谢您的指正, 我们已经在实验 1 的实验程序 (位于正文 2.3 部分) 中对此进行了补充。

②实验二和实验四之所以只选择 8-2 的分配结果, 主要是出于以下考虑:

首先, 如果在实验二中采取重复博弈, 则前一轮博弈 (例如 8-2 分配) 的规范提示可能会影响后一轮博弈 (例如 9-1 分配) 提示前的估计与惩罚, 因此实验二采用了单轮次博弈; 而实验四选择了与实验二相同的分配结果, 部分原因是希望能够与实验二的结果形成印证。

其次, 更重要的是, 我们在三种不公平分配方案中 (10-0、9-1 和 8-2) 做了不公平程度 (编码方式同实验 1, 即 10-0 记为 5, 表示完全不公平, 以此类推) 对规范错觉的回归分析, 不论是在实验一中 ( $\beta = -0.03, t = -0.76, p = 0.450$ ) 还是实验三中 ( $\beta = -0.03, t = -0.70, p = 0.486$ ), 不公平程度对规范错觉的回归系数都是不显著的, 因此我们认为在实验二和实验四中选择任意一种不公平分配结果都是可以接受的, 而不会对分析结论有实质性的影响。

最后, 现有文献中有大量研究选择 8-2 作为一个典型的不公平分配方案 (e.g., Chen et al., 2020;

Csukly et al., 2011; Przepiorka & Liebe, 2016), 这也在一定程度上说明该选择的合理性, 较符合主流的研究范式, 因此, 在实验二和实验四中最终遵照前人的研究范式, 选择了 8-2 作为不公平分配方案的代表。

### 实验二正文 3.2 部分补充内容如下:

但实验 2 中被试只需针对 8-2 的分配方案进行估计和实施惩罚, 主要是出于以下两个方面的考虑: 1) 若此处选择重复博弈, 前一轮博弈的规范提示可能会影响后一轮博弈中提示前的惩罚和估计; 2) 实验 1 的结果显示, 对于三种不公平分配方案而言, 分配方案的不公平程度在很大程度上不会影响规范错觉的强弱: 我们在实验 1 三种不公平分配方案中 (10-0、9-1 和 8-2) 做了不公平程度 (编码方式同实验 1, 即 10-0 记为 5, 表示完全不公平, 以此类推) 对规范错觉的回归分析, 结果显示不公平程度对规范错觉没有显著影响 ( $\beta = -0.03$ ,  $t = -0.76$ ,  $p = 0.450$ ), 因此本文参考现有文献选取了 8-2 的分配结果 (Chen et al., 2020; Przepiorka & Liebe, 2016)。

### 实验四正文 5.2 部分补充内容如下:

但实验 4 中被试只需针对 8-2 的分配方案进行估计和实施惩罚, 一方面是因为实验 3 中不公平分配方案的不公平程度基本不会影响规范错觉的强弱 ( $\beta = -0.03$ ,  $t = -0.70$ ,  $p = 0.486$ ), 另一方面也是为了和实验 2 的结果相呼应。

### 参考文献

- Chen, H., Zeng, Z., & Ma, J. (2020). The source of punishment matters: Third-party punishment restrains observers from selfish behaviors better than does second-party punishment by shaping norm perceptions. *PloS One*, 15(3), e0229510.
- Csukly, G., Polgár, P., Tombor, L., Réhelyi, J., & Kéri, S. (2011). Are patients with schizophrenia rational maximizers? Evidence from an ultimatum game study. *Psychiatry Research*, 187(1-2), 11–17.
- Dumas, T. M., Davis, J. P., & Neighbors, C. (2019). How much does your peer group really drink? Examining the relative impact of overestimation, actual group drinking and perceived campus norms on university students' heavy alcohol use. *Addictive Behaviors*, 90, 409–414.
- Enge, S., Mothes, H., Fleischhauer, M., Reif, A., & Strobel, A. (2017). Genetic variation of dopamine and serotonin function modulates the feedback-related negativity during altruistic punishment. *Scientific Reports*, 7, 2996.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785–791.
- Jordan, J. J., McAuliffe, K., & Rand, D. G. (2016). The effects of endowment size and strategy method on third party punishment. *Experimental Economics*, 19(4), 741–763.
- Kenney, S. R., Ott, M., Meisel, M. K., & Barnett, N. P. (2017). Alcohol perceptions and behavior in a residential peer social network. *Addictive Behaviors*, 64, 143–147.
- Larimer, M. E., Parker, M., Lostutter, T., Rhew, I., Eakins, D., Lynch, A., ... Duran, B. (2020). Perceived

descriptive norms for alcohol use among tribal college students: Relation to self-reported alcohol use, consequences, and risk for alcohol use disorder. *Addictive Behaviors*, 102, 106158.

Przepiorka, W., & Lieke, U. (2016). Generosity is a sign of trustworthiness—the punishment of selfishness is not. *Evolution and Human Behavior*, 37(4), 255-262.

.....

**审稿人 2 意见：**

该研究具有一定的理论意义和创新性，研究设计严谨，数据分析和论文写作规范，总体质量较高。当然，文章也有一些值得完善的地方，列举如下，供作者参考。

**意见 1：**该研究基本是“从文献中来，到文献中去”。作者需要思考、论述、明确指出该研究主题对于社会或实际生活的意义。

**修改说明：**感谢专家的宝贵建议，考虑到篇幅问题，我们确实没有在文中充分展示该研究对于社会或实际生活的意义。我们将从第三方惩罚现实意义与本文研究结论的现实意义两个角度分别回答您提出的问题。

**①第三方惩罚的现实意义。**一方面，来自实验室的研究表明，第三方惩罚很好地解释了在单次互动中合作为什么会频繁发生（Fehr & Gächter, 2002），这可能是由于惩罚的存在能够提升潜在违规者的惩罚预期，及其对他人合作水平的预期，进而提升了合作水平（Lergetporer, et al., 2014）。在集体行动中可能表现为：1）搭便车行为会遭到严厉的惩罚，这降低了个体搭便车的动机；2）个体认为他人也会因此减少搭便车行为，这又进一步提升了个体的合作行为；3）惩罚激活了受罚者内化的合作规范，从而提高了在后续互动中的合作水平（陈思静，邢懿琳等，2021）。另一方面，来自现实社会的研究同样发现了上述现象。Kosfeld 和 Rustagi（2015）的研究以公地管理为背景，说明了以实现公平和效率为目标的惩罚在很大程度上提高了团队的合作水平；Mathew 和 Boyd（2011）则以东非游牧民族的 88 次冲突为研究对象，在一定程度上说明了对搭便车者（战争中的逃兵）的惩罚能够维系群体内部的高度合作。综合上述两个方面，我们有理由认为惩罚能够在一定程度上提升人类社会的合作水平，而本文以第三方惩罚为研究对象，具有重要的现实意义。关于这一部分的论述，我们已经在引言中进行了补充（由于这部分的补充内容与意见 2 的补充内容存在一些相似之处，因此专家可参看正文引言部分，或意见 2 的修改部分）。

