Received: 16 November 2017 | Revised: 27 February 2018 | Accepted: 18 March 2018

DOI: 10.1002/hbm.24062

RESEARCH ARTICLE

WILEY

Common and distinct neural correlates of self-serving and prosocial dishonesty

Narun Pornpattananangkul[1,2,3] | Shanshan Zhen[1,2] | Rongjun Yu[1,2]

[1]School of Psychology, Center for Studies of Psychological Application and Key Laboratory of Mental Health and Cognitive Science of Guangdong Province, South China Normal University, Guangzhou, China

[2]Department of Psychology, National University of Singapore, Singapore, Singapore

[3]Mood Brain & Development Unit, Emotion & Development Branch, National Institute of Mental Health National Institutes of Health, Bethesda, Maryland

Correspondence
Narun Pornpattananangkul and Rongjun Yu, Department of Psychology, National University of Singapore, 9 Arts Link, Singapore, 117570.
Email: narunzhang@gmail.com; psyyr@nus.edu.sg

Funding information
National Natural Science Foundation of China, Grant/Award Number: 81771186; NMRC, Grant/Award Number: NMRC/OFYIRG/0058/2017; MOE Tier 2, Grant/Award Number: MOE2016-T2-1-01

Abstract

People often anticipate certain benefits when making dishonest decisions. In this article, we aim to dissociate the neural–cognitive processes of (1) dishonest decisions that focus on overall benefits of being dishonest (regardless of whether the benefits are self-serving or prosocial) from (2) those that distinguish between self-serving and prosocial benefits. Thirty-one participants had the opportunity to maximize their monetary benefits by voluntarily making dishonest decisions while undergoing functional magnetic resonance imaging (fMRI). In each trial, the monetary benefit of being dishonest was either self-serving or prosocial. Behaviorally, we found dissociable patterns of dishonest decisions: some participants were dishonest for overall benefits, while others were primarily dishonest for self-serving (compared with prosocial) benefits. When provided an opportunity to be dishonest for either self-serving or prosocial benefits, participants with a stronger overall tendency to be dishonest had stronger vmPFC activity, as well as stronger functional connectivity between the vmPFC and dlPFC. Furthermore, vmPFC activity was associated with decisions to be dishonest both when the benefits of being dishonest were self-serving and prosocial. Conversely, high self-serving-biased participants had stronger striatum activity and stronger functional connectivity between the striatum and middle-mPFC when they had a chance to be dishonest for self-serving (compared with prosocial) benefits. Altogether, we showed that activity in (and functional connectivity between) regions in the valuation (e.g., vmPFC and Str) and executive control (e.g., dlPFC and mmPFC) systems play a key role in registering the social-related goal of dishonest decisions.

KEYWORDS
executive control, deception, dishonest decision making, prosocial, self-serving

1 | INTRODUCTION

Oftentimes people expect some form of benefit when they decide to act dishonestly (e.g., deceiving others with false information or withholding some information from others) (Gneezy, 2005; Tenbrunsel, 1998). Yet, these benefits are not always self-serving. In times of natural disasters or terrorist attacks, for instance, government officials may decide to give false information to prevent panic among the public (Perry & Lindell, 2003). They may do so not only for the sake of their own careers, but also for public safety. In other words, this type of dishonest decision provides benefits for both self and others. In contrast, some dishonest decisions are purely self-serving. Some brokers, for example, recommend stocks to clients based not on their clients' best interests, but on the commission they can receive from trading these stocks (McDonald, 2002; Sanford, 2014). In fact, in this example, self-serving dishonesty provides self-serving benefits at the cost of others. While numerous cognitive-neuroscience studies have examined the neural basis of dishonest decision-making using techniques such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and optical imaging (Abe & Greene, 2014; Baumgartner, Fischbacher, Feierabend, Lutz, & Fehr, 2009; Ding, Gao, Fu, & Lee, 2013; Garrett, Lazzaro, Ariely, & Sharot, 2016; Greene & Paxton, 2009; Hu, Pornpattananangkul, & Nusslock, 2015; Maréchal, Cohn, Ugazio, & Ruff, 2017; Sun, Chan, Hu, Wang, & Lee, 2015; Yin & Weber, 2016), it was not until recently that researchers started to investigate the extent to which social-related goals of being dishonest (e.g., self-serving vs. prosocial) modulate neural cognitive processes of dishonest decision-making (Cui et al., 2018; Yin, Hu, Dynowski, Li, & Weber, 2017).

In an fMRI study, Yin et al. (2017) modified the Sender-Receiver Game (Gneezy, 2005) to study the modulatory role of social-related goals. In this study, participants decided whether to send a false message to another person for a higher payoff. In one condition, the higher payoff was for the participant him/herself (i.e., having a self-serving benefit), whereas in another condition, the higher payoff was for a charity (i.e., having a prosocial benefit). They found participants were more likely to be dishonest for a prosocial benefit than for a self-serving benefit. Moreover, being dishonest for a self-serving benefit, compared with a prosocial benefit, was associated with a stronger activity in the anterior insula (AI). Their study design, however, was criticized by Cui et al. (2018). According to Cui et al. (2018), the Sender-Receiver game may make participants too concerned about their self-image and reputation. They argued that, in this game, if participants decided to be dishonest, participants had to record their dishonesty on the computer while being observed by experimenters. This is unnatural and may divert participants from the incentives of being dishonest. Seeking to improve the dishonesty paradigm, Cui et al. (2018) employed the Coin-Guessing task (Greene & Paxton, 2009) in their EEG study. In this task, participants made dishonest decisions privately, undetected by experimenters. The study reported a reduction in the N2 component when participants were dishonest for prosocial, compared with self-serving, benefits. Critically, unlike Yin et al. (2017), Cui et al. (2018) found a higher propensity to be dishonest for self-serving benefits than for prosocial benefits. Because scalp-recorded EEG used in Cui et al.' (2018) study has a poor spatial resolution and may not detect signals from the brain areas that are further away from scalp (e.g., ventromedial prefrontal cortex [vmPFC] and striatum [Str]), the exact brain regions in which activity is modulated by social-related goals is still unknown.

More importantly, despite the progress made by these two recent studies (Cui et al., 2018; Yin et al., 2017), little attention has been paid to dissociate (1) the neural–cognitive processes underlying dishonest decisions made for overall benefits (regardless of whether the benefits are self-serving or prosocial) and (2) the processes underlying dishonest decisions made selectively for self-serving (as opposed to prosocial) benefits. Establishing this dissociation would allow us to examine different mechanisms of how people decide to be dishonest as a function of social goals. For instance, the neural-cognitive processes that are common across dishonest decisions for both self-serving and prosocial benefits may underlie decisions that do not concern whether the self was the beneficiary, such as when the government officials decide to provide somewhat false information about national disaster to prevent public panic (Perry & Lindell, 2003). The neural–cognitive processes that are specific to self-serving (compared with prosocial) benefits may instead underlie dishonesty behaviors people make strategically for self-serving benefits, such as brokers providing false recommendations (as mentioned above) (McDonald, 2002; Sanford, 2014).

Most cognitive–neuroscience studies on dishonest decision-making have focused exclusively on executive-control processing during decision-making, as reflected by enhanced activity in areas such as the dorsolateral prefrontal cortex (dlPFC), dorso/middle-medial prefrontal cortex (dmPFC/mmPFC), and inferior parietal lobe (IPL) (for meta-analysis see Lisofsky, Kazzer, Heekeren, & Prehn, 2014). For instance, Greene and Paxton (2009) showed an association between stronger activity in the dlPFC and a stronger tendency to be dishonest when being dishonest provides financial gains for the self. However, recent research has suggested an important role of reward and valuation processing, in addition to executive-control processing, in making dishonest decisions (Abe & Greene, 2014; Hu et al., 2015; Mazar & Ariely, 2006). Enhanced BOLD activity in the reward-related area of the striatum (Str) during a reward-processing task is associated with a higher frequency of dishonest decisions in a separate task (Abe & Greene, 2014). Similarly, enhanced reward-related event-related potentials (ERPs) predict a higher propensity to make dishonest decisions (Hu et al., 2015). For valuation-processing, enhanced BOLD activity in a valuation-related area, the ventromedial prefrontal cortex (vmPFC), is reliably shown across deception-related tasks (Mameli et al., 2016). One positron emission tomography (PET) study, for instance, showed an enhanced activity in the vmPFC when participants deceived the interrogator (Abe, Suzuki, Mori, Itoh, & Fujii, 2007). Yet, it is still unclear the extent to which social-related goals of being dishonest modulate reward and valuation processing during dishonest decision-making.

