

自恋人格调节信任博弈的结果评价*

王益文¹ 付超¹ 任相峰¹ 林羽中¹
郭丰波¹ 张振¹ 黄亮² 袁博³ 郑玉玮⁴

(¹ 福州大学心理与认知科学研究所, 福州 350116) (² 闽南师范大学应用心理研究所, 漳州, 363000)

(³ 宁波大学心理学系暨研究所, 宁波 315211) (⁴ 济南大学教育与心理科学学院, 济南 250022)

摘要 自恋人格个体表现出亲社会行为的减少和对他人反馈敏感。信任是人类社会互动行为的基础, 自恋人格特质可能影响他们的信任决策。为了探究自恋水平如何影响信任博弈中结果评价的大脑活动, 本研究采用事件相关电位技术记录了 38 名被试完成单次信任博弈时的脑电波。行为结果发现相对高自恋者的信任选择率显著低于相对低自恋者的信任选择率。脑电结果表明, 相对高自恋者信任损失反馈减互惠获利反馈所得的 FRN 差异波 (dFRN) 显著地大于相对低自恋者, 并且相对高自恋者结果评价诱发的 P300 波幅显著地大于相对低自恋者。本研究提供了自恋人格个体信任博弈中结果评价的初步神经电生理学证据。

关键词 信任博弈; 自恋人格; 结果评价; dFRN; P300

分类号 B849:C91; B845

1 引言

自恋(narcissism)是一种与傲慢、权力感和剥削他人的意愿有关的人格特质(Raskin & Terry, 1988; Campbell & Foster, 2007), 它对个体的认知、情感和行为具有重要影响。新近研究表明东西方文化中自恋倾向均越来越明显(Cai, Kwan, & Sedikides, 2012; Twenge, Campbell, & Gentile, 2012)。自恋与不同的自我特征有关, 如夸张性、崇拜的需要、缺乏共情和述情障碍(Campbell & Foster, 2007), 而自恋的夸张性和崇拜的需要源于扭曲的自我观点和自我提升(Morf & Rhodewalt, 2001; Maxwell, Donnellan, Hopwood, & Ackerman, 2011)。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、楔前叶和前脑岛参与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)。Morf 和 Rhodewalt (2001)认为自恋者是自我中心的、夸张的, 且对与他人互动结果的反馈敏感。自恋者聚焦于个人利益, 较少考虑个人行为对他人的影响, 表现出对人际关系、公共价值和亲社会行为兴趣的消退(Campbell & Foster, 2007; Cisek, Hart, & Sedikides, 2008)。最近研究发现自恋者表现出个人利益为中心的决策特点(Campbell, Bush, Brunell, & Shelton, 2005; Brunell, Tumbli, & Buelow, 2014), Campbell 等(2005)采用资源困境(resource dilemmas)范式探讨自恋者的行为决策, 被试代表一个林业公司与其他公司共同从可再生森林中采伐木材。结果发现自恋者在初始阶段选择采伐更多木材, 导致资源快速耗竭, 个人短期获益但整体收益更少。这表明自恋者的决策以个人利益为中心, 而对他人造成长期的破坏性。社会决策中个体的策略倾向与其社会偏好有关, 亲社会者具有合作的内在意愿, 而亲自我者的合作行为需要有个个人获益的外在动机信号(Boone, Declerck, & Kiyonari, 2010; Emonds, Declerck,

收稿日期: 2015-12-21

* 国家自然科学基金(31371045)、新世纪优秀人才支持计划(NCET-11-1065)、三年千人计划和中青年骨干创新人才资助。

通讯作者: 张振, E-mail: zhangzhenpsy@126.com; 袁博, E-mail: yuanbopsy@gmail.com

Boone, Seurinck, & Achten, 2014; Kuss et al., 2015)。信任是人类社会互动的关键成分,它是合作的基础,对个人在复杂社会环境中成功运作至关重要。