**②本文研究结论的现实意义。**总体而言，本研究可以概括为：在第三方惩罚中引入了规范错觉，并首次证明了基于违规行为的第三方惩罚中存在着低估的规范错觉，即个体倾向于低估他人的惩罚水平（实验 1），这种低估进一步降低了个体的惩罚行为（实验 2）；较低的公正

世界信念带来了低估的规范错觉（实验 3 和实验 4），但随着社会距离的拉近，公正世界信念对规范错觉的影响不再显著（实验 3）。现实意义如下：

**1) 从提升惩罚行为的可能性来看。**现有文献显示 (Amialchuk et al., 2019; Dumas et al., 2019)：

对他人积极行为的低估减少了个体实施这类行为的可能性。结合本文实验 1 研究结论（即第三方惩罚中存在着低估的规范错觉），这也带来了如下启示：如果我们能够通过某种干预方法，改变第三方惩罚中规范错觉的方向，或许能够在一定程度上提升惩罚水平，进而维护社会规范。

**2) 从不同干预方式的可操作性来看。**（1）实验 2 从规范提示的角度为我们提供了第一种干预方式：当被试得知自己低估了他人的惩罚后，会相应地产生高估的规范错觉，进而提升惩罚水平。遗憾的是，现实中我们无法针对每一次违规行为进行规范提示。Bursztyn 等（2020）的研究表明，对女性劳动参与支持程度的规范提示，增加了被试让其妻子报名参与驾驶课程的可能性。这也意味着规范提示可能存在溢出效应，因此一种潜在的替代方案是：对现实生活中见义勇为的行为进行及时的宣传报道，进而影响其他违规情境下人们的规范错觉与惩罚行为。但这一方案在第三方惩罚中是否真正可行，还需进一步的研究才能得出结论。（2）实验 4 则为我们提供了另一种较为简便的干预方式：当被试阅读完相应的文字材料、激活公正世界信念后，规范错觉的方向由低估转为高估，且惩罚行为也得到了提升。这也意味着即便在现实中人们倾向于低估他人的惩罚水平，我们仍然可以通过恰当的宣传教育提升人们的公正世界信念 (Correia et al., 2007; Jiang et al., 2018)，进而影响规范错觉和惩罚行为。

**3) 从干预方式的应用场景来看。**实验 3 中社会距离的调节效应指明了第二种干预方式的适用场景，即公正世界信念主要是在相对陌生的环境中对规范错觉和惩罚行为发挥作用。一个合理的猜想是，公正世界信念的提升是否会使人们在公共场所（如商场、火车站）目睹违规行为时，感知到更高的惩罚水平，进而提升惩罚行为？当然，这一猜想本文无法直接给出结论，需要未来更多的研究才能给出明确答案。

**这一部分的文字补充在正文 6.1 部分：**

本研究主要有以下三方面的现实意义。第一，实验 2 表明，通过规范信息提示我们可以在一定程度上影响规范错觉的方向与程度，这意味着在现实中规范信息干预 (normative information intervention) 可能是影响被试规范错觉进而校正被试行为的一种可行手段。第二，实验 2 和实验 4 的结果为我们提供了两种潜在的干预方式。一方面，虽然在现实中，我们难以针对每一次违规行为进行规范提示，但由于规范提示存在溢出效应 (Bursztyn et al., 2020)，我们可以通过对现实生活中见义勇为的行为进行及时的宣传报道，进而影响其他违规情境下

人们的规范错觉与惩罚行为；另一方面，一种较为简便的干预方式是：通过恰当的宣传教育提升人们的公正世界信念（Correia et al., 2007; Jiang et al., 2018），进而影响规范错觉和惩罚行为。第三，社会距离的调节效应也指出了第二种干预方式的适用范围，即公正世界信念主要是在相对陌生的环境中对规范错觉和惩罚行为发挥作用，因此一个合理的猜想是，在以陌生人为主的公共场合中提升公正世界信念的措施可能会发挥更好的作用。

**意见 2：**作者在引言第一段论述了合作与第三方惩罚的关系，但是阅读下来感觉并没有论述清楚。例如，为什么“第三方惩罚在一定程度上解释了人类社会中的合作行为”？现在第一段的逻辑比较跳跃，建议梳理出一个更通顺的逻辑，并且最好用实际生活中的例子来解释“第三方惩罚”的内涵、应用范围、研究价值等，否则本文的读者将仅限于熟悉“第三方惩罚”概念的“专家”。

**修改说明：**感谢专家富有建设性的建议，您的建议有助于我们更好地将本文呈现给不同领域的读者。在您的建议下，我们重新阅读了引言的第一段，发现确实存在您指出的问题，因此我们重新撰写了这部分的文字，修改后的文字如下（位于正文引言部分第 1 段和第 2 段）：

Fehr 和 Gächter（2002）提出的第三方惩罚（third-party punishment）——由利益无关者对违规者实施的惩罚——在一定程度上解释了人类社会中的合作行为：存在惩罚机制的情况下，1) 个体预期搭便车行为会遭到严厉的惩罚，这降低了个体搭便车的动机（Molho et al., 2020）；2) 个体认为他人也会因此减少搭便车，这又进一步提升了个体的合作行为（Lergetporer et al., 2014）；3) 惩罚激活了受罚者内化的合作规范，从而提高了在后续互动中的合作水平（陈思静，邢懿琳等，2021）。来自现实社会的研究同样支持上述观点：不论是在公共资源管理中（Kosfeld & Rustagi, 2015），还是在战争与冲突中（Mathew & Boyd, 2011），惩罚都在很大程度上提高了团队的合作水平。

**意见 3：**本文应该是有 4 个假设。研究问题 RQ1 和 RQ2 也应该有对应的 H1 和 H2（或者不需要写 RQ，直接写假设即可），那么原来的 H1 和 H2 就是 H3 和 H4。另外，最后一个假设需要有方向，不能只说“调节了”，需要具体说明假设的调节作用方向。

**修改说明：**感谢专家的所指出的问题，我们已经遵照您的建议，将原文的 RQ1 和 RQ2 改为 H1 和 H2，也将 H1 和 H2 相应地调整为了 H3 和 H4，并在正文中调整了相应的表述。同时也在最后一个假设处加入了调节作用的方向，原文处相应地修改为：“社会距离调节了公正世界信念对规范错觉的影响，具体表现为社会距离的拉近削弱了公正世界信念对规范错觉的



促进作用。”