The valuation system plays a major role in reward and valuation processing (Bartra, McGuire, & Kable, 2013; Ruff & Fehr, 2014). The vmPFC and Str are two key regions in this system (O'Doherty, 2004; Schultz, Dayan, & Montague, 1997). Activity in the vmPFC is thought to represent decision-value signals (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008). That is, when making decisions, activity in the vmPFC usually correlates positively with subjective values of choice options (Bartra et al., 2013). Additionally, according to the "common currency" account (Chib, Rangel, Shimojo, & O'Doherty, 2009), the vmPFC employs a similar computation and representation of values across domains of choice options. For instance, stronger vmPFC activity corresponds to a decision to choose highly preferred food or merchandises as well as a decision to gamble when the potential payoff is subjectively high (Chib et al., 2009; Tom, Fox, Trepel, & Poldrack, 2007). In terms of prosocial decisions, stronger vmPFC activity is associated with decisions to donate to highly preferred charitable foundations (Hare, Camerer, Knoepfle, O'Doherty, & Rangel, 2010). Moreover, vmPFC traces subjective values of stimuli that are of high value from the perspective of oneself and others. For instance, when making choices for another person, participants' vmPFC activity is enhanced toward the options preferred by the other person, as opposed to the options preferred by the participants themselves (Nicolle et al., 2012). Thus, it might be reasonable to expect the vmPFC to trace the value of dishonest decisions in terms of the benefits they provide, regardless of whether the benefits are self-serving or prosocial.

As for the Str, researchers have reliably showed enhanced activity in the Str during anticipation and outcome phases of reward-processing (Diekhof, Kaps, Falkai, & Gruber, 2012). More recently, research has shown the enhancement of Str activity both when people obtain rewards for themselves and when they observe others obtain rewards (Braams et al., 2014; Mobbs et al., 2009; Ruff & Fehr, 2014). While this suggests the involvement of the Str in vicarious neural representation, the Str is also sensitive to the distinction between rewards for the self and rewards for others. For instance, although self-reported

pleasure from obtaining rewards for oneself and from observing others gain rewards correlate with activity in the overlapped region in the Str, this relationship is weaker when observing others gain rewards (Mobbs et al., 2009). Additionally, observing disliked others gain rewards reduces Str activity, and, conversely, observing disliked others lose rewards enhances Str activity (Braams et al., 2014). Thus, regarding dishonest decision-making, it might be reasonable to predict that people who decide to make dishonest decisions more predominantly for self-serving benefits (than for prosocial benefits) may elicit stronger Str activity in response to their own rewards, compared with others' rewards. If we take the aforementioned brokers' behaviors as an example (McDonald, 2002; Sanford, 2014), brokers who decide to make dishonest decisions for the sake of their own commissions may have stronger Str activity related to their own commissions than to their clients' benefits.

Altogether, both executive-control and valuation systems seem to play a part in dishonest decision-making. Therefore, it is possible that these two systems interact with each other when an individual decides whether to be dishonest. Recent social psychological theories have suggested that making dishonest decisions involves weighing economic benefits and psychological costs (e.g., being dishonest may challenge one's moral self) (Barkan, Ayal, Gino, & Ariely, 2012; Shalvi, Gino, Barkan, & Ayal, 2015). In cognitive neuroscience, the signals from regions in the executive-control system, such as the dlPFC, have been viewed as representing the psychological costs of being dishonest (Abe & Greene, 2014; Ding et al., 2013; Greene & Paxton, 2009; Hu et al., 2015; Lisofsky et al., 2014), while the signals from regions in the valuation system, especially the vmPFC, have been viewed as representing the evaluation of potential economic benefits (Hare et al., 2008; O'Doherty, 2004; Schultz et al., 1997). Thus, if the weighing of economic benefits and psychological costs occur during dishonest decision-making, we should observe higher functional connectivity between the valuation and executive control systems, such as between the vmPFC and dlPFC. Additionally, when taking the social-related goals of being dishonest into account (self-serving vs. prosocial benefits), neural activity in the valuation system that traces self-serving benefits, such as activity in the Str (Braams et al., 2014; Mobbs et al., 2009), may also functionally connect with activity in the executive control system. This conjecture is in line with the framework recently proposed by Ruff and Fehr (2014) that when making social-related decisions, activity in the valuation system that involves basic reward and value representation is further interconnected with activity in the higher-cognitive areas, including regions in the executive control system.

We aim to dissociate the neural correlates of (1) dishonest decisions concerned about overall benefits of being dishonest (regardless of whether the benefits are self-serving or prosocial) from (2) those that distinguish between self-serving or prosocial benefits. To separate these two types of dishonest decisions, we modified an established dishonest decision-making task, the Coin-Guessing task (Greene & Paxton, 2009), such that the benefits of being dishonest in each trial could either go to the participants themselves (i.e., self-serving) or a charity (i.e., prosocial). Specifically, participants were free to choose to be predominantly honest, predominantly dishonest for both self-serving and prosocial benefits or strategically dishonest for self-serving or for prosocial benefits.

We expected the involvement of the valuation and executive control systems in making dishonest decisions. First, we predicted the pattern of the vmPFC based on the findings that vmPFC traces decision values for both oneself and for others (Nicolle et al., 2012) and across domains of choice options (Chib et al., 2009). Accordingly, participants who had a higher frequency of making dishonest decisions (regardless of whether the benefits of the dishonesty were self-serving or prosocial) should have a stronger activity in the vmPFC when having an opportunity to make dishonest decisions. More specifically, both the frequencies of decisions to be dishonest for self-serving and for prosocial benefits should be positively associated with vmPFC activity. We also expected to observe stronger functional connectivity between the vmPFC and other regions in the executive control systems when making dishonest decisions among participants with a higher frequency of dishonest decisions. This connectivity pattern would reflect the weighing of economic benefits and psychological costs during dishonest decision-making (Barkan et al., 2012; Shalvi et al., 2015). Second, we predicted the pattern of Str activity based on its differential sensitivity toward rewards for the self and rewards for others (Braams et al., 2014). Specifically, participants who decide to be selectively dishonest for self-serving (compared with prosocial) benefits should elicit stronger Str activity when being dishonest for self-serving (compared with prosocial) benefits. Similar to vmPFC activity, we also examined the changes in functional connectivity between the Str and regions in the executive control system as a function of biases toward self-serving benefits during dishonest decision-making. We expected stronger functional connectivity among those who were more dishonest for self-serving (compared with for prosocial) benefits. This pattern of functional connectivity would reflect an interplay between basic reward representation and activity in the higher-cognitive areas when making social-related decisions (Ruff & Fehr, 2014).

2 | METHODS

2.1 | Participants

Thirty-one right-handed, paid volunteers (18 females; age $M = 20.35$ years, $SD = 2.21$) participated in this study. This sample size is consistent with five previous studies that employed a similar (dis)honest decision-making paradigm ($M = 27.40$ subjects, $SD = 6.23$) (Abe & Greene, 2014; Ding et al., 2013; Greene & Paxton, 2009; Hu et al., 2015; Shalvi & De Dreu, 2014). Participants were given Chinese Yuan (CNY) 20 for their participation in addition to a monetary bonus (M = CNY 14.19, $SD = 17.66$) for completing the Coin-Guessing task (see below). Participants were screened for neurological history and had normal or corrected-to-normal vision. The study was approved by the Institutional Review Board at South China Normal University, and participants provided written consent prior to the experiment.

