

Self-oriented or other-oriented empathic concern behind altruism

Zih-Yun Yan

Supervisor: Joe Kable, Kristin M. Brethel-Haurwitz

CAPSTONE PAPER

Abstract

It is hypothesized that empathic concern evokes altruistic motivation (Batson, 1991). As we can see in our daily life, stimulating empathy to the suffering is a common advertising strategy for charitable donation. While empathizing, we adopt the perspective of others and share their feelings so we can understand their need. Then, these empathic responses motivate us to concern others' well-being and save them from any negative outcomes. However, whether altruistic behaviors are truly other-oriented or actually self-benefit motivated is still controversial. In this study, we focus on the empathy network in the human brain and use Multi-Voxel Pattern Analysis (MVPA) to provide new evidence in this debate. Adapting an established protocol of empathy-for-pain studies (Singer et al. 2004, 2006; Hein et al. 2010), we tested whether the neural activities of empathy can predict altruistic behaviors and how kin relationship modulate the willingness to take altruistic actions. In the experiment, daughters faced two types of conditions: in "Forced Choices" trials, subjects either passively received the shock or observed their mothers or strangers receiving the shock; in "Free Choices" trials, daughters had to actively decide whether to receive the shock themselves or to defer the shock to mothers and strangers. We find that when daughter chose to sacrifice themselves to receive the shock, the neural pattern in empathy network is more similar to when daughters themselves were in pain rather than observing others in pain. These finding suggest that altruistic choices are self-oriented process. We do not find a distinct neural pattern when subjects had to make the altruistic choices facing their mother or a stranger, however, the shock deferring rate to stranger is significantly higher than mother at the behavior level.

Introduction

When natural disasters or civic wars happen and destroy the countries, we can always endlessly see and hear the tragedies reported on the media. Every time during this period, people around the world share their distressed feelings with the suffering and donate the money in order to help them and rescue them from the struggle. However, this helping behavior has been thought of as a bizarre behavior that ever happen in the animal world as the acts seems not benefiting the self but instead adding the cost to the self. Therefore, scientists are curious about what motivates these ostensibly other-regarding behaviors. Based on evolution theory, helping behaviors happen when the individual we are helping with share the relatedness with the helper (Hamilton, 1964) or when considering that temporarily helping other can receive the reciprocal return in the later time (Trivers, 1971). Social psychologists have tried to understand why human beings help others by proposing two contrasting motivations: egoistic motives and true altruism. Viewed from the egoistic perspective, helping behaviors are produced in order to make themselves feel better explained by the negative state relief model (Cialdini, Darby, & Vincent, 1973; Cialdini, Kenrick, & Baumann, 1982). The negative state relief model suggested that witnessing others' suffering can produce negative affect, a temporal sorrow. Thus, to restore the mood, helping behaviors had been instrumentally used to solve the negative state. Also suggested by the arousal cost reward model, to reduce the vicarious emotional arousal, after weighing the cost of helping others, people decide to help or not, thus if some other present and are able to help, they are more unlikely to help (Piliavin et al., 1981). These models hold a premise that the so-called pro-social behaviors are only out of self-interest.

Other researchers suggest that helping behaviors out of the concern of other's well-being rather than self-benefit truly exist, which is referred to as true altruism. Batson and his colleagues

propose the empathy-altruism model to argue that the cause of helping is the concern of the victim's own good which is evoked by our empathic responses. In contrast to the relief of personal distress, empathic concern makes us to put ourselves in others' shoes and to imagine the negative situation the victim is in. At this point, the victim's well-being becomes the primary concern rather than ourselves'. A distinct difference between the arousal cost reward model and the empathy-altruism model is that the empathy-altruism model can explain why someone would sacrifice themselves to help others, because under the arousal cost reward model, if perceiving no cost to escape from helping others, individuals would rather not engage in the helping behaviors as escaping can also reduce the aversive arousal. To account for these other potential explanations of helping behaviors, Baston conducted a series of researches to tackle each of the self-oriented motives behind the altruism. He found that individuals who had high empathy would still choose to help others even if escaping is an easy option (Baston et al., 1981; Fultz, Batson, Fbrtenbach, McCarthy, & Varney, 1986). In their another experiment, they further tested whether helping behaviors are caused by social reward (honor, praise) or social punishment (guilt, shame). They found when the subjects encountered someone in need, they would still feel better if the victim is not relieved due to their own help and when they realized that there would not be any socially-mediated punishment to the failure to help, the helping behaviors were also not diminished (Baston et al., 1988) .