**意见 4:** 实验 3 提出的有调节的中介模型，调节变量只调节了 X->M 路径，但这个模型也存在可能的竞争模型，比如调节变量同时也调节 X->Y 路径甚至 M->Y 路径。建议论述为什么只假设这个模型，也可以实际检验竞争模型，使结论更有说服力。

**修改说明:** 感谢专家的宝贵建议，您的建议极大地提高了本文数据分析的严谨性。在实验 3 提出的有调节的中介模型中，我们只检验了调节变量（社会距离）对 X->M 路径的影响，主要是因为本文重点关注社会距离在公正世界信念影响规范错觉这一路径中的调节作用。诚如您所言，该模型也存在可能的竞争模型，即社会距离可能会调节 X->Y 路径或者 M->Y 路径。因此根据您的建议，我们使用 PROCESS3.3 (Bootstrap  $N = 5000$ , Model = 59) 检验了调节变量对不同路径的调节效应，结果显示：将惩罚作为因变量时，自变量 $\times$ 调节变量（即公正世界信念 $\times$ 社会距离， $\beta = -0.03 \sim 0.01$ ,  $t = -0.71 \sim 0.26$ ,  $p = 0.480 \sim 0.912$ ）、中介变量 $\times$ 调节变量（即规范错觉 $\times$ 社会距离， $\beta = -0.11 \sim 0.05$ ,  $t = -1.82 \sim 0.91$ ,  $p = 0.071 \sim 0.453$ ）的回归系数在三种不公平方案下均不显著，这也在很大程度上说明了，调节变量（社会距离）对 X->Y 路径以及 M->Y 路径难以起到调节作用。我们也在正文中添加了关于上述结果论述（具体见正文 4.4.2 第 2 段）：

我们进一步检验了该模型的竞争模型，即同时考虑社会距离对公正世界信念-惩罚行为、规范错觉-惩罚行为的调节效应，使用 PROCESS3.3 (Bootstrap  $N = 5000$ , Model = 59) 进行分析后结果显示：将惩罚作为因变量时，公正世界信念 $\times$ 社会距离（ $\beta = -0.03 \sim 0.01$ ,  $t = -0.71 \sim 0.26$ ,  $p = 0.480 \sim 0.912$ ）、规范错觉 $\times$ 社会距离（ $\beta = -0.11 \sim 0.05$ ,  $t = -1.82 \sim 0.91$ ,  $p = 0.071 \sim 0.453$ ）的回归系数在三种不公平方案下均不显著，这在很大程度上说明社会距离在上述两个路径中不存在调节效应。

**意见 5:** 实验 4 多重比较，请用 Bonferroni 或其他 p 值校正方法，LSD 相当于没有对多重比较进行 p 值校正。

**修改说明:** 感谢专家的宝贵建议，我们遵循您的建议，在实验 4 的多重比较中使用了 Bonferroni 法，并在文中进行了修改（见正文 5.4 第 2 段）：

多重比较（Bonferroni 法）的结果表明，激活组的公正世界信念显著高于抑制组（ $p < 0.001$ ,  $95\%CI = [0.46, 0.95]$ ）和对照组（ $p = 0.002$ ,  $95\%CI = [0.12, 0.63]$ ），抑制组则显著低于对照组（ $p = 0.006$ ,  $95\%CI = [-0.59, -0.08]$ ），这说明我们对被试公正世界信念的操纵是有效的。此

外，抑制组中存在显著低估的规范错觉，而激活组的规范错觉不显著（尽管单纯从数字上来看，存在高估的规范错觉），且两组间存在显著差异（ $p < 0.001$ , 95%CI = [0.24, 0.79]）。最后，激活组的惩罚水平也显著高于抑制组（ $p = 0.042$ , 95%CI = [0.01, 0.62]）。

**意见 6：**作者在研究 3-4 里面考察了公正世界信念的作用，但只在研究 3 里面检验了一个有调节的中介模型（条件过程模型）。因此，副标题“基于公正世界信念的条件过程模型”这一表述是否足够准确？尤其是，“基于公正世界信念”会让读者以为公正世界信念是中介变量，但实际上在模型中它是自变量。而且，既然标题中提到了条件过程模型，读者自然好奇调节变量是什么。同时，主标题和副标题都没有提及惩罚行为，但无论是研究 3-4，还是所谓的条件过程模型，因变量都是惩罚行为而非规范错觉。总之，标题需要再考虑调整一下，使之更准确、更严谨。

**修改说明：**非常感谢专家所指出的问题，如您所言，本文的标题确实存在一些问题：在条件过程模型中，公正世界信念作为本文的自变量，采用“基于公正世界信念的条件过程模型”这一表述稍有不妥。本文的逻辑主要是：1）在第三方惩罚中发现了规范错觉的存在（实验 1），且规范错觉在一定程度上会影响惩罚行为（实验 2）；2）公正世界信念是影响规范错觉的重要原因（实验 3 和实验 4）。因此我们遵循您的建议，并且经过反复思考，最终将标题确定为“第三方惩罚中的规范错觉：基于公正世界信念的解释”。

**意见 7：**文章写作方面，请减少冗长、复杂、隐晦、绕来绕去的表述。

**修改说明：**感谢专家的宝贵建议，这有利于我们更清楚地表达本文的观点。遵循您的建议，我们修改了文章中存在的一些表述问题，减少了一些冗长和复杂的表述，再次感谢您的宝贵意见！

**意见 8：**第二段“个体的惩罚行为与感知到的惩罚强度呈正相关，换言之，感知惩罚规范将直接影响个体的惩罚行为”，需要注意“正相关”并不能换言之“直接影响”，相关不是因果。

**修改说明：**感谢专家的细心阅读和宝贵建议。我们在此处的表述确实存在不妥之处，我们已经在原文处进行了修改，删去了“换言之”之后的内容，修改后的文字如下：“[Kamei \(2020\)](#)指出了感知惩罚规范（即对他人惩罚行为的估计）的重要性，他们的研究表明，个体的惩罚行为与感知到的惩罚强度呈正相关。”

**意见 9：**建议在表 1 表头部分加一行，标明“先惩罚后估计”、“先估计后惩罚”、“只惩罚”、“只估计”；图 1~3 最好也标出这些信息，不要只用代号去区分。

**修改说明：**非常感谢专家的宝贵建议。我们遵循您的建议，将图 1-3 的标题修改为了：“先惩罚后估计中估计 1 和惩罚 1 的比较”，“先估计后惩罚中估计 2 和惩罚 2 的比较”，“只估计和只惩罚中估计 4 和惩罚 3 的比较”。此外，我们还在表 1 的上方增加了一行，用以标明组别，修改后的表 1 如下所示：

表 1 四种实验条件下惩罚和估计的描述性统计 ( $M \pm SD$ )

