Causal role of the right temporoparietal junction in selfishness depends on the social partner

Kristin M. Brethel-Haurwitz, [1,2] Desmond J. Oathes, [2,3] and Joseph W. Kable[1,2]

[1]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]MindCORE, University of Pennsylvania, Philadelphia, PA 19104, USA
[3]Department of Psychiatry and Center for Neuromodulation in Depression and Stress, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Correspondence should be addressed to Kristin M. Brethel-Haurwitz, Department of Psychology, University of Pennsylvania, 433 South University Avenue, Kable Lab - Goddard 5, Philadelphia, PA 19104, USA. E-mail: kbreth@sas.upenn.edu.

Abstract

The right temporoparietal junction (rTPJ) is a hub of the mentalizing network, but its causal role in social decisions remains an area of active investigation. While prior studies using causal neurostimulation methods have confirmed the role of the rTPJ in mentalizing and strategic social interactions, most of the evidence for its role in resource-sharing decisions comes from correlational neuroimaging studies. Further, it remains unclear if the influence of the rTPJ on decisions about sharing resources depends on whether the other person is salient and identifiable. To clarify the causal role of the rTPJ in social decision making, we examined the effects of putatively inhibitory rTPJ transcranial magnetic stimulation (TMS) on Dictator Game behavior with one partner that was physically present and one that was only minimally identified. Under control conditions, participants tended to create more advantageous inequity toward the partner that was only minimally identified, selfishly keeping more resources themselves. rTPJ TMS reduced this differential treatment of the two partners. Clarifying prior mixed findings, results suggest that the rTPJ may play a role in differentiating between others when deciding how equitably to divide resources, but may not play a general role in reducing selfishness by promoting aversion to advantageous inequity.

Key words: temporoparietal junction; transcranial magnetic stimulation; social decision making

Mentalizing allows us to imagine, anticipate and react to the thoughts, feelings and intentions of others. Also known as 'theory of mind' or 'cognitive empathy', mentalizing is the process by which we infer the content of other minds and is thought to support large-scale cooperation and intricate social networks (Frith and Singer, 2008; Frith and Frith, 2010, 2012). Extensive research has implicated the right temporoparietal junction (rTPJ) as a causally important hub in the mentalizing network (Saxe and Kanwisher, 2003; Saxe and Powell, 2006; Young *et al.*, 2007, 2010b; Santiesteban *et al.*, 2012; Kanske *et al.*, 2015, 2016). However, although mentalizing has been hypothesized to play an important role in decisions to share resources with others, evidence for the causal role of rTPJ in such decisions is limited. In the current study, we used putatively inhibitory repetitive transcranial magnetic stimulation (TMS) to examine the role of the rTPJ in decisions involving a tradeoff between one's own outcome and the outcome of another person, when that other person was either someone that was briefly introduced and physically present or someone that was only minimally identified and not physically present.

Decisions that involve such a tradeoff between one's own outcome and the outcome of another appear to be influenced by at least two distinct motives, according to a well-supported economic model of social preferences (Fehr and Schmidt, 1999). One motive is aversion (or lack thereof) to advantageous inequity, or having more resources than another person, while a second motive is aversion (or lack thereof) to disadvantageous inequity, or having fewer resources than another person. This distinction between advantageous and disadvantageous inequity aversion is well supported by behavioral evidence (Charness and Rabin, 2002; Bruhin *et al.*, 2019). Neural evidence also further supports this distinction, as brain activity differs for decisions about advantageous *vs* disadvantageous inequity (Gao *et al.*, 2018) and TPJ gray matter volume is associated with advantageous but not disadvantageous inequity aversion (Morishima *et al.*, 2012). Further, a similar distinction is made in studies of selfishness, which distinguish adaptive forms that involve obtaining a baseline level of resources from pathological forms that involve hoarding resources when one has plenty (Raine and Uh, 2019). To try to distinguish between advantageous and disadvantageous inequity aversion, we used a

---

previously developed Dictator Game paradigm that manipulates the cost of giving so that there is an incentive to create disadvantageous inequity on some trials (i.e. resources are more valuable when shared than when kept). This paradigm has previously been used to generate independent estimates of advantageous and disadvantageous inequity aversion at the group level (Sáez *et al.*, 2015).

Previous research suggests two hypotheses about the potential role of the rTPJ in aversion to advantageous or disadvantageous inequity. One body of evidence suggests that the TPJ plays a role in how we weight the value of outcomes for others relative to ourselves (Hutcherson *et al.*, 2015; Morishima *et al.*, 2012; Tankersley *et al.*, 2007; Tusche *et al.*, 2016; Soutschek *et al.*, 2016; Obeso *et al.*, 2018). The pattern of neural responses in the TPJ predicts prosocial donation behavior, as well as the degree to which people report mentalizing, rather than emotional empathizing, when deciding whether to donate (Tusche *et al.*, 2016). Further, larger gray matter volume (Morishima *et al.*, 2012) and greater functional activation (Tankersley *et al.*, 2007) in the TPJ are associated with a greater degree of altruism (although TPJ activation may reflect conflict between selfish and altruistic motives, see Morishima *et al.*, 2012; Hutcherson *et al.*, 2015). Since the association with TPJ structure is specific to altruism under conditions of advantageous inequity (Morishima *et al.*, 2012), and donation studies can be reasonably construed to be under conditions of advantageous inequity, these results suggest that rTPJ reduces selfishness by promoting aversion to advantageous inequity. However, *causal* evidence for the role of the TPJ in advantageous inequity aversion is mixed. Soutschek *et al.* (2016) found that inhibitory TMS to the rTPJ increased social discounting or the rate at which one becomes more selfish toward more socially distant relationships. In contrast, Obeso *et al.* (2018) found that inhibition of rTPJ decreased the weight of selfish monetary compared to moral concerns and did not affect giving in general or changes in giving in response to one's decisions being observed by another person.

