Published in final edited form as:

Neuropsychologia. 2015 May ; 71: 165–172. doi:10.1016/j.neuropsychologia.2015.04.003.

Damage to the insula is associated with abnormal interpersonal trust

Amy M. Belfi^{a,b}, Timothy R. Koscik^c, and Daniel Tranel^{a,b,d}

^aInterdisciplinary Graduate Program in Neuroscience, University of Iowa, 356 MRC, Iowa City, Iowa, USA

^bDepartment of Neurology, University of Iowa College of Medicine, 2155 RCP, Iowa City, Iowa, USA

^cDepartment of Psychology, University of Toronto, 100 St. George Street, Toronto, ON, Canada

^dDepartment of Psychology, University of Iowa, E11 Seashore Hall, Iowa City, Iowa, USA

Abstract

Reciprocal trust is a crucial component of cooperative, mutually beneficial social relationships. Previous research using tasks that require judging and developing interpersonal trust has suggested that the insula may be an important brain region underlying these processes (King-Casas et al., 2008). Here, using a neuropsychological approach, we investigated the role of the insula in reciprocal trust during the Trust Game (TG), an interpersonal economic exchange. Consistent with previous research, we found that neurologically normal adults reciprocate trust in kind, i.e., they increase trust in response to increases from their partners, and decrease trust in response to decreases. In contrast, individuals with damage to the insula displayed abnormal expressions of trust. Specifically, these individuals behaved benevolently (expressing misplaced trust) when playing the role of investor, and malevolently (violating their partner's trust) when playing the role of the trustee. Our findings lend further support to the idea that the insula is important for expressing normal interpersonal trust, perhaps because the insula helps to recognize risk during decision-making and to identify social norm violations.

Keywords

Insula; brain lesion; trust; decision-making; social cognition

© 2015 Published by Elsevier Ltd.

Corresponding Author: Amy Belfi, University of Iowa Hospitals and Clinics, Department of Neurology, 200 Hawkins Drive, 2155 RCP, Iowa City, Iowa 52242, Phone: 319-384-5088, amy-belfi@uiowa.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

The ability to form cooperative, mutually beneficial relationships with others is an integral part of human life. Establishing reciprocal relationships can improve an individual's success in their career, family, and social life. Moreover, overall quality of life can be greatly affected by one's ability to work cooperatively with others. To establish cooperative relationships, individuals engage in repeated interactions with one another. Through these interactions they develop an understanding of whether their partner is trustworthy, which informs further interactions. A partner who reciprocates one's trust can be trusted in the future. However, an inability to reciprocate appropriately can result in adverse outcomes. For example, if an individual trusts an untrustworthy partner, that individual might get exploited. Conversely, if they do not trust a trustworthy partner, they may jeopardize a potentially beneficial relationship.

Successfully engaging in these types of interpersonal exchanges involves many social and emotional processes. During interpersonal exchanges, individuals often must accept rejection, celebrate success, judge the trustworthiness of others, and assess whether their behaviors violate social norms. A brain region that has been strongly implicated in these types of processes is the insular cortex (insula). Insula activity is associated with processes ranging from risk tracking (Preuschoff, Quartz, & Bossaerts, 2008) to social emotions (Singer et al., 2006) to judgments of trustworthiness (King-Casas et al., 2008; Todorov, Baron, & Oosterhof, 2008; Winston, Strange, O'Doherty, & Dolan, 2002). Here, we aim to elucidate the role of the insula in the process of interpersonal trust by studying individuals with damage to this area.

Historically, the insula has been considered an important structure for negatively-valenced emotional states, including pain (Ostrowsky et al., 2002), distress (Kross, Egner, Ochsner, Hirsch, & Downey, 2007), and disgust (Phillips et al., 1997). However, the insula is also involved in emotions and feelings of both positive and negative valence (A. Damasio et al., 2000). For example, individuals with insula lesions show reduced arousal in response to both unpleasant and pleasant visual stimuli, suggesting a role for the insula in affective processes regardless of valence (Berntson et al., 2011). Additionally, the insula is important for social emotions (for a review see Lamm & Singer, 2010), including empathy (Singer et al., 2006) and social rejection (Eisenberger, Lieberman, & Williams, 2003). In line with its involvement in social emotions, the insula is also important for social evaluations. For example, increased insula activation is observed in response to viewing untrustworthy faces (Todorov et al., 2008; Winston et al., 2002). However, increased insula activation is also observed in response to faces of individuals who performed cooperatively in a partnered game (Singer, Kiebel, Winston, Dolan, & Frith, 2004). These findings may point to a general role of the insula in interpersonal trust and distrust, regardless of the direction of this value.

An additional way the insula processes emotions is through its involvement in interoception and perception of feeling states (Craig, 2002). Conscious experience of interoception can influence decision-making (Bechara & Damasio, 2005; A. Damasio, 1996). Notably, individuals with damage to the insula show abnormal decision-making under uncertainty

and risk (Bar-On, Tranel, Denburg, & Bechara, 2003; Clark et al., 2008). These findings support a critical role for the insula, and the emotional processes it supports, in decision-making.

Consistent with the idea that emotions influence decision-making, it is no surprise that insula activity is implicated in many types of decision-making tasks. Functional neuroimaging research indicates insula involvement in economic decision-making tasks including the Prisoner's Dilemma (Singer et al., 2004), the Ultimatum Game (Sanfey, Rilling, & Aronson, 2003), the Cups Task (Xue, Lu, Levin, & Bechara, 2010), and the Iowa Gambling Task (d'Acremont, Lu, Li, Van der Linden, & Bechara, 2009; Krawitz, Fukunaga, & Brown, 2010). The results from these studies suggest that the insula is involved in risk, risk prediction error (Bossaerts, 2010; Preuschoff et al., 2008), and uncertainty (Critchley, Mathias, & Dolan, 2001). Similarly, in nonhuman primates (*Macaca mulatta*), the insula encodes expectations of rewards (Mizuhiki, Richmond, & Shidara, 2012).

