Blunted Neuroeconomic Loss Aversion in Schizophrenia

Currie, James; Waiter, Gordon D.; Johnston, Blair; Feltovich, Nick; Steele, J. Douglas

Published in:
Brain Research

DOI:
10.1016/j.brainres.2022.147957

Publication date:
2022

Blunted Neuroeconomic Loss Aversion in Schizophrenia

James Currie[1,2], Gordon D. Waiter[3], Blair Johnston[4], Nick Feltovich[5], J. Douglas Steele[6]

[1]Department of Psychiatry, Hamad Medical Corporation, Doha, Qatar

[2]School of Medicine, University of Aberdeen, UK

[3]Biomedical Imaging Centre, University of Aberdeen, UK

[4]MRI Physics, Department of Physics and Bioengineering, NHS Greater Glasgow and Clyde, UK

[5]Department of Economics, Monash University, Clayton VIC 3800, Australia

[6]Division of Imaging Science and Technology, School of Medicine, University of Dundee, UK

Correspondence

Prof. Douglas Steele
MD PhD MIPEM FRCPsych
Honorary Consultant Psychiatrist
School of Medicine, University of Dundee
Ninewells Hospital and Medical School
Dundee, UK, DD1 9SY
Tel: (44) (0)1382 383579
Email: d.steele@dundee.ac.uk

Abstract words: 248
Total words: 4441
Tables (2) + Figures (3)

Keywords: schizophrenia, RDoC, neuroeconomics, fMRI, loss aversion

ABSTRACT

Background: Abnormal social decision-making is prominent in schizophrenia. Antipsychotic medication often improves interpersonal functioning but this action is poorly understood. Neuroeconomic paradigms are an effective method of investigating social decision-making in psychiatric disorders that can be adapted for use with neuroimaging. Using a neuroeconomic approach, it has been shown that healthy humans reproducibly alter their behavior in different contexts, including exhibiting loss aversion: a higher sensitivity to loss outcomes compared to gains of the same magnitude.

Methods: Here, using a novel loss aversion task and fMRI, we tested three hypotheses: controls exhibiting normal behavioral loss aversion show changes in brain activity consistent with previous studies on healthy subjects; behavioral loss aversion is significantly reduced in schizophrenia and associated with abnormal activity in the same brain regions activated in controls during loss aversion behavior; and for the patient group alone, there is a significant correlation between increased psychotic symptoms, blunted loss aversion and abnormal brain activity. These hypotheses were tested in patients with schizophrenia and healthy controls using a loss aversion paradigm and fMRI.

Results: The results support the hypotheses, with patients exhibiting significantly blunted behavioral loss aversion compared to controls. Controls showed a robust loss aversion brain activation pattern in the medial temporal lobe, insula and dopaminergic-linked areas, which was blunted in schizophrenia.

Conclusions: Our results are consistent with blunted loss aversion being a reproducible feature of schizophrenia, likely due to abnormal dopaminergic and medial temporal lobe function, suggesting a route by which antipsychotics could influence interpersonal behavior.

1. Introduction

Psychiatric disorders are the leading cause of years of life lived with disability worldwide (Whiteford et al., 2013). Severe and enduring illnesses, including schizophrenia, are associated with a reduction in lifespan of about a decade (Chang et al., 2011). However, it is widely recognized that clinical practice in psychiatry has not fundamentally changed in over half a century (Stephan et al., 2016a; Stephan et al., 2016b). Better understanding of illness mechanisms is required to bridge the gap between the phenomenological and biological understanding of these illnesses (Montague et al., 2012).

Abnormalities in social interactions occur in many psychiatric disorders including schizophrenia. This is notable given increasing adoption of Research Domain Criteria (RDoC) which aims to link symptoms, including abnormalities of social processes, to specific brain systems (Insel et al., 2010). Suspicion and paranoia are common symptoms of schizophrenia associated with abnormal interpersonal behavior including social withdrawal. Antipsychotic medication often improves interpersonal functioning (Swartz et al., 2007) but the mechanism of action is poorly understood. Neuroeconomic paradigms (Glimcher, 2003) are an effective, but as yet rarely used, method of investigating social decision-making in psychiatric disorders, that can be adapted for use with neuroimaging (Hasler, 2012; Robson et al., 2020). A substantial body of evidence from decision-making studies indicates that healthy humans exhibit loss aversion, a higher sensitivity to losses compared to gains of the same magnitude (Kahneman and Tversky, 1979), with some studies reporting losses weighted twice as strongly as gains (Cachon and Camerer, 1996; Feltovich et al., 2012; Rydval and Ortmann, 2005; Tversky and Kahneman, 1992).

The dopamine hypothesis has been the leading theory of schizophrenia for more than forty years (Howes et al., 2017) and seminal work on healthy humans has linked loss aversion to dopamine function (Tom et al., 2007). Given this link and robust evidence for medial temporal lobe and striatal abnormalities in schizophrenia (e.g., van Erp et al., 2016), it is notable that behavioral economic studies have reported blunted behavioral loss aversion in schizophrenia (Brown et al., 2013; Tremeau et al., 2008)[1]. Recently, using the Iterated Prisoners' Dilemma (IPD) paradigm in a non-imaging behavioral study, we replicated blunted loss aversion in schizophrenia, reporting the magnitude of blunting correlated with Positive

[1] However, both Tremeau et al. (2008) and Brown et al. (2013) appear to have used proxy measures of loss aversion; see the discussion in the Introduction of our previous publication (Currie et al., 2017).

and Negative Syndrome Scale (PANSS) ratings of illness severity (Currie et al., 2017). Reduced loss aversion in schizophrenia could be related to abnormalities in the salience of gains and losses (Tremeau et al., 2008) and it has been argued that abnormal salience is related to positive psychotic symptoms and dopaminergic abnormalities (Kapur, 2003).