Another body of evidence suggests that the TPJ plays a role in distinguishing between social partners. In strategic social interactions, the TPJ is sensitive to features of one's interaction partner. In humans, the pattern of BOLD activity in the TPJ makes a unique contribution to predicting bluffing behavior in a poker game against a human, but not a computer, and this difference in predictive power was more pronounced when participants rated the human as a more effective opponent (Carter *et al.*, 2012). In monkeys, neurons in a TPJ homologue selectively respond to cooperative rewards in a strategic social decision-making paradigm, in a manner that depends on whether the opponent is another monkey or a computer (Ong *et al.*, 2020). Interestingly, inhibitory TMS to rTPJ did not affect the amount of money participants were willing to forgo to share with their closest social relationships, but rather escalated the rate at which selfishness increased as a function of increasing social distance (Soutschek *et al.*, 2016). These results suggest that the causal role of the rTPJ in advantageous or disadvantageous inequity aversion may depend on the social relevance and salience of the interaction partner. Although we know that giving to others can be modulated by the amount of information provided about the recipient (Small and Loewenstein, 2003; Genevsky *et al.*, 2013), no study has specifically looked at the causal role of the TPJ in this differentiation.

In the current study, participants made resource allocation decisions in this Dictator Game paradigm for two different social partners: one who was physically present and another who was only minimally identified and not physically present. In a within-subjects design, participants made these decisions twice, once after putatively inhibitory TMS to rTPJ and once after TMS to a control site (vertex). In this experimental context, we investigated two major research questions: (i) Does modulation of the rTPJ have a general effect on the creation of advantageous or disadvantageous inequity, regardless of the social partner? (ii) Does the effect of rTPJ modulation on advantageous or disadvantageous inequity depend on the salience and physical presence of the social partner?

Methods

Participants

Twenty-seven adults (17 female, 9 male, 1 non-binary) aged 19–46 years (*M*: 30, *SD*: 7.6) completed this study as TMS participants. The sample size was selected to detect the same size effect of TMS to rTPJ as observed on social discounting in Soutschek *et al.* (2016). The required sample size to have 90% power to detect an effect size of 0.653 (the effect size for the increase in social discounting between TMS to rTPJ *vs* vertex in Soutschek *et al.*, 2016) in a within-subjects design in a two-tailed test with an alpha of 0.05 is 27 participants. Participants who had previously completed a neuroimaging study in the laboratory of JWK or DJO were recruited so that previously collected MRI scans could be used for TMS neuronavigation. An additional four women aged 24–27 years (*M*: 25, *SD*: 1.5) from the Penn community were recruited as study partners for the Dictator Game.

TMS participants were paid $30/h plus payoffs from the Dictator Game described below. Study partners were paid $10/h plus payoffs from the Dictator Game described below. All participants provided informed consent and all study procedures were approved by the Institutional Review Board at the University of Pennsylvania.

Procedure

Each TMS participant completed three study sessions. The first was an introductory session for consent, confirmation of tolerance of TMS and completion of self-report trait questionnaires and demographic questions. In the second and third sessions, participants received TMS to the rTPJ or vertex and completed two iterations of the Dictator Game (Figure 1). The order of stimulation of different brain regions was counterbalanced across participants, and the second and third sessions were conducted at least 1 day apart (*M* = 3.89, *SD* = 2.78).

TMS

Putatively inhibitory repetitive TMS was applied using a Magventure Cool B65 coil and MagVenture X100 Stimulator. TMS was administered as continuous theta burst stimulation (cTBS), with triplets of TMS pulses at 50 Hz, delivered in 5 Hz bursts (every 200 ms). Each cTBS session involved 200 bursts (a triplet of pulses) over 40 s for a total of 600 stimulations, with stimulation intensity set to 70% of resting motor threshold. A simple non-parametric algorithm (adaptive PEST) determined resting motor threshold (visually discernible *abductor pollicis brevis* or first *dorsal interosseous* muscle twitch) through successive TMS probes (Ah Sen *et al.*, 2017). cTBS was administered immediately prior to task administration. Instructions and practice trials of the Dictator Game were conducted prior to cTBS to ensure comprehension. Stimulation sites for the rTPJ were determined using structural MRI scans collected through a prior protocol and neuronavigation via Brainsight software, which matches fiduciary points on the scalp with points detectable in the MRI image so that TMS can be delivered to specific brain areas of interest. Neuronavigation assistance was
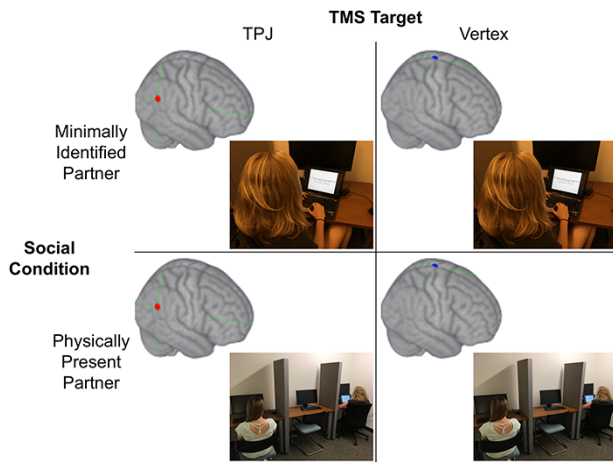
Fig. 1. The four within-subjects conditions in our study design. In each of the two TMS sessions, participants received continuous theta burst stimulation to the rTPJ or vertex, then completed two separate blocks of the Dictator Game with respect to a minimally identified interaction partner indicated by a subject number on the computer screen and a physically present interaction partner. TMS target was counterbalanced between the two TMS sessions and order of interaction partner was counterbalanced within participants across the two sessions.