Given the involvement of the insula in emotions, social judgments, and economic decision-making, we expect that the insula will also play a critical role in tasks that integrate these processes, such as decision-making during a social exchange. The Trust Game (TG) is an economic decision-making task that incorporates a social aspect. The TG is an economic exchange between two players: one player invests their money with another player, and the other player can choose to reciprocate this investment. The amount of money earned by one player is partially dependent upon the reciprocation from the other player. Therefore, by deciding how to reciprocate their investments, players can exhibit trust and/or distrust in their partners. Mutual trust between the two players results in successful economic exchanges (increased profits) for both parties, whereas misplaced trust or distrust results in unsuccessful exchanges. The TG involves multiple aspects of decision-making typically associated with the insula, including assessments of risk, uncertainty, and appropriate social behavior.

Insula activity is often found during the TG. Using a multi-round, multiplayer version of the TG, King-Casas and colleagues (2008) observed insula activation when participants both received and paid small (as opposed to large) investments. They concluded that the insula responds to violations of social norms, i.e., small investments, during these interpersonal exchanges. Using a one-shot version of the TG, van den Bos, van Dijk, Westenberg, Rombouts, & Crone (2009) found greater insula activation when players received small investments from their partners, but chose to reciprocate (as opposed to defect), which suggests that insula activity represents social norm violations. Consistent with the involvement of the insula in interoception, increased insula activation has been observed during TG performance when subjects were physically cold, compared to warm (Kang, Williams, Clark, Gray, & Bargh, 2011). However, other neuroimaging studies of TG behavior have not found insula activation (King-Casas et al., 2005; Krueger et al., 2007).

In the current study, we sought to investigate the role of the insula for normal performance on the TG by studying individuals with lesions to this structure. We used a multi-round, multiplayer Trust Game (TG) to investigate the performance of individuals with lesions to the insula, and contrasted their behavior to that of individuals with lesions outside the insula

(brain damaged comparison participants) and individuals with no lesions (healthy comparison participants). We predicted that damage to the insula would be associated with abnormal behavior, i.e., misplaced trust and distrust, on the TG.

2. Method

2.1. Participants

Participants were grouped based on lesion location and included individuals with damage to the insula (INS; $N = 11$, 7 left hemisphere, 4 right hemisphere; 7 men, 4 women; see Figure 1), brain-damaged comparison participants with focal brain damage outside the insula (and also excluding the amygdala and medial prefrontal cortex, given the potential importance of these structures for TG performance) (BDC; $N = 15$, 9 left hemisphere, 5 right hemisphere, 1 bilateral; 5 men, 10 women; see Figure 1), and neurologically normal comparison participants (NC; $N = 31$, 17 men, 14 women). Brain-damaged participants were recruited from the Patient Registry of the Division of Behavioral Neurology and Cognitive Neuroscience of the Department of Neurology at the University of Iowa. NCs were recruited through local advertisements. Participants were screened to ensure that none had general intellectual impairment (as determined by Wechsler Adult Intelligence Scale; Wechsler, 2008). Full-scale IQ (FSIQ) in the NC group was estimated using the Wechsler Test of Adult Reading (Wechsler, 2001). Although the NC group has a modestly higher FSIQ numerically, there were no significant differences between groups on demographic attributes, including age [$F(2, 54) = 0.47$, $p = 0.62$], education [$F(2, 54) = 0.54$, $p = 0.58$], and FSIQ [$F(2, 43) = 2.42$, $p = 0.10$] (Table 1).

Brain-damaged participants have been extensively characterized both neuroanatomically and neuropsychologically under the standard protocols of the Benton Neuropsychology Laboratory and the Laboratory of Brain Imaging and Cognitive Neuroscience (Tranel, 2007). As noted, individuals with damage to the amygdala were excluded, as damage to this region may affect performance on the trust game (Koscik et al., 2011). Additionally, individuals were excluded if they scored above cutoff levels for mild depression using the Beck Depression Inventory-II (13). Lesion etiologies differed slightly between groups. For the INS group, all individuals had lesions due to stroke. For the BDC group, the etiologies were: stroke ($n = 12$), benign tumor (meningioma) resection ($n = 2$), and resection for medically-intractable epilepsy ($n = 1$). The INS and BDC groups did not significantly differ in lesion chronicity, [$t(12.60) = 0.62$, $p = 0.54$], or on most neuropsychological variables, including Performance IQ [$t(23) = -0.45$, $p = 0.96$], Verbal IQ [$t(22) = -0.16$, $p = 0.87$], Beck Anxiety Inventory [$t(14) = 0.05$, $p = 0.96$], Beck Depression Inventory-II [$t(24) = 1.49$, $p = 0.14$], Rey Auditory-Verbal Learning Test 30-minute recognition [$t(10.62) = -1.83$, $p = 0.13$], Wisconsin Card Sorting Task number of categories completed [$t(21) = -0.30$, $p = 0.76$], Facial Discrimination Test [$t(22) = -0.63$, $p = .50$], Controlled Oral Word Association Test [$t(23) = -1.25$, $p = 0.22$], and Boston Naming Test [$t(12.05) = -1.20$, $p = 0.29$] (Table 2). All participants provided informed consent before participating in the experiment, following the protocol of the University of Iowa Institutional Review Board.

