

# 《心理学报》审稿意见与作者回应

题目：分配意图与上行间接互惠：来自行为与 ERP 的证据

作者：王婷，赵梁佛，杨金朋，张丹丹，雷震

## 第一轮

### 审稿人 1 意见：

该文稿作者采用实验的方法，考察资源分配中的结果公平性和分配意图对上行间接互惠行为的影响，同时结合事件相关电位技术，考察被试加工分配方案过程中的时间特征。结果显示，与电脑决定的方案相比，在获得对家提出的不公平分配方案之后，被试分配给新对家的金额更低，相反，与电脑决定的方案相比，在获得对家提出的公平分配方案之后，被试分配给新对家的金额更高。在脑电方面，真人分配比电脑分配诱发了更大的 N1 波幅，不公平分配比公平分配诱发更大的 FRN，P3 则对分配结果的公平性和分配意图的交互作用敏感。该研究针对社会行为中的间接互惠这一重要主题，选题具有意义和一定的新颖性，方法较为得到，结果呈现和论述较为清楚。主要问题如下：

**意见 1：**请解释行为指标“间接互惠水平”如何定义以及背后的逻辑。前言中提到上行间接互惠的定义为“A 帮助/伤害 B，B 以同样方式对待 C”，那么，在“(a)接受不公平方案之后提出不公平方案（即传递不公平）”和“(b)接受不公平方案之后提出较公平方案（即未传递不公平）”相比，谁的间接互惠水平更高？似乎是(a)更符合间接互惠的定义。但是，结果部分的间接互惠水平是由“间接互惠分配金额”作为指标，即被试在第二阶段分配给他人的金额，按照该指标，似乎(b)更可能被认定为间接互惠水平更高。

**回应：**我们非常感谢审稿专家提到的这一点，因为涉及到本文的创新点以及核心逻辑，确实需要更清晰的阐明，这促使我们就此做了大量文献和思考，并对文章做了大量修改。我们从以下两个层次来阐明。

(1) 首先要回答的问题是“上行间接互惠”中的互惠的是什么。

文献基本上都以等分作为公平与非公平或者慷慨与贪婪的分水岭，核心假说是：B 得到 A 大于或等于等分的分配时，B 看成是得到了 A 的善意，考察 B 是否将这个“善意”传递给第三方 C；反之亦反。也就是说，文献主要考察的是基于等分分配的互惠。但是，为什么将等分分配作为分水岭却缺乏充分论证。

根据 DG 实验文献的元分析，几乎所有实验研究均发现：分配大于 50% 的份额给他人的被试极其少见，被试平均分配 28% 给他人 (Engel, 2011)。文献中大于 50% 的实验设计与被试在真实世界中难以遇到高于 50% 的分配额的经验相冲突。因此，已有文献中的实验设计难以刻画人们面对真实世界中分配信号所做出的反应，所得到的间接互惠的结论也就难以通过外部有效性检验（详见修改版论文的讨论部分）。

我们认为，人们对他人的行为做出善意与非善意解读更可能是相对于社会均值这一参照点而言的。因为，A 给 B 分配一个等于或接近社会均值的分配时，B 通常不会感到惊讶，而 A 向上或向下大幅偏离社会平均分配值则更有可能被 B 解读为善意或非善意。我们这样的看法与上述已有文献有着非常大的区别。试想，A 实际平均分 28% 给 B，而且没有人分给

B 大于 50%，如果 B 得到 A 的分配为 49%，按照我们的理解，B 大概率会认为 A 是善意的。我们所认为的善意与非善意是基于“是怎样”的描述性规范，而文献主要是基于“应该是什么”的命令性社会规范。基于此，我们提出本文的第一个核心假说 H1。

**H1:** 人们得到的分配值高（低）于社会均值时，倾向于给第三方一个高（低）于社会均值的分配值。

(2) 其次要回答的问题是指标“间接互惠水平”该如何定义。

大部分上行间接互惠文献没有考虑收入效应这一竞争性假说。也就是说，人们分给第三方更有可能是由自己收入增加导致的。

与 B 得到一个由电脑随机分配的低收入相比，当接受到电脑分配的高收入时，B 分配给 C 的值更高，我们认为这主要是来自正向的收入效应，这意味着，在间接互惠中，与 B 收到一个由真人分配的低收入相比，当 B 接受到真人分配的高收入时，B 不仅因为收入增加分配给 C 的值更高（正向收入效应），更重要的是，B 能感知到真人对他/她的善意，进而比接受到电脑的高收入分配时分配给 C 的值更多，这时我们就说 B 传递了正向的间接互惠。