组别	先惩罚后估计		先估计后惩罚		只惩罚	只估计
方案	惩罚 1	估计 1	惩罚 2	估计 2	惩罚 3	估计 4
10-0	4.19 $\pm$ 0.98	3.94 $\pm$ 1.08	4.14 $\pm$ 1.05	3.90 $\pm$ 1.09	4.28 $\pm$ 0.97	3.85 $\pm$ 1.11
9-1	3.46 $\pm$ 1.01	3.22 $\pm$ 1.01	3.48 $\pm$ 0.95	3.27 $\pm$ 0.99	3.50 $\pm$ 0.94	3.13 $\pm$ 1.03
8-2	2.82 $\pm$ 0.94	2.63 $\pm$ 0.91	2.80 $\pm$ 0.85	2.64 $\pm$ 0.93	2.83 $\pm$ 0.77	2.69 $\pm$ 0.94
7-3	2.00 $\pm$ 0.91	1.97 $\pm$ 0.85	2.16 $\pm$ 0.77	2.07 $\pm$ 0.83	2.01 $\pm$ 0.79	2.00 $\pm$ 0.90
6-4	1.06 $\pm$ 0.99	1.16 $\pm$ 0.83	1.18 $\pm$ 0.81	1.17 $\pm$ 0.79	1.17 $\pm$ 0.85	1.05 $\pm$ 0.71
5-5	0.11 $\pm$ 0.44	0.16 $\pm$ 0.50	0.20 $\pm$ 0.50	0.19 $\pm$ 0.44	0.18 $\pm$ 0.48	0.17 $\pm$ 0.42

**意见 10：**图 4、图 6 需要添加坐标轴标题。

**修改说明：**感谢专家的细心阅读，您的建议有助于我们更好地向读者传递本文的研究结果。图 4 的横坐标为规范错觉，纵坐标为惩罚行为；图 6 的横坐标为社会距离，纵坐标为公正世界信念对规范错觉的回归系数。我们已经根据您的建议在图 4 中增加了坐标轴标题，在图 6 中增加了坐标轴的说明。

**意见 11：**作者在配对样本 t 检验中均报告了 r 值，指的是两个重复测量之间的相关性。但是一般来说，这个指标并不需要汇报，报告了反而容易造成读者误解（比如理解成效应量 r）。如果一定要报告，也需要区分开或做必要的标注，因为目前结果报告里面 r 和后面的 t、d 等统计量是放在一起的，不做区分则会有歧义。

**修改说明：**非常感谢专家的细心和严谨。遵循您的建议，我们删去了配对样本 t 检验中的 r 值。

**意见 12:** 表 6、表 9 中的“主效应”“中介效应”表述不当。规范错觉不是中介效应，而是中介变量，这里的结果是路径  $b$  而非  $a*b$ 。其次，由于正文已经报告了表 6 中的这几个系数和  $R^2$  等结果，建议不必再放表 6（属于冗余信息）。

**修改说明:** 感谢专家的宝贵建议，诚如您所言，我们在正文中已经报告了表 6 中的信息，因此遵循您的建议，我们已经将表 6 删去，并相应地调整了下文中表格的序号以及文中的表述。此外，我们已经将表 9（调整序号后为表 8）中的主效应修改为了“自变量”，而将中介效应修改为了“中介变量”。

**意见 13:** 所有回归分析表需要在标题或表头部分写明因变量。

**修改说明:** 非常感谢专家的建议，我们仔细检查了本文的回归分析表（含表 2、表 3、表 4、表 7、表 8 和表 10），其中表 8 和表 10 未写明因变量。这两个表格是对中介效应和调节效应的检验，因变量分别为惩罚行为与规范错觉，遵循您的建议，我们在标题后添加了因变量。

**意见 14:** 全文在报告贝叶斯因子时，如果是为了支持零假设，可以报告  $BF_{01}$ （即  $BF_{10}$  的倒数），这样更直观。

**修改说明:** 感谢专家的宝贵建议，遵循您的建议，我们已经重新在 JASP0.14.1 中重新计算了  $BF_{01}$ ，并将正文中的  $BF_{10}$  修改为了  $BF_{01}$ ，也相应地修改了对这一结果的解释，例如“因此我们用 JASP0.14.1 计算了相应的贝叶斯因子，结果显示当前数据更加支持零假设。”

## 参考文献

- 陈思静, 邢懿琳, 翁异静, 黎常. (2021). 第三方惩罚对合作的溢出效应: 基于社会规范的解释. *心理学报*, 53(7), 758–772.
- Amialchuk, A., Ajilore, O., & Egan, K. (2019). The influence of misperceptions about social norms on substance use among school-aged adolescents. *Health Economics*, 28(6), 736–747.
- Bursztyn, L., González, A. L., & Yanagizawa-Drott, D. (2020). Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *The American Economic Review*, 110(10), 2997–3029.
- Correia, I., Vala, J., & Aguiar, P. (2007). Victim's innocence, social categorization, and the threat to the belief in a just world. *Journal of Experimental Social Psychology*, 43(1), 31–38.
- Dumas, T. M., Davis, J. P., & Neighbors, C. (2019). How much does your peer group really drink? Examining the relative impact of overestimation, actual group drinking and perceived campus norms on university students' heavy alcohol use. *Addictive Behaviors*, 90, 409–414.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.
- Jiang, R., Liu, R. D., Ding, Y., Zhen, R., Sun, Y., & Fu, X. (2018). Teacher justice and students' class identification: Belief in a just world and teacher-student relationship as mediators. *Frontiers in psychology*, 9, 802.

- Kosfeld, M., & Rustagi, D. (2015). Leader punishment and cooperation in groups: Experimental field evidence from commons management in Ethiopia. *American Economic Review*, 105(2), 747–783.
- Lergetporer, P., Angerer, S., Glätzle-Rützler, D., & Sutter, M. (2014). Third-party punishment increases cooperation in children through (misaligned) expectations and conditional cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), 6916–6921.
- Mathew, S., & Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences of the United States of America*, 108(28), 11375–11380.
- Molho, C., Tybur, J. M., Van Lange, P. A., & Balliet, D. (2020). Direct and indirect punishment of norm violations in daily life. *Nature Communications*, 11(1), 3432.
- 

## 第二轮

**审稿人 2 意见:**修改稿较原稿进行了针对性修改，所提及问题都有了明显改善，文章具有一定的创新性和理论及应用意义，总体质量较高，同意发表。

**回复:**感谢专家的肯定，这极大地鼓舞了我们！

.....

**编委意见:**文章的理论意义明确，具有一定的创新性。研究设计严谨，对修改意见有较好的回应，修改稿有较大的提升。同意发表。

**回复:**感谢编委老师对我们的肯定，这对我们是很大的鼓励！

.....

