

《心理学报》审稿意见与作者回应

题目：直觉还是推理？亲社会行为的决策模式与人性感知

作者：王婧嫣，张洪

第一轮

审稿人 1 意见：本文有以下问题需要进行探讨。

意见 1：实验操纵中对高低情绪唤起度的界定不明确，区分不明显。实验 1 只考察高情感唤起情境，采用的实验材料的情绪唤起度 $M=5.75$ ，研究者认为是高情绪唤起。实验 2 和 3 同时考察高低情感唤起情境，采用 t 检验来确定高低组：实验 2 中高情感唤起场景 $M = 5.79$ ，低情感唤起场景 $M = 4.73$ ；实验 3 中高情感唤起组 $M = 5.55$ ，低情感唤起组 $M = 5.16$ 。实验 4 只考察低情感唤起情境，采用的实验材料的情绪唤起度 M 为 5.13，研究者认为是低情绪唤起。既然所有情境的情绪唤起度都超过中心值，之间的差异很小，这似乎削弱了研究情绪唤起变量的重要性。

回应：从描述统计的结果来看高、低情感唤起的差值确实不大，并且即使是低情绪唤起组，其情感唤起的程度也超过了量表的中心值。对于这些问题，我们从以下几个方面进行解释：

1) 所有情境均涉及有人需要帮助的情况，因此即便是相对来说情感唤起较低的情境，仍然比中性情境更能够唤起情感。所以，我们所采用的所有情境，其情感唤起度看上去都比较高，且超过了中心值，这是合理的、可解释的，也在我们的预料之中。

2) 即便如此，我们仍然能够在情感唤醒度上对情境加以区分。我们在正式实验之前通过预实验测量了不同情景的情感唤起度，并通过比较对情景进行了分类，而非在实验中主观臆断情感唤起类型，这些结果在我们的论文中有所报告。从统计上看，实验 2 和 3 都达到了非常显著的差异水平 $ps < .001$ 。

3) 此外，为了更进一步证明我们对于情境的分类是有效的，我们补充收集了一份数据。该数据共包含 100 个有效样本，其中男性 27 名，女性 73 名，平均年龄 $M = 23.13$ 岁， $SD = 6.20$ 岁。此调查为单因素被试内设计，向被试随机呈现本论文实验中用到的 6 个情景，要求他们根据自己的感受对情景的唤醒度进行迫选分类(高情感唤起 vs. 低情感唤起)。结果表明，所有情景的区分度均达到了统计显著的水平，具体结果请参见下表。我们认为，这一补充数据能够表明，被试对我们采用的不同情感唤起水平的情境可以进行高低水平的区分。我们在文中用脚注的形式补充了这一调查的结果。

	分类-高情感唤起(N)	分类-低情感唤起(N)	χ^2	df	p
实验 1-高情感唤起情景	78	22	31.36	1	< .001
实验 2-高情感唤起情景	87	13	54.76	1	< .001
实验 2-低情感唤起情景	29	71	17.64	1	< .001
实验 3-高情感唤起情景	72	28	19.36	1	< .001
实验 3-低情感唤起情景	30	70	16.00	1	< .001
实验 4-低情感唤起情景	37	63	6.76	1	= .009

4) 综上所述，情感唤起在本研究中是可以区分的，也确实是目前研究中较为重要的一

个变量。此外，这一变量的重要性还在于它与其他变量相比所体现的独特性，这在我们对第二点的回复中有详细的解释。

意见 2: 没有控制好实验材料中亲社会行为的成本和要承受的风险、事件的紧急性等因素，可能是这些因素影响亲社会行为的人性评价，而非情绪唤起度。

当亲社会行为的成本和风险高、事件不紧急时，推理可能并不会影响人们对其 HN 的评价。而亲社会行为的成本和风险低、事件紧急时，如果一个人帮助前还要思考很久，这似乎会损害人们对其 HN 的评价。

回应: 确如审稿专家所言，亲社会行为中存在着“成本和要承受的风险、事件的紧急性等因素”，且这些因素可能干扰研究结果。我们在进行文献综述以及实验的过程中都充分考虑了这些因素，也在文章中有所提及。

1) 审稿专家所提及的这些因素，大致上可分为两类，一类与事件本身的类型有关，比如紧急性、严重性、情感唤醒度等，另一类与做出帮助行为对助人者的可能影响有关，比如代价、风险等。而这两类变量内部，往往纠缠在一起。在实验 2 中，我们选择了预实验中代价和风险上的差异不显著的两个情境(所有统计值见下表)。但这两类情境在严重性和紧急程度上与情感唤起上一样，是存在显著差异的。事实上我们认为，在现实生活里的助人情境中，要将紧急性、严重性和情感唤起相区分，可能是比较困难的。一个需要帮助的情境越危急，同时也往往越能引发人的情感反应。但从实验设计的角度，确实需要更为精准、没有混淆的设计，因此，我们设计了实验 3。

