

Published in final edited form as:

J Neurosci. 2011 April 6; 31(14): 5244–5252. doi:10.1523/JNEUROSCI.3138-10.2011.

Neural segregation of objective and contextual aspects of fairness

Nicholas D Wright, Mkael Symmonds, Stephen M Fleming, and Raymond J Dolan

Wellcome Trust Centre for Imaging Neuroscience, Institute of Neurology, University College, London, 12 Queen Square, London, WC1N 3BG, UK

Abstract

Perception of fairness can influence outcomes in human exchange. However, an inherent subjectivity in attribution renders it difficult to manipulate fairness experimentally. Here using a modified Ultimatum Game, within a varying social context, we induced a bias in human subjects' acceptance of objectively identical offers. To explain this fairness-related behaviour we use a computational model to specify metrics for the objective and contextual aspects of fairness, testing for correlations between these model parameters and brain activity determined using fMRI. We show that objective social inequality, as defined by our model, is tracked in posterior insula cortex. Crucially, this inequality is integrated with social context in posterior and mid-insula, consistent with construction of a fairness motivation that flexibly adapted to the social environment. We suggest the dual importance of objective and contextual aspects to fairness we highlight might explain seemingly inconsistent societal phenomena, including public attitudes to income disparities.

Keywords

Game theory; fMRI; fairness; Ultimatum Game; insula; prefrontal

Introduction

Fairness is of interest to sociologists (Homans, 1961), economists (Akerlof, 1979; Kahneman et al., 1986) and neuroscientists (Sanfey et al., 2003). Fairness reflects objective features of how people share resources, classically elicited in the Ultimatum Game (UG) where one player (the Proposer) is given an endowment (e.g. £10) and proposes a division (e.g. keep £6/offer £4) to a second player (the Responder), who can accept (both get the proposed split) or reject (both get nothing) the offer (Guth et al., 1982). Behaviourally, low offer proportions are routinely rejected (Camerer, 2003). However, fairness attribution varies between contexts and individuals: labourers' wages might not seem unfair considered alongside colleagues, yet extremely unfair alongside executives' salaries. Here, we tease apart objective and contextual components of fairness, dissociating their neural substrates. Importantly, we define the contextual component of fairness as a choice bias, leaving open the question of whether subjects are subjectively aware of this shift (Pronin, 2007).

Responders in the classic UG are reported to show greater activity in anterior insula and dorsolateral prefrontal cortex (DLPFC) for lower compared to higher offers, a finding interpreted as reflecting fairness and cognitive-control respectively (Sanfey et al., 2003).

However, as the classic UG cannot dissociate objective and contextual fairness, alternative approaches have endeavoured to isolate components of fairness in the UG. One attempt to unconfound fairness from offer amount treated it as synonymous with offered endowment proportion, implicating lateral PFC in cognitive control (Tabibnia et al., 2008). An alternative strategy manipulated the stimuli used, where by changing Proposer intentionality anterior insula cortex was implicated in fairness responses (Guroglu et al., 2010). Outside the UG framework others have investigated reward comparison (Fliessbach et al., 2007) and fairness in third-party decisions (Hsu et al., 2008), with the latter demonstrating that posterior (but not anterior) insula tracked an objective measure of fairness, namely inequality. However, an isolated fairness manipulation has not as yet been reported.

Here, we contextually manipulate the fairness of a set of offers from a group of Proposers by presentation: alone; interleaved with higher offers from different Proposers; or interleaved with lower offers. Using a formal inequality aversion model (Fehr and Schmidt, 1999; Messick and McClintock, 1968) we isolated objective and contextual fairness components together with economic self-interest, and apply this model's parameters to fMRI data. Behaviourally, we predicted increased acceptance when offers were contextually perceived as fairer, despite being objectively identical. Neurally, we predicted anatomical dissociations between cognitive control in prefrontal and fairness in insula cortex. Within insula itself, a broad-based framework suggests a role for posterior insula encoding more primary quantities, mid-insula in contextual integration (Craig 2002, 2009) and anterior insula in introspective awareness of emotion and bodily-state (Critchley et al., 2004; Paulus and Stein, 2006; Singer et al., 2009). Also within insula, previous studies suggested objective inequality is encoded posteriorly (Hsu et al., 2008) and an integrated fairness metric anteriorly (Sanfey et al., 2003). Therefore, within insula we predicted objective and contextual aspects of fairness would map to posterior and mid/anterior regions respectively.

Materials and Methods

Subjects

32 healthy, right-handed subjects participated in the study, 16 of whom played an Ultimatum Game (UG) during fMRI scanning (10 male; age 18-30; mean age 21.8) and 16 undertook exactly the same task outside the scanner (4 male; age 18-26; mean age 20.6). Two further subjects were excluded from the study, one because he could not tolerate the MRI scanner and the other because he did not understand the task. All subjects provided informed consent and the study was approved by the Institute of Neurology (University College, London) Research Ethics Committee.

Experimental design

Subjects underwent fMRI scanning or behavioural testing as Responder in an UG. The general form of the UG with two players is as follows. Initially, one player (the Proposer) is given an endowment (e.g. £8) and makes an offer to the second player (the Responder) about how to split the endowment (e.g. £6 for the Proposer and £2 for the Responder). The Responder then chooses to accept the offer (both get the split as proposed) or to reject it (both get nothing).

To reinforce the social nature of the task, two subjects always attended each experimental session and at the outset were seated together. Prior to data collection the pair were assigned to separate testing rooms. Subjects understood that they were responding to offers made by past and present participants, who had been placed into one of three coloured Proposer groups (blue, yellow and orange) based on their answers to two questionnaires. The questionnaires were chosen as they quantify the tension between self- and other-regarding

motivations (Machiavellianism IV (Christie and Geis, 1970; Van Lange et al., 1997); and Social Value Orientation (ref)). In each trial, subjects were told which group an offer came from but were not provided with information regarding which specific member. Subjects were told each group comprised the six preceding participants classified into that group and, additionally, that one group would also contain the other attendee that day. Subjects were shown photographs of group members and asked if any were known to them, and subjects also had their own photograph taken for use in future testing with other participants.

This group format was motivated by a need to ensure that, as far as possible, subjects treated each trial individually thereby preventing temporal dependencies in choice with small numbers of individual Proposers gaining reputations over repeated proposals (e.g. earlier lower offers leading to later rejections of higher offers to punish that individual). The group format also avoided a need to present a less plausible scenario that large numbers of subjects had previously attended the experiment. However, in reality the three Proposer groups comprised three sets of 25 offer proportions. Each set spanned a full range from around 0.10 to 0.50 of the endowment with the intention that subjects consider each individual offer and not deterministically accept or reject all offers from a particular group. The behavioural regularity from experimental economics is that offers below 0.25 are rejected about half the time (Camerer, 2003). The “M set” offers were concentrated around this point to maximise our sensitivity to contextual changes in acceptance rates. The “L set” had a mean offer proportion of 0.21; the “M set” had a mean offer proportion of 0.30 (see footnote below); and the “H set” had a mean offer proportion of “0.40” (see Supplementary Materials for the full L and H sets). The means of the “L set” and “H set” were chosen to induce the context effects on the “M set” described below.