also used to keep the TMS coil focused on the brain target in real time during cTBS administration. The target in right posterior TPJ (MNI: 60, −58, 31) was selected based on studies of social discounting using both fMRI (Strombach *et al.*, 2015) and TMS (Soutschek *et al.*, 2016). A 10 mm diameter sphere at this MNI coordinate was transformed to the native space of each participant's structural scan, using estimates from spatial normalization of the structural scan to a standard MNI template via FNIRT in FSL and manually checked for accuracy. The rTPJ TMS target was placed in the center of this sphere on the cortical surface for each TMS participant. These coordinates are within several millimeters of other coordinates in right posterior TPJ that have been targeted by TMS to disrupt mentalizing and moral reasoning (Young *et al.*, 2010a) and strategic social decision making (Hill *et al.*, 2017). Further, these coordinates fall within an association map region of interest defined by Neurosynth (Yarkoni *et al.*, 2011) for 'theory (of) mind' and within a posterior region of the TPJ defined by Mars *et al.* (2012) as more likely to be associated with mentalizing than non-social attention processes (see Supplementary Figure S1).

Dictator Game

We used a Dictator Game paradigm with variable exchange rates originally developed by Sáez *et al.* (2015). On each of 20 trials, participants were asked how they would share an endowment of tokens. Importantly, the value of each token in points varied depending on if a token was kept or shared. In some cases, the rate was 1:1, which is equivalent to a typical Dictator Game. Other possible self:other exchange rates were 3:1, 2:1, 1:2 and 1:3, in which the token was worth two or three times more depending on whether it was kept or shared. These exchange rates manipulated the cost of giving, such that giving was less costly to the self on some trials (i.e. the 1:3 rate) than others (i.e. the 3:1 rate). Each point was worth $0.10. Participants completed the Dictator Game twice per TMS session: once with respect to an abstract and minimally identified (i.e. by a random subject number displayed on the computer screen) future participant and once with respect to a fellow participant physically present in the room

who was previously unknown to the participant (i.e. a stranger), counterbalanced in order within participants across TMS sessions (Figure 1).

The physically present interaction partner was recruited specifically for that role and had previously completed the Dictator Game. TMS participants were briefly introduced to the study partner by name at the beginning of the session. During study sessions, the study partner was seated in another cubicle in the same testing room, but did not interact with the TMS participant beyond the brief introduction. A different partner was physically present for each TMS session and all study partners were demographically similar (females in their mid-twenties). Three of the study partners participated in 15 sessions each, while the fourth participated in only 9 sessions due to scheduling conflicts. The participation of each of the four study partners was evenly distributed between the first and second study sessions, and between the two TMS targets, $\chi^2(9) = 5.58$, $P = 0.781$. It was confirmed at the end of each session that participants had no prior familiarity with these study partners.

To minimize potential effects of reciprocity and reputation, individual decisions were kept private, with only total payoffs from randomly selected trials revealed at the end of each session. Decisions were incentive compatible and no deception was used: one decision each from the TMS participant and study partner was randomly selected to be paid out in each session, such that payoffs for each participant were affected by both one of their own choices and one choice of another participant. Specifically, across the two games played by each TMS participant in each session, one of the 40 choices was randomly selected, and the TMS participant received any amount kept and the relevant partner received any amount passed. The TMS participant also received the amount passed from one randomly selected trial from one of the partners, either the partner who was physically present (who had previously completed the DG once) or a random trial from the last TMS participant to have participated in the study. The study partner also always received the amount they kept from one randomly selected trial and the amount they passed on that trial was added to the TMS participant payment if the random trial selected for the TMS participant above was a trial in which they interacted with the study partner.

Analysis

Analysis of this within-subjects repeated measures 2×2 design included both planned (an analysis pre-registration was created partway through data collection before beginning analysis: https://osf.io/xenrb) and exploratory approaches. Only the first valid response to each trial was analyzed (i.e. participants occasionally entered held and passed amounts that did not total the amount of tokens on a given trial, at which point the trial was automatically repeated until a valid set of allocations was made). Planned statistical models included a 2 (TMS condition: rTPJ *vs* vertex) × 2 (social condition: minimally identified *vs* physically present interaction partner) repeated measures ANOVA for each dependent variable of interest and post-hoc paired *t*-tests to test potential main effects of TMS condition, main effects of social condition and interactions between TMS condition and social condition. A main effect of TMS condition would support a general role for the TPJ, for example, in reducing selfishness by promoting aversion to advantageous inequity, whereas a TMS condition by social condition interaction would support a role for the TPJ that depends on the salience and physical presence of the social

partner. Hedge's g_{av} is reported as a less biased effect size than Cohen's d for mean comparisons (Lakens, 2013).

We focused on two dependent variables of interest, the amount of advantageous inequity created and the amount of disadvantageous inequity created. Advantageous inequity was the difference between the value kept and the value shared when this favored the self, and disadvantageous inequity was the difference between the value kept and the value shared when this favored the other. Both measures were summed across the 20 trials of the task and separate ANOVAs and related post-hoc comparisons were run for each measure.