2) 在实验 3 中，我们将高情感和低情感唤起设为了同样的帮助行为(即让座)，使得除了情感唤起之外，两组在其他变量上均得到有效控制。实验 3 所涉及的高低情感唤起情境，在严重性、代价和风险上的差异均不显著，仅在紧急程度上存在较小差异。在此情况下，我们的研究假设仍然得到了验证。因此我们认为，是情感唤起差异，而非其它变量，影响了研究的结果。

		高情感 <i>M(SD)</i>	低情感 <i>M(SD)</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>	95% CI
实验 2	严重性	6.44 (0.92)	4.58 (1.27)	8.90	47	<.001	1.284	[0.90, 1.66]
	紧急性	6.42 (0.99)	4.83 (1.24)	7.46	47	<.001	1.077	[0.72, 1.43]
	代价	4.48 (1.52)	4.38 (1.50)	0.40	47	.690	0.058	[-0.23, 0.34]
	风险	4.58 (1.93)	4.46 (1.62)	0.34	47	.733	0.050	[-0.23, 0.33]
实验 3	严重性	3.12 (1.59)	2.93 (1.39)	1.38	462	.169	0.128	[-0.05, 0.31]
	紧急性	3.65 (1.55)	3.36 (1.51)	1.97	462	.049	0.183	[0.001, 0.37]
	代价	3.34 (1.62)	3.33 (1.56)	0.05	462	.961	0.005	[-0.18, 0.19]
	风险	2.96 (1.60)	3.00 (1.55)	-0.24	462	.810	-0.022	[-0.20, 0.16]

3) 考虑到文章的逻辑性和连贯性，我们未报告其他在预实验中测量的条目结果。现在，我们在脚注中对此进行了补充说明，以便读者能够更好的理解研究结果的可靠性。

审稿专家提到：“当亲社会行为的成本和风险高、事件不紧急时，推理可能并不会影响人们对其 HN 的评价。而亲社会行为的成本和风险低、事件紧急时，如果一个人帮助前还要思考很久，这似乎会损害人们对其 HN 的评价。”我们非常同意您的这一说法。成本/风险和紧急/严重/情感唤起这两类因素可能会协同作用。您所指出的，是在较没有争议的情况下的判断，但这些因素很有可能相互冲突，比如成本和风险高，但事情很紧急或严重。整个研究做下来后，我们的感受是，尽管在成本和风险较低的时候，如果情感唤起低，事情也不严重不紧急的话，人们不一定更偏好直觉助人者，但是，当亲社会行为的成本或风险过高时，人

们会更青睐做出理性决策的个体，而此时紧急性或严重性的作用会被削弱(尽管我们认为此时情感唤起仍然会有一些的作用)。比如，考虑借大笔金额的钱这种情况。不假思索地将钱借出去，就算事情紧急，但除非是极高情感唤起的情况，否则，可能仍会被评价为幼稚的、轻率的，从而未必能获得 HN 上更高的评价。考虑到借钱这个助人行为有些特殊，因为金钱本就可能唤起理性思维，所以我们未将其包含在当前研究中。但任何帮助行为，只要代价足够大，都可能在理性上压倒感性。

本研究中因为 HN 与直觉和情感之间的关联，我们将关注的重点放在了情感唤起程度上。从上面的说明您可以看到，我们的研究对代价和风险做了比较好的控制，也能够说明情感唤起度在其中的作用。

意见 3: 研究问题的理论支撑不够充分，需要进一步丰富。

回应: 本文的引言部分基于理论对自变量和因变量进行了论述，具体如下：基于认知理论中的双加工模型(dual-processing model)提出直觉与推理两种决策类型，并阐述了以往研究者在该模型下对亲社会行为的相关探索(见论文“1.2 亲社会行为决策模式与人性感知”第一、二段)；基于 Haslam(2006)的双维度模型对因变量人的本质属性(HN)和独特属性(HU)两大人性维度进行了阐述，此外还阐述了与此相关的去人性化领域的其他理论，如 Leyens 等人(2000)的降人性化理论(infrahumanization)，Gray 等人(2007)感受性-能动性(experience-agency)心智感知理论(见论文“1.1 亲社会行为与人性”第二段)。

您所指出的理论支撑不充分，应该是在假设的推出上，没有能够非常好地加以梳理。我们也非常认同这一点，并在这次修改中进行了更清晰的阐述。请参考文中的修改部分。

意见 4: P10 讨论部分“此外，我们还发现，对推理者的 HN 评价甚至也高于 HU”：这里高于 HU 有误？应该是高于直觉者的 HN，不是高于 HU(结果中是：推理亲社会行为者的 HN 高于直觉亲社会行为者)。