Loss aversion in healthy humans is associated with activation of the amygdala (Canessa et al., 2013; Chandrasekhar Pammi et al., 2015; De Martino et al., 2010; Sokol-Hessner et al., 2013), hippocampus (Canessa et al., 2013), and insula (Canessa et al., 2013; Sokol-Hessner et al., 2013). In loss aversion it has been reported that gains are associated with activation of dopaminergic-linked regions such as the striatum and losses with decreased activity in the same regions (Tom et al., 2007). Consistent with these findings, during successful instrumental reward learning, healthy humans exhibit robust activation of dopaminergic regions such as the ventral striatum, rostral-subgenual anterior cingulate and ventral tegmental area (Gradin et al., 2014; Johnston et al., 2015b; Rupprechter et al., 2020). In contrast aversive events, such as experienced loss during unsuccessful loss avoidance (different from loss aversion) in instrumental learning or aversive classical conditioning, are associated with activation of the amygdala (Buchel et al., 1998; Michely et al., 2020), insula and periaqueductal grey, and with deactivation of dopaminergic regions such as the ventral striatum (Gradin et al., 2014; Johnston et al., 2015b; Tolomeo et al., 2020). Structural and functional abnormalities in these brain regions have been reported in schizophrenia (Gradin et al., 2011; van Erp et al., 2016).

In the present study, as a test of loss aversion during fMRI, patients with schizophrenia and healthy controls each played two versions of the IPD: the ‘gain frame’ involving only payoffs with gains (Figure 1); and the ‘loss frame’ involving payoffs with both gains and losses, derived from the gain frame payoffs by subtracting a constant. As with our previous study (Currie et al., 2017), the difference in cooperation rates between the gain frame and the loss frame is a measure of loss aversion (see Materials and Methods, and game theoretical model, Supplementary Material).

We tested three hypotheses: (i) controls exhibiting healthy behavioral loss aversion show changes in brain activity, consistent with previous neuroimaging studies investigating loss aversion in healthy humans (Canessa et al., 2013; Sokol-Hessner et al., 2013; Tom et al., 2007); (ii) behavioral loss aversion is significantly reduced in schizophrenia (Brown et al., 2013; Currie et al., 2017; Tremeau et al., 2008) and this reduction is associated with abnormal activity in brain regions identified with healthy human loss aversion; and (iii) for the patient group alone, there is a significant correlation between increased psychotic symptoms and

blunted loss aversion and abnormal brain activity related to blunted loss aversion. On the basis of previous neuroeconomic studies on healthy humans, brain regions of particular interest were the medial temporal lobe (amygdala-hippocampal complex), insula and striatum (Canessa et al., 2013; Sokol-Hessner et al., 2013).

2. Results

2.1. Between-groups behavioral analyses and between-study replication

As expected, controls in the present study exhibited loss aversion, with cooperate choices in 40.3% of the gain frame compared to 35.6% in the loss frame which is significantly different (p = 0.03, n = 16), corresponding to a change in the frequency of cooperate choices (loss aversion score) of 4.7 percentage points; in contrast, patient choices were not significantly different between the two frames, 48.0% and 48.5% for the gain and loss frames respectively (n = 16, p = 0.66), corresponding to a loss aversion score of -0.5 percentage points, significantly lower than for controls (p = 0.046) (see Figure S3a, Supplementary Material). Probit tests for possible confounding variables including antipsychotic dosage did not significantly alter these findings (see Supplementary Material).

These behavioral results *replicate* findings from our previous independent behavioral-only study (Currie et al., 2017), with a change in frequency of cooperate choices (loss aversion score) of 8.3 percentage points for controls (n = 16, p = 0.01) and 0.5 percentage points for patients (n = 20, p = 0.35), which was significantly lower for patients than controls (p = 0.02) (see Figure S3b, Supplementary Material).

Pooling the behavioral data from the present study and our previous study, controls (32 in total) showed cooperate choices in 40.6% of the gain frame and 34.2% of the loss frame, significantly lower in the loss frame (p = 0.003, n = 32), corresponding to a mean control loss aversion score of 6.4 percentage points; in contrast, pooled patient data (36 patients in total) showed patient frequencies of cooperate choices that were not significantly different between the two frames, 43.1% and 43.0% for the gain and loss frames respectively (p = 0.49, n = 36), corresponding to a mean loss aversion score of 0.1 points, significantly lower than for controls (p = 0.005) (see Figure S3c, Supplementary Material).

2.2. Correlations between behavior and illness severity measures

The relationships between patient loss aversion scores and PANSS symptom scores were significant for PANSS total (R^2 = 0.48, p = 0.003) and positive symptom scores (R^2 = 0.64, p = 0.0002). This indicates that more severely ill patients exhibited more blunting of loss aversion behavior, particularly relating to psychotic symptoms. Probit tests of possible confounding variables did not significantly alter these findings (see Supplementary Material). These results *replicate* findings from our previous behavioral-only study (Currie et al., 2017),

which also found a significant correlation between patient behavioral loss aversion scores and PANSS total ($R^2 = 0.32$, $p = 0.01$ using least squares regression) and positive symptom scores ($p = 0.00001$ using Probit regression).