2.2 Lesion Analysis

Neuroanatomical analysis was based on MR or CT images obtained in the chronic epoch of recovery. Each lesion was reconstructed using Brainvox (Frank, Damasio, & Grabowski, 1997) and manually warped to a normal template brain using the MAP-3 technique (H. Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Fiez, Damasio, & Grabowski, 2000). After manual transfer to the normal template space, the template brain was warped to the MNI152 standard 1 mm T1-weighted atlas (Collins, Neelin, Peters, & Evans, 1994; Evans, Dai, Collins, Neelin, & Marrett, 1991; Mazziotta et al., 2001) using BRAINSDemonWarp (Johnson & Zhao, 2009). This transform, from the lesion template to the MNI152 template, was then applied to each of the lesion maps. Lesion maps were then used to create overlap maps, calculated by the sum of n lesions overlapping at any single voxel. In addition to manual selection of participants with focal insula damage from the Iowa Patient Registry, we included only those participants with at least 5% of this structure damaged in our insula-damaged group. This criterion reduces the likelihood that participants have lesions that only peripherally include the insula while ensuring that even small but focal insular lesions are included, given the nature of naturally occurring lesions that do not respect functional or neuroanatomical boundaries (see Gaznick, Tranel, McNutt, & Bechara, 2014; Koscik & Tranel, 2012, 2013).

2.3. Procedure

Participants completed a multi-round, multiplayer version of the Trust Game (TG; King-Casas et al., 2005; Koscik & Tranel, 2011). In the TG, one player is the “Investor” and another player is the “Trustee.” At the beginning of each round, the Investor is given \$20. The Investor decides how much of that \$20 to keep and how much to give to the Trustee, in whole dollar amounts. The amount the Investor gives to the Trustee is then tripled. The Trustee is then given the same choice. They decide how much of the received money to keep and how much to return to the Investor. The Investor is informed of the Trustee’s decision and a new round begins by the Investor receiving another \$20 to divide.

Participants first played the role of Trustee for 20 rounds and then played the role of the Investor for 20 rounds. Between roles, the participants were told that they would “switch roles with the other player.” All participants played against a computer-simulated opponent. During the instructions and the game, the opponent was always referred to as the “other player.” If the participant inquired about the other player’s identity they were told that they would be informed of their identity once the task was completed. All participants played against the same simulated opponent – this opponent responded dynamically in a “Tit-for-Tat” manner to the behavior of the participant (for more details on how this simulated opponent responds, see Koscik et al., 2011).

In the first version, when the participant was the Trustee, the participant saw a text description of how much the Investor had given to them and how much it became after it tripled. The participant also saw how much money was retained by the investor. Then, the participant decided how much to return to the Investor. Participants selected how much to return to the Investor by sliding a horizontal scroll bar from left to right. The default position at the far left signified keeping all of the money; the further the participant scrolled the bar

to the right, the more money they gave to the investor. On the screen at all times was a running tally for both the total amounts earned by the Investor and Trustee, and the amount of the previous investment and previous return.

In the second version, when the participant was the Investor, the participant was first given \$20 and asked to decide how much to give to the Trustee. This decision was made using a horizontal slider similar to the one used in the role of Trustee. The participant received feedback on how much they gave and the amount when it was tripled, and then waited for the Trustee to decide how much to return (~6 s). The participant was then informed how much the Trustee returned to them and the next round began. The TG was implemented in MATLAB (r2007b, The Mathworks) and displayed on a 17 inch LCD monitor. All responses were made via mouse using the participant's preferred hand.

Following completion of both versions of the TG, participants answered a short questionnaire to assess their opinions and perceptions of the other player. Participants responded to six queries using a visual analog scale. The scales were each presented as a continuum with one anchor on each side. There were a total of six scales, consisting of the following ratings: greedy vs. generous, scrooge-like vs. saint-like, human-like vs. computer-like, trusting vs. distrusting, fair vs. unfair, and quick vs. slow. Values were calculated as the proportion of the length of the line where the participant indicated their marker. For example, a rating falling at the center between human-like and computer-like would be scored as 0.5.

2.4. Measures

Based on previous literature (Koscik et al., 2011) we are using two dependent variables of interest: Investor Reciprocity and Trustee Reciprocity. These measures capture the amount of change in one player's response in response to the amount of change in the other player's response. "Investor Reciprocity" is the amount of money the investor changed their investment based on the change in the amount the trustee returned to them. "Trustee Reciprocity" is the change in the Trustee's return in response to the amount the investor changed their investment. As in previous literature (Koscik et al., 2011), we classify reciprocity strategies into four groups: 1. Tit-for-tat reciprocity is when the participant's change in money sent to the other player mirrors the changes in the amount sent from the other player. 2. Benevolent reciprocity is when the participant increases the amount of money sent in response to a decrease or no change in the amount sent from the other player. 3. Ambivalent reciprocity is when the participant does not change the amount of money sent despite changes in the amount sent from the other player. 4. Malevolent reciprocity is when the participant decreases the amount of money sent in response to an increase or no change in the amount sent from the other player.

A tit-for-tat strategy is the most common reciprocity strategy used by normal, healthy adults (King-Casas et al., 2005; Koscik & Tranel, 2011). This strategy exemplifies mutual trust or distrust. That is, when a player increases the amount they send and the other player increases the amount they return (mutual trust), or when a player decreases the amount they send and the other player decreases the amount they return (mutual distrust), they are exhibiting a tit-for-tat strategy. The other strategies (benevolent, ambivalent, and malevolent) exemplify

inappropriate trust. A benevolent strategy signifies inappropriately high trusting of the other player, while a malevolent strategy signifies inappropriately high distrust of the other player.

Participants typically do not use the same reciprocity strategy for every trial. Instead, the type of strategy used varies throughout the task. For example, an individual may behave benevolently on one trial and tit-for-tat on a subsequent trial. We computed the frequency of use for each strategy, expressed as a percentage of the total number of trials when the participant was both an investor and a trustee.

In addition to our main variables of interest, we calculated the total amount earned by the participant and by their computer player for both trustee and investor roles. We calculated the growth rate of trust over the course of the game as the average change in investment amount between adjacent trials. However, we would not necessarily anticipate group differences in total earnings, as the computer was programmed to minimize differences in earnings between the participants and the simulated computer player, to reduce the impact of this as a driving force in participants' decision-making.

