

公平加工的情境依赖性：来自 ERP 的证据*

吴燕¹ 周晓林²

(¹ 杭州师范大学教科院心理学系, 杭州 310026) (² 北京大学心理学系, 北京 100871)

摘要 公平是人类社会生活的重要概念。大量研究采用最后通牒博弈, 发现人们具有不公平厌恶倾向, 即宁愿牺牲个人的经济利益, 也要拒绝不公平的分配提议。已有研究表明, 损失情境会增强不公平厌恶, 但其神经机制尚不清楚。本实验采用 ERP 技术, 运用最后通牒博弈范式, 考察两个情境因素: 域(损失或获益)和博弈对象(人或计算机)对公平加工的影响。发现损失域下对不公平分配的拒绝率更高, 而博弈对象对决策行为无影响; 获益域下, 对家为人比对家为计算机诱发了更负的 N1; 人机博弈时, 获益域比损失域诱发了更大的 P2 和 LPP, 损失域比获益域、不公平提议比公平提议诱发了更大的 N350, 而公平提议比不公平提议诱发了更大的 LPP; 人际博弈时, 这些差异均不显著。这些结果表明, 对分配提议的大脑加工受博弈对象的调节, 人际博弈时, 对损益域、公平与不公平提议的加工类似, 而人机博弈时, 损失域和不公平提议涉及更多的抑制加工和冲突解决, 获益域和公平提议则更富动机性意义, 证实公平加具有情境依赖性。

关键词 公平; 损益域; 最后通牒博弈; ERP; N1; N350; LPP

分类号 B842; B849:C91

1 前言

公平(fairness)是人类社会生活的重要追求。我国自古就有“不患寡而患不均”的思想。大量现场和实验室研究证实人们具有不公平厌恶倾向(inequity aversion), 当受到不公平对待时, 人们会表现出不满和怨恨, 并不惜牺牲个体的经济利益以惩罚对方进而达到公平(Fehr & Gächter, 2002; Henrich et al., 2006)。

公平加工的实验室研究常采用最后通牒博弈范式(Ultimatum Game) (Güth, Schmittberger, & Schwarze, 1982)。该博弈有两个角色: 提议者(proposer)和反应者(responder)。博弈的规则是博弈双方被赋予一笔钱, 首先由提议者提出分配的方案, 然后由反应者决定是否接受这一分配方案: 如果反应者接受, 则双方按照分配方案来分配这笔钱; 如果反应者拒绝, 则双方都无所获。根据传统经济学理论的理性人假设, 作为反应者会接受一切大于 0 的分配方案, 因为聊胜于无; 而提议者知道反应者

的这一倾向, 他所提议的分给对方的部分应无限接近于 0。但实证研究却发现, 提议方案大多是 5-5 分, 平均分给对方的部分在总额的 30%~40%之间; 少于 20%的分配常会被拒绝; 当总额提高, 不公平的提议也有较高收入时, 同样会遭到拒绝(Kahneman, Knetsch, & Thaler, 1986; Thaler, 1988; Weg & Smith, 1993)。尽管实验结果较为一致地表明人们具有公平偏好和不公平厌恶的倾向, 但公平偏好和不公平厌恶并非一成不变, 它们会受到各种因素的调节, 如博弈规则的呈现方式、博弈双方关系、博弈的背景信息等。

不公平厌恶现象与理性人假设不符, 引起了研究者对其神经机制的关注。Sanfey 等(2003)最早用 fMRI 技术探讨了不公平厌恶的神经相关, 发现不公平提议激活了与认知控制有关的背外侧前额叶(dorsal lateral prefrontal cortex, DLPFC)和与情绪加工有关的前脑岛(anterior insula), 以及负责冲突解决的前扣带回(anterior cingulate cortex, ACC); 前脑岛激活越强的被试越倾向于拒绝不公平的分配

收稿日期: 2011-06-27

* 国家自然科学基金项目(31100744)、浙江省自然科学基金资助项目(Y2110381)和杭州师范大学勤慎研究项目资助。

通讯作者: 吴燕, E-mail: lewuyan@126.com

方案,表明情绪对决策的影响(Sanfeiy, Rilling, Aronson, Nystrom, & Cohen, 2003)。相对于不公平提议,公平提议则激活了腹侧纹状体(ventral striatum)、杏仁核(amygdala)、腹内侧前额叶、眶额叶(orbitofrontal cortex, OFC)、黑质附近的中脑区等经典的奖赏加工区,表明公平分配的意义相当于奖赏能诱发被试正面的快乐反应,且公平的加工可能相对自动化和本能化,因为腹侧纹状体、杏仁核、腹内侧前额叶都与自动和本能的加工有关(Tabibnia & Lieberman 2007)。脑损伤病人研究和局部经颅磁刺激(TMS)研究证实了认知控制脑区和情绪调节脑区在公平决策中的重要作用(Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006; Koenigs & Tranel, 2007)。最近一项研究则发现不公平厌恶的神经反应会根据情境发生变化:当对方没有其它选择只能做出不公平提议时,拒绝不公平分配方案会激活内侧前额叶前部(anterior medial prefrontal cortex)和颞顶联合区(temporoparietal junction, TPJ);而当对方有其它公平选择却做出不公平提议时,拒绝不公平分配方案会激活脑岛和背内侧前额叶(dorsal medial prefrontal cortex)(Güroğlu, van den Bos, Rombouts, & Crone, 2010)。

已有脑成像研究揭示了加工公平偏好和不公平厌恶的相关脑区,ERP研究则能反映不公平厌恶加工的时间进程。如Polezzi等(2008)的研究分析了最后通牒博弈中反应者对不同提议额(分得总额的1/10、2/10、3/10、4/10、5/10)的脑电反应,发现在提议额呈现后的200~300 ms时间窗口内,3/10和1/10的方案诱发了更大的反馈负波(feedback-related negativity, FRN),在300~400 ms时间窗口内,3/10的方案诱发了更大的N350。对FRN和N350的源分析发现了左侧颞上回(left superior temporal gyrus)和左顶下小叶(left inferior parietal lobule)这两个经典的心理理论加工区。研究者认为,FRN反映了对不理想结果的评价,N350反映了认知冲突,而源分析结果则表明博弈过程中涉及对他人心理的揣测与加工(Polezzi et al., 2008)。然而,这个ERP实验采用了重复博弈的方式,即被试跟同一个对象连续进行200轮博弈,这种设计使得对每一轮提议的加工不独立。最近几项ERP研究采用重复单轮(repeated one-shot)的博弈范式,都发现相对于公平提议,不公平提议诱发了更大的FRN,且FRN幅值与个体对公平的关注感呈正相关(Boksem & De Cremer, 2010),也与被试对不公平提议的拒绝率相