Our prime interest in this experiment related to subjects’ responses to the “M set” where our key manipulation of contextual fairness meant that subjects saw the “M set” of 25 offers in three different contexts (Fig. 1, panel b). This change in context was expected to bias subjects’ choices in relation to what was otherwise an objectively identical M set. The specific contexts were: 1) “M alone” in which the full set of 25 M offers was shown on its own in random order with no contextual manipulation; 2) “M-in-H” in which the full set of 25 M offers was shown interleaved with the full set of 25 H offers (i.e. there were 50 offers shown in random order; note that on each individual trial subjects were told which set the offer emanated from); and 3) “M-in-L” with the full set of 25 M offers interleaved with the full set of 25 L offers. The main testing session comprised three runs of trials as Responder. In each run the full set of 25 M offers was presented three times, once in each context (Fig. 1), and the order of these conditions was counterbalanced within and between subjects. Therefore, in this main session there were a total of 375 trials consisting of: 75 M-in-L; 75 M-in-H; 75 M-alone; 75 L-in-M; and 75 H-in-M trials.

During each trial subjects (Responders) first saw from which group a proposal emanated and then saw the proposed division of the endowment in that trial (Fig. 1). They then indicated, via a button press, their decision to either accept (both get the split as proposed) or reject (both get nothing) the offer. The endowment was varied in increments of £0.10 from around £7.70 to £9.80.

For each subject the experiment comprised four consecutive phases. Initially, subjects completed questionnaires and were then informed of the nature of the task. Secondly, subjects made 20 proposals. Thirdly, to learn the reputations of the three groups (L, M and H), subjects (as Responders) played against the full set of 25 offers in each group separately. Fourthly, the main testing session comprised three runs of trials as Responder and a total of 375 trials. To provide subjects with rest periods of approximately six seconds and to mitigate against fatigue, subjects saw a brief screen with “REST” displayed every 8-9 contiguous

trials. After each rest period an introductory screen announced the group or groups whose offers would be presented next (i.e. M group only; M and H groups; M and L groups). Of the two subjects attending each experimental session the one with fewer rejections in the learning phase continued the main testing phase in the behavioural testing room whilst the second subject conducted this main phase in the MRI scanner. This was to increase power to detect fairness related activation independently of choice and also to avoid scanning subjects with a deterministic strategy of accepting all offers. Subjects were informed that this selection had been made at random. At the end of the experimental session subjects were debriefed and all reported believing that the proposals they faced were made by present and past participants. The subject's payment was determined by responses and proposals chosen at random. Subjects received on average around £30 (~\$50). All statistical tests were two-tailed.

Behavioural modelling

We fit data on an individual subject basis both according to a psychometric model and an economic model. In our psychometric analysis we modelled changes in the proportion of acceptance as offer proportion increases with logistic regression:

$$P(\text{accept}) = \frac{1}{1 + e^{-(b_0 + b_1 z)}}$$

In this model z is the offer proportion (binned into groups of 5 trials). We estimate the model separately for "M" offers in each of the three offer contexts (M-alone, M-in-L and M-in-H).

In our economic analysis, we modelled individual choices using a binary logistic regression utility model. Various utility functions with other regarding preferences have been proposed (Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Fehr and Schmidt, 1999; Messick and McClintock, 1968) that make similar predictions in the UG and we chose the following formulation:

$$P(\text{accept}) = \frac{1}{1 + e^{-\frac{U}{\lambda}}}$$

where $U = X_{\text{self}} - \alpha^* (X_{\text{other}} - X_{\text{self}})$, $\alpha \geq 0; \lambda \geq 0$

x_{self} is the amount the Proposer offered in the trial and x_{other} is the amount the Proposer keeps, α is an 'envy' parameter (reflecting a tradeoff between inequality and self interest), and λ reflects choice randomness. In this model, there is no constant term as we assume a utility of 0 represents indifference between acceptance and rejection of an offer (i.e. rejection of an offer has a utility of 0). For illustration, in a single trial as Responder, the utility of the offer is calculated by combining the self-regarding component (amount to self) and the other-regarding component (the weighted impact of inequality). This utility is then compared to the utility of rejecting (zero), with the offer being accepted if greater and rejected if lesser (with stochasticity in action selection captured by λ). To further illustrate the other-regarding component of the utility function, we used an objective metric for unfairness (inequality) weighted by the α ("envy") parameter. For example, if $\alpha=0$ the subject is entirely self-regarding (they will accept any offer however small) and if α is large they will reject lower offers. Because in this study x_{self} did not exceed x_{other} the "guilt" parameter contained in some formulations was not used (Fehr and Schmidt, 1999). We optimised subject-specific α and λ parameters across all trials using nonlinear optimization implemented in Matlab for maximum likelihood estimation.

We also estimated parameters for each subject in the M-in-H and in the M-in-L conditions to examine context effects. 30 of the 32 subjects were included as 2 behavioural subjects accepted essentially all offers such that parameters could not be estimated. No difference was found when a paired t-test was used to compare either α or λ from the two contexts obtained when both parameters were freely estimated for each subject in both conditions. As we were interested in the relative contributions of both parameters, we fixed λ as the mean λ in the two conditions for that subject. As an additional check, we confirmed that this result held when λ was fixed at the average for these two conditions across all subjects. Finally, we performed the same procedure but fixed α and estimated λ .

fMRI data acquisition

Images were acquired using a 3T Allegra scanner (Siemens, Erlangen, Germany). BOLD sensitive functional images were acquired using a gradient-echo EPI sequence (48 transverse slices; TR, 2.88 secs; TE, 30 ms; 3×3 mm in-plane resolution; 2 mm slice thickness; 1 mm gap between adjacent slices; z-shim -0.4 mT/m; positive phase encoding direction; slice tilt -30 degrees) optimised for detecting changes in the OFC and amygdala (Weiskopf et al., 2006). Four runs of 284-286 volumes were collected for each subject, followed by a T1-weighted anatomical scan. Local field maps were also acquired.

fMRI data analysis

Functional data were analysed using standard procedures in SPM5 (Statistical Parametric Mapping; www.fil.ion.ucl.ac.uk/spm). fMRI timeseries were regressed onto a composite general linear model (GLM) containing delta (stick) functions representing the onsets of the offer. These delta functions were convolved with the canonical HRF and its temporal derivative. The stimulus delta functions were separated into five regressors depending on the Proposer type (L-in-M, M-in-L, M-alone, M-in-H and H-in-M). Each was then parametrically modulated by two orthogonalised regressors entered in the following order: offer amount followed by inequality ($x_{other} - x_{self}$). A second GLM with a 2 (accept, reject) by 5 (Proposer type) factorial design matrix was also constructed to determine the betas for the components of the interaction of choice and context. Throughout the analysis of the imaging data the main effects were calculated using the data across all five Proposer types, whilst the effect of context was probed within the comparison of M-in-L and M-in-H. These two conditions were fully matched in terms of offer set (i.e. M offers) and task demands (unlike M-alone offers, which were not interleaved with another trial type).

Cluster-based statistics were used to define significant activations both on their intensity and spatial extent (Friston et al., 1994). Clusters were defined using a threshold of $P < 0.005$ and corrected for multiple comparisons using family-wise error correction and a threshold of $P < 0.05$. For presentation purposes, images are displayed at $P < 0.001$ uncorrected unless otherwise stated. We report all activations at $P < 0.05$ that survive whole brain correction using family-wise error at the cluster level, unless otherwise stated.

