ELECTROPHYSIOLOGICAL MARKERS OF FAIRNESS AND SELFISHNESS REVEALED BY A COMBINATION OF DICTATOR AND ULTIMATUM GAMES

A bioRχiv PREPRINT

Ali M. Miraghaie *
Neuroheuristic Research Group
Faculty of Business and Economics
University of Lausanne, CH 1015 Lausanne, Switzerland
& Faculty of Psychology
Shahid Beheshti University
Deneshjou Bouleverd, Evin Tehran, Iran

Hamidreza Pouretemad †
Dept. of Clinical & Health Psychology
Shahid Beheshti University
Tehran, Iran

Alessandro E. P. Villa ‡
Neuroheuristic Research Group
Faculty of Business and Economics
University of Lausanne, CH 1015 Lausanne, Switzerland

Mohammad A. Mazaheri
Dept. of Clinical & Health Psychology
Shahid Beheshti University
Tehran, Iran

Reza Khosrowabadi
Institute for Cognitive and Brain Science
Shahid Beheshti University
Tehran, Iran

Alessandra Lintas
Neuroheuristic Research Group & LABEX §
Faculty of Business and Economics
University of Lausanne, CH 1015 Lausanne, Switzerland

February 3, 2022

ABSTRACT

Event Related Potentials (ERPs) were recorded from 39 participants who played the role of Allocators in a Dictator Game (DG) and Responders in an Ultimatum Game (UG). Most participants expressed very low levels of altruistic decision making, and two homogeneous groups could be identified, one formed by fair ($N = 10$) individuals and another by selfish ($N = 8$) individuals. At fronto-central cortical sites, the ERP early negativity (N1) was reduced in selfish participants with a latency about 10 ms earlier than in fair participants. In fair DG players, the features of the subsequent positive wave P2 suggested that more cognitive resources were required when they allocated the least gains to the other party. P2 latency and amplitude in the selfish group supported the hypothesis that these participants tended to maximize their profit, as expected by a rational *Homo economicus*. During UG, we observed that a medial frontal negativity (MFN) occurred earlier and with greater amplitude when selfish participants rejected less favorable endowment shares. In this case, all players received zero payoffs, which showed that MFN in selfish participants was associated with a spiteful punishment. At posterior-parietal sites we found that the greater the selfishness, the greater the amplitude of the late positive component (LPC). Our results bring new evidence to the existence of specific somatic markers associated with the activation of distinct cerebral circuits by the evaluation of fair and unfair proposals in participants characterized by different expressions of perceived fairness, thus suggesting that particular brain dynamics could be associated with moral decisions.

*Email: sam.miraghaie@gmail.com
†Email: pouretemad.h@gmail.com
‡Email: alessandro.villa@unil.ch
§Email: alessandra.lintas@unil.ch

Keywords Cooperation, EEG, N1, P2, MFN, Late positive potential, Spiteful punishment, Costly punishment

1 Introduction

The maximization of individual's own gain is considered to represent the main behavioral drive during an economic transaction. However, the balance of individuals' self-interest and acceptance of others' benefit is a common observation based on culturally shaped moral phenomena such as fairness and altruism (Rachlin, 2002; Fehr and Rockenbach, 2004; Altman, 2005; Henrich et al., 2010). Unrelated individuals try to ensure fairness to others in daily life, and hope to get fair treatment and cooperation in return (Rand and Nowak, 2013; Kocher et al., 2012). Strong reciprocity model, as one of the strategies for maintaining and urging people to commit to the norms, has many practical and effective use in the light of cooperation (Gintis et al., 2003; Stallen and Sanfey, 2013) and people act differently in diverse societies to apply it (Herrmann et al., 2008). Cooperation induced by the sacrifice of one's own self-interest in punishing violations of the social norm defines the so-called altruistic punishment (Fehr and Gächter, 2002; Fowler, 2005; Du and Chang, 2015; Balafoutas et al., 2016). A different kind of punishment, so-called spiteful punishment, appears when an individual spends personal resources to punish one or several other individuals who behave against his presumed interest within the context of collective norms and rules (Jensen, 2010; Marlowe et al., 2011; Brañas-Garza et al., 2014; Yamagishi et al., 2017).

The Ultimatum Game (UG) has been one of the most prolific experimental tasks aimed at investigating the nature of human fairness over the last decades (Güth et al., 1982). A typical UG involves two players, the Proposer, who offers how to share an endowment in two parts, and the Responder, who can either accept the offer to share it accordingly or to reject it with both players receiving a zero payoff. Empirical studies have demonstrated that Proposers typically offered approximately 40% (fair offers) of the total amount at stake, with Responders being more likely to reject 20% or less of positive, albeit low (unfair) offers (Camerer, 2003; Sanfey et al., 2003). Rejection of an unfair offer at UG might be triggered by cooperative and popular motivations or it can be considered to be an expression of a costly punishment aimed at enforcing the social norm of fairness (Fehr and Schmidt, 1999; Henrich et al., 2006; Fehr and Fischbacher, 2004; Hewig et al., 2011). This argument has been used to support the strong reciprocity model of the evolution of cooperation (Bear and Rand, 2016; Righi and Takács, 2018; Chen et al., 2019), but competitive and spiteful motivation associated with less prosocial attitude have also been suggested (Kirchsteiger, 1994; Brañas-Garza et al., 2014). The Dictator Game (DG) task (Kahneman et al., 1986; Artinger et al., 2010) is basically a single player game because the Proposer (called the Allocator in DG) determines at his own will whether to send a fraction (ranging from nothing to all) of the initial endowment to the Responder (called the Recipient in DG). The Recipient plays a passive role and has no influence over the outcome of the game. It has been observed that the vast majority of Allocators give something to the Recipients, usually about 20% of the amount at stake, despite the fact that Allocators have no reason to give up some of the money (Artinger et al., 2010; Camerer, 2003).

Contactless monetary transactions are rapidly increasing in all socio-cultural environments and in the absence of effective face-to-face negotiations it seems crucial to gain a better understanding of decision-making processes. The observation that individuals characterized by well-defined impairments in decision making, as they tend to take decisions not advantageous for their personal lives and for their social status, may also be characterized by biophysical measurements associated with brain activity led to define the Somatic Markers Hypothesis (Bechara et al., 1994; Bechara and Damasio, 2005; Rilling and Sanfey, 2011; Villa et al., 2012). The observation of such somatic markers is important to investigate the neural basis of decision making processes. In UG and DG, the key question is searching of where and how a "decision" is taken in a recursive way. Functional magnetic resonance imaging (fMRI) has been mainly used to investigate the localization of brain activity associated with the violations of social norms and expectations in decision-making tasks (Sanfey et al., 2003; Spicer et al., 2007). The search of 'how' requires a fine grain temporal resolution (in the order of 1 ms time scale), which is a feature of electroencephalography (EEG) by means of recording brain activity in a non invasive way with external electrodes placed over many standard locations of the scalp (Boksem and De Cremer, 2010; Wu et al., 2011a; Luo et al., 2014; Yu et al., 2015; Mesrobian et al., 2015; Peterburs et al., 2017; Cui et al., 2019).

Transient electric potentials associated in time with a sensory or mental occurrence are termed event-related potentials (ERPs) (Picton et al., 2000). A fundamental prediction of this model is that multiple occurrences of the same event evoke a similar pattern of activity in the same cerebral circuits. Hence, the usual analysis of ERPs requires to average the activity triggered by multiple repetitions of the same event. Among the ERP wave components observed in cognitive studies, we introduce those that have been associated to experimental protocols of interest with respect to our study. The most common triggering events are the stimulus onset corresponding, in UG, to the time of the decision (acceptance/rejection) taken by the Responder following the presentation of the offer by the Proposer and, in DG, to the time of the decision (which endowment share offered to the Recipient) taken by the Allocator. In the chronological order, we introduce N1 (peaking near 130 ms after stimulus onset), P2 (or P200, peaking at 220 ms), medial-frontal

negativity (MFN, peaking at 310 ms), and a late positive component (LPC, extending between 400 and 650 ms). In several experimental paradigms, ERPs can be stimulus-locked at the time of presentation of the feedback information on the decision made. This is not the case in the current study and ERPs triggered by the onset of the feedback phase are not further discussed.

In a UG protocol, the amplitude of frontal N1 was modulated by the facial attractiveness of the player (Ma and Hu, 2015; Weiss et al., 2020) and following observed nonverbal social interactions (friendly vs. nonfriendly) with Proposers (Moore et al., 2021). Parietal N1 amplitude was reported in association with the process of emotional stimuli regardless of trustworthiness of the player (Mei et al., 2020). In UG, it was observed that an increase in P2 is larger in response to high-value over low-value offers (Weiss et al., 2020), in participants informed by the opponent's low social status (Hu et al., 2014). A similar observation was reported in participants who feel gratification in response to rewarding results but also by seeing the opponent player's bad luck (Falco et al., 2019). The MFN amplitude was larger following unfair as compared to fair offers (Polezzi et al., 2008; Boksem and De Cremer, 2010; Wu et al., 2011a; Alexopoulos et al., 2012; Mesrobian et al., 2015). In DG, MFN is observed in Recipient's ERP after receiving disadvantageous inequitable allocations (Li et al., 2020), but it is also dependent on the recipients' social status and altruistic tendency (Sun et al., 2015; Cui et al., 2019; Mayer et al., 2019). A late positive component (LPC) observed at latencies of 450-650 ms over central and parietal areas was sometimes termed late positive potential and often regarded as a sustained P300 response to be associated with emotion-driven decision making (Ito et al., 1998; Wu et al., 2011b; Cui et al., 2019). In UG, the P300/LPC amplitude was larger in conditions with higher risk-tendency, such as following responses to a fair offer made by an unknown Proposer in comparison with the same payoff offered by a socially known and close individual (Polezzi et al., 2010; San Martín, 2012; Yu et al., 2015; Lintas et al., 2017; Harris et al., 2020). This amplitude also tended to be larger and localized more frontally after rejection of unfair offers (Mesrobian et al., 2015) and after receiving fair over unfair offers (Xu et al., 2020).

An unanswered question from previous studies, which usually pooled participants in decision-making tasks in one group, is whether somatic markers are different in individuals with different profiles of fairness towards accepting/rejecting wretched offers in UG. This question has been investigated in this study with an original design combining DG and UG. The cooperation is tested with the same participant playing the roles of Allocator in DG and Responder in UG, thus allowing to separate individuals' profiles following an index of *altruism* determined during UG and an index of *selfishness* determined during DG. The experimental prediction is that the combination of these games will reveal the existence of different groups of individuals characterized by differences in selfishness and altruism. To our knowledge, this is the first study to carry out a systematic comparison of ERPs following the behavior of the same participants performing both DG and UG neuroeconomic games. The fronto-central N1 wave component is expected to be associated with top-down prefrontal control of decision making, meaning that N1 should be different in groups, if any, clearly separated by selfish and altruistic behaviors. Rewarding choices are expected to evoke larger P2 in participants who feel gratification, meaning that P2 is expected to be larger in altruistic and fair participants rather than in selfish and conceit participants. Being associated with disadvantageous inequitable offers and with altruistic tendency, MFN is expected to be larger in altruistic participants. The amplitude of posterior-parietal LPC is associated with motivation and subsequent allocation of attention, meaning that it is expected to be large in selfish Allocators playing DG and in altruistic Responders playing UG. An additional careful analysis of N1, P2, MFN, and LPC latencies for distinct behavioral groups is expected to point out interacting factors, if any, associated with the outcome of decision making.

2 Method

2.1 Participants

An *a priori* power analysis was conducted to compute the required sample size on the basis of repeated measures and within-between interaction (Faul et al., 2007). We planned 6 measurements per value (taken by three experienced users at two adjacent electrode sites) and two factors, each one with two levels (i.e. the accept/reject response and the UG/DG game), using a desired effect size of 0.33 (to represent a medium effect) with a power of 0.95 (1-β error probability). The power analysis indicated we need a sample size of 44 to achieve this result. Forty-four volunteer healthy young male adults were recruited via posted announcement at the university campus of the Shahid Beheshti University of Tehran (Iran), but only 39/44 completed the protocol. Then, our participant count is restricted to $N = 39$ volunteers aged between 18 and 32 years old (M=22.5, SEM=0.5). This sample size corresponds to a power of 0.92. All participants followed an academic education (8 undergraduate, 22 postgraduate and 9 PhD students) and all were right-handed with normal or corrected-to-normal vision. The research protocol was approved by the mandatory Ethics Committees requested by Shahid Beheshti University and with written informed consent from all participants, in accordance with the latest version of the Declaration of Helsinki (World Medical Association., 2000). Prior to the experimental session, all participants were interviewed to asses that they were not reporting any neurological or neuropsychiatric diseases and

not taking any psychoactive medication. The whole experiment lasted between 2 and 2 and half hours. The participants were compensated for their time with a fixed amount of cash (200,000 IRR, Iranian Rials, approximately corresponding to 4 USD). In addition, the participants received performance-based cash rewards (comprised between 200,000 and 800,000 IRR) corresponding to the cumulative outcome of the UG and DG earned at the end of the experimental session.

2.2 Experimental procedure

The participants were seated comfortably in a sound-attenuated, electrically shielded, and dimly lit cabin with the instruction to maintain their gaze on a white fixation cross at the center of a 32-inch computer screen at a viewing distance of 100 cm. At the begin of the recording session, the EEG was recorded during 2 min while the participants kept the eyes closed and during 2 min while they fixated a cross on the center of the computer screen. The participants were randomly assigned in two groups, those who played UG at first and those who played DG at first. Each game started with 10 trials of practice. After completing the first game, a short break (5-10 min) was set to allow the participant to get acquainted with the rules of the second game.

Each game was composed of 5 blocks with 48 trials each (240 trials for each game) and the participants took a break after each block. Each trial started with a fixation cross appearing for 800-1200 ms. Then, a circular pie chart representing the endowment at stake divided in two colored pieces, was presented with the diameter of 8.0 visual degree centered on a 32-inch computer screen at a 100 cm viewing distance. The pie chart appeared on the screen for 1200 ms. The red piece of the pie corresponded always to the participant's payoff of the endowment share. The participant had up to 3000 ms to make the decision either to accept the endowment subdivision represented by the pie chart —by pressing the 'K' letter of the computer keyboard with the medium finger of the right hand– or to reject it —by pressing the 'H' letter with the index finger of the same hand. During this interval, a schematic drawing appeared on the computer screen with a 'thumb up' emoji above a letter 'K' on the right and a 'thumb down' emoji above a letter 'H' on the left. After making the decision, the players fixated the cross for 800-1200 ms.

At the end of each trial, the outcome of the game showing the payoff earned by the participant and his opponent, was presented on the computer screen for 2000 ms. The stimulus presentation and data acquisition were performed using PsychLab and EEG8 amplifiers (Precision Instruments, Glastonbury, Somerset, BA6 8AQ, UK). The statistical analyses were performed using the R language and environment for statistical computing (R Core Team, 2021; Venables and Ripley, 2002).

2.3 Dictator Game

In the Dictator Game of this study, the Allocator/Proposer could not impose *a priori* the amount to be divided with the other player (the Recipient/Responder), but he had to refuse or agree with a randomly selected allocation appearing on the computer screen. As in a usual Dictator Game, the Recipient/Responder was completely passive and had to accept whatever amount was imposed by the Allocator/Proposer. Both players received zero payoff if the Allocator/Proposer refused the randomly selected allocation. All participants played the role of Allocator/Proposer and they were informed that the Recipient/Responder was an anonymous player located elsewhere. The participants played 240 trials subdivided into five blocks of 48 trials. The allocations were randomly selected from among five ([Allocator/Proposer] : [Recipient/Responder]) possibilities $\{$(50:50), (60:40), (70:30), (80:20), (90:10)$\}$ —a fraction (80:20) meaning 80% of the amount allocated to the participant and 20% allocated to the other player. A single block of 48 trials was not necessarily balanced for the five possible fractions, but the complete sequence of 240 (5×48) trials was uniformly distributed, such that the sequence contained the same number of trials for each pre-determined allocation. In this version of DG, the performance-based cash reward of the Participant (always playing the role of Allocator/Proposer) was determined by the average payoff of all accepted proposals. Figure 1 presents a schematic illustration of the time course of two DG trials.