## 主编意见:

该文章的研究问题具有较好的创新性，且实验设计严谨丰富，样本量大。作者较好地回应了编委老师和外审专家的问题，但仍有以下问题可以再进行思考：

**意见 1:**研究者在进行错觉方向的假设（H1）时其实是互斥假设，因为在假设推导过程中并没有明确驳斥某一方向假设的成立，但在具体写假设时却只写了低估规范这一种假设。当然这可能是因为提出后续假设的方便，但仍会带给读者数据驱动的感受。是否可以考虑将 H1 写成互斥假设，然后根据不同方向假设成立的可能性对后续假设进行陈述？

**修改说明:**感谢主编老师的宝贵建议，您的建议极大地提升了本文提出问题的合理性，也增强了本文的可读性！诚如您所言，我们在进行实验研究前，并未能从已有文献中获取充分的证据来证明第三方惩罚中的规范错觉是低估的。而我们直接在假设 1 中指明个体倾向于低估规范，在表述上确实存在一定的问题。因此遵循您的建议，站在读者的角度进行思考后，我

们选择以研究问题的形式提出原先的假设 1 和假设 2，并将假设 3 和假设 4 修改为假设 1 和假设 2，并调整了正文中的相关表述。修改后的研究问题和假设如下所示：

**RQ1:** 第三方惩罚中是否存在规范错觉？如果存在的话，规范错觉的方向如何？

**RQ2:** 规范错觉如何影响个体自身的惩罚行为？

**RQ3:** 违规行为的严重程度是否影响了人们对惩罚的规范错觉？

**H1:** 公正世界信念影响了第三方惩罚中的规范错觉。

**H2:** 社会距离调节了公正世界信念对规范错觉的影响。

**意见 2:** 针对人们对积极行为，如亲环境行为具有低估的错觉，可以再举一些其他积极行为的例子，并加以机制的解释。

**修改说明:** 非常感谢主编老师的建议，遵循您的建议，我们在引言第 3 段补充了积极行为的例子，包括：1）青少年低估了其同龄人对校园欺凌行为的反对程度（Dillon & Lochman, 2019）；2）沙特阿拉伯的已婚男性低估了其他相似男性对女性劳动参与的支持（Bursztyn et al., 2020）。此外，我们还增加了关于产生这类规范错觉的机制探讨。修改后的内容如下所示（见正文引言第 3 段）：

出于对积极自我形象的追求（Leary, 2007），个体往往存在着自我服务偏差（self-serving bias）（Zhang et al., 2018），对他人的贬损恰好可以满足这一心理需求（Rau et al., 2019）。这也使得规范错觉总是表现为一种系统性的认知偏差：人们倾向于低估他人的积极行为，例如公众低估了他人的亲环境行为（Bouman et al., 2020; Bouman & Steg, 2019），青少年低估了同龄人对校园欺凌行为的反对程度（Dillon & Lochman, 2019），沙特阿拉伯的已婚男性低估了其他男性对女性劳动参与的支持（Bursztyn et al., 2020）；而倾向于高估他人的消极行为，如食物浪费（陈思静，濮雪丽等, 2021）和酗酒（Amialchuk et al., 2019; Dumas et al., 2019）。

**意见 3:** 关于采用规范错觉程度（PO-MPS）还是感知到的规范程度（PO）作为变量，虽然作者在编委初审意见中进行了回复，但从文章内容来看读者可能仍会产生这样的疑问，不知道是否有方式可以在文中主动回应一下作者的潜在疑问？错觉并非是一个个体能够直接感知到的变量，个体只能对规范本身有所感知，而意识不到自己的感知谬差，虽然只是是否减去恒定值 MPS 的问题，但 PO 似乎更容易理解。

**修改说明:** 感谢主编老师的宝贵建议，您的建议有助于我们更好地向读者阐述文章的思路！由于我们修改过程中的疏忽，在回复完编委老师的意见后，未能在正文中增加对这一问题的



阐述。为了进一步增强文章的可读性，遵循您的建议，在上一轮意见回复的基础之上，我们在实验 1 的 2.2 部分补充了相关的论述，并在研究不足中进行了说明。

#### 正文 2.2 部分补充：

需要说明的是，本文采用规范错觉（ $PO - M_{PS}$ ）而非感知规范（ $PO$ ）作为关键变量，主要是出于以下两方面的考虑。第一，不论是从统计还是定义的角度来看，规范错觉和感知规范都能够较好地解释个体的惩罚行为。第二，规范错觉相较于感知规范而言，可以表达出更为丰富的内涵：1）规范错觉能够更好地展示个体对他人惩罚水平的估计是否与实际水平发生偏离，这也是本文的重点关注之一；2）在解释惩罚行为时，采用规范错觉可以更直观地展示当规范错觉存在或不存在时个体惩罚行为的变化。

#### 正文 6.2 部分补充：

第四，在本文的四个实验中，都采用了平均惩罚来代替真实的惩罚规范，进而计算了被试的规范错觉，这也是目前研究中较为普遍的方式（e.g., 陈思静，濮雪丽等, 2021; Duong & Parker, 2018; Sandstrom et al., 2013）。但平均惩罚和真实的惩罚规范之间可能存在细微的差异，如何更好地衡量存在于集体中的真实规范，是值得未来研究者思考的一个问题。

**意见 4：**低估错觉的产生具有边界条件，例如文中所多次关注的公平程度，但综述部分缺失对这部分探讨。

**修改说明：**非常感谢主编老师所指出的问题，如您所言，在第三方惩罚中产生规范错觉需要满足一定的条件。虽然本文重点关注的是规范错觉的方向，但由于作为亲社会行为的第三方惩罚针对的应当是某类违规行为，因此从理论上来说，分配方案的公平程度也会影响错觉是否存在。而实验 1 之所以设置了 6 种分配方案（从完全不公平的 10-0 分配，到完全公平的 5-5 分配），也是出于这一考虑。目前在综述部分的确缺少了对这一问题的探讨。重新阅读这部分文字后，我们在综述中增加了相应的表述并提出了新的研究问题，修改后内容如下：

另一方面，违规行为的程度有高有低（Balafoutas et al., 2016），例如在独裁者博弈中，分配方案的公平程度在一定程度上体现了分配者是否违反社会规范以及违规的程度（Csukly et al., 2011）。那么一个值得探讨的问题是违规程度是否影响了人们对他人惩罚的规范错觉，换言之，规范错觉的存在是否具有一定的边界条件？基于此，本文提出第三个研究问题：

**RQ3：**违规行为的严重程度是否影响了人们对惩罚的规范错觉？

**意见 5：**应再加强对 H3 的推导逻辑，现在的表述还没有落到规范错觉上。

**修改说明：**感谢主编老师一针见血的点评，这大大增强了本文研究假设提出的逻辑性！诚如您所言，由于我们表述上的问题，在提出 H3（现为 H1）时只是强调了公正世界信念可能会对惩罚的规范感知产生影响，并未落到规范错觉上。遵循您的建议，我们增加了相应的表述（见正文引言部分）：