2.2 | Procedure

To examine neural correlates of dishonest decision-making for self-serving and prosocial benefits, we modified the Coin-Guessing task

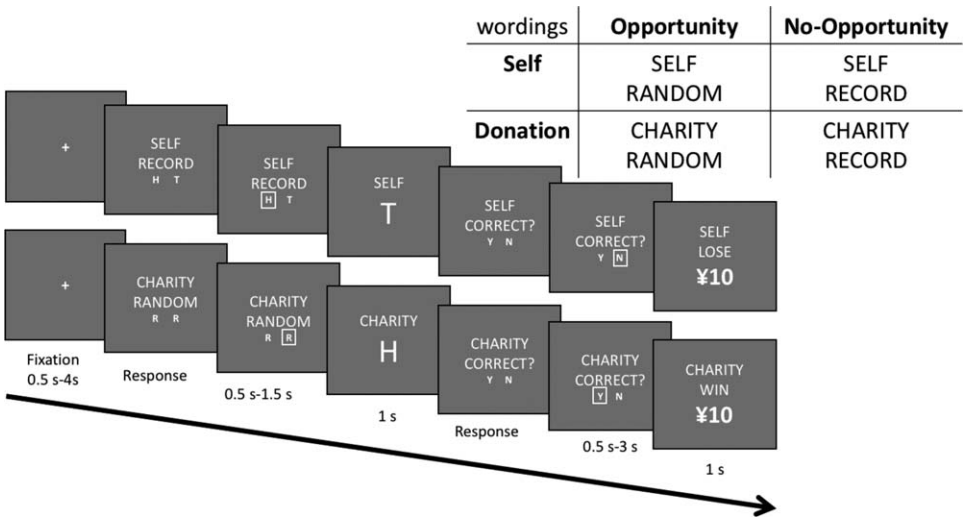

FIGURE 1 Task structure of the Coin-Guessing task (translated from Chinese). In each trial, participants predicted the outcome of a coin flip. Correct predictions corresponded to winning CNY 10, while incorrect predictions corresponded to losing CNY 10. During the No-Opportunity trials (signified by the word, "RECORD"), participants had to enter their prediction as either "Heads" or "Tails" by pressing either the "H" or the "T" key. During the Opportunity trials (signified by the word, "RANDOM"), participants had to randomly press one of the two "R" keys to control for motor activity. Earnings during the Self trials (signified by the word, "SELF") would go to the participants themselves, while earnings during Donation trials (signified by the name of a chosen charitable organization, written down here as "CHARITY") would be donated to a chosen charitable organization

(Greene & Paxton, 2009) using an fMRI event-related design (Figure 1). In each trial, we instructed participants to predict a coin-flip outcome, in which a correct (incorrect) prediction corresponded to gaining (losing) CNY 10 for that trial. There were four unique types of trials based on a 2 Opportunity (Opportunity vs. No-Opportunity) × 2 Self-Serving (Self vs. Donation) design. During the Opportunity trials, participants had an opportunity to engage in dishonest gain if they decided to over-report their performance, while during the No-Opportunity trials they could not do so. Moreover, during the Self trials, monetary incentive (earned through reported accuracy) would go to the participants themselves, whereas during the Donation trials, monetary incentive would go to a charitable foundation of the participants' choice. We pseudo-randomized the trial order using optseq2 (https://surfer.nmr.mgh. harvard.edu/optseq/), and presented each of the four unique trial 70 times, for a total of 280 trials. There were four blocks of 70 trials, separated by breaks of participant-determined length. On average, the task lasted 44.86 min ($SD = 3.82$). Participants completed 10 practice trials outside the scanner.

The trial started with a fixation ITI, jittered between 0.5 and 4.0 s. On No-Opportunity trials, we presented the word "RECORD" following the ITI. In these trials, participants had to record their prediction about the upcoming coin flip by pressing a button labeled "H" or "T" if they predicted "heads" or "tails," respectively, for that trial. Requiring participants to record their prediction during No-Opportunity trials prevented them from being dishonest confidentially (without being exposed) during these trials. On the contrary, in Opportunity trials, we presented the word "RANDOM" following the ITI. When the word "RANDOM" appeared on the screen, participants were instructed to make a prediction in their mind about the upcoming coin flip, but they did not have to record their prediction by pressing an external button. To justify this manipulation, we informed participants that, based on an established protocol used commonly in previous research (Abe & Greene, 2014; Ding et al., 2013; Greene & Paxton, 2009; Hu et al., 2015), people's ability to predict the future (i.e., a coin flip) might be better if they made the predictions privately to themselves. To balance motor activity across Opportunity and No-Opportunity trials, participants were instructed to randomly press one of two buttons, both labeled "R" (random), during Opportunity trials. After pressing, the chosen choice was highlighted for 0.5–1.5 s.

Next, we presented the outcome of the coin flip (a letter "H" or "T" for a head or tail outcome, respectively) for 1 second. The question "CORRECT?" then appeared on the monitor, prompting participants to indicate whether their prediction was accurate or not. For No-Opportunity (i.e., RECORD) trials, we instructed participants to press either a "YES" button (i.e., correct prediction) or a "NO" button (i.e., incorrect prediction) based on their previously recorded responses. For Opportunity (i.e., RANDOM) trials, we instructed participants to press either the "YES" button (i.e., correct prediction) or the "NO" button (i.e., incorrect prediction) based on their prior, nonrecorded predictions. The fact that participants did not declare and record their predictions during Opportunity trials afforded them the opportunity to over-report their performance to increase their possible winnings (i.e., make dishonest decisions). We then highlighted their answers for 0.5–3.0 s, followed by a 1-s screen confirming whether participants won or lost CNY 10 for that trial.

For the manipulation of the Self-Serving conditions, during Self trials we presented the word "SELF" on top of every screen except for the fixation screen. This signified that the money earned in these trials would go to participants themselves. On the contrary, during the Donation trials, we presented the name of a charity of participants' choice instead of the word "SELF." This name was picked out of six famous charities in China by participants themselves before the experiment. Participants also had the choice to select an organization that was not on the list; however, none of participants chose to do so. We told participants that the money earned in the Donation trials would be donated to

the chosen charity. At the end of the experiment, we randomly picked 10 trials, and the earnings in these 10 trials were either paid to participants or donated to charities based on the trial type (i.e., Self or Donation). We did not take away any money from participants if the total randomly chosen earnings in their Self trials were less than zero.

2.3 | Behavioral indices of dishonesty

Because we instructed participants to predict the outcome of a coin flip, the expected reported accuracy for honest participants, regardless of the types of trials, should be around the chance level (i.e., ~50%). Thus, a higher self-reported % accuracy from trials when participants had an opportunity to over report accuracy (i.e., both Opportunity-Self and Opportunity-Donation trials) would suggest a higher likelihood of dishonesty (Abe & Greene, 2014; Greene & Paxton, 2009). Accordingly, we defined *Overall Dishonesty* as total self-reported % accuracy across both Opportunity-Self and Opportunity-Donation trials. That is, when having an opportunity to be dishonest, participants who had *Overall Dishonesty* close to 100% were dishonest for both self-serving and prosocial benefits, while participants who had this index close to 50% were honest for both benefits. We also defined *Opportunity-Self Dishonesty* and *Opportunity-Donation Dishonesty* as self-reported % accuracy in Opportunity-Self and Opportunity-Donation trials, respectively. As such, separately *Opportunity-Self* reflects the degree of dishonesty for self-serving benefits, and *Opportunity-Donation Dishonesty* reflects the degree of dishonesty for prosocial benefits. Note that following original fMRI studies (Abe & Greene, 2014; Greene & Paxton, 2009), we did not use the self-reported % accuracy in the No-Opportunity trials because % accuracy largely depended on the randomization of a coin flip. This randomization was generated by the computer, making the correct response in the No-Opportunity trials varied from trial-to-trial and from participant-to-participant. The self-reported % accuracy in the Opportunity trials, however, largely depended on the decisions to be dishonest, and therefore were controlled by participants themselves.