2.3. Neural loss aversion, within and between-groups analyses

For controls, a one group t-test using the loss aversion contrast revealed significant activations, as hypothesized, in the amygdala, hippocampus and parahippocampal gyrus, (26, -12, -24) and (-28, -22, -20) (Figure 2A). For patients, a one group t-test showed no significant activations (Figure 2B). In comparison, a between-groups analysis using a two-group t-test showed that loss aversion signals in patients were significantly blunted in the amygdala, hippocampus and parahippocampal gyrus, (24, -18, -22) and (-28, -22, -20) (Figure 2C). Results for controls, patients and between groups are also presented in Tables S3, S4 and S5 respectively (see Supplementary Material).

2.4. Neural loss aversion, correlations with illness severity

When the PANSS Positive score was entered as a regressor for the patient group alone, a significant correlation between PANSS Positive score and loss aversion was found in the amygdala, hippocampus and parahippocampal gyrus, (22, -8, -12) and (-14, -14, -20) (Figure 2D). This indicates that more severely ill patients exhibited more blunting of neural signals related to loss aversion, consistent with the between-groups analyses. Results are also presented in Table S6 (see Supplementary Material).

2.5. Correlation between behavioral loss aversion and neural loss aversion

The behavioral loss aversion measure was included in the neural loss aversion analyses as a covariate for controls and patients (see Tables S7 and S8, Supplementary Material). Region of interest analyses then compared patients and controls, using 10 mm diameter spherical regions of interest centered at maximally significant voxels for controls (voxels taken from Table S7, Supplementary Material). For controls there was a significant correlation between the behavioral loss aversion measure and the neural loss aversion imaging contrast in dopamine-rich areas: ventral tegmental area, medial temporal lobe (amygdala-hippocampal complex and parahippocampal gyrus), insula, anterior cingulate and paracingulate gyri, and striatum (see Table 2). In contrast, for patients a significant correlation was only found in the

striatum in those same regions of interest, while significant differences between controls and patients were found in the insula, right medial temporal lobe (amygdala-hippocampal complex and parahippocampal gyrus), and ventral tegmental area, but not the striatum, or anterior cingulate and paracingulate gyri (see Table 2).

These results for the ventral tegmental area and right medial temporal lobe (amygdala-hippocampal complex and parahippocampal gyrus) are also shown for patients and controls in Figure 3.

As is common in studies of schizophrenia, patients had a slightly lower average IQ score than controls and most were receiving antipsychotics. We therefore tested for a correlation between the loss aversion contrast and IQ, and separately for a correlation between the loss aversion contrast and chlorpromazine equivalents and did not find significant signals in the brain regions of interest.

2.6. Neural loss aversion, within-study replication

Consistency of the neural loss aversion imaging results was tested using machine learning cross-validation (within-study replication) approach (Johnston et al., 2015a; Johnston et al., 2015b). Diagnostic status was correctly predicted in 91% of participants (sensitivity 88%, specificity 94%, $\chi^2 = 18.1$, $p < 0.0001$), indicating a consistent pattern of abnormal brain activity in patients.

3. Discussion

Consistent with hypotheses, we found healthy controls who exhibited behavioral loss aversion showed changes in brain activity consistent with previous neuroimaging studies. Behavioral loss aversion was robustly reduced in schizophrenia and this was associated with reduced activity in the same brain regions. For the patient group alone there was a significant correlation between increased psychotic symptoms and blunted behavioral loss aversion, and between abnormal brain activity and blunted behavioral loss aversion.

In an influential study on the neural basis of loss aversion in decision-making, Tom and colleagues reported activation of dopaminergic regions including the striatum as gains increased, with losses represented by decreasing activity in gain-sensitive regions, and with neural loss aversion linked to targets of the dopamine system (Tom et al., 2007). They speculated that individual differences in behavioral and neural loss aversion may be caused by naturally occurring differences in dopamine function, suggesting neuroeconomic studies might shed light on neuropsychiatric disorders (Tom et al., 2007).

Tom and colleagues had also expected amygdala and insula activation due to negative emotional states, but instead found changes in activity in dopamine-rich regions (Tom et al., 2007). Later studies also found evidence linking neural loss aversion to dopamine-rich regions but additionally amygdala involvement (Canessa et al., 2013; Sokol-Hessner et al., 2013). A recent review has proposed two pathways contributing to loss aversion; a dopaminergic pathway from the ventral tegmental area for rewarding gain information, and a noradrenergic pathway from the amygdala processing aversive loss information, both converging on the striatum (Sokol-Hessner and Rutledge, 2018). Consistent with this suggestion, there is neuroimaging evidence for aversive information processing in healthy humans activating medial temporal lobe structures such as the amygdala (Buchel et al., 1998; Michely et al., 2020) and hippocampus (Gradin et al., 2014; Johnston et al., 2015b; Tolomeo et al., 2020). Our results for controls are consistent with these research findings.

There are reasons to expect loss aversion behavior may be abnormal in schizophrenia. An influential theory posits that psychosis occurs due to abnormal salience of internal and external representations (Kapur, 2003). This may result in an imbalance in the weighting given to losses and gains, and therefore disrupt loss aversion decision making. Using an instrumental reward learning paradigm, previously we reported blunted encoding of expected value in the medial temporal lobe of patients with schizophrenia (Gradin et al., 2011). Kraepelin speculated hallucinations and thought disorder were due to impairment of the

temporal lobes and since then many investigators have highlighted similarities between schizophrenia and schizophrenia-like psychoses caused by medial temporal lobe lesions (Arnold, 1997). A recent large meta-analysis of patients with schizophrenia and controls reported robust reductions in the volumes of the hippocampus, amygdala and accumbens (van Erp et al., 2016). In individuals at risk of developing psychosis, it has been proposed that medial temporal lobe structures such as the hippocampus are linked to dysregulation of dopamine and glutamate, leading to aberrant salience signaling, whereas disorganized speech and language impairments are linked to lateral temporal lobe dysfunction (Allen et al., 2019). In our study, patients exhibited blunted loss aversion behavior, and increased blunting of amygdala-hippocampal complex activity was associated with increased psychotic symptoms. This implies the neural basis for blunted loss aversion behavior in schizophrenia is abnormalities in the dopaminergic system and medial temporal lobes.