关(Hewig et al., 2011),且FRN效应受分配总额大小的影响(van der Veen & Sahibdin, 2011),也受即时诱发情绪的调节,诱发正性情绪下的公平与不公平提议诱发的FRN差异更明显(Wang et al., 2011)。

本研究采用最后通牒范式,考察不公平厌恶的大脑加工如何受损益域(domain)及博弈对象的调节。已有行为研究发现,相对于获益域,损失域下对不公平提议的拒绝率更高(Zhou & Wu, 2011),损失情境下更加关注公平。但损失情境强化不公平厌恶的神经机制尚不清楚。因损失域涉及更多的抑制加工,预期损失域下的不公平提议会诱发更大的负成分。另外,为考察公平加工是否受博弈对象的影响,特加入人机博弈条件,以与人际博弈进行对比。已有研究曾发现,即使明知博弈对象是计算机,被试也会拒绝不公平的提议(Sanfeiy et al., 2003),表明公平加工的跨对象特性。但人机博弈是否与人际博弈具有类似的大脑活动模式,目前较少研究探讨。Rilling等(2004)的UG脑成像研究发现,人际博弈和人机博弈均激活了经典心理理论(theory of mind)区,如前扣带回和颞上沟后部(posterior superior temporal sulcus, posterior STS),人际博弈下这些区域的激活更强(Rilling, Sanfeiy, Aronson, Nystrom, & Cohen, 2004)。我们预期,与人际博弈相比,人机博弈下对提议公平性的加工将会减弱,损益域的差异也相对较小。

2 方法

2.1 被试

20名在校非心理系、非经济系本科生或研究生。男、女生各10人,平均年龄23.5岁,均身心健康,右利手,视力正常或矫正后正常,没有色盲、色弱等疾病。其中4名被试因脑电数据的伪迹较多而剔除,最后16名被试的数据进入统计分析。支付所有被试底金30元,博弈任务中的实际决策将决定每个人最终获得的报酬。

2.2 实验设计和刺激

实验采用2(损失域/获益域)×2(博弈对象:人/计算机)×2(提议公平水平:不公平(1/9、2/8)、公平(4/6、5/5))被试内设计,3/7为填充材料,共10种分配提议:获益域下1/9、2/8、3/7、4/6、5/5;损失域下-9/-1、-8/-2、-7/-3、-6/-4、-5/-5,斜线前数字表示分配给反应者的部分,斜线后数字表示提议者自己的部分。因变量为反应者对每类分配提议的拒绝率,以及看到每种分配方案时的脑

电成分。

被试均为反应者, 完成多轮最后通牒博弈。由计算机程序决定每轮博弈是总体赢钱(+¥10)还是赔钱(-¥10), 然后将提议者提出的分配方案呈现给反应者, 反应者决定是否接受。在总体赢钱情况下, 如果被试拒绝了方案, 则双方都得 0 元; 在总体赔钱的情况下, 如果拒绝了方案, 则双方都扣去 10 元。如果接受了该轮方案, 双方则按方案来分钱或承担后果。被试最后的报酬在参与实验应得的报酬的基础上根据其在博弈中的选择作相应额度的增减。

实验刺激为呈现在黑色背景下的白色汉字或数字。数字字体 Courier, 字号 32。分配提议和最终结果显示的字体宋体, 字号 32, 分两行呈现在屏幕中间。博弈对象的图片(人或计算机)为灰度图片, 89×102 像素, 其中人物为剪影, 以消除性别、面孔吸引力等因素的影响。

2.3 实验程序

实验开始前被试仔细阅读指导语。向被试强调分配提议是由不同的对家提出的, 但是由于无法在统一的时间地点同时进行博弈, 所以事先收集了他们的数据。由于每一轮是跟不同的对家进行博弈, 所以被试每一轮的选择不影响接下来的分配提议,

也不影响下一轮是一起赢钱还是赔钱。但是被试的选择决定了他自己最后的收入, 也决定了对家同学最后的收益。这种操纵是重复的单轮博弈(repeated one-shot game), 可以实现与同一个对家进行单轮博弈的同时, 满足每种提议条件下有多个重复试次的需求, 并强化博弈情境的真实性。被试正确理解博弈规则后开始实验。正式实验前进行 20 轮练习, 以帮助熟悉博弈规则和按键操作。

每一轮博弈开始屏幕中央呈现注视点, 持续 800 ms~1200 ms 之间不等。然后呈现该轮博弈的博弈对象, 是人或计算机, 图片呈现 1000 ms 后消失。间隔 500 ms 的黑屏后呈现对方的分配提议 1200 ms, 这段时间内被试不可做出按键反应。在呈现按键提示屏后, 被试在 2s 内按键反应(反应对应按键在被试间平衡)。一段 300 ms~700 ms 不等的间隔后呈现该轮博弈的结果 1500 ms, 若被试选择接受, 则结果与分配提议一样, 若被试选择拒绝, 则在赢 ¥10 条件下双方都得 0, 在赔 ¥10 条件下双方都-10。若被试在规定时间内未完成按键, 视为反应错误, 则该轮博弈被试将得 0 元或赔 10 元, 对家的金额不受影响。注视点出现表示下一轮博弈开始。脑电成分分析被试看到提议时刺激锁定的相关活动。单轮博弈流程以及重要控制如图 1 所示。