回应: 非常抱歉这里出现了笔误。我们已经在正文中对此书写错误进行了更正：“此外，我们还发现，对推理者的 HN 评价甚至也高于直觉者。”

意见 5: P15 讨论部分“积极后果主要提高的是直觉决策模式下的人性感知”的论述需要在结果部分增加另一个交互作用方向的数据作为支撑。

回应: 感谢您指出这个重要的问题，讨论的内容确实需要另一方向的交互作用结果作为支撑。我们已在实验 4 的结果部分补充了另一方向的简单效应分析结果，以说明讨论中的内容。具体如下：“在直觉和推理决策模式下，不同后果之间的 HU 评分差异均显著， $F_s(2, 607) = 26.05$ 和 21.17 ， $p_s < .001$ ， η_p^2 s = .079 和 0.065，90% CIs [0.0463, 0.1129]和[0.0354, 0.0969]。具体来说，在直觉决策模式下，积极后果组的 HU 评分显著高于控制组， $t(200) = 3.78$ ， $p_{\text{bonf}} < .001$ ， $d = 0.512$ ，95% CI [0.2506, 0.8120]和消极后果组， $t(206) = 6.96$ ， $p_{\text{bonf}} < .001$ ， $d = 0.988$ ，95% CI [0.6773, 1.2524]，且后两者之间的差异显著， $t(208) = 3.14$ ， $p_{\text{bonf}} = .006$ ， $d = 0.442$ ，95% CI [0.1592, 0.7068]；在推理决策模式下，消极后果组的 HU 评分显著低于积极组， $t(207) = 5.79$ ， $p_{\text{bonf}} < .001$ ， $d = 0.775$ ，95% CI [0.5182, 1.0820]和控制组， $t(197) = 5.84$ ， $p_{\text{bonf}} < .001$ ， $d = 0.823$ ，95% CI [0.5382, 1.1184]，后两者之间差异不显著， $t(196) = -0.20$ ， $p_{\text{bonf}} = 1.000$ ， $d = -0.030$ ，95% CI [-0.3074, 0.2505]。”

意见 6: 研究中用于删除被试的质量检测题是什么题目？请说明。

回应: 我们在实验 1 的被试部分已经补充上了所使用的质量检测题“该题请选择‘完全不符’”，在后面的实验中使用的也是同样的检测条目，并在后面的实验中注明了质量检测题

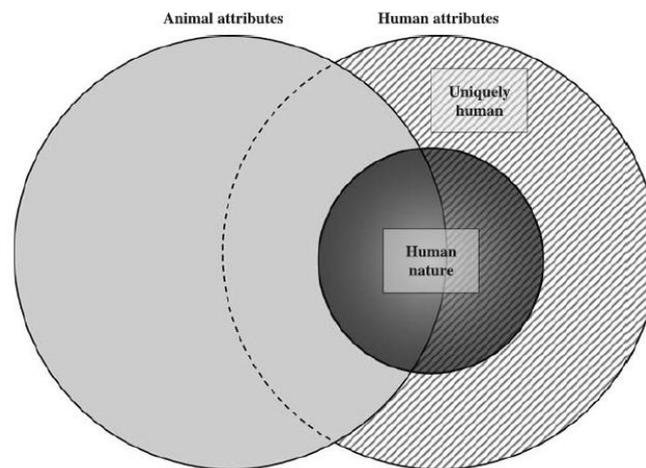
“(同实验 1)”的说明。

审稿人 2 意见：该研究通过 4 个实验，探究了不同情境下人们对直觉亲社会者和推理亲社会者人性判断的差异。研究主题新颖且具有现实意义。但是，该研究的假设逻辑和研究方法还存在一些问题，可能影响实验结果的可信度和研究结论。具体问题如下：

意见 1：引言

作者在假设 1 和 2 都提出“情感和理性效应的相互抵消”，但是其中的理论依据和逻辑并不清楚。HN 和 HU 是相互独立还是相互作用的两个维度？情感和理性效应如何分别或共同影响 HN 和 HU？

回应：感谢您指出这一重要的理论问题，我们确实没有完全阐述清楚。HN 和 HU 是 Haslam 等研究者关于去人性化的理论模型中的两个维度，这两个维度无论从概念和测量结果来看，均具有一定的相关性，但两者仍然是可以区分开的，如下图所示：



(from Haslam et al., 2007)

HN 代表了人的本质属性，是人与无生命物体相区分的那部分特征，如主动性、灵活性、人际温暖、情感反应等，而 HU 代表了人的独特属性，是人相对于其他动物所独有的那些特征，如自我控制、文化、高级认知和情感功能等。因此，相对来讲，HN 与先天性、自发性的情感反应更为相关，而 HU 则与后天发展的理性、逻辑等更为相关。尽管如此，HN 和 HU 与情感和理性并非是完全对应的，HN 中也包含了较为初级的自发思考能力，而 HU 也涉及较为高级的情感，如骄傲、羞愧等。关于这一点，我们在文中进行了补充。