类似的，与 B 得到一个由电脑随机分配的高收入相比，当接受到电脑分配的低收入时 B 分配给 C 的值更低，我们认为这主要是来自负向的收入效应，这意味着，在间接互惠中，与 B 收到一个由真人分配的高收入相比，当 B 接受到真人分配的低收入时，B 不仅因为收入减少分配给 C 的值更低（负向收入效应），更重要的是，B 能感知到真人对他/她的非善意，进而比接受到电脑的低收入分配时分配给 C 的值更少，这时我们就说 B 传递了负向的间接互惠。

所以，“间接互惠水平”可以由上述两个方向差额的大小来度量，而不能简单由 B 给 C 分配的多少来度量。本文重点考察（我们所定义的）上行间接互惠是否存在，是基于真人分配和电脑分配条件下给第三方分配值的对比，这是一个序数概念。

此外，结合审稿专家的意见，我们修改了文章的标题，以更好地突出本文的贡献。

**意见 2:** 对 Hu 等（2018）研究的批判不够严谨。Hu 等采用的分配方案迫选方式是公平决策研究中一种常用的方法，被试仍然可以按照自己的真是想法进行分配，“该分配设计导致被试无法按照真实的想法进行分配”的批评略显偏颇。

**回应:** 我们赞同“该分配设计导致被试无法按照真实的想法进行分配”这句的表述不严谨，我们不仅删除了这句话，更做了深入研究。

我们认为，Hu 等（2018）最主要的问题还是选择了将等分分配作为了参照点，导致了关于分配意图的行为假说未得到支持，即：B 给 C 分配行为数据在真人与电脑条件间无显著差别。基于此，我们做出如下修改：

Hu 等（2018）虽同样探讨了分配结果与分配意图对上行间接互惠行为的影响，但如前文所述，现有 DG 实验发现人们几乎不会将大于等分的分配给接受者（Engel, 2011），而 Hu 等（2018）将大于等分的分配方案的出现次数与小于等分的分配方案的出现次数设置为相等，或许会让被试在收到大于等分的真人分配方案时难以相信其真实性。更重要的是，当被试收到小于等分的所谓贪婪分配时，由于小于等分的所谓贪婪分配包含本文所识别的低于社会均值的分配和高于社会均值的分配这两种分配方案，使得被试既因收到低于社会均值的分配而倾向于给第三方一个较低的分配，又因收到高于社会均值而给他人一个较高的分配，两种相反的分配模式被归入同一分配条件来衡量间接互惠，最终导致 Hu 等（2018）发现真人和电脑分配条件下人们的间接互惠分配在行为上没有显著差异，即在行为上基于意图的间接互惠假说未得到支持可能是因为沿用了等分分配作为善意和非善意参照点的研究框架和实验设

计。（详见修改稿的引言和讨论部分）

**意见 3:** 正如作者提到，第一阶段 DG 的分配者为真人时，真人面孔的出现可能引发被试更多的选择性注意，那么真人分配和电脑分配之间的差异，究竟是由于意图的不同导致的，还是由于刺激不同带来的选择性注意差异导致的？

**回应:** 我们接受审稿专家的建议，将真人和电脑照片分别替换为“真人”和“电脑”两个词。本轮修改，我们重新做了实验、采集了脑电和行为数据，因此避免了视觉差异对实验结果的干扰。实验刺激如图 1 所示。本次新实验共招募了在校本科生共计 42 人（最新实验内容及结果详见文章的方法和结果部分）。