RESULTS

Behavior

To confirm consistency of our results with well known behavioural regularities in the UG (Camerer, 2003), we first collapsed across all Proposer group sets. In both the learning and main sessions all subjects accepted almost all offers of half the endowment, and 29 of the 32 subjects rejected almost all offers under one tenth of the endowment. In the main session the mean acceptance rate was 0.50 (s.d. 0.20) (Fig. 2). Using an analytic framework derived from psychophysics we show a graded relationship between increasing offer size and increasing acceptance rate, at both individual (Fig. 3) and group levels (Supplemental Fig.

2). Moreover, such a framework predicts that reaction times (RTs) should be greatest at the point of equality in value as this represents the point of maximum decision uncertainty (Grinband et al., 2006), which is exactly what we observed (Fig. 3; Supplemental Fig. 2).

To test our key behavioural prediction of a contextual bias in fairness perception, we focused on the critical M set trials from the main session. Here the set of M offers was presented alone (“M-alone”); interleaved with H offers (“M-in-H”; contextually less fair) or interleaved with the L offers (“M-in-L”; contextually more fair). When all 32 subjects were included in the analysis we observed effects of social context on the acceptance rate of otherwise identical offers (one way ANOVA; $F(2,62)=4.88$, $P=0.013$), driven by a highly significant difference between M-in-L trials (seen as “fairest”; mean 48.5%) relative to M-in-H trials (seen as “less fair”; mean 45.8%; paired two-tailed t-test, $t(31)=2.86$; $P=0.007$; Fig. 2).

16 of the 32 subjects who comprised our behavioural sample also performed the task in the fMRI scanner. The same pattern was seen when the 16 scanned subjects were analysed separately for the key comparison of M-in-L with M-in-H ($t(15)=2.64$, $P=0.019$; M-in-L mean 39.3%; M-in-H mean 36.3%) with trend-level significance for a one-way ANOVA across all three contexts ($F(2,30)=3.21$, $P=0.074$). This contextual bias was primarily driven by a lower rate of acceptance in the “less fair” (M-in-H) condition, with a significant difference between M-in-H and M-alone (all subjects $t(31)=2.18$, $P=0.037$; scanned subjects $t(15)=2.54$, $P=0.023$) but not between M-in-L and M-alone ($P>0.1$).

Modelling behavior

We next tested if a formal economic model predicted subjects’ choices (Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Fehr and Schmidt, 1999; Messick and McClintock, 1968). The intention here was to derive model parameters as a bridge to underlying neural mechanisms. The model is specified by the following formalism:

$$U = X_{\text{self}} - \alpha^* (X_{\text{other}} - X_{\text{self}})$$

In this model, the utility (U) a subject derives from a proposal is given by the subject’s payoff (x_{self}) minus the weighted inequality of the proposal, with inequality being the amount kept by the Proposer (x_{other}) less the amount offered. The α (“envy”) parameter quantifies how much a particular subject cares about inequality (i.e. how much they weight this social component, e.g. if $\alpha=0$ they are entirely self-regarding). We combined this model of fairness preference with a logistic model of stochastic choice with a noise parameter, λ , and estimated both parameters using maximum likelihood estimation.

We first collapsed across all trial types. Across all 32 subjects mean α was 0.89 (s.d. 0.54; range 0-2.59) and mean λ was 0.54 (s.d. 0.31; range 0-1.26). For the 16 scanned subjects, mean α was 1.08 (s.d. 0.55; range 0.38-2.59) and mean λ was 0.51 (st. dev. 0.31; range 0.20-1.26). These observations are consistent with previous studies using similar models (Fehr and Schmidt, 1999; Krajbich et al., 2009). A statistical measure of how well the model predicted subjects’ choice is provided by a likelihood ratio test: comparing the full to a reduced model without the inequality term was highly significant (all 32 subjects $X^2(1)>10^4$, $P<0.001$; 16 scanned subjects $X^2(1)>6000$, $P<0.001$).

We next examined the contextual bias in fairness perception by estimating α and λ in the M-in-H and M-in-L conditions for 30 subjects (2 subjects in the behavioural group accepted essentially all offers; the remaining 30 subjects had mean $\alpha=0.95$, mean $\lambda=0.56$ when collapsed across all trial types). There was a significant difference in α between the M-in-H

and M-in-L conditions when the λ parameter used for estimation was fixed at either the mean λ in the two conditions for that subject (mean $\alpha_{M-in-H}=1.01$; mean $\alpha_{M-in-L}=0.91$; $t(29)=2.246$, $P=0.032$) or the average for these conditions across all subjects (mean $\alpha_{M-in-H}=0.97$; mean $\alpha_{M-in-L}=0.89$; $t(29)=2.157$, $P=0.039$). λ did not significantly differ between conditions when α was fixed in this fashion.

Neuroimaging

16 of the subjects performed the task in the fMRI scanner and we report only activation surviving whole brain correction at the cluster level unless otherwise stated. Initially, in a conventional factorial analysis we determined whether our behavioural bias in choice (accept or reject) driven by a change in social context (M-in-L or M-in-H) was reflected in brain activity. We computed the interaction term $[(Acc_{M-in-H} + Rej_{M-in-L}) - (Acc_{M-in-L} + Rej_{M-in-H})]$, which revealed an effect in right DLPFC (middle frontal gyrus; Table 1). Examination of simple effects showed that this interaction was driven by greater activity for rejecting contextually fairer offers (i.e. M-in-L) than rejecting contextually less fair offers (i.e. M-in-H; Fig. 4; paired two-tailed tests rejection $t(15)=3.3$, $P=0.005$, acceptance $t(15)=0.04$, $P>0.9$).

We next used a parametric analysis to isolate activity correlating with specific components of our formal economic model and their relationship to social context. Offer amount and inequality were included as parametric regressors, and the subject-specific envy (a) parameter was included as a second level covariate on the inequality regressor. We orthogonalised inequality with respect to offer amount in order to identify its independent contribution to the BOLD signal having accounted for activity related to the offer magnitude (visually illustrated in Supplemental Fig. 3). We examined the main effect of inequality, our objective metric of fairness, by collapsing across all trials (L; H; M alone; M-in-L; M-in-H) and this analysis showed a negative correlation with inequality in right posterior insula (i.e. greater activation for a more equal allocation; Table 2 and Fig. 5). A similar pattern was evident on the left which did not survive cluster level correction for the whole brain ($x=-45$, $y=-15$, $z=3$, $Z=4.01$, 58 voxels at $P<0.005$ uncorrected).

We next asked where in the brain inequality is integrated with context to produce changes in contextual fairness. The key contrast in this analysis was to compute the interaction between the parametric regressor for inequality in the more (M-in-L) versus the less fair (M-in-H) context [$inequality_{M-in-L} > inequality_{M-in-H}$]. This interaction showed activation in bilateral mid-insula and rolandic operculum, extending on the right into posterior insula and inferior parietal cortex (Table 2 and Fig. 5). Intuitively, this interaction can be expressed as a difference in regression slope of activity under both levels of the categorical factor (Toga and Mazziotta, 2002) (see Fig. 5). Interestingly, in light of our inequality aversion model, inequality modulated insula activity in the condition when offers are viewed as more aversive (M-in-H; Fig. 5). In fact, the observation that the M-in-H condition drove the interaction effect exactly mirrors our behavioural finding where this factor was the principal driver of a contextual bias (Fig. 2).