However, there is still some controversy about whether a true boundary between self and other, the most crucial assumption behind the empathy-altruism hypothesis, exists. In the experiment of Cialdini, Briwn, Lewis, and Neuberg (1997), they found stronger closeness of self and other's relationship can lead to greater empathic concern and predict helping behaviors. Based on this result, they argued that the stronger closeness for the other person can be viewed as

perceiving more of themselves in the other, which provides an alternative explanation of a self-oriented process that can also become the source of empathic responses. To further resolve this Self and Other debate, later on with the help of the neuroscientific tool fMRI, social neuroscientists started to find the common and distinct neural network of the perception of self and other as well as empathy neural network (e.g. Decety & Chaminade, 2003, Singer et al., 2004, Ochsner et al., 2008). This emerging line of research can potentially illuminate the nature of altruism. A particular empathy-for-pain experiment was designed to explore empathy for pain and self-other representation in the fMRI (Hein et al., 2010; Singer et al., 2004, 2006). The neural correlates of empathy for pain show a greater overlap of neural networks of self-experienced pain and observing other-experienced pain (Jackson et al., 2007; Oschner et al., 2008; Zaki et al., 2007) in dorsal anterior cingulate cortex (dACC) and anterior insula (AI), as well as regions associated with imagining others' emotions in parietal cortex.

Following Batson's work in finding out how empathy leads to egocentric or altruistic helping behaviors, recent neural studies focus on further conceptualization of empathy network in the brain. Recent studies found helping behavior that is considered mainly caused by "empathy care" (care for others well-being) rather than "empathy distress" (viciously experiences victims' feelings) has a distinct brain system which involves nucleus accumbens and medial orbitofrontal cortex, whereas, the empathy distress is preferentially associated with premotor and somatosensory cortical activity (Hare et al., 2010; Ashar et al., 2017).

Built upon the previous studies, in our study, we want to utilize the finding of distinct neural response between self versus others in pain to examine which one is more similar to the brain activity experienced when one decides to sacrifice on behalf of the other. Therefore, we defined seven regions-of-interest (ROIs; dACC, bilateral AI, bilateral Parietal cortex, bilateral,

right striatum and left occipital cortex) through a meta-analysis of empathy-for-pain literatures to further isolate neural similarities among the experimental conditions and used them as our “brain marker”. Later, we use machine-learning based analysis to decode how the neural pattern in these brain markers can predict the altruistic behaviors. In the present research, we explore subjects’ willingness to sacrifice to others with varied social distance. Our subjects had to either passively received a shock or see someone else be shocked, or they decided whether to take a shock on behalf of either her mother or the stranger. By this experimental design, we can find out whether it’s self-oriented or other-oriented brain activities more similar to helping behaviors and how these so-called empathy sharing regions are in association with the helping actions which is taken later on.

Methods

Participants

Seventeen adult female subjects (mean age = 24.5) participated with their mother (mean age = 52) and stranger who matched their mother’s age and race. Two subjects were excluded from the fMRI analysis because of excessive movement.