信任是个体在不确定对方的动机、意图和行为时,乐意将自身所拥有的资源交予对方处置并承受相应风险的意愿(Mayer, Davis, & Schoorman, 1995; Moretto, Sellitto, & di Pellegrino, 2013)。信任博弈(Trust Game, TG)是研究信任行为的经典实验范式(Berg, Dickhaut, & McCabe, 1995),它是在两个匿名玩家之间进行的一轮游戏(one-shot game),委托人(trustor)首先被分配一些代币,他可以选择全部保留或投资部分代币给受托人(trustee),投资给受托人的代币将会翻倍,然后受托人决定是保留还是返还部分给委托人。研究表明信任博弈中存在普遍的合作行为,即委托人会选择投资部分金钱,而受托人也会返还部分金钱(Tzieropoulos, 2013)。当前对信任的神经生理机制研究发现,委托人做出信任行为时会激活内侧前额叶皮层(medial prefrontal cortex, mPFC),背外侧前额叶皮层(dorsolateral prefrontal cortex, dLPFC)和右侧颞顶联合区(right temporoparietal junction, rTPJ)等大脑皮层(Delgado, Frank, & Phelps, 2005; King-Casas et al., 2005; Krueger et al., 2007)。研究者认为 mPFC 和 rTPJ 参与心理理论,如推测他人的意图或互动策略,dLPFC 激活可以促使个体克服自私利益诱惑,选择互利合作行为(Declerck, Boone, & Emonds, 2013)。最近研究者从个体社会偏好角度对信任行为进行研究(Kanagaretnam, Mestelman, Nainar, & Shehata, 2009; Komiya & Mifune, 2015), Derks, van Scheppingen, Lee 和 Krabbendam (2015)采用信任博弈任务探讨了不同社会价值取向个体的信任行为,结果发现亲自我者比亲社会者具有更低的信任和可信度水平。个体做出信任行为时需要克服自私利益的诱惑,进而选择互惠利他行为,但自恋者的行为决策以自我利益为中心,因此自恋者在信任博弈过程中可能做出更少的信任决策。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,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, Greig, Herrmann, & Zeckhauser, 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, Anderson, Schlösser, Ehlebracht, & Fetchenhauer, 2014),而高自恋者的行为决策(赌博和投资)具有明显风险寻求倾向(Campbell, Goodie, & Foster, 2004; 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。

已有研究发现结果评价相关的脑电成分为:反馈相关负波(feedback related negativity, FRN),它是反馈结果呈现后 250 ~ 350 ms 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)。研究者认为 FRN 反映的是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 FRN 波幅更大(San Martín, 2012; Oliveira, McDonald, & Goodman, 2007)。也有研究者认为 FRN 反映的是个体对反馈刺激情绪动机意义的评价(Gehring & Willoughby, 2002; Wang et al., 2013)。当前信任博弈的 ERP 研究中主要采用 FRN 差异波(dFRN)为指标(Chen et al., 2012; Ma, Meng, & Shen, 2015),它由损失反馈减获利反馈所获得。Chen 等(2012)采用信任博弈任务探讨面孔吸引力对合作行为的影响,结果表明相对于低面孔吸引力对手,被试与高面孔吸引力对手互动时信任博弈结果反馈的 dFRN 更大,这表明高面孔吸引力增加了被试的奖励期待。Ma 等(2015)采用单次信任博弈任务探讨承诺对委托人决策的影响,结果发现违背承诺时结果反馈的 dFRN 更大,这说明承诺增加了个体的互惠期待。

综上所述,本研究采用事件相关电位(Event Related Potentials, ERP)技术和信任博弈任务,考察不同自恋倾向个体在信任博弈中的行为特征及其结果评价时的大脑活动。鉴于自恋者对亲社会行为兴趣的消退(Campbell & Foster, 2007; Cisek et al., 2008),且表现出个人利益为中心的行为特征(Campbell et al., 2005),本研究假设一:高自恋者的信任选择率低于低自恋者。对于结果评价的脑电结果,FRN 反映的是反馈结果与事先预期之间的偏离程度,意料之外的损失诱发的 FRN 波幅更大(San Martín, 2012)。并且已有研究认为自恋个体是自我为中心的,且与他人互动结果的反馈很敏感(Morf & Rhodewalt, 2001)。本研究假设二:与低自恋者相比,高自恋者对信任决策的结果反馈的 dFRN 更大。

2 方法

2.1 被试

38名(20名女生)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。

2.2 任务和程序

本研究采用信任博弈游戏,根据 Berg 等(1995)的投资/信任游戏改编而来的,被试作为委托人与其他玩家一起完成 150 轮游戏,同时记录其脑电活动。每轮游戏前,委托人和受托人各有 10 点出场费,被试决定是否将 10 点出场费交予受托人。如果被试选择保留,则该轮游戏结束,双方各获 10 点出场费;如果选择将这 10 点出场费交予受托人,则这些点数将会翻 3 倍交予受托人,此时,由受托人决定如何分配当前其所拥有的点数(10 点出场费加翻倍后的 30 点),其拥有平分或独吞两种分配方案,无论选择哪种分配方案,当前一轮游戏结束且双方获得相应的点数。

当前信任游戏采用一次性(one-shot)匿名互动方式进行,即每轮游戏对手均不相同。具体操作如下:实验时告知被试在脑电实验前实验者选取了 400 名具有社会代表性的成人,要求他们指出“在一次性匿名互动游戏中,如果其被他人信任且给予了 30 点,那么他们将会平分还是独吞这些点数?”然后将这 400 人的选择贮存于计算机之中;在脑电实验中每轮互动时,计算机都会从中随机挑选出一个选择作为受托人的选择以完成该轮游戏。事实上,所有受托人的选择都是由实验前所编写的实验程序控制的,以保证整个任务中结果强化率(信任选择后正性反馈的比例)为 50%。