For comprehensiveness we also tested for areas correlating with inequality in the matched M-in-H and M-in-L conditions alone. Here, we demonstrated a similar negative correlation with activation in posterior insula to that seen when including all trials ($P < 0.05$, cluster-level whole-brain corrected on the left, Table 2; 68 vox. at $P < 0.005$ unc. on the right at $x=42$ $y=-24$ $z=27$, $Z=3.51$) with additional activation seen in STS close to the location of activation reported in “theory of mind” (TOM) tasks (Gobbini et al., 2007; Grezes et al., 2004). Separately, we observed an interaction of offer amount with social context in right DLPFC [$offer\ amount_{M-in-L} - offer\ amount_{M-in-H}$] Table 2), which in light of the fact that

choice is highly correlated with offer magnitude converges with the pattern identified by the interaction term in our factorial design (Fig. 4).

Finally, we examined neural correlates of how much a particular subject cares about what others receive, regardless of the particular contextual manipulation. In our economic model this is captured by the α (“envy”) parameter derived from each subject’s behaviour, which specifically weights the inequality component of the utility function. The magnitude of the α parameter correlated across subjects with activity in the precuneus ($P < 0.05$, voxel-level whole brain corrected), left frontopolar region ($P < 0.05$, voxel-level whole brain corrected) and left temporo-parietal junction (cluster level corrected as above).

Discussion

Our principal aim in this study was a behavioural and neural characterisation of objective and contextual aspects of fairness. We defined the contextual component of fairness as a shift in choices in response to otherwise identical offers, while remaining agnostic to the question of conscious awareness of this shift. Our finding of a marked context-dependence provides a perspective on fairness as a relative rather than absolute quantity, echoing findings in relation to other high-level quantities such as valuation (Ariely et al., 2006; Seymour and McClure, 2008; Vlaev et al., 2009). However, our neural data also highlight a fundamental role for objective social inequality that accords with effects seen in the UG across diverse cultures (Henrich et al., 2004), in human infants (Fehr et al., 2008) and in similar tasks in non-human primates (Brosnan and De Waal, 2003, but note Jensen et al., 2007). Our data highlights how these objective and contextual aspects interact to construct a fairness motivation with sufficient flexibility to enable appropriate responses to the social environment.

Fairness relates to how intentional agents should divide resources amongst potentially entitled recipients (Kahneman et al., 1986). Inequality aversion quantifies how this motivation influences choice (Messick and McClintock, 1968). Here, we assume choice is the outcome of processes whose neural implementation may involve social computations such as prediction errors (e.g. Behrens et al., 2008; Hampton et al., 2008). We also characterise objective and contextual components of fairness in decision-making by combining a standard economic inequality-aversion model (Fehr and Schmidt, 1999; Messick and McClintock, 1968) with psychophysical methods and the psychological concept of cognitive control (Gilbert and Burgess, 2008; Miller and Cohen, 2001).

Our neuroimaging data strongly support inequality aversion models: first, we find a main effect of inequality in posterior insula; second, between subjects the envy parameter correlates with activity in the precuneus, left TPJ and frontopolar cortex; third, inequality modulated posterior and mid-insula activity more strongly when inequality is psychologically more aversive (M-in-H; Fig. 5). Our findings extend current inequality aversion models, demonstrating the behavioural and neural flexibility to avoid knee-jerk aversion to inequality.

We found no neural correlate for a self-interested component of the utility function, a result that accords with previous UG studies that, as indeed was also the case here, did not provide reward-related feedback during the task (Guroglu et al., 2010; Sanfey et al., 2003; Halko et al., 2009). Conversely, robust reward-related activity has been seen in social comparison tasks when feedback was given in an estimation task (Fleissbach et al., 2007) or where, analogous to feedback, subjects were given an outcome in every trial (Tricomi et al., 2010). Our data tentatively link the behavioural economic concept of “other regarding preferences” (Fehr and Camerer, 2007) and the psychological concept of “Theory of Mind” (TOM) (Frith

and Frith, 2006; Premack and Woodruff, 1978). Subjects' α (envy) parameter correlated with activity in a subset of TOM-related areas including TPJ, implicated in representing others' intentions, and precuneus involved in perspective taking (Van Overwalle and Baetens, 2009).

Neurally, our results implicate insula cortex in fairness motivation and, combined with previous work, suggest functional segregation in this extensive (over 5cm long) and cytoarchitecturally diverse cortical region (Flynn, 1999; Varnavas and Grand, 1999). Both here, and in Hsu et al. (2008), posterior insula activity negatively correlated with inequality (see Fig 4 in Hsu et al., 2008). Hsu and colleagues asked subjects to choose between distributions of meals for African children, varying in inequality and amount. Our concordant findings are striking, as Hsu used decisions about third-parties rather than first-party decisions (e.g. in the UG), a difference markedly affecting choice in behavioural experiments (Camerer, 2003). We also find a mid-insula peak for integration of context with inequality in our UG. Anterior insula activity has been reported as higher for rejected versus accepted offers in the UG (Sanfey et al, 2003), a result replicated in a task-matched study (Halko et al., 2009), although the same contrast in other UG studies shows little activity in this region (Guroglu et al. 2010; Tabibnia et al 2008; this study). Indeed, recent work shows anterior insula activity depends on Proposer intentionality in the UG (Guroglu et al., 2010).

One resolution to these diverse findings is that distinct fairness-related processes are expressed in segregated regions of the insula. Thus, a negative correlation with inequality in posterior insula occurs when subjects can form predictions about inequality, having previously experienced group offers (in our study) or the distribution of experimental allocations (in Hsu et al., 2008). Given predictions for inequality, a more equal division than expected could engender a positive prediction error and a more unequal division a negative prediction error (see also Paulus and Stein 2006; Singer et al., 2009). This explanation for the observed negative correlation can also explain why it may not have been seen when each UG Proposer is encountered only once and fewer trials are played, as predictions cannot be formed (Halko et al., 2009; Sanfey et al., 2003). Capacity of posterior insula for high-level computation is suggested by involvement in other high-level tasks, albeit different to fairness, for example inter-temporal choice (Tanaka et al., 2004; Wittmann et al., 2007) and language perception (Jones et al., 2010). Note the peak activity for our contextual manipulation is in bilateral mid-insula, which has a role integrating representations of physiological state or feelings from the body with activity associated with awareness (Craig, 2002; Craig, 2009; Farrer et al., 2003; Tsakiris et al., 2007). Finally, anterior insula activity in some UG studies may reflect processing of disgust (Sanfey et al., 2003) or aversion to norm-violation (Guroglu et al., 2010), consistent with its role in introspective awareness of emotion (Craig 2009). Such emotional impact may be attenuated with many more trials (hundreds versus 10 human Proposals in Sanfey et al., 2003), with unfairness directed at third-parties (Hsu et al., 2008), or where much shorter trials reduce scope for introspection (36secs in Sanfey et al., 2003).

In DLPFC we noted an interaction between choice and context, driven by greater activity for rejecting contextually fairer offers than less fair offers (Fig. 4, Tables 1 and 2). Against a background of multiple competing motivations in the UG, our results appear consistent with previous suggestions of a role for DLPFC in cognitive control during the UG (Sanfey et al. 2003; Knoch et al. 2006). In light of this, one interpretation of the data is that increased activity in DLPFC during rejection of contextually more fair offers reflects enhanced difficulty of implementing these rejections. Such an interpretation is consistent with evidence that rTMS to right (but not left) DLPFC increases acceptance of lower proportion offers (Knoch et al. 2006), where such disruption would be expected to diminish rejection of contextually fairer (M-in-L) offers. However, this interpretation has to be tempered by

consideration of the fact that if indeed DLPFC activity is specific to rejection, then this would predict uniformly greater activity in DLPFC for rejection compared to acceptance, which was not what we observed (Fig. 4). Understanding the dynamics of DLPFC activity in the UG is clearly highly complex and an issue for further investigation using more refined paradigms.