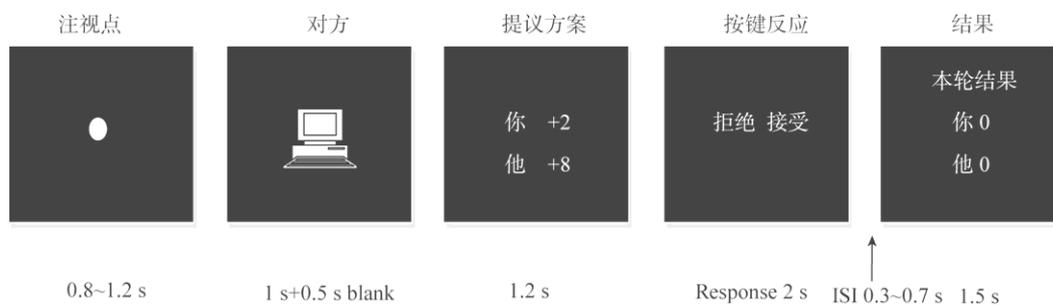


图 1 实验单轮博弈流程图

共进行 440 轮博弈, 分 10 个区组, 每个区组有 44 个试次。每类公平(包括 5/5、-5/-5、4/6、-6/-4)与不公平(包括 1/9、-9/-1、2/8、-8/-2)提议各重复 50 次, 中等公平(3/7、-7/-3)条件各重复 20 次, 对家为人和对家为计算机各半。分配提议以伪随机方式呈现。为避免呈现序列的影响, 损失和获益域连续出现不超过 4 次, 同一种分配提议连续出现不超过 3 次, 对家为人或对家为计算机的试次连续出现不超过 4 次。

2.4 ERP 记录与分析

实验仪器为 NeuroScan ERP 工作站。记录电极

固定于 64 导电极帽。根据国际 10-20 系统, 以位于左右眼眶外侧 1.5 cm 的电极记录水平眼电(HEOG), 位于左眼上下约 1 cm 的电极记录垂直眼电(VEOG)。实验中采用单耳乳突作为参考电极, 离线数据处理时以左右乳突的平均电位作为重参考。头皮与电极之间的阻抗小于 5 kΩ。信号经 AC 放大, 记录连续 EEG 和 EOG, 滤波带通为 0.05~70Hz, 采样频率为 500 Hz/导, 增益为 500, 离线式(off-line)叠加处理。

离线处理采用 NeuroScan 4.3 软件提供的算法进行眼电伪迹校正, 分析时程(epoch)选择提议呈

现阶段提议刺激呈现前 200 ms 到提议刺激呈现后 800 ms, 并以刺激前的均值校正基线。采用 30Hz 低通无相移滤波, 自动校正垂直眼电并排除其它伪迹, 波幅大于 $\pm 70\mu\text{V}$ 在叠加时被去除。

按分配提议的类型进行叠加平均, 取得每个被试看到每类提议时的脑电波幅。对总平均波幅的视觉检测发现各条件间的波幅潜伏期比较一致。根据总平均波幅的特征, 提取了 N1 (50~150 ms 时窗内的最小峰值)、P2 (200~400 ms 时窗内的最大峰值)、N350 (300~450 ms 时窗内的最小峰值)、LPP (late positive potential, 400~700 ms 时窗内的平均波幅) 的幅值, 进行 2 (损失域/获益域) \times 2 (博弈对象: 人/计算机) \times 2 [提议公平水平: 不公平(1/9、2/8)、公平(4/6、5/5)] 重复测量的方差分析, 电极位置二因素同时纳入分析: 5 (从前到后五排电极: F、FC、C、CP、P) \times 3 [电极位置: 左(3)、中(z)、右(4)]。统计结果非球形时采用 Greenhouse-Geisser 校正, 多重比较采用 Bonferroni 校正。

3 结果

剔除超时或错误反应的数据, 剔除率为总数据的 0.5%。运用 SPSS 15.0 进行数据统计与分析。

3.1 行为结果——拒绝率

对拒绝率进行 2 (损失域/获益域) \times 2 (博弈对象: 人/计算机) \times 2 (提议公平水平: 不公平(1/9、2/8)、公平(4/6、5/5)) 重复测量的方差分析。损益域的主效应显著, $F(1, 15) = 14.25, p < 0.005$, 损失域下的拒绝率(28%)高于获益域(19%); 提议公平水平的主效应显著, $F(1, 15) = 27.10, p < 0.001$, 不公平提议的拒绝率(45%)高于公平提议(1%); 博弈对象的主效应不显著, $F(1, 15) = 0.02, p > 0.1$, 且与

损益域和提议公平水平的交互作用均不显著; 损益域与提议公平水平之间有交互作用, $F(1, 15) = 9.72, p < 0.01$, 表现为公平提议下, 损益域间的拒绝率无差异, 不公平提议下, 损失域的拒绝率高于获益域(见图 2)。

3.2 脑电成分结果

N1

对 N1 幅值的方差分析发现, 损益域的主效应不显著, $F(1, 15) = 0.56, p > 0.1$; 提议公平水平的主效应不显著, $F(1, 15) = 0.27, p > 0.1$; 博弈对象的主效应显著, $F(1, 15) = 10.54, p < 0.01$, 对家为人($-3.02\mu\text{V}$)比对家为计算机($-2.56\mu\text{V}$)诱发了更负的 N1; N1 幅值从额叶区到顶叶区依次递减, $F(4, 60) = 11.24, p < 0.005$, F 排的幅值最大; 中部电极($-3.18\mu\text{V}$)的 N1 大于左边($-2.80\mu\text{V}$)和右边($-2.38\mu\text{V}$)电极, $F(2, 30) = 10.93, p < 0.001$ 。方差分析发现了博弈对象、损益域、提议公平水平三因素间的交互作用, $F(1, 15) = 4.60, p < 0.05$ 。简单效应分析发现, 在获益域下, 博弈对象的主效应显著, $F(1, 15) = 5.64, p < 0.05$, 对家为人($-3.07\mu\text{V}$)比对家为计算机($-2.37\mu\text{V}$)诱发了更负的 N1, 提议公平水平无主效应, 二者的交互作用也不显著; 而在损失域下, 博弈对象和提议公平水平的主效应不显著, 二者的交互作用边缘显著, $F(1, 15) = 3.25, p = 0.09$, 对家为计算机时公平提议与不公平提议的差异($-2.42\mu\text{V}$ vs. $-3.05\mu\text{V}$)大于对家为人($-3.05\mu\text{V}$ vs. $-2.89\mu\text{V}$)。