“情感和理性效应的相互抵消”这部分的逻辑确实需要展开。其基本假设是：在高情感唤起的情况下，人们对助人者的人性判断整体上会受到情感唤起的主导。此时，直觉型助人代表了个体出于自发的情感反应去帮助他人，因而会提高 HN 上的判断，但是在 HU 上，直觉型相比推理型又显得较缺乏深入思考，因此，直觉助人者情感上对 HU 的提升可能会被理性上的劣势所抵消，因而在 HU 的评价上相比推理型助人没有优势。相反的，在低情感唤起的情况下，人们对助人者的人性判断整体上会更受到理性的主导。此时，在 HU 上，理性助人者的评分显然会高于直觉助人者。但在 HN 上，相比直觉助人者，推理型助人者会出现理性上对 HN 的提升受到情感上对 HN 的降低的效应的抵消(因为理性助人者显得较为缺少情感)，所以，可能造成两者在 HN 上的差异不显著。

这个逻辑尽管看上去颇为复杂，但其无论与理论假设，还是与我们的研究数据之间，均具有高度的一致性。因此，我们认为这一逻辑是成立的。我们在论文中做了相应修改，尝试

将这一逻辑进一步阐述清楚。请您参考文中的修改部分。

意见 2: 方法

情境的高、低情感唤起是研究的一个重要变量,但是,实验当中使用的低情感唤起情境的情感唤起的程度都在 5 分左右(最高 7 分),而高情感唤起情境的情感唤起的程度也不超过 6 分,尽管两者在统计上差异显著,但是“低”情感唤起情境似乎只是不那么高,而不是真的低。作者有必要选取情感唤起程度更低的情境检验现有的结论。

回应: 这确实可能是一个非常严重的方法问题,实际上,可能挑战我们的理论假设和研究逻辑。与您一样,审稿专家 1 也提出了同样的问题,因此,在前面我们已经对这一问题给予了详细的解答。请您参考我们对审稿专家 1 的第一个问题的回应,这里请允许我们不再赘述,但我们仍想再强调和补充几点:

1) 因为所有情境均涉及有人需要帮助的情况,因此有一定的情感唤起是合理的。比如,即使是像借车这样在助人情境中不算能引发很高情感反应的情况,其情感唤起可能也比中性情境要更高。

2) 尽管如此,我们仍然能够在助人情境下,区分出高低情感唤起(麻烦您参考给审稿专家 1 第一条意见的回复)。

3) 更为重要的是,在助人情境下,情感唤起只要不足够高,也就是我们所用的“低”情感唤起情景,就足以让人们推理亲社会行为更加认可了。也就是说人们认为直觉地进行亲社会行为是否更人性的取决于情景的情感唤起够不够高,而不在于够不够低。只要情感唤起不足够高,人们就会更认可理性助人。

4) 另一方面,在高情感唤起度情境下,情感唤起仍然没有超出 6 分,这里面也有我们的考虑。因为如果一个情境情感唤起过高,则很可能在其他维度上无法与情感唤起度低的情境相匹配。但我们也相信,如果在高低情感唤起区分度更明显的情景中(这意味着高情感唤起情境的情感唤起分数更高,比如 6 分及以上,而不是低情感唤起情境的分数更低),我们的假设将得到更有力的证明。

意见 3: 对于 HN 维度的评价,实验 1、3 和实验 2、4 结果不一致,目前文章并没有合理的解释。对于作者提出的“一个可能的原因是研究 2 中所采用的情境(借车)相比该研究中的情境(让座)代价和风险更大,因而理性所扮演的角色更为重要,不仅影响了 HU,也影响了 HN 评价。”一方面,没有数据支持借车情境比让座情境代价和风险更大,另一方面实验 4 的结果也不能支持这一推断。行为结果对人性判断的影响与情境的风险程度是不同的概念。作者需要提供更有说服力的解释和数据支持。

回应: 我们未提供数据来证明借车相比让座代价和风险更大,尽管这在直觉上是成立的,但确实不够严谨。事实上,我们是有这个数据的,只是由于行文逻辑上的要求,未能全部展示。这里,还麻烦您参考我们给审稿专家 1 的第二个问题的答复中所给出的数据。借车比让座的代价和风险差不多高出了 1 分。

实验 4 中借手机的行为,其代价和风险(代价 $M = 3.42$, $SD = 1.67$, 风险 $M = 3.65$, $SD = 1.59$),虽不及借车,但比让座还是高了一些。其在 HN 上的结果(推理组高于直觉组)与实验 2 的结果较为一致,表明在低情感唤起情境下,人们对代价和风险的敏感性可能主导了对不同决策者的人性判断。