There are potential limitations to the study. Although we included a limited number of participants, the numbers are acceptable for a neuroimaging study (Friston, 2012), although independent replication of results in larger studies is always indicated. The study also recruited predominantly male participants which may limit generalizability to women with schizophrenia. These are avenues for future work. Our study has a number of strengths, including an in-depth analysis for potential confounding factors, such as learning effects, IQ and antipsychotic dosage, which did not alter the findings. Most notably, there is replication of previous literature results – blunted behavioral loss aversion in schizophrenia (Brown et al., 2013; Currie et al., 2017; Tremeau et al., 2008), and using a different paradigm, neural loss aversion findings in healthy subjects (Tom et al., 2007; Canessa et al., 2013; Sokol-Hessner et al., 2013) – with the novel finding of abnormal brain activity in dopaminergic and medial temporal lobe regions of schizophrenia patients linked to blunted loss aversion.

In conclusion, abnormalities in social interactions occur in many psychiatric disorders including schizophrenia. Using a neuroeconomic approach to test RDoC Social Process hypotheses, our results are consistent with blunted loss aversion behavior being a reproducible feature of schizophrenia, likely linked to abnormal dopaminergic and medial temporal lobe function, suggesting a route by which antipsychotics could influence interpersonal functioning. Further study of schizophrenia using neuroeconomics methods is indicated.

4. Materials and Methods

4.1. Participants

The study was approved by the local NHS Research Ethics Committee (REC ref: 09/S0802/114). Patients were recruited from Community Mental Health Teams (CMHTs) based at the Royal Cornhill Hospital, Aberdeen, UK. The present behavioral-neuroimaging study acquired data from thirty-two subjects comprising sixteen patients and sixteen controls. Our earlier independent non-imaging behavioral study (no patients or controls from this earlier study took part in the present neuroimaging study) acquired data from thirty-six subjects, comprising twenty patients and sixteen controls (Currie et al., 2017). This work was conducted as part of the PhD completed by JC at the University of Aberdeen.

For the present neuroimaging study, inclusion criteria for patients were a diagnosis of schizophrenia confirmed by two psychiatrists according to DSM-IV (First et al., 2002) and by using OPCRIT (McGuffin et al., 1991). Inclusion criteria for controls were no current or historical psychiatric illness, confirmed by the Mini-International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998). Exclusion criteria were contraindications for MRI, current or previous history of alcohol or substance misuse, significant head injury, difficulty understanding the game paradigm and inability to provide informed consent. Control and patient groups were matched based on age and sex.

Schizophrenia symptom severity was assessed using the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987), with ratings obtained for co-morbid symptoms (depression, anxiety, anhedonia), pre-morbid IQ, and dose of antipsychotic medications (Beck et al., 1996; Hamilton, 1960; Nelson and Wilson, 1991; Snaith et al., 1995; Spielberger, 1983). This is summarized in Table 1 (also Supplementary Material).

4.2. Iterated Prisoners' Dilemma paradigm

The original description of the Prisoners' Dilemma (Tucker and Straffin, 1983) involves two individuals, both charged with an offense and held so they cannot communicate. Each can betray by testifying against the other or cooperate by remaining silent. There are four payoff contingencies, depending on each participant's decision, with some outcomes better than others. The 'dilemma' lies in the fact that the optimal Nash equilibrium (Nash, 1950) for each individual is to betray, even though mutual cooperation leads to a better outcome for both, compared to mutual betrayal.

The IPD is a variation on the Prisoners' Dilemma and involves participants repeatedly playing over time, making choices with memory of previous encounters, which tends to result in cooperative strategies being more successful (Axelrod and Hamilton, 1981). Here we used two different versions of the IPD: a 'gain frame' which had all four outcomes as gains, and a 'loss frame' derived from the gain frame by subtracting a constant (5 points) from all outcomes such that both gain and loss outcomes were possible. Gain frame payoffs (+2, +4, +6, +8) are shown in Figure 1. The loss frame payoffs (-3, -1, +1, +3) were derived by subtracting a constant of 5 from all gain frame payoffs. Notably, this subtraction does not affect the Nash equilibrium, so should not affect rational decision-making. However, healthy individuals have been found to cooperate less in the loss frame compared to the gain frame; one possible explanation is that this an adaptive response conducive to individual survival since in times of plenty a betrayal or poor gamble will result in no worse than a "smaller dinner", while in times of scarcity a similar bad outcome could result in starvation (Currie et al., 2017). This difference in cooperation rates between gain and loss frames is a measure of loss aversion (see game theoretical model, Supplementary Material).