P2

对 P2 幅值的方差分析发现, 损益域的主效应不显著, $F(1, 15) = 0.70, p > 0.1$; 提议公平水平的主效应不显著, $F(1, 15) = 1.47, p > 0.1$; 对家身份的

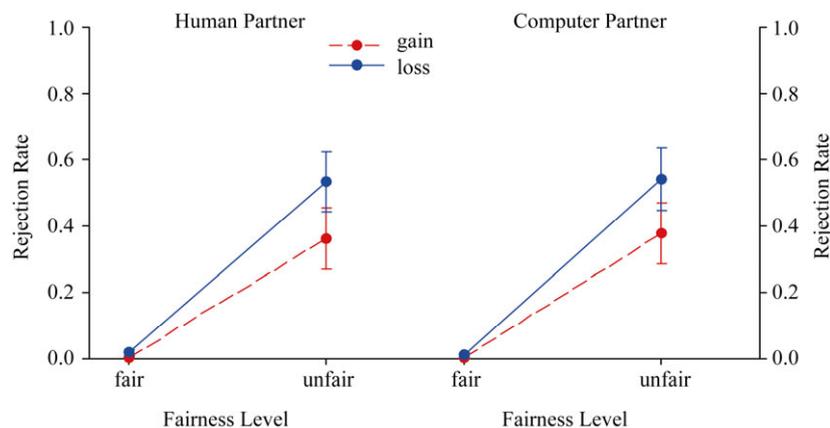


图 2 博弈对象为人和为计算机时对损益情境下公平与不公平提议的拒绝率

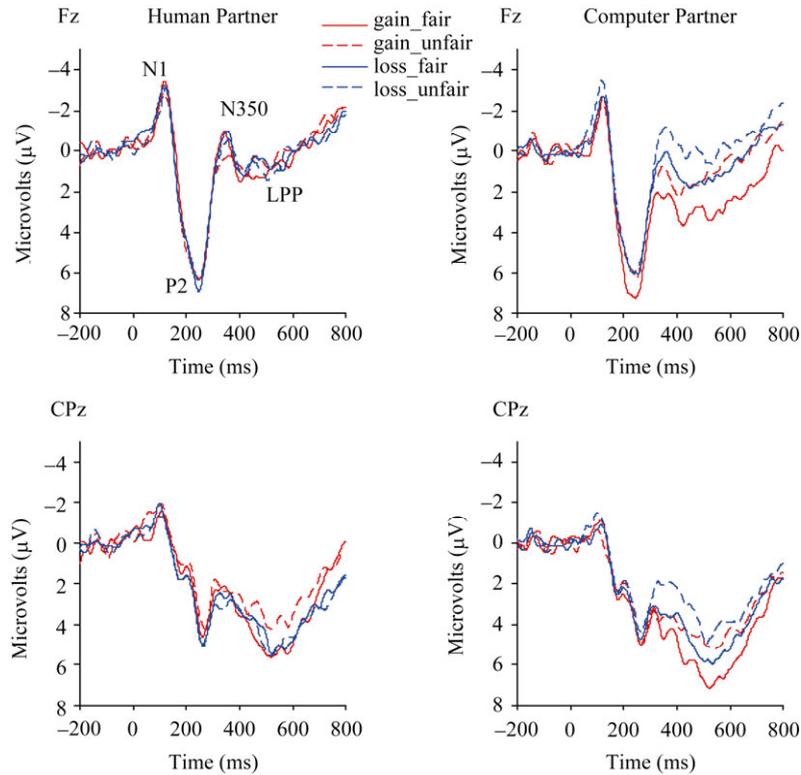


图3 博弈对象为人或计算机时, 损益域下公平与不公平提议诱发的总平均波幅。0 ms 表示提议开始呈现的时间。

主效应不显著, $F(1, 15) = 0.99, p > 0.1$; P2 幅值从额叶区到顶叶区依次递减, $F(4, 60) = 14.14, p < 0.001$, F 排的幅值最大; 左、中、右边电极的差异不显著, $F(2, 30) = 1.36, p > 0.1$ 。方差分析发现了博弈对象与损益域的交互作用, $F(1, 15) = 9.26, p < 0.01$, 简单效应分析发现, 对家为计算机时, 获益域($6.92 \mu\text{V}$)比损失域($6.30 \mu\text{V}$)诱发了更大的 P2, $F(1, 15) = 6.81, p < 0.05$, 对家为人时, 损益域无差异($6.30 \mu\text{V}$ vs. $6.49 \mu\text{V}$, $F(1, 15) = 0.38, p > 0.1$)。

N350

对 N350 幅值的方差分析发现, 损益域的主效应不显著, $F(1, 15) = 2.08, p > 0.1$; 提议公平水平的主效应不显著, $F(1, 15) = 2.81, p > 0.1$; 博弈对象的主效应边缘显著, $F(1, 15) = 4.41, p = 0.053$, 对家为人($-0.16 \mu\text{V}$)比对家为计算机($0.5 \mu\text{V}$)诱发了略微更负的 N350; N350 幅值从额叶区到顶叶区依次递减, $F(4, 60) = 7.36, p < 0.05$, F 排的幅值最大; 中部电极($-0.43 \mu\text{V}$)的 N350 大于左边($0.28 \mu\text{V}$)和右边($0.67 \mu\text{V}$)电极, $F(2, 30) = 6.85, p < 0.01$ 。方差分析发现, 博弈对象与损益域的交互作用显著, $F(1, 15) = 6.22, p < 0.05$, 简单效应分析发现, 对家为计算机时, 损失域($0.05 \mu\text{V}$)比获益域($0.95 \mu\text{V}$)诱发了更负的 N350, $F(1, 15) = 9.02, p < 0.01$, 对家为人时,