被试到达实验室后,实验者首先向其介绍游戏规则,并填写个人基本信息和知情同意书。随后,被试被带入电磁屏蔽良好的脑电实验室,舒适地坐在距离电脑屏幕 100 cm 左右的椅子上,并开始脑电实验的准备工作。正式实验中被试需要完成 150 轮信任游戏,每轮游戏的刺激呈现顺序如下,首先呈现信任游戏的简易决策树(1500 ms),提示被试当前任务中可能的选项与结果。在一个 800 ~ 1000 ms 随机呈现的“+”注视点之后,屏幕上呈现一个决策选项图(2000 ms);被试需要在决策选项图呈现时间内进行按键做出信任或不信任的抉择,信任按“1”,不信任按“3”,按键方式在被试间平衡,被试若在 2000 ms 内未按键,则提示其未做出按键反应,并要求其重新开始这一轮游戏。随后,在一个 800 ~ 1200 ms 随机呈现的黑屏后,屏幕上呈现当前一轮游戏的结果(1200 ms)。游戏分两个部分,每部分结束后呈现总收益(2000 ms)。

2.3 问卷测量

2.3.1 自恋人格问卷

采用简版自恋人格问卷(A brief version of Narcissistic Personality Inventory, NPI-16)测量被试的自恋人格倾向(Ames, Rose, & Anderson, 2006),它在 NPI-40 (Raskin & Terry, 1988)的基础上编制的。该量表用于对一般人群(general population)自恋的测量,其内部一致性系数为 0.72,重测信度为 0.85。为了获取被试更多的反应可变性(response variability) (Pincus et al., 2009),本研究采用 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。

2.3.2 贝克抑郁量表

为了排除抑郁作为混淆变量的可能性,采用贝克抑郁量表(Beck Depression Inventory-II, BDI-II)测量被试的抑郁倾向(Fan et al., 2011)。BDI-II 是根据精神障碍诊断与统计手册第四版(DSM-IV)中抑郁症诊断标准,对贝克抑郁量表第 1 版进行修订重新编制而成。该量表由 21 个项目组成,包括一般、认知-情感和躯体症状三个因子。贝克抑郁量表中文版(BDI-II-C)在非临床青少年样本中内部一致性系数为 0.89,重测信度为 0.76 (杨文辉等, 2014)。

2.4 脑电记录与分析

采用 NeuroScan ERP 记录和分析系统(NeuroScan Inc., USA),以国际 10-20 系统扩展的 64 导电帽记录 EEG。脑电记录时所有电极以置于左侧乳突的电极作为参考,离线分析时再以置于右侧乳突的电极进行再参考(re-reference),即从各导联信号中减去 1/2 该参考电极所记录的信号,转化为以双侧乳突平均值为参考。双眼外侧安置电极记录水平眼电(HEOG),左眼上下安置电极记录垂直眼电(VEOG)。每个电极与头皮之间的阻抗始终都保持在 5 k Ω 以下。滤波带通为 0.05 ~ 100 Hz, AC 采样,采样率为 1000 Hz/导。对所采集的脑电数据进行离线分析,用 Scan 软件矫正 VEOG 和 HEOG,自动排除其他波幅大于 $\pm 75 \mu\text{V}$ 的伪迹信号,所得 ERP 波形进行 30 Hz 低通滤波。

本研究主要关注结果反馈的脑电波,分析时程为结果反馈呈现前 200 ms (基线)到呈现后 1000 ms,分别叠加被试做出信任抉择后损失反馈(Loss)和收益反馈(Gain)两种条件下的 ERP 波形图。我们主要聚焦于结果评价相关的 ERP 成分:dFRN 和 P300, dFRN 测量的是信任损失减互惠获利所得的波形在 230 ~ 280 ms 内的平均波幅, P300 测量的是反馈呈

现后 300 ~ 600 ms 的峰值。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 1), 我们选取脑中线 5 点(Fz/FCz/Cz/CPz/Pz)进行分析。即对 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。对 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。其中自恋水平为被试间变量, 反馈条件和电极分布为被试内变量。对所有主效应和交互作用的 p 值均采用 Greenhouse-Geisser 法校正, 事后配对比较采用 Bonferroni 法校正。

3 结果

3.1 问卷测评结果

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, $t(36) = 0.56$, $p = 0.58$ 。

3.2 行为结果

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($60.3\% \pm 13\%$)显著低于相对低自恋组的信任选择率($67.5\% \pm 5.3\%$), $t(36) = 2.25$, $p < 0.05$ 。

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 ± 227 ms 和 555.8 ± 192 ms; 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 ± 227.7 ms 和 518.5 ± 199 ms。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作用均不存在显著性($ps > 0.05$)。

3.3 ERP 结果

3.3.1 dFRN

对 dFRN (230 ~ 280 ms)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 波幅差异显著, $F(1, 36) =$