In addition to *Overall Dishonesty*, *Opportunity-Self Dishonesty*, and *Opportunity-Donation Dishonesty*, we also computed *Self-Serving Dishonesty* as the self-reported % accuracy during Opportunity-Self trials minus the accuracy during Opportunity-Donation trials. *Self-Serving Dishonesty* indicated the extent to which participants strategically chose to over-report accuracy for self-serving benefits (i.e., during Opportunity-Self trials) than for prosocial benefits (i.e., during Opportunity-Donation trials). That is, the higher *Self-Serving Dishonesty* was, the more likely that the participants selectively over-reported accuracy when they were the beneficiary. Altogether, participants who were dishonest regardless of who were the beneficiary would be high in *Overall Dishonesty*, while participants who were selectively dishonest for self-serving benefits would be high in *Self-Serving Dishonesty*. The fMRI analyses below were designed to examine the BOLD contrasts of interest that corresponded to these four behavioral indices.

2.4 | fMRI acquisition

We conducted MRI scanning on a 3-Tesla Tim Trio Magnetic Resonance Imaging scanner (Siemens, Germany) using a standard 12-channel head-coil system. We acquired functional images by using T2*-weighted, gradient echo, echo planar imaging (EPI) sequences (31 oblique axial slices, 3 mm-thickness; TR = 2,000 ms; TE = 30 ms; flip angle = 90°; FOV = 224 mm; voxel size: 3 × 3 × 3 mm). For coregistration and normalization, we also obtained a high-resolution anatomical T1-weighted image at a resolution of 1 × 1 × 1 mm.

2.5 | fMRI analyses

We preprocessed and analyzed fMRI data using SPM8 (www.fil.ion.ucl.ac.uk/spm/). The first three volumes were discarded due to unsteady magnetization. Then, to correct for motion for each participant, the remaining volumes were realigned spatially to the first nondiscarded volume. Images were then resliced and a mean image was created. After a high-resolution image was coregistered onto the mean image, all volumes were normalized to the Montreal Neurological Institute (MNI) space using the MNI International Consortium for Brain Mapping (ICBM) 125 template. The normalized images were then spatially smoothed with a 6-mm Gaussian kernel. Finally, we applied a high-pass temporal filter with a cutoff of 128 s to remove low-frequency drifts.

After preprocessing, we conducted whole-brain statistical analyses for each subject using the general linear model (GLM) (Friston et al., 1994). At the first level, we convolved each trial with a canonical hemodynamic-response function, using the onset of the coin-flip outcome as an event of interest. This outcome phase was the moment when participants evaluated their prediction (e.g., accurate or not). For Opportunity trials, this is the time participants decided whether to make a dishonest decision to increase their earnings by claiming a correct prediction for trials in which they actually made an incorrect prediction. This reasoning is supported by a recent ERP study showing (1) that the Opportunity conditions modulated ERPs locked to this coin-flip outcome, and (2) that these changes in ERPs predicted greater likelihood of engaging in overall dishonest decisions (Hu et al., 2015). We estimated a GLM for every subject with autoregressive order 1. We had four task-related regressors [Opportunity-Self, Opportunity-Donation, No-Opportunity-Self, No-Opportunity-Donation] in our GLM design matrix.[1] To control for motion artifact, we also added six

[1]Note that unlike previous fMRI studies (Abe & Greene, 2014; Greene & Paxton, 2009), we (1) did *not* explicitly model Accurate (i.e., Win) and Inaccurate (i.e., Loss) trials in the first-level analysis and (2) did *not* separately model dishonest and honest participants (i.e., depending on whether their self-reported accuracy in Opportunity trials were higher than chance level) in the second-level analysis. We decided not to do so to avoid differences in signal-to-noise ratio between dishonest and honest participants in Accurate and Inaccurate trials (Hu et al., 2015). That is, dishonest participants would have fewer Inaccurate trials in the Opportunity condition, and thus their statistical models may not be stable. In fact, previous studies (Abe & Greene, 2014; Greene & Paxton, 2009) needed to exclude some dishonest participants because they had too few Inaccurate trials in the Opportunity condition, even though the goal of this Coin-Guessing task was to examine dishonesty. Lumping Accurate and Inaccurate trials avoided this issue. Even though we did not separately model dishonest and honest participants, we were still able to investigate the neural patterns related to making dishonest decisions by using behavioral indices of dishonesty (i.e., self-reported % accuracy) as continuous-variable covariates in our second-level analyses.

head-motion regressors based on SPM's realignment estimation routine to our design matrix.

At the second level, we treated subjects as random effects (Penny & Holmes, 2004), used the default settings in SPM8 for the design specification of our model, and did not specify grand mean scaling. Because participants made decisions whether to be dishonest in private in the Opportunity trials, it is difficult to identify in which trials they honestly reported their performance and in which trials they over-reported it. Accordingly, at the second level, we decided to use behavioral indices as covariates to capture individual variability in dishonest tendencies along with contrasts of interest that corresponded to these behavioral indices. We first normalized these four behavioral indices by ranking them across participants. We then conducted whole-brain regression analyses using these ranked covariates and associated contrasts that were estimated from the first level.

We planned our analyses to focus on two aims. Our first main aim was to examine the relationship between the actual decisions to over-report accuracy overall (regardless of whether the benefits were self-serving or prosocial) and neural activity when participants had an opportunity to over-report accuracy overall. The behavioral index that corresponded to these decisions was Overall Dishonesty, and the first-level estimated contrasts that corresponded to this neural activity were the [Opportunity vs. No-Opportunity] contrasts (collapsed across Self and Donation trials). The relationship revealed from this whole-brain regression analysis would reflect neural activity associated with dishonest decisions made for overall benefits, regardless of whether the self was the beneficiary. To further assess whether the relationship between the actual decisions to over-report accuracy and the neural activity when having an opportunity to over-report accuracy was common across self-serving and for prosocial benefits, we examined the relationships in the Self and in the Donation trials separately. Specifically, we conducted two additional whole-brain regression analyses: (1) using ranked Opportunity-Self Dishonesty as a covariate and Opportunity-Self vs. No-Opportunity-Self as first-level contrasts and (2) using ranked Opportunity-Donation Dishonesty as a covariate and Opportunity-Donation versus No-Opportunity-Donation as first-level contrasts. Activity in areas that demonstrated significant positive relationships across these the two regression analyses should represent the value of making dishonest decisions, regardless of whether benefits of the decisions were self-serving or prosocial.

In addition to conducting the whole-brain regression analyses, we further examined the task-dependent functional connectivity for the first aim. Given that we found the relationship between Overall Dishonesty and the [Opportunity > No-Opportunity] effect at the vmPFC (see Section 3), we investigated the task-dependent functional connectivity that the vmPFC had with other regions using the PsychoPhysiological Interaction (PPI) (Friston et al., 1997; O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). We first created a seed region using a 6-mm diameter sphere at the vmPFC [3 57 −6] as defined by the whole-brain regression analysis without the PPI. At the first-level analysis, we applied a generalized form of context-dependent PsychoPhysiological Interaction (gPPI) (McLaren, Ries, Xu, & Johnson, 2012) as follows. First, we extracted the first mean time series within the vmPFC seed. We then created two separate PPI terms using element-by-element products of the extracted, deconvolved vmPFC time series and each task regressor [Opportunity and No-Opportunity]. These two PPI terms were then reconvolved with the canonical hemodynamic response function and entered as PPI regressors along with the task (psychological), vmPFC time series (physiological), and head-motion (nuisance) regressors. The contrast between the PPI regressors [Opportunity > No-Opportunity] for each participant was then used at the second-level analysis with ranked Overall Dishonesty as a covariate.