2.4 Ultimatum Game

All participants played the role of Responder, who could either accept or reject the offers appearing on the computer screen. The participants were informed that the Proposer was an anonymous player located elsewhere. In case of acceptance, each player received a payoff according to the proposed split, otherwise both players did not receive anything. The overall sequence of 240 trials was subdivided into five blocks of 48 trials with five possibilities of (Proposer:Responder) endowment shares $\{$(50:50), (60:40), (70:30), (80:20), (90:10)$\}$ —a fraction (80:20) meaning 80% of the endowment paid to the Proposer (i.e., the opponent player) and 20% paid to the Participant, who always played the role of Responder. The trials were uniformly randomly distributed such that the complete sequence contained the same number of trials for each pre-determined endowment share. The performance-based cash reward of the

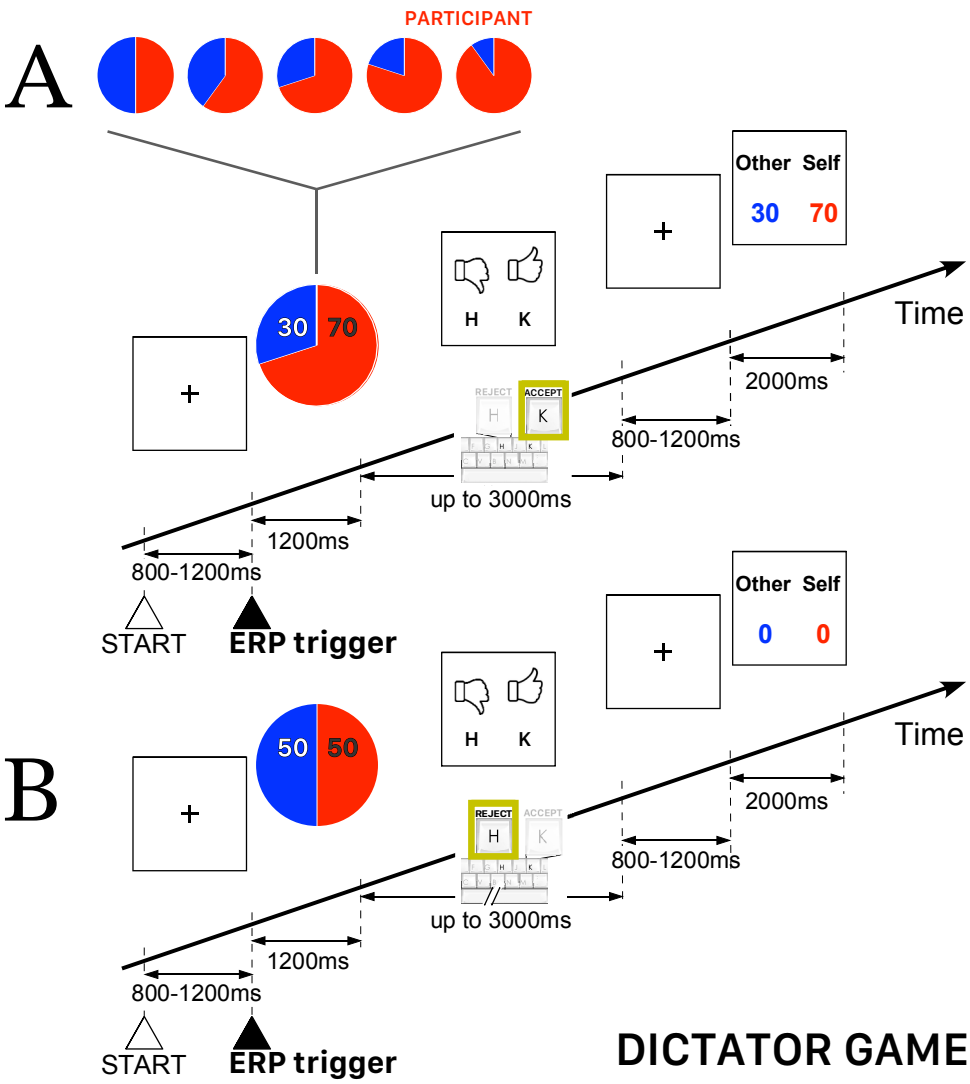

Figure 1: Schematic illustration of the time course of the Dictator Game task. The participant always played the role of Allocator/Proposer with the possibility to accept or refuse an allocation randomly presented on the computer screen. The red piece of the pie corresponds to participant's payoff. **A.** Example of a trial when the Allocator/Proposer agreed with the randomly proposed allocation —in this example, 70% for the participant and 30% for the other player. **B.** Example of a trial when both players ended with a zero payoff because the Allocator/Proposer refused the randomly proposed allocation (50:50).

participant was determined, as usually in UG, by the total payoff of all accepted endowment shares. Figure 2 presents a schematic illustration of the time course of two UG trials.

2.5 EEG recording and ERP processing

The electroencephalogram (EEG) of each Participant was recorded using 32 scalp dry titanium/titanium nitride (Ti/TiN) electrodes, mounted on a headcap (international 10/20 layout) referenced to the left mastoid bone. Electrophysiological signals were sampled at $1024\ Hz$ and filtered with a band-pass of $0.1 - 200$ Hz, with electrode impedances kept near $5\ k\Omega$ for all recordings. The analysis was performed using the EEGLAB_v12 (Delorme and Makeig, 2004)) and ERPLAB_v7 toolboxes (Lopez-Calderon and Luck, 2014)) of MATLAB 2017b (The Mathworks, Inc., Natick, MA, USA). Raw data were preprocessed with a band-pass IIR Butterworth filter from 0.1 to 32 Hz ($-36dB$/octave roll off). Blink, saccade and eyelid artifact components were corrected or set to zero, based on their respective shape

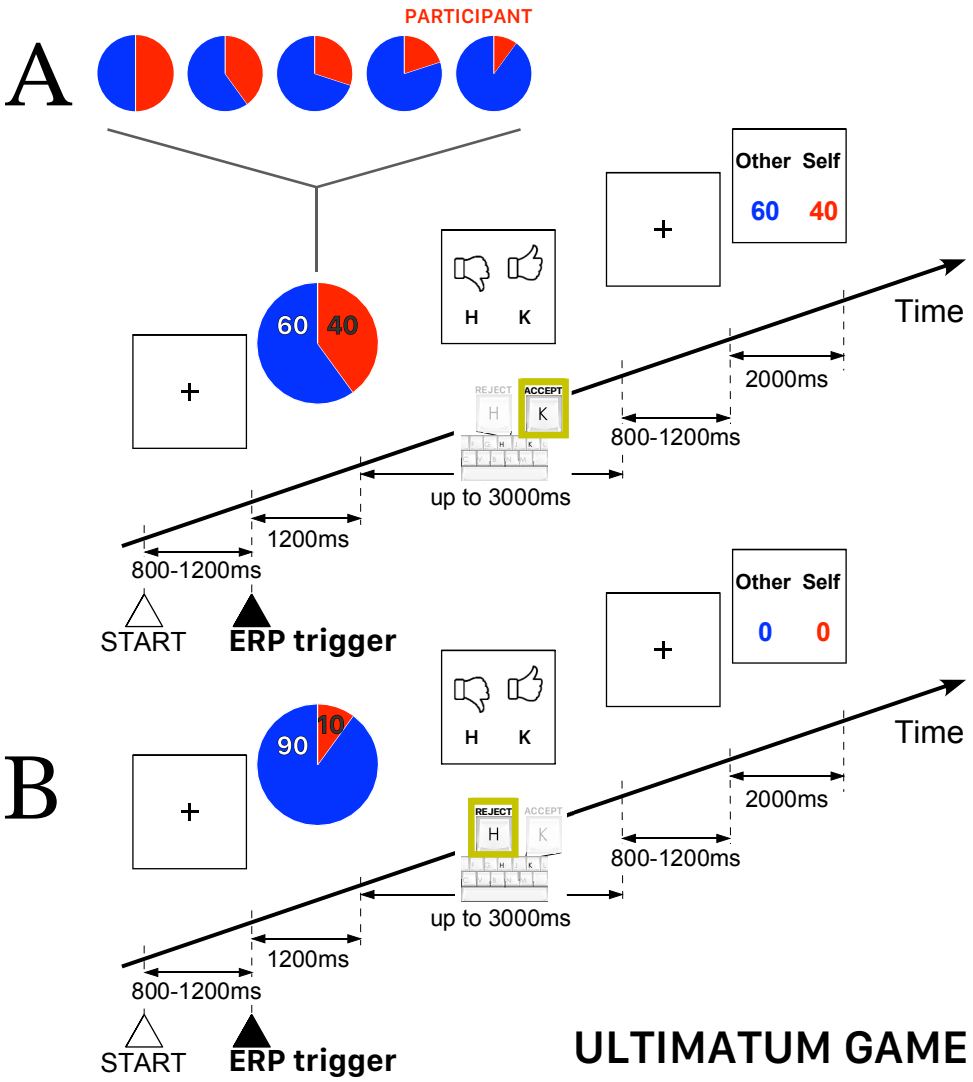

Figure 2: Schematic illustration of the time course of the Ultimatum Game task. The Participant always played the role of Responder and could either accept or reject the endowment share randomly presented on the computer screen. The red piece of the pie corresponds to participant's payoff. **A.** Example of a trial when a Participant accepted the randomly proposed share—in this example, 60% for the other player and 40% for the Participant playing the role of Responder. **B.** Example of a trial when a Participant rejected the randomly proposed share—in this example, 90% for the the other player and 10% for the Participant—thus ending with a zero payoff for both players.

and topography after using Independent Component Analysis (ICA) (Jung et al., 2000; Plöchl et al., 2012). Epochs time-locked to markers were obtained after off-line segmentation of the continuous EEG. The epochs were defined for the interval between -200 through 800 ms around the trigger and corrected to baseline 200 ms prior to marker. Noisy data were rejected after further visual inspection of the epochs for contamination by muscular or electrode artifacts. We analyzed four ERP components defined by the early negative peak near 110–160 ms after the trigger (N1), followed by a positive wave peaking in the 180–240 ms interval post trigger (P2), a negative component observed in the 260-400 ms interval (MFN), and the most positive peak observed in the 430-630 ms interval (LPC). For each participant and at each electrode site, measurements of each peak latency and amplitude, if peaks of wave components were clearly identified, were taken by three independent trained experimenters. In this way, we could perform repeated measurements (three at maximum) for each wave component at each electrode site.

2.6 Statistical analyses

The statistical software R v4.0.5 (R Core Team, 2021). was used for all the analyses with packages `effectsize` (Makowski et al., 2019), `rstatix` (Kassambara, 2020) and `sjstats` (Lüdecke, 2021). In general, the grouped values were reported as (median, Mean ±SEM). The null hypothesis of homoscedasticity (i.e., equal variances) in the data samples was tested with the Levene's test. Unpaired comparisons between groups were performed by Student t-test if variances were equal and by Welch t-test, otherwise. For multiple comparisons, the p-value of t-test was adjusted following Bonferroni correction. The factorial analysis was performed using linear mixed effects models (with package `lmerTest`; Kuznetsova et al., 2017) and goodness of fit was assessed by F-statistics based on Satterthwaite's method for denominator degrees-of-freedom and F-statistic). The generalized eta squared (η^2_G) was used to estimate the effect size of main and interaction factors (*negligible* if $\eta^2_G < 0.01$; *small* if $0.01 \leq \eta^2_G < 0.06$; *medium* if $0.06 \leq \eta^2_G < 0.14$; *large* if $0.14 \leq \eta^2_G$). For independent two-samples comparisons, the effect size of t-test was assessed by Cohen's d (*negligible* if $d < 0.2$; *small* if $0.2 \leq d < 0.5$; *medium* if $0.5 \leq d < 0.8$; *large* if $0.8 \leq d$).

3 Results

3.1 Behavioral analysis

On average (M±SEM), the participants to the Dictator Game completed 237.3±0.7 out of 240 trials. This difference is due to the fact that, in a few trials, some participants did not take any decision (to agree or refuse the proposed allocation) within the maximum interval of 3000 ms. The valid DG trials were subdivided in three sets: *(i) DG selfish behavior* including the trials when the participant playing the role of Allocator/Proposer agreed with the most favorable proposals of allocations for himself {(90:10), (80:20)} (i.e. 90% or 80% of the amount in favor of the participant) and refused the least favorable amounts {(60:40), (50:50)}; *(ii) DG neutral behavior* including all trials (both agreed and refused) with an allocation (70:30); *(iii) DG fair behavior* including the trials when the participant playing the role of Allocator/Proposer agreed with the least favorable allocations for himself {(60:40), (50:50)} (i.e. 60% or 50% of the amount in favor of the participant) and refused the most favorable allocations {(90:10), (80:20)}. Hence, on the basis of the number (Nb.) of trials belonging to these three sets, we computed the index $DG_{selfishness}$ (in the range [-1,+1]), such that the higher the index the more selfish the behavior, defined as follows:

$$DG_{selfishness} = \frac{(\mathrm{Nb}_{(DG\ selfish\ behavior)} - \mathrm{Nb}_{(DG\ fair\ behavior)})}{(\mathrm{Nb}_{(DG\ selfish\ behavior)} + \mathrm{Nb}_{(DG\ fair\ behavior)})} \quad . \tag{1}$$

In the Ultimatum Game, all participants played the role of Responder and completed 238.1±0.4 trials. The behavioral analysis during UG was based on the subdivision of valid trials in the next three sets: *(i) UG altruistic behavior* including the trials when the Responder accepted the least favorable offers for himself {(90:10), (80:20)} (i.e. 10% or 20% of the share for himself) and rejected the equitable splits {(60:40), (50:50)}; *(ii) UG neutral behavior* including all trials (both accepted and rejected) offering an endowment share (70:30); *(iii) UG conceit behavior* including the trials when the Responder accepted the equitable offers {(60:40), (50:50)} and rejected the least favorable offers {(90:10), (80:20)}. Notice that in this version of UG, the equitable splits {(60:40), (50:50)} corresponded also to the most favorable payoffs that the participant could expect to receive. In a way similar to DG, we computed an index termed $UG_{altruism}$ (in the range [-1,+1]) based on the the number of trials in each set, such that the higher the index the more altruistic (i.e., the less conceit) the behavior, defined as follows:

$$UG_{altruism} = \frac{(\mathrm{Nb}_{(UG\ altruistic\ behavior)} - \mathrm{Nb}_{(UG\ conceit\ behavior)})}{(\mathrm{Nb}_{(UG\ altruistic\ behavior)} + \mathrm{Nb}_{(UG\ conceit\ behavior)})} \quad . \tag{2}$$

On the basis of both indices $DG_{selfishness}$ and $UG_{altruism}$, we performed an agglomerative hierarchical clustering using the function `hclust` set with a metric Euclidean distance and the complete-linkage option of the standard `stats` package of R (R Core Team, 2021). At the begin, the procedure considered the possibility of 39 clusters, i.e. one cluster for each participant. Then, the algorithm analyzed iteratively the outcome of the clustering such that the number of clusters decreased at each step and and the procedure finally stabilized and stopped with four clusters, labeled 'GrpS', 'GrpB', 'GrpC', 'GrpF'. The procedure was repeated with different random seeds at the initialization. The outcome of repeating the procedure was that the attribution of three individuals to groups 'GrpS' or 'GrpB' depended on the initial random seed. Then, we defined an additional cluster labeled 'GrpA' formed by those three individuals.

Table 1 reports the behavioral features of the clusters of participants and cluster 'GrpA' was composed by three participants whose behavior was intermediate between 'GrpS' and 'GrpB'. Participants who expressed a conceit

Table 1: Relative frequencies (median, Mean±SEM) of the behavioral responses to the Dictator and Ultimatum Game and corresponding behavioral indices $DG_{selfishness}$ and $UG_{altruism}$ of the five groups of participants sorted after an agglomerative hierarchical clustering procedure described in the text.

	$N = 39$	'GrpS' ($n = 8$)	'GrpA' ($n = 3$)	'GrpB' ($n = 10$)	'GrpC' ($n = 8$)	'GrpF' ($n = 10$)
Dictator Game						
Selfish behavior (%)		79.2 77.4±1.3	65.0 67.9±5.7	69.4 68.5±2.6	45.0 49.5±4.1	3.1 8.6±2.9
Fair behavior (%)		0.8 2.7±1.3	16.4 12.6±5.8	10.6 11.5±2.7	34.9 30.4±4.1	77.0 71.5±2.9
$DG_{selfishness}$		0.98 0.93±0.03	0.60 0.69±0.14	0.73 0.71±0.07	0.13 0.24±0.10	-0.92 -0.79±0.07
Ultimatum Game						
Altruistic behavior (%)		1.3 4.5±2.0	12.2 12.7±3.8	23.9 245±1.9	28.4 29.4±7.2	1.3 2.2±0.6
Conceit behavior (%)		78.7 75.4±2.0	67.5 67.1±3.9	56.1 55.4±1.9	51.5 50.5±7.2	78.6 77.8±0.51
$UG_{altruism}$		-0.97 -0.89±0.05	-0.69 -0.68±0.10	-0.40 -0.39±0.05	-0.29 -0.26±0.18	-0.97 -0.95±0.01

behavior during UG but a low level of selfishness during DG formed the stable cluster of fair participants (GrpF). Participants with a very high value of $DG_{selfishness}$ and the lowest level of $UG_{altruism}$ formed the stable cluster of selfish participants (GrpS). Cluster GrpB corresponded to participants with a medium–low $UG_{altruism}$ index and medium–high $DG_{selfishness}$ index. GrpB participants could be considered a group of more altruistic/less conceit participants within the context of this study. The most heterogeneous group of participants (GrpC) was characterized by variable $UG_{altruism}$ and medium–low $DG_{selfishness}$. Figure 3 shows the scatterplot of the 39 participants distributed on a 2D feature space defined by their corresponding values of $UG_{altruism}$ and $DG_{selfishness}$.