损益域无差异($-0.02 \mu\text{V}$ vs. $-0.30 \mu\text{V}$, $F(1, 15) = 0.67, p > 0.1$)。博弈对象与提议公平水平的交互作用显著, $F(1, 15) = 5.99, p < 0.05$, 简单效应分析发现, 对家为计算机时, 不公平提议($0.05 \mu\text{V}$)比公平提议($0.96 \mu\text{V}$)诱发了更负的 N350, $F(1, 15) = 8.36, p < 0.05$, 对家为人时, 公平提议与不公平提议无差异($-0.20 \mu\text{V}$ vs. $-0.12 \mu\text{V}$, $F(1, 15) = 0.07, p > 0.1$)。

LPP

对 LPP 幅值的方差分析发现, 损益域的主效应不显著, $F(1, 15) = 2.48, p > 0.1$; 提议公平水平的主效应显著, $F(1, 15) = 5.10, p < 0.05$, 公平提议($2.96 \mu\text{V}$)比不公平提议($2.21 \mu\text{V}$)诱发了更大的 LPP; 对家身份的主效应不显著, $F(1, 15) = 2.98, p > 0.1$; LPP 幅值从额叶区到顶叶区依次增加, $F(4, 60) = 13.37, p < 0.005$, CP 排的幅值最大; 左、中、右边电极的幅值无差异, $F(2, 30) = 0.02, p > 0.1$ 。方差分析发现, 博弈对象与损益域的交互作用显著, $F(1, 15) = 6.72, p < 0.05$, 简单效应分析发现, 对家为计算机时, 获益域($3.40 \mu\text{V}$)比损失域($2.37 \mu\text{V}$)诱发了更大的 LPP, $F(1, 15) = 12.55, p < 0.01$, 对家为人时, 损益域无差异($2.23 \mu\text{V}$ vs. $2.43 \mu\text{V}$, $F(1, 15) = 0.25, p > 0.1$)。博弈对象与提议公平水平的交互作用显著, $F(1, 15) = 5.27, p < 0.05$, 简单效应分析发现, 对家

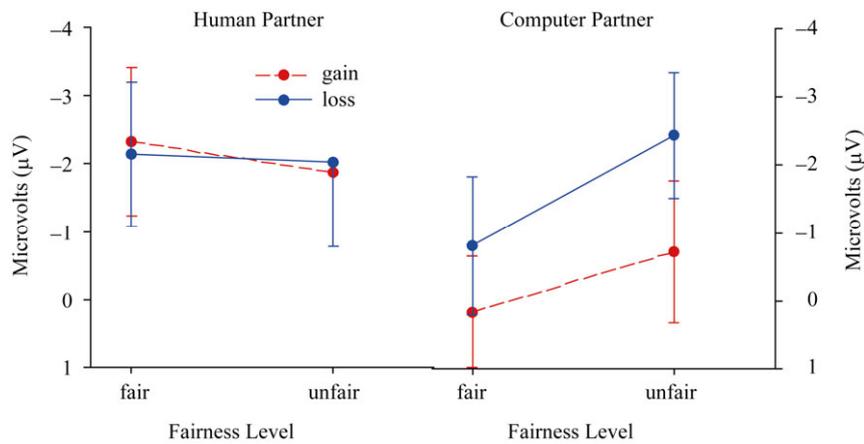


图4 Fz 电极处博弈对象为人或计算机时, 损益域下公平与不公平提议诱发的 N350 波幅。

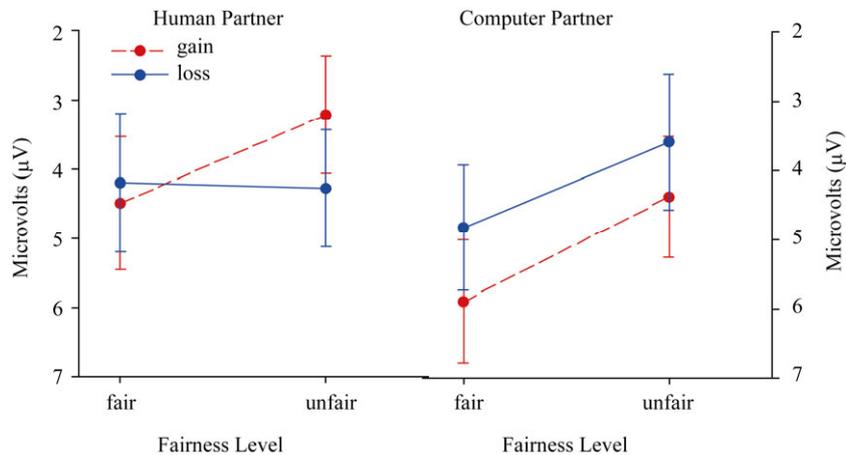


图5 CPz 电极处博弈对象为人或计算机时损益域下公平与不公平提议诱发的 LPP 波幅。

为计算机时, 公平提议(3.47 μV)比不公平提议(2.30 μV)诱发了更大的 LPP, $F(1, 15) = 9.25, p < 0.01$, 对家为人时, 公平提议与不公平提议无差异(2.51 μV vs. 2.16 μV , $F(1, 15) = 0.87, p > 0.1$)。

4 讨论

已有对公平加工的 ERP 研究没有探讨损益域和博弈对象这两个情境因素的影响。本研究发现, 损失域下被试对不公平提议的拒绝率更高, 这与已有行为研究的结果一致(Zhou & Wu, 2011)。尽管博弈对象不影响被试对提议的拒绝行为, 但在脑电活动上有所不同: 对家为人比对家为计算机诱发了更负的 N1, 但只在获益域下; 人机博弈时, 获益域比损失域诱发了更大的 P2 和 LPP, 损失域比获益域、不公平提议比公平提议诱发了更大的 N350, 而公平提议比不公平提议诱发了更大的 LPP; 人际博弈

时, 这些差异均不显著。这些结果表明大脑对分配提议的加工受博弈对象的调节, 人际博弈时, 对损益域、公平与不公平提议的加工类似, 而人机博弈时, 损失域和 unfair 提议涉及更多的抑制加工和冲突解决, 获益域和公平提议则更富动机性意义。