尽管行为结果对人性判断的影响与情境的风险程度是不同的概念,但是,风险是造成负面结果的一个原因。通过设置结果为好的而非坏的,便可以将风险因素排除在外,表明在低风险的时候,人们对直觉助人行为者的 HN 评价相对更好。我们在文中加上了一定的解释。

意见 4: 目前的 4 个实验使用了不同的情境, 这样的好处是可以为研究结论提供汇聚性的证据, 但是目前出现了不一致的结果。另外, 如作者所说, 除了高、低情感唤起, 情境的其他属性也可能影响人性判断。建议作者对其他可能影响判断的情境属性, 例如代价、风险, 进行测量, 在实验中进行匹配或者控制, 排除无关因素的影响。另外, 建议每个实验条件下使用至少两个不同的情境, 使用混合效应模型进行数据分析, 把实验情境作为随机效应, 使研究结论可以推广至其他情境(而不是仅限于实验所用的情境)。

回应: 1) 针对您所提出的第一个问题, 情境的其他属性也可能影响人性判断, 我们有过充分的考虑。因为审稿专家 1 也提出了同样的疑问, 我们已经在前面做出详细的解答, 这里请允许我们不再重复。总体来说, 实验 2 和 3 中的高低情感唤起情境均不存在代价和风险上的显著差异, 具体情况请您参考我们对审稿专家 1 的第二条意见的答复。在当前修改中, 我们用脚注对这些情况给予了说明。

2) 我们认真考虑了您所建议的混合效应模型设计方案。理想地来看, 能够在每个实验条件下加入多个情景刺激是很好的想法, 这可以解决目前研究中可能存在的问题, 并减少误差, 提高研究结果的推广性。但是, 从现实条件来看, 可能当前研究并不太适合采用混合效应模型设计方案。

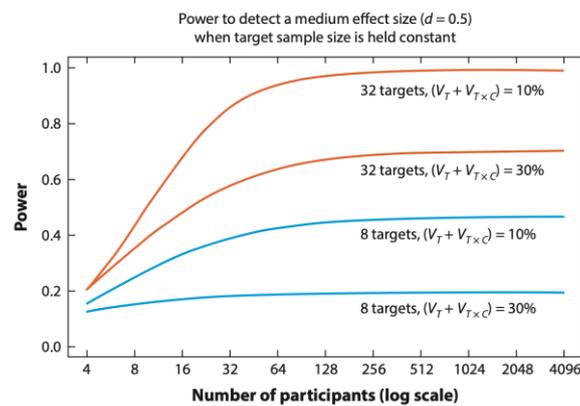


Figure 2

Plot of statistical power as a function of the total number of participants for the CNC (P and T crossed, P crossed with C, and T nested in C) design. The number of targets has been set to either 8 or 32. The other variance components are set to the following values: $V_E = 0.3$, $V_P \times T = 0.1$, $V_P \times T \times C = 0$, $V_P = 0.2$, and $V_P \times C = 0.1$. Note that these other variance components affect only the rate at which the power functions converge to their asymptotes; they do not affect the maximum attainable power values, which depend only on the effect size, number of targets, and the target variance components. Abbreviations: C, fixed condition factor; E, error; P, random participant factor; T, random target factor.

图源于原文中的 Figure 2 (Judd et al., 2017)

对于在哪些条件下可以将刺激作为随机因子, Judd 等人(2017)曾给出建议。在综合考虑实验中的效应量、统计检验力等因素后, 他们通过严格的统计检验, 得到上图所示的结果。从图中我们可以看到, 在中等效应量条件下, 当每个实验条件(固定效应)下的刺激(随机效应, 即图中的 targets)为 8 时, 即便被试量无限增大, 统计效应量依然只能达到 0.4 左右。而当刺激增加到 32 个, 则在样本量足够大时, 可以获得较好的统计效力。作者明确指出, 刺激数量的增加能够有效提高统计检验力。

因此, 从严格的实验设计角度考虑, 使用混合效应模型需要我们在每个实验条件下设置较多的情境进行重复测量, 才能够达到一般要求下 0.8 的统计检验力。但是这在我们的研究中是不太现实的。事实上, 采用这种研究设计要求每个实验试次(trial)的持续时间较短, 对因变量的测量快速而简单。这种设计和分析方法较常用于认知行为范式的研究中, 比如重复多样的图片或词语刺激, 记录被试的快速打分或反应速度等(潘晚珂等, 2023; Judd et al., 2012)。而我们研究中的因变量为人性, 该变量的测量需要完整的二维量表, 多次的重复情