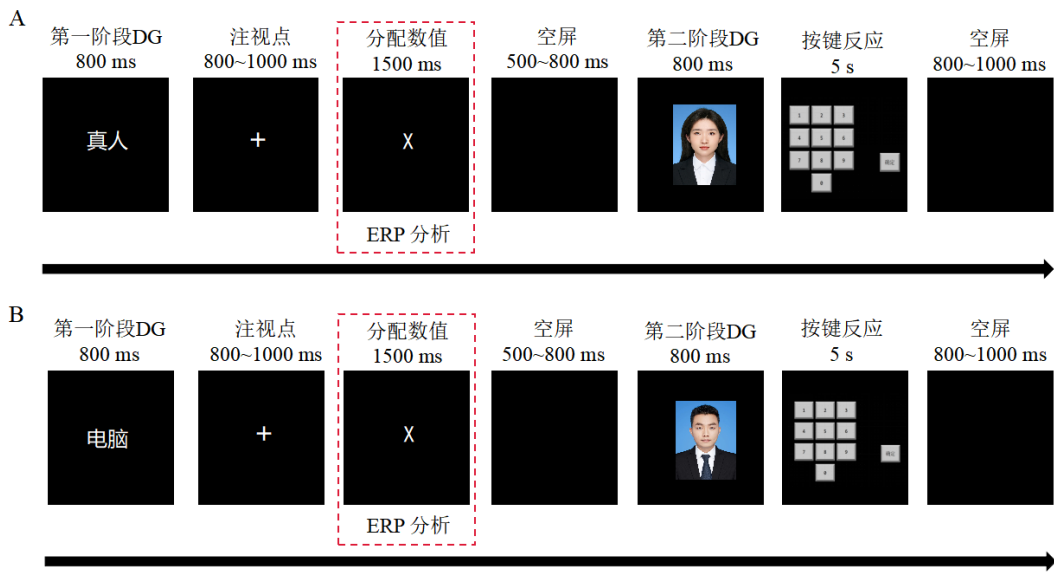


图 1 实验刺激材料。A，单试次的真人分配间接互惠任务。B，单试次的电脑分配间接互惠任务。

**意见 4:** 测算样本量时，效应量是如何确定的？

**回应:** 本研究根据相关研究 (Liu et al., 2022) 报告的效应量最小值 ( $\eta_p^2 = 0.09$ )，使用 G\*Power 3.1.9 进行样本量估计 ( $f = 0.31$ ,  $\alpha = 0.05$ , 方差分析: 重复测量, 被试内因素)，需要 32 名被试即可达到 99% 的统计检验力 (power)。为了避免实验中的各种不可预见因素，我们在此基础上增加了 10 名被试，因此本研究共招募被试 42 名。这一内容我们已经添加到了 2.1 被试部分。

**意见 5:** 建议呈现脑电结果波形图时标注出分析的时间窗范围。

**回应:** 感谢审稿专家的建议。我们在呈现脑电结果波形图时已明确标出脑电成分的时间窗 (具体见文章图 3-图 5)。

**意见 6:** 建议作者在解释 ERP 成分结果和推论时更加谨慎，避免过度反向推断。例如，“真人分配比电脑分配诱发更大 P3 波幅，该结果表明被试在真人分配中监测到更强的情感显著

性信息”，实际上 P3 成分和注意等很多其他心理过程也有关系，直接解释为“监测到更强的情感显著性信息”这样的反向推断是不准确的，也缺乏情绪相关数据的直接支持。

回应：感谢审稿专家提出的建议。我们在文章的讨论部分修改了对 P3 成分结果的讨论，具体内容详见讨论部分。

.....

审稿人 2 意见：

意见 1：该研究试图结合认知实验和脑电技术探讨社会分配中的结果与意图对间接互惠行为的影响，研究主题较有意义。然而实验设计存在缺陷和不足。首先，电脑分配与真人分配的操纵如何能够区分分配意图？这里的分配意图如何定义？前言部分过于强调脑电成分的介绍，而对分配意图等实验主要研究问题和逻辑的介绍并不清楚。

回应：非常感谢审稿专家上述审稿意见。我们针对审稿专家提到问题做了深入细致的修改，重点对“电脑分配与真人分配的操纵如何能够区分分配意图”以及研究问题的逻辑做了清晰的阐明。我们从以下两个层次来阐明。

(1) 分配意图的定义以及为什么采用电脑分配与真人分配的操纵来区分分配意图。

所谓分配意图是指个体在进行分配决策时从一系列可选项中主动选择（不）利于接受者的选项的主观意愿（Falk & Fishbacher, 2006）。Falk 等（2008）专门探索了分配意图的重要性，实验中 A、B 分别作为月光博弈（moonlighting game）的分配者和响应者，当分配值是由真人 A 做出时，B 的返还额随分配值的增加而增加，但当分配值是电脑随机产生时，B 的返还额却不随分配值增加而增加，两种情况返还额的差就代表了 B 对 A 意图的反应。由于随机决定的分配结果不是经过人主动选择的，它不暗含人的分配意图（Stanca et al., 2009），大量研究将由真人决定的与由随机抽取方式决定的决策进行比较，以研究分配意图对个体互惠行为的影响（Rabin, 1993; Blount, 1995; Charnes & Rabin, 2002; Falk et al., 2008; Stanca, 2010）。

