

《心理学报》审稿意见与作者回应

题目：自恋人格调节信任博弈的结果评价

作者：王益文 付超 任相峰 郭丰波 张振 黄亮 袁博 郑玉玮

第一轮

审稿人 1 意见：作者利用事件相关电位技术考察了自恋人格对信任博弈任务中的结果评价过程的影响。该研究比较有意义。写作也比较流畅，是一个不错的稿件。下面提出的一些问题，主要是从一个与作者交流的角度所提，并不是意味着希望作者按照意见进行修改，请作者自行思考。

意见 1：被试数量。被试只有 27 名，这样的量只适合将脑电指标与问卷数据做个相关，而不适合进行分组。

回应：非常感谢审稿专家对本研究和稿件的肯定及所提出的宝贵意见。在认真思考了两名审稿专家的意见后，我们努力对本研究的被试样本量进行了随机取样补充，现将研究的被试扩充至 38 人，其中男生 18 人，女生 20 人。具体修改和补充的内容如下：“38 名(20 名女生，18 名男生)在校大学生参与实验，年龄为 19.7 ± 1.5 岁。所有被试均为右利手，视力或者矫正视力正常，无精神病史。实验前签署知情同意书，实验后根据被试在博弈任务中的最终收益，给予一定报酬。”详见文中被试描述部分的红色字体，请审稿专家予以审查。

意见 2：神经机制方面。作者介绍了信任行为的神经成像研究，但是没怎么介绍自恋人格的神经成像研究。如能将信任行为和自恋人格的神经机制联系起来进行讨论，是不是有助于读者更好的理解研究意义和提出研究预期？

回应：针对审稿专家提出的引言部分并未介绍自恋人格的神经成像研究，我们在引言中对这一部分内容进行了补充。具体内容如下：“研究者从自恋的自我特征角度对其神经生理机制进行探讨，发现额中回、内侧额叶、眶额叶皮层、楔前叶和前脑岛涉及自我提升和自我评估(Kwan et al., 2007; Barrios et al., 2008; Beer, & Hughes, 2011); 前脑岛、右侧背外侧前额叶皮层和右侧后扣带回皮层与缺乏共情有关(Fan et al., 2011; Schulze et al., 2013; Nenadic et al., 2015)。”详见文中引言部分的红色字体，请审稿人审查。

意见 3：研究预期的提出。目前的稿子给人的印象正如作者在文中所言是鉴于当前对自恋个体信任行为的实验研究尚不充分，但是仅仅这么提对于研究意义是一种伤害。应当进一步指明开展这项研究有助于人

们认识某个问题，或者解决先前研究中的某种不一致。也即是说要让读者有一种强烈的感觉：之前的研究没搞清楚某个问题，本研究有助于搞清楚该问题。

回应：针对问题提出我们进行了如下修改：“研究者认为具有相同自然风险态度的个体在人际信任情境中可能做出不同决策(Bohnet & Zeckhauser, 2004; Dunning et al., 2014)，一方面，个体可能受到背叛厌恶或后悔厌恶的影响而抑制信任行为的产生(Bohnet et al., 2008)；另一方面，个体可能出于纯粹利他性或尊重规范而选择信任他人(Dunning et al., 2014; Schlösser, Mensching, Dunning, & Fetchenhauer, 2015)，而高自恋者的行为(赌博和投资)决策具有明显风险寻求(risk taking)倾向(Campbell, Goodie, & Foster, 2010; Foster, Reidy, Misra, & Goff, 2011)。因此，对具有风险寻求的自恋者的信任行为进行研究有助于扩展人们对自恋个体在人际信任情境中决策特征的认识。”详见文中引言部分的红色字体，请审稿人审查。

意见 4：FRN 的测量。差异波与原始波都是可行的方法。虽然有研究者近年来主张奖赏正波，但是近年来仍有许多研究采用差异波。不过，采用差异波的方法的确丢失了一些有意义的信息。例如，对于自恋人格到底是对负性结果更敏感还是对正性结果更敏感这样的问题，差异波恐怕无法回答。另外，结合作者的结果图，至少能确定在正性结果上两组被试的差异是很明显的，负性结果的组间差异有待作者进一步检验。所以，不强求作者改用原始波，但是建议作者在回复里将输赢分别对待，报告一下统计结果。