3.2 Reaction Times

Reaction times were measured between the time of the presentation of the allocation proposal in DG (presentation of the endowment share in UG) and the time of pressing the letter key on the computer keyboard. In either game, separately, the RTs for all 39 participants were faster when the participants agreed *vs.* refused the allocation during DG (median RT 381.0, 492.0 ± 4.9 ms *vs.* 408.0, 514.2 ± 5.1 ms; Student $t(9093) = -3.146$, $p = .002$, $d = -0.07$) or accepted *vs.* rejected the proposed endowment share during UG (median RT 403.0, 500.3 ± 4.3 ms *vs.* 426.5, 534.1 ± 6.4 ms; Student $t(8961) = -4.516$, $p < .001$, $d = -0.10$). Despite these differences, RTs were distributed as usually, i.e. a positively skewed distribution with a long tail which reflected occasional slow responses. These long-tailed distributions of RTs introduced a bias that was amplified by the differences between single participants. For this reason, to account within and between participants variance of raw RTs, a mixed model ANOVA with repeated measurements did not reveal the main effect of response choice (DG: $F(1, 76) = 2.519$, $p = .117$, $\eta^2_G = 0.03$; UG: $F(1, 74) = 1.694$, $p = .197$, $\eta^2_G = 0.02$). Instead of raw RTs, z-scores of RTs computed for each participant and each game separately were considered further for factorial analysis using mixed models with repeated (i.e., single trials) measurements. With z-scores of RTs the effect of factor $response$ was significant in each game (DG: $F(1, 76) = 14.47$, $p < .001$, $\eta^2_G = 0.16$; UG: $F(1, 74) = 24.62$, $p < .001$, $\eta^2_G = 0.25$).

We focused on less conceit/more altruistic (GrpB), fair (GrpF), and selfish (GrpS) participants, which corresponded to the three most homogenous groups revealed by the behavioral analysis. For each game separately, we run an analysis with factors *response* (2 levels: Agreed/Accepted and Refused/Rejected in DG/UG) and *group* (3 levels: GrpB, GrpF and GrpS) with repeated measurements. A linear mixed model fit by maximum likelihood was performed due to the presence of unbalanced samples (Table 2). A strong main effect of Response was observed in both games, with reaction times faster when the participants agreed with the allocation or accepted the proposed endowment share irrespective of the participants' group.

During DG we categorized four sets of trials according to the behavior: (i) the *Selfish trials* with the Allocator/Proposer in 'agreement' with the randomly selected allocations {(90:10), (80:20)} (i.e. 90% or 80% of the share for the participant); (ii) the *Selfish trials* with the Allocator/Proposer 'refusing' the least favorable (in the current task design) allocations for himself, i.e. {(60:40), (50:50)}; (iii) the *Fair trials* with the Allocator/Proposer in 'agreement'

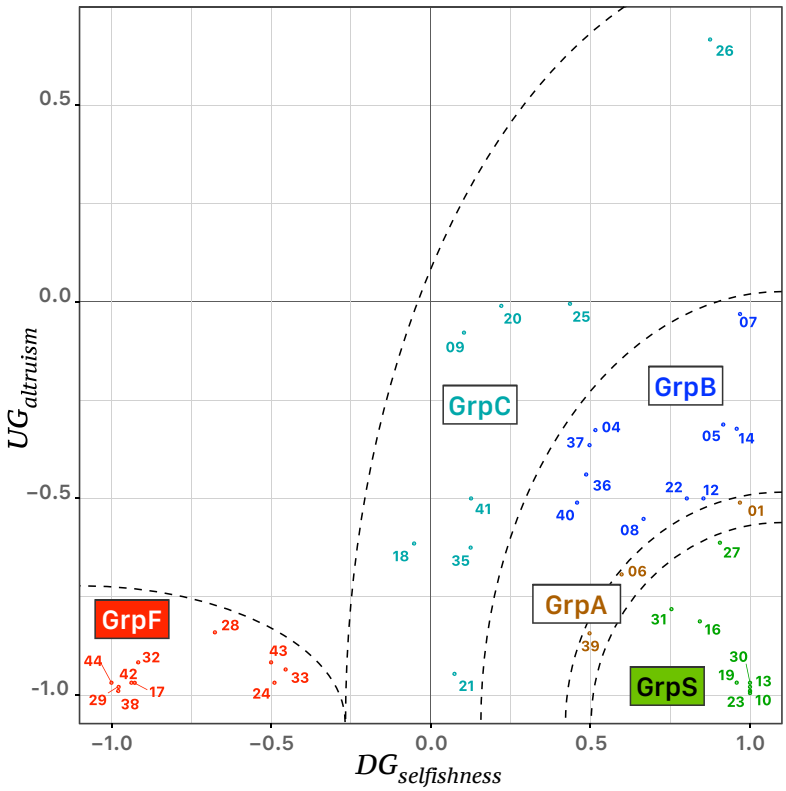

Figure 3: Scattergram of the behavioral analysis of 39 participants, identified by their identification tag, distributed in a 2D feature space defined by the corresponding values of $UG_{altruism}$ and $DG_{selfishness}$. Five clusters of colored points were identified on the bases of an agglomerative hierarchical clustering procedure, i.e. 'GrpS' (green), 'GrpA' (brown), 'GrpB' (blue), 'GrpC' (turquoise), and 'GrpF' (red). The dashed lines correspond to ideal separatrix lines in the feature space.

with the allocations {(60:40), (50:50)}; (iv) the *Fair trials* with the Allocator/Proposer 'refusing' the most favorable allocations for himself, i.e. {(90:10), (80:20)}. Along a similar scheme, four sets of trials were categorized during UG: (i) the *Altruistic trials* when the Responder 'accepted' the least favorable outcomes that granted only 10% or 20% of the share to the participant (i.e., shares {(90:10), (80:20)}); (ii) the *Altruistic trials* when the Responder 'rejected' equitable offers (i.e., shares {(60:40), (50:50)}); (iii) the *Conceit trials* when the Responder 'accepted' the equitable offers (i.e., the participant received 40% or 50% of the share that are also the most advantageous payoffs in the current task design); (iv) the *Conceit trials* when the Responder 'rejected' the least favorable payoff. Note that during UG any trial ended without a payoff for the participant if he rejected the offer (either altruistic or conceit).

For both games, the factorial analysis (Table 2) was run with factors *response* (2 levels) and *behavior* (2 levels). During UG, we observed a very strong effect of behavior (RTs during conceit trials shorter than during altruistic trials) $F(1, 89) = 4.111$, $p = .045$, $\eta_G^2 = 0.21$) and confirmed the main effect (although weaker) of the response choice ($F(1, 89) = 23.53$, $p < .001$, $\eta_G^2 = 0.04$). On the contrary, during DG no main effects of response and behavior were observed in groups (GrpB, GrpF, GrpS).

A multiple regression was conducted to see if the indices $DG_{selfishness}$ and $UG_{altruism}$ predicted the RTs in a specific set of trials. During DG, it was found that the level of $DG_{selfishness}$ alone always explained a significant amount of the variance in RTs in all sets of trials. We considered both linear and quadratic fits for raw RTs and z-scaled RTs as explained above, but quadratic fits always provided better results. During selfish trials (Figure 4A), irrespective of the response choice, RTs tended to be faster for participants whose $DG_{selfishness}$ index was high ($F(2, 53) = 17.00$, $p < .001$, $R^2 = .391$, $R^2_{Adjusted} = .368$). During fair trials, it is interesting to note a specular curve (Figure 4B), characterized by faster RTs for participants whose $DG_{selfishness}$ index was low ($F(2, 53) = 26.66$, $p < .001$, $R^2 = .447$, $R^2_{Adjusted} = .430$). During UG, $UG_{altruism}$ alone explained most of the variance only in a specific set of

Table 2: **Reaction times**. The table reports measurements of the reaction times (RTs in ms) (median, mean ±SEM values and number of trials n) when participants agree (or disagree) with the proposed allocation during the Dictator Game, and when they accept (or reject) the proposal of endowment share during the Ultimatum Game. Statistics of Levene's test for heteroscedascity (homogeneity of variances test) and ANOVA-like tables for a linear mixed model fit by maximum likelihood with corresponding main factors and interaction are reported for z-scores of RTs computes for each participant separately.

	Dictator Game: Allocator		**Ultimatum Game: Responder**	
Factor *Response choice*	Agreed	Refused	Accepted	Rejected
Participants' Group				
	$n = 855$	$n = 1032$	$n = 1481$	$n = 395$
GrpB ($N = 10$)	355.0	373.5	390.0	409.0
	423.5±9.1	464.5±9.5	502.1±8.7	556.2±20.3
(*z-score*)	*(-0.031±0.029)*	*(0.024±0.029)*	*(-0.027±0.022)*	*(0.110±0.052)*
	$n = 763$	$n = 1077$	$n = 890$	$n = 922$
GrpF ($N = 10$)	398.0	432.0	400.0	430.0
	515.1±12.5	543.6±11.1	470.9±9.5	520.9±11.4
(*z-score*)	*(-0.066±0.033)*	*(0.047±0.027)*	*(-0.064±0.028)*	*(0.065±0.031)*
	$n = 769$	$n = 738$	$n = 724$	$n = 692$
GrpS ($N = 8$)	391.0	415.0	440.0	426.5
	488.3±11.0	521.8±12.1	521.7±11.8	512.6±13.0
(*z-score*)	*(-0.064±0.030)*	*(0.066±0.035)*	*(-0.032±0.031)*	*(0.036±0.036)*
Homogeneity of variances *Levene's test:*	$F(5, 50) = 0.921, p = .475$		$F(5, 50) = 9.008, p < .001$	
Anova-like table Effect: *Response*	$F(1, 50) = 12.41, p < .001$		$F(1, 50) = 23.76, p < .001$	
Group	$F(1, 50) = 0.225, p = .800$		$F(2, 50) = 2.688, p = .078$	
Response × Group	$F(1, 50) = 0.537, p = .588$		$F(2, 50) = 3.108, p = .053$	
Participants' Behavior	Fair trials		Altruistic trials	
	$n = 862$	$n = 1099$	$n = 622$	$n = 77$
	395.5	429.0	463.0	298.0
	518.6±12.3	533.2±10.2	597.8±16.5	625.3±78.5
(*z-score*)	*(-0.047±0.035)*	*(0.020±0.029)*	*(0.242±0.048)*	*(0.573±0.217)*
	Selfish trials		Conceit trials	
	$n = 1525$	$n = 1748$	$n = 2473$	$n = 1932$
	364.0	390.0	393.0	428.5
	448.3±6.9	494.2±8.0	472.6±5.7	521.0±7.7
(*z-score*)	*(-0.070±0.023)*	*(0.063±0.026)*	*(-0.145±0.016)*	*(0.047±0.023)*
Homogeneity of variances *Levene's test:*	$F(3, 88) = 1.150, p = .334$		$F(3, 89) = 18.36, p < .001$	
Anova-like table Effect: *Response*	$F(1, 88) = 0.196, p = .659$		$F(1, 89) = 4.111, p = .045$	
Behavior	$F(1, 88) = 1.537, p = .218$		$F(1, 89) = 23.53, p < .001$	
Response × Behavior	$F(1, 88) = 0.312, p = .578$		$F(1, 89) = 1.018, p = .316$	

trial. i.e. when participants rejected the proposed endowment share during conceit trials ($F(1, 35) = 16.53, p < .001$, $R^2 = .321$, $R^2_{Adjusted} = .301$).

3.3 Electrophysiological results

Figure 5 shows the grand average ERPs for all trials recorded at fronto-central (FCz and Fz) and posterior-parietal sites (Pz and CPz) triggered by the stimulus onset in both games and followed by a response (accept or reject) for all 39 participants grouped together. The visual inspection of the grand-averaged waveforms showed several components, which were generally observed in both games and at various scalp locations, but their amplitudes varied greatly as a function of the participant and of the recording site. The earliest component was a negative wave (N1) ranging in latency between 110 and 160 ms post trigger, particularly visible at fronto-central sites, and at FCz in particular. A posterior-parietal N1 was often observed, but it was tiny and noisy in the ERP of several participants, which led us to discard this wave from further quantitative analysis. This wave was followed by P2, a positive component much sharper at fronto-central sites peaking at approximately 220 ms post trigger and by a negative wave peaking within the interval

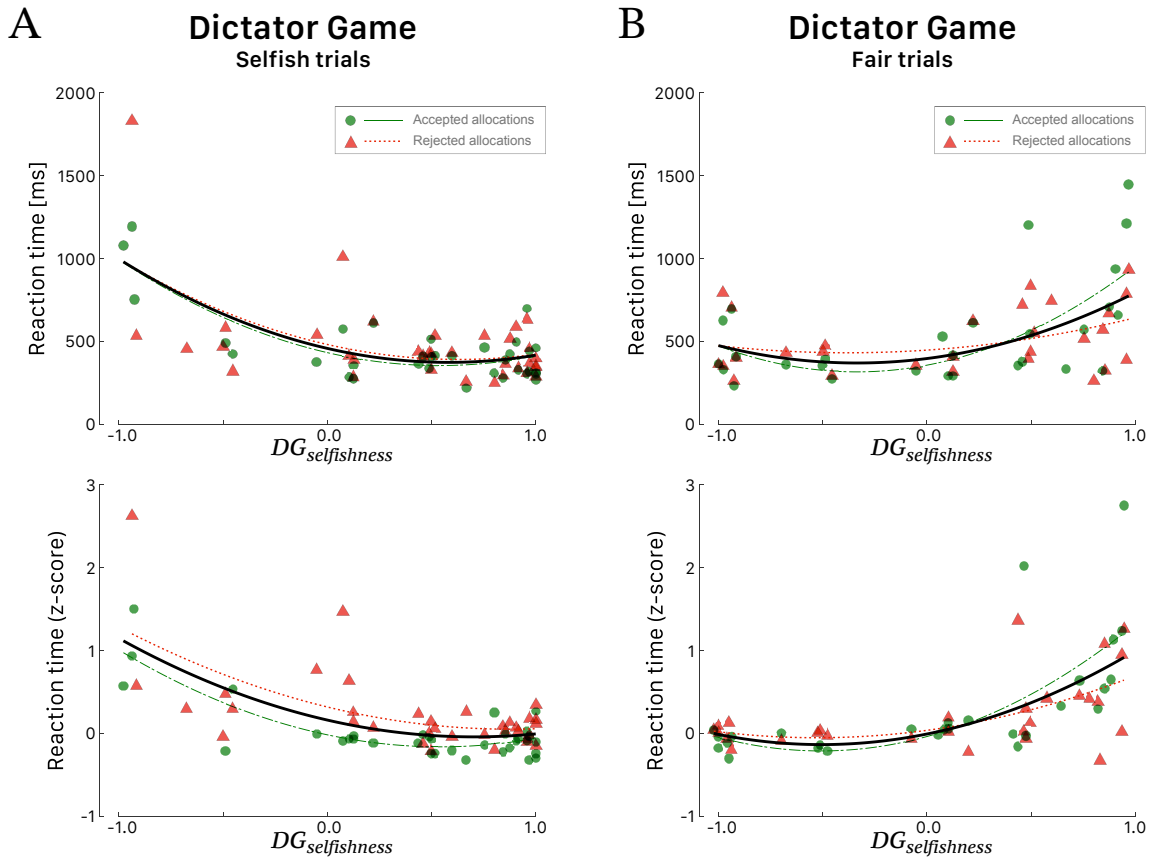

Figure 4: Reaction times (in ms, top panels; corresponding z-score in bottom panels) for all 39 participants during the Dictator Game as a function of the $DG_{selfishness}$ behavioral index. All participants played the role of Allocator. RTs during *Selfish trials* **(A)** and *Fair trials* **(B)**. Green dots correspond to participants' RTs when they agreed with the proposed allocation and red triangles when they refused the allocation. Green (long dashed) and red curves correspond to quadratic fitted regressions. The black curves correspond to quadratic fits irrespective of the response choice.

260-400 ms after trigger onset. Note that the amplitude of this negative component was deeper at fronto-central sites for both games, which led us to identify it as the medial frontal negativity (MFN). At posterior-parietal sites (Figure 5), the form and latency of this negativity were different and the component was called N2. A positive wave extending from 400 to 650 ms after stimulus onset followed MFN/N2 at all sites. This extended positivity was characterized by two successive peaks, which might be identified as the putative P3 and the late positive potential. At posterior-parietal sites, the amplitude of this positivity was larger than at other sites and the second peak was larger than the first peak. Then, the second peak at posterior-parietal sites appeared as being the most representative and was identified as the peak of the late positive component (LPC).

A comparative analysis of the grand averaged ERPs during DG and UG, with all trials pooled irrespective of payoff and participants' decision, was carried out on peak latencies and amplitudes. The analysis was focused on less conceit/more altruistic (GrpB), fair (GrpF), and selfish (GrpS) participants, which corresponded to the three most homogenous groups revealed by the behavioral analysis. We run an analysis with factors *Game* (2 levels: UG and DG) and *Group* (3 levels: GrpB, GrpF and GrpS) with repeated measures (i.e., two electrodes for each area and three measurements at each site, as described in the Methods section). Due to the presence of unbalanced groups we performed the factor analysis with a linear mixed model fit by maximum likelihood and present the results in an Anova-like table (Table 3). Note that wave peaks were not always visible in all participants and in both games, and therefore the number of measurements (n in Table 3) for each variable was not always a regular multiple of the number of participants. For example, in group GrB with $N = 10$ participants, the maximum number of measurements for a variable was $n = 60$, but it could actually be as low as $n = 48$ (e.g., N1 peak of GrpB during DG) if a peak could be uniquely identified at all electrodes.