In conclusion, we provide evidence that objective inequality of social distributions fundamentally influences fairness behaviour; and that this inequality is flexibly integrated with social context. This account might explain otherwise hard to reconcile social phenomena. A role for objective inequality helps explain the trans-cultural phenomenon of UG rejections (Henrich et al., 2004), the importance of workplace inequality (Akerlof, 1982) and the historical attraction of political ideas stressing equality (Rousseau, 1754). However, whilst within a particular context inequality is key, relativity of fairness helps explain the importance of comparator groups in labour negotiations (Akerlof and Shiller, 2009); and the muted public response as US executive pay gradually increased over the past 20 years from around 60 to 160 times median US income (The Economist, 2007). That fairness inherently involves both objective and contextual aspects can also inform wide-ranging social debates, from labour disputes to fair structuring of tax systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was carried out under a Wellcome Trust Programme Grant to R.J.D. N.D.W. is in receipt of a Wellcome Trust Research Training Fellowship. S.M.F. is in receipt of an MRC Studentship. We thank Chris Frith and Bahador Bahrami for their helpful comments.

Reference List

- Akerlof GA. The Case against Conservative Macroeconomics: An Inaugural Lecture. *Economica*. 1979; 46:219–237.
- Akerlof GA. Labor Contracts as Partial Gift Exchange. *The Quarterly Journal of Economics*. 1982; 97:543–569.
- Akerlof, GA.; Shiller, RJ. *Animal Spirits*. Princeton University Press; Princeton: 2009.
- Ariely D, Loewenstein G, Prelec D. Tom Sawyer and the construction of value. *Journal of Economic Behavior & Organization*. 2006; 60:1–10.
- Behrens TE, Hunt LT, Woolrich MW, Rushworth MF. Associative learning of social value. *Nature*. 2008; 456:245–249. [PubMed: 19005555]
- Bolton GE, Ockenfels A. ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review*. 2000; 90:166–193.
- Brosnan SF, De Waal FB. Monkeys reject unequal pay. *Nature*. 2003; 425:297–299. [PubMed: 13679918]
- Camerer, CF. *Behavioral Game Theory: experiments on strategic interaction*. Princeton UP; Princeton: 2003.
- Charness G, Rabin M. Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics*. 2002; 117:817–869.
- Christie, R.; Geis, F. *Studies in Machiavellianism*. Academic Press; New York: 1970.
- Craig AD. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci*. 2002; 3:655–666. [PubMed: 12154366]
- Craig AD. How do you feel--now? The anterior insula and human awareness. *Nat Rev Neurosci*. 2009; 10:59–70. [PubMed: 19096369]

- Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ. Neural systems supporting interoceptive awareness. *Nat Neurosci.* 2004; 7:189–195. [PubMed: 14730305]
- Farrer C, Franck N, Georgieff N, Frith CD, Decety J, Jeannerod M. Modulating the experience of agency: a positron emission tomography study. *Neuroimage.* 2003; 18:324–333. [PubMed: 12595186]
- Fehr E, Camerer CF. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn Sci.* 2007; 11:419–427. [PubMed: 17913566]
- Fehr E, Bernhard H, Rockenbach B. Egalitarianism in young children. *Nature.* 2008; 454:1079–1083. [PubMed: 18756249]
- Fehr E, Schmidt KM. A Theory Of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics.* 1999; 114:817–868.
- Fliessbach K, Weber B, Trautner P, Dohmen T, Sunde U, Elger CE, Falk A. Social comparison affects reward-related brain activity in the human ventral striatum. *Science.* 2007; 318:1305–1308. [PubMed: 18033886]
- Flynn FG. Anatomy of the insula functional and clinical correlates. *Aphasiology.* 1999; 13:55–78.
- Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping.* 1994; 1:210–220.
- Frith CD, Frith U. How we predict what other people are going to do. *Brain Res.* 2006; 1079:36–46. [PubMed: 16513098]
- Gilbert SJ, Burgess PW. Executive function. *Curr Biol.* 2008; 18:R110–R114. [PubMed: 18269902]
- Gobbini MI, Koralek AC, Bryan RE, Montgomery KJ, Haxby JV. Two takes on the social brain: a comparison of theory of mind tasks. *J Cogn Neurosci.* 2007; 19:1803–1814. [PubMed: 17958483]
- Grezes J, Frith CD, Passingham RE. Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage.* 2004; 21:744–750. [PubMed: 14980577]
- Grinband J, Hirsch J, Ferrera VP. A neural representation of categorization uncertainty in the human brain. *Neuron.* 2006; 49:757–763. [PubMed: 16504950]
- Guroglu B, van den Bos W, Rombouts SA, Crone EA. Unfair? It depends: Neural correlates of fairness in social context. *Soc Cogn Affect Neurosci.* 2010
- Guth W, Schmittberger R, Schwarze B. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization.* 1982; 3:367–388.
- Halko ML, Hlushchuk Y, Hari R, Schurmann M. Competing with peers: mentalizing-related brain activity reflects what is at stake. *Neuroimage.* 2009; 46:542–548. [PubMed: 19385019]
- Hampton AN, Bossaerts P, O'Doherty JP. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc Natl Acad Sci U S A.* 2008; 105:6741–6746. [PubMed: 18427116]
- Henrich, J.; Boyd, R.; Bowles, S.; Camerer, C.; Fehr, E.; Gintis, H. *Foundations of Human Sociality.* Oxford University Press; Oxford: 2004.
- Homans, GC. *Social Behavior: Its Elementary Forms.* Harcourt Brace; New York: 1961.
- Hsu M, Anen C, Quartz SR. The right and the good: distributive justice and neural encoding of equity and efficiency. *Science.* 2008; 320:1092–1095. [PubMed: 18467558]
- Jensen K, Call J, Tomasello M. Chimpanzees are rational maximizers in an ultimatum game. *Science.* 2007; 318:107–109. [PubMed: 17916736]
- Jones CL, Ward J, Critchley HD. The neuropsychological impact of insular cortex lesions. *J Neurol Neurosurg Psychiatry.* 2010; 81:611–618. [PubMed: 20522870]
- Kahneman D, Knetsch JL, Thaler R. Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *The American Economic Review.* 1986; 76:728–741.
- Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science.* 2006; 314:829–832. [PubMed: 17023614]
- Krajbich I, Adolphs R, Tranel D, Denburg NL, Camerer CF. Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J Neurosci.* 2009; 29:2188–2192. [PubMed: 19228971]
- Messick DM, McClintock CG. Motivational Bases of Choice in Experimental Games. *Journal of Experimental Social Psychology.* 1968; 4:1–25.

- Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci.* 2001; 24:167–202. [PubMed: 11283309]
- Paulus MP, Stein MB. An insular view of anxiety. *Biol Psychiatry.* 2006; 60:383–387. [PubMed: 16780813]
- Premack D, Woodruff G. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences.* 1978; 1:515–526.
- Pronin E. Perception and misperception of bias in human judgment. *Trends Cogn Sci.* 2007; 11:37–43. [PubMed: 17129749]
- Rousseau, JJ. *A Discourse on Inequality.* Penguin Books Ltd.; London: 1754.
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. The neural basis of economic decision-making in the Ultimatum Game. *Science.* 2003; 300:1755–1758. [PubMed: 12805551]
- Seymour B, McClure SM. Anchors, scales and the relative coding of value in the brain. *Curr Opin Neurobiol.* 2008; 18:173–178. [PubMed: 18692572]
- Singer T, Critchley HD, Preusschoff K. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn Sci.* 2009; 13:334–340. [PubMed: 19643659]
- Tabibnia G, Satpute AB, Lieberman MD. The sunny side of fairness: preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychol Sci.* 2008; 19:339–347. [PubMed: 18399886]
- Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci.* 2004; 7:887–893. [PubMed: 15235607]
- The Economist. In the Money. Jan 18th.2007
- Toga, AW.; Mazziotta, JC. *Brain Mapping: The Methods.* Academic Press; New York: 2002.
- Tricomi E, Rangel A, Camerer CF, O’Doherty JP. Neural evidence for inequality-averse social preferences. *Nature.* 2010; 463:1089–1091. [PubMed: 20182511]
- Tsakiris M, Hesse MD, Boy C, Haggard P, Fink GR. Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb Cortex.* 2007; 17:2235–2244. [PubMed: 17138596]
- van ’t Wout M, Kahn RS, Sanfey AG, Aleman A. Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *Neuroreport.* 2005; 16:1849–1852. [PubMed: 16237340]
- Van Lange PA, Otten W, De Bruin EM, Joireman JA. Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *J Pers Soc Psychol.* 1997; 73:733–746. [PubMed: 9325591]
- Van Overwalle F, Baetens K. Understanding others’ actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage.* 2009; 48:564–584. [PubMed: 19524046]
- Varnavas GG, Grand W. The insular cortex: morphological and vascular anatomic characteristics. *Neurosurgery.* 1999; 44:127–136. [PubMed: 9894973]
- Vlaev I, Seymour B, Dolan RJ, Chater N. The price of pain and the value of suffering. *Psychol Sci.* 2009; 20:309–317. [PubMed: 19254237]
- Weiskopf N, Hutton C, Josephs O, Deichmann R. Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. *Neuroimage.* 2006; 33:493–504. [PubMed: 16959495]
- Wittmann M, Leland DS, Paulus MP. Time and decision making: differential contribution of the posterior insular cortex and the striatum during a delay discounting task. *Exp Brain Res.* 2007; 179:643–653. [PubMed: 17216152]

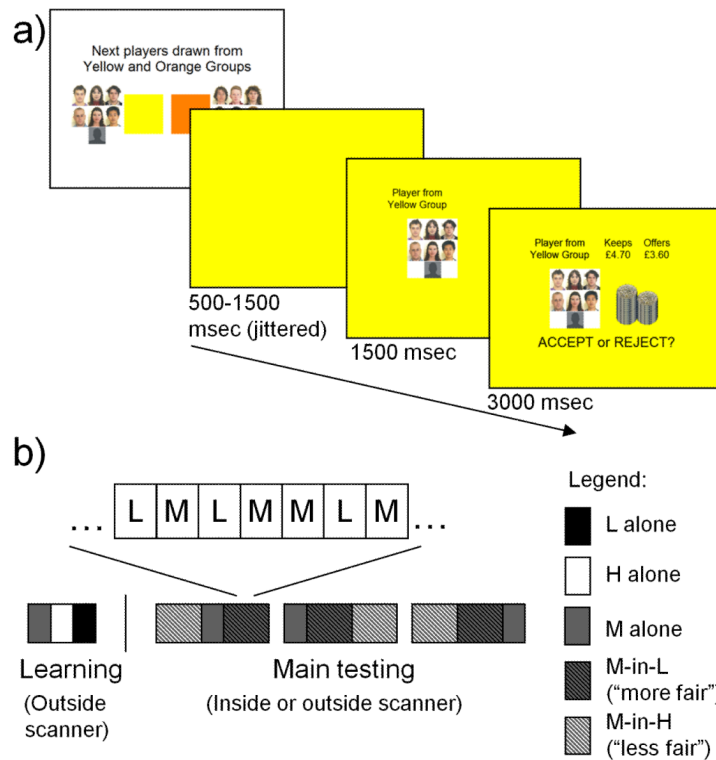


Figure 1. Illustration of experimental design.

a) Timeline of trials. Illustrated is the task from the perspective of the Responder: firstly a blank screen is presented for 500-1500msec (mean 1000msec) in the colour of the Proposer group (L, M or H); second, a panel containing the photographs of the Proposer group was added for 1500 msec; and third the proposal was then shown for 3000msec (denoted both numerically and visually by the height of the coin stacks) along with the instruction to accept or reject (side counterbalanced between subjects). Subjects understood that the silhouette represented the other subject attending that session, who had been placed in one of the three Proposer groups. During the 3000msec in which the proposal was shown, subjects had to decide by a button press whether to accept or reject the offer. Subjects saw a brief screen with "REST" displayed every 8-9 contiguous trials before an introductory screen announced the group or groups whose offers would be presented next (i.e. M group only; M and H groups; M and L groups). **b)** Order of conditions. Initially, in a reputation learning session performed outside the scanner subjects responded to the full set of 25 offers from the M Proposer group alone (grey in panel b), the 25 H offers alone (white in panel b) and the 25 L offers alone (black in panel b). Subjects then underwent the main testing session in the scanner or behaviourally, which comprised 3 runs. In each run of 125 trials, subjects responded to the M set of offers in 3 contexts: alone ("neutral"); interleaved with the L set (M-in-L; "more fair"); and interleaved with the H set (M-in-H; "less fair").

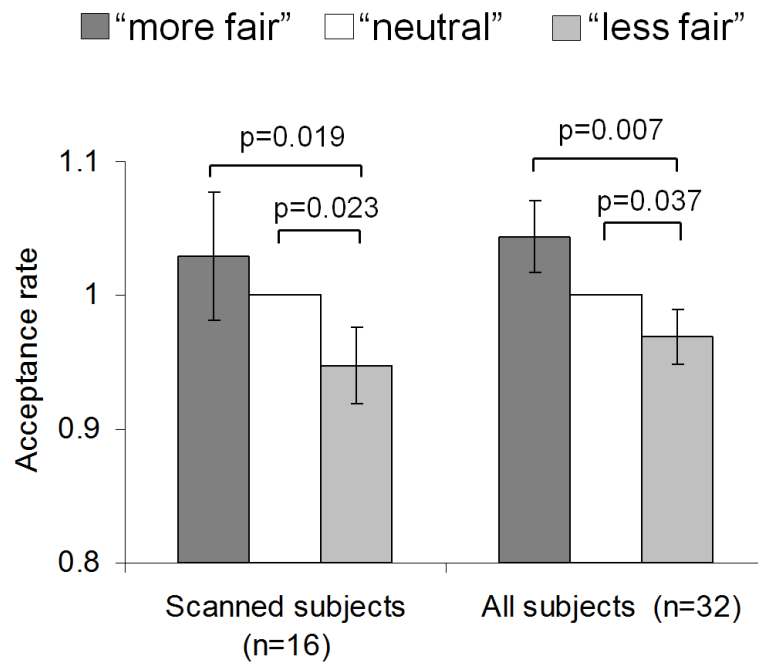


Figure 2. Biased acceptance of objectively identical offers by contextual manipulation. Each bar represents rate of acceptance (\pm s.e.m) of the “M” set of offers that spans the full range of offer proportions from 0.08 to 0.50. The M set is presented in three different contexts, representing a manipulation of contextual fairness: M-in-L (“more fair”, interleaved with lower offers); M-alone (“neutral”, presented alone); and M-in-H (“less fair”, interleaved with higher offers). The rate of acceptance is normalised with respect to the rate of acceptance in the neutral M-alone condition. Data is shown for the main session for the scanned group ($n=16$) and the combined data from all subjects including both the scanned subjects and those who solely underwent behavioural testing ($n=32$).