回应：我们采用了审稿专家针对“FRN 的测量”所提出的宝贵意见，补充了 FRN 的测量分析。针对 FRN 的测量，依据已有关于 FRN 的研究文献和对本研究中 ERP 总平均波形图的观测，我们选取的时间窗口为反馈呈现后 230~280ms，电极点为脑中线 5 点(Fz、FCz、Cz、CPz、Pz)。鉴于 FRN 可能会受到晚期正成分(P300)的影响，我们也对 FRN 差异波(dFRN)进行了测量，其测量的是损失反馈减获利反馈所获得的波形在 230~280ms 内的平均波幅。在对补充被试后的实验数据分析中，我们分别对 FRN 的原始波和差异波进行测量和分析。

我们对 FRN 的原始波和差异波的测量进行了如下描述：“FRN 测量的是反馈呈现后 230~280ms 的平均波幅，P300 测量的是反馈呈现后 300~600ms 的峰值。鉴于 FRN 可能会受到晚期正成分(P300)的影响，我们也对 FRN 差异波(dFRN)进行了测量，其测量的是损失反馈减获利反馈所获得的波形在 230~280ms 内的平均波幅。根据已有研究(Gehring & Willoughby, 2002; Yeung & Sanfey, 2004)，以及对 ERP 总平均波形和脑地形图的视觉检测(见图 2)，我们选取脑中线 5 点(Fz、FCz、Cz、CPz、Pz)进行分析。即对 FRN 和 P300 进行 2(自恋水平：高、低)×2(反馈条件：信任损失、互惠获利)×5(电极分布：Fz、FCz、Cz、CPz、Pz)的三因素混合方差分析，对 dFRN 进行 2(自恋水平：高、低)×5(电极分布：Fz、FCz、Cz、CPz、Pz)的二因素混合方差分析。”

我们对 FRN 的原始波进行了如下的分析：“对 FRN(230~280ms)的平均波幅进行 2(自恋水平)×2(反馈条件)×5(电极分布)的三因素混合方差分析，结果表明不同反馈条件的 FRN 平均波幅差异显著， $F(1,$

36)=65.56, $p<.001$, $\eta^2=.65$, 信任损失的 FRN 平均波幅($9.65\pm 1.06\mu V$)显著大于互惠获利的 FRN 平均波幅($14.11\pm 95\mu V$): 反馈条件与自恋水平的交互作用显著, $F(1, 36)=5.94$, $p<.05$, $\eta^2=.14$, 简单效应分析发现, 在相对高自恋者中, 信任损失的 FRN 波幅($9.21\pm 1.49\mu V$)显著大于互惠获利的 FRN 波幅($15.02\pm 1.34\mu V$), $F(1, 36)=55.48$, $p<.001$, $\eta^2=.61$; 在相对低自恋者中, 信任损失的 FRN 波幅($10.08\pm 1.49\mu V$)显著大于互惠获利的 FRN 波幅($13.2\pm 1.34\mu V$), $F(1, 36)=16.02$, $p<.001$, $\eta^2=.31$, 相对高自恋者的差异量更大, p 值更显著, 这也就与 dFRN 的统计结果相一致。”

我们对 FRN 的差异波进行了如下的分析:“对 dFRN(230~280ms)的平均波幅进行 2(自恋水平) \times 5(电极分布)的二因素混合方差分析, 结果表明不同水平自恋者的 dFRN 平均波幅差异显著, $F(1, 36)=5.64$, $p<.05$, $\eta^2=.14$, 相对高自恋者的 dFRN 平均波幅($-5.81\pm .79\mu V$)比相对低自恋者的 dFRN 平均波幅($-3.19\pm .79\mu V$)更大。电极位置与自恋水平的交互作用显著, $F(4, 144)=3.34$, $p=.042$, $\eta^2=.085$, 简单效应分析发现在额区(Fz)、额中区(FCz)和中央区(Cz)相对高自恋者的 dFRN 平均波幅均显著大于相对低自恋者的 dFRN 平均波幅($ps<.05$), 但在中央顶区(CPz)和顶区(Pz)上两组被试的 dFRN 平均波幅不存在显著差异($ps>.05$)。”详见文中“脑电记录和分析”和 ERP 结果的 FRN 和 dFRN 部分的红色字体, 请审稿人审查。

意见 5: 自恋问卷的得分。审稿人不清楚该问卷是否有常模, 或者中国人一般得分是多少? 本研究中的高低自恋组得分分别是 3 点多和 2 点多, 这与其它研究中的得分比较起来是个什么情况?