Experimental Design

During the experiments, all subjects (daughter, mother and stranger) received shock at four different levels that are calibrated with each subjects’ tolerance of pain. However, only the daughters were in the scanner. The two other participants (the mother and stranger) were in the control room of the MRI scanner. There were two types of trials, Forced Choice trials and Free Choice trials. In Forced Choice trials, the shock was passively delivered to the designated

subject. In the Free Choice trials, the daughter would have to choose between either themselves versus their mother to receive the shock or themselves versus the stranger to receive the shock. The goal of this design is to collect the neural signal when anticipating receiving the shock themselves and shock in others as well as when the daughter is making altruistic choices: sacrificing themselves to receive the shock or deferring the shock to others.

There were 142 trials presented across four runs, in which 78 trials are Forced Choice trials and 64 trials are free choice trials. In Forced Choice trials, the subject would see two options: the recipient of the shock with the shock level and “no choice”, they were instructed to select the assigned button. Forced Choice trials were presented at every shock level (five level-1, five level-2, seven level-3, and nine level-4; 26 total), and were identical for participants, their mothers and the strangers.

In Free Choice trials, the daughters chose whether to sacrifice themselves and receive a shock or to defer the shock to the other subject (either the mother or the stranger). Every comparison of each level between the daughter and the other subject (e.g., daughter level-1 vs. mother level-2) was made twice (32 trials each). The mother and stranger were always presented on the left. The daughter was presented on the left in Forced Choice trials and on the right in Free Choice trials. The two types of trials were mingled within each run.

Each trial was approximately 20 seconds. At the beginning of the trial, participants had a 1-second fixation, and then would have a choice period. Their name and shock intensities were displayed on the screen. The choice they chose were highlighted in red and then a pre-shock jitter was shown before the shocks were delivered. The shock was delivered to the pre-determined recipient in Forced Choice trials or to the choice of subjects in Free Choice trials. Shocks were sent following a warning cue to the chosen participant for approximately 1 second.

Images (either a lightning bolt for the daughter, or live video feed of the mother's hand and stranger's hand) were displayed for 4 seconds.



Figure 1. The design of a trial

fMRI Image acquisition

The MRI images were acquired using a 3 Tesla MR scanner at the University of Pennsylvania Hospital with a 32-channel head coil. In all runs, the blood oxygenation level-dependent (BOLD) signals were acquired with following parameters: interleaved acquisition of axial slices covering the whole brain; slice thickness of 3mm; TR=3s; TE=30; flip angle; field of view= 192mm; matrix size=64x64.

fMRI Image data preprocessing

The functional MRI data were processed using FSL (FMRIB Software Library, The University of Oxford, UK) with an event-related model. Images of each subject were realigned to the first image to correct for head movements. Two subjects were removed for excess movement. The images were normalized with a 3x3x3 mm voxel size and were smoothed with a full width at half maximum of 10 mm Gaussian kernel. A high-pass temporal filter cut-off of 150s was applied to remove any low-frequency drifts. Finally, we used interleaved slice-timing to correct the temporal shift between data acquisition.

fMRI Data Analysis

To determine whether an altruistic choice is a self-oriented or other-oriented process in terms of empathy response in human brain, we collected BOLD response from forced trials while subjects were anticipating either themselves or others to receive the shock. In the Free Choice trial, we collected signal while subjects were making choices about whether to sacrifice and received the shock themselves or defer the shock to others. We also collected the brain signal while the shock was delivered to either subjects or others. Particularly, we examined the brain pattern in pre-defined ROI brain areas. These ROI brain areas were obtained through coordinate-based meta-analysis. They are the results of a group of papers that included the brain coordinate data about neural signal contrast when observing others in physical pain versus observing others not in pain. The brain clusters are the dorsal anterior cingulate cortex (dACC), right and left anterior insula (AI), right striatum (R striatum), right and left parietal cortex (R Parietal, L Parietal), and the left lateral occipital cortex (L Occipital).

After data collection, we used multivoxel pattern analysis (MVPA) to compare the neural pattern between different experimental conditions. MVPA is a supervised classification technique to analyze spatially distributed patterns for specific functional brain activities. We created several classifiers to understand the relationship between the BOLD signals in predefined empathy neural network and experimental conditions. Our goal is to use these classifiers as a model of cognitive state to predict subjects' behavioral choices.