本研究正是基于 Falk & Fishbacher（2006）分配意图的定义并借鉴了以往研究的做法（Rabin, 1993; Blount, 1995; Falk et al., 2008; Zhang et al., 2016; Hu et al., 2018），设置了真人分配和电脑分配两种方式，用以刻画意图的影响（详见文章引言部分）。

(2) 本文研究逻辑介绍

首先，我们先解释本文为何研究上行间接互惠。

间接互惠突破了直接互惠的封闭体系，是促进人类社会大规模合作的重要力量，识别间接互惠的内在机制具有重要的理论和现实意义。所谓上行间接互惠，通常是指“B 接受到 A 的善意（非善意）后，B 以同样方式对待第三方 C”（Nowak & Sigmund, 2005）。根据新古典经济学的经济人假设，在完全匿名的一次性博弈中，B 接受了 A 的（非）善意后，由于无法从重复博弈的合作中获益而不会以（非）善意对待 C，导致间接互惠的链条难以形成（Sigmund, 2010）。然而，上行间接互惠得到大量行为实验研究的验证，并与真实世界中的观察一致（Engelmann & Fischbauer, 2009; Horita et al., 2016; Rutte & Taborsky, 2007; 孙熠譞等, 2022），这促使大量心理学、行为科学领域学者对上行间接互惠进行关注并展开研究。然而，至今其内在机制尚不明晰（详见文章引言部分）。

其次，我们来阐明为何本文认为社会均值是人们评判上行间接互惠中善意与非善意的参

照点。

回顾 DG 实验文献发现,几乎所有实验研究发现,分配大于 50%的份额给他人的被试极少,被试平均分配 28%给他人 (Engel, 2011)。然而,现有研究基本上以等分分配作为上行间接互惠中善意与非善意的分水岭。核心假说是: B 得到 A 大于或等于等分的分配时,看成是得到了 A 的善意,考察 B 是否将这个“善意”传递给第三方 C; 反之亦反。也就是说,文献主要考察的是基于等分分配的互惠。但是,为什么将等分分配作为分水岭却缺乏充分论证。

我们认为,人们对他人的行为做出善意与非善意解读更可能是相对于社会均值这一参照点而言的。因为, A 给 B 分配一个等于或接近社会均值的分配时 B 通常不会感到惊讶,而 A 向上或向下大幅偏离社会平均分配值则更有可能被 B 解读为善意或非善意。我们这样的看法与上述已有文献有着非常大的区别,试想, A 实际平均分 28%给 B,而且没有人分给 B 大于 50%,如果 B 得到 A 的分配为 49%,按照我们的理解, B 大概率会认为 A 是善意的。我们所认为的善意与非善意是基于“是怎样”的描述性社会规范,而文献主要是基于“应该是什么”的命令性社会规范。基于此,我们提出本文的第一个核心假说 1 (详见文章引言部分)。

假说 1: 人们得到的分配值高(低)于社会均值时,倾向于给第三方一个高(低)于社会均值的分配值。

再次,我们来说明为何要单独研究分配意图对上行间接互惠的影响。

如果人们在得到高(低)于社会均值的分配值后,确实倾向于给第三方一个高(低)于社会均值的分配值,这一实验结果很可能被另一竞争性假说——收入效应假说——解释,即当人们的收入更高时倾向于给他人分配更多。那么,被试给第三方的分配变化便不是由高(低)于社会均值的分配值所代表的善意(非善意)触发的,而是由不包含意图的纯粹收入效应触发的。

有鉴于此,本文有必要考察控制收入效应后是否还存在因意图所触发的间接互惠。我们认为,相对于低于社会均值的电脑分配, B 接收到高于社会均值的电脑分配时,给 C 分配更多,这可由收入效应来解释;但是,相对于低于社会均值的真人分配, B 接收到高于社会均值的真人分配时, B 可能不仅感受到高收入,而且感受到来自真人 A 的善意,则 B 给 C 的分配值就有可能比在电脑分配的情形时更高,高出的部分正是本文要识别的间接互惠。反之亦反。于是,我们提出本文的第二个核心假说 2 (详见文章引言部分)。

假说 2: 当人们得到的分配值高(低)于社会均值时,人们在真人分配比电脑分配时分给第三方的值更高(低)。

最后,我们采用 ERP 技术对上述假说做进一步验证。

实验中的行为数据支持了上述两个假说,通过考察人们在做决策时的脑电数据可以进一步打开人们决策的“黑箱”,针对此我们提出了 3 个对脑电结果的预测。本文借助 ERP 技术采集了被试的 N1、FRN 和 P300 数据,结果发现,脑电数据符合我们的预测,因此从脑神经活动的时间进程角度验证了上行间接互惠假说。