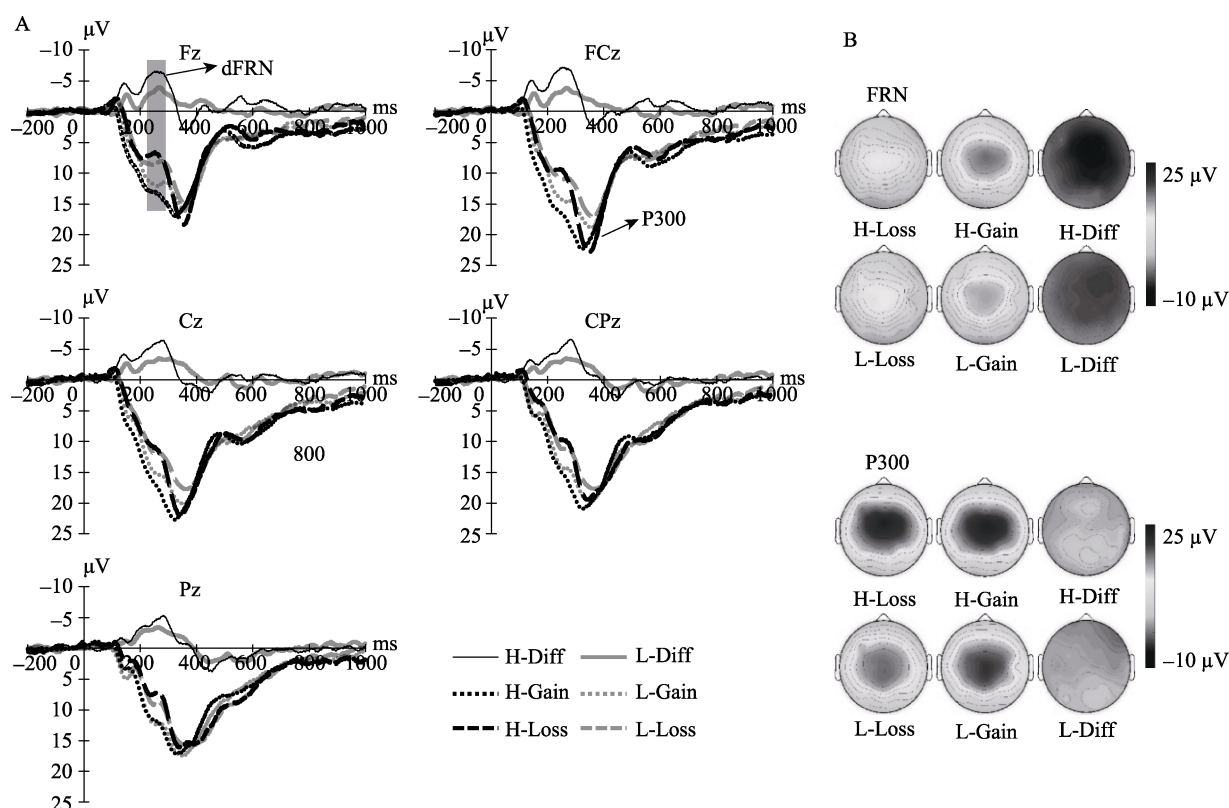


图 1 A 为两组被试在不同反馈条件下的 ERP 波形图; B 为 FRN (230~280 ms)平均波幅和 P300 (300~600 ms)峰值的脑地形图(H 代表相对高自恋组, L 代表相对低自恋者, Loss 为信任损失反馈, Gain 为互惠获利反馈, Diff 为信任损失反馈减互惠获利反馈的差值)。

5.64, $p < 0.05$, $\eta^2 = 0.14$, 相对高自恋者的 dFRN 波幅($M = -5.81 \mu V$, $SE = 0.79$)比相对低自恋者的 dFRN 波幅($M = -3.19 \mu V$, $SE = 0.79$)更负。电极分布的主效应显著, $F(1, 36) = 3.916$, $p < 0.05$, $\eta^2 = 0.1$, 额中区(FCz)的 dFRN 波幅($M = -5.09 \mu V$, $SE = 0.59$)比中央区(Cz) ($M = -4.44 \mu V$, $SE = 0.65$)和顶区(Pz) ($M = -3.86 \mu V$, $SE = 0.57$) 的 dFRN 波幅更负($ps < 0.05$); 电极分布与自恋水平的交互作用显著, $F(4, 144) = 3.34$, $p < 0.05$, $\eta^2 = 0.09$, 简单效应分析发现在额区(Fz)、额中区和中央区相对高自恋者的 dFRN 波幅比相对低自恋者的 dFRN 波幅更负($ps < 0.05$), 但在中央顶区(CPz)和顶区上两组被试的 dFRN 波幅不存在显著差异($ps > 0.05$)。

3.3.2 P300

对 P300 (300 ~ 600 ms)峰值进行 2(自恋水平) \times 2(反馈条件) \times 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, $F(4, 144) = 13.79$, $p < 0.001$, $\eta^2 = 0.28$, 表现为 P300 峰值从大脑前部到后部呈倒 U 型, 额区峰值最小($M = 18.04 \mu V$, $SE = 1.42$), 中央区峰值最大($M = 22.14 \mu V$, $SE = 1.25$), 顶区峰值次小($M = 18.46 \mu V$, $SE = 0.86$); 电极分布和自恋水平的交互作用显著, $F(4, 144) = 3.67$, $p < 0.05$, $\eta^2 = 0.09$, 简单效应分析发现在 FCz 上相对高自恋者的 P300 峰值($M = 24.43 \mu V$, $SE = 1.94$)与相对低自恋者的 P300 峰值($M = 19.03 \mu V$, $SE = 1.94$)之间的差异边缘显著($p = 0.058$), 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($ps > 0.05$); 反馈的主效应不显著, $F(1, 36) = 1.39$, $p = 0.25$, $\eta^2 = 0.04$, 互惠获利的 P300 峰值($M = 20.71 \mu V$, $SE = 1.22$)与信任损失的 P300 峰值($M = 19.89 \mu V$, $SE = 1.14$)无显著差异。