These analyses deviated from the analysis pre-registration in two ways. First, we had originally planned to evaluate TMS effects on four measures: advantageous inequity, disadvantageous inequity, total inequity (advantageous plus disadvantageous) and generosity (the total amount shared). The other two measures, total inequity and generosity, are functions of both advantageous and disadvantageous inequity. Total inequity increases with both advantageous and disadvantageous inequity, whereas generosity decreases with advantageous and increases with disadvantageous inequity. Our plan to analyze all four measures was based the results of Sáez et al. (2015), in which each of the four measures could be distinguished from the others. However, in contrast to Sáez et al. (2015), few of our participants consistently created disadvantageous inequity, and so total inequity and generosity were highly correlated with each other and with advantageous inequity in our study. As such, examining all four measures independently was no longer appropriate, and we thus chose to focus on advantageous and disadvantageous inequity given the extensive theoretical and empirical support for these constructs (Fehr and Schmidt, 1999; Charness and Rabin, 2002; Morishima et al., 2012; Gao et al., 2018; Bruhin et al., 2019). Results for total inequity and generosity are presented in the Supplement.

Second, we had originally planned to analyze model-based measures of advantageous and disadvantageous inequity aversion, as in Sáez et al. (2015). However, we found that these model parameters could not be estimated accurately in some individual subjects. (Sáez et al. (2015) only estimated these parameters at the group level, not the individual subject level as we had planned.) Therefore, we focus on the simpler, model-free estimates of advantageous and disadvantageous inequity described above and report the originally planned model-based measures in the Supplement. We observe the same pattern of results with the model-based measures.

We also report several exploratory analyses in the Supplement. We report an exploratory robustness check on our central analyses in which we control for participant gender. The reported pattern of results holds with this covariate. Also, as an alternative to repeated measures ANOVA, we report exploratory analyses using generalized estimating equations (GEE). The reported pattern of results holds with this alternate method of analysis.

Model-free analyses were conducted in R version 3.6.0. Data and analysis scripts are available at https://osf.io/5ky3x/.

Results

We first confirmed that participants showed the expected effects of the exchange rate in the task, creating the greatest unequal splits benefiting themselves when the exchange rate was favorable to themselves and the greatest unequal splits benefiting the partner when the exchange rate was favorable to the

partner. We examined how participant behavior varied as a function of the exchange rate in the task averaged across all four conditions. There were significant effects of exchange rate on both advantageous inequity, $F(1.24, 32.33) = 42.97$, $P < 0.001$, $\eta_p^2 = 0.622$, and disadvantageous inequity, $F(1.11, 28.90) = 8.28$, $P < 0.01$, $\eta_p^2 = 0.241$ (Figure 2). Advantageous inequity decreased significantly from the 3:1 to the 1:3 exchanges rates. Disadvantageous inequity was flat at about 0 for exchange rates favoring the participant, but then increased significantly from the 1:1 to 3:1 exchanges rates (Figure 2B). Thus, our participants responded to exchange rates in the expected manner, and in a similar direction as the participants in Sáez et al. (2015). However, our participants created less disadvantageous inequity than those in Sáez et al. (2015), with only a minority of participants creating relatively small amounts of disadvantageous inequity at the 1:2 and 1:3 exchange rates favoring the partner. This difference led us to depart from our pre-registered analysis plan, as described in the methods, although all originally planned analyses are reported in the Supplement.

Next, we turned to our planned analyses of the effects of TMS target and social condition on advantageous inequity and disadvantageous inequity. These analyses examined overall levels of these measures, collapsed across exchange rates.

We found no overall effects of TMS or partner on advantageous inequity, but we did find evidence that TMS reduced differences in advantageous inequity created for the two partners. For advantageous inequity, there was no significant main effect of TMS target, $F(1,26) = 0.55$, $P = 0.464$, $\eta_p^2 = 0.021$. Thus, TPJ TMS did not have a general effect on advantageous inequity created independent of the social partner. There was also no significant main effect of social condition on advantageous inequity, $F(1,26) = 1.03$, $P = 0.320$, $\eta_p^2 = 0.038$, although there was a significant TMS target × social condition interaction, $F(1,26) = 5.94$, $P = 0.022$, $\eta_p^2 = 0.186$. This effect was driven by differences between the partners in the control condition that were not present in the TPJ TMS condition. More advantageous inequity was created for the minimally identified than the physically present partner in the control condition, $t(26) = 2.24$, $P = 0.034$, Hedge's $g_{av} = 0.064$, but there were no differences in advantageous inequity created for the two partners following TPJ TMS, $t(26) = 0.73$, $P = 0.470$, Hedge's $g_{av} = 0.019$ (Figure 3). Thus, participants differentiated between the two partners in the control condition, acting more selfishly toward the minimally identified partner than the physically present partner, but this difference appears to have been eliminated with TMS to the TPJ.

Our participants created relatively little disadvantageous inequity, hindering our ability to detect any effects on disadvantageous inequity. We found no overall effects of TMS or partner, or interaction between TMS and partner. For disadvantageous inequity, there was no main effect of TMS target, $F(1,26) = 0.54$, $P = 0.469$, $\eta_p^2 = 0.020$, no main effect of social condition, $F(1,26) = 1.12$, $P = 0.299$, $\eta_p^2 = 0.041$ and no TMS target × social condition interaction, $F(1,26) = 0.15$, $P = 0.704$, $\eta_p^2 = 0.006$ (Figure 4).