4.1 获益域下人际博弈捕获更多的早期注意

前部 N1 是视觉刺激呈现后约 100 ms 左右出现在额部中央区的负波(Luck, 1995)。它不同于也是由视觉刺激诱发的后部 N1 (或 N170), 后者在颞枕电极处达到最大(Hillyard & Anllo-Vento, 1998)。已有研究表明, 前部 N1 成分受注意的影响, 被注意的刺激比非注意的刺激诱发了更大的 N1 (Clark & Hillyard, 1996)。后部 N1 被认为反映了视觉的辨别加工, 而前部 N1 则反映了与反应准备有关的准备性运动加工(Voge & Luck, 2000)。一项人际决策研究发现, 二人利益冲突比利益一致条件下的视觉刺

激诱发了更大的前部 N1 (Boudreau, McCubbins, & Coulson, 2009)。

本研究发现, 人际博弈比人机博弈诱发了更负的前部 N1, 且差异仅存在获益域下。获益是种正性事件, 可能会起到诱发正性情绪的作用(我们的另一个实验让被试评价获益和损失域下公平与不公平提议诱发的情绪值, 发现对获益域提议的评价比损失域显著更偏正性)。已有研究曾表明, 正性情绪可提高个体在任务切换中的认知灵活性(Dreisbach & Goschke, 2004)。而王益文等(2011)的研究也发现, 公平提议与不公平提议诱发的 FRN 差异只存在正性情绪诱发下(Wang et al., 2011)。因而, 在提议呈现的早期阶段, 正性情境下可以发现博弈对象的差异。而损失域作为一种负性事件, 会诱发负性偏向(negative bias) (Rozin & Royzman, 2001), 占据较多的认知资源而无法对博弈对象作出区分。当正性情境下可对博弈对象作出区分时, 人际博弈比人机博弈诱发了更负的前部 N1, 可能是因为在这一早期视觉注意阶段, “他”比“它”捕获了更多的注意资源, 得到了更多的编码表征, 即人所提出的提议比计算机所提出的提议引发了被试更多的外源性注意(自下而上)和更大的内源性关注(自上而下)。表明早在 100ms 左右, 大脑就可对博弈对象做出区分, 比起人机交互, 人际交互更受关注。

4.2 初级评价: 人机博弈时获益域的动机性含义更强

提议呈现后 250 ms 左右, 博弈对象与损益域交互影响着波幅的峰值。人所提出的提议在损益域下诱发的 P2 波幅相当, 而计算机提出的提议在获益域比损失域诱发了更大的 P2 波幅。作为初级评价成分, 有研究认为 P2 反映了对某些属性的高级知觉加工(Kranczioch, Debner, & Engel, 2003)。在目标检测范式下, 以多种视觉属性来定义目标与非目标, 发现目标刺激比非目标刺激诱发了更大的 P2 (Hillyard & Muentz, 1984), 它可能反映了对多维度特征的检测加工(Luck & Hillyard, 1994)。由于 P2 受内隐和外显目标检测范式的调节, 因而有研究者认为 P2 主要对知觉信息的任务相关性敏感, 即反映了对动机性信息和知觉信息的整合加工(Potts, 2004)。在对奖赏加工的研究中发现, 结果比预期更好的反馈刺激诱发了更大的 P2 成分(Potts, Martin, Burton, & Montague, 2006)。Boudreau 等(2009)的决策研究发现, 二人利益一致比利益冲突条件下的刺

激诱发了更大的 P2, 研究者认为利益一致比利益冲突条件具有更强的动机和知觉突显性(Boudreau et al., 2009)。

博弈对象与损益域在 P2 波幅上的交互作用, 再次表明博弈对象在决策加工中的影响。我们认为, P2 反映了以目标为导向的自上而下的加工, 对与目标一致的动机性信息敏感。在本研究中, 当对家为计算机时, 获益域比损失域诱发了更大的 P2, 可能是因为与计算机博弈时, 相对于损失域, 获益域这一信息与个体趋利避害的天性一致, 并更符合经济利益最大化这一具体目标, 因此具有更强的动机性含义和凸显性, 从而能够被及时提取, 并被 P2 表征。而当对家为人时, 个体除了经济利益最大化这一目标外, 还有追求社会公平这一动机(Knoch et al., 2006)。在双重动机的影响下, 一起分担损失和一起获利, 对个体而言, 具有同等重要的影响, 因而在 P2 波幅上无差异。这些结果提示, 人与人交往时, 损失分摊和利益分享都具有重大的意义, 而人机博弈时, 赢钱的意义大于赔钱。

4.3 人机博弈时损失域、不公平提议涉及更多的抑制加工

大约 350 ms 左右, 提议呈现诱发了一个负成分, 在额叶中央达到最大。已有 UG ERP 研究多将这个成分视为 FRN/MFN (Boksem & De Cremer, 2010; Hewig et al., 2011; van der Veen & Sahibdin, 2011; Wang et al., 2011), 发现不公平提议诱发了更大的 FRN/MFN, 反映了对公平这一社会规范的违反或对不公平提议诱发的负性情绪的评价。而最早发表的一篇 UG ERP 研究除报告 FRN 外, 还报告了 N350 (Polezzi et al., 2008)。鉴于本实验结果所得到的波形, 我们认为这个波峰出现在 350 ms 左右、在额叶电极幅值达到最大的负成分从潜伏期、幅值和头皮分布上均更接近于 N350 (de Pascalis, Strelau, & Zawadzki, 1999)。

N350 多见于视认知研究, 反映了继初级自发快速的自下而上的视加工之后的知觉编码过程, 是对客体信息的进一步整合加工, 包括客体模型的选择和匹配(Schendan & Lucia, 2009)。在一个客体识别任务中, 被试从寻常和不寻常的视角来观察物体, 发现与寻常视角相比, 从不寻常的视角识别客体具有更多的反应时和更大的 N350 幅值, 即 N350 可能与加工难度有关(Schendan & Kutas, 2003)。但在一个清醒与睡眠的研究中, 被试在阅读、躺在床上、轻度睡眠、慢波睡眠和快速眼动睡眠时呈现声音脉