回应: 本研究使用的简版自恋人格问卷(A brief version of Narcissistic Personality Inventory, NPI-16) (Ames, Rose, & Anderson, 2006), 它由 16 个二选一的迫选项目构成的单因素结构量表, 分数越高代表自恋水平越高。NPI-16 是 Ames 等(2006)在 NPI-40 (Raskin, 1988)基础上编制而成的。在本研究中为了获取被试更多的反应可变性(response variability) (Pincus et al., 2009), 将原来的二选一的迫选问题调整为单一选择问题, 并采用 5 点计分, “1”表示“完全不赞同”, “5”表示“完全赞同”, 得分越高表示自恋倾向越高。

当前在我国采用类似问卷测量自恋人格的得分如下:

郑涌和黄藜(2005)在自恋人格问卷(NPI-40)和自恋人格障碍量表(NPDS)的基础上编制了自恋人格问卷, 该问卷由 28 个题项构成, 采用 5 级评定计分, 文理科被试在该量表上的平均得分为 3.15 ± 0.54 和 3.0 ± 0.51

周晖等(2009)根据开放式问卷、访谈和参考已有自恋问卷(NPI-40), 编制了自恋人格问卷。该量表由 34 个项目构成, 采用 6 点计分, 男女生被试在该量表上的平均得分为 3.80 ± 0.70 和 3.52 ± 0.59 。

Cai, Kwan, 和 Sedikides (2012)使用自恋人格问卷(NPI-40), 采用 6 点计分, 城市和农村的平均得分为 3.57 ± 0.75 和 3.5 ± 0.75

王晓燕(2008)使用 NPI-16 原始量表, 采用二分计分法, 统计被试的总分, 独生和非独生子女的得分分别为 6.16 ± 2.19 和 5.61 ± 2.59

将上述不同研究的被试得分进行 5 点计分转换，结果发现被试在自恋人格问卷上的得分均在 3 分左右。我们将被试分组命名改为“**相对高自恋组**”、“**相对低自恋组**”，并对问卷测评结果部分进行了如下修改：“将 38 名被试在自恋人格问卷上的平均得分进行 T 分数转换后由低到高排序，以 50% 百分位数的分数为标准，将分数在 50% 以下的 19 名被试(9 男)被划分为**相对低自恋组**，分数在 50% 以上的 19 名被试被划分为**相对高自恋组**。**相对高自恋组**的分数(57.05 ± 7.38)显著高于**相对低自恋组**的分数(42.95 ± 6.79)， $t(36)=6.14, p<.001$ 。”详见文中问卷测评结果部分的红色字体，请审稿人审查。

意见 6: 即使本研究中两组人自恋得分不同，也不宜过于强调两组人的共情差异和自私差异。因为没有进行具体的测量。

回应: 在修改后的稿件中，我们删除了自恋个体之间的信任选择率差异是由两组人的共情差异和自私性差异导致的。此外，根据另一名审稿专家的建议，我们将“未对共情进行测量”列入到研究局限之中。详见文中研究局限部分的红色字体，请审稿人审查。

意见 7: 实验设计中，被试只有两秒钟的时间思考，是不是太短了？本实验中被试做出决策的时间在 500 毫秒左右，请报告一下三五个同类研究中的决策时间。

回应: 大量脑电研究中实验设置了 2000ms 左右的被试抉择时间，被试需要在 2000ms 内做出信任或不信任抉择，被试绝大多数反应均在 2s 内完成。在补充被试后，当前研究中被试的决策时间如下：“**相对低自恋者**做出信任和**不信任决策的平均反应时**分别为 $583.5 \pm 227ms$ 和 $555.8 \pm 192ms$ ；**相对高自恋者**做出信任和**不信任决策的平均反应时**分别为 $528.8 \pm 227.7ms$ 和 $518.5 \pm 199ms$ 。”当前被试的决策时间与已有采用信任博弈范式的研究结果类似。现列举如下研究中研究者设置的决策时间和被试的反应时供审稿专家查阅：