Our second main aim was to examine the relationship between the propensity to over-report accuracy for self-serving, relative to prosocial, benefits, and neural activity when deciding whether to be dishonest for self-serving benefits versus for prosocial benefits. The behavioral index that corresponded to these decisions was Self-Serving Dishonesty, and the first-level estimated contrasts that corresponded to this neural activity were [Opportunity-Self vs. Opportunity-Donation] contrasts.[2] The relationship revealed from this whole-brain regression analysis would reflect neural processes underlying dishonest decisions made selectively for self-serving benefits. For completeness, we also ran two additional whole-brain regression analyses with ranked Self-Serving Dishonesty as a covariate using (1) [Self vs. Donation] contrasts (collapsing across Opportunity and No-Opportunity conditions) and (2) [No-Opportunity-Self vs. No-Opportunity-Donation] contrasts. This allowed us to investigate whether the relationship was specific to situations when people made dis/honest decisions (i.e., [Opportunity-Self vs. Opportunity-Donation] contrasts) or when people evaluated the outcome of their prediction (i.e., [No-Opportunity-Self vs. No-Opportunity-Donation] contrasts).

Similar to the first aim, we examined the task-dependent functional connectivity for the second aim. Specifically, given the relationship between Self-Serving Dishonesty and Str activity from [Opportunity-Self > Opportunity-Donation] contrasts (see Section 3), we conducted

[2]As opposed to using the full-interaction contrasts (e.g., [(Opportunity-Self vs. Opportunity-Donation) vs. (No-Opportunity-Self vs. No-Opportunity-Donation)]), we decided to use the [Opportunity-Self vs. Opportunity-Donation] contrasts, which were simple effect contrasts, in our design because of the following reasons. First, the [Opportunity-Self vs. Opportunity-Donation] contrasts closely represented the processes underlying the behavioral index of interest, Self-Serving Dishonesty. Self-Serving Dishonesty was defined as (dis)honest decisions made (reflected by the self-reported % accuracy) during Opportunity-Self trials compared to decisions made during Opportunity-Donation trials. Thus, the (dis)honest decisions underlying Self-Serving Dishonesty only occurred during the Opportunity trials. Second, in the full-interaction contrasts, the No-Opportunity part (i.e., [No-Opportunity-Self vs. No-Opportunity-Donation]) was used to control the Opportunity part (i.e., [Opportunity-Self vs. Opportunity-Donation]). However, there might be fundamental differences in neural–cognitive processes between the two parts, possibly making the No-Opportunity part an unappropriated control. This is especially concerning because the No-Opportunity trials only required the confirmation of the outcome. Contrasting Opportunity-Donation against Opportunity-Self conditions in the simple-effect contrasts should control for unrelated neural processes associated with having the opportunity to be dishonest. Thus, using the simple-effect contrasts should allow us to focus on the modulation role of social-related benefits (i.e., self-serving vs. prosocial) on making dishonest decisions when having an opportunity to do so.

another PPI analysis to examine the task-dependent functional connectivity that the Str had with other regions. Here, as with the first aim, we first created a seed region using a 6-mm diameter sphere at the Str [−15 27 12] as defined by the whole-brain regression analysis without the PPI. Then, we conducted the gPPI analysis (McLaren et al., 2012) using [Opportunity-Self > Opportunity-Donation] contrasts with the Str as a seed and ranked Self-Serving Dishonesty as a covariate.

For both aims, we conducted additional analyses to confirm the specificity of the relationships found. First, we simultaneously included both Overall Dishonesty and Self-Serving Dishonesty behavioral indices as covariates in the main analyses of both aims. Thus, the relationship that was explained by one behavioral index would be statistically controlled for by the other behavioral index. For the first aim, the two behavioral indices were used with the [Opportunity vs. No-Opportunity] contrasts. For the second aim, the two behavioral indices were used with the [Opportunity-Self vs. Opportunity-Donation] contrasts. Second, we also conducted analyses to explore the specificity of the seeds for the functional connectivity analyses. Specifically, we used the seeds of interests, the vmPFC and Str, in the functional connectivity analyses of both aims.

In addition to conducting the whole-brain regression analyses for the two main aims, we also performed whole-brain one-sample *t*-test analyses on the contrasts estimated from the first level. We conducted this set of analyses to overview whole-brain activation patterns that were consistent across participants, regardless of their level of dishonesty. For the first aim, we examined the consistency in the effect of the [Opportunity vs. No-Opportunity] contrasts (collapsing across Self and Donation conditions) across participants. For the second aim, we looked at the consistency in the effect of the [Opportunity-Self vs. Opportunity-Donation] contrasts. For completeness, we also reported the [Self vs. Donation], (collapsing across Opportunity and No-Opportunity conditions), [No-Opportunity-Self vs. No-Opportunity-Donation] and full-interaction contrasts.

To control for multiple statistical testing for the whole-brain second-level analyses, we employed Monte-Carlo simulations using the "3dClustSim-ACF" command [10,000 iterations; cluster-forming threshold (CFT) = .005; bi-sided thresholding; first-nearest neighbor clustering] in AFNI version 16.3.05 (http://afni.nimh.nih.gov) (Cox, 1996). To account for the noise smoothness structure, we used a mixed model for estimating a non-Gaussian spatial autocorrelation function ("3dFWHMx-ACF" command) (Cox, Chen, Glen, Reynolds, & Taylor, 2017). A recent empirical study (Cox et al., 2017) showed that, for event-related design studies, using this recently developed estimation method with a CFT of .005 can control for a nominal familywise error rate (*p*FWE) $\sim=$ .05 for clusterwise inference, which was problematic in older methods (Eklund, Nichols, & Knutsson, 2016). Thus, we only reported clusters higher than 99 voxels that survived a clusterwise correction *p*FWE < .05 across the whole brain using this method.

3 | RESULTS

3.1 | Behavioral results

Figure 2 and Table 1 show behavioral results. We first investigated the discrepancy between self-reported % accuracy and the actual prediction performance during No-Opportunity trials. Overall, the occurrence of this discrepancy was quite rare ($M = 4.79\%$, $Mdn = 2.14\%$, $SD = 8.4$), suggesting that participants largely reported their true performance in the No-Opportunity trials.[3] We then examined individual differences in Overall Dishonesty and Self-serving Dishonesty indices. In general, our participants varied in both Overall Dishonesty (total self-reported % accuracy across both Opportunity-Self and Opportunity-Donation trials, $M = 74.14\%$, $SD = 16.86$, Figure 2a) and Self-serving Dishonesty indices (self-reported % accuracy during Opportunity-Self trials minus during Opportunity-Donation trials, $M = 5.53\%$, $SD = 9.3$, Figure 2b). To statistically examine if self-reported accuracy during Opportunity trials was higher than chance level (50%), we conducted one-tailed binomial tests using a $p < .001$ threshold on each participant's overall dishonesty (Abe & Greene, 2014; Greene & Paxton, 2009). We found that 16 (out of 31) participants had improbably high levels of accuracy at the individual level (see Figure 2a). This confirms the heterogeneity of our participants in terms of dishonesty. The correlation between Overall Dishonesty and Self-serving Dishonesty was not significant ($r(29) = -.15$, $p = .42$, Figure 2c), suggesting that these two indices tap onto two distinct and dissociable behavioral tendencies.