境呈现和重复测量会让实验设计变得冗长且繁杂，且加重被试的练习、测试和迁移效应。

虽然这是一种遗憾，但我们在研究中已尽量使用了多种不同的情境，并对情境进行严格的操控。我们相信，所得到的主要结果上的稳健性足以证明我们的假设。

参考文献

- 潘晚珂, 温秀娟, 金海洋. (2023). 贝叶斯混合效应模型: 基于 brms 的应用教程. *心理技术与应用*, 11(10), 577–598. <https://doi.org/10.16842/j.cnki.issn2095-5588.2023.10.001>
- Haslam, N. , Loughnan, S. , Reynolds, C. , & Wilson, S. (2007). Dehumanization: a new perspective. *Social and Personality Psychology Compass*, 1(1), 409–422. <https://doi.org/10.1111/j.1751-9004.2007.00030.x>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54–69. <https://doi.org/10.1037/a0028347>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, 68, 601–625. <https://doi.org/10.1146/annurev-psych-122414-033702>

第二轮

审稿人 1 意见：论文仍然存在以下问题需要探讨。

意见 1：前言部分对预期结果的推理缺乏依据，不够有说服力，尤其是高情感唤起情境下 HU 维度上理性和情感效应的抵消的推理，以及，低情感唤起情境下 HN 维度上理性和情感效应的抵消的推理。

在高情感唤起情境的 HU 维度上，文中提到“在与理性思维较为匹配的 HU 维度上，直觉行为者因为情感所带来的提升可能会被理性上的不足所抵消，因而相比推理助人者没有优势。也就是说，因为情感和理性效应的相互抵消，我们可能看不到两种亲社会行为在 HU 上的差异”。“直觉行为者因为情感所带来的提升”是指提升 HU 判断吗？这种提升和直觉导致的理性降低相互抵消？为何直觉行为者情感的提升会影响与理性较匹配的 HU 维度？

在低情感唤起情境的 HN 维度上，文中提到“因为推理助人者相比直觉助人者显得较为不真诚和缺少情感，所以其与情境相匹配的理性行为对 HN 的提升可能受到低情感对 HN 的负面效应的抵消”。为何理性行为会提升与情感较匹配的 HN 维度？进而理性和情感相互抵消？

回应：感谢您的意见。您的困惑是合理的，实际上，最初我们也有这样的设想，也就是说，情感提高 HN 而不影响 HU，理性提高 HU 而不影响 HN。但等到具体研究设计时，我们发现，这一设想有点过于简单了。毕竟，亲社会行为所涉及的情境是非常广泛的，而不同的情境，可能对行为的要求也有所不同，这就使得与情境更为匹配的行为可能不但在人性的某一维度上评价为更高，也同样会在另一维度上得到更高评价。

首先，尽管人性的两个维度，HN 和 HU 与情感和理性之间存在一定的对应关系，但两者并非是完全对应的。为了更好地说明这一点，我们在“1.1 亲社会行为与人性”的末尾部分更加详细地阐述了 Haslam 的双维度人性模型，以及 HN、HU 两个维度与理性和直觉之间的关系。所加入的内容具体如下“该研究基于 Haslam 等的 HN-HU 双维度人性模型，并将情境对情感的唤起程度也纳入进来。在 Haslam 等人(Haslam, 2006; Haslam et al., 2007)的理论框架中，HN 维度代表了人与无生命物体相区分的那部分特征，如主动性、灵活性、人际温

暖、情感反应等，而 HU 则代表了人相对于其他动物所独有的那些特征，如自我控制、文化、高级认知和情感功能等。因此，相对来讲，HN 与先天性、自发性的情感反应更为相关，而 HU 则与后天发展的理性、逻辑等更为相关。尽管如此，HN 和 HU 与情感和理性并非是完全对应的，HN 中也包含了较为初级的自发思考能力，而 HU 也涉及较为高级的情感，如骄傲、羞愧等。HN 和 HU 两个维度尽管在概念上有所区别，但在测量上存在高度相关，因此有时研究者会将二者合并为总体的人性水平 (Bastian et al., 2012)。本研究涉及感性和理性之间的对比，因此为了更清晰的揭示不同模式的亲社会行为对人性的影响，我们将 HN 和 HU 分开进行分析。”