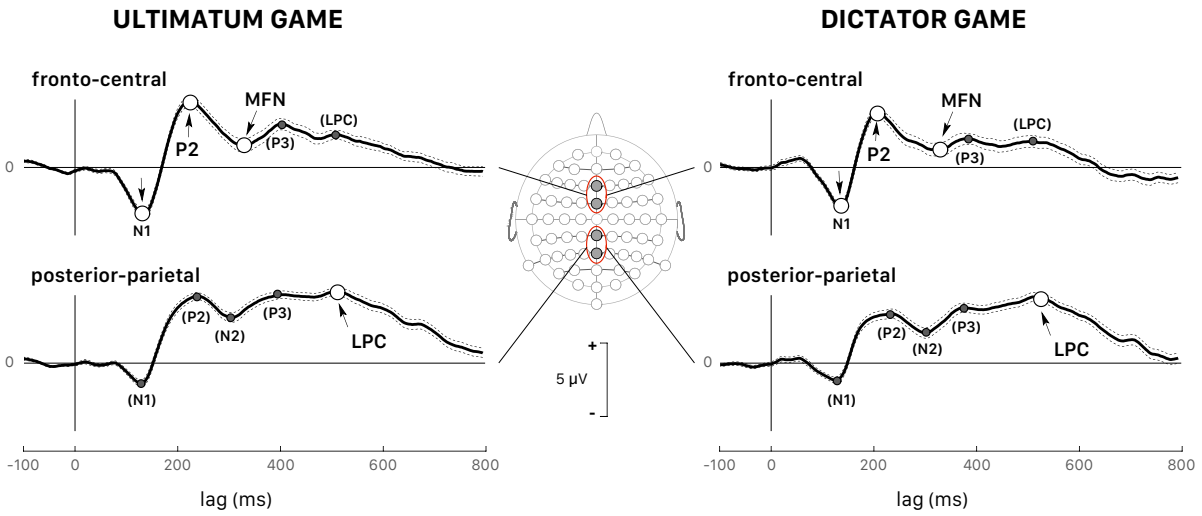

Figure 5: Grand average (all 39 participants and *all trials pooled together*) of the ERPs triggered by stimulus presentation during the Ultimatum Game (left panels) and the Dictator Game (right panels). The dotted lines show the confidence intervals determined by the standard error of the mean (SEM). All peaks indicated with arrows correspond to the wave components whose latency and peak were analyzed in detail: P2 and medial frontal negativity (MFN) at the fronto-central sites (Fz and FCz, upper panels) and the late positive component (LPC) at the posterior-parietal sites (Pz and CPz, lower panels). All peaks identified within parentheses indicate the wave components that were not considered in the quantitative analysis.

In selfish participants (dashed green curves in Figure 6), the amplitude of N1 peak negativity was reduced in both games, which accounted for the main effect of factor *Group* (Figure 6). In fair participants (solid red curves in Figure 6), P2 peak latency was much shorter during DG (lower panel) than UG (upper panel) ($t(110) = -6.198$, $p < .001$, $d = -1.17$). During DG (lower panel), the negativity of MFN in fair participant was reduced (GrpF *vs.* GrpB: $t(118) = 3.819$, $p = .001$, $d = 0.70$; GrpF *vs.* GrpS: $t(106) = 5.828$, $p < .001$, $d = 1.13$). At posterior-parietal sites, the main effect of factor *Group* for LPC amplitude (Table 3) was determined by larger LPC wave of GrpS during UG (GrpS *vs.* GrpB: $t(106) = 6.733$, $p = .001$, $d = 1.30$; GrpS *vs.* GrpF: $t(106) = 9.562$, $p < .001$, $d = 1.85$). Note that at the fronto-central sites of selfish participants, a sustained late positive wave also appeared during UG (dashed green curves in Figure 6, upper panel). Overall, these observations highlight different brain dynamics between fair and selfish participants, but these curves were obtained with a super pool of trials, which means all trials regardless of the amount of the endowment share, and of the decision-making process with related payoff, if any. In the next two subsections, we present the most salient results on the ERP wave components for each game, separately.

3.3.1 Dictator Game

We considered separately four sets of ERP recorded trials following the same categories used for the analysis of reaction times: *Selfish trials* when the participant (i) 'agreed' with the highest allocations for himself and (ii) 'disagreed' with the least advantageous allocations; *Fair trials* when the participant (iii) 'agreed' with the least advantageous allocations and (iv) 'disagreed' with the most favorable allocations. For each participant, all single trials of one set were averaged together at each electrode site. The four sets of trials were analyzed by three observers at two electrode sites (Fz and FCz for fronto-central ERP waves and Pz and CPz for LPC). We focused our analysis on the three main groups of participants, which represented a total of 28 participants. Then, the theoretical maximum number of measurements of a kind for the whole set of four trials was $n = 28 \text{ participants} \times 4 \text{ sets of trials} \times 2 \text{ electrodes} \times 3 \text{ observers} = 672$ for each ERP wave.

We analyzed the effect of factors *response* (2 levels: Agreed, Refused) and *behavior* (2 levels: Selfish trials, Fair trials) on latencies and amplitudes of major ERP peaks by the linear mixed model with repeated measures (Table 4). In different trials and in different participants, ERP wave peaks could not always be distinguished by all observers at all electrode sites, which ultimately represented an overall number (n) of measurements between 476 and 513. In order to analyze the effect of these factors independently, the table reports the values measured for the response type irrespective of the behavior and the values measured for the behavior irrespective of the response. Note that a significant main effect of participant's *behavior* was observed on both latency and amplitude for all ERP waves (Table 4). The main effect

Table 3: **ERP wave components**. The table reports measurements of the peak latency (in ms) and amplitude (in μV) (median, mean $\pm$SEM values and number of measurements n) of N1, P2, and MFN recorded at the fronto-central sites (Fz and FCz) and LPC recorded at the posterior-parietal sites (Pz and CPz) irrespective of the endowment share and of participants' decision. The table reports Levene's test for heteroscedascity (homogeneity of variances test) and type III Anova-like table for a linear mixed model fit by maximum likelihood with factors *Grp* and *Game*, and unpaired *t*test comparisons for values of only fair and selfish groups of participants.

	fronto-central N1				fronto-central P2			
	Latency $[ms]$		Amplitude $[\mu V]$		Latency $[ms]$		Amplitude $[\mu V]$	
Factor *Game*	DG: Allocator	UG: Responder	DG: Allocator	UG: Responder	DG: Allocator	UG: Responder	DG: Allocator	UG: Responder
Participants' Group								
GrpB ($N = 10$)	$n = 48$ 139.1 139.3±1.7	$n = 54$ 130.8 130.0±1.2	$n = 48$ -4.04 -4.67±0.36	$n = 54$ -3.94 -3.92±0.20	$n = 60$ 222.1 217.2±2.4	$n = 58$ 222.7 223.4±2.0	$n = 60$ 5.01 4.81±0.26	$n = 58$ 4.72 4.77±0.14
GrpF ($N = 10$)	$n = 60$ 138.4 139.2±1.2	$n = 60$ 140.7 141.0±2.0	$n = 60$ -2.83 -3.14±0.28	$n = 60$ -3.53 -3.65±0.17	$n = 56$ 201.3 201.8±2.5	$n = 56$ 226.0 227.1±3.2	$n = 56$ 4.40 4.40±0.26	$n = 56$ 4.83 4.98±0.22
GrpS ($N = 8$)	$n = 39$ 128.2 128.9±2.3	$n = 48$ 127.8 131.9±2.0	$n = 39$ -1.66 -1.40±0.19	$n = 48$ -2.33 -2.66±0.30	$n = 48$ 205.9 207.0±3.5	$n = 48$ 211.3 212.5±2.4	$n = 48$ 3.31 4.15±0.34	$n = 48$ 3.74 4.12±0.32
unpaired t-test (GrpF, GrpS)	$t(97) = 4.357$	$t(106) = 3.159$	$t(97) = -4.623$	$t(106) = -2.993$	$t(102) = -1.239$	$t(102) = 3.515$	$t(102) = 0.603$	$t(102) = 2.254$
p-value	$p < .001$	$p = .012$	$p < .001$	$p = .018$	$p = .218$	$p = .004$	$p = .548$	$p = .156$
Cohen's d	$d = 0.90$	$d = 0.61$	$d = -0.95$	$d = -0.58$	$d = -0.24$	$d = 0.69$	$d = 0.12$	$d = 0.44$
Homogeneity of variances *Levene's test:*	$F(5, 46) = 1.080, p = .384$		$F(5, 46) = 0.896, p = .492$		$F(5, 50) = 0.595, p = .704$		$F(5, 50) = 1.188, p = .329$	
Anova-like table Effect: $Game$	$F(1, 46) = 0.334, p = .566$		$F(1, 46) = 0.566, p = .456$		$F(1, 50) = 4.109, p = .048$		$F(1, 50) = 0.019, p = .892$	
$Group$	$F(2, 46) = 2.087, p = .136$		$F(2, 46) = 6.055, p = .005$		$F(2, 50) = 1.130, p = .331$		$F(2, 50) = 0.427, p = .655$	
$Game \times Group$	$F(2, 46) = 1.016, p = .370$		$F(2, 46) = 1.334, p = .273$		$F(2, 50) = 1.068, p = .351$		$F(2, 50) = 0.083, p = .920$	

	fronto-central MFN				posterior-parietal LPC			
	Latency $[ms]$		Amplitude $[\mu V]$		Latency $[ms]$		Amplitude $[\mu V]$	
Factor *Game*	DG: Allocator	UG: Responder	DG: Allocator	UG: Responder	DG: Allocator	UG: Responder	DG: Allocator	UG: Responder
Participants' Group								
GrpB ($N = 10$)	$n = 60$ 326.5 323.1±4.3	$n = 60$ 328.1 319.7±3.8	$n = 60$ -2.23 -0.91±0.58	$n = 60$ 1.07 0.10±0.42	$n = 60$ 524.3 525.6±3.8	$n = 60$ 531.1 537.5±6.2	$n = 60$ 5.12 4.45±0.32	$n = 60$ 3.66 3.88±0.20
GrpF ($N = 10$)	$n = 60$ 314.6 307.8±4.2	$n = 56$ 338.1 336.5±2.4	$n = 60$ 1.26 1.56±0.29	$n = 56$ 0.62 0.84±0.17	$n = 60$ 520.6 520.1±4.3	$n = 60$ 581.0 574.3±5.9	$n = 60$ 4.19 3.97±0.32	$n = 60$ 2.14 2.20±0.30
GrpS ($N = 8$)	$n = 48$ 292.4 307.2±5.0	$n = 42$ 318.3 315.9±4.5	$n = 48$ -2.13 -1.38±0.43	$n = 42$ 0.68 -0.30±0.41	$n = 48$ 551.4 558.2±5.9	$n = 48$ 530.5 535.7±3.5	$n = 48$ 4.96 4.67±0.21	$n = 48$ 5.75 5.82±0.21
unpaired t-test (GrpF, GrpS)	$t(106) = 0.096$	$t(96) = 4.309$	$t(106) = 5.828$	$t(96) = 2.818$	$t(106) = -5.360$	$t(106) = 5.293$	$t(106) = -1.726$	$t(106) = -9.562$
p-value	$p = .924$	$p < .001$	$p < .001$	$p = .036$	$p < .001$	$p < .001$	$p = .261$	$p < .001$
Cohen's d	$d = 0.02$	$d = 0.88$	$d = 1.13$	$d = 0.58$	$d = -1.04$	$d = 1.02$	$d = -0.33$	$d = -1.85$
Homogeneity of variances *Levene's test:*	$F(5, 49) = 0.861, p = .514$		$F(5, 49) = 2.310, p = .058$		$F(5, 50) = 1.503, p = .206$		$F(5, 50) = 0.890, p = .495$	
Anova-like table Effect: $Game$	$F(1, 49) = 1.682, p = .201$		$F(1, 49) = 0.416, p = .522$		$F(1, 50) = 1.843, p = .181$		$F(1, 50) = 0.510, p = .479$	
$Group$	$F(2, 49) = 0.561, p = .574$		$F(2, 49) = 2.490, p = .093$		$F(2, 50) = 0.969, p = .386$		$F(2, 50) = 4.876, p = .012$	
$Game \times Group$	$F(2, 49) = 1.275, p = .289$		$F(2, 49) = 0.350, p = .706$		$F(2, 50) = 4.173, p = .021$		$F(2, 50) = 2.228, p = .118$	

of *response* choice was significant for MFN$_{latency}$ and LPC$_{amplitude}$. This finding extends the differences between the groups of fair and selfish participants observed for the latency of MFN and amplitude of LPC when all trials were pooled together.

It is interesting also to note that the interaction effects between behavior and response were statistically significant for P2 and MFN. We investigated this interaction following response choices (i.e., agreed *vs.* refused proposed allocations) recorded in the most representative trials for each group of participants, that were *Selfish trials* in selfish participants (GrpS) and *Fair trials* in fair participants (GrpF). Selfish participants who refused fair allocations during DG were characterized by longer P2$_{latency}$ ($t(61) = 9.134$, $p < .001$, $d = 2.70$) and greater P2$_{amplitude}$ ($t(61) = 16.727$, $p < .001$, $d = 4.95$) over refusal of selfish payoffs. This suggests that refusal of fair payoffs was a rewarding choice for selfish participants. In fair participants (GrpF), longer P2$_{latency}$ ($t(88) = 3.673$, $p < .01$, $d = 0.82$) was observed when they refused selfish allocations over fair ones. This suggests that refusal of selfish gains was a rewarding choice for fair participants. Accordingly, when fair participants refused selfish allocations (Figure 7A, left panel, dashed line),

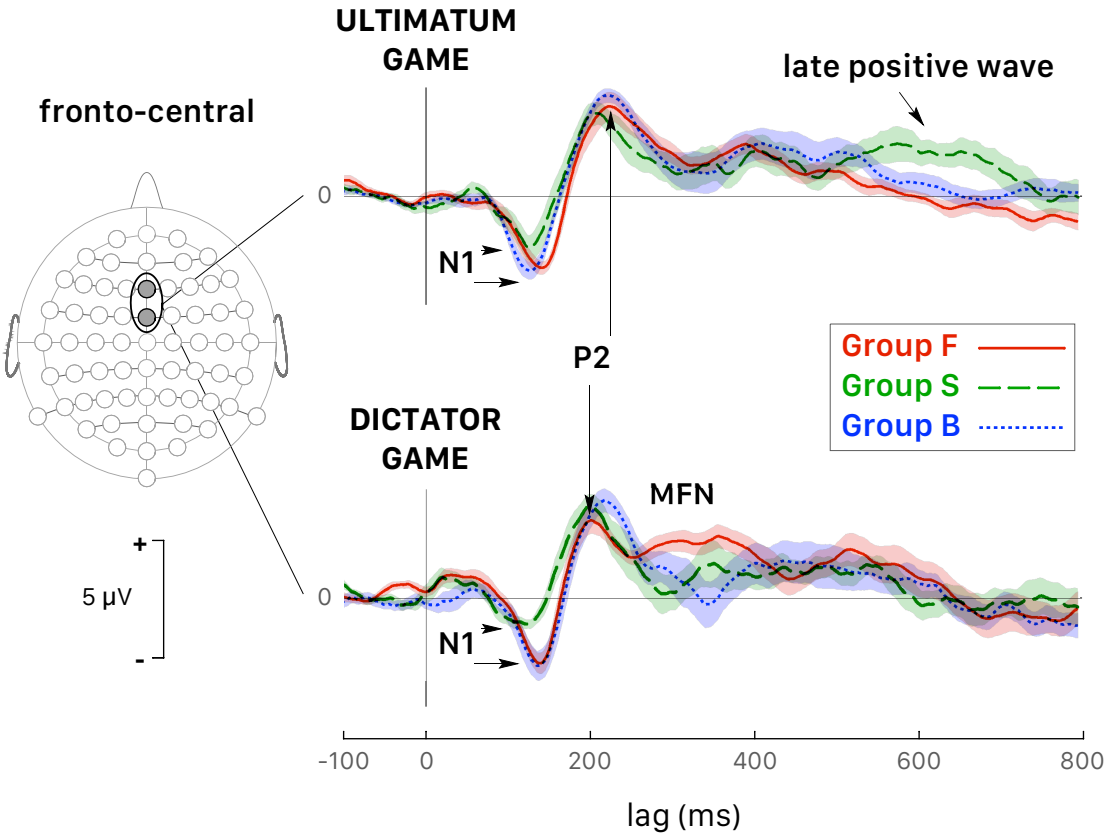

Figure 6: Grand average ERPs (*all trials pooled together*) recorded at fronto-central sites (Fz and FCz pooled together) during the Ultimatum Game (top panel) and the Dictator Game (bottom panel). Colored curves correspond to fair (GrpF, solid red curves), selfish (GrpS, dashed green curves) and less conceit (GrpB, dotted blue curves) participants with confidence intervals determined by the standard error of the mean (SEM). The arrows point out at several significant differences emphasized in the text. See also Table 3.

fronto-central $P2_{amplitude}$ was greater ($t(118) = 3.43$, $p < .01$, $d = 0.63$) than after acceptance of least favorable gain for themselves (Figure 7A, left panel, solid line).