4 讨论

当前采用 ERP 技术探讨自恋者在信任博弈中结果评价时大脑活动特点的研究尚显不充分。本研究结合 ERP 技术和信任博弈(TG)任务, 试图揭示不同自恋倾向个体的人际信任特征及其信任决策后结果评价的大脑活动特点。本研究发现, 相对高自恋者的信任选择率显著低于相对低自恋者; 脑电结果发现, 相对于相对低自恋者, 相对高自恋者对信任决策后结果反馈的 FRN 差异波(dFRN)更大, 且相对高自恋者对结果反馈诱发的 P300 波幅更大。

4.1 自恋者的自我中心性导致信任行为的减少

信任是人类社会互动过程中的必要元素, 它在

社会和经济活动中起着润滑剂的作用, 且被视为合作行为产生的前提。当前研究中被试的信任选择率均高于 50%, 这与已有研究一致(Ma et al., 2015; 王益文等, 2015), 即人们一般在信任博弈过程中表现出广泛的合作行为。本研究中相对高自恋者的信任选择率显著低于相对低自恋者。已有研究表明自恋者是自我中心的(Morf & Rhodewalt, 2001); 关注个人利益, 很少考虑个人行为对他人的影响, 且存在对亲社会行为兴趣的消退(Campbell & Foster, 2007; Cisek et al., 2008), 最近研究表明自恋者的行为决策以自我利益为中心(Campbell et al., 2005; Brunell et al., 2014), Brunell 等(2014)发现自恋者参与志愿服务是为了个人利益而非人道主义价值。新近对不同社会价值取向个体的研究也发现亲自我者比亲社会者具有更低的信任水平和可信度(Kanagaretnam et al., 2009; Derks et al., 2015; Kuss et al., 2015; Komiya & Mifune, 2015), Komiya 和 Mifune (2015)采用信念游戏(faith game)和标准赌博游戏(gambling game)探讨了背叛厌恶和社会价值取向的关系, 结果发现亲自我者对他人信任的水平显著低于亲社会者。因此, 自恋者的自我中心性及对亲社会行为兴趣的消退可能使他们在信任博弈过程中表现出更少的信任行为。

4.2 dFRN 反映了自恋者对意料之外的反馈更敏感

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)认为反馈相关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 Martín, 2012; Oliveira et al., 2007)。已有研究发现对社会规则的违背会产生更大的反馈相关负波(Luo et al., 2014; Mussel, Hewig, Allen, Coles, & Miltner, 2014; Osinsky, Mussel, Öhrlein, & Hewig, 2014)。自恋者具有自我的夸张性, 对与他人互动结果的反馈很敏感(Morf & Rhodewalt, 2001), 这可能是导致高自恋者 dFRN 更大的原因。当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会更强烈, 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。

情绪动机理论认为反馈相关负波反映了对负性反馈引起的情绪动机意义的评价(Gehring & Willoughby, 2002; Wang et al., 2013), 具有更强情绪动机意义的结果可以导致更大的反馈负波。自恋者在行为决策中表现出个人利益为中心的特征(Campbell et al., 2005), 且聚焦于个人利益(Campbell & Foster, 2007)。当高自恋者在信任博弈中选择信任时, 为了个人利益最大化, 他们也更期待对方做出互惠决策, 因此, 如果对方选择独吞时他们可能会体验到更强烈的背叛感, 这也可能导致高自恋者对结果反馈产生更大的 dFRN 波幅。

4.3 P300 反映了自恋者对他人反馈的关注

结果评价的 ERP 结果表明, 在额中区(FCz)高自恋者比低自恋者出现更大的 P300 波幅。已有研究表明额叶/中央区的 P300 与刺激的新异性有关, 它与事件评估行为有关(Polich, 2007), 更大的 P300 波幅反映了个体对于该刺激更敏感。由于自恋者对与他人互动结果的反馈很敏感(Morf & Rhodewalt, 2001), 高自恋者在额中区更大的 P300 波幅表明高自恋个体在结果评价的早期阶段更关注他人的反馈(评价)。此外, 本研究未发现 P300 的反馈效价效应, 这与已有研究结果一致(王益文等, 2015; Yeung & Sanfey, 2004), 即 P300 对反馈强度而非效价敏感。独立编码模型认为 P300 对结果效价不敏感, 而对结果大小敏感(Yeung & Sanfey, 2004), 更大的结果反馈诱发出更正的 P300 波幅(Gu et al., 2011; Kreussel et al., 2012), 研究者认为 P300 反映了结果评价中注意资源分配及动机/情感评估(Leng & Zhou, 2010; 王益文等, 2015)。本研究中选择信任就意味着可能获得对方独吞或互惠反馈, 然而相对于不信任决策(基线), 两类反馈的得失金额是一致的(均为 10 点)。也就是说这两种反馈对于被试具有同等的动机/情感意义, 所需消耗的注意资源也相似。因此, 两种反馈所诱发的 P300 波幅没有显著性差异。