换言之，公正世界信念影响了人们对惩罚的规范感知，而存在于集体中的真实规范在某个时点上相对稳定的（Laniga-Wijnen et al., 2018），因此公正世界信念对惩罚规范感知的影响也可以视为对规范错觉（规范感知与真实规范之间的差异，具体定义参见 2.2 部分）的影响。基于上述推理，本文提出以下假设，并在实验 3 和实验 4 中进行检验：

**H1：**公正世界信念影响了第三方惩罚中的规范错觉。

**意见 6：**H4 表述有问题，公正世界信念对规范错觉不是促进作用，因为这里默认的错觉方向是低估。

**修改说明：**非常感谢主编老师指出的问题，您的建议对提升本文的严谨性有很大的帮助！根据您的意见 1，同时也为了增强文章的合理性和可读性，我们在提出研究问题 1 时并未明确规范错觉的方向。为了与研究问题 1 相呼应，我们将 H4（现为 H2）修改为：“社会距离调节了公正世界信念对规范错觉的影响。”

**意见 7：**研究一的结果有如下问题需要考虑：1) 描述统计部分可以先讲样本总体的描述性统计的结果，再讲不同人口统计学变量之间的差异，且分配方案之间的比较应放在人口统计学比较之后比较有条理。2) 表 1 下面一段中提到“个体对他人平均惩罚的估计显著低于自身实施的惩罚，即存在低估的规范错觉”表述有误，应将“自身”改为“实际的”；3) 报告多个 p 值时应使用 ps。4) 表 2 中整体模型的 R2 也是显著的，应该进行标注 5) 既然前面已经验证估计和做出行为决策的顺序没有影响，为何表 3 和表 4 不直接合并样本？6) 讲完结论之后不应该又呈现结果，读起来不顺畅，可以调整一下顺序。

**修改说明：**非常感谢主编老师的细心阅读，您的建议帮助我们更好地呈现本文的研究结果。根据您的建议，我们对研究一进行了逐一修改，以下是我们的修改说明。

1) 遵循您的建议，在描述性统计方面我们进行了如下修改：①我们在实验 1 的描述性统计部分添加了总体样本的描述性统计结果，并将这部分的表述放在了基于人口统计学变量的比较之前；②针对性别的平均数差异检验显示，整体上男性对他人惩罚的估计高于女性，而这种差异主要体现在 9-1 分配与 7-3 分配中，根据您的建议，我们适当调整了这部分结果在文

中的表述，先陈述了总体上的差异，然后进行了分方案的比较。修改后的原文如下（见正文 2.4 第 1 段）：“总体上被试的平均惩罚为  $2.31 \pm 1.61$ ，平均估计为  $2.17 \pm 1.53$ 。基于性别的平均数差异检验表明，男性和女性的惩罚水平不论是在整体上（ $t(2050) = 0.14, p = 0.887$ ），还是在不同方案中（ $t(340) = -0.51 \sim 0.75, ps = 0.455 \sim 0.988$ ），均不存在显著差异。然而男性对他人惩罚的估计显著高于女性（ $t(1972) = 2.16, p = 0.031, d = 0.10, 95\%CI = [0.01, 0.29]$ ），这种差异主要体现在 9-1 分配（ $t(327) = 2.42, p = 0.016, d = 0.27, 95\%CI = [0.05, 0.49]$ ）和 7-3 分配中（ $t(327) = 2.02, p = 0.045, d = 0.23, 95\%CI = [0.005, 0.39]$ ）。”

- 2) 根据您的建议，我们将表 1 下方第一段的表述改为了“个体对他人平均惩罚的估计显著低于实际的惩罚水平”。
- 3) 我们仔细检查了全文，并将涉及多个  $p$  值的地方都用  $ps$  进行了表示。
- 4) 我们仔细检查了全文，并补充了表 2 以及正文其他表格中  $R^2$  的显著性水平。
- 5) 根据您的建议，我们将组 1 和组 2 的数据进行了合并。合并后的回归结果如下表所示：

表 3 规范错觉对惩罚行为的回归分析

方案	$B$	$SE$	$\beta$	LLCI	ULCI	Adj- $R^2$
10-0	0.75***	0.04	0.80	0.68	0.82	0.64***
9-1	0.76***	0.04	0.78	0.68	0.84	0.60***
8-2	0.68***	0.05	0.70	0.58	0.77	0.48***

注：  $n = 222$ ， \*\*\*  $p < 0.001$ 。

由于图 4 是表 3（原表 3 和表 4）的补充说明，因此我们也相应地对图 4 进行了修改，修改后的图 4 如下所示：

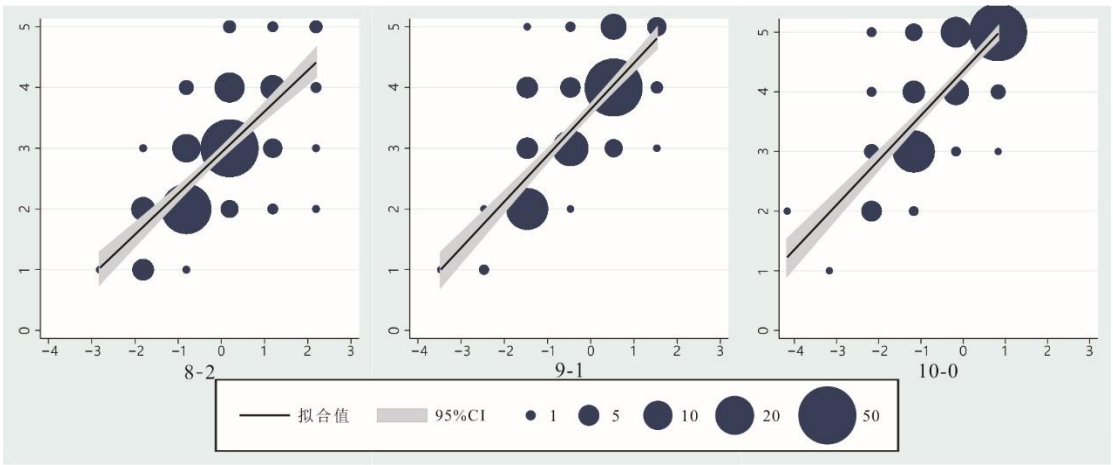


图 4 不同分配方案下规范错觉与惩罚行为的气泡图

6) 由于图 4 是表 3 的补充说明, 因此根据您的建议, 我们选择将图 4 直接放在表 3 后呈现。并将图表前的表述改为了“为了更直观地展示在违规情境中规范错觉对惩罚行为的影响, 我们将组 1 和组 2 合并后, 根据三种不公平的分配方案进行规范错觉对惩罚行为的回归分析, 结果如表 3 所示; 图 4 则十分直观地展现了在不同分配方案下规范错觉(横轴)和自身惩罚行为(纵轴)之间的联系(气泡大小表示该位置上的被试数量)。”