1、Kang, Y., Williams, L. E., Clark, M. S., Gray, J. R., & Bargh, J. A. (2010). Physical temperature effects on trust behavior: the role of insula. *Social Cognitive and Affective Neuroscience*, nsq077. 该研究采用信任博弈任务，实验中被试的决策时间为 2s，被试需要在该时间段内做出决策。

2、Boudreau, C., McCubbins, M. D., & Coulson, S. (2009). Knowing when to trust others: An ERP study of decision making after receiving information from unknown people. *Social Cognitive and Affective Neuroscience*, 4(1), 23-34. 该研究中被试的反应时为 800ms 左右

3、Tortosa, M. I., Lupi áñez, J., & Ruz, M. (2013). Race, emotion and trust: An ERP study. *Brain research*, 1494, 44-55. 该研究采用信任博弈任务，研究者并未报告被试的反应时，但实验中被试的决策时间为 1.5s，被试需要在该时间段内做出决策。

4、王益文，张振，原胜，郭丰波，何少颖，& 敬一鸣. (2015). 重复信任博弈的决策过程与结果评价. *心理学报*(8). 该研究采用重复信任博弈任务，被试为在校大学生，实验中被试的决策时间为 2000ms，被

试选择信任和不信任的平均反应时分别为 $511 \pm 22\text{ms}$ 和 $502 \pm 22\text{ms}$ 。

意见 8: 稿子中有个别文字错误，比如引言第四段中，也研究者应为也有研究者。像 ERP、FRN 这些词首次出现时都没全称。

回应: 有关稿件中文字错误问题，我们对稿件进行了细致的阅读，对稿件中的文字错误进行了修改。针对英文缩写（ERP、FRN）首次出现时没有全称的问题，我们在稿件中已进行了修改。针对稿件中文字错误的修改，详见文中红色字体部分，请审稿人审查。

审稿人 2 意见:

意见 1: 本文探讨了一个有趣的研究问题，即自恋者的反馈加工敏感性问题，审稿人对研究的设计并无异议。

回应: 感谢审稿专家对本研究的意义和价值的肯定。

意见 2: 研究的样本量组成有偏差问题。对于一个研究分组和个体差异的研究而言，仅 27 个被试，每组被试不足 15 人，代表性有欠缺，对于这样一个有趣的研究主题，审稿人希望看到更大样本量的研究。考虑到本研究中男性被试仅 7 人，而男女被试在社会决策中的性别差异是一个重复率较高的结果，建议样本量增加到每组 20 人，尤其是男性被试；

回应: 针对研究的样本量问题，在认真思考了两名审稿专家的意见后，我们努力对本研究的被试样本量进行了随机取样补充，将研究的被试数量增加至 38 人(增加了 11 名男生被试)，其中男生 18 人，女生 20 人。我们对本研究的数据进行重新分析，以下是我们对被试、问卷测量结果和行为结果的描述和分析。

对被试描述的修改和补充如下：“38 名(20 名女生，18 名男生)在校大学生参与实验，年龄为 19.7 ± 1.5 岁。所有被试均为右利手，视力或者矫正视力正常，无精神病史。实验前签署知情同意书，实验后根据被试在博弈任务中的最终收益，给予一定报酬。”详见文中被试描述部分的红色字体。

对问卷测量结果的分析如下：“将 38 名被试在自恋人格问卷上的平均得分进行 T 分数转换后由低到高排序，以 50%百分位数的分数为标准，将分数在 50%以下的 19 名被试(9 男)被划分为相对低自恋组，分数在 50%以上的 19 名被试被划分为相对高自恋组。相对高自恋组的分数(57.05 ± 7.38)显著高于相对低自恋组的分数(42.95 ± 6.79)， $t(36)=6.14$ ， $p<.001$ 。对两组被试在贝克抑郁量表上的得分进行独立样本 t 检验，结果表明两组被试的抑郁水平不存在显著差异， $t(36)=.56$ ， $p=.58$ 。”详见稿件中结果的问卷测评结果部分。