We estimated brain activation from both choice period and shock period. The brain activation was estimated through trial-by-trial beta values, in which the beta values for each trial was obtained through a General linear model with a regressor for that trial, another regressor for all other trials and the other regressor for all trials in shock receiving period for choice period signal estimation, and all trials in choice period for shock period signal estimation (Mumford,

Turner, Ashby, & Poldrack, 2012). These regressors were impulse responses which convolved with a double gamma hemodynamic response function (HRF). For each trial, we also included six motion parameters as control regressors. We repeated this estimation process for every trial to obtain all beta values and used these numbers to perform MVPA.

We implemented ROI analysis, meaning we only extracted voxels in each ROI to run the MVPA. The algorithm we used for classification is support vector machine (SVM). We used the toolbox LIBSVM (Chang & Lin, 2011) to perform the analysis. During the analysis, we split the data into three-fold. Two folds were used as training data set and the other fold was used as testing data set. We repeated this process three times for each model training. To increase the accuracy of SVM classifier, we test the optimal regularization parameter c within [0.001, 0.01, 0.1, 1, 10, 100, 1000] and found the optimal parameter that maximize the average accuracy across three folds.

Results

Forced Choice

Average predictive accuracy across 15 subjects in the regions determined by the meta-analysis were analyzed using a one-sample t-test. We tested the accuracy of the classifier trained with the neural response in the context of anticipation of self-received shock or observed other in shock. During the anticipatory phase of forced choice trial, the classifier has an above chance accuracy to predict the anticipation of self-versus other shock neural pattern from the highest to the lowest accuracy in L Parietal, R Insula, L Insula, Dacc, R Parietal, R Striatum and L Occipital ($p = 0.004$, $p = 0.01$, $p = 0.01$, $p = 0.02$, $p = 0.02$, $p = 0.04$, $p = 0.08$).

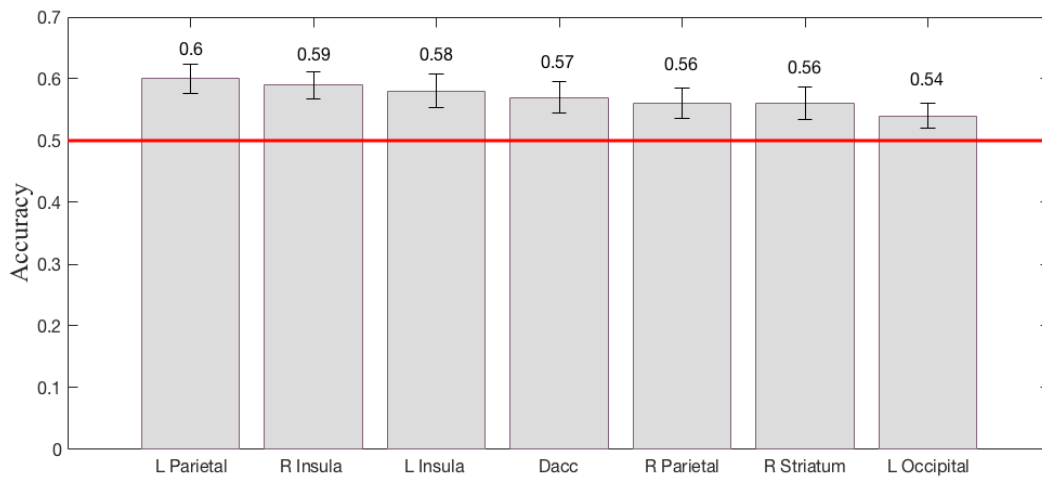


Figure 2 The predictive accuracy in distinguishing anticipation of self or other receiving the shock

Free Choice

Behavioral results

Across the fifteen daughters, 53% of them sacrificed more to receive the shock in Free Choice trial. They showed more altruistic concern to their mother than to strangers. Two of the subjects defer all free choices to other and did not sacrifice themselves at all.