**意见 8:** 研究二的结果有如下问题需要考虑: 1) 3.2 部分“实验 2 为 2 组间因子设计”表述不准确。 2) 这个操纵方法很难说是操纵了什么, 个人感觉不是操纵的错觉, 而是操纵的对规范的感知, 原因同问题 3。

**修改说明:** 感谢主编老师的细心阅读与中肯的建议, 遵循您的建议, 我们对研究二进行了如下修改。

1) 根据您的建议, 我们删除了 3.2 部分原先不准确的表述, 修改后的表述为: “实验 2 根据规范提示的方向设置了两种实验条件”。

2) 诚如您所言, 实验 2 的操纵方法更准确的表述应该是“通过操纵个体对规范的感知, 有针对性地改变了规范错觉”, 或者是“通过有针对性的规范提示, 改变了被试的规范错觉”。我们仔细检查了全文, 发现涉及这一表述的地方共有 4 处: 引言的最后 1 段、实验 2 的第 1 段、研究结论及其意义的第 3 段、研究不足的第 3 点。修改后的表述分别如下所示(绿色部分为此次修改的内容):

#### 引言最后 1 段:

实验 2 则为实验 1 的稳健性检验, 我们通过操纵个体对规范的感知, 有针对性地改变了规范错觉, 进而考察规范错觉与惩罚行为之间的因果关系。

#### 实验 2 第 1 段:

因此实验 2 主要通过有针对性的规范提示来改变被试的规范错觉, 进一步考察两者间的因果关系, 从而与实验 1 的结果形成印证。

#### 研究结论及其意义第 3 段:

为了更好地建立两者的因果关系, 我们还在实验 2 中通过有针对性的规范提示改变了被试的规范错觉。

#### 研究不足第 3 点:

第三, 在实验 2 和实验 4 中, 我们通过有针对性地改变被试的规范错觉以及操纵其公正世界信念, 进而考察了两者对各自结果变量的影响。

**意见 9：研究三：** 1) 可以补充一下个人公正世界信念和一般公正世界信念各自的定义和辨析； 2) 研究一结果显示，选取 10-0、9-1 和 8-2 这三个不公平分配方案时违规情境下不公平程度对规范错觉的回归系数不显著，为何不对三种方案进行合并分析？

**修改说明：**感谢主编老师的建议，这有助于我们更直观地向读者阐述本文的研究结果。以下是研究三的修改说明：

1) 遵循您的建议，我们在 4.2 变量与设计部分对公正世界信念进行了详细的解释，补充了个人公正世界信念和一般公正世界信念的定义。绿色部分是我们补充的内容（见正文 4.2 第 2 段）：

公正世界信念可以划分为两个维度：个人公正世界信念（personal belief in a just world）和一般公正世界信念（general belief in a just world）（Wu et al., 2011）。前者即相信世界对自身而言是公正的，往往预测了个体的主观幸福感（Sutton & Winnard, 2007）；后者则是相信世界对他人而言是公正的，也被称为他人公正世界信念（belief in a just world to others，本文将其统称为一般公正世界信念）（周春燕，郭永玉，2013），主要影响个体对社会现象的判断以及对社会环境的解释（Testé & Perrin, 2013）。由于研究者常将后者与感知惩罚（Bai et al., 2014）和惩罚态度（Bègue & Bastounis, 2003）相联系，因此将一般公正世界信念作为本文的自变量。

2) 诚如您所言，在违规情境下，分配方案的不公平程度并不会影响本文的分析结果。因此遵循您的建议，我们将三种方案进行了合并分析，并重新撰写了正文 4.4.1 和 4.4.2。由于涉及内容过多，烦请主编老师参看正文 4.4.1 和 4.4.2 部分。

**意见 10：**所有报告结果时都应标注检验的自由度。统计方法上有些冗余，有些于逻辑上并没有增益，是可以考虑去掉的。

**修改说明：**非常感谢主编老师的建议，这提升了本文数据分析的严谨性！遵循您的建议，增加了正文中 t 检验和 F 检验的自由度。另外，我们仔细阅读全文后，删去了实验 3 中的稳健性检验部分，尽可能地避免了统计方法上的冗余。

## 参考文献

陈思静，濮雪丽，朱玥，汪昊，刘建伟. (2021). 规范错觉对外出就餐中食物浪费的影响：心理机制与应对策略. *心理学报*, 53(8), 904-918.

- 周春燕, 郭永玉. (2013). 公正世界信念——重建公正的双刃剑. *心理科学进展*, 21(1), 144–154.
- Amialchuk, A., Ajilore, O., & Egan, K. (2019). The influence of misperceptions about social norms on substance use among school-aged adolescents. *Health Economics*, 28(6), 736–747.
- Balafoutas, L., Nikiforakis, N., & Rockenbach, B. (2016). Altruistic punishment does not increase with the severity of norm violations in the field. *Nature Communications*, 7, 13327
- Bouman, T., & Steg, L. (2019). Motivating society-wide pro-environmental change. *One Earth*, 1(1), 27–30.
- Bouman, T., Steg, L., & Zawadzki, S. J. (2020). The value of what others value: When perceived biospheric group values influence individuals' pro-environmental engagement. *Journal of Environmental Psychology*, 71, 101470.
- Bursztyrn, L., González, A. L., & Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *The American Economic Review*, 110(10), 2997–3029.
- Csukly, G., Polgár, P., Tombor, L., Réhelyi, J., & Kéri, S. (2011). Are patients with schizophrenia rational maximizers? Evidence from an ultimatum game study. *Psychiatry Research*, 187(1-2), 11–17.
- Dillon, C. E., & Lochman, J. E. (2019). Correcting for norm misperception of anti-bullying attitudes. *International Journal of Behavioral Development*. DOI: <https://doi.org/10.1177/0165025419860598>.
- Dumas, T. M., Davis, J. P., & Neighbors, C. (2019). How much does your peer group really drink? Examining the relative impact of overestimation, actual group drinking and perceived campus norms on university students' heavy alcohol use. *Addictive Behaviors*, 90, 409–414.
- Duong, H. T., & Parker, L. (2018). Going with the flow. *Journal of Social Marketing*, 8(3), 314–332.
- Enge, S., Mothes, H., Fleischhauer, M., Reif, A., & Strobel, A. (2017). Genetic variation of dopamine and serotonin function modulates the feedback-related negativity during altruistic punishment. *Scientific Reports*, 7, 2996.
- Laninga-Wijnen, L., Harakeh, Z., Dijkstra, J. K., Veenstra, R., & Vollebergh, W. (2018). Aggressive and prosocial peer norms: Change, stability, and associations with adolescent aggressive and prosocial behavior development. *The Journal of Early Adolescence*, 38(2), 178–203.
- Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, 58(1), 317–344.
- Rau, R., Nestler, S., Geukes, K., Back, M. D., & Dufner, M. (2019). Can other-derogation be beneficial? Seeing others as low in agency can lead to an agentic reputation in newly formed face-to-face groups. *Journal of Personality and Social Psychology*, 117(1), 201–227.
- Sandstrom, M., Makover, H., & Bartini, M. (2013). Social context of bullying: Do misperceptions of group norms influence children's responses to witnessed episodes? *Social Influence*, 8(2-3), 196–215.
- Sutton, R. M., & Winnard, E. J. (2007). Looking ahead through lenses of justice: The relevance of just-world beliefs to intentions and confidence in the future. *British journal of social psychology*, 46(3), 649–666.
- Testé, B., & Perrin, S. (2013). The impact of endorsing the belief in a just world on social judgments. *Social Psychology*, 44(3), 209–218.
- Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-serving bias in memories. *Experimental Psychology*, 65(4), 236–244.