This pattern of results, and particularly the reduction of differential advantageous inequity toward the two social partners after rTPJ TMS, was robust across several alternative analyses. This includes an exploratory analysis controlling for the participant gender, an exploratory analysis using GEE rather than repeated measures ANOVA and analysis using model-based rather than model-free estimates of advantageous inequity aversion (see the Supplement).
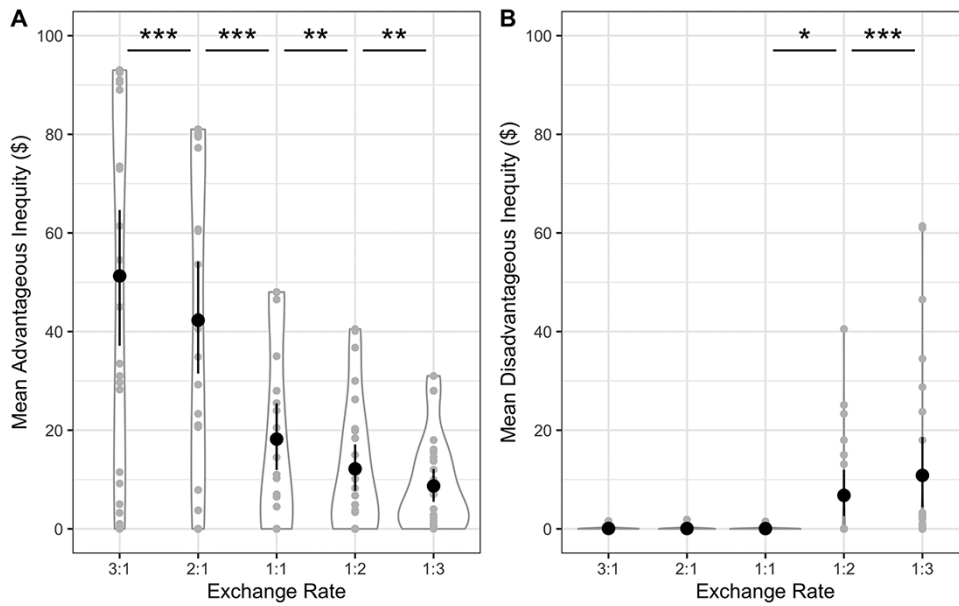
Fig. 2. Mean advantageous inequity and disadvantageous inequity, summed within condition and then averaged across the four conditions by exchange rate. Individual participant data points, densities, and means with bootstrapped 95% confidence intervals are displayed. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ (non-parametric comparisons).
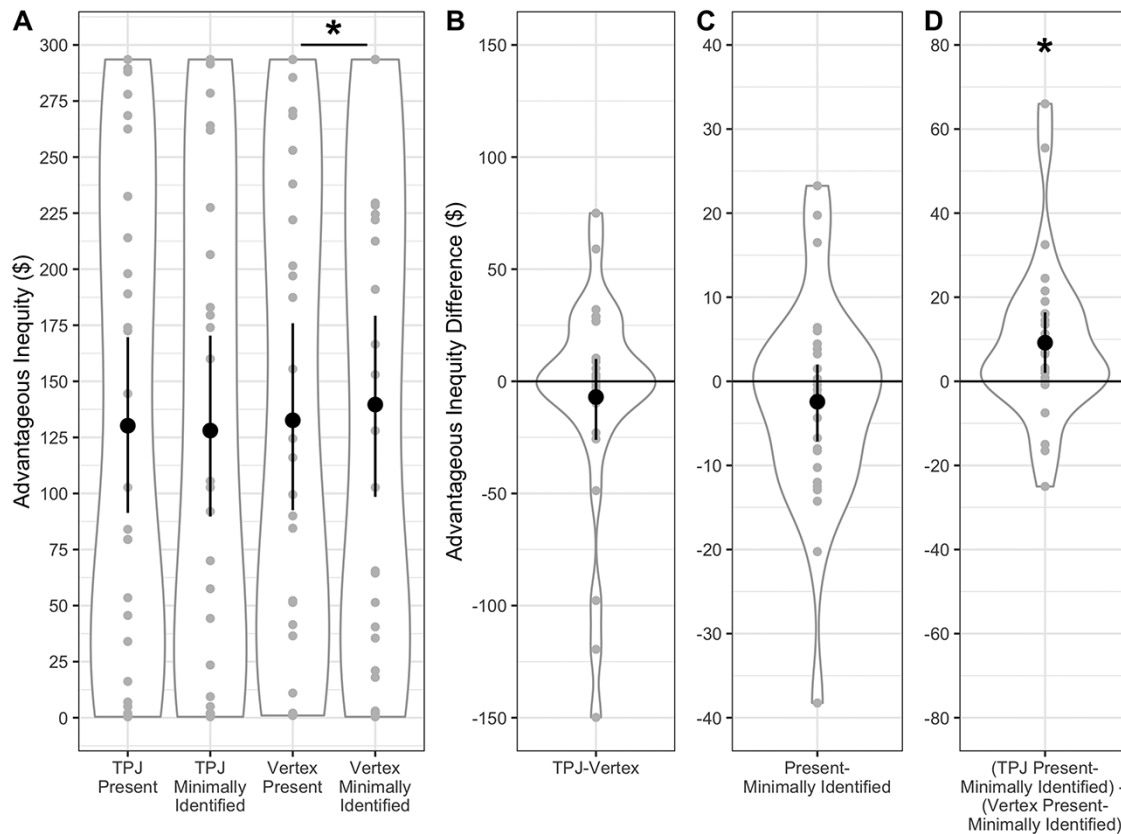