冲,同时记录其脑电活动。发现各个条件下均出现了 N350,且在轻度睡眠时达到最大值。研究者认为 N350 反映了与减少信息加工有关的抑制过程(Kállai, Harsh, & Voss, 2003)。关于 N350 的加工难度假设与抑制信息假设并不冲突,因为在复杂的加工过程中,加工难度越大,可能涉及对信息的抑制越多。例如 Polezzi 等(2008)的 UG ERP 研究发现,中等公平水平的提议(3/7)比公平提议(5/5)和不公平提议(1/9)诱发了更大的 N350,这可能是因为中等公平水平的提议构成了两难境地,诱发了更强的决策冲突——利动机驱动接受提议与追求公平动机驱动的拒绝提议之间的冲突,需要更多的抑制加工,同时也增强了加工难度(Polezzi et al., 2008)。

本研究发现了博弈对象与损益域、博弈对象与提议公平水平在 N350 上的交互作用:对方为计算机时,损失域、不公平提议诱发了更大的 N350,而对方为人时,损益域、公平/不公平提议诱发的 N350 无差异。已有研究曾表明,相对于损益域和公平提议,损失域和不公平提议诱发了更强的负性情绪和决策冲突(Zhou & Wu, 2011),因而需要更多的抑制加工。作为抑制加工的指标, N350 效应仅表现于人机博弈而消失于人际博弈,这可能与人际博弈加工的特殊性有关。人际博弈时,涉及更多心理理论的推理加工,比如对对方意图和目标的揣测。相比之下,对损益域以及提议公平或不公平等信息的加工相对弱化。如 Rilling 等(2004)的 UG 脑成像研究发现,人际博弈和人机博弈均激活了经典心理理论(theory of mind)区,如前扣带回和颞上沟后部(posterior superior temporal sulcus, posterior STS),但人际博弈下这些区域的激活显著更强(Rilling et al., 2004)。Polezzi 等(2008)的 ERP 研究的源分析结果也发现,中等公平水平的提议比公平提议诱发了更强的颞上沟区域的活动(Polezzi et al., 2008)。这两个研究证实,相对于人机博弈,人际博弈时具有更多的心理理论加工的参与。当认知资源被心理理论加工占据时,分配到损失域和不公平提议上的抑制加工资源减少,导致 N350 上差异不显著。而人机博弈时,认知资源较多地分配到加工提议信息本身,损失、不公平等信息被有效提取,其引发的负性情绪和决策冲突能得到表征和抑制,从而在 N350 上表现出差异。

4.4 再评价:人机博弈时损益域、公平提议的意义凸显

各类提议在大约 400 ms 以后诱发了一个正成

分 LPP,该成分在中央顶叶区(CPz)达到最大。晚期正成分一般与评价加工有关。例如在情绪图片加工中,当负性刺激呈现在正性刺激背景下,或反之,都会诱发 LPP(Cacioppo, Crites, & Gardner, 1996)。正性刺激和负性刺激诱发的 LPP 波幅相当,即 LPP 与情绪效价无关,但与刺激的唤醒程度有关,刺激越具有动机性意义,其诱发的 LPP 越大(Schupp et al., 2000)。LPP 的非效价特异性(valence-specific)表明, LPP 可能反映的并非评价加工本身,而是对有机性意义的刺激检测,或者反映了对评价系统的输出结果(初级评价)的再评价加工过程(Cunningham, Espinet, DeYoung, & Zelazo, 2005)。

在这个晚期再评价阶段,再次发现了博弈对象与损益域、博弈对象与提议公平水平的交互作用:对方为计算机时,损益域、公平提议诱发了更大的 LPP,而对方为人时,损益域、公平/不公平提议诱发的 LPP 无差异。这个结果可能是因为,与计算机博弈时,由于趋利避害的天性和偏好公平的倾向性,相对于损失和不公平提议而言,获益和公平提议对个体的动机性含义更强,它们与自我利益最大化的目标一致,并直接增长经济利益,因此这些正性信息起到正强化的作用,在晚期再评价阶段得到凸显;而人际博弈时,由于心理理论加工的存在,关注于对对方意图的揣测,从而使得对其它信息的再评价加工减弱,因而无论获益还是损失,公平提议还是不公平提议,在 LPP 波幅上无差异。

最后,值得一提的是,尽管博弈对象影响对公平加工的脑电反应,却不影响行为决策。这一相分离的结果可能是因为脑电指标比行为指标更敏感,也可能是因为脑电成分只能反映信息加工的某一阶段,而行为结果则是整个信息加工过程的最终输出结果,二者无法一一对应。人机博弈与人际博弈的决策结果无异,其原因可能有二:(1)被试采取了简单而稳定的决策策略,比如接受公平提议拒绝不公平提议。一旦形成这种稳定的策略和倾向,决策就不易受到博弈对象的影响了。对每个被试每种条件下前 10 个试次和最后 10 个试次的决策行为的比较分析,发现二者几乎完全一致,表明在经过练习程序后,被试即形成了稳定的策略和决策倾向。(2)计算机对象拟人化。如 Rilling 等(2004)提出,被试对计算机对象与人一视同仁(imbue),即人机博弈激发了类似于人际博弈的情感和行为(Rilling et al., 2004)。目前将人际博弈与人机博弈直接进行对比的研究比较少,尚需更多的研究来探讨二者的异同。

5 结论

公平加工具有情境依赖性: 对提议的拒绝行为受损益情境的调节, 损失情境下更加注重分配公平; 尽管拒绝率不受博弈对象的影响, 但脑电活动却反映出博弈对象的特异性; 人机博弈时, 损失域和不公平提议涉及更多的抑制加工, 获益域和公平提议则具有更强的动机性含义; 人际博弈时, 损益域和公平与不公平提议吸引了同等水平的评价加工。