We also conducted a 2 Opportunity (Opportunity vs. No-Opportunity) × 2 Self-Serving (Self vs. Donation) repeated-measure ANOVA on self-reported % accuracy (see Figure 2d and Table 1 for descriptive statistics), using generalized-η^2 (Bakeman, 2005) for effect sizes. While the interaction was not statistically significant ($F(1,30) = 3.68$, $p = .065$, generalized-$\eta^2 = .008$), there was a main effect of Opportunity ($F(1,30) = 39.21$, $p < .001$, generalized-$\eta^2 = .27$). This indicated that there was a

[3]In our fMRI analyses, we did not exclude No-Opportunity trials in which participants misreported their performance for the following reasons. First, the first-level contrasts in our fMRI analyses that involved No-Opportunity trials (e.g., Opportunity > No-Opportunity, Opportunity-Self > No-Opportunity-Self, etc.) did not concern if participants were actually (dis)honest. Rather, they concerned the differences in neural activity when they had (vs. did not have) *opportunities* to be dishonest without being caught in the corresponding trials. What determined whether participants were actually dishonest were the behavioral indices that were employed as covariates in the second-level analysis (*Overall Dishonesty*, *Opportunity-Self Dishonesty*, *Opportunity-Donation Dishonesty*, and *Self-Serving Dishonesty*). Because these behavioral indices were calculated using participants' reports in the Opportunity trials (Abe & Greene, 2014; Greene & Paxton, 2009), by definitions, misreports in the No-Opportunity trials did not influence these behavioral indices and, therefore, did not influence our analyses of neural correlates of actual dishonesty. As mentioned above, to calculate behavioral indices, we did *not* use the accuracy reports in the No-Opportunity trials because they largely depended on the computer's coin randomization, which varied across participants. Additionally, while misreports in the No-Opportunity trials in some cases may indicate dishonesty, it is difficult to rule out honest mistakes (e.g., participants forgot their first report, or pressed the wrong button). This is especially important given the rare occurrence of misreports in the No-Opportunity trials. Another benefit of not excluding these trials was that the number of Opportunity trials and No-Opportunity trials were equal to each other, making the estimation of betas and contrasts between the two types of trials similar in their signal-to-noise ratio at the first level. This makes the use of behavioral indices at the second level more appropriate. The similar approach of not excluding misreports in the No-Opportunity trials is commonly employed in past research (Abe & Greene, 2014; Cui et al., 2018; Greene & Paxton, 2009; Hu et al., 2015).

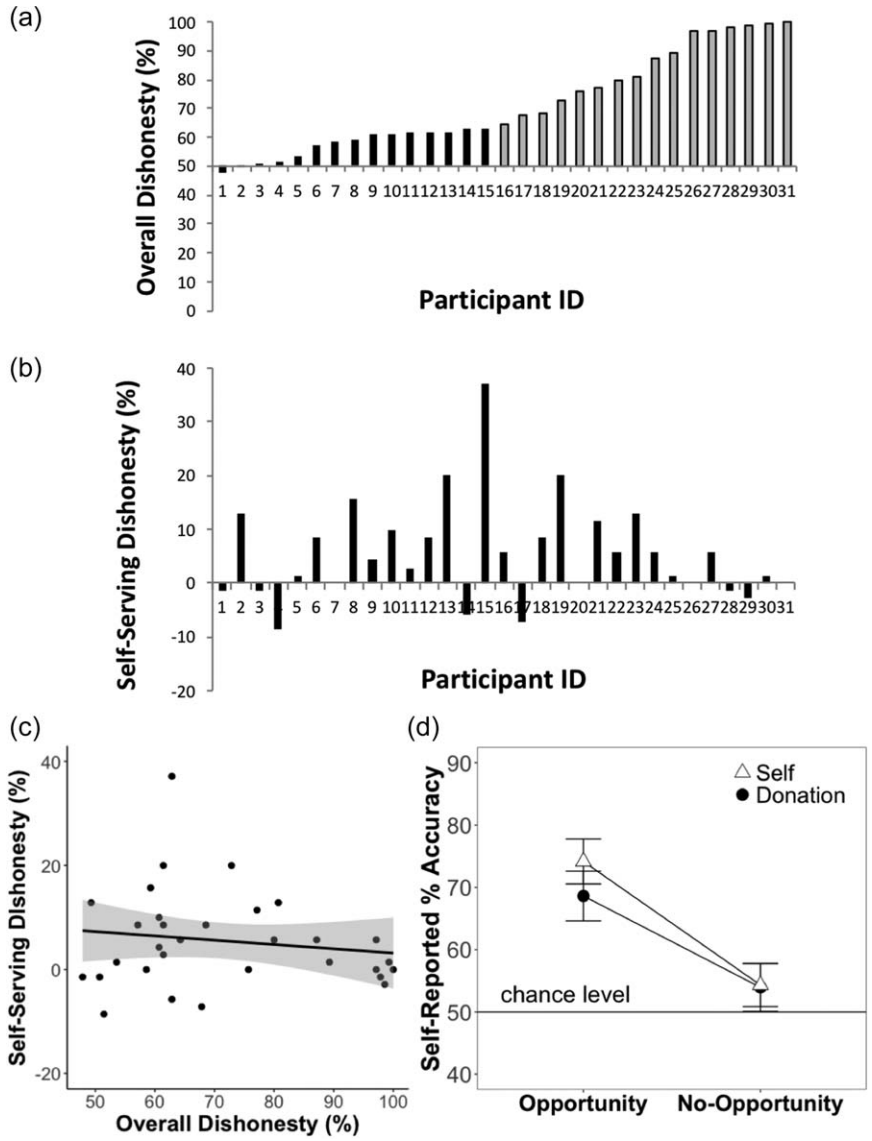

FIGURE 2 Behavioral results of the Coin-Guessing task. Figure 2a shows individual differences in Overall Dishonesty, defined by total self-reported % accuracy across both Opportunity-Self and Opportunity-Donation trials. If participants honestly reported their prediction, their overall dishonesty should be around 50%. Sixteen participants (ID 16–31; represented by gray bars) reported improbably high levels of accuracy at the individual level, as revealed by a one-tailed binomial test, $p < .001$ (Greene & Paxton, 2009). Figure 2b shows individual differences in Self-Serving Dishonesty, defined by self-reported % accuracy during Opportunity-Self trials minus % accuracy during Opportunity-Donation trials. Figure 2c shows a scatter plot between Overall Dishonesty and Self-serving Dishonesty. This plot indicates a nonsignificant relationship between the two indices ($r(29) = -.15$, $p = .42$). The gray shaded area in the scatterplot represents 95% CIs around the linear regression line. Figure 2d shows self-reported % accuracy as a function of Opportunity and Self-Serving conditions. Error bars represent within-subject 95% CIs (Morey, 2008)

TABLE 1 Means, *SDs*, and CIs of self-reported % accuracy as a function of Opportunity and Self-Serving conditions

	Opportunity	No-Opportunity
Self	74.1 (*SD* = 9.81, 95% CI = 3.6)	54.3 (*SD* = 9.39, 95% CI = 3.44)
Donation	68.6 (*SD* = 10.86, 95% CI = 3.98)	54.0 (*SD* = 10.51, 95% CI = 3.85)

CIs represent within-subject 95% CIs calculated via the "summarySEwithin" command in the Rmisc library. This calculation is based on a previously established method (Morey, 2008).

higher self-reported % accuracy during Opportunity trials than during No-Opportunity trials. Additionally, a main effect of Self-Serving was also statistically significant ($F(1,30) = 14.50$, $p < .001$, generalized-$\eta^2 = .01$), indicating that there was a higher self-reported % accuracy during Self trials than during Donation trials.