此外,结合审稿专家的意见,我们修改了文章的标题,以更好地凸显本文的贡献。

综上,本研究采用两阶段独裁者博弈实验框架探讨上行间接互惠,将第一阶段 DG 中高(低)于社会均值而非等分的分配视为(非)善意分配,在此基础上通过操控第一阶段 DG 的分配者角色(电脑或真人)来刻画分配者意图,并借助 ERP 技术考察分配结果与分配意图对上行间接互惠的决策行为和脑神经活动的影响。

## 参考文献

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences.

*Organizational Behavior and Human Decision Processes*, 63, 131–144.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117, 817–869.

Hu, Y., He, L., Zhang, L., Wölk, T., Dreher, J., & Weber, B. (2018). Spreading inequality: Neural computations underlying paying-it-forward reciprocity. *Social Cognitive and Affective Neuroscience*, 13, 578–589.

Falk, A., Fehr, E., & Fishbacher, U. (2008). Testing theories of fairness-Intentions matter. *Games and Economic Behavior*, 62, 287–303.

Falk, A., & Fishbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54, 293–315.

Stanca, L., Bruni, L., & Corazzini, L. (2009). Testing theories of reciprocity: Do motivations matter? *Journal of Economic Behavior & Organization*, 71(2), 233–245.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83, 1281–1302.

Zhang, Y., Yu, H., Yin, Y., & Zhou, X. (2016). Intention Modulates the Effect of Punishment Threat in Norm Enforcement via the Lateral Orbitofrontal Cortex. *The Journal of Neuroscience*, 36(35), 9217–9226.

**意见 2:** 其次，如果没有第一阶段 DG（这里被试作为接受者，应该是 UG）的干扰，被试自身的分配会是怎么样并不清楚。如果要说明间接互惠的影响，需增加无第一阶段的条件作为控制。

**回应:** 我们非常感谢审稿专家提出的建议。我们根据专家建议重新搜集了实验数据，增加了第 1 阶段独裁者博弈。

增加第 1 阶段 DG 的主要有以下 4 个目的：首先，检验本实验样本的分配行为是否与以往 DG 实验文献中被试的分配行为一致，进而判断本实验样本是否有偏；其次，获取样本平均分配值，便于用该分配值来代表社会平均分配值；再次，被试作为分配者进行分配后，增加被试在间接互惠实验中看到第一阶段 DG 中他人给与的分配值时的可信度；最后，便于考察不同利他水平被试间接互惠行为的差异。（详见 2.2 实验设计和实验材料、3.1 行为结果）。

此外，需要说明的是，本文采用 2（分配结果：低于社会均值分配 vs. 高于社会均值分配）× 2（分配意图：真人分配 vs. 电脑分配）的被试内实验设计，这样设计的好处是保证了四个条件被试分布是完全相同的，进而便于重点关注被试在四个条件间的间接互惠差异。

---

## 第二轮

**审稿人 1 意见:**

尚有少量文字需要修改:

**意见 1:** 根据正文，摘要中“而高于社会均值的真人分配和电脑分配所诱发的 P3 没有显著差异。”应修改为：“高于社会均值的真人分配比电脑分配诱发了更大的 P3。”

**回应:** 感谢审稿专家的意见。根据您的意见，我们对摘要进行了修改。

**意见 2:** “作为本研究的社会平均分配值；”建议修改为“便于检验本研究将 3 作为区隔善意与非善意的合理性；”

**回应：**感谢审稿专家的意见，已经按照您的意见在正文部分进行了修改。具体见 2.2 实验过程和实验材料。

**审稿人 2 意见：**我已经没什么意见了。

**回应：**感谢审稿专家为本稿件付出宝贵的时间！

---

### 第三轮

**编委意见：**

两位审稿人的建议都是对文章进行简单修改后接受，虽然我个人觉得该研究的神经证据和对于意图的分类稍显粗糙，但是还是尊重审稿人的意见，建议编辑部接受该稿件。

**主编意见：**

赞同编委意见，ERP 结果的呈现过于简单，波形图是基于很多个电极的平均幅值，没有清楚显示地形图的分布，需要补上，便于读者判断。

**回应：**非常感谢主编的建议。如您的意见，我们丰富了 ERP 结果，分别补充了 N1/FRN/P3 的地形图，具体见正文的图 3、图 4 和图 5。