In fair participants, the depth of MFN negativity was reduced compared to the other groups (see Table 3). In GrpF, $MFN_{latency}$ was shorter by approximately 60 ms in fair trials ($t(118) = -11.333$, $p < .001$, $d = -2.07$) when they accepted fair (bringing the least favorable gains) over refusal of selfish (with the most advantageous payoffs) allocations (Figure 7A, left panel). In the group formed by less conceit/more altruistic participants (GrpB), $MFN_{latency}$ was also shorter by 60 ms ($t(76) = -8.089$, $p < .001$, $d = -1.88$) and with greater negativity ($t(76) = 3.175$, $p < .05$, $d = 0.74$) after acceptance of the least advantageous allocations during fair trials. When these participants accepted the least advantageous allocations $MFN_{latency}$ was also shorter by approximately 70 ms ($t(88) = -7.567$, $p < .001$, $d = -1.69$) and with greater negativity ($t(88) = 5.029$, $p < .001$, $d = 1.12$) over acceptance of selfish (with the most advantageous payoffs) allocations. Then, it appears that it is rather the proposal of the fair and least advantageous allocations that evoked larger MFN in GrpB participants.

In these participants (GrpB, the less conceit/more altruistic), the amplitude of posterior-parietal LPC (Figure 7A, middle panel) was greater in agreed (solid line) than refused (dashed line) most advantageous (selfish) trials ($t(118) = 5.72$, $p < .001$, $d = 1.04$). When these participants refused selfish allocations the $LPC_{latency}$ was shorter by approximately 60 ms over refusal of fair allocations ($t(81.7) = -12.878$, $p < .001$, $d = -2.539$). Among fair Allocators, the most relevant finding was a $LPC_{latency}$ longer by approximately 40 ms after fair proposals, irrespective of the response. Overall, these findings emphasize that different dynamics distinguished the groups of participants based on behavioral responses. Fair participants were mainly characterized by ERP markers at fronto-central sites during the early stages of decision making in our DG task, while less conceit and more altruistic participants showed most salient markers in the mid-late components of the ERPs.

Table 4: **Dictator Game**. All participant played the role of Allocators. The table reports measurements of each peak's latency and amplitude (median, mean ±SEM values and number of measurements n) for trials independently sorted according to factors *response* and *behavior*. Statistics and significance are reported for Levene's test for heteroscedascity (homogeneity of variances test) and type III Anova-like table for a linear mixed model fit by maximum likelihood.

Dictator Game	**Allocator**					
	fronto-central P2 ($n = 476$)		fronto-central MFN ($n = 513$)		posterior-parietal LPC ($n = 507$)	
	Latency $[ms]$	Amplitude $[\mu V]$	Latency $[ms]$	Amplitude $[\mu V]$	Latency $[ms]$	Amplitude $[\mu V]$
Factor						
Response choice						
	$n = 227$	$n = 227$	$n = 252$	$n = 252$	$n = 243$	$n = 243$
Agreed	210.6	4.99	310.6	-1.13	540.3	5.64
	216.4±1.6	4.82±0.22	315.6±2.9	-1.24±0.35	544.7±3.6	4.81±0.24
	$n = 249$	$n = 249$	$n = 261$	$n = 261$	$n = 264$	$n = 264$
Refused	219.7	4.84	343.0	-0.52	535.4	3.92
	219.9±1.5	5.13±0.24	342.9±2.4	-1.05±0.28	544.3±2.6	3.53±0.19
Participants' behavior						
	$n = 221$	$n = 221$	$n = 237$	$n = 237$	$n = 228$	$n = 228$
Fair trials	213.3	5.37	330.4	-0.47	582.7	3.87
	221.2±1.8	5.54±0.25	327.2±2.8	-1.54±0.35	566.6±4.0	3.66±0.25
	$n = 255$	$n = 255$	$n = 276$	$n = 276$	$n = 279$	$n = 279$
Selfish trials	215.6	4.61	327.8	-1.05	530.9	4.84
	215.7±1.4	4.50±0.21	331.5±2.7	-0.80±0.28	526.4±1.6	4.54±0.18
Homogeneity of variances test						
Levene's test:	$F(3, 77) = 0.203$	$F(3, 77) = 0.205$	$F(3, 87) = 1.164$	$F(3, 87) = 0.656$	$F(3, 84) = 8.064$	$F(3, 84) = 1.093$
	$p = .894$	$p = .893$	$p = .328$	$p = .581$	$p < .001$	$p = .357$
Anova-like table						
Effect: *Response*	$F(1, 451) = 3.366$	$F(1, 455) = 1.424$	$F(1, 485) = 125.5$	$F(1, 488) = 1.585$	$F(1, 482) = 0.081$	$F(1, 481) = 23.72$
	$p = .067$	$p = .233$	$p < .001$	$p = .209$	$p = .776$	$p < .001$
Behavior	$F(1, 463) = 15.79$	$F(1, 471) = 10.40$	$F(1, 492) = 5.586$	$F(1, 499) = 16.19$	$F(1, 502) = 76.49$	$F(1, 497) = 6.018$
	$p < .001$	$p = .001$	$p = .018$	$p < .001$	$p < .001$	$p = .015$
Response × Behavior	$F(1, 447) = 6.605$	$F(1, 450) = 35.24$	$F(1, 485) = 120.6$	$F(1, 488) = 6.189$	$F(1, 482) = 2.973$	$F(1, 480) = 3.131$
	$p = .010$	$p < .001$	$p < .001$	$p = .013$	$p = .085$	$p = .077$

3.3.2 Ultimatum Game

We considered separately four sets of ERP recorded trials following the same categories used for the analysis of reaction times during UG: *Altruistic trials* when the Responder (i) 'accepted' the least favorable offers and (ii) 'rejected' equitable offers and received zero payoff; *Conceit trials* when the Responder (iii) 'accepted' the equitable offers (that are the most advantageous payoff in the current task design) and (iv) 'rejected' with the least favorable endowment shares and received zero payoff. All single trials of each set were averaged, pooled at fronto-central and posterior-parietal sites and analyzed following the same procedure used for the Dictator Game. We analyzed P2, MFN and LPC peaks, and the overall number of measurements (n) was 448, 518 and 514, respectively. Table 5 reports the latencies and amplitudes of major ERP peaks measured for the response type irrespective of the behavior and the values measured for the behavior irrespective of the response.

In the Ultimatum Game, the factors for the analysis with the linear mixed model with repeated measures were *response* (2 levels: Accepted, Rejected) and *behavior* (2 levels: Altruistic trials, Conceit trials). In the absence of main effects, the strong interaction effect observed for P2 suggested more detailed analysis. In selfish participants, it is interesting that P2$_{latency}$ was longer by approximately 35 ms ($t(64) = 5.841$, $p < .001$, $d = 1.44$) after acceptance of the least favorable outcomes (paying only 10% or 20% of the share to the participant) over acceptance of equitable shares (that granted 40% or 50% of the share). This comparison was weaker but held also for P2$_{amplitude}$, which was greater after accepting the least favorable outcomes ($t(53.7) = 3.204$, $p < .05$, $d = 0.80$). This suggests that even acceptance of the least favorable payoffs represents a rewarding choice for selfish participants, in agreement with the hypothesis that selfish participants tend to maximize their gain at all circumstances. GrpB participants (i.e., the least conceit/the most altruistic) showed also an interesting effect on P2. In this group, the latency of P2 was longer by approximately 40 ms ($t(70) = 4.610$, $p < .001$, $d = 1.46$) and P2 amplitude was greater ($t(12.5) = 3.422$, $p < .05$, $d = 1.25$) after acceptance of least favorable over equitable shares. This supports the rationale that accepting less favorable payoffs was a rewarding choice for altruistic participants.

For MFN, significant main effects of participant's *behavior* and *response* choice were observed on both peak latency and amplitude (Table 5). When selfish participants rejected the least favorable payoff over the acceptance of the equitable (most advantageous in this task design) offers (Figure 7B, left panel), MFN$_{latency}$ was shorter by 30 ms ($t(94) = -5.968$, $p < .001$, $d = -1.22$) and MFN negativity was greater (Welch's $t(93.2) = 4.003$, $p < .001$,

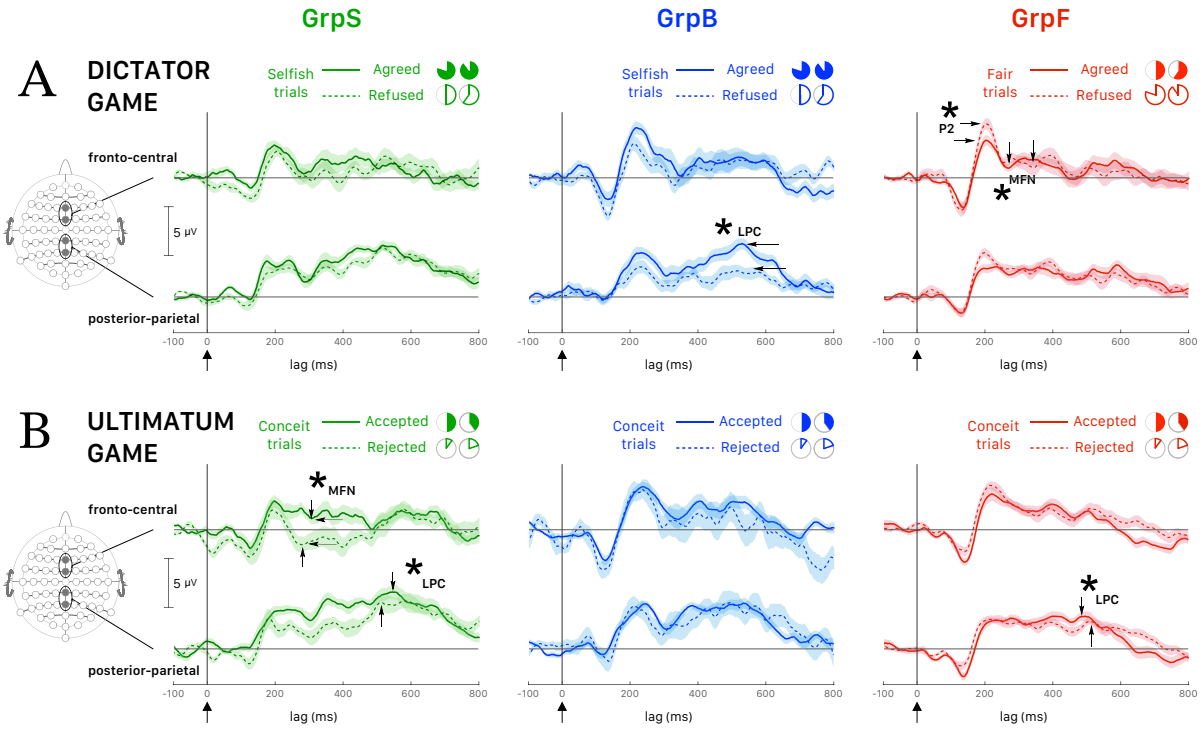

Figure 7: Averaged ERPs (and ±SEM bands) recorded at fronto-central sites (Fz, FCz) and posterior-parietal sites (Pz, CPz). **A.** Dictator Game: all participants played the role of Allocator/Proposer. The most representative trials for GrpS (left panel) and GrpB (middle panel) participants were *Selfish trials* and *Fair trials* for GrpF (right panel). Time zero corresponds to the presentation of the endowment share to the Allocator. Solid lines correspond to trials when the participant agreed with the proposal and dashed lines when the participant refused the proposal. **B.** Ultimatum Game: all participants played the role of Responder. The most representative trials for all groups of participants were *Conceit trials*. Time zero corresponds to the presentation of the endowment share to the Responder. Solid lines correspond to trials when the participant accepted the proposal and dashed lines when the participant rejected the proposal. Asterisks and labeled wave peaks indicate the most noticeable differences also commented in the text.

$d = 0.82$). In fair participants, after acceptance of least favorable over acceptance of equitable (most advantageous) payoffs, the latency of MFN was also shorter by approximately 30 ms ($t(94) = -6.684$, $p < .001$, $d = -1.41$) and MFN negativity was greater (Welch's $t(39.7) = 6.903$, $p < .001$, $d = 1.59$).

The similar pattern observed for MFN at fronto-central sites between fair and selfish participants did not held any more at posterior-parietal sites. The main effect of participant's behavior was very strong for latencies and amplitudes of LPC peak (Table 5) and the pattern for LPC was opposite between fair and selfish participants (Figure 7B). $\text{LPC}_{latency}$ during conceit trials, was shorter in selfish participants after rejection of least advantageous over acceptance of most advantageous (equitable) offers (Welch's $t(81.8) = -3.945$, $p < .001$, $d = 0.82$), while it was shorter in fair participants after acceptance of most advantageous payoffs over rejection of least advantageous offers (Welch's $t(104.7) = 4.266$, $p < .001$, $d = 0.71$).

During UG, fewer *Altruistic trials* were recorded than *Conceit trials* because participants in all groups expressed little altruism and more conceitedness. However, it is worth reporting some noticeable differences in LPC wave between the groups of participants during altruistic trials. In selfish participants, acceptance of most favorable payoffs corresponded to a shorter $\text{LPC}_{latency}$ by about 90 ms (Welch's $t(57.8) = 11.220$, $p < .001$, $d = 2.82$). On the opposite, in fair participants it is the rejection of most favorable payoffs that corresponded to a shorter $\text{LPC}_{latency}$ by about 130 ms (Welch's $t(51.84) = 18.891$, $p < .001$, $d = 5.12$). Participants showing less conceitedness and more altruism (GrpB) showed an intermediate pattern in comparison with selfish and fair participants. In this group, we observed shorter latency ($t(110.7) = -5.427$, $p < .001$, $d = -1.01$) and greater amplitude of LPC ($t(108.2) = 3.584$, $p < .001$, $d = 0.67$) when they accepted equitable (but also most advantageous) offers over acceptance of inequitable splits of the endowment. Overall, the analysis of ERPs during the Ultimatum Game showed selfish participants being

Table 5: **Ultimatum Game**. All participant played the role of Responders. The table reports measurements of each peak's latency and amplitude (median, mean ±SEM values and number of measurements n) for trials independently sorted according to factors *response* and *behavior*. Same legend of Table 4.

Ultimatum Game	**Responder**					
	fronto-central P2 ($n = 448$)		fronto-central MFN ($n = 518$)		posterior-parietal LPC ($n = 514$)	
	Latency $[ms]$	Amplitude $[\mu V]$	Latency $[ms]$	Amplitude $[\mu V]$	Latency $[ms]$	Amplitude $[\mu V]$
Factor						
Response choice						
	$n = 264$	$n = 264$	$n = 300$	$n = 300$	$n = 288$	$n = 288$
Accepted	223.4	5.28	322.3	-0.05	517.3	5.59
	226.3±1.5	5.26±0.24	318.6±1.9	-0.65±0.32	521.8±2.2	5.56±0.20
	$n = 184$	$n = 184$	$n = 218$	$n = 218$	$n = 226$	$n = 226$
Rejected	218.1	4.67	302.0	-0.72	513.6	5.41
	218.3±1.7	4.61±0.35	303.4±2.0	-1.31±0.42	516.0±3.5	5.55±0.35
Participants' behavior						
	$n = 158$	$n = 158$	$n = 188$	$n = 188$	$n = 184$	$n = 184$
Altruistic trials	223.0	5.48	289.8	-3.43	527.9	6.34
	222.9±2.4	5.67±0.49	296.7±2.3	-2.90±0.56	530.0±4.2	6.57±0.41
	$n = 290$	$n = 290$	$n = 330$	$n = 330$	$n = 330$	$n = 330$
Conceit trials	219.9	4.82	319.8	0.67	514.4	4.86
	223.1±1.2	4.63±0.16	321.0±1.6	0.19±0.22	513.3±1.9	4.99±0.18
Homogeneity of variances test						
Levene's test:	$F(3, 72) = 0.409$	$F(3, 72) = 6.421$	$F(3, 83) = 1.360$	$F(3, 83) = 7.078$	$F(3, 83) = 5.632$	$F(3, 83) = 4.075$
	$p = .747$	$p < .001$	$p = .261$	$p < .001$	$p < .01$	$p < .01$
Anova-like table						
Effect: *Response*	$F(1, 430) = 3.689$	$F(1, 441) = 1.738$	$F(1, 498) = 63.19$	$F(1, 499) = 5.624$	$F(1, 499) = 0.032$	$F(1, 499) = 5.896$
	$p = .055$	$p = .188$	$p < .001$	$p = .018$	$p = .859$	$p = .016$
Behavior	$F(1, 429) = 1.057$	$F(1, 439) = 5.864$	$F(1, 491) = 143.2$	$F(1, 492) = 59.10$	$F(1, 489) = 16.51$	$F(1, 489) = 36.77$
	$p = .304$	$p = .016$	$p < .001$	$p < .001$	$p < .001$	$p < .001$
Response × Behavior	$F(1, 430) = 13.80$	$F(1, 441) = 4.713$	$F(1, 497) = 2.605$	$F(1, 498) = 0.271$	$F(1, 501) = 0.282$	$F(1, 501) = 11.16$
	$p < .001$	$p = .030$	$p = .107$	$p = .603$	$p = .595$	$p = .001$

strongly characterized at the early stages of the decision making and with a pattern of brain activity opposite to fair participants at later stages.