3.2 | fMRI results

Table 2 and Figure 3 list the results from the analyses of the first main aim. To examine the relationship between the actual decisions to over-

TABLE 2 Neural activity of the Opportunity vs. No-Opportunity, Opportunity-Self vs. No-Opportunity effects as a function of ranked Overall Dishonesty, Opportunity-Self Dishonesty, and Opportunity-Donation Dishonesty

Contrast	Region	R/L/M	BA	MNI coordinates			t-score	Voxels
				x	y	z		
Opportunity vs. No-Opportunity contrasts collapsing across Self and Donation trials with ranked Overall Dishonesty as a covariate								
Opportunity > No-Opportunity	Superior temporal gyrus	L	22	−66	−18	6	4.48	165
	Primary motor cortex	R	4	36	−30	69	4.42	1,161
	Premotor cortex	L	6	−60	−6	36	4.27	125
	Somatosensory cortices	R	2	60	−21	15	4.08	114
	Ventromedial prefrontal cortex	R	11	3	57	−6	3.63	215
No-Opportunity > Opportunity	Occipital cortex and cerebellum	R	18	27	−78	−21	4.24	488
	Anterior cingulate cortex and dorsomedial prefrontal cortex	R	10	9	27	27	3.47	104
Opportunity vs. No-Opportunity contrasts in Self trials with ranked Opportunity-Self Dishonesty as a covariate								
Opportunity-Self > No-Opportunity-Self	Intraparietal sulcus	R	19	39	−81	36	5.08	147
	Primary motor cortex	R	4	36	−24	72	5.03	164
	Ventromedial prefrontal cortex	L	11	−9	39	−15	4.58	107
	Primary motor cortex	L	4	−12	−27	57	4.21	308
No-Opportunity-Self > Opportunity-Self	Occipital cortex	R	19	6	−87	−9	4.16	216
Opportunity vs. No-Opportunity contrasts in Donation trials with ranked Opportunity-Donation Dishonesty as a covariate								
Opportunity-Donation > No-Opportunity-Donation	Superior parietal lobe	R	7	27	−51	54	5.63	2,536
	Ventromedial prefrontal cortex	M	11	0	45	−9	4.98	342
	Premotor cortex	L	6	−60	−6	36	4.87	496
	Premotor cortex	L	6	−24	−24	63	4.67	258
	Occipital cortex	L	19	−18	−72	−3	4.64	183
	Occipital cortex	L	19	−42	−78	15	4.11	179
	Corpus callosum	L		−15	−48	12	3.92	162
No-Opportunity-Donation > Opportunity-Donation	No supra-threshold clusters were found							

Overall Dishonesty is defined by total self-reported % accuracy across both Opportunity-Self and Opportunity-Donation trials. Opportunity-Self Dishonesty and Opportunity-Donation Dishonesty are defined by self-reported % accuracy in Opportunity-Self and Opportunity-Donation trials, respectively. The results were based on whole-brain regression analyses [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. Significant positive t-scores reflect positive associations. BA, Brodmann areas.

report accuracy overall and neural activity when having an opportunity to over-report accuracy overall, we used Overall Dishonesty as a covariate to modulate the effect based on the [Opportunity vs. No-Opportunity] contrasts (collapsing across Self and Donation conditions). We found that people who had a stronger overall tendency to over-report accuracy (reflected by higher Overall Dishonesty) had stronger activity in the vmPFC among other areas when given an opportunity to over-report accuracy overall [Opportunity > No-Opportunity] (see Figure 3a). Additionally, when simultaneously including both Overall Dishonesty and Self-Serving Dishonesty as covariates in this model, the effect of Overall Dishonesty at the vmPFC still remained statistically significant, while the effect of Self-Serving Dishonesty did not pass the threshold (see Supporting Information Table S1). This suggests that the effect of Overall Dishonesty on the [Opportunity > No-Opportunity] contrast at the vmPFC could not be explained by Self-Serving Dishonesty. Moreover, when separately analyzing the data in the Self [Opportunity-Self > No-Opportunity-Self with Opportunity-Self Dishonesty as a covariate] and Donation [Opportunity-Donation > No-Opportunity-Donation with Opportunity-Donation Dishonesty as a covariate] trials, we found that the significant relationships overlapped at the vmPFC (see Figure 3b). In other words, regardless of whether the benefits of being dishonest were self-serving or pro-social, people who had a stronger tendency to over-report accuracy (reflected by the three indices) had a stronger activity in the vmPFC when given an opportunity to over-report accuracy. Table 3 and Figure 3a show the results from the gPPI analysis using the [Opportunity > No-Opportunity] contrast with the vmPFC as a seed and ranked Overall Dishonesty as a covariate. We found that participants with higher Overall Dishonesty had stronger functional-connectivity strength between the vmPFC and bilateral dlPFC during Opportunity compared with No-Opportunity trials. We found no supra-threshold clusters when we used the Str instead of the vmPFC as a seed region in this model.

Table 4 and Figure 4 list the results from the analyses of the second main aim. To examine the relationship between the tendency to over-report accuracy for self-serving (relative to prosocial) benefits and associated neural activity, we used Self-Serving Dishonesty as a covariate to modulate the effect based on the [Opportunity-Self vs. Opportunity-Donation] contrasts. We found that people with higher Self-Serving Dishonesty had stronger activity in both dorsal and ventral parts of the striatum (Str) when making decisions to be (dis)honest for

self-serving [Opportunity-Self] benefits relative to for prosocial [Opportunity-Donation] benefits. Additionally, when simultaneously including both Overall Dishonesty and Self-Serving Dishonesty as covariates in this model, the effect of Self-Serving Dishonesty at the Str still remained statistically significant, while the effect of Overall Dishonesty did not pass the threshold (see Supporting Information Table S2). This suggests that the effect of Self-Serving Dishonesty on the [Opportunity-Self > Opportunity-Donation] contrast at the Str could not be explained by Overall Dishonesty. Moreover, while Self-Serving Dishonesty modulated both the [Opportunity-Self > Opportunity-Donation]