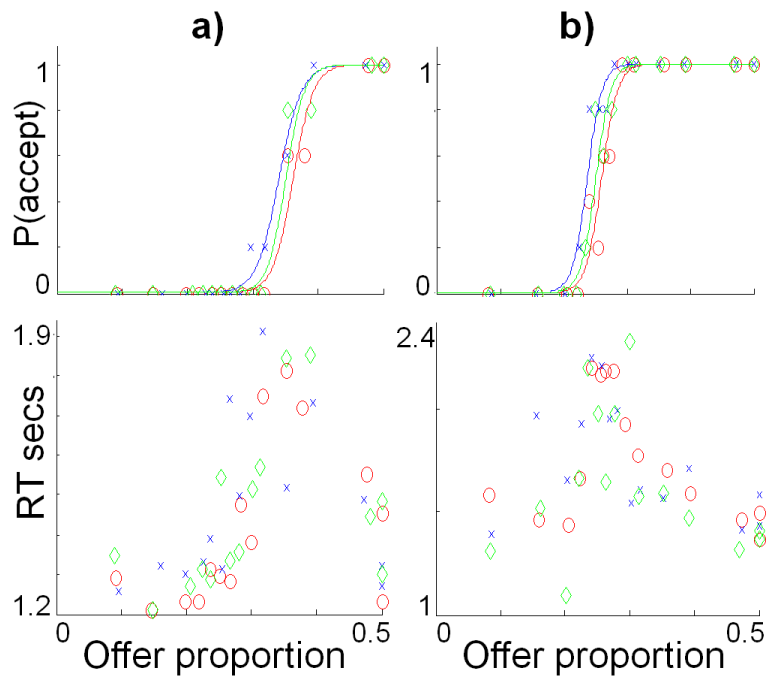


Figure 3. Psychometric analysis of individual subject data.

Data from the scanning session is shown for two exemplar subjects (a and b). For each subject the upper panels show the probability of acceptance and the lower panels show reaction times (RTs), both plotted against offer proportion (each data point being the mean of five trials). Data is shown for the “M” offers in the three contexts: M-alone (green; “neutral”); M-in-L (blue; “more fair”); and M-in-H (red; “less fair”). The upper panels also show probability of acceptance as a logistic function fitted to those points, demonstrating a contextual bias evident in a shift to the left for contextually more fair (M-in-L) and a shift to the right for the less fair (M-in-H) relative to the control condition (M-alone). In the lower panels it can be seen that the point of indifference (probability of acceptance is 0.5) corresponds to peak RTs, consistent with choice difficulty being greatest at this point and arguing against either acceptance or rejection as being a “default” choice.

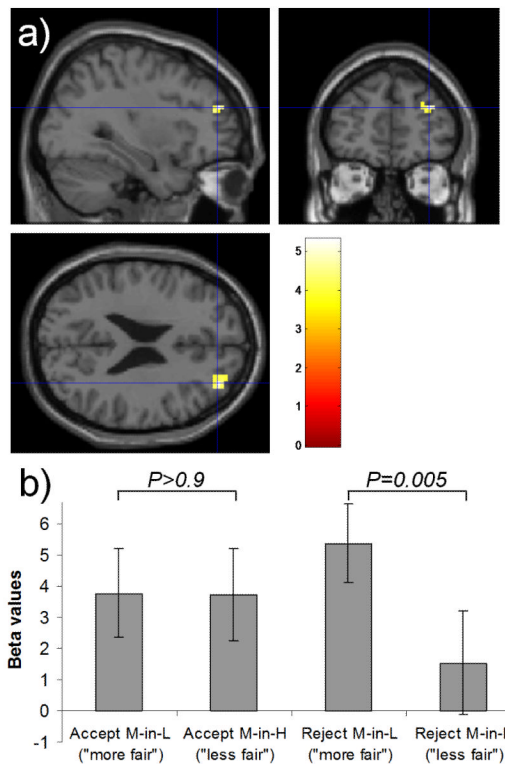


Figure 4. Interaction of choice and context in right dorsolateral prefrontal cortex

a) In our factorial analysis of the fMRI data there was an effect in right DLPFC for the interaction of choice (accept > reject) and context (M-in-H > M-in-L). This activation survived whole-brain cluster-level correction ($P < 0.05$ FWE corrected; threshold of $P < 0.005$ used to define the cluster) and is displayed at $P < 0.001$ (uncorrected) on slices through the peak voxel ($x=33, y=48, z=24$). **b)** To illustrate which differences drive the interaction we plot the regression coefficients (beta values \pm s.e.m.) for each condition versus baseline at the peak voxel for the interaction. Greater activation is seen for rejecting offers perceived as more fair than for those perceived as less fair ($t(15)=3.3, P=0.005$, two-tailed), with no difference during acceptance ($t(15)=0.04, P > 0.9$, two-tailed).

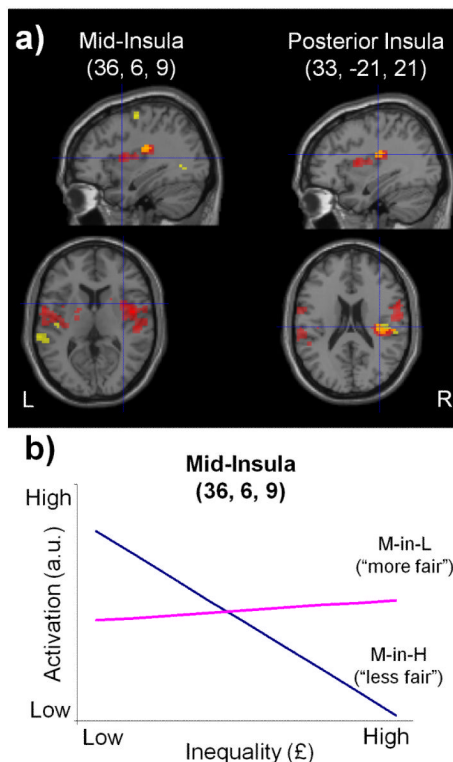


Figure 5. Neural responses to inequality and its interaction with social context in insular cortex. *a)* Sections showing activation related to processing of inequality in the insula. In yellow we show posterior insula is negatively correlated with inequality when collapsed across all trials (L, M-in-L, M-alone, M-in-H, H). In red are shown areas of differential activity in mid and posterior insula when contrasting responses to inequality in the more fair (M-in-L) versus the less fair (M-in-H) context. Areas of overlap are shown in orange. The right panels are at the peak of activation for the negative correlation, whilst the left panels show slices at the peak for the interaction. The activations including these peaks survive whole-brain cluster-level correction ($P < 0.05$ FWE corrected; threshold of $p < 0.005$ used to define the clusters) and are displayed at $p < 0.005$ (uncorrected, 10 voxel threshold). *b)* Visualisation of the effects driving the interaction of inequality with social context at the peak voxel in mid-insula ($x=36, y=6, z=9$). The slope of the regression line for activation (in arbitrary units) against inequality (£) is shown for each social context (M-in-H and M-in-L). The difference in slopes illustrates that inequality modulates insular activity only when offers are viewed more aversively (M-in-H). The slope of each regression line is the parameter estimate (beta) for each condition obtained from the general linear model used to analyse the fMRI data, averaged across subjects (M-in-H beta = -2.92 with s.e.m. = 0.64 ; M-in-L beta = 0.32 with s.e.m. = 0.63).