3.3.3 Dimensional analysis of ERP peaks

The differences between the participants' groups observed in Tables 4,5 suggest that the ERP peaks might also be correlated with the behavioral indices, which defined the participants' groups. This analysis requires samples of data with behavioral ratings distributed throughout the values range, but the range of $UG_{altruism}$ was limited to values between -1 and 0 due to the bias towards conceitedness. The most significant correlations were observed with peak latencies, that are characterized by less variance than amplitudes. During fair trials of DG, fronto-central P2 peak latency correlated linearly with $DG_{selfishness}$ (Pearson correlation coefficient $r = 0.63$, 95% CI $[0.38, 0.79]$, $t(36) = 4.778$, $p < .001$) when the participant agreed or refused the allocation (Figure 8A). Irrespective of the response choice, the slopes of both regression curves were similar and the smaller the selfishness, the shorter the latency of P2 peak ($F(1, 40) = 17.45$, $p < .001$, $R^2 = .304$, $R^2_{Adjusted} = .286$). This results should be viewed taking into account the correlation between RTs and selfishness during DG (i.e., the smaller the selfishness, the shorter the RTs, Figure 4B).

A positive correlation was also observed during altruistic trials of UG (Figure 8B) between the values of selfishness and the posterior-parietal LPC peak latency ($r = 0.92$, 95% CI $[0.69, 0.98]$, $t(8) = 6.657$, $p < .001$) when the participants rejected the proposed payout ($F(1, 8) = 44.31$, $p < .001$, $R^2 = .847$, $R^2_{Adjusted} = .828$). The only ERP wave that correlated well with altruism is the fronto-central MFN (Figure 8C,D) when the participants agreed with the proposed allocation. Interestingly, the smaller the altruism, the shorter the latency of MFN peak ($r = 0.67$, 95% CI $[0.36, 0.84]$, $t(22) = 4.186$, $p < .001$) during selfish trials ($F(1, 22) = 17.52$, $p < .001$, $R^2 = .443$, $R^2_{Adjusted} = .418$) and the larger the altruism, the deeper the amplitude of MFN peak ($r = -0.58$, 95% CI $[-0.82, -0.16]$, $t(16) = -2.860$, $p < .001$ during fair trials ($F(1, 16) = 8.180$, $p = .011$, $R^2 = .338$, $R^2_{Adjusted} = .297$).

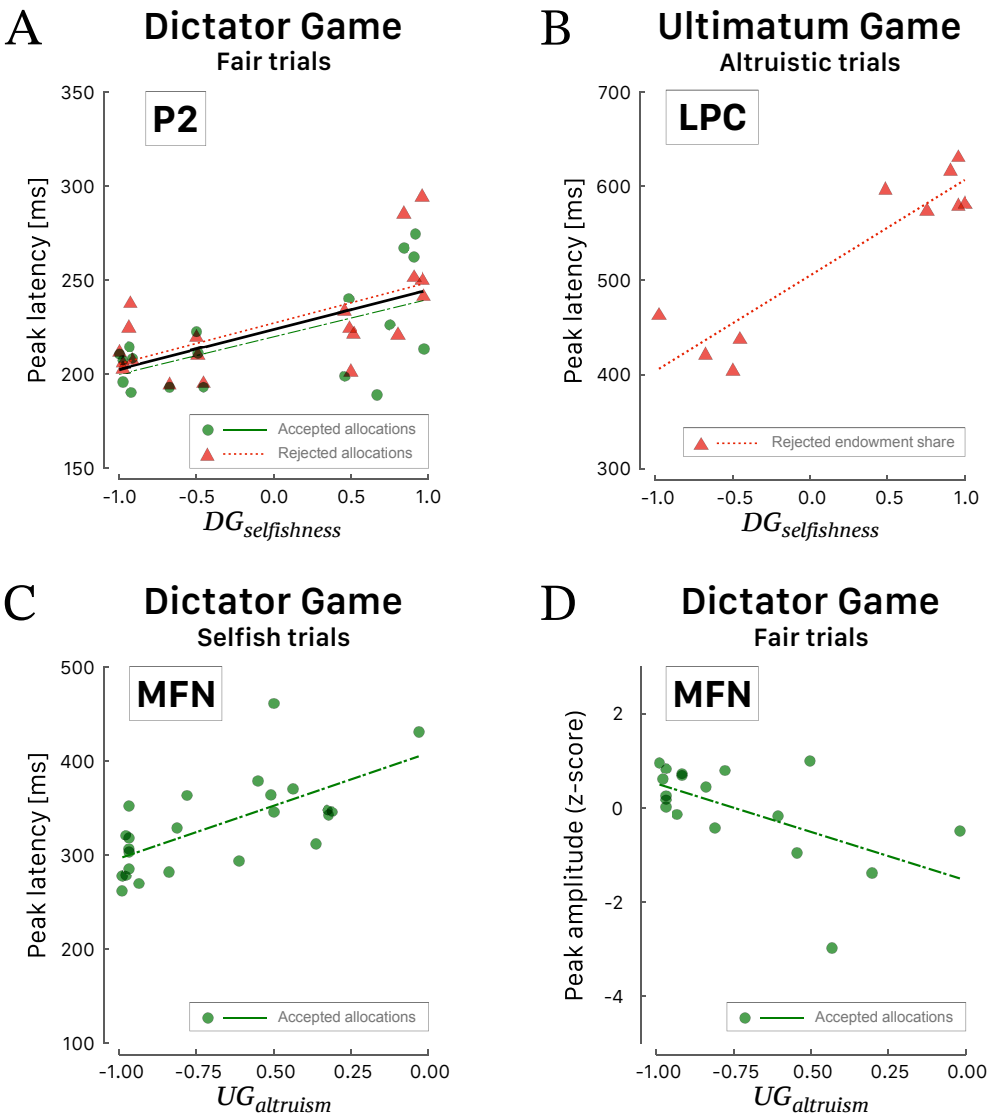

Figure 8: Scattergrams of significant correlations between ERP peaks and behavioral indices. **A.** Latency of P2 peak *vs.* $DG_{selfishness}$ during fair trials of DG. The black curve corresponds to the linear fit irrespective of the response choice. **B.** Latency of LPC peak *vs.* $DG_{selfishness}$ during altruistic trials of UG. **C.** Latency of MFN peak *vs.* $UG_{altruism}$ during selfish trials of DG. **D.** Amplitude of MFN peak *vs.* $UG_{altruism}$ during fair trials of DG. Green dots correspond to participants' peak latencies when they agreed with the proposed allocation and red triangles when they refused the allocation. Green (long dashed) and red curves correspond to linear fitted regressions.

4 Discussion

Human life is characterized by situations where the maximization of the individual payoff in face-to-face negotiations is challenged by the socio-cultural environment that contributed to shape moral phenomena such as altruism and fairness (Altman, 2005) and by the social distance to the other party, which affects the subjective perception of justice (Yu et al., 2015). In the context of an increasing number of contactless monetary transactions, our study aimed to test the hypothesis whether the characteristic patterns of brain activity—recorded by ERPs—were somatic markers associated with the behavioral profiles of individuals determined by a combination of Dictator and Ultimatum Game performances in the absence of effective human interactions between the parties. The analysis of the major ERP wave components showed that our observations were in agreement with several findings reported in the literature, which were mainly focused on only one of those classical neuroeconomic games (Nagel, 2001; De Martino et al., 2006; Fellner et al., 2009;

Güth, 2010; De Neys et al., 2011; Fiori et al., 2013). The identification of two homogeneous groups of participants, corresponding to selfish and fair participants, allowed to report more precisely the activity patterns associated with clearly modeled behaviors characterized by spiteful and costly punishments, which are respectively positively and negatively correlated with impulsive choices (Sanfey et al., 2003; Rilling and Sanfey, 2011; Rodrigues et al., 2018).

4.1 Behavioral results

Distinct groups of individuals were identified following a hierarchical cluster analysis based on indices associated with conceitedness/altruism and selfishness/fairness determined by performance during Ultimatum and Dictator Game. The definitions of conceitedness/altruism and selfishness/fairness are operational and we do not pretend that they necessarily overlap with semantic definitions that are culturally biased. The most representative and homogeneous groups were formed by fair participants (GrpF, individuals expressing a conceit behavior during UG and with a low level of selfishness during DG), by individuals characterized by less conceitedness and medium-high selfishness (GrpB), and selfish participants (GrpS, individuals with highest values of selfishness and lowest levels of altruism). The validity of the categorization of these participants' groups is further confirmed by the observation that reaction times and, even more importantly, several characteristics of the ERP peaks —i.e., the somatic markers— were correlated with the behavioral indices. Hence, these indices are likely to reflect what they claim to reflect but they should not be extended beyond our experimental protocol without careful consideration.

In the UG, it is rationale to expect that the Proposer offers the smallest possible amount and the Responder accepts any offer. In this study, we could not fully evaluate a true altruistic attitude in the behavioral performance. All participants played the role of Responder and expressed conceitedness to some extent, because the most favorable payoffs for the Responder were equitable offers (i.e., with 40% or 50% of the share for the participant himself). The rejection of an unfair offer by the Responder in the UG is often explained by a bias towards profit maximization in association with positive social factors, such as common ethical principles and friendship, by negative factors, such as fear of the perceived consequences of having one's offer rejected, and from the sense of guilt linked to worries about the opponent's outcome (Carver and Miller, 2006; Marchetti et al., 2011; Gaertig et al., 2012). To test a Responder's attitude from selfless/altruistic to truly greedy, even in the absence of a face-to-face interaction, UG's task design would have to include endowment shares paying 60%–70% of the amount in favor of the Responder, as well as the most inequitable endowment shares (i.e., splits that provide 80%–90% of the amount to the Responder). We are aware of the limitations of the task design chosen in this study, which was driven by a compromise between the total duration of the experiment and the number of test repetitions that are necessary to get meaningful ERPs for both neuroeconomic games.

During DG, all participants played the role of of Allocator, who is the player who imposes an endowment share on the other party, and the behavioral performance distinguished two opposite groups, selfish (GrpS) and fair (GrpF) participants. However, during UG all participants played the role of Responder and both groups GrpS and GrpF tended to reject unequal offers (Table1). The rejection of the least favorable offers by the participants (i.e., payoffs of 10%-20% of the amount at stake) is considered to be the expression of a punishment towards a selfish Proposer (Forber and Smead, 2014; Ma and Hu, 2015) and was interpreted either as a spiteful punishment (Marlowe et al., 2011) or as costly —rather than altruistic— punishment (Brethel-Haurwitz et al., 2016; Yamagishi et al., 2017). The differences between the groups that during UG, when Responders have a passive position expecting to receive fair offers (Polezzi et al., 2008; Boksem and De Cremer, 2010), selfish participants (GrpS) were driven by spiteful punishment and fair participants (GrpF) by costly punishment. GrpA and GrpB participants also expressed a tendency towards selfishness, thus suggesting that spiteful behavior was predominant. This is not surprising because a single altruistic decision is likely to lack any reward for an individual, unless its repetition over time proves it is truly beneficial. The time frame of our experiment is very limited and the feature of learned behavior is in contrast to the spiteful punishment exercised by selfish participants who feel emotionally aroused when exposed to an unfair proposal even in a short lasting experimental context.

4.2 Socio-cultural dimension

Spiteful punishment is a kind of anti-social punishment applied in various ways that reflect the psychological and social dimensions of regional-cultural values (Chuah et al., 2007; Sylwester et al., 2013). It is recognized that the evolution of fairness is promoted by spitefulness and inhibited by altruism (Fehr and Gächter, 2002; Fehr and Rockenbach, 2004; Zhang and Fu, 2018). Behavioral differences have been observed in association with cross-country differences in socio-economic status and respect for authority (Inglehart, 2000; Oosterbeek et al., 2004; Chuah et al., 2007). High levels of cooperation and altruistic behavior can be observed in societies where people face economic disparities and poor social standing, and where they are exposed to punitive behavior if they violate certain norms and are treated as inappropriate (Fehr and Gächter, 2002; Nowak, 2006; Egas and Riedl, 2008).

In the present study, the socio-cultural dimension of our sample of participants is homogenous and made up of young Iranian males with an academic education. The choice to test participants without including both genders avoided the introduction of a gender effect discussed elsewhere in the literature (Eckel and Grossman, 2001; Servátka, 2009; García-Gallego et al., 2012; Li et al., 2018). Despite the evidence of altruistic tendencies in the Iranian people, such as willingness to pay for health services (Javan-Noughabi et al., 2017), organ donation (Abbasi et al., 2018), blood donation (Javadzadeh Shahshahani et al., 2006), moral sensitivity of nurses (Borhani et al., 2017; Amiri et al., 2020), the identification of the mechanisms associated with the perception of fairness has never been carefully investigated. Our results show that despite spitefulness appeared to prevail over fairness, we have observed a sizable sample of participants expressing costly/altruistic punishment and a tendency towards less conceit, somehow altruistic, behavior. It has also been established that in many experimental neuroeconomic contexts the game appears to reflect common interactional patterns of daily life, with greater behavioral variability between social groups than has been found among societies (Henrich et al., 2005), and the multiple play of UG does not affect rates of rejection (Oosterbeek et al., 2004). Therefore, we are confident that the participants included in this study did not introduce any strong behavioral bias towards an extreme behavioral strategy due to the homogeneous socio-demographic composition of our group of participants.

4.3 Fronto-central N1

In ERP studies, the latency and amplitude of a cortical wave component vary with the processing speed and the amount of cognitive resources required to reach the necessary degree of information processing for completion a certain mental task (Fabiani et al., 2000). The earliest ERP wave component that we have analyzed is the N1 negative component, much larger at fronto-central sites, in the 110–160 ms time window after trigger onset. It is important to distinguish the frontal N1 discussed here from the well known and described N1 component at occipital and temporal scalp sites associated with visual-spatial selective attention (Kutas et al., 1994; Mangun, 1995). The fronto-central N1 wave component is expected to be associated with a visual processing of a functional representation of the stimulus (Antal et al., 2000; Proverbio et al., 2007) and when participants expected a stimulus for decision making on whether to act or not to act (Di Russo et al., 2019). Such top-down prefrontal control of decision making (Fuster, 2015) suggests that N1 reflects a proactive cognitive control to access the procedural knowledge necessary to enable an active motor response (Petit et al., 2006; Proverbio et al., 2007; Di Russo et al., 2019).

We observed that fronto-central N1 did not vary much between the games, but it was strongly associated with the group of participants. The greater the selfishness, the shorter the latency of N1 and in the group of selfish participants this component peaked about 10 ms earlier than fair participants and with smaller amplitude. In our tasks, the visual cues carry a cognitive value associated with the costs and benefits determined by the behavioral/personality traits of each individual. Hence, we can interpret the N1 difference between fair and selfish participants as the first evidence in ERPs between participants who have otherwise expressed costly and spiteful forms of punishment. Longer N1 latencies and larger amplitudes in fair participants are compatible with a top-down control of a rather rich and less stereotyped set of initial activity patterns. Our results might also be interpreted in agreement with an anterior N1 enhancement assumed to reflect the cost of a modality change and readjustment of attentional weight-setting in order to optimize target detection (Töllner et al., 2009; Hsieh and Wu, 2011).

4.4 Fronto-central P2

The P2 wave peaked in the range of 180 to 260 ms and was most salient in the fronto-central area, where a top-down process using previous experience and expectations is indicative of stimulus evaluation rather than response production (Potts, 2004; Rauss and Pourtois, 2013; Fuster, 2015). Increased fronto-central P2 is expected in response to rewarding contextual outcomes, in participants who feel gratification following the amount of cognitive resources activated by a reward signal in the brain, which may include satisfaction when observing the opponent's misfortune (Falco et al., 2019; Weiss et al., 2020). In previous decision-making studies, it was also suggested that P2 is modulated by the degree of predictability of the outcomes (Polezzi et al., 2008) and by the social context (Hu et al., 2014; Liu and Huang, 2015). Our results are in agreement with the rationale that refusal of fair payoffs is a rewarding choice for selfish participants. and refusal of selfish payoffs is a rewarding choice for fair participants. These findings support the hypothesis that in DG a selfish player required more cognitive resources when he refused to allocate a fair share and a fair player required more cognitive resources when he agreed to assign a selfish profit. In UG, the analysis of P2 wave suggested that selfish participants tend to maximize their gain with greater and longer P2 even after acceptance of less favorable payoffs. Moreover, we found evidence that accepting less favorable payoffs was a rewarding choice for the group of more altruistic participants (or less conceit, GrpB).