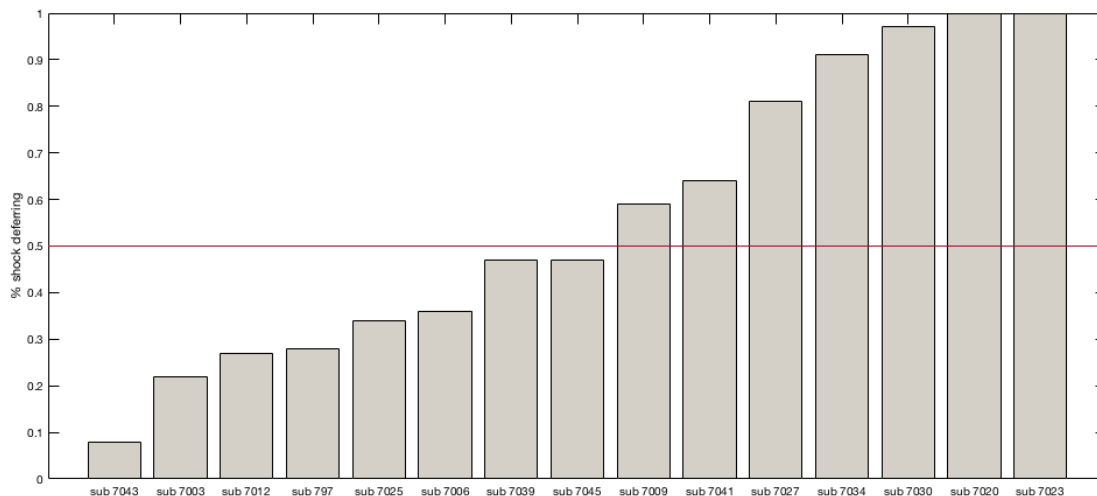


Figure 3 Individual difference in % of shock deferring

As predicted, subjects were more altruistic towards their own mothers than the strangers in which they deferred fewer shocks to their mothers ($p = .01$).

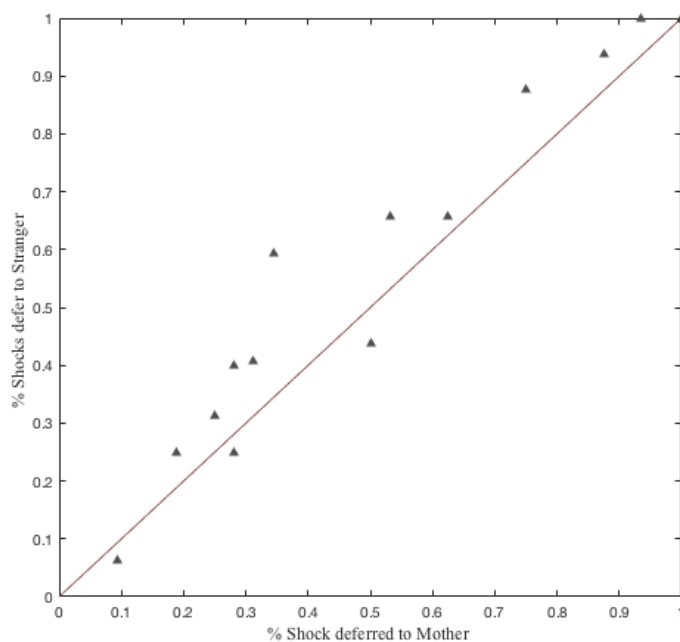


Figure 4 proportion of shock deferring between stranger and mother

Among the Fifteen daughters, two of them did not yield different choices in the free choice trials, thus they were not included into the free choice classification analysis due to the lack of labels. Average predictive accuracy of thirteen subjects in the regions determined by the meta-analysis were analyzing using a one-sample t-test. We tested the accuracy of the classifier trained with the neural response in the context of choosing to sacrifice themselves to receive the shock or defer the shock to others. During the anticipatory phase of free choice trial, the classifier has an above chance accuracy to predict the choices of sacrificing and receiving the shock and deferring the shock to others in L Insula, L Parietal, R Parietal, dACC, R Insula, R Striatum and L Occipital ($p = 0.0006$, $p = 0.03$, $p = 0.0009$, $p = 0.005$, $p = 0.07$, $p = 0.03$, $p = 0.46$).