1) 在高情感唤起情境，之所以直觉行为者情感的提升会影响与理性较匹配的 HU 维度，然后又形成抵消，需要通过以下几步逻辑来理解：a. 在 HN-HU 模型中，这两维度是高相关的，而不是独立的，并且，这两者与直觉和理性也并非完全对应关系；b. 在高情感唤起情境中，“人们此时对助人者的人性判断整体上会受到情感唤起的主导”，也就是说，直觉相比推理亲社会行为因为情感带来的提升是总体人性上的，包含 HN 和 HU 两方面。（在此，我们假设一个虚拟的数值来代表提升度以便于理解，假设直觉相比推理因为情感对于 HN 和 HU 各提升了“+10”）；c. 虽然此时直觉行为对于两维度都有提升，但是因为 HU 与理性思维更为相关，而理性思维与直觉冲动是冲突的，所以此时直觉行为理性评估上的不足会减弱 HU。（假设 HU 因此比推理低“-10”）；d. 总结起来，在高情感唤起情境，直觉行为所体现的情感属性对 HN 和 HU 有共同提升，但与 HU 中理性冲突带来 HU 减弱，因此，我们会观察到，直觉行为相比推理行为会得到更高的 HN 评价，但两者的 HU 评价可能没有差异。（直觉相对推理，HN 上“+10”，但 HU 上“+10-10=0”）

2) 与 1) 的逻辑推论方式同理。在低情感唤起情景下，理性占主导，推理行为相比直觉行为会因为与情境的匹配得到更高的 HN 和 HU 评价（比如，各高“10”），但是推理行为可能降低情感判断因而会在 HN 维度上有一个负面的影响（因此推理亲社会行为者相比直觉亲社会者 HN 评价变成“+10-10=0”，而 HU 评价仍然为 HU “+10”）。

以上即我们推出两项假设的逻辑，确实有些复杂。也许，有两点值得再强调一下，一是人性的两个维度并非完全独立，且它们与情感和理性之间也不是完全的对立关系。二是情境的属性可能会改变人们对采用不同决策方式做出亲社会行为者的评价。高情感唤起情境下，行为所体现的情感性会占主导，而低情感唤起情境下，行为所体现的理性思维能力可能会占主导。

希望我们的回答能解答您的疑惑。我们也已在论文中进行了相应说明。如果您对我们的解释仍然有疑问，我们非常乐意继续与您探讨。

意见 2: 实验 2~4 在低情感唤起情境中 HN 维度的结果也并不一致。文中提到“在低情感唤起情境中，推理亲社会行为者人性中的独特属性维度高于直觉亲社会行为者，本质属性上的结果则依赖于情境(实验 2~4)”，所谓的依赖于情境到底是取决于什么呢？

回应: 感谢您的意见。实验 2~4 在低情感唤起情境中 HN 维度上的结果确实不一致。这也与我们最初的研究假设不太符合。我们最初认为，即使在低情感唤起情境下，推理亲社会行为在 HN 维度上的评分也不会高于直觉亲社会行为，毕竟，直觉行为被认为是发自内心的情感，而 HN 又表征了人性中偏先天的、情感的那部分。但实际上却发现，在某些情况下，推理行为可能在 HN 上得到更高评价（比如当前的实验 2）。而因为实验 2~3 在低情感唤起情境中 HN 维度的结果不一致，所以我们才进一步做了实验 4。

我们认为，之所以会出现这种情况，是因为人们可能对直觉和推理行为有不同的结果预期，而对结果的预期会对他们的判断产生影响。相比推理行为，直觉行为可能被认为是冲动的，可能引发负面结果的。因此，在实验 4 中，我们关注了低情感唤起情境下，不同的行为

结果(积极 vs. 控制 vs. 消极)可能产生的影响。我们发现, 当不呈现行为结果时, “在 HN 维度上, 本研究重复了研究 2 的结果, 即人们同样给予推理亲社会行为者更高的 HN 评价。尽管本研究中借手机行为的风险和代价不及研究 2 中借车, 但比研究 3 中的让座还是高了一些。其在 HN 上的结果表明在低情感唤起情境下, 人们对代价和风险的敏感性可能主导了对不同决策者的人性判断, 因而推理被赋予了更高权重, 足以抵消其所反映的“不真实”和冷漠。”当加入行为后果之后, 如我们所预测, “积极后果使得两种决策模式所导致的人性感知差异不复存在。具体而言, 积极后果主要提高的是直觉决策模式下的人性感知, 表明人们确实认为直觉的决策模式可能蕴含风险和代价, 因而当被告知后果是好的时候这种对于风险和代价的担忧就消失了, 从而提升了对直觉亲社会行为者的人性评价。消极后果则同时降低了两种决策模式下的人性水平, 不过此时推理模式的 HU 评价仍然高于直觉模式。”

所以, 我们在这里所指的情境因素, 包括行为者需要付出的代价和对其的可能后果等。当然, 也可能存在其他我们未考察的因素。因为这句话是摘要中的, 似乎不太合适在文中做进一步解释。请您谅解, 也希望我们已经将这一问题阐述清楚。

意见 3: 不同场景对情感唤起的程度不同的本质到底是什么? 似乎是事件的严重性和紧急性?