2.5. Statistical Analyses

Demographic variables were compared between all groups using one-way ANOVAs. Neuropsychological variables were compared between the INS and BDC groups using independent samples t-tests, Bonferroni-corrected for multiple comparisons. Reciprocity strategy was analyzed using two separate three-by four mixed-ANOVAs with reciprocity strategy as a within-subjects variable and group membership as a between-subjects variable. Separate ANOVAs were performed for Investor Reciprocity and Trustee Reciprocity.

3. Results

In response to our manipulation check, there were no group differences in terms of how greedy/generous [$F(2, 35) = 0.76, p = 0.47$], scrooge-like/saint-like [$F(2, 33) = 0.21, p = .80$], human-like/computer-like [$F(2, 32) = 1.11, p = 0.34$], trusting/distrusting [$F(2, 33) = 0.60, p = 0.55$], fair/unfair [$F(2, 34) = 0.32, p = 0.72$], or quick/slow [$F(2, 33) = 0.07, p = 0.92$] the other player was. Participants tended to rate the other player as falling near the center of the continuum ranging between the two anchors (all group means between .40 and .61).

There were no differences between groups in the total earnings for themselves in the trustee [$F(2, 54) = 1.28, p = 0.28$] or investor role [$F(2, 54) = 1.53, p = 0.22$]. There were also no group differences in total earnings for the computer during the trustee [$F(2, 54) = 1.01, p = .36$] or the investor role [$F(2, 54) = 2.69, p = 0.07$], although there was a trend-level effect for the BDC group's computer partner to have greater earnings than the NC and INS groups. In addition, there were no group differences in growth rate of trust in the role of trustee [$F(2, 54) = 0.32, p = 0.72$] or investor [$F(2, 54) = 1.61, p = 0.20$].

For Investor Reciprocity, a 3×4 mixed ANOVA with Reciprocity Strategy as a within-subjects variable and Group as a between-subjects variable revealed a significant group by reciprocity strategy interaction, [$F(6,162) = 2.19, p = 0.04, \eta_p^2 = 0.07$] (Figure 2). To

follow-up on this interaction, we conducted post-hoc tests of simple main effects to investigate pairwise comparisons between the three groups. The INS group reciprocated benevolently a significantly greater proportion of trials ($M = 0.32$, 95% CI: [0.27, 0.38]) than the NC group ($M = 0.22$, 95% CI: [0.19, 0.25]), $p = 0.004$. There was no significant difference between the INS group and the BDC group ($M = 0.26$, 95% CI: [0.21, 0.31]), $p = 0.12$. The NC group performed significantly more in a tit-for-tat reciprocity strategy ($M = 0.36$, 95% CI: [0.31, 0.40]) than the INS group ($M = 0.23$, 95% CI: [0.14, 0.31]), $p = .01$. Although the NC group did not perform significantly different than the BDC group, there was a trend towards the NC group performing more frequently in a tit-for-tat manner than the BDC group ($M = 0.28$, 95% CI: [0.21, 0.35]), $p = .07$. There were no other significant differences between groups for reciprocity strategies.

For trustee reciprocity, a 3×4 mixed ANOVA with reciprocity strategy as a within-subjects variable and group as a between-subjects variable revealed no significant group by reciprocity strategy interaction, $F(6,168) = 1.12$, $p = 0.33$, $\eta_p^2 = 0.04$ (Figure 3). Although there was not a significant interaction at the group-level, we conducted post hoc tests to investigate any differences between groups. Post-hoc tests of simple main effects revealed at the trend level that the INS group tended to perform more malevolently ($M = 0.27$, 95% CI: [0.20, 0.33]) than the BDC group ($M = 0.19$, 95% CI: [0.14, 0.24]), $p = 0.06$. There were no other significant or trend-level differences between groups for trustee reciprocity.

To further investigate the relationship between damage to the insula and performance on the TG, we ran correlations between the proportion of insula damage and all reciprocity strategies for both trustee and investor reciprocity. Amount of insula damage was positively correlated with malevolent reciprocity style during the role of trustee ($r = 0.38$, $p = 0.01$). Table 3 displays correlations between amount of insula damage and all reciprocity strategies

Additionally, we were interested in examining the effects of sex on performance strategies. First, we divided the participants into two groups based on sex (irrespective of brain injury status). We then performed two 2×4 mixed ANOVAs with reciprocity strategy as a within-subjects variable and sex as a between-subjects variable. There was no significant interaction between sex and reciprocity strategy for investor reciprocity [$F(3, 165) = 2.40$, $p = .06$, $\eta_p^2 = 0.04$]. Post-hoc tests of simple main effects revealed that women performed significantly more malevolently ($M = 0.25$, 95% CI:[0.21, 0.28]) than men ($M = 0.19$, 95% CI: [0.15, 0.22], $p = .04$). To follow up on this analysis, we performed a $2 \times 2 \times 4$ mixed ANOVA including reciprocity strategy as a within-subjects variable and lesion group membership and sex as between-subjects variables. There was no significant interaction between lesion group, sex, and reciprocity strategy [$F(6, 153) = 0.60$, $p = 0.72$, $\eta_p^2 = 0.01$]. There was no significant interaction between sex and reciprocity strategy for trustee reciprocity [$F(3, 165) = 0.79$, $p = .49$, $\eta_p^2 = 0.02$]. We then performed these same analyses looking at brain lesion laterality, by dividing the groups into those with left hemisphere lesions and those with right hemisphere lesions. There was no significant interaction between lesion side and reciprocity strategy for investor reciprocity [$F(6, 69) = 0.82$, $p = 0.55$, $\eta_p^2 = 0.06$]. There was also no significant interaction between lesion side and reciprocity strategy for trustee reciprocity [$F(6, 69) = 1.60$, $p = 0.16$, $\eta_p^2 = 0.12$].