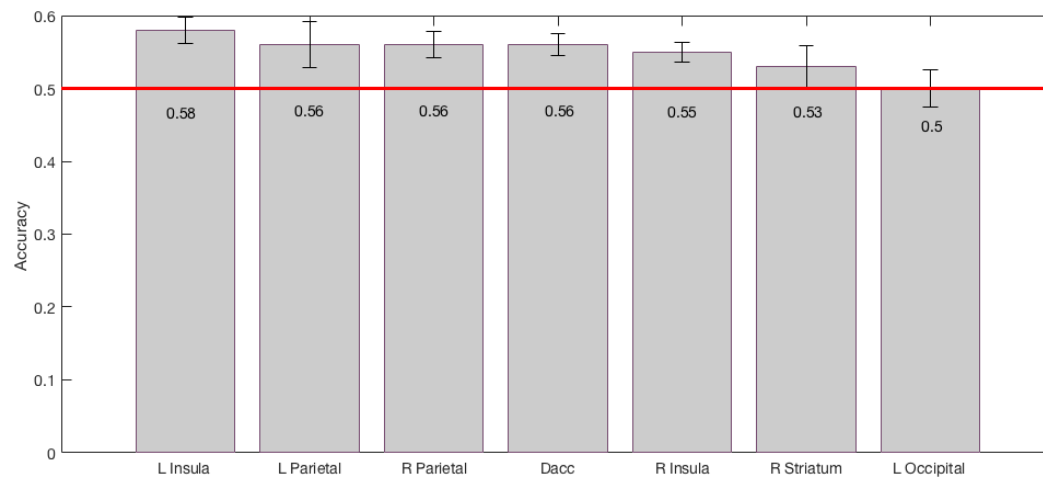


Figure 5 Predictive accuracy in distinguishing deciding self or other to receive the shock

Mother versus stranger

The classifier could not distinguish the neural pattern between the self-versus mother choices and self-versus stranger choices in ROIs. ($p > 0.1$)

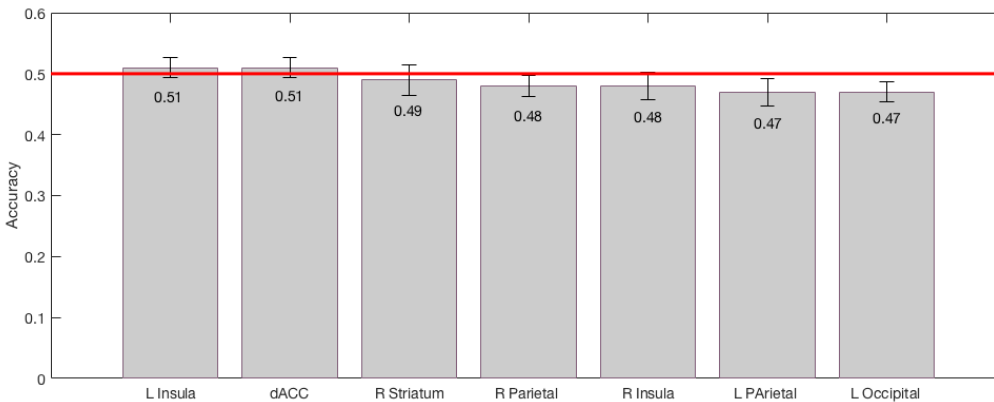


Figure 6 Predictive accuracy in distinguishing self & mother or self & stranger decision trials