Fig. 3. Advantageous inequity, by TMS and social condition. Note that plots of choice behavior in each of the four conditions emphasize the variability between participants due to known individual differences in social preferences (e.g. Morishima *et al.*, 2012; Yamagishi *et al.*, 2014). For this reason, we also present the within-subjects comparisons on which inference is based in our within-subjects design. (A) Advantageous inequity, summed by condition. Maximum advantageous inequity in a given condition is $293.50. All participants created at least some advantageous inequity in each condition (i.e. all graphed values are greater than 0). Condition differences for the main effect of TMS target (B), main effect of social condition (C) and TMS target by social condition interaction (D). A significant TMS target by social condition interaction was driven by a significant difference between interaction partners in the vertex condition but not the TPJ condition. Individual participant data points, densities and means with bootstrapped 95% confidence intervals are displayed. *$P < 0.05$.
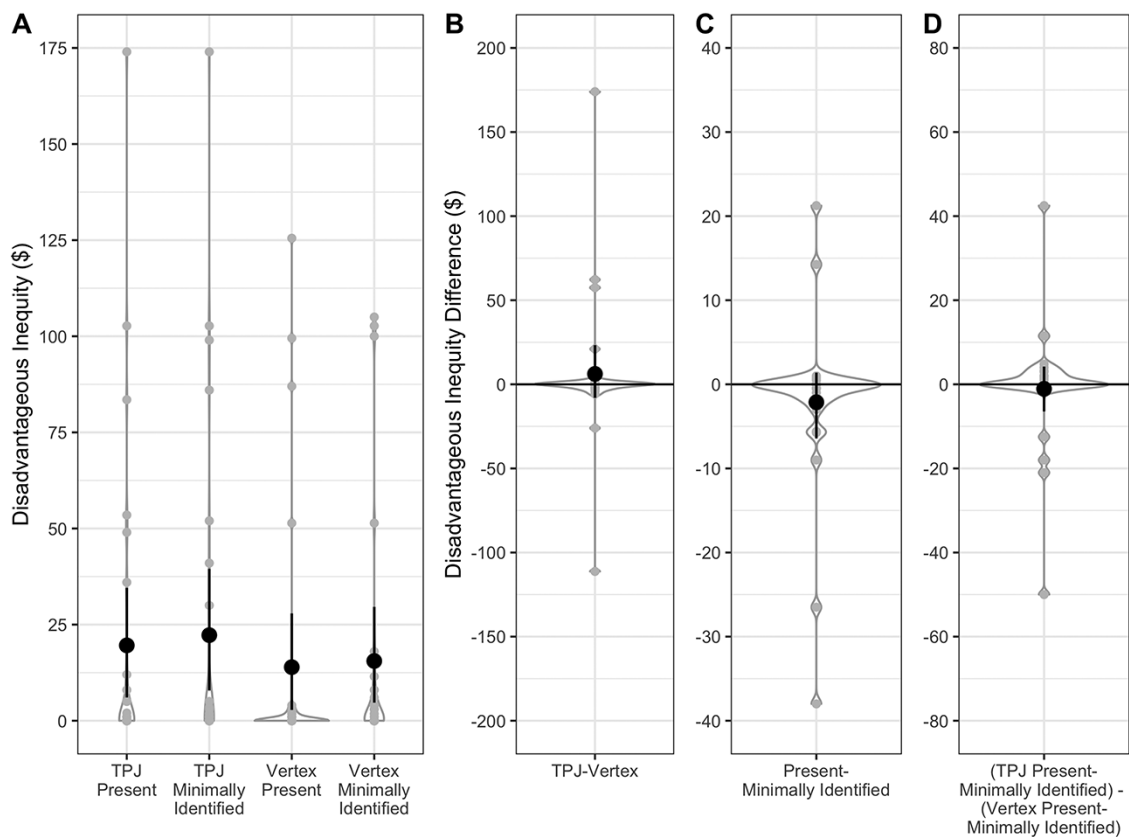
Fig. 4. Disadvantageous inequity, by TMS and social condition. (A) Disadvantageous inequity, summed by condition. Maximum disadvantageous inequity in a given condition is $293.50. Condition differences for the main effect of TMS target (B), main effect of social condition (C) and TMS target by social condition interaction (D). No main effects or interactions were observed. Individual participant data points, densities and means with bootstrapped 95% confidence intervals are displayed.

Discussion

The TPJ has been implicated in mentalizing across an extensive body of research, but the causal role of the TPJ in decisions about sharing resources with others has been less clear. Here, we used putatively inhibitory repetitive TMS to examine the causal role of the rTPJ in a modified Dictator Game. rTPJ TMS did not affect overall levels of advantageous inequity, when one selfishly holds on to more resources for oneself. Instead, we found evidence that rTPJ TMS reduces differences in the advantageous inequity created for different social partners (physically present *vs* minimally identified). Under control conditions, participants tended to selfishly hold on to more money when interacting with an abstract minimally identified partner than when interacting with a salient physically present partner. After rTPJ TMS, this effect of social partner on advantageous inequity was reduced—participants tended to treat the two interaction partners more similarly following modulation of the rTPJ. Building on prior work, this suggests that the effect of the rTPJ on social decisions may depend on the salience or identifiability of the 'other' with which one is interacting.