In light of these results, it was of interest to determine whether there were real-world correlates of impaired social decision-making in the INS group. Our patients have been well characterized on various neuropsychological measures, and from this dataset we identified a measure that captures changes in real-world behaviors after brain injury. The Iowa Scales of Personality Change (ISPC; Barrash et al., 2011) assesses personality changes that occur after brain injury. One scale included in the ISPC is “Social Inappropriateness,” which measures the extent to which the individual “might not act properly in social settings.” Impaired judgments of social norms might be associated with abnormal behavior on the trust game, making this scale particularly relevant to the current investigation. We found that individuals in the INS group ($M=3.75$, $SD=2.4$) were rated as significantly higher on this dimension than individuals in the BDC group ($M=1.14$, $SD = 0.37$), $t(7.38) = 2.98$, $p = .01$. Anecdotally, one patient’s spouse reported “increased anger, road rage” and “decreased personal grooming,” following brain injury, which may indicate impaired assessments of and compliance with social norms. Another patient’s spouse reported that the patient had “become clingy, always wants [them] around, and has become dependent,” possibly reflecting abnormal trust.

4. Discussion

Our results support a role for the insula in the process of complex social decision-making. There are several aspects of decision-making during the TG that make contact with functions of the insula. Performance in the trust game depends on accurate evaluations of a partner’s trustworthiness, cooperation between partners, and evaluations of risk during economic decisions. Successful interactions during the TG are dependent on the ability to accurately assess trustworthiness and create mutually beneficial social interactions.

Consistent with previous studies (King-Casas et al., 2005; Kosciak & Tranel, 2011), we report here that neurologically normal adults use a tit-for-tat reciprocity strategy in the TG. Normal, healthy adults respond in kind to their partners – betrayals are reciprocated with punishment, while expressions of trust are reciprocated in kind.

Conversely, we found that individuals with damage to the insula did not respond in a tit-for-tat manner. Instead, these participants tended to differ from this normative reciprocity style. As investors, individuals with insula damage tended to respond in a benevolent manner, expressing trust in their partner even when their partner had betrayed their trust. As trustee, individuals with insula damage tended to respond in a malevolent manner, betraying their partners even when their partners had expressed trust.

Participants with brain damage excluding the insula (BDCs) displayed a slightly different manner of behavior than both participants with insula damage and normal, healthy adults. The behavior of these participants also differed based on the role they were playing. As investor, these participants tended to behave in a tit-for-tat manner less frequently than the NC group but more frequently than the INS group. As trustee, these participants tended to perform in a tit-for-tat manner similar to normal healthy adults. Further studies could look for systematic brain-behavior relationships amongst patients with lesions to other structures.

Our results support a critical role of the insula for interpersonal economic decision-making. Previous studies using various economic decision-making games have found involvement of the insula during risky or uncertain decisions. A recent meta-analysis found that activity in the anterior insula reflects both anticipation of risk and risky decisions (Mohr, Biele, & Heekeren, 2010). The TG contains an element of risk (van den Bos et al., 2009). For example, if the investor invests a large amount of their money, they are taking a risk that the trustee may not reciprocate. In this way, benevolence as the investor may be seen as a risky choice. Here, we found that individuals with insula damage tended to act benevolently when playing the role of the investor. This means that participants with insula damage consistently made investments that were higher than the reciprocation they received from the trustee. Therefore, by being benevolent as the investor, the individuals with insula damage consistently took risky choices that were not being rewarded.

Insula activity has been found in response to trials when participants made risky decisions that were not rewarded (Paulus, Rogalsky, Simmons, Feinstein, & Stein, 2003). When participants took a risky investment and subsequently lost, they showed increased insula activation. Additionally, the greater the insula activation in response to these losses, the more likely the participant was to choose the safe bet on the next turn. Therefore, activity in the insula may signal a risky choice – if the risky choice results in a loss, increased activity in the insula may then be interpreted by the participant as an indicator to not make this risky choice again. In individuals with insula damage, this mechanism may be impaired, and they may not be able to understand that they are making a risky choice. So, when their partner responds with a decrease in their return, they do not similarly decrease their investment, but instead act benevolently. Our findings are consistent with previous lesion research that has shown that individuals with insula damage show reduced risk adjustment, i.e., do not place lower bets on riskier outcomes (Clark et al., 2008). Additionally, individuals with insula damage have shown higher levels of bankruptcies in economic decision-making games (Clark et al., 2008). Our data are therefore consistent with the idea that insula activity reflects risky decisions, and that damage to the insula impairs judgments of risk. In the case of the TG, individuals with insula damage acted benevolently as the investor, a decision that places them at risk of being exploited.

While individuals with insula damage tended to act benevolently when in the role of investor, they tended to act malevolently when in the role of trustee. This may reflect differences in the two roles: what constitutes a risky choice in the role of trustee may be different from what constitutes a risky choice in the role of investor. Seemingly, malevolence is a ‘safe’ decision: an individual keeps the money for themselves. However, this type of behavior ignores the social aspect of the trust game, and the necessity of building trust with one’s partner in order to facilitate advantageous future transactions. Neurologically normal adults perform in a tit-for-tat manner – when their partners increase their investments, the trustees also increase their investments. This facilitates trust between the partners and builds a relationship to enhance future interactions. The trustee is reliant on the investor to receive money, so if they do not reciprocate in kind, the investor may not invest with them again in the future. Thus, being malevolent as the trustee is a risky decision, because it does not facilitate the formation of a positive relationship and jeopardizes future economic gain. This is different from when in the role of the investor,

because the trustee is solely reliant on the investor for his income. Individuals with insula damage may behave malevolently because they do not recognize the risk involved.