Forced choice predicts free choice

The main idea of our study is to test whether the neural response when viewing someone's in pain or self in pain can predict the helping behavior or not. Therefore, we trained a classifier with the Forced Choices trial neural pattern and used this classifier to predict the neural pattern in Free Choice trial. Namely, with the accuracy, we can infer which brain pattern is much more similar in helping behaviors. What we found in this analysis is that the accuracy is higher when used anticipation of self-pain in Force Choices to sacrificing behavior in Free Choices and anticipation of other-pain in Forced Choices to deferring behavior in Free Choices. (p = 0.003, p = 0.03, p = 0.04, p = 0.29, p = 0.0002, p = 0.0003, p = 0.04)

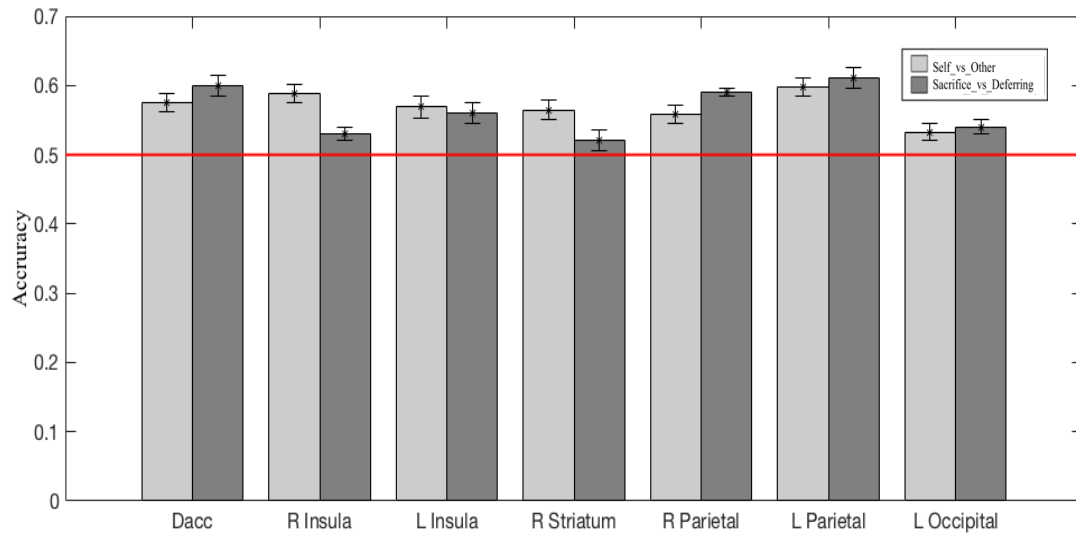


Figure 7 The predictive accuracy of using forced choices trials classifier in distinguishing sacrifice and deferring in free choices trials

Discussion

Here we found that the bilateral parietal cortex, the bilateral insula and the dACC showed distinct neural pattern when anticipating receiving the shock daughters themselves or observing others in shock in Forced Choices trials as well as when deciding to sacrifice and receive the shock or deferring the shock to others in Free Choices trials. Further, the classifier trained with the neural patterns during the anticipation of daughters receiving the shock and others receiving the shock was also able to predict the neural pattern in sacrificing or deferring choices. This result showed that when mapping the anticipation of receiving the shock to sacrificing rather than mapping it to deferring the shock to others would give us a higher than average predictive accuracy, suggesting that altruistic behaviors are associated with the self-oriented empathic response. This result provided another solid evidence in the debate about self versus other-oriented altruism debate. As previous literatures only identified the discrimination in neural response of self versus others in pain, they did not link this neural response to the decision of the altruistic behavior itself.