本研究为理解具有自恋型人格特质个体的信任行为特征提供了初步行为和脑电证据, 揭示了自恋者在人际互动过程中具有更低信任倾向。神经电生理证据表明自恋者很关注与他人互动结果的反馈, 在信任博弈中更期待对方做出互利共赢抉择。然而, 本研究也存在一些不足之处。首先, 一次性匿名信任博弈虽然控制了声誉、报复等因素对玩家信任行为的影响, 但被试可能无法有效地从互动过程中获取对方声誉信息, 或根据反馈结果调整随后的行为决策, 因此本研究并不能完全有效地揭示自

恋者在真实人际互动中的信任特征, 及其对决策后反馈阶段的大脑活动特征。其次, 已有研究显示自恋者存在共情缺陷, 而在信任决策过程中需要个体共情能力的参与, 这也可能导致自恋者具有较低的信任水平, 但本研究并未对被试的共情进行测量。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。未来研究应更好的控制和改善上述不足, 本研究结果应该谨慎推论。

参 考 文 献

- Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality*, 40(4), 440-450.
- Barrios, V., Kwan, V. S. Y., Ganis, G., Gorman, J., Romanowski, J., & Keenan, J. P. (2008). Elucidating the neural correlates of egoistic and moralistic self-enhancement. *Consciousness and Cognition*, 17(2), 451-456.
- Beer, J. S., & Hughes, B. L. (2011). Self-enhancement: A social neuroscience perspective. In M. D. Alicke & C. Sedikides (Eds.), *Handbook of self-enhancement and self-protection* (pp. 49-68). New York, NY: Guilford Press.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122-142.
- Bohnet, I., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, turkey, and the United States. *The American Economic Review*, 98, 294-310.
- Boone, C., Declerck, C., & Kiyonari, T. (2010). Inducing cooperative behavior among proselves versus prosocials: The moderating role of incentives and trust. *Journal of Conflict Resolution*, 54(5), 799-824.
- Brunell, A. B., Tumbler, L., & Buelow, M. T. (2014). Narcissism and the motivation to engage in volunteerism. *Current Psychology*, 33(3), 365-376.
- Cai, H. J., Kwan, V. S. Y., & Sedikides, C. (2012). A sociocultural approach to narcissism: The case of modern China. *European Journal of Personality*, 26(5), 529-535.
- Campbell, W. K., Bush, C. P., Brunell, A. B., & Shelton, J. (2005). Understanding the social costs of narcissism: The case of the tragedy of the commons. *Personality and Social Psychology Bulletin*, 31(10), 1358-1368.
- Campbell, W. K., & Foster, J. D. (2007). The narcissistic self: Background, an extended agency model, and ongoing controversies. In C. Sedikides & S. J. Spencer (Eds.), *The self* (pp. 115-138). New York, NY, US: Psychology Press.