Aside from economic-based explanations, it is important to take into account the social aspect of the TG and the role of the insula in social behavior. Previous functional imaging research has shown insula activity in response to social norm violations during the TG (King-Casas et al., 2008; van den Bos et al., 2009). King-Casas and colleagues (2008) found that insula activity was negatively associated with investment level, so that higher insula activity was associated with lower investments. They suggest that insula activity when making and receiving low offers signals violations of social norms, i.e., low offers. Thus, although malevolence may be monetarily beneficial for the trustee, it is a violation of social norms. Here, we found that individuals with insula damage displayed malevolent reciprocity when in the role of the trustee. These individuals may not appropriately recognize a norm violation and continue to repeatedly behave in a way that violates social norms. While malevolent behavior as the trustee is a violation of social norms, benevolence may be socially beneficial. In fact, King-Casas et al. (2008) describe benevolence as “coaxing;” when normal, healthy adults receive a low offer, they often return a higher amount in order to “coax” their partner back into a more reciprocal relationship. Therefore, instead of behaving in a socially appropriate way, here we see individuals with insula damage violating the trust of their partners and behaving malevolently.

When in the role of investor, individuals with damage to the insula displayed a benevolent reciprocity strategy, in contrast with their malevolent strategy use in the role of trustee. However, displaying benevolence as an investor is also a form of social norm violation. Acting benevolently puts the investor in a position which may let their partner take advantage of them. When one repeatedly behaves benevolently, they risk being exploited by their partner. As the investor, an appropriate social norm would be to protect oneself from others until an accurate judgment could be made about how they will reciprocate. When individuals with insula damage repeatedly invest benevolently, they do not realize that they are being taken advantage of, i.e., they do not recognize this violation of normal social behavior. In this way, benevolence as the investor both violates normative behavior during the TG and has a component of risk – both processes supported by the insula. Consistent with the idea of the insula as a critical region for identifying social norm violations, we found that individuals with damage to the insula display greater levels of inappropriate social behavior in real-world situations.

Additionally, the insula has been implicated as an important neural structure underlying betrayal aversion. For example, insula activity is increased in betrayal-averse individuals when they make a trusting decision (Aimone, Houser, & Weber, 2014). When trusting a partner, a person opens him or herself to a potential betrayal of that trust. The insula may be important for signaling a potential betrayal and influencing a person’s decision not to trust. Damage to the insula may remove this inhibition of trust and lead to increased benevolence as an investor. Similarly, if the insula signals one’s own betrayals as well, removal of this would increase malevolence, as seen in our current data.

At the completion of the experiment, participants were asked to describe any particular strategy they used during the task. Anecdotally, participants in the NC and BDC group tended to respond that their strategy was to perform in a tit-for-tat manner, giving responses such as “if they gave more money I would give more back,” and “when the other person gave me a very small amount I would give less.” Individuals in the INS group did not as frequently express a tit-for-tat strategy, instead suggesting that they “tried to keep the investor happy” and “attempted to be ‘nice’ to the other player,” although one INS group member reported that they would “give more if the balance was out of whack,” suggesting that they were attempting to maintain a fair balance. These anecdotal data support our account that individuals with damage to the insula display abnormal strategies during interpersonal decision-making.

While these findings are consistent with a role of the insula in social decision-making, there are several issues inherent in studying patients with brain damage that may present limitations to these conclusions. First of all, patients with focal, circumscribed insula lesions are rare, as evidenced by our current sample. While these lesions all include insula damage, they tend to cover other areas of cortex, including regions outside of the insula. While our maximum region of overlap is located in the insula region, this additional lesion involvement may limit the specificity of our conclusions. We have aimed to address this concern by including a BDC group of individuals with brain damage outside the insula. The BDC group provides a control for general effects of brain damage. Including a BDC group helps address the possibility that brain damage to any region could cause the behavioral effects seen here—this does not appear to be the case, based on our findings.

However, including this BDC group introduces an additional source of variability between brain-damaged groups. As described above, participants in the BDC group have lesions due to a variety of etiologies. To address this concern, we grouped patients together by etiology (stroke vs. resection) and conducted independent-samples t-tests to investigate differences in our dependent variables. We found no significant differences between etiology groups. An additional limitation of this study may stem from the variability in our participants due to individual difference variables (e.g., education, socioeconomic status, age). These variables may play an important role in decision-making during the TG. However, we found no statistical differences between groups on demographic or neuropsychological variables. Another limitation of the current approach is the fact that the caudate is also frequently damaged in individuals with insula damage, as seen in our present group. This issue is hard to approach definitively, given the small size of the structure and its frequent collateral damage with the insula.

In conclusion, our findings support the role of the insula in complex social decision-making, including judgments of risk and social norm violations. In the TG, risky decisions can differ based on the role the participant plays. Here, we found that individuals with insula damage performed benevolently as the investor and malevolently as the trustee, reflecting impaired decision-making in this task. Additionally, these findings support the role of the insula as responsive to both positive and negative affective states (e.g., benevolence and malevolence). Insula damage may result in impaired social decision-making due to an

inability to evaluate both the risk as well as the normative social behavior that occurs during a reciprocal exchange.

Acknowledgments

We would like to acknowledge Ellen van der Plas for her help with statistical analyses, Ken Manzel for help with the neuropsychological data, Joel Bruss for his help with the preparation of Figure 1, and Ruth Henson for help in scheduling the patients. This study was supported by the James S. McDonnell Foundation (grant #220020387), the Kiwanis Foundation, and NINDS P01 NS19632.