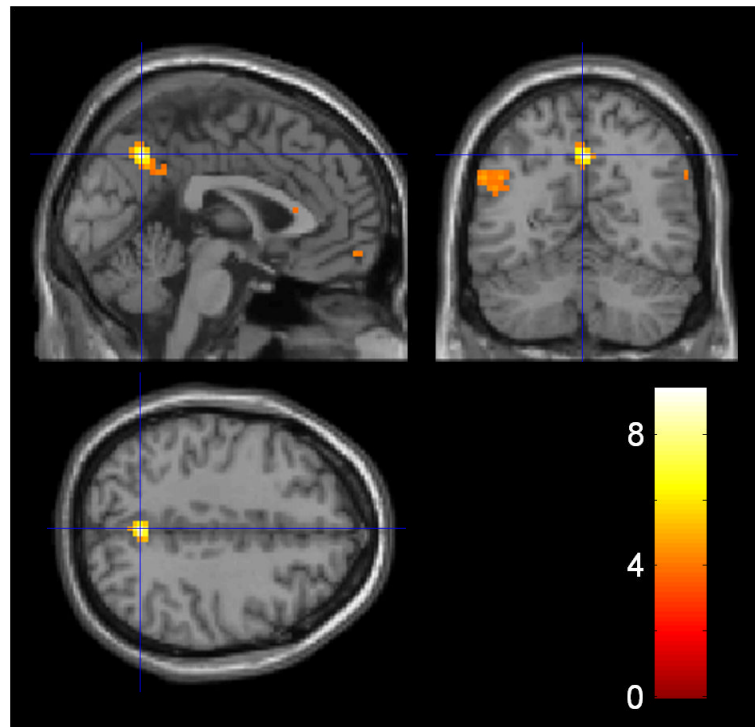


Figure 6. Inequality and envy.

For each subject we calculated an α (envy) parameter by fitting an inequality aversion model to behavioural data of the Responder. In this economic model, utility is the offer amount minus inequality (defined as amount to other minus amount to self) and inequality is weighted by α . Given their close relationship, the envy parameter was used as a second level covariate on the main effect of inequality. Areas highlighted are those that correlate between subjects with the degree to which the Responder cared about how much money the other person stood to gain relative to themselves. Activation surviving whole brain correction at the voxel level ($P < 0.05$ FWE) was seen in precuneus ($x=0, y=-60, z=42$) and L frontopolar cortex ($x=-18, y=60, z=18$), whilst activation surviving cluster level correction ($P < 0.05$ FWE corrected; threshold of $P < 0.005$ used to define the clusters) is seen in the L angular gyrus (inferior parietal cortex) ($x=-54, y=-60, z=30$). The data are displayed at $P < 0.001$ (uncorrected) at the peak voxel of the precuneus activation.

Table 1

Results using a factorial model for analysis of the fMRI data.

This table shows all activation that survived cluster level correction ($P < 0.05$ FWE corrected; threshold of $P < 0.005$ used to define the clusters) for the following contrasts: main effects of choice (collapsed across all trials; M, L and H); main effects of context (between the two matched conditions; M-in-L and M-in-H); and interactions between choice and context [$(\text{AccM-in-H} + \text{RejM-in-L}) - (\text{AccM-in-L} + \text{RejM-in-H})$]. For each cluster is shown: the three constituent peaks with the highest Z-scores; the number of voxels at $P < 0.005$ (uncorrected); and the P -value of the cluster after FWE correction across the whole brain. (SMA = Supplementary Motor Area; SFG = Superior Frontal Gyrus; MFG = Middle Frontal Gyrus).

Area	L/R	x	y	z	Z score	# vox, $p < 0.005$	Corr. p-value
<i>Choice (accept > reject): Main effect across all trial types</i>							
SMA and pre-SMA	L	-9	3	54	3.66	158	0.001
		-9	12	57	3.52		
SFG	R	24	15	54	3.65		
<i>Choice (reject > accept): Main effect across all trial types</i>							
Nil whole brain corrected.							
<i>Interaction of choice and social context</i>							
MFG	R	33	48	24	3.92	108	0.005
		27	45	15	3.54		
		15	54	30	3.44		
SFG	R	18	15	51	3.68	79	0.028
		24	9	51	3.60		
SMA	R	6	6	66	3.41		

Table 2

Results using a parametric model for analysis of the fMRI data.

This table shows all activation that survived cluster level correction ($P < 0.05$ FWE corrected; threshold of $P < 0.005$ used to define the clusters) for the following contrasts: main effects of offer amount (collapsed across all trials; M, L and H); main effects of inequality orthogonalised with respect to offer amount (firstly collapsed across all trials and secondly collapsed across the matched M-in-L and M-in-H trials); and the effects of context (M-in-L v. M-in-H) on each of these parametric regressors. For each cluster is shown: the three constituent peaks with the highest Z-scores; the number of voxels at $P < 0.005$ (uncorrected); and the P -value of the cluster after FWE correction across the whole brain. (SMA = Supplementary Motor Area; SFG = Superior Frontal Gyrus; MFG = Middle Frontal Gyrus; MCC = Middle Cingulate Cortex; STS = Superior Temporal Sulcus; MTG = Middle Temporal Gyrus; STG = Superior Temporal Gyrus).

Area	L/R	x	y	z	Z score	# vox. $p < 0.005$	Corr. p-value
Offer amount: Main effect across all trial types							
Postcentral gyrus	R	39	-30	45	4.91	536	<0.001
		33	-33	36	4.16		
SMA and pre-SMA	L	-9	3	51	4.41		
Offer amount: Interaction with social context (M-in-H > M-in-L)							
MFG	R	30	45	18	4.44	113	0.006
Inequality: Negative main effect across all trial types							
Posterior Insula	R	33	-21	21	3.73	114	0.006
		42	-24	24	3.43		
IPC	R	54	-18	3	3.38		
Inequality: Negative main effect, matched trials (M-in-H & M-in-L)							
Posterior Insula / STG	L	-45	-18	0	3.87	80	0.043
		-36	-12	3	3.25		
		-54	-12	6	3.24		
STS / MTG	R	66	-33	-6	3.55	137	0.002
		57	-30	6	3.49		
		63	-12	0	3.36		
Inequality: Interaction with social context (M-in-L > M-in-H)							
Mid-Insula / Rolandic operc.	R	48	-3	9	5.05	728	<0.001
		36	6	9	4.12		

Area	L/R	x	y	z	Z score	# vox. p<0.005	Corr. p-value
Precentral gyrus	R	60	3	18	4.00		
MCC	L	-12	-6	39	4.46	135	0.002
SMA / MCC	R	9	-27	51	3.89		
		9	-3	45	3.45		
Supramarginal gyrus	L	-51	-27	27	4.38	512	<0.001
Mid-Insula / Rolandic operc.	L	-48	-9	12	3.93		
Precentral gyrus	L	-60	0	9	3.91		