The concept of loss aversion is not inherently social, rather it was first outlined as a behavioral economic phenomenon in single-person settings, involving individual decisions with known outcomes and probabilities (Kahneman and Tversky, 1979), and demonstrated in the literature using gambling tasks (Kahneman and Tversky, 1979, Tversky and Kahneman, 1992, Tom et al., 2007, De Martino et al., 2010, Canessa et al., 2013, Sokol-Hessner et al., 2013, Chandrasekhar Pammi et al., 2015). However, with the exception of gambling, such single-person decision-making settings are extremely rare in the real world. Most real settings involve probabilities that are not specified, often due to strategic uncertainty from lack of information about what others are going to do. Thus, for loss aversion to be considered a valid behavioral economic construct, it should also be demonstrable in interactive social settings. Indeed a substantial strand of work has analyzed loss aversion in these more realistic interactive settings, including two-player games and multi-player contests such as auctions and tournaments (Feltovich, 2011; Currie et al., 2017; Chowdhury et al., 2018; Zheng et al., 2018; Rosato and Tymula, 2019). The IPD paradigm was specifically selected in this study of decision-making in schizophrenia due to its relative simplicity and intuitive mechanics (involving simple binary choices), both of which are desirable given the cognitive impairments associated with schizophrenia (Aylward et al., 1984), to minimize the potential for IQ to become a confounding factor. While the IPD involves a two-player interaction, the only difference between the gain and loss frames was a subtraction of a constant from the

payoffs of the gain frame to derive the payoffs of the loss frame, and this meant that the task remained a pure test of loss aversion, despite the social element.

4.3. Neuroimaging

During scanning, participants played 100 rounds of the IPD, 50 rounds in each of the gain and loss frames. To standardize the play received by each participant, participants played against a computer algorithm. Each round started with a prompt "Please make your choice" that began the 9-second choice stage, with a 5-second timer counting down, ending with the decision of the participant (Player 1) being displayed for the remainder of the 9 seconds. This was followed by a 5-second feedback stage, with the decision of the computer algorithm (Player 2) being displayed along with the resultant Player 1 payoff, or the payoff for Player 1 making no choice (Figure 1). Following this, there was presentation of a fixation point of varying duration ('jitter') before proceeding to the next round.

In each block of 50 rounds, the program played 25 rounds using a 'neutral' strategy, against which neither cooperation nor betrayal from the participant was systematically better, and 25 rounds using a 'cooperative' strategy against which a cooperative approach was advantageous. Counterbalancing of the gain and loss frames and the two computer strategies was used to control for order and learning effects, and changes in cooperative choices over time were analyzed in probit regression with rounds as a covariate, to further control for any learning effects (Supplementary Material).

Participants were told before playing that the other player was another human player although decision outcomes were programmed according to computer algorithm. They were not told of the game strategies that had been programmed. Participants agreed to maximize their overall score and told they would receive a payment proportional to their overall score in the range £5 to £15. Performance based payments in economics have long been considered important (Siegel and Goldstein, 1959) to encourage serious decision-making and reduce variance in the data (Camerer and Hogarth, 1999).

4.4. Behavioral analyses and between-study replication

Here we did two behavioral analyses using: (i) the behavioral data observed from the present study and (ii) an analysis pooling the behavioral data from the present study and our previous behavioral-only study (Currie et al., 2017) to maximise the power of the statistical

test of the hypothesis of blunted behavioral loss aversion in schizophrenia. For the patient group the relation between illness severity (measured by PANSS total and positive symptom scores) and loss aversion was tested using linear regression. Probit regression was used with cooperative choices over time as a dependent variable, allowing examination of the influence of possible confounding variables, such as IQ and antipsychotic dose, as well for any learning effects over time with rounds as a covariate (Supplementary Material).

4.5. Statistical analyses

Analyses were done using Stata (StataCorp), 'R' including the 'Cocor' package for comparing correlation coefficients (Diedenhofen and Musch, 2015) and SPM (Friston et al., 2007).

4.5.1. Image acquisition and pre-processing

Scanning was done at the Aberdeen Biomedical Imaging Centre, UK. Image data was acquired using a 3T Philips Achieva TX MRI scanner with an 8-channel phased-array head coil to acquire high-resolution gradient echo 3D volumetric images and fMRI data in the axial plane with a T2*-weighted single shot, gradient-echo, echo-planar pulse sequence using the following parameters: FOV, 24cm; TR/TE 2500/30 ms; flip angle, 78 degrees, slices, 30; slice thickness, 5mm; matrix, 96x96. For pre-processing, images were realigned to the first image in each time series and visually inspected to ensure no head movement greater than voxel dimensions. The mean realigned image for each participant was used to determine spatial normalization transformations to the SPM template, then these transformations were applied to each image in the time series. Smoothing was done using an 8-mm Gaussian kernel.

4.5.2. Neuroimaging analyses and within-study replication

An event-related random-effects design was implemented using SPM. For first-level analyses, events were modelled as truncated delta functions at times when payoff outcomes from the decisions by the participant and computer player (cooperation or betrayal) were revealed. Six motion realignment terms were included as covariates of no interest in the design matrix to account for residual movement artefacts not removed by pre-processing realignment. For neural loss aversion, individual contrast images were computed for the

contrast [cooperation - betrayal], loss frame versus gain frame. Whilst for loss avoidance [loss>neutral] may be the appropriate contrast (e.g., Johnston et al., 2015b; Waltz et al., 2018), for loss aversion [loss>gain] is the appropriate contrast as, by definition, loss aversion involves comparing losses and gains, with greater weighting given to losses.

First level contrast images were used to test hypotheses in second-level tests of within-group activations using one-group t-tests. Hypothesized between-group differences were tested using two-group t-tests. For all analyses, regions were considered significant at a whole brain $p < 0.05$ cluster level, achieved by a simultaneous requirement for a voxel threshold of $p < 0.005$ and a minimum cluster size of 40 contiguous voxels, with these parameters determined using a popular Monte Carlo method (Slotnick et al., 2003).