References

- Aimone JA, Houser D, Weber B. Neural signatures of betrayal aversion: an fMRI study of trust. *Proceedings of the Royal Society B*. 2014;281.
- Bar-On R, Tranel D, Denburg NL, Bechara A. Exploring the neurological substrate of emotional and social intelligence. *Brain*. 2003; 126:1790–1800. [PubMed: 12805102]
- Barrash J, Asp E, Markon K, Manzel K, Anderson SW, Tranel D. Dimensions of personality disturbance after focal brain damage: Investigation with the Iowa Scales of Personality Change. *Journal of Clinical and Experimental Neuropsychology*. 2011; 33:833–852.
- Bechara A, Damasio AR. The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*. 2005; 52:336–372.
- Berntson GG, Norman GJ, Bechara A, Bruss J, Tranel D, Cacioppo JT. The insula and evaluative processes. *Psychological Science*. 2011; 22:80–86. [PubMed: 21148459]
- Bossaerts P. Risk and risk prediction error signals in anterior insula. *Brain Structure & Function*. 2010; 214:645–653. [PubMed: 20512378]
- Clark L, Bechara A, Damasio H, Aitken MRF, Sahakian BJ, Robbins TW. Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain*. 2008; 131:1311–1322. [PubMed: 18390562]
- Collins DL, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*. 1994; 18(192)
- Craig AD. How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*. 2002; 3:655–666.
- Critchley HD, Mathias CJ, Dolan RJ. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*. 2001; 29:537–545. [PubMed: 11239442]
- d'Acremont M, Lu ZL, Li X, Van der Linden M, Bechara A. Neural correlates of risk prediction error during reinforcement learning in humans. *NeuroImage*. 2009; 47:1929–1939. [PubMed: 19442744]
- Damasio A. The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society*. 1996:1413–1420.
- Damasio A, Grabowski TJ, Bechara A, Damasio H, Ponto LL, Parvizi J, Hichwa RD. Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*. 2000; 3:1049–1056.
- Damasio H, Tranel D, Grabowski T, Adolphs R, Damasio A. Neural systems behind word and concept retrieval. *Cognition*. 2004; 92:179–229. [PubMed: 15037130]
- Eisenberger NI, Lieberman MD, Williams KD. Does rejection hurt? An fMRI study of social exclusion. *Science*. 2003; 302:290–292. [PubMed: 14551436]
- Evans A, Dai W, Collins L, Neelin P, Marrett S. Warping of a computerized 3-D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis. *Proceedings of the International Society of Optical Engineering (SPIE): Medical Imaging*. 1991; 1445:236–246.
- Fiez JA, Damasio H, Grabowski TJ. Lesion segmentation and manual warping to a reference brain: Intra- and interobserver reliability. *Human Brain Mapping*. 2000; 9(4):192–211. [PubMed: 10770229]

- Frank RJ, Damasio H, Grabowski TJ. Brainvox: An interactive, multimodal visualization and analysis system for neuroanatomical imaging. *NeuroImage*. 1997; 5:13–30. [PubMed: 9038281]
- Gaznick N, Tranel D, McNutt A, Bechara A. Basal Ganglia plus insula damage yields stronger disruption of smoking addiction than basal ganglia damage alone. *Nicotine & Tobacco Research*. 2014; 16:445–453. [PubMed: 24169814]
- Johnson H, Zhao Y. BRAINSDemonWarp: An application to perform demons registration. *The Insight Journal*. 2009
- Kang Y, Williams LE, Clark MS, Gray JR, Bargh JA. Physical temperature effects on trust behavior: the role of insula. *Social Cognitive and Affective Neuroscience*. 2011; 6:507–515. [PubMed: 20802090]
- King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague PR. The rupture and repair of cooperation in bordering personality disorder. *Science*. 2008; 321:806–810. [PubMed: 18687957]
- King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR. Getting to know you: Reputation and trust in a two-person economic exchange. *Science*. 2005; 308:78–83. [PubMed: 15802598]
- Koscik TR, Tranel D. The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia*. 2011; 49:602–611. [PubMed: 20920512]
- Koscik TR, Tranel D. The human ventromedial prefrontal cortex is critical for transitive inference. *Journal of Cognitive Neuroscience*. 2012; 24:1191–1204. [PubMed: 22288395]
- Koscik TR, Tranel D. Abnormal causal attribution leads to advantageous economic decision-making: A neuropsychological approach. *Journal of Cognitive Neuroscience*. 2013; 25:1372–1382. [PubMed: 23574584]
- Krawitz A, Fukunaga R, Brown JW. Anterior insula activity predicts the influence of positively framed messages on decision making. *Cognitive, Affective & Behavioral Neuroscience*. 2010; 10:392–405.
- Kross E, Egner T, Ochsner K, Hirsch J, Downey G. Neural dynamics of rejection sensitivity. *Journal of Cognitive Neuroscience*. 2007; 19:945–956. [PubMed: 17536965]
- Krueger F, McCabe K, Moll J, Kriegeskorte N, Zahn R, Strenziok M, Grafman J. Neural correlates of trust. *Proceedings of the National Academy of Sciences*. 2007; 104:20084–20089.
- Lamm C, Singer T. The role of anterior insular cortex in social emotions. *Brain Structure & Function*. 2010; 214:579–91.
- Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Mazoyer B. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*. 2001; 356:1293–1322.
- Mizuhiki T, Richmond BJ, Shidara M. Encoding of reward expectation by monkey anterior insular neurons. *Journal of Neurophysiology*. 2012; 107:2996–3007. [PubMed: 22402653]
- Mohr PNC, Biele G, Heekeren HR. Neural processing of risk. *Journal of Neuroscience*. 2010; 30:6613–6619. [PubMed: 20463224]
- Ostrowsky K, Magnin M, Ryvlin P, Isnard J, Guenot M, Mauguière F. Representation of pain and somatic sensation in the human insula: A study of responses to direct electrical cortical stimulation. *Cerebral Cortex*. 2002; 12:376–385. [PubMed: 11884353]
- Paulus MP, Rogalsky C, Simmons A, Feinstein JS, Stein MB. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage*. 2003; 19:1439–1448. [PubMed: 12948701]
- Phillips ML, Young AW, Senior C, Brammer M, Andrew C, Calder A, David AS. A specific neural substrate for perceiving facial expressions of disgust. *Nature*. 1997; 389:495–498. [PubMed: 9333238]
- Preuschoff K, Quartz SR, Bossaerts P. Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*. 2008; 28:2745–2752. [PubMed: 18337404]
- Sanfey AG, Rilling JK, Aronson JA. The neural basis of economic decision-making in the ultimatum game. *Science*. 2003; 300:1755–1758. [PubMed: 12805551]