4.5 Medial Frontal Negativity

The MFN is a wave component that has been observed in association with some assessment of fairness (Botvinick et al., 1999; Gehring and Willoughby, 2002; Alexopoulos et al., 2012; Li et al., 2020), social status and altruistic bias (Van der Veen and Sahibdin, 2011; Sun et al., 2015; Cui et al., 2019; Mayer et al., 2019), but only one study has reported quantitative ERP results in Allocators during DG (Li et al., 2020). However, these studies did not clearly distinguish between behavioral groups. In our study, we observed that in both games the negativity of MFN tended to be greater when the latency was shorter. In general, the greater the selfishness, the deeper the negativity of MFN. Being associated with disadvantageous inequitable offers and with altruistic tendency, MFN is expected to be larger in altruistic participants. This was indeed our observation for GrpB participants playing the Dictator Game, who expressed less conceit and more altruistic behavior. This observation was further confirmed by the correlation of MFN amplitude and the index $UG_{altruism}$ in the fair trials of DG when the participants playing the role of Allocator accepted the proposed amount.

In fair participants, the latency of MFN was shorter during DG after accepting to allocate a fair amount to the other player than after rejection of a selfish allocation, and during UG the latency of MFN was longer after acceptance of most advantageous over acceptance of least favorable splits of the endowment share. In selfish participants, spiteful punishment was revealed by shorter latency and reduced MFN negativity after rejection of the least favorable payoffs over acceptance of most advantageous offers. These results support the hypothesis that in UG the MFN amplitude of Responders is likely to be associated with an emotional reaction to the expectation of fairness by the Proposers. Thus, the perception of the offers by the Responders —all participants of this study played this role— depended on the behavioral group, and not on the absolute value of the endowment share (Van der Veen and Sahibdin, 2011; Mayer et al., 2019; Spapé et al., 2019; Li et al., 2020). In general, our results may be interpreted in accordance with the hypothesis of a short latency of MFN as a sign of an emotional response, which required little cognitive evaluation and expressed reckless mental processing (Brookhuis et al., 1981; Kok, 1997; Codispoti et al., 2007; Fukushima and Hiraki, 2009; Boksem et al., 2011). This is also in agreement with the presumed generation of MFN in the anterior cingular cortex in association with emotional regulation (Etkin et al., 2011; Stevens et al., 2011; Kanske and Kotz, 2011; García-Cabezas and Barbas, 2017).

4.6 Posterior-parietal late positive component

Evidence reported in the literature indicates that LPC-associated activity is likely to be generated by a widespread network of cerebral sources including the frontal lobe, the cingulate cortex and the parietal regions involved in attention control (Sabatinelli et al., 2007; McDonald and Green, 2008; Scharmüller et al., 2011; Sun et al., 2017). LPC is closely related to relevant motivational stimuli, to the subsequent sustained allocation of attention (Ito et al., 1998; Wu et al., 2011b; Bamford et al., 2015), and to the saliency of an outcome also in social comparisons (Sun et al., 2020; Zhang et al., 2021; Sharif et al., 2021). It could be hypothesized that a large LPC wave should characterize selfish Allocators playing the Dictator Game. In both games, we have indeed observed greater LPC (and at shorter latency in UG) in selfish over fair participants. Selfish and fair participants were characterized by quiet opposite patterns of activity recorded at the posterior-parietal area. Recent evidence provided by the analysis of a modified Dictator Game exploring the role of players' social status (Cui et al., 2019) revealed that LPC amplitude was more sensitive to the interaction of the recipient's status and fairness— i.e., with a tendency of greater amplitude triggered by unfair endowment shares for a high-status Recipient and greater amplitude triggered by fair splits for a medium-status Recipient. It is interesting to note that selfish participants were also characterized during UG by a sustained late positive wave at the unusual fronto-central sites. This observation suggests that in these participants a potential source probably not located at the cortical level, but deeper in the brain, might contribute to a late positive wave visible at fronto-central as well as at posterior-parietal areas.

It is rationale that altruistic players expect equitable shares of the endowment during the Ultimatum Game. Being associated with motivation and subsequent allocation of attention, the amplitude of posterior-parietal LPC is expected to be grater in altruistic Responders playing the Ultimatum Game. In agreement with this hypothesis, we observed greater amplitude of LPC in GrpB participants when they accepted equitable over inequitable splits of the sum at stake. Moreover, during DG the latency of LPC was shorter in GrpB Allocators when they refused more advantageous allocations over refusal of fair allocations. Overall, the differences between selfish and fair participants support the hypothesis that LPC at posterior-parietal sites reflects a processing stage that evaluates norms of fairness (Wu et al., 2011b) in association with saliency interacting with context (Minnix et al., 2013; Stolz et al., 2019; Sun et al., 2020; Bauer et al., 2021).

4.7 Limitations

A critical limitation of our study is common to most neuroeconomic studies, that is whether studies based on experimental protocols in laboratory settings can really give insights in human brain activity outside real-world situations (Clithero et al., 2008; Huettel and Kranton, 2012; Konovalov and Krajbich, 2019). There is no doubt that interaction between real people plays a key role in business relationships, which are based on a much more complex set of contextual signals with an individual realm of emotions and feelings that help shape individual personalities. In our protocol, the participants were informed that they were playing against a human party located elsewhere, without specifying whether the other party was the experimenter or other participants. This is a workaround, but it is important that the participants were not simply playing for nothing because they ultimately received a real amount of money that depended on the performance accumulated at the games. Future studies should also require the participants to fill questionnaires defining personality traits (e.g., HEXACO; Ashton et al., 2014) but a validated Persian version of this inventory scale (Basharpoor et al., 2019) was not yet avaibale at the time our study was run and completed.

The semantic definition of benefit in the economic context set by neuroeconomic games refers to a rational choice aimed at achieving the optimal maximization of gain (Bernheim, 2009). The definitions of altruism, selfishness and fairness are very subjective, but we have defined the sets of conceit/altruistic and selfish/fair trials within a precise operational context defined by our experimental protocol, based on the same participant playing both games. The observation of irrational choices, i.e. low levels of altruism and opposing levels of selfishness (GrpF and GrpS), indicate that one's own assessment of the coherence of actions can lead to beneficial outcomes that are biased compared to a rational evaluation of monetary payment. To this respect, another limitation of the current study is the lack of neurophenomenological interview (Bockelman et al., 2013; Høffding et al., 2021; Klinke and Fernandez, 2022) in order to understand the subject's own perception of his/her spiteful or costly punishment.

Last but not least, we recognize methodological limitations associated with the sample size when we consider that the number of trial repetitions needed for averaging ERPs was limited by the duration of one session and the unbalanced number per set determined by individual strategies that could not be foreseen before the experiment. We are aware that peak latencies and amplitudes reported in our study were sensitive to high-frequency noise. The signal-to-noise level is low in our study due to averaging over a small amount of trials and this brings some methodological considerations. The limits of the time windows of an ERP wave are extremely difficult to set firmly for all participants because of the individual differences in the adopted strategy and may vary between participants across the various sets of trials (selfish trials: accepted vs. rejected, fair trials: accepted/rejected). These time limits could be considered only after a careful post-hoc analysis, which is contradictory with the measurement of the mean amplitude (instead of peak amplitude) and fractional area latency (instead of peak latency) that are usually considered more meaningful values in itself (Fabiani et al., 2007; Woodman, 2010; Luck, 2014). Our aim was not the establishment of standardized ERP values for the experimental protocols of neuroeconomic games, but to determine if somatic markers could be reliably observed in participants who clearly adopted different strategies. In the absence of a large number of trials at hand to overcome the difficulties in measuring peak values with a jackknife-based approach (Kiesel et al., 2008), we have adopted the method of repeated measurements taken by several observers. In future studies, we believe that the newly presented standardized measurement error for amplitude and latency values carries the potential for greater replicability of ERP studies in neuroeconomic experimental protocols, making it possible to identify participants with low-quality data and noisy channels (Luck et al., 2021).

4.8 Conclusion

This study has presented a novel combination of Dictator and Ultimatum Game with the aim of investigating whether ERP markers support the hypothesis that spiteful and costly punishments can be differentiated. We have presented evidence that different groups of participants can be identified on the basis of their performance at the combination of DG and UG. Two groups of participants, both with little tendency towards altruism, were characterized by a selfish or fair attitude during Dictator Game. An early ERP negative component (N1), peaking in the interval of 110 to 160 ms at fronto-central sites, distinguished fair participants from selfish ones. Analysis of middle frontal negativity (MFN) suggested that selfish participants were likely to exhibit spiteful punishment positively related to impulsive choice. The differences in late positive component (LPC) observed between the groups could be interpreted as that costly punishment behavior, enabled by fair participants, must be stable to make sense and requires less attention than selfish participants, whose cognitive-behavioral contradiction requires more attention and activation of a larger brain network to complete the information processing necessary to make a decision. In conclusion, this study brings an original contribution and new evidence to the existence of different circuits activated by the evaluation of fair and unfair proposals in participants characterized by different expressions of perception of fairness.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

Conceptualization, AMM, HP, MAM, RK; formal analysis, HP, MAM, AL; funding acquisition, HP, AL; clinical investigation, MAM; methodology, AMM, HP, AL; project administration, MAM, HP, AL; supervision, HP, AL, AEPV; validation, HP, AL, AEPV, AL; visualization, AEPV, AL; writing original draft, AMM, AEPV, AL; writing-review and editing, AEPV, AL.

Funding

This research was partially supported by the Swiss National Science Foundation grant n. IZSEZ0_183401 to AEPV.

Acknowledgments

The authors extend their gratitude and acknowledgements to all study participants and study team members for their time and energy spent on this project.

References

Abbasi, M., Kiani, M., Ahmadi, M., and Salehi, B. (2018). Knowledge and Ethical Issues in Organ Transplantation and Organ Donation: Perspectives from Iranian Health Personnel. *Ann Transplant*, 23:292–299.

Alexopoulos, J., Pfabigan, D. M., Lamm, C., Bauer, H., and Fischmeister, F. P. (2012). Do we care about the powerless third? An ERP study of the three-person ultimatum game. *Front Hum Neurosci*, 6:e00059.

Altman, M. (2005). The ethical economy and competitive markets: Reconciling altruistic, moralistic, and ethical behavior with the rational economic agent and competitive markets. *J. Econ. Psychol.*, 26(5):732–757.

Amiri, E., Ebrahimi, H., Namdar Areshtanab, H., Vahidi, M., and Asghari Jafarabadi, M. (2020). The Relationship between Nurses' Moral Sensitivity and Patients' Satisfaction with the Care Received in the Medical Wards. *J Caring Sci*, 9(2):98–103.

Antal, A., Kéri, S., Kovács, G., Janka, Z., and Benedek, G. (2000). Early and late components of visual categorization: an event-related potential study. *Brain Res Cogn Brain Res*, 9(1):117–9.

Artinger, F., Exadaktylos, F., Koppel, H., and Sääksvuori, L. (2010). Fairness in simple bargaining games: the role of empathy and theory of mind. *Jena Economic Research Paper*, 2010-037.

Ashton, M. C., Lee, K., and de Vries, R. E. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality factors: a review of research and theory. *Pers Soc Psychol Rev*, 18(2):139–152.

Balafoutas, L., Nikiforakis, N., and Rockenbach, B. (2016). Altruistic punishment does not increase with the severity of norm violations in the field. *Nat Commun*, 7:13327.

Bamford, S., Broyd, S. J., Benikos, N., Ward, R., Wiersema, J. R., and Sonuga-Barke, E. (2015). The late positive potential: a neural marker of the regulation of emotion-based approach-avoidance actions? *Biol Psychol*, 105:115–123.

Basharpoor, S., Taherifard, M., and Mohammadi, G. (2019). Psychometric Properties of Persian Version of Brief HEXACO Inventory in University Students. *Training Measurement*, 9(36):65–89.

Bauer, E. A., Thomas, G., and MacNamara, A. (2021). The controllability of aversive and neutral pictures: an event-related potential study. *Behav Brain Res*, 408:113265.

Bear, A. and Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proc Natl Acad Sci U S A*, 113(4):936–41.

Bechara, A. and Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games Econ Behav*, 52(2):336 – 372. Special Issue on Neuroeconomics.

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1-3):7–15.

Bernheim, B. D. (2009). On the Potential of Neuroeconomics: A Critical (but Hopeful) Appraisal. *AEJ: Microeconomics*, 1(2):1–41.

Bockelman, P., Reinerman-Jones, L., and Gallagher, S. (2013). Methodological lessons in neurophenomenology: Review of a baseline study and recommendations for research approaches. *Front Hum Neurosci*, 7:608.

Boksem, M. A., Kostermans, E., and De Cremer, D. (2011). Failing where others have succeeded: Medial frontal negativity tracks failure in a social context. *Psychophysiology*, 48(7):973–979.

Boksem, M. A. S. and De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, 5(1):118–125.

Borhani, F., Abbaszadeh, A., Mohamadi, E., Ghasemi, E., and Hoseinabad-Farahani, M. J. (2017). Moral sensitivity and moral distress in iranian critical care nurses. *Nurs Ethics*, 24(4):474–482.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., and Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, 402(6758):179–181.

Brañas-Garza, P., Espín, A. M., Exadaktylos, F., and Herrmann, B. (2014). Fair and unfair punishers coexist in the ultimatum game. *Sci Rep*, 4:6025–6025.

Brethel-Haurwitz, K. M., Stoycos, S. A., Cardinale, E. M., Huebner, B., and Marsh, A. A. (2016). Is costly punishment altruistic? Exploring rejection of unfair offers in the Ultimatum Game in real-world altruists. *Sci Rep*, 6:18974.

Brookhuis, K., Mulder, G., Mulder, L., Gloerich, A., Van Dellen, H., Van Der Meere, J., and Ellermann, H. (1981). Late positive components and stimulus evaluation time. *Biol Psychol*, 13:107–123.

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, Princeton, NJ.

Carver, C. S. and Miller, C. J. (2006). Relations of serotonin function to personality: Current views and a key methodological issue. *Psychiatry Res*, 144(1):1–15.

Chen, W., Wu, T., Li, Z., and Wang, L. (2019). Evolution of fairness in the mixture of the Ultimatum Game and the Dictator Game. *Physica A: Statistical Mechanics and its Applications*, 519:319–325.

Chuah, S.-H., Hoffmann, R., Jones, M., and Williams, G. (2007). Do cultures clash? Evidence from cross-national Ultimatum Game experiments. *J Econ Behav Organ*, 64(1):35–48.

Clithero, J. A., Tankersley, D., and Huettel, S. A. (2008). Foundations of Neuroeconomics: From Philosophy to Practice. *PLoS Biol*, 6(11):e298.

Codispoti, M., Ferrari, V., and Bradley, M. M. (2007). Repetition and event-related potentials: distinguishing early and late processes in affective picture perception. *J Cogn Neurosci*, 19(4):577–86.

Cui, F., Wang, C., Cao, Q., and Jiao, C. (2019). Social hierarchies in third-party punishment: A behavioral and ERP study. *Biol Psychol*, 146:107722.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, 313(5787):684–7.

De Neys, W., Novitskiy, N., Geeraerts, L., Ramautar, J., and Wagemans, J. (2011). Cognitive control and individual differences in economic ultimatum decision-making. *PLoS One*, 6(11):e27107.

Delorme, A. and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*, 134(1):9–21.

Di Russo, F., Berchicci, M., Bianco, V., Perri, R. L., Pitzalis, S., Quinzi, F., and Spinelli, D. (2019). Normative event-related potentials from sensory and cognitive tasks reveal occipital and frontal activities prior and following visual events. *Neuroimage*, 196:173–187.

Du, E. and Chang, S. W. C. (2015). Neural components of altruistic punishment. *Front Neurosci*, 9:26.

Eckel, C. C. and Grossman, P. J. (2001). Are Women Less Selfish Than Men? Evidence From Dictator Experiments. *The Economic Journal*, 108(448):726–735.

Egas, M. and Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proc Biol Sci*, 275(1637):871–8.

Etkin, A., Egner, T., and Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci*, 15(2):85–93.

Fabiani, M., Gratton, G., and Coles, M. G. H. (2000). Event-related brain potentials: Methods, theory and application. In *Handbook of psychophysiology*, pages 53–84. Cambridge University Press.

Fabiani, M., Gratton, G., and Federmeier, K. D. (2007). Event-related brain potentials : Methods, theory and application. In Cacioppo, J. T., Tassinary, L. G., and Berntson, G., editors, *Handbook of Psychophysiology*, chapter 4, pages 85–119. Cambridge University Press.

Falco, A., Albinet, C., Rattat, A.-C., Paul, I., and Fabre, E. (2019). Being the chosen one: social inclusion modulates decisions in the ultimatum game. An ERP study. *Soc Cogn Affect Neurosci*, 14(2):141–149.

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*, 39(2):175–91.

Fehr, E. and Fischbacher, U. (2004). Third-party punishment and social norms. *Evol Human Behav*, 25(2):63–87.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–40.

Fehr, E. and Rockenbach, B. (2004). Human altruism: economic, neural, and evolutionary perspectives. *Curr Opin Neurobiol*, 14(6):784–90.

Fehr, E. and Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868.

Fellner, G., Güth, W., and Maciejovsky, B. (2009). Satisficing in financial decision making -a theoretical and experimental approach to bounded rationality. *J Math Psychol*, 53(1):26–33.

Fiori, M., Lintas, A., Mesrobian, S., and Villa, A. E. P. (2013). Effect of Emotion and Personality on Deviation from Purely Rational Decision-Making. In Guy, V. T., Kárný, M., and Wolpert, D., editors, *Decision Making and Imperfection*, pages 129–161. Springer, Berlin, Heidelberg.