To test the consistency of our neural loss aversion results, we used a combination of feature selection, machine learning and within-study replication (cross-validation) to make unbiased inferences about each individual's diagnostic category (schizophrenia vs. controls) as described in previous work (Johnston et al., 2015a; Johnston et al., 2015b). Briefly, feature selection involved automated identification of brain regions, including only data from the training set and a Support Vector Machine (Schwaighofer) approach was used with each participant's diagnostic category predicted using leave one out cross-validation (LOOCV). This involves leaving one participant out, performing feature selection and training a predictive model on the remaining participants' data, then predicting the diagnostic category of the left-out participant. The process was repeated until the category for every participant had been predicted allowing estimation of accuracy, sensitivity and specificity. The feature selection in this machine-learning analysis used the individual contrast images for neural loss aversion, from the above-mentioned contrast [cooperation - betrayal], loss frame versus gain frame.

Acknowledgements

We thank all participants for their time and effort in taking part in this study and radiographers for their invaluable work in the MRI procedures. This study was supported by funding from the Millar-McKenzie Trust and the BMA Margaret Temple grant.

Declaration of Competing Interest

The authors have declared there are no conflicts of interest.

References

Allen, P., Moore, H., Corcoran, C.M., Gilleen, J., Kozhuharova, P., Reichenberg, A., Malaspina, D., 2019. Emerging Temporal Lobe Dysfunction in People at Clinical High Risk for Psychosis. Front Psychiatry 10, 298.

Arnold, S.E., 1997. The medial temporal lobe in schizophrenia. J Neuropsychiatry Clin Neurosci 9, 460-470.

Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390-1396.

Aylward, E., Walker, E., Bettes, B., 1984. Intelligence in schizophrenia: meta-analysis of the research. Schizophr Bull 10, 430-459.

Beck, A.T., Steer, R.A., Ball, R., Ranieri, W., 1996. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. J Pers Assess 67, 588-597.

Brown, J.K., Waltz, J.A., Strauss, G.P., McMahon, R.P., Frank, M.J., Gold, J.M., 2013. Hypothetical decision making in schizophrenia: the role of expected value computation and "irrational" biases. Psychiatry Res 209, 142-149.

Buchel, C., Morris, J., Dolan, R.J., Friston, K.J., 1998. Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 20, 947-957.

Cachon, G., Camerer, C., 1996. Loss-Avoidance and Forward Induction in Experimental Coordination Games. 111 1, 165-194.

Camerer, C.F., Hogarth, R.M., 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework. Journal of Risk and Uncertainty 19, 7-42.

Canessa, N., Crespi, C., Motterlini, M., Baud-Bovy, G., Chierchia, G., Pantaleo, G., Tettamanti, M., Cappa, S.F., 2013. The functional and structural neural basis of individual differences in loss aversion. J Neurosci 33, 14307-14317.

Chandrasekhar Pammi, V.S., Pillai Geethabhavan Rajesh, P., Kesavadas, C., Rappai Mary, P., Seema, S., Radhakrishnan, A., Sitaram, R., 2015. Neural loss aversion differences between depression patients and healthy individuals: A functional MRI investigation. Neuroradiol J 28, 97-105.

Chang, C.K., Hayes, R.D., Perera, G., Broadbent, M.T., Fernandes, A.C., Lee, W.E., Hotopf, M., Stewart, R., 2011. Life expectancy at birth for people with serious mental illness and other major disorders from a secondary mental health care case register in London. PLoS One 6, e19590.

Chowdhury, S.M., Jeon, J.Y., Ramalingam, A., 2018. Property Rights and Loss Aversion in Contests. Economic Inquiry 56, 1492-1511.

Currie, J., Buruju, D., Perrin, J.S., Reid, I.C., Steele, J.D., Feltovich, N., 2017. Schizophrenia illness severity is associated with reduced loss aversion. Brain Res 1664, 9-16.

De Martino, B., Camerer, C.F., Adolphs, R., 2010. Amygdala damage eliminates monetary loss aversion. Proc Natl Acad Sci U S A 107, 3788-3792.

Diedenhofen, B., Musch, J., 2015. cocor: a comprehensive solution for the statistical comparison of correlations. PLoS One 10, e0121945.

Feltovich, N., 2011. The effect of subtracting a constant from all payoffs in a hawk-dove game: experimental evidence of loss aversion in strategic behavior. Southern Economic Journal 77, 814-826.

Feltovich, N., Iwasaki, A., Oda, S.H., 2012. Payoff levels, loss avoidance and equilibrium selection in games with multiple equilibria: an experimental study. Economic Inquiry 50, 932-952.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 2002. Structured clinical interview for DSM-IV-TR axis I disorders, Research Version, Patient Edition. New York State Psychiatric Institute, New York.

Friston, K., 2012. Ten ironic rules for non-statistical reviewers. Neuroimage 61, 1300-1310.

Friston, K.J., Ashburner, J.T., Kiebel, S.J., Nichols, T.E., Penny, W.D., 2007. Statistical Parametric Mapping. Academic Press, London.

Glimcher, P.W., 2003. Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics. MIT Press, Cambridge, Massachusetts.

Gradin, V.B., Baldacchino, A., Balfour, D., Matthews, K., Steele, J.D., 2014. Abnormal brain activity during a reward and loss task in opiate-dependent patients receiving methadone maintenance therapy. Neuropsychopharmacology 39, 885-894.

Gradin, V.B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., Steele, J.D., 2011. Expected value and prediction error abnormalities in depression and schizophrenia. Brain 134, 1751-1764.

Hamilton, M., 1960. A rating scale for depression. J Neurol Neurosurg Psychiatry 23, 56-62.

Hasler, G., 2012. Can the neuroeconomics revolution revolutionize psychiatry? Neuroscience and Biobehavioral Reviews 36, 64-78.