- Campbell, W. K., Goodie, A. S., & Foster, J. D. (2004). Narcissism, confidence, and risk attitude. *Journal of Behavioral Decision Making*, 17(4), 297–311.
- Chen, J., Zhong, J., Zhang, Y. X., Li, P., Zhang, A. Q., Tan, Q. B., & Li, H. (2012). Electrophysiological correlates of processing facial attractiveness and its influence on cooperative behavior. *Neuroscience Letters*, 517(2), 65–70.
- Cisek, S. Z., Hart, C. M., & Sedikides, C. (2008). Do narcissists use material possessions as a primary buffer against pain? *Psychological Inquiry*, 19(3-4), 205–207.
- Declerck, C. H., Boone, C., & Emonds, G. (2013). When do people cooperate? The neuroeconomics of prosocial decision making. *Brain and Cognition*, 81(1), 95–117.
- Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, 8(11), 1611–1618.
- Derks, J., van Scheppingen, M. A., Lee, N. C., & Krabbendam, L. (2015). Trust and mindreading in adolescents: The moderating role of social value orientation. *Frontiers in Psychology*, 6, 965.
- Dunning, D., Anderson, J. E., Schlösser, T., Ehlebracht, D., & Fetchenhauer, D. (2014). Trust at zero acquaintance: More a matter of respect than expectation of reward. *Journal of Personality and Social Psychology*, 107(1), 122–141.
- Emonds, G., Declerck, C. H., Boone, C., Seurinck, R., & Achten, R. (2014). Establishing cooperation in a mixed-motive social dilemma: An fMRI study investigating the role of social value orientation and dispositional trust. *Social Neuroscience*, 9(1), 10–22.
- Fan, Y., Wonneberger, C., Enzi, B., De Greck, M., Ulrich, C., Tempelmann, C., ... Northoff, G. (2011). The narcissistic self and its psychological and neural correlates: An exploratory fMRI study. *Psychological Medicine*, 41(8), 1641–1650.
- Foster, J. D., Reidy, D. E., Misra, T. A., & Goff, J. S. (2011). Narcissism and stock market investing: Correlates and consequences of cocksure investing. *Personality and Individual Differences*, 50(6), 816–821.
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 295(5563), 2279–2282.
- Gu, R. B., Lei, Z. H., Broster, L., Wu, T. T., Jiang, Y., & Luo, Y. J. (2011). Beyond valence and magnitude: A flexible evaluative coding system in the brain. *Neuropsychologia*, 49(14), 3891–3897.
- Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2010). Hypersensitivity to reward in problem gamblers. *Biological Psychiatry*, 67(8), 781–783.
- Kanagaretnam, K., Mestelman, S., Nainar, K., & Shehata, M. (2009). The impact of social value orientation and risk attitudes on trust and reciprocity. *Journal of Economic Psychology*, 30(3), 368–380.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, 308(5718), 78–83.
- Komiya, A., & Mifune, N. (2015). An individual difference in betrayal aversion: Prosociality predicts more risky choices in social but not natural domains. *Letters on Evolutionary Behavioral Science*, 6(1), 5–8.
- Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G. H., & Miltner, W. H. R. (2012). The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. *Psychophysiology*, 49(2), 207–219.
- Krueger, F., McCabe, K., Moll, J., Krieschke, N., Zahn, R., Strenziok, M., ... Grafman, J. (2007). Neural correlates of trust. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), 20084–20089.
- Kuss, K., Falk, A., Trautner, P., Montag, C., Weber, B., & Fliessbach, K. (2015). Neuronal correlates of social decision making are influenced by social value orientation- an fMRI study. *Frontiers in Behavioral Neuroscience*, 9, 40.
- Kwan, V. S. Y., Barrios, V., Ganis, G., Gorman, J., Lange, C., Kumar, M., ... Keenan, J. P. (2007). Assessing the neural correlates of self-enhancement bias: A transcranial magnetic stimulation study. *Experimental Brain Research*, 182(3), 379–385.
- Leng, Y., & Zhou, X. L. (2010). Modulation of the brain activity in outcome evaluation by interpersonal relationship: An ERP study. *Neuropsychologia*, 48(2), 448–455.
- Liu, W. H., Wang, L. Z., Shang, H. R., Shen, Y., Li, Z., Cheung, E. F. C., & Chan, R. C. K. (2014). The influence of anhedonia on feedback negativity in major depressive disorder. *Neuropsychologia*, 53, 213–220.
- Luo, Y., Wu, T. T., Broster, L. S., Feng, C. L., Zhang, D. D., Gu, R. L., & Luo, Y. J. (2014). The temporal course of the influence of anxiety on fairness considerations. *Psychophysiology*, 51(9), 834–842.
- Ma, Q. G., Meng, L., & Shen, Q. (2015). You have my word: Reciprocity expectation modulates feedback-related negativity in the trust game. *PLoS One*, 10(2), e0119129.
- Maxwell, K., Donnellan, M. B., Hopwood, C. J., & Ackerman, R. A. (2011). The two faces of Narcissus? An empirical comparison of the Narcissistic Personality Inventory and the Pathological Narcissism Inventory. *Personality and Individual Differences*, 50(5), 577–582.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734.
- Moretto, G., Sellitto, M., & di Pellegrino, G. (2013). Investment and repayment in a trust game after ventromedial prefrontal damage. *Frontiers in Human Neuroscience*, 7, 593.
- Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry*, 12(4), 177–196.
- Mussel, P., Hewig, J., Allen, J. J., Coles, M. G. H., & Miltner, W. (2014). Smiling faces, sometimes they don't tell the truth: Facial expression in the ultimatum game impacts decision making and event-related potentials. *Psychophysiology*, 51(4), 358–363.
- Nenadic, I., Güllmar, D., Dietzek, M., Langbein, K., Steinke,