对行为结果的分析如下：“对两组被试的信任选择率进行独立样本 t 检验，结果发现相对高自恋组的

信任选择率(60.3±13%)显著低于相对低自恋组的信任选择率(67.5±5.3%), $t(36)=2.25, p<.05$ 。

相对低自恋者做出信任和信任决策的平均反应时分别为 583.5±227ms 和 555.8±192ms; 相对高自恋者做出信任和信任决策的平均反应时分别为 528.8±227.7ms 和 518.5±199ms。对被试的反应时进行 2(自恋水平: 相对高自恋, 相对低自恋)×2(信任抉择: 信任, 不信任)混合方差分析, 结果发现, 自恋水平和信任抉择的主效应, 以及它们之间的交互效应均不存在显著性($ps>.05$)。”详见稿件中结果的行为结果部分。以上修改内容, 详见文中相应内容的红色字体部分, 请审稿人审查。

意见 3: 从地形图来看, P3 最大的位置在中间位置, 分析选点 FCz 有待商榷。且 P300 的效应似乎在 loss 的时候有组间差异, 希望后续的分析能验证一下这一推论;

回应: 针对审稿专家对 P300 成分提出的意见, 我们将样本量补充到 38 人(增加了 11 名男生被试)后, 对本研究的数据进行了重新分析。P300 波幅最大的位置在中间位置(Cz 点), P300 在实验条件之间的最大差异在 FCz 点, 这两个点我们都纳入了测量分析。以下是我们对 P300 分析结果。

对 P300 的结果分析如下:“对 P300(300~600ms)峰值进行 2(自恋水平)×2(反馈条件)×5(电极分布)的三因素混合方差分析, 结果表明电极分布的主效应显著, $F(4, 144)=13.79, p<.001, \eta^2=.277$, 表现为 P300 峰值从大脑前部到后部呈倒 U 型, 额区(Fz)峰值最小(18.04±1.42μV), 中央区(Cz)峰值最大(22.14±1.25μV), 顶区(Pz)峰值次小(18.46±.86μV); 电极分布和自恋水平的交互作用显著, $F(4, 144)=3.67, p<.05, \eta^2=.092$, 简单效应分析发现在 FCz 上相对高自恋者的 P300 峰值(24.43±1.94μV)与相对低自恋者的 P300 峰值(19.03±1.94μV)之间的差异边缘显著($p=.058$), 而在其他电极位置两组被试的 P300 峰值之间不存在显著性差异($ps>.05$); 反馈的主效应不显著, $F(1, 36)=1.39, p=.246, \eta^2=.04$, 互惠获利的 P300 峰值(20.71±1.22μV)与信任损失的 P300 峰值(19.89±1.14μV)无显著差异。”详见稿件中结果的 P300 结果部分。

我们也对大脑中线 5 个电极的 P300 峰值分别进行 2(自恋水平)×2(反馈条件)的二因素混合方差分析, 结果发现在 FCz 上相对高自恋者的 P300 峰值(24.43±1.94μV)与相对低自恋者的 P300 峰值(19.04±1.95μV)之间的差异边缘显著($p=.058$), 其他主效应和交互效应不显著($ps>.05$)。而在 Fz、Cz、CPz、Pz 上自恋水平和反馈条件的主效应, 及二者之间的交互效应均不存在显著性($ps>.05$)。

此外, 我们还对 P300 峰值在信任损失(loss)和互惠获利(gain)上的组间差异分别进行了独立样本 T 检验, 结果发现, 在 FCz 点存在信任损失的组间差异, $t(36)=2.063, p<.05$, 相对高自恋者的 P300 峰值(24.3±1.92μV)显著大于相对低自恋者的 P300 峰值(18.43±2.1μV), 而各电极点上均不存在互惠获利的组间差异($p>.05$)。

意见 4: 作者在讨论中一部分提到共情缺失, 另外一个讨论部分是关于 FRN 反映了自恋个体对意料之外的反馈更敏感。对于前者, 没有采用 IRI 问卷测试被试的共情稍显不足, 可列为研究局限性; 对于后者意