### 第三轮

**主编意见：**作者对终审意见的回复较为充分，但仍存在可修改的空间。终审意见第一条中，关于如何提出和表述备择假设，作者现在的修改方法仍不是最完美的，例如现在的假设中没



有体现出对主效应——错觉方向的假设、三条研究问题与假设之间是否存在冗余等。建议作者查找类似的包含备择假设的文献，并参考其处理方式，再次完善假设提出的过程和表述的方式。

**修改说明：**非常感谢主编老师对我们上一轮修改的肯定，以及此次提出的宝贵建议，您的建议有助于我们更清晰地提出研究假设！遵循您的建议，我们阅读了国内外有关心理学、经济学等学科的相关文献（杜勇等, 2017; 姜付秀等, 2017; 张杰, 郑文平, 2018; 胡国柳等, 2019; 张明等, 2020; Anseel & Lievens, 2007; Humberg et al., 2019），参考了他们对假设提出的处理方式，并修改了原文中相应的表达方式，从而使表述更为精简清楚。修改内容包括：

- ①我们将目前的研究问题 1 和 2 以互斥假设的方式重新提出；
- ②考虑到您指出的研究问题之间的冗余，我们删去了涉及分配方案公平程度的研究问题 3，但同时也保留了对规范错觉边界条件的探讨，我们将这部分的内容放到了两组互斥假设提出之后；
- ③对后文中相应的内容也重新进行了修改，但这些修改较为零散，因此不在这里赘述，我们在原文中已经将修改内容标为**橙色**，以方便主编老师审阅。

以下是我们主要进行的修改（从引言第 5 段开始，**橙色**部分表示此次修改的内容）：

我们首先探讨的问题是：第三方惩罚中是否存在规范错觉？如果存在的话，其方向如何？上文的分析表明，规范错觉相关的多数研究认为个体倾向于高估他人的消极行为（陈思静，濮雪丽等, 2021; Amialchuk et al., 2019; Davis et al., 2019），而低估他人的积极行为（Anthenien et al., 2018; Dempsey et al., 2018），那么可以推测惩罚中存在着低估的规范错觉。如果参考来自合作领域的研究，更多地考虑策略性动机的影响，则会得到完全相反的推论。在此基础上，本文提出以下竞争性假设，并在实验 1 中进行检验：

**H1a:** 人们系统性地低估了他人的第三方惩罚（低估的规范错觉）。

**H1b:** 人们系统性地高估了他人的第三方惩罚（高估的规范错觉）。

需要进一步思考的是，规范错觉将如何影响惩罚行为？Grimm 等（2017）指出，对他人行为的感知往往会影响到自身的行为，这可能是因为人们会通过声称自己的行为符合规范来证明其道德合理性（Schlag et al., 2015），例如对他人亲社会行为的低估会减少自身相应的行为（Ganz et al., 2020），而对他人物质使用行为（包括酗酒、吸烟和吸毒）的高估则增加了从事这类行为的可能性（Amialchuk et al., 2019）。我们推测，规范错觉对惩罚行为也有类似的影响。基于上述推理，本文提出以下竞争性假设，并在实验 1 和实验 2 中进行检验：

**H2a:** 对他人惩罚行为的低估进一步降低了自身的惩罚行为。

**H2b:** 对他人惩罚行为的高估进一步提升了自身的惩罚行为。

需要说明的是，在典型的第三方惩罚中，受罚者往往是因为违反了社会规范而受到他人惩罚，而违反社会规范的程度有高有低（Balafoutas et al., 2016），例如在独裁者博弈中，分配方案的公平程度在一定程度上体现了分配者是否违反社会规范以及违规的程度（Csukly et al., 2011），因此在检验上述两组竞争性假设时，我们还考察了违规程度是否影响人们对他人惩罚行为的错觉，作为对本研究结果的一个重要补充。

## 参考文献

- 杜勇, 张欢, 陈建英. (2017). 金融化对实体企业未来主业发展的影响: 促进还是抑制. *中国工业经济*(12), 113–131.
- 胡国柳, 赵阳, 胡珺. (2019). D&O 保险、风险容忍与企业自主创新. *管理世界*, 35(08), 121–135.
- 姜付秀, 王运通, 田园, 吴恺. (2017). 多个大股东与企业融资约束——基于文本分析的经验证据. *管理世界*, 33(12), 61–74.
- 张杰, 郑文平. (2018). 创新追赶战略抑制了中国专利质量么? *经济研究*, 53(05), 28–41.
- Amialchuk, A., Ajilore, O., & Egan, K. (2019). The influence of misperceptions about social norms on substance use among school-aged adolescents. *Health Economics*, 28(6), 736–747.
- Anseel, F., & Lievens, F. (2007). The relationship between uncertainty and desire for feedback: A test of competing hypotheses. *Journal of Applied Social Psychology*, 37(5), 1007–1040.
- Balafoutas, L., Nikiforakis, N., & Rockenbach, B. (2016). Altruistic punishment does not increase with the severity of norm violations in the field. *Nature Communications*, 7, 13327
- Csukly, G., Polgár, P., Tombor, L., Ráthelyi, J., & Kéri, S. (2011). Are patients with schizophrenia rational maximizers? Evidence from an ultimatum game study. *Psychiatry Research*, 187(1–2), 11–17.
- Humberg, S., Dufner, M., Schönbrodt, F. D., Geukes, K., Hutteman, R., Küfner, A. C., ... Back, M. D. (2019). Is accurate, positive, or inflated self-perception most advantageous for psychological adjustment? A competitive test of key hypotheses. *Journal of Personality and Social Psychology*, 116(5), 835–859.