- Singer T, Kiebel SJ, Winston JS, Dolan RJ, Frith CD. Brain responses to the acquired moral status of faces. *Neuron*. 2004; 41:653–662. [PubMed: 14980212]
- Singer T, Seymour B, O'Doherty JP, Stephan KE, Dolan RJ, Frith CD. Empathic neural responses are modulated by the perceived fairness of others. *Nature*. 2006; 439:466–469. [PubMed: 16421576]
- Todorov A, Baron SG, Oosterhof NN. Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience*. 2008; 3:119–127. [PubMed: 19015102]
- Tranel, D. Theories of clinical neuropsychology and brain-behavior relationships: Luria and beyond. In: Morgan, JE.; Ricker, JH., editors. *Textbook of Clinical Neuropsychology*. New York: Taylorand Francis; 2007. p. 27-39.
- Van den Bos W, van Dijk E, Westenberg M, Rombouts SARB, Crone EA. What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Social Cognitive and Affective Neuroscience*. 2009; 4:294–304. [PubMed: 19304843]
- Wechsler, D. Wechsler Test of Adult Reading. San Antonio, TX: Psychological Corporation; 2001.
- Wechsler, D. Wechsler Adult Intelligence Scale. 4. San Antonio, TX: Pearson; 2008.
- Winston JS, Strange BA, O'Doherty J, Dolan RJ. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*. 2002; 5:277–283.
- Xue G, Lu Z, Levin IP, Bechara A. The impact of prior risk experiences on subsequent risky decision-making: The role of the insula. *NeuroImage*. 2010; 50:709–716. [PubMed: 20045470]

Highlights

Individuals with brain damage and healthy adults participated in a Trust Game (TG)

Individuals with damage to the insula performed abnormally on the TG

The insula is a critical structure for decision-making in the TG

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Lesion overlap map of individuals with insula lesions (orange) and brain-damaged comparison participants (blue). Panels a–k depict axial cuts through the insula region. The three panels in the bottom right of the image depict peri- and midsagittal views illustrating the planes of the axial cuts. Images are shown in radiological convention, with the left hemisphere on the right. The color bars indicate maximum lesion overlap, with “hotter” colors representing higher numbers of overlap (orange to yellow for INS; dark blue to light blue for BDC). Maximum overlap ($n = 5$) in the INS group is located in the left anterior insula. Max overlap in the BDC group ($n = 4$) is located in the right occipital cortex.

Figure 2.

Proportion of turns spent in each reciprocity strategy for investor reciprocity. The INS group used significantly more benevolent and less tit-for-tat than the NC group. Error bars are 95% confidence intervals.

Figure 3.

Proportion of turns spent in each reciprocity strategy for trustee reciprocity. The INS group was significantly more benevolent and less tit-for-tat than the NC group. Error bars are 95% confidence intervals.

Table 1

There are no significant differences between groups on any demographic variables. Values are means (SDs). Handedness: R = right-handed, L = left-handed; FSIQ = Full Scale Intelligence Quotient. Lesion size is proportion of the entire brain.

Demographics			
	Group		
	INS (n = 11)	BDC (n = 15)	NC (n = 31)
Age (years)	61.1 (7.6)	54.9 (16.1)	56.5 (16.8)
Handedness	9 R, 2 L	12 R, 3 L	31 R, 0 L
Sex	4 F, 7 M	10 F, 5 M	14 F, 17 M
Education (years)	14.9 (3.1)	15.1 (2.7)	15.7 (2.4)
FSIQ	104.4 (11.1)	105.6 (13.6)	112.3 (8.9)
Chronicity (years)	13.1 (15.5)	10.0 (6.5)	--
Lesion Size	0.03 (0.03)	0.02 (0.02)	--

Table 2

Participant groups did not differ significantly on any of the neuropsychological measures. WAIS, Wechsler Adult Intelligence Scale scores (VIQ, Visual IQ; PIQ, Performance IQ); BDI-II, Beck Depression Inventory-II; BAI, Beck Anxiety Inventory; AVLT - 5, Trial 5 of the Auditory-Verbal Learning Test scores (raw score #/15, an index of verbal learning); WCST, Wisconsin Card Sorting Test Categories completed (#/6, an index of reasoning and concept formation); FRT, Benton Facial Discrimination Test (raw score #/54, an index of visual perception); COWA, Controlled Oral Word Association Test (raw score, an index of verbal fluency); BNT, Boston Naming Test (raw score #/60, an index of confrontation naming).

Neuropsychology		
	Group	
	INS	BDC
WAIS - PIQ (SD)	104.6 (14.7)	104.8 (9.7)
WAIS - VIQ (SD)	105.3 (15.8)	106.5 (19.3)
BDI-II (SD)	6.1 (3.0)	4.3 (3.1)
BAI (SD)	3.8 (3.2)	3.7 (2.9)
AVLT - 5 (SD)	13.1 (2.3)	14.3 (0.8)
WCST (SD)	5.2 (1.9)	5.4 (1.2)
Facial Disc. (SD)	46.4 (4.5)	47.6 (3.8)
COWA (SD)	33.8 (11.0)	40.4 (14.4)
BNT (SD)	52.3 (9.6)	55.7 (3.4)

Table 3

Correlations between amount of insula damage and reciprocity strategies.

	Investor		Trustee	
	r-value	p-value	r-value	p-value
Tit-for-Tat	−0.233	0.15	−0.23	0.14
Benevolent	−0.13	0.48	0.12	0.46
Ambivalent	0.06	0.70	−0.08	0.59
Malevolent	−0.03	0.81	0.38	0.01