Howes, O.D., McCutcheon, R., Owen, M.J., Murray, R.M., 2017. The Role of Genes, Stress, and Dopamine in the Development of Schizophrenia. Biol Psychiatry 81, 9-20.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., Sanislow, C., Wang, P., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 167, 748-751.

Johnston, B.A., Steele, J.D., Tolomeo, S., Christmas, D., Matthews, K., 2015a. Structural MRI-Based Predictions in Patients with Treatment-Refractory Depression (TRD). PLoS One 10, e0132958.

Johnston, B.A., Tolomeo, S., Gradin, V., Christmas, D., Matthews, K., Steele, J.D., 2015b. Failure of hippocampal deactivation during loss events in treatment-resistant depression. Brain 138, 2766-2776.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica 47, 263-292.

Kapur, S., 2003. Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. American Journal of Psychiatry 160, 13-23.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 13, 261-276.

McGuffin, P., Farmer, A., Harvey, I., 1991. A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. Arch Gen Psychiatry 48, 764-770.

Michely, J., Rigoli, F., Rutledge, R.B., Hauser, T.U., Dolan, R.J., 2020. Distinct Processing of Aversive Experience in Amygdala Subregions. Biol Psychiatry Cogn Neurosci Neuroimaging 5, 291-300.

Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational psychiatry. Trends Cogn Sci 16, 72-80.

Nash, J.F., 1950. Equilibrium Points in N-Person Games. Proc Natl Acad Sci U S A 36, 48-49.

Nelson, H.E., Wilson, J.R., 1991. The revised national adult reading test - test manual. NFER-Wilson, Winsor.

Robson, S.E., Repetto, L., Gountouna, V.E., Nicodemus, K.K., 2020. A review of neuroeconomic gameplay in psychiatric disorders. Mol Psychiatry 25, 67-81.

Rosato, A., Tymula, A.A., 2019. Loss aversion and competition in Vickrey auctions: Money ain't no good. Games and Economic Behavior 115, 188-208.

Rupprechter, S., Romaniuk, L., Series, P., Hirose, Y., Hawkins, E., Sandu, A.L., Waiter, G.D., McNeil, C.J., Shen, X., Harris, M.A., Campbell, A., Porteous, D., Macfarlane, J.A.,

Lawrie, S.M., Murray, A.D., Delgado, M.R., McIntosh, A.M., Whalley, H.C., Steele, J.D., 2020. Blunted medial prefrontal cortico-limbic reward-related effective connectivity and depression. Brain 143, 1946-1956.

Rydval, O., Ortmann, A., 2005. Loss avoidance as selection principle: evidence from simple stag-hunt games. Economics Letters 88, 101-107.

Schwaighofer, A., SVM toolbox.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. The Journal of Clinical Psychiatry 59 Suppl 20, 22-33;quiz 34-57.

Siegel, S., Goldstein, D.A., 1959. Decision making behavior in a two-choice uncertain outcome situation. J Exp Psychol 57, 37-42.

Slotnick, S.D., Moo, L.R., Segal, J.B., Hart, J., Jr., 2003. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. Brain research. Cognitive brain research. 17, 75-82.

Snaith, R.P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., Trigwell, P., 1995. A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. Br J Psychiatry 167, 99-103.

Sokol-Hessner, P., Camerer, C.F., Phelps, E.A., 2013. Emotion regulation reduces loss aversion and decreases amygdala responses to losses. Soc Cogn Affect Neurosci 8, 341-350.

Sokol-Hessner, P., Rutledge, R.B., 2018. The Psychological and Neural Basis of Loss Aversion. Current Directions in Psychological Science 28, 20-27.

Spielberger, C.D., 1983. State-trait anxiety inventory for adults, STAI Form Y-1. Mind Garden, Palo-Alto.

Stephan, K.E., Bach, D.R., Fletcher, P.C., Flint, J., Frank, M.J., Friston, K.J., Heinz, A., Huys, Q.J.M., Owen, M.J., Binder, E.B., Dayan, P., Johnstone, E.C., Meyer-Lindenberg, A., Montague, P.R., Schnyder, U., Wang, X.J., Breakspear, M., 2016a. Charting the landscape of priority problems in psychiatry, part 1: classification and diagnosis. Lancet Psychiatry 3, 77-83.

Stephan, K.E., Binder, E.B., Breakspear, M., Dayan, P., Johnstone, E.C., Meyer-Lindenberg, A., Schnyder, U., Wang, X.J., Bach, D.R., Fletcher, P.C., Flint, J., Frank, M.J., Heinz, A., Huys, Q.J.M., Montague, P.R., Owen, M.J., Friston, K.J., 2016b. Charting the

landscape of priority problems in psychiatry, part 2: pathogenesis and aetiology. Lancet Psychiatry 3, 84-90.

Swartz, M.S., Perkins, D.O., Stroup, T.S., Davis, S.M., Capuano, G., Rosenheck, R.A., Reimherr, F., McGee, M.F., Keefe, R.S., McEvoy, J.P., Hsiao, J.K., Lieberman, J.A., 2007. Effects of antipsychotic medications on psychosocial functioning in patients with chronic schizophrenia: findings from the NIMH CATIE study. Am J Psychiatry 164, 428-436.

Tolomeo, S., Macfarlane, J.A., Baldacchino, A., Koob, G.F., Steele, J.D., 2020. Alcohol Binge Drinking: Negative and Positive Valence System Abnormalities. Biol Psychiatry Cogn Neurosci Neuroimaging.

Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. Science 315, 515-518.

Tremeau, F., Brady, M., Saccente, E., Moreno, A., Epstein, H., Citrome, L., Malaspina, D., Javitt, D., 2008. Loss aversion in schizophrenia. Schizophr Res 103, 121-128.