- J., & Gaser, C. (2015). Brain structure in narcissistic personality disorder: A VBM and DTI pilot study. *Psychiatry Research: Neuroimaging*, 231(2), 184–186.
- Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: Expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, 19(12), 1994–2004.
- Osinsky, R., Mussel, P., Öhrlein, L., & Hewig, J. (2014). A neural signature of the creation of social evaluation. *Social Cognitive and Affective Neuroscience*, 9(6), 731–736.
- Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G. C., & Levy, K. N. (2009). Initial construction and validation of the pathological narcissism inventory. *Psychological Assessment*, 21(3), 365–379.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148.
- Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology*, 54(5), 890–902.
- San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, 6, 304.
- Schulze, L., Dziobek, I., Vater, A., Heekeren, H. R., Bajbouj, M., Renneberg, B., ... Roepke, S. (2013). Gray matter abnormalities in patients with narcissistic personality disorder. *Journal of Psychiatric Research*, 47(10), 1363–1369.
- Twenge, J. M., Campbell, W. K., & Gentile, B. (2012). Generational increases in agentic self-evaluations among American college students, 1966–2009. *Self and Identity*, 11(4), 409–427.
- Tzieropoulos, H. (2013). The Trust Game in neuroscience: A short review. *Social Neuroscience*, 8(5), 407–416.
- Wang, Y. W., Roberts, K., Yuan, B., Zhang, W. X., Shen, D. L., & Simons, R. (2013). Psychophysiological correlates of interpersonal cooperation and aggression. *Biological Psychology*, 93(3), 386–391.
- Wang, Y. W., Zhang, Z., Yuan, S., Guo, F. B., He, S. Y., & Jing, Y. M. (2015). The decision-making and outcome evaluation during a repeated trust game. *Acta Psychologica Sinica*, 47(8), 1028–1038.
- [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.]
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience*, 24(28), 6258–6264.
- Yang, W. H., Liu, S. L., Zhou, T., Peng, F., Liu, X. M., Li, L., ... Yi, J. Y. (2014). Reliability and validity of Chinese version of the beck depression inventory-II in Chinese adolescents. *Chinese Journal of Clinical Psychology*, 22(2), 240–245.
- [杨文辉, 刘绍亮, 周烜, 彭芳, 刘细梅, 李莉, ... 蚁金瑶. (2014). 贝克抑郁量表第2版中文版在青少年中的信效度. *中国临床心理学杂志*, 22(2), 240–245.]

Narcissistic personality modulates outcome evaluation in the trust game

WANG Yiwen¹; FU Chao¹; REN Xiangfeng¹; LIN Yuzhong¹; GUO Fengbo¹;
ZHANG Zhen¹; HUANG Liang²; YUAN Bo³; ZHENG Yuwei⁴

(¹ Institute of Psychological and Cognitive Sciences, Fuzhou University, Fuzhou 350116, China)

(² Institute of Applied Psychology, Minnan Normal University, Zhangzhou 363000, China)

(³ Department of Psychology, Ningbo University, Ningbo 315211, China)

(⁴ School of Education and Psychology, University of Jinan, Jinan 250022, China)

Abstract

Narcissism is a personality characteristic involving arrogance, a feeling of entitlement, and the willingness to exploit others. Narcissistic people tend to be self-centered, grandiose and sensitive to feedback from others. They are also extremely selfish, focusing solely on themselves, lacking regard for others, and showing deficits in pro-social behavior and emotional empathy. Research on Social Value Orientation (SVO) has found that proself people are less trusting and trustworthy than prosocial people. Because of his or her egoism and lack of emotional empathy, the narcissist may tend to not trust other people, especially unfamiliar ones. Relatedly, narcissism may also modulate the outcome evaluation of trust decisions. Previous social neuroscience research has revealed two ERP components related to outcome evaluation, including feedback related negativity (FRN) and P300. In this study, we aimed to investigate how narcissistic personality modulates outcome evaluation in the one-shot trust game.

38 healthy undergraduate participated in our experiment. To assess their narcissism and control for potential depression, all participants completed a brief version of narcissistic personality inventory (NPI-16) and

Beck depression inventory-II (BDI-II) before the behavioral task. Participants then played the role of trustor in the one-shot trust game, while their electroencephalograms (EEGs) were recorded. In the trust game, each participant decide whether to trust or distrust different partners over 150 trials. The partners' reciprocation strategies were pre-programed by the experimenter (the overall reciprocating rate was 50%). All participants were provided with post-decision feedback about the outcome of their decisions (gain or loss game points) in each trail. We analyzed their behavioral responses at the decision-making stage and ERP components at the outcome evaluation stage.

Based on NPI-16 scores, 19 participants were classified as relatively low narcissists (scores below the 50th percentile), while the other 19 participants were classified as relatively high narcissists (scores above the 50th percentile); these two groups differed significantly in NPI-16 scores, but not significantly in BDI-II scores (depression). Behavioral results revealed that the high-narcissistic group made less trusting choices than did the low-narcissistic group in the trust game. ERP results indicated that the feedback-related negativity (FRN) was more negative going in response to loss feedback compared to gain feedback. This may indicate that participants typically experienced expectation violation when their trust was exploited. More importantly, the FRN difference wave (dFRN, loss-FRN minuses gain-FRN) was larger for the high narcissistic group than the low narcissistic group. Relatedly, the P300 amplitudes following outcome feedback were larger for the high narcissistic group than the low narcissistic group, especially at the frontocentral site (FCz). Both dFRN and P300 effects suggested that more narcissistic people may be more sensitive to the outcome feedback of trust decisions. The present study provides preliminary evidence that individual differences in narcissism modulate trust decision-making and outcome evaluation.

Key words Trust game; narcissistic personality; outcome evaluation; differentiated feedback related negativity; P300