料之外的反馈，应该更多指的对负性反馈更敏感，因此还是建议除差异波外，区分为正性负性反馈进行分析。

回应：针对审稿专家对讨论提出的意见，本研究中并未测量被试的共情，我们删除了有关共情的直接推论，并在修改后的稿件中我们将此列为研究的局限性。

针对审稿专家提出的 FRN 的测量和分析问题，在对补充被试后的实验数据分析中，我们分别对 FRN 的原始波和差异波进行测量和分析。对 FRN 的原始波和差异波的测量进行了如下描述：“FRN 测量的是反馈呈现后 230~280ms 的平均波幅，P300 测量的是反馈呈现后 300~600ms 的峰值。鉴于 FRN 可能会受到晚期正成分(P300)的影响，我们也对 FRN 差异波(dFRN)进行了测量，其测量的是损失反馈减获利反馈所获得的波形在 230~280ms 内的平均波幅。根据已有研究(Gehring & Willoughby, 2002; Yeung & Sanfey, 2004)，以及对 ERP 总平均波形和脑地形图的视觉检测(见图2)，我们选取脑中线5点(Fz、FCz、Cz、CPz、Pz)进行分析。即对 FRN 和 P300 进行 2(自恋水平：高、低)×2(反馈条件：信任损失、互惠获利)×5(电极分布：Fz、FCz、Cz、CPz、Pz)的三因素混合方差分析，对 dFRN 进行 2(自恋水平：高、低)×5(电极分布：Fz、FCz、Cz、CPz、Pz)的二因素混合方差分析。”

我们对 FRN 的原始波进行了如下的分析：“对 FRN(230~280ms)的平均波幅进行 2(自恋水平)×2(反馈条件)×5(电极分布)的三因素混合方差分析，结果表明不同反馈条件的 FRN 平均波幅差异显著， $F(1, 36)=65.56, p<.001, \eta^2=.65$ ，信任损失的 FRN 平均波幅($9.65\pm 1.06\mu V$)显著大于互惠获利的 FRN 平均波幅($14.11\pm 95\mu V$)；反馈条件与自恋水平的交互作用显著， $F(1, 36)=5.94, p<.05, \eta^2=.14$ ，简单效应分析发现，在相对高自恋者中，信任损失的 FRN 波幅($9.21\pm 1.49\mu V$)显著大于互惠获利的 FRN 波幅($15.02\pm 1.34\mu V$)， $F(1, 36)=55.48, p<.001, \eta^2=.61$ ；在相对低自恋者中，信任损失的 FRN 波幅($10.08\pm 1.49\mu V$)显著大于互惠获利的 FRN 波幅($13.2\pm 1.34\mu V$)， $F(1, 36)=16.02, p<.001, \eta^2=.31$ ，相对高自恋者的差异量更大，这也就与 dFRN 的统计结果相一致。”

我们对 FRN 的差异波进行了如下的分析：“对 dFRN(230~280ms)的平均波幅进行 2(自恋水平)×5(电极分布)的二因素混合方差分析，结果表明不同水平自恋者的 dFRN 平均波幅差异显著， $F(1, 36)=5.64, p<.05, \eta^2=.14$ ，相对高自恋者的 dFRN 平均波幅($-5.81\pm .79\mu V$)比相对低自恋者的 dFRN 平均波幅($-3.19\pm .79\mu V$)更大。电极位置与自恋水平的交互作用显著， $F(4, 144)=3.34, p=.042, \eta^2=.085$ ，简单效应分析发现在额区(Fz)、额中区(FCz)和中央区(Cz)相对高自恋者的 dFRN 平均波幅均显著大于相对低自恋者的 dFRN 平均波幅($p<.05$)，但在中央顶区(CPz)和顶区(Pz)上两组被试的 dFRN 平均波幅不存在显著差异($p>.05$)。”详见文中“脑电记录和分析”和 ERP 结果的 FRN 和 dFRN 部分的红色字体，请审稿人审查。

第二轮

审稿人 1 意见:

意见 1: 作者对原始波的分析不能令人满意, 重点不是差异量, 而是两组间的差异是不是只存在于赢条件而不存在于输条件, 或者两组间的差异在赢条件和输条件都存在。从波幅值和地形图上看, 很可能组间差异只存在于赢条件而不是输条件。如果真是如此, 那么作者很可能需要重新讲述这个故事了。有些研究个体差异的文献的确是只发现组间差异只反映在赢钱条件而在输钱条件下无差异, 例如, 病理性赌博 Hewig et al., 2010; 抑郁个体 Liu et al., 2014。但是, 如果作者打算采用差异波的计算方法, 那就把原始波的统计分析删掉吧, 也完全可以。