回应: 感谢您的意见。这确实是个非常重要的理论问题, 我们在上一轮的回复中也已有所解释。在此, 我们将更清晰地加以说明。尽管事件的严重性和紧急性可能与情感唤起有关(比如, 越严重和紧急的情况下, 人们的情感唤起也会更高), 但不同场景对于情感唤起的程度不同的本质是情感反应本身, 而非严重性和紧急性(就像孟子说的, 看到一个婴儿爬到井边, 人们都会起恻隐之心。一些场景比如罹患疾病需要照顾, 比另一些场景比如因紧急事务需借车出行, 也许严重性和紧急性并没有很大差异, 但情感唤起是不同的)。我们在实验 2 中所使用的两个场景对代价和风险进行了较好控制, 但严重性和紧急性确实存在显著差异(请参考下表)。因此, 为了分离出情感唤起的独特影响, 而避免其他因素的干扰, 我们设计了实验 3。从下表可以看到, 实验 3 中严重性、紧急性、代价、风险都得到了良好的控制(但在情感唤起的维度有非常显著的差异, $t(462) = 3.77, p < .001, d = 0.350, 95\% \text{ CI } [0.1668, 0.5336]$)。在这种严格控制下, 我们依然验证了我们的假设, 这足以表明我们所假设的差异是由于情感唤起而不是其他因素的影响。

		高情感 $M(SD)$	低情感 $M(SD)$	t	df	p	d	95% CI
实验 2	严重性	6.44 (0.92)	4.58 (1.27)	8.90	47	<.001	1.284	[0.90, 1.66]
	紧急性	6.42 (0.99)	4.83 (1.24)	7.46	47	<.001	1.077	[0.72, 1.43]
	代价	4.48 (1.52)	4.38 (1.50)	0.40	47	.690	0.058	[-0.23, 0.34]
	风险	4.58 (1.93)	4.46 (1.62)	0.34	47	.733	0.050	[-0.23, 0.33]
实验 3	严重性	3.12 (1.59)	2.93 (1.39)	1.38	462	.169	0.128	[-0.05, 0.31]
	紧急性	3.65 (1.55)	3.36 (1.51)	1.97	462	.049	0.183	[0.001, 0.37]
	代价	3.34 (1.62)	3.33 (1.56)	0.05	462	.961	0.005	[-0.18, 0.19]
	风险	2.96 (1.60)	3.00 (1.55)	-0.24	462	.810	-0.022	[-0.20, 0.16]

意见 4: 积极和消极的行为后果的影响很有可能不仅仅存在于低情感唤起情境下, 在高情感唤起情境下可能也有影响。

回应: 我们非常同意您的想法, 并且, 我们也认为这个结果可能会相当有趣。比如, 一种可能性是, 在高情感唤起情境下, 如果结果是好的, 那么直觉亲社会行为相比推理亲社会行为在 HN 上的得分可能会比在结果是中性的或坏的情况下更高, 而坏的结果则可能会降低两类

行为在 HN 上的得分差异。另一种可能性则是，高情感唤起情境调动人们本能中对于他人的关怀，在此情境下人们会将关注的重点放在行为意图而非行为后果上。虽然此时行为后果也可能有影响，但是情感的高度唤起可能会掩盖行为后果的影响。这种情况下，不管行为后果如何，人们均会将直觉亲社会行为者评价为更有人性。

我们之所以没有在当前研究中关注这个情况，是因为当前研究的重点放在对直觉和推理两类行为的人性判断差异上。之所以在实验 4 中引入行为后果，也是因为前面的实验发现的结果上的不一致。为什么在低情感唤起情境下，推理亲社会行为得到的 HN 评价有时也会高于直觉亲社会行为呢（请参见前面我们对您第 2 个问题的答复）？我们想解决这个问题。而通过当前研究，我们基本上已经说清楚了两类行为在人性判断上的不同。所以，尽管高情感唤起情境下，行为决策方式与行为结果间的交互也是非常有趣的问题，但要将亲社会决策模式与行为结果之间的交互作用讲清楚，恐怕又得开启新的系统性研究了。对此，我们在文中进行了解释：“最后，尽管实验 4 为结果预期的作用提供了一定的线索，我们的探讨并不系统（比如，未涉及高情感唤起情境下，决策模式与行为后果的可能交互影响），且未对不同决策模式导致不同人性判断的机制进行深入研究，未来研究可系统性探讨情境、决策模式和行为后果之间的交互作用，并进一步探索共情、动机、真实性等在其中的影响。”我们将在未来的研究中关注这一问题，希望今后仍然有机会与您共同探讨。

.....

审稿人 2 意见：无其他意见。

第三轮

审稿人 1 意见：作者较好地解答了之前的问题。

编委意见：同意审稿人意见，建议发表。

主编意见：同意编委意见，建议发表。