However, we could not use the neural pattern in these ROIs to predict the altruistic choices between when the Free Choices trials were self-versus mother or self-versus stranger trials, whether with single-voxel analysis, multi-voxel pattern analysis or searchlight analysis. As one of the major arguments for the self-oriented empathy perspective actually comes from the behavioral evidence that people usually sacrifice more to family members or close friends. Even though our behavioral results also reproduce this observation that they tend to sacrifice more for their mother than the stranger, at the neural level, we could not prove that the neural patterns of observing mother versus strangers are different. These results might be due to the fact that we only had thirteen subjects in our analysis and our approach would require more data to increase

the predictive accuracy originally. The number of data can be a really determining factor in our results. In future research, we should recruit more subjects and potentially cluster them into several subgroup with similar level of altruistic tendencies, then we can analyze their neural patterns at a more homogeneous level.

Another future direction of this research is to test the generality of our results beyond the empathy for pain paradigm. The relationship between empathy and altruism were studied with different designed empathy-related tasks as well as different forms of helping behaviors. Our meta-analysis was specifically targeted at the empathy-for-pain paradigm, therefore, it would also be interesting to formulate a more comprehensive meta-analysis with a broader range of empathy-related paradigm and test their predictive power on the altruistic behaviors. For example, we can also examines the regions that are reported to be preferentially in associated with “empathy care” and brain regions associated with “empathy distress” and to observe their predictive power individually.

In sum, the current study sheds light on the cognitive and neural mechanisms underlying empathic response in altruistic choices. It also contributes to our understanding of helping behaviors. These findings may aid in the development of the empathy research.

References

- Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of personality and Social Psychology*, *40*(2), 290.
- Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological inquiry*, *2*(2), 107-122.
- Batson, C. D., Bolen, M. H., Cross, J. A., & Neuringer-Benefiel, H. E. (1986). Where is the altruism in the altruistic personality? *Journal of Personality and Social Psychology*, *50*(1), 212.
- Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., & Cialdini, R. B., Darby, B. L., & Vincent, J. E. (1973). Transgression and altruism: A case for hedonism. *Journal of Experimental Social Psychology*, *9*(6), 502-516.
- Batson, C. D., Batson, J. G., Griffitt, C. A., Barrientos, S., Brandt, J. R., Sprengelmeyer, P., & Bayly, M. J. (1989). Negative-state relief and the empathy—altruism hypothesis. *Journal of Personality and Social Psychology*, *56*(6), 922.
- Cialdini, R. B., Kenrick, D. T., & Baumann, D. J. (1982). Effects of mood on prosocial behavior in children and adults. In *The development of prosocial behavior* (pp. 339-359). Academic Press.
- Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, *41*(2), 127-138.
- Griffitt, C. (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of personality and social psychology*, *55*(1), 52.

Hein, G., & Singer, T. (2010). Neuroscience meets social psychology: An integrative approach to human empathy and prosocial behavior. *Prosocial motives, emotions, and behavior: The better angels of our nature*, 109-125.

Hein, G., & Singer, T. (2010). Neuroscience meets social psychology: An integrative approach to human empathy and prosocial behavior. *Prosocial motives, emotions, and behavior: The better angels of our nature*, 109-125.

Neuberg, S. L., Cialdini, R. B., Brown, S. L., Luce, C., Sagarin, B. J., & Lewis, B. P. Does empathy lead to anything more than superficial helping? Comment on Batson et al. (1997).

Piliavin, J. A., Dovidio, J. F., Gaertner, S. L., & Clark, R. D., 111. (1981). *Emergency intervention*. New York: Academic.

Singer, T., Seymour, B., O'doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661), 1157-1162.

Singer, T., Seymour, B., O'doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661), 1157-1162.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, 439(7075), 466.

Ochsner, K. N., & Gross, J. J. (2008). Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. *Current directions in psychological science*, 17(2), 153-158.

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, 1156(1), 81-96.

Zaki, J., Ochsner, K. N., Hanelin, J., Wager, T. D., & Mackey, S. C. (2007). Different circuits for different pain: patterns of functional connectivity reveal distinct networks for processing pain in self and others. *Social neuroscience*, 2(3-4), 276-291.