The rTPJ did not appear to play a general role in advantageous inequity aversion, as rTPJ TMS did not affect overall levels of advantageous inequity. This finding is seemingly at odds with prior work implicating this region in altruistically sharing resources, particularly with strangers (Tankersley *et al.*, 2007; Morishima *et al.*, 2012; Hutcherson *et al.*, 2015; Strombach *et al.*, 2015; Soutschek *et al.*, 2016; Tusche *et al.*, 2016). However, some previous imaging evidence is subject to alternative explanations,

for example that TPJ activation reflects conflict between motives rather than altruistic motives *per se* (Morishima *et al.*, 2012; Hutcherson *et al.*, 2015), and previous studies that have used TMS to examine the causal role of the rTPJ have not found consistent effects. One previous TMS study found that inhibiting rTPJ increased selfishness toward more socially distant relationships (Soutschek *et al.*, 2016), whereas another found that it reduced weighting of selfish material relative to moral concerns (Obeso *et al.*, 2018). Rather than the TPJ, the drive to altruistically share resources may depend more critically on affective empathy regions, including the anterior insula and amygdala (Tusche *et al.*, 2016). While such medial regions are not easily accessible by brain stimulation techniques, indirect modulation of these regions through lateral functional connections (e.g. Oathes *et al.*, 2021) may be a useful future approach for determining the causal mechanisms for altruism.

Instead, we found some evidence that rTPJ TMS tended to reduce the differences in advantageous inequity created for different social partners. It is important to note that this effect was relatively small, and strict correction for the multiple outcomes we had originally planned to examine would yield a marginal result. That the effect appears consistently across multiple alternative analysis approaches provides some confidence, but the small effect size highlights the clear need for a replication of this finding.

Further research would be needed to determine the process by which the rTPJ influences selfishness depending on the social context. The task used here did not require strategic decision

making, since reciprocity and reputation influences were eliminated by the one-shot nature of the decisions. Yet, the rTPJ may still have influenced differences in selfishness toward physically present *vs* minimally identified partners via second-order mentalizing. Other work has highlighted the necessary role of the TPJ in considering others' intentions and beliefs in strategic decisions (Hill *et al.*, 2017). Here, the rTPJ may be necessary for differentially considering what the interaction partners knew or believed about the TMS participant, influencing the participant's differential aversion to holding on to a greater portion of the resources to be shared. Alternatively, it has been argued that the TPJ helps to resolve the conflict between material (self-benefit) and moral (fairness) concerns in resource allocation decisions (Obeso *et al.*, 2018), and it is possible that this conflict resolution process might differ for more *vs* less salient others.

Participants did not have substantially more experience with the physically present interaction partner than the minimally identified partner—only a brief interaction and sitting in the same room with this stranger while the Dictator Game was completed. Thus, simply putting a name and face to an interaction partner and having this person nearby during decisions made participants somewhat more averse to creating advantageous inequity. This mirrors the identifiable victim effect, in which people are more generous in their donation behavior when provided with a photo of the recipient (Small and Loewenstein, 2003; Genevsky *et al.*, 2013). Our work thus provides initial evidence that the TPJ may be necessary for the effect of identifiability on decisions to share resources. Such an effect would be consistent with other work on parochial biases in punishment. Specifically, rTPJ TMS reduces a bias toward greater third-party punishment of outgroup relative to ingroup members for non-cooperative behavior (Baumgartner *et al.*, 2014). The TPJ may thus be especially relevant for social decision making based on real or perceived social closeness (Baumgartner *et al.*, 2014; Soutschek *et al.*, 2016).

Several limitations should be noted. Few participants consistently created resource allocations in which they had less money than their partner, limiting our ability to determine the effects of TPJ TMS on disadvantageous inequity. Further, we did not directly measure the effects of TMS on underlying neural activity. Although the TPJ is a relatively functionally heterogenous structure (Mars *et al.*, 2012), our study was designed to target the TPJ subregion responsible for mentalizing processes. We have further confidence that rTPJ TMS did not influence attention, rather than mentalizing, given that variability in choices did not differ across conditions (see model-based results in the Supplementary material). Future work should extend our findings by pursuing individualized functional targeting of the mentalizing network.

Based on the current results, the TPJ may not be critical for overall levels of aversion to advantageous inequity, but rather it may be critical for differentiating between social partners in resource allocation decisions. Even in the non-strategic task employed here, the rTPJ appears to play a causal role in selfishly keeping more resources when interacting with a partner who was not physically present compared to one who was. Excitatory stimulation of brain hubs in empathic networks has been suggested as a promising avenue for treatment of psychiatric conditions characterized by empathic deficits (Wassermann and Zimmermann, 2012; Schuwerk *et al.*, 2014; Cotter *et al.*, 2017). However, the current results highlight that excitatory stimulation of a cognitive empathy hub like the TPJ could increase parochialism (differential treatment of different social partners) rather than decrease selfishness. This potential role of the TPJ in biasing selfishness

toward abstract relative to in-person interactions holds particular relevance in a modern social world in which many interactions occur remotely with the decreased identifiability offered by avatars, screennames and email addresses.

Conflict of interest

The authors declared that they had no conflict of interest with respect to their authorship or the publication of this article.

Supplementary data

Supplementary data is available at *SCAN* online.

Open practices statement

The analysis plan for this study was pre-registered on the Open Science Framework partway through data collection but before analysis began: https://osf.io/xenrb.

Data and analysis scripts are also available through the Open Science Framework: https://osf.io/5ky3x/.

References

Ah Sen, C.B., Fassett, H.J., El-Sayes, J., Turco, C.V., Hameer, M.M., Nelson, A.J. (2017). Active and resting motor threshold are efficiently obtained with adaptive threshold hunting. *PLoS One*, **12**(10), e0186007.

Baumgartner, T., Schiller, B., Rieskamp, J., Gianotti, L.R., Knoch, D. (2014). Diminishing parochialism in intergroup conflict by disrupting the right temporo-parietal junction. *Social Cognitive and Affective Neuroscience*, **9**(5), 653–60.