回应: 我们认真思考了审稿专家的意见, 确实发现近年来有关信任博弈的 ERP 研究采用 dFRN 为指标(Chen et al., 2012; Ma et al., 2015)。因此, 综合审稿专家意见和已有相关研究, 我们将 FRN 原始波的统计分析删除, 采用差异波的计算方法(详见正文结果部分), 这样也能够更好地解决 P3 对 FRN 的混淆。另外, 我们将审稿专家提出的宝贵建议列入了本研究的局限和未来研究方向, 具体如下“第三, 当前在极端个体(Hewig et al., 2010)或临床个体(Liu et al., 2014)中对 FRN 的研究表明在奖励条件下存在 FRN 组间差异。然而, 本研究在一般人群中以二分方式选取研究对象并未发现 FRN 在奖励或惩罚条件中的组间差异。将来亦可在一般人群中选取极端自恋个体或在临床自恋个体中探讨自恋人格对信任博弈结果评价的影响”。请审稿专家予以审查。

意见 2: P12 最后一段第二行关于 FRN 的时间范围介绍的太窄了, 建议作者修改一下; 一些数字符号应与两侧的数字之间有个空格, 例如 $528.8 \pm 227.7\text{ms}$ 。

回应: 我们对 FRN 时间范围进行了如下修改: “反馈相关负波(feedback related negativity, FRN), 它是反馈结果呈现后 250~350ms 左右出现于额中区(frontocentral)的负走向成分, 输钱比赢钱的波幅更大, 其源定位于前扣带回附近(anterior cingulate cortex, ACC)(Gehring & Willoughby, 2002)。”有关审稿专家提出的“数字符号应与两侧的数字之间有个空格”的问题, 文中已进行了相应的修改。详见文中的绿色字体部分, 请审稿专家予以审查。

审稿人 2 意见:

意见 1: 分组问题的表述, 由于 38 个被试全进入分析, 全进行分组, 19 个高 19 个低。那么 T 分数转换则没什么意义。

回应: 遵照审稿专家对分组问题表述所提出的意见。我们采用原始分数, 对分组的表述进行了如下修改:

“将 38 名被试的自恋人格问卷平均得分进行由低到高排序，以 50%百分位数的分数为标准，将分数在 50% 以下的 19 名被试(9 男)划分为相对低自恋组，分数在 50%以上的 19 名被试划分为相对高自恋组。相对高自恋组的平均得分(3.52 ± 0.38)显著高于相对低自恋组的平均得分(2.8 ± 0.35)， $t(36) = 6.14, p < 0.001$ 。对两组被试在贝克抑郁量表上的得分进行独立样本 t 检验，结果表明两组被试的抑郁水平不存在显著差异， $t(36) = 0.56, p = 0.58$ 。”详见文中问卷测评结果中的绿色字体部分，请审稿专家予以审查。

意见 2：原始波的分析是否比较过 H-gain vs. L-gain 和 H-loss vs. L-loss，这一结果决定了作者的讨论部分是否有理有据。还是与初稿一样的疑问，FRN 反映了个体对意料之外的反馈更敏感，究竟是哪一类反馈更敏感呢？这一推论需要有研究结果支持。目前只有高自恋者更少信任他人的行为数据。

而作者讨论 FRN 结果时也忽略了这两个组间对比究竟是否有差异。如“当高自恋者决定将自身所拥有的金钱投资给对方时，期待对方做出互惠决策的愿望也更强烈，因此，当反馈结果为信任损失时相对高自恋者体验到更大的预期偏差，从而导致相对高自恋者产生更大的 dFRN 波幅。”这里已经是对差异波进行解释，而差异波是输赢一起的结果，这里合理的结果实际上是 H-loss 的 FRN 比 L-loss 的 FRN 更负才能支持作者这一讨论。

回应：非常感谢审稿专家对本文所提出的宝贵意见。结合另一位审稿专家的建议，而且也确实发现近年来有关信任博弈的 ERP 研究采用 dFRN 为指标(Chen et al., 2012; Ma et al, 2015)。因此，综合审稿专家意见和已有相关研究，我们将 FRN 原始波的统计分析删除，采用差异波的计算方法（详见正文结果部分），这样也能够更好地解决 P3 对 FRN 的混淆。