Tucker, A.W., Straffin, P.D., 1983. The Mathematics of Tucker: A Sampler. The Two-Year College Mathematics Journal 14, 228-232.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297-323.

van Erp, T.G., Hibar, D.P., Rasmussen, J.M., Glahn, D.C., Pearlson, G.D., Andreassen, O.A., Agartz, I., Westlye, L.T., Haukvik, U.K., Dale, A.M., Melle, I., Hartberg, C.B., Gruber, O., Kraemer, B., Zilles, D., Donohoe, G., Kelly, S., McDonald, C., Morris, D.W., Cannon, D.M., Corvin, A., Machielsen, M.W., Koenders, L., de Haan, L., Veltman, D.J., Satterthwaite, T.D., Wolf, D.H., Gur, R.C., Gur, R.E., Potkin, S.G., Mathalon, D.H., Mueller, B.A., Preda, A., Macciardi, F., Ehrlich, S., Walton, E., Hass, J., Calhoun, V.D., Bockholt, H.J., Sponheim, S.R., Shoemaker, J.M., van Haren, N.E., Pol, H.E., Ophoff, R.A., Kahn, R.S., Roiz-Santianez, R., Crespo-Facorro, B., Wang, L., Alpert, K.I., Jonsson, E.G., Dimitrova, R., Bois, C., Whalley, H.C., McIntosh, A.M., Lawrie, S.M., Hashimoto, R., Thompson, P.M., Turner, J.A., 2016. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. Mol Psychiatry 21, 585.

Waltz, J.A., Xu, Z., Brown, E.C., Ruiz, R.R., Frank, M.J., Gold, J.M., 2018. Motivational Deficits in Schizophrenia Are Associated With Reduced Differentiation Between Gain and Loss-Avoidance Feedback in the Striatum. Biol Psychiatry Cogn Neurosci Neuroimaging 3, 239-247.

Whiteford, H.A., Degenhardt, L., Rehm, J., Baxter, A.J., Ferrari, A.J., Erskine, H.E., Charlson, F.J., Norman, R.E., Flaxman, A.D., Johns, N., Burstein, R., Murray, C.J., Vos, T., 2013. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 382, 1575-1586.

Zheng, H., Xu, B., Hao, L., Lin, Z., Te'Eni, D., Katsamakas, E., 2018. Reversed loss aversion in crowdsourcing contest. European Journal of Information Systems 27, 434-448.

Table 1 Demographics and clinical details

Mean	Patients	Controls
Age in years	42.5 (SD 12.8)	47.0 (SD 8.5)
Males (percentage)	94% (15/16)	94% (15/16)
PANSS Positive	19.4 (SD 4.5)	-
PANSS Negative	23.2 (SD 4.1)	-
PANSS General	41.3 (SD 7.7)	-
PANSS Total	83.9 (SD 14.0)	-
Hamilton Rating Scale for Depression	8.3 (SD 5.1)	0.3 (SD 0.6)
Beck Depression Inventory	17.1 (SD 12.7)	2.4 (SD 3.6)
Snaith-Hamilton Pleasure Scale	1.8 (SD 2.5)	0.4 (SD 0.6)
Spielberger State-Trait Anxiety Inventory	40.8 (SD 11.2)	31.4 (SD 7.2)
National Adult Reading Test IQ	110.6 (SD 10.9)	114.1 (SD 5.7)
Antipsychotic dose, mg (chlorpromazine-equivalent)	618.3 (SD 305.6)	-

Table 2 Behavioral and neural loss aversion, controls vs. patients

Behavioral and neural loss aversion	MNI coordinates			Controls		Patients		Difference between groups
Brain region	**x**	**y**	**z**	***r* value**	***p* value**	***r* value**	***p* value**	***p* value**
Ventral Tegmental Area	0	-16	-8	0.720	0.002	0.084	0.758	0.036
Amygdala, hippocampus, and parahippocampal gyrus	24	0	-22	0.724	0.002	-0.326	0.218	0.001
Insula	-38	12	-10	0.664	0.005	0.017	0.950	0.046
Anterior cingulate and paracingulate gyri	14	30	16	0.651	0.006	0.119	0.660	0.094
Dorsal striatum (caudate)	8	20	8	0.595	0.015	0.548	0.028	0.858
Insula	36	10	-14	0.595	0.015	-0.137	0.614	0.036
Amygdala, hippocampus, and parahippocampal gyrus	-30	-6	-24	0.571	0.021	0.279	0.296	0.356

Figure 1 Loss aversion task. Each round started with a prompt for Player 1 (patient or control) to cooperate or betray, followed by the outcome of Player 2 (computer algorithm) being revealed to the participant, or the outcome from making no choice. Gain frame payoffs are shown (+2, +4, +6, +8). The loss frame was derived by subtracting a constant of 5 from all gain frame payoffs (-3, -1, +1, +3).

Figure 2 Loss aversion contrast activations in the amygdala-hippocampal complex. (A) controls, (B) absence of signal in patients, (C) contrast controls > patients, and (D) correlations with PANSS Positive ratings for the patients alone. All brain regions significant at $p < 0.05$.

Figure 3 Correlations between behavioral loss aversion measure and loss aversion contrast. (A) Controls with region of interest volume centered at the maximally significant voxels in the amygdala-hippocampal complex (24, 0, -22), showing significantly different loss aversion correlations for controls compared to patients (B). (C) Controls with region of interest volume centered at the maximally significant voxels in the ventral tegmental area (0, -16, -8), showing significantly different loss aversion correlations for controls compared to patients (D). Correlation coefficients and *p*-values for these two areas are presented in Table 2.