Bruhin, A., Fehr, E., Schunk, D. (2019). The many faces of human sociality: uncovering the distribution and stability of social preferences. *Journal of the European Economic Association*, **17**(4), 1025–69.

Carter, R.M.K., Bowling, D.L., Reeck, C., Huettel, S.A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science*, **337**(6090), 109–11.

Charness, G., Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, **117**(3), 817–69.

Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C.Y., Barnett, J.H. (2017). Social cognitive dysfunction as a clinical marker: a systematic review of meta-analyses across 30 clinical conditions. *Neuroscience and Biobehavioral Reviews*, **84**, 92–9.

Fehr, E., Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, **114**(3), 817–68.

Frith, C.D., Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, **63**, 287–313.

Frith, C.D., Singer, T. (2008). The role of social cognition in decision making. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **363**(1511), 3875–86.

Frith, U., Frith, C. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **365**(1537), 165–76.

Gao, X., Yu, H., Sáez, I., *et al.* (2018). Distinguishing neural correlates of context-dependent advantageous- and disadvantageous-inequity aversion. *Proceedings of the National Academy of Sciences*, **115**(33), E7680–9.

Genevsky, A., Vastfjall, D., Slovic, P., Knutson, B. (2013). Neural underpinnings of the identifiable victim effect: affect shifts preferences for giving. *Journal of Neuroscience*, **33**(43), 17188–96.

Hill, C.A., Suzuki, S., Polania, R., Moisa, M., O'Doherty, J.P., Ruff, C.C. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nature Neuroscience*, **20**(8), 1142–9.

Hutcherson, C.A., Bushong, B., Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron*, **87**(2), 451–62.

Kanske, P., Bockler, A., Trautwein, F.M., Singer, T. (2015). Dissecting the social brain: introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and theory of mind. *NeuroImage*, **122**, 6–19.

Kanske, P., Böckler, A., Trautwein, F.M., Parianen Lesemann, F.H., Singer, T. (2016). Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Social Cognitive and Affective Neuroscience*, **11**(9), 1383–92.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, **4**, 863.

Mars, R.B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., Rushworth, M.F. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. *Cerebral Cortex*, **22**(8), 1894–903.

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C.C., Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, **75**(1), 73–9.

Oathes, D.J., Zimmerman, J.P., Duprat, R., *et al.* (2021). Resting fMRI-guided TMS results in subcortical and brain network modulation indexed by interleaved TMS/fMRI. *Experimental Brain Research*, **239**, 1165–78.

Obeso, I., Moisa, M., Ruff, C.C., Dreher, J.C. (2018). A causal role for right temporo-parietal junction in signaling moral conflict. *Elife*, **7**, e40671.

Ong, W.S., Madlon-Kay, S., Platt, M.L. (2020). Neuronal correlates of strategic cooperation in monkeys. *Nature Neuroscience*, **24**(1), 116–28.

Raine, A., Uh, S. (2019). The selfishness questionnaire: egocentric, adaptive, and pathological forms of selfishness. *Journal of Personality Assessment*, **101**(5), 503–14.

Sáez, I., Zhu, L., Set, E., Kayser, A., Hsu, M. (2015). Dopamine modulates egalitarian behavior in humans. *Current Biology*, **25**(7), 912–9.

Santiesteban, I., Banissy, M.J., Catmur, C., Bird, G. (2012). Enhancing social ability by stimulating right temporoparietal junction. *Current Biology*, **22**(23), 2274–7.

Saxe, R., Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in "theory of mind". *NeuroImage*, **19**(4), 1835–42.

Saxe, R., Powell, L.J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science*, **17**(8), 692–9.

Schuwerk, T., Langguth, B., Sommer, M. (2014). Modulating functional and dysfunctional mentalizing by transcranial magnetic stimulation. *Frontiers in Psychology*, **5**, 1309.

Small, D.A., Loewenstein, G. (2003). Helping a victim or helping the victim: altruism and identifiability. *Journal of Risk and Uncertainty*, **26**, 5–16.

Soutschek, A., Ruff, C.C., Strombach, T., Kalenscher, T., Tobler, P.N. (2016). Brain stimulation reveals crucial role of overcoming self-centeredness in self-control. *Science Advances*, **2**(10), e1600992.

Strombach, T., Weber, B., Hangebrauk, Z., *et al.* (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences*, **112**(5), 1619–24.

Tankersley, D., Stowe, C.J., Huettel, S.A. (2007). Altruism is associated with an increased neural response to agency. *Nature Neuroscience*, **10**(2), 150–1.

Tusche, A., Bockler, A., Kanske, P., Trautwein, F.-M., Singer, T. (2016). Decoding the charitable brain: empathy, perspective taking, and attention shifts differentially predict altruistic giving. *Journal of Neuroscience*, **36**(17), 4719–32.

Wassermann, E.M., Zimmermann, T. (2012). Transcranial magnetic brain stimulation: therapeutic promises and scientific gaps. *Pharmacology and Therapeutics*, **133**(1), 98–107.

Yamagishi, T., Li, Y., Takagishi, H., Matsumoto, Y., Kiyonari, T. (2014). In search of *Homo economicus*. *Psychological Science*, **25**(9), 1699–711.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, **8**, 665–70.

Young, L., Cushman, F., Hauser, M., Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, **104**(20), 8235–40.

Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R. (2010a). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, **107**(15), 6753–8.

Young, L., Dodell-Feder, D., Saxe, R. (2010b). What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia*, **48**(9), 2658–64.