参 考 文 献

- Boksem, M. A. S., & De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, 5(1), 118–128.
- 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.
- Cacioppo, J. T., Crites, S. L., Jr., & Gardner, W. L. (1996). Attitudes to the right: Evaluative processing is associated with lateralized late positive event-related brain potentials. *Personality and Social Psychology Bulletin*, 22(12), 1205–1219.
- Clark, V. P., & Hillyard, S. A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience*, 8, 387–402.
- Cunningham, W. A., Espinet, S. D., DeYoung, C. G., & Zelazo, P. D. (2005). Attitudes to the right and left: Frontal ERP asymmetries associated with stimulus valence and processing goals. *NeuroImage*, 28(4), 827–834.
- de Pascalis, V., Strelau, J., & Zawadzki, B. (1999). The effect of temperamental traits on event-related potentials, heart rate and reaction time. *Personality and Individual Differences*, 26, 441–465.
- Dreisbach, G., & Goschke, T. (2004). How positive affect modulates cognitive control: Reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 343–353.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.
- Güroğlu, B., van den Bos, W., Rombouts, S. A. R. B., & Crone, E. A. (2010). Unfair? It depends: Neural correlates of fairness in social context. *Social Cognitive and Affective Neuroscience*, 5(4), 414–423.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4), 367–388.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., et al. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–1770.
- Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., et al. (2011). Why humans deviate from rational choice. *Psychophysiology*, 48(4), 507–514.
- Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences USA*, 95, 781–787.
- Hillyard, S. A., & Münte, T. (1984). Selective attention to color and location: an analysis with event-related brain potentials. *Perception & Psychophysics*, 36, 185–198.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *The Journal of Business*, 59(4), S285–S300.
- Kállai, I., Harsh, J., & Voss, U. (2003). Attention to external stimuli during wakefulness and sleep: Evoked 40-Hz response and N350. *Psychophysiology*, 40(6), 955–966.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, 314(5800), 829–832.
- Koenigs, M., & Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the ultimatum game. *Journal of Neuroscience*, 27(4), 951–956.
- Kranczioch, C., Debner, S., & Engel, A. K. (2003). Event-related potential of the attentional blink phenomenon. *Cognitive Brain Research*, 17, 177–187.
- Luck, S. J. (1995). Multiple mechanisms of visual-spatial attention: recent evidence from human electrophysiology. *Behavioral Brain Research*, 71, 113–123.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3), 291–308.
- Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G., et al. (2008). Mentalizing in economic decision-making. *Behavioural Brain Research*, 190(2), 218–223.
- Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain & Cognition*, 56, 5–13.
- Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: The medial frontal cortex and the allocation of processing resources. *Journal of Cognitive Neuroscience*, 18, 1112–1119.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, 22(4), 1694–1703.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296–320.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626), 1755–1758.
- Schendan, H. E., & Kutas, M. (2003). Time course of processes and representations supporting visual object identification and memory. *Journal of Cognitive Neuroscience*, 15, 111–135.
- Schendan, H. E., & Lucia, L. C. (2009). Visual object cognition precedes but also temporally overlaps mental rotation. *Brain Research*, 1294, 91–105.
- Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, 37(2), 257–261.
- Tabibnia, G., & Lieberman, M. D. (2007). Fairness and cooperation are rewarding: Evidence from social cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1118(1), 90–101.
- Thaler, R. H. (1988). Anomalies: The ultimatum game. *The Journal of Economic Perspectives*, 2(4), 195–206.

- van der Veen, F., & Sahibdin, P. P. (2011). Dissociation between medial frontal negativity and cardiac responses in the ultimatum game: Effects of offer size and fairness. *Cognitive, Affective, & Behavioral Neuroscience, 11*, 516–525.
- Voge, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology, 37*, 190–203.
- Wang, Y. W., Wang, Y., Lin, C. D., Chen, X. Y., Yuan, B., & Shen, D. L. (2011). Modulation of conscientiousness on medial frontal negativity in negative emotions: an ERP study on ultimatum game. *Science China: Life Science, 41*(4), 320–331.
- Weg, E., & Smith, V. (1993). On the failure to induce meager offers in ultimatum game. *Journal of Economic Psychology, 14*(1), 17–32.
- Zhou, X. L., & Wu, Y. (2011). Sharing losses and sharing gains: Increased demand for fairness under adversity. *Journal of Experimental Social Psychology, 47*(3), 582–588.

The Context-Dependency of Fairness Processing: Evidence from ERP Study

WU Yan¹; ZHOU Xiao-Lin²

(¹ Department of Psychology, School of Educational Science, Hangzhou Normal University, Hangzhou 310026, China)

(² Department of Psychology, Peking University, Beijing 100871, China)

Abstract

In human society, sharing losses is at least as common as sharing gains. Although the psychological and neural processes underlying the latter have been investigated in depth, those related to the former are not clear. Our recent study demonstrates an increased demand for fairness under adversity (e.g. loss sharing). Here we investigated how our brain encodes unfairness in the loss and gain domains using event-related potentials (ERP) technique. We adopted the Ultimatum Game (UG) to probe the processes related to fairness consideration in either gain or loss domain. In UG, two players, the proposer and the responder, bargain on how to divide a certain amount of money endowed by the experimenter. The proposer suggests a division policy, on which the responder evaluates and decides whether to accept. Upon acceptance, the money is divided as suggested; while rejection results in both players going empty-handed. Participants, as responders, were required to decide whether to accept an offer that was either fair (equal or nearly equal division) or unfair in both gain and loss domain. Offers were either made by the human partner or by the computer partner. Behavioral results replicated our previous findings that the rejection rate of unfair offers was higher in the loss than in the gain domain. ERP results revealed that the N1 amplitude was more pronounced for human partners compared with computer partners, however, this effect was only observed in the gain domain. When interacting with computer partners, unfair offers and offers in the loss domain were associated with larger N350 compared with fair offers and offers in the gain domain, and offers in the gain domain elicited larger P2 than offers in the loss domain, whereas fair offers and offers in the gain domain were associated with larger LPP than unfair offers and offers in the loss domain. In addition, these differences in ERP responses were diminished when the interacting partners were humans. These findings suggest that fairness processing is modulated by the property of the partner and gain-loss domain. In human-computer interaction, unfair offers and offers in the loss domain elicit more inhibition and conflict resolving process, while fair offers and offers in the gain domain are more motivationally significant to human. The present findings support the view that fairness processing is context-dependent, in which factors like gain-loss domain and the property of the partner play a role.

Key words fairness; gain-loss domain; Ultimatum Game; ERP; N1; N350; LPP