针对审稿专家提出的关于讨论部分的意见，我们对讨论部分进行了修改，删除了针对 FRN 原始波结果的句子，更为谨慎地讨论如下：“结果评价的 ERP 结果显示，相对高自恋者的 dFRN 波幅显著大于相对低自恋者的 dFRN 波幅。预期偏差理论(expectancy deviation hypothesis)认为反馈负波是由个体实际得到的反馈与预期不一致所诱发的，它反映了结果反馈与事先预期之间的偏离程度，意料之外的反馈诱发的波幅更大(San Mart ín, 2012; Oliveira, McDonald, & Goodman, 2007)。有研究发现对社会规则的违背会产生更大的波幅(Luo et al., 2014; Mussel et al., 2014; Osinsky et al., 2013)。自恋者具有自我夸张性，对与他人互动结果的反馈更敏感(Morf, & Rhodewalt, 2001)，有可能是导致高自恋者 dFRN 波幅更大的原因。高自恋者决定信任他人时，对互动对象做出互惠决策的期待可能也更强烈，当这种期待未得到满足时高自恋者体验到的预期偏差更大，从而导致高自恋者产生更大的 dFRN 波幅。这一定程度上反映了具有更高自恋特质的个体在人际互动过程中更关注他人的反馈，对来自他人的反馈更敏感。”但上述讨论所能够得出的结论较为虚弱，需要在未来研究中用量表直接测量被试在输赢不同条件下的期望和敏感程度。我们下一步的研究也从审稿专家的建议中获得启发。

意见 3: 讨论部分多次提到自恋者对他人反馈的关注,也可能是表述问题。审稿人理解为,对他人反馈的关注需要呈现他人的结果,而研究中不是如此。作者想讨论的可能是自恋者更在乎于他人互动的结果。

回应: 在讨论中提及的“自恋者对他人反馈的关注”的目的正是审稿专家说的“自恋者更在乎于他人互动的结果”,我们已统一更改为“.....与他人互动的结果”。正如审稿专家认为的“对他人反馈的关注需要呈现他人的结果”,本研究中并未呈现他人的结果。当前已有的采用信任博弈范式(trust game)对信任或合作的研究中为了使结果呈现更加简洁,在被试做出信任与否的决策后均呈现的是双方互动导致的被试收益情况(Chen et al., 2012; Ma et al, 2015; 王益文等, 2015),基于上述研究,在本研究中实验程序的结果反馈阶段也只呈现被试在双方互动中产生的结果。

Chen, J., Zhong, J., Zhang, Y., Li, P., Zhang, A., Tan, Q., & Li, H. (2012). Electrophysiological correlates of processing facial attractiveness and its influence on cooperative behavior. *Neuroscience Letters*, 517(2), 65-70.

Ma, Q., Meng, L., & Shen, Q. (2015). You Have My Word: Reciprocity Expectation Modulates Feedback-Related Negativity in the Trust Game. *PloS One*, 10(2), e0119129.

王益文, 张振, 原胜, 郭丰波, 何少颖, 敬一鸣. (2015). 重复信任博弈的决策过程与结果评价. *心理学报*, 47(8), 1028-1038.

意见 4: 一些表述性问题, 1、“研究者从自恋的自我特征角度对其神经生理机制进行探讨,发现额中回、内侧额叶、眶额叶皮层、楔前叶和前脑岛涉及自我提升和自我评估(Kwan et al., 2007; Barrios et al., 2008; Beer, & Hughes, 2011); ...”翻译痕迹过重,还是应注意中文表述。此处“涉及”可改为“参与”。

2、“信任者(trustor)首先被...”,此处 trustor 也不宜直接翻译为信任者,“委托人/信托人”

3、信任损失的 FRN 平均波幅($9.65 \pm 1.06 \mu\text{V}$)显著大于互惠获利的 FRN 平均波幅($14.11 \pm 9.95 \mu\text{V}$);此处显著大于的表述有歧义,建议将类似表述都改为有更负的 FRN

4、这与大量的已有研究一致(San Mart ín, 2012)。既然是大量已有研究,应该多增加一些参考文献为依据。

回应: 针对这些表述性问题,我们已根据审稿专家进行了修改,详见文中的绿色字体部分,请审稿专家予以审查。