Forber, P. and Smead, R. (2014). The evolution of fairness through spite. *Proc Biol Sci*, 281(1780):20132439.

Fowler, J. H. (2005). Altruistic punishment and the origin of cooperation. *Proc Natl Acad Sci U S A*, 102(19):7047–9.

Fukushima, H. and Hiraki, K. (2009). Whose loss is it? Human electrophysiological correlates of non-self reward processing. *Soc Neurosci*, 4(3):261–75.

Fuster, J. M. (2015). *The Prefrontal Cortex*. 460 pages. Academic Press, 5th edition.

Gaertig, C., Moser, A., Alguacil, S., and Ruz, M. (2012). Social information and economic decision-making in the ultimatum game. *Front Neurosci*, 6:103–103.

García-Cabezas, M. Á. and Barbas, H. (2017). Anterior Cingulate Pathways May Affect Emotions Through Orbitofrontal Cortex. *Cereb Cortex*, 27(10):4891–4910.

García-Gallego, A., Georgantzís, N., and Jaramillo-Gutiérrez, A. (2012). Gender differences in ultimatum games: Despite rather than due to risk attitudes. *J Econ Behav Organ*, 83(1):42–49.

Gehring, W. J. and Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 295(5563):2279–2282.

Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2003). Explaining altruistic behavior in humans. *Evol Human Behav*, 24(3):153–172.

Güth, W. (2010). The Generosity Game and Calibration of Inequity Aversion. *Journal of Socio-Economics*, 39:155–157.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J Econ Behav Organ*, 3(4):367–388.

Harris, A., Young, A., Hughson, L., Green, D., Doan, S. N., Hughson, E., and Reed, C. L. (2020). Perceived relative social status and cognitive load influence acceptance of unfair offers in the Ultimatum Game. *PLoS One*, 15(1):e0227717.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., and Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behav Brain Sci*, 28(6):795–815.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., and Ziker, J. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, 327(5972):1480–4.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., and Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781):1767–70.

Herrmann, B., Thöni, C., and Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868):1362–7.

Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G., Holroyd, C. B., and Miltner, W. H. (2011). Why humans deviate from rational choice. *Psychophysiology*, 48(4):507–514.

Høffding, S., Martiny, K., and Roepstorff, A. (2021). Can we trust the phenomenological interview? Metaphysical, epistemological, and methodological objections. *Phenom. Cogn. Sci.*, 21(1):33–51.

Hsieh, S. and Wu, M. (2011). Electrophysiological correlates of preparation and implementation for different types of task shifts. *Brain Res*, 1423:41–52.

Hu, J., Cao, Y., Blue, P. R., and Zhou, X. (2014). Low social status decreases the neural salience of unfairness. *Front Behav Neurosci*, 8:402.

Huettel, S. A. and Kranton, R. E. (2012). Identity economics and the brain: uncovering the mechanisms of social conflict. *Philos Trans R Soc Lond B Biol Sci*, 367(1589):680–91.

Inglehart, R. (2000). Culture and democracy. In Harrison, L. and Huntington, S., editors, *Culture Matters: How Values Shape Human Progress*, pages 80–97. Basic Books, New York.

Ito, T. A., Larsen, J. T., Smith, N. K., and Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *J Pers Soc Psychol*, 75(4):887–900.

Javadzadeh Shahshahani, H., Yavari, M. T., Attar, M., and Ahmadiyeh, M. H. (2006). Knowledge, attitude and practice study about blood donation in the urban population of Yazd, Iran, 2004. *Transfus Med*, 16(6):403–9.

Javan-Noughabi, J., Kavosi, Z., Faramarzi, A., and Khammarnia, M. (2017). Identification determinant factors on willingness to pay for health services in Iran. *Health Econ Rev*, 7(1):40.

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philos Trans R Soc Lond B Biol Sci*, 365(1553):2635–50.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., and Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–78.

Kahneman, D., Knetsch, J. L., and Thaler, R. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *Am Econ Rev*, 76(4):728–741.

Kanske, P. and Kotz, S. A. (2011). Emotion triggers executive attention: anterior cingulate cortex and amygdala responses to emotional words in a conflict task. *Hum Brain Mapp*, 32(2):198–208.

Kassambara, A. (2020). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. R package version 0.6.0.

Kiesel, A., Miller, J., Jolicoeur, P., and Brisson, B. (2008). Measurement of ERP latency differences: a comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, 45(2):250–74.

Kirchsteiger, G. (1994). The role of envy in ultimatum games. *J Econ Behav Organ*, 25(3):373–389.

Klinke, M. E. and Fernandez, A. V. (2022). Taking phenomenology beyond the first-person perspective: conceptual grounding in the collection and analysis of observational evidence. *Phenom. Cogn. Sci.*, (10.1007/s11097-021-09796-1).

Kocher, M., Martinsson, P., and Visser, M. (2012). Social background, cooperative behavior, and norm enforcement. *J Econ Behav Organ*, 81(2):341–354.

Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: a review and synthesis. *Biol Psychol*, 45(1-3):19–56.

Konovalov, A. and Krajbich, I. (2019). Over a Decade of Neuroeconomics: What Have We Learned? *Organ. Res. Methods*, 22(1):148–173.

Kutas, M., Iragui, V., and Hillyard, S. A. (1994). Effects of aging on event-related brain potentials (ERPs) in a visual detection task. *Electroencephalogr Clin Neurophysiol*, 92(2):126–39.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of statistical software*, 82(1):1–26.

Li, J., Pan, J., and Zhu, C. (2020). Inter-brain Synchronization Is Weakened by the Power to Reject Offers in Bilateral Bargaining Games. *SSRN eLibrary*, page e3617843.

Li, S., Qin, X., and Houser, D. (2018). Revisiting gender differences in ultimatum bargaining: experimental evidence from the US and China. *Journal of the Economic Science Association*, 4(2):180–190.

Lintas, A., Vysyaraju, S. C., Jaquerod, M., and Villa, A. E. P. (2017). Event Related Potentials Reveal Fairness in Willingness-to-share. *Lect Notes Comput Sci*, 10613:191–198.

Liu, Y. and Huang, J. (2015). Does Power Corrupt? The Evidence from Event-Related Potentials. *Open Journal of Social Sciences*, 03(01):109–116.

Lopez-Calderon, J. and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci*, 8:213.

Luck, S. J. (2014). *An Introduction to Event-Related Potentials and Their Neural Origins*. 416 pp. MIT Press, Massachusetts Institute of Technology, 2nd edition.

Luck, S. J., Stewart, A. X., Simmons, A. M., and Rhemtulla, M. (2021). Standardized measurement error: A universal metric of data quality for averaged event-related potentials. *Psychophysiology*, 58(6):e13793.

Lüdecke, D. (2021). *sjstats: Statistical Functions for Regression Models (Version 0.18.1)*.

Luo, Y., Wu, T., Broster, L. S., Feng, C., Zhang, D., Gu, R., and Luo, Y. J. (2014). The temporal course of the influence of anxiety on fairness considerations. *Psychophysiology*, 51(9):834–842.

Ma, Q. and Hu, Y. (2015). Beauty matters: social preferences in a three-person ultimatum game. *PLoS One*, 10(5):e0125806.

Makowski, Dominique, Lüdecke, and Daniel (2019). The report package for R: Ensuring the use of best practices for results reporting. *CRAN*. R package.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, 32(1):4–18.

Marchetti, A., Castelli, I., Harlé, K. M., and Sanfey, A. G. (2011). Expectations and outcome: The role of Proposer features in the Ultimatum Game. *J Econ Psychol*, 32(3):446 – 449.

Marlowe, F. W., Berbesque, J. C., Barrett, C., Bolyanatz, A., Gurven, M., and Tracer, D. (2011). The 'spiteful' origins of human cooperation. *Proc Biol Sci*, 278(1715):2159–64.

Mayer, S. V., Rauss, K., Pourtois, G., Jusyte, A., and Schönenberg, M. (2019). Behavioral and electrophysiological responses to fairness norm violations in antisocial offenders. *Eur Arch Psychiatry Clin Neurosci*, 269(6):731–740.

McDonald, J. J. and Green, J. J. (2008). Isolating event-related potential components associated with voluntary control of visuo-spatial attention. *Brain Res*, 1227:96–109.

Mei, D., Zhang, W., and Yin, L. (2020). Neural responses of in-group "favoritism" and out-group "discrimination" toward moral behaviors. *Neuropsychologia*, 139:107375.

Mesrobian, S. K., Bader, M., Götte, L., Villa, A. E. P., and Lintas, A. (2015). Imperfect Decision Making and Risk Taking Are Affected by Personality. In Guy, T. V., Kárný, M., and Wolpert, D. H., editors, *Decision Making: Uncertainty, Imperfection, Deliberation and Scalability*, volume 538 of *Studies in Computation Intelligence*, chapter 6, pages 145–184. Springer, Cham.

Minnix, J. A., Versace, F., Robinson, J. D., Lam, C. Y., Engelmann, J. M., Cui, Y., Brown, V. L., and Cinciripini, P. M. (2013). The late positive potential (LPP) in response to varying types of emotional and cigarette stimuli in smokers: a content comparison. *Int J Psychophysiol*, 89(1):18–25.

Moore, M., Katsumi, Y., Dolcos, S., and Dolcos, F. (2021). Electrophysiological Correlates of Social Decision-making: An EEG Investigation of a Modified Ultimatum Game. *J Cogn Neurosci*, pages 1–25.

Nagel, R. (2001). Rationality and emotions in ultimatum bargaining: Comment. *Annales d'Economie et de Statistique*, pages 33–39.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805):1560–3.

Oosterbeek, H., Sloof, R., and Van de Kuilen, G. (2004). Cultural Differences in Ultimatum Game Experiments : Evidence from a Meta-Analysis. *Exp Econ*, 7(2):171–188.

Peterburs, J., Voegler, R., Liepelt, R., Schulze, A., Wilhelm, S., Ocklenburg, S., and Straube, T. (2017). Processing of fair and unfair offers in the ultimatum game under social observation. *Sci Rep*, 7:44062.

Petit, L. S., Pegna, A. J., Harris, I. M., and Michel, C. M. (2006). Automatic motor cortex activation for natural as compared to awkward grips of a manipulable object. *Exp Brain Res*, 168(1-2):120–30.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., and Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, 37(2):127–152.

Plöchl, M., Ossandón, J. P., and König, P. (2012). Combining eeg and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Front Hum Neurosci*, 6:278–301.

Polezzi, D., Lotto, L., Daum, I., Sartori, G., and Rumiati, R. (2008). Predicting outcomes of decisions in the brain. *Behav Brain Res*, 187(1):116–22.

Polezzi, D., Sartori, G., Rumiati, R., Vidotto, G., and Daum, I. (2010). Brain correlates of risky decision-making. *Neuroimage*, 49(2):1886–1894.

Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain Cogn*, 56(1):5–13.

Proverbio, A. M., Del Zotto, M., and Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indices of animal vs. artifact recognition. *BMC Neurosci*, 8:24.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 4.0.5 edition.

Rachlin, H. (2002). Altruism and selfishness. *Behav Brain Sci*, 25(2):239–296.

Rand, D. G. and Nowak, M. A. (2013). Human cooperation. *Trends Cogn Sci*, 17(8):413–25.

Rauss, K. and Pourtois, G. (2013). What is Bottom-Up and What is Top-Down in Predictive Coding? *Front Psychol*, 4:276.

Righi, S. and Takács, K. (2018). Social Closure and the Evolution of Cooperation via Indirect Reciprocity. *Sci Rep*, 8(1):11149.

Rilling, J. K. and Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annu Rev Psychol*, 62:23–48.

Rodrigues, J., Nagowski, N., Mussel, P., and Hewig, J. (2018). Altruistic punishment is connected to trait anger, not trait altruism, if compensation is available. *Heliyon*, 4(11):e00962.

Sabatinelli, D., Lang, P. J., Keil, A., and Bradley, M. M. (2007). Emotional perception: correlation of functional MRI and event-related potentials. *Cereb Cortex*, 17(5):1085–91.

San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Front Hum Neurosci*, 6:304.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, 300(5626):1755–1758.

Scharmüller, W., Leutgeb, V., Schäfer, A., Köchel, A., and Schienle, A. (2011). Source localization of late electrocortical positivity during symptom provocation in spider phobia: an sLORETA study. *Brain Res*, 1397:10–18.

Servátka, M. (2009). Separating reputation, social influence, and identification effects in a dictator game. *Europ. Econ. Rev.*, 53(2):197–209.

Sharif, L., Marusak, H. A., Peters, C., Elrahal, F., and Rabinak, C. A. (2021). Trustworthiness and electrocortical processing of emotionally ambiguous faces in student police officers. *Psychiatry Res Neuroimaging*, 307:111237.

Spapé, M., Harjunen, V., Ahmed, I., Jacucci, G., and Ravaja, N. (2019). The semiotics of the message and the messenger: How nonverbal communication affects fairness perception. *Cogn Affect Behav Neurosci*, 19(5):1259–1272.

Spicer, J., Galvan, A., Hare, T. A., Voss, H., Glover, G., and Casey, B. (2007). Sensitivity of the nucleus accumbens to violations in expectation of reward. *Neuroimage*, 34(1):455–61.

Stallen, M. and Sanfey, A. G. (2013). The cooperative brain. *Neuroscientist*, 19(3):292–303.

Stevens, F. L., Hurley, R. A., and Taber, K. H. (2011). Anterior cingulate cortex: unique role in cognition and emotion. *J Neuropsychiatry Clin Neurosci*, 23(2):121–5.

Stolz, C., Endres, D., and Mueller, E. M. (2019). Threat-conditioned contexts modulate the late positive potential to faces-A mobile EEG/virtual reality study. *Psychophysiology*, 56(4):e13308.

Sun, L., Tan, P., Cheng, Y., Chen, J., and Qu, C. (2015). The effect of altruistic tendency on fairness in third-party punishment. *Front Psychol*, 6:820.

Sun, S., Yu, R., and Wang, S. (2020). Outcome saliency modulates behavioral decision switching. *Sci Rep*, 10(1):14288.

Sun, S., Zhen, S., Fu, Z., Wu, D.-A., Shimojo, S., Adolphs, R., Yu, R., and Wang, S. (2017). Decision ambiguity is mediated by a late positive potential originating from cingulate cortex. *Neuroimage*, 157:400–414.

Sylwester, K., Herrmann, B., and Bryson, J. J. (2013). Homo homini lupus? explaining antisocial punishment. *J Neurosci Psychol Econ*, 6(3):167–188.

Töllner, T., Gramann, K., Müller, H. J., and Eimer, M. (2009). The anterior N1 component as an index of modality shifting. *J Cogn Neurosci*, 21(9):1653–69.

Van der Veen, F. and Sahibdin, P. (2011). Dissociation between medial frontal negativity and cardiac responses in the ultimatum game: Effects of offer size and fairness. *Cogn Affect Behav Neurosci*, 11:516–525.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Statistics and Computing. Springer-Verlag New York, 498 pp., 4th edition.

Villa, A. E. P., Missonnier, P., and Lintas, A. (2012). Neuroheuristics of Decision Making: from neuronal activity to EEG. In Guy, T. V., Kárný, M., and Wolpert, D. H., editors, *Decision Making with Imperfect Decision Makers*, volume 28 of *Intelligent Systems Reference Library*, chapter 7, pages 159–194. Springer, Berlin, Heidelberg.

Weiss, M., Mussel, P., and Hewig, J. (2020). The value of a real face: Differences between affective faces and emojis in neural processing and their social influence on decision-making. *Soc Neurosci*, 15(3):255–268.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Atten Percept Psychophys*, 72(8):2031–46.

World Medical Association. (2000). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 284(23):3043–3045.

Wu, Y., Leliveld, M. C., and Zhou, X. (2011a). Social distance modulates recipient's fairness consideration in the dictator game: an erp study. *Biol Psychol*, 88(2-3):253–62.

Wu, Y., Zhou, Y., van Dijk, E., Leliveld, M. C., and Zhou, X. (2011b). Social Comparison Affects Brain Responses to Fairness in Asset Division: An ERP Study with the Ultimatum Game. *Front Hum Neurosci*, 5:131.

Xu, Q., Yang, S., Huang, Q., Chen, S., and Li, P. (2020). A sense of unfairness reduces charitable giving to a third-party: Evidence from behavioral and electrophysiological data. *Neuropsychologia*, 142:107443.

Yamagishi, T., Li, Y., Fermin, A. S. R., Kanai, R., Takagishi, H., Matsumoto, Y., Kiyonari, T., and Sakagami, M. (2017). Behavioural Differences and Neural Substrates of Altruistic and Spiteful Punishment. *Sci Rep*, 7(1):14654.

Yu, R., Hu, P., and Zhang, P. (2015). Social distance and anonymity modulate fairness consideration: An ERP study. *Sci Rep*, 5:13452.

Zhang, H., Gu, R., Yang, M., Zhang, M., Han, F., Li, H., and Luo, W. (2021). Context-based interpersonal relationship modulates social comparison between outcomes: an event-related potential study. *Soc Cogn Affect Neurosci*, 16(4):439–452.

Zhang, Y. and Fu, F. (2018). Strategy intervention for the evolution of fairness. *PLoS One*, 13(5):e0196524.