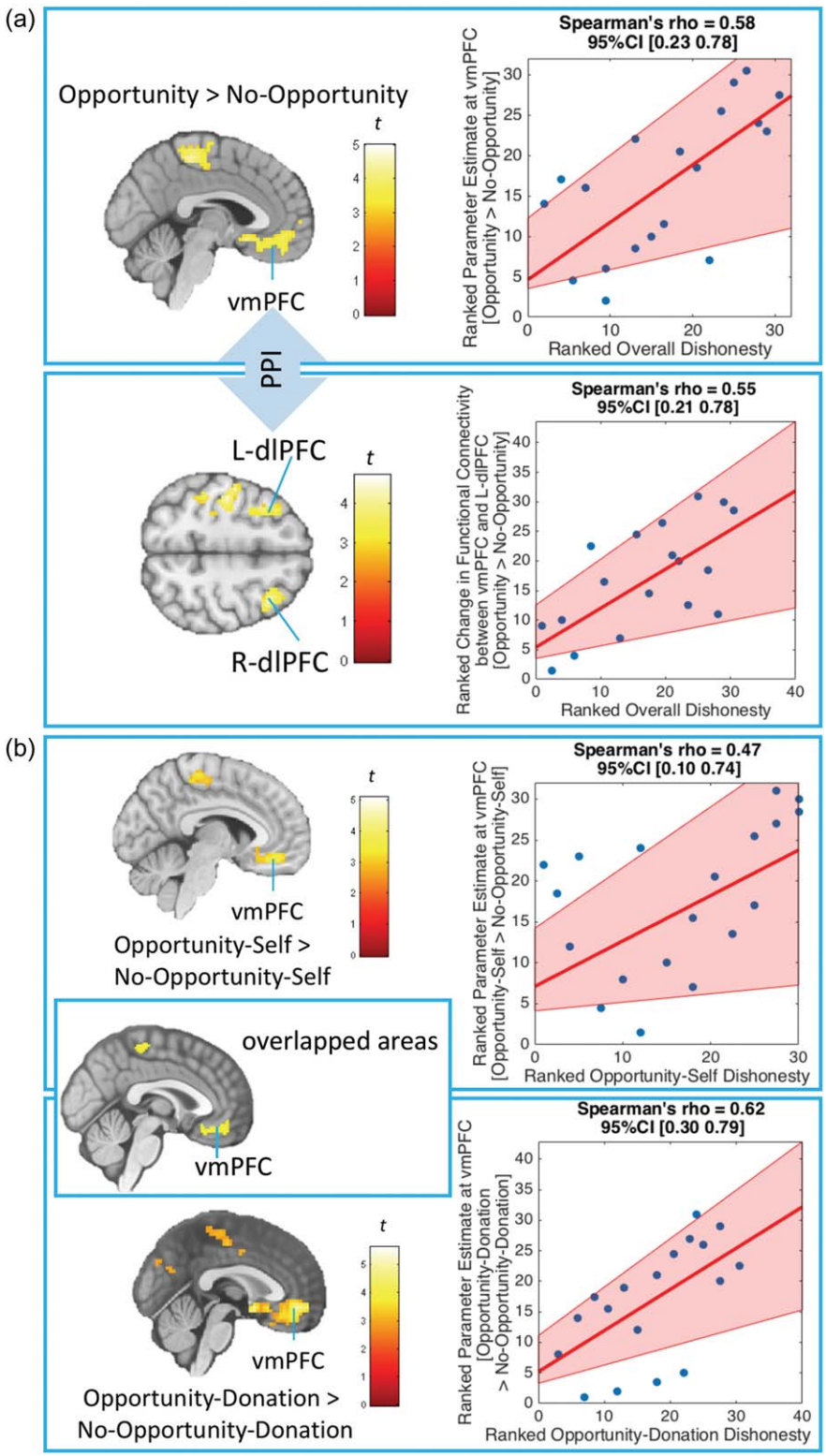

FIGURE 3.

and [Self > Donation] effects at the Str, it did not modulate the [No-Opportunity-Self > No-Opportunity-Donation] effect. Thus, this relationship was specific to situations when people had a chance to make dis/honest decisions [Opportunity-Self > Opportunity-Donation], and did not apply to situations where there was no chance [No-Opportunity-Self- > No-Opportunity-Donation]. In other words, when having an opportunity to over-report accuracy in Opportunity trials, participants who had stronger activity in the Str when deciding whether to over-report accuracy for themselves (compared with for donation) [Opportunity-Self- > Opportunity-Donation] were more likely to selectively over-report accuracy for themselves (compared with for donation). Table 5 and Figure 4 show the gPPI results using Opportunity-Self > Opportunity-Donation contrasts with the Str as a seed and ranked Self-Serving Dishonesty as a covariate. We found that people with higher Self-Serving Dishonesty had stronger functional-connectivity strength between the Str and middle-medial prefrontal cortex (mmPFC) during Opportunity-Self compared with Opportunity-Donation trials. We found no supra-threshold clusters when we used the vmPFC instead of the Str as a seed region in this model.

Table 6 and Figure 5 list a summary of neural activity during the coin-guessing task across participants regardless of their dishonesty level. For the first aim, we examined the effect of the [Opportunity vs. No-Opportunity] contrasts (collapsing across Self and Donation conditions) across participants. We found that having an opportunity to over-report accuracy [Opportunity > No-Opportunity] was associated with enhanced activity in executive-control areas (Niendam et al., 2012), such as the dorsolateral prefrontal cortex (dlPFC), dorso/middle-medial prefrontal cortex (dmPFC/mmPFC), and inferior parietal lobe (IPL). For the second aim, we examined effect of the [Opportunity-Self vs. Opportunity-Donation] contrasts across participants. We found enhanced activity in the dmPFC and IPL when participants processed the outcome for oneself (compared with for donation) in Opportunity [Opportunity-Self > Opportunity-Donation] trials. On the other hand, processing the outcome for donation (compared with for oneself) in Opportunity trials [Opportunity-Donation > Opportunity-Self] was associated with stronger activity in the other regions in the dmPFC, anterior cingulate cortex (ACC), and premotor cortices, among other areas.

4 | DISCUSSION

Our main goal was to separately investigate the neural-cognitive processes underlying (1) dishonest decisions made for overall benefits, regardless of whether the self was the beneficiary, and (2) dishonest decisions made selectively for self-serving benefits. Behaviorally, we found support for dissociable patterns in individual variations for these two types of dishonest decisions in our task. Some participants were dishonest for overall benefits, irrespective of whether their decisions had self-serving or prosocial benefits (i.e., having high Overall Dishonesty). Others were dishonest more selectively for self-serving benefits (i.e., having high Self-Serving Dishonesty). More importantly, as predicted, these two patterns of dishonesty were separately associated with activity in the two key regions in the valuation system, the vmPFC and Str, and their functional connectivity with the executive-control system.

On one hand, when provided an opportunity to make dishonest decisions, vmPFC activity was stronger among people who had a stronger tendency to be dishonest overall. We found this pattern both when the benefits of dishonesty were self-serving and prosocial. This is consistent with the idea that vmPFC activity represents decision-value signals and is positively correlated with subjective values of choices (Bartra et al., 2013; Hare et al., 2008). Given the monetary benefits from being dishonest in our task, the vmPFC may trace the values of these benefits. Our findings are also consistent with the common currency account, in which the vmPFC encodes values across domains of choice options (Chib et al., 2009). That is, the vmPFC represents decision-value signals not only for food choices, nonfood merchandises, and gamble choices found previously (Chib et al., 2009; Tom et al., 2007), but also for moral choices in our study. Additionally, because the vmPFC traced a propensity to be dishonest for prosocial benefits, our results are also in line with the findings that the vmPFC represents decision values that are of high value for others (Hare et al., 2010; Nicolle et al., 2012). Finally, the stronger functional connectivity between the vmPFC and dlPFC among participants with higher Overall Dishonesty also provides a more complete picture of how the valuation (vmPFC) and executive-control (dlPFC) systems (Lisofsky et al., 2014) interacted with each other when deciding whether to behave dishonestly.

FIGURE 3 Neural activity of the Opportunity vs. No-Opportunity effects as a function of ranked Overall Dishonesty, Opportunity-Self Dishonesty and Opportunity-Donation Dishonesty. Figure 3a shows neural activity of the Opportunity vs. No-Opportunity effects across both Self and Donation trials as a function of ranked Overall Dishonesty. Overall Dishonesty is defined by total self-reported % accuracy across both Opportunity-Self and Opportunity-Donation trials. The top section shows a positive relationship between Overall Dishonesty and neural activity in the ventromedial prefrontal cortex (vmPFC) when participants had an opportunity to over-report accuracy [Opportunity > No-Opportunity]. The bottom section shows a positive relationship between Overall Dishonesty and the functional connectivity between the vmPFC and bilateral dorsolateral prefrontal cortex (dlPFC). The functional-connectivity analysis was conducted using a PPI between Opportunity and No-Opportunity conditions collapsing across Self and Donation trials with the vmPFC [3 57 −6] as a seed. Figure 3b shows neural activity of the Opportunity vs. No-Opportunity effects separately for Self and Donation trials as a function of ranked Opportunity-Self Dishonesty and Opportunity-Donation Dishonesty, respectively. Opportunity-Self Dishonesty and Opportunity-Donation Dishonesty are defined by self-reported % accuracy in Opportunity-Self and Opportunity-Donation trials, respectively. This figure shows positive relationships between self-reported % accuracy and neural activity in the ventromedial prefrontal cortex (vmPFC) when participants had an opportunity to over-report accuracy in both Self [Opportunity-Self Dishonesty and Opportunity-Self > No-Opportunity-Self contrasts] and Donation [Opportunity-Donation Dishonesty and Opportunity-Donation > No-Opportunity-Donation contrasts] trials. The images were based on whole-brain regression analyses [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. The pink shaded area in the rank-transformed scatterplot (higher value = higher rank) represents bootstrapped 95% CIs around the linear regression line (Pernet, Wilcox, and Rousselet, 2013) [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Functional connectivity of the Opportunity > No-Opportunity contrasts as a function of ranked Overall Dishonesty

Contrast	Region	R/L/M	BA	MNI coordinates			t-score	Voxels
				x	y	z		
Opportunity > No-Opportunity	primary motor cortex	L	4	−42	−16	57	4.68	135
	occipital cortex	R	18	6	−66	−3	4.41	106
	dorsolateral prefrontal cortex	L	8/9	−33	18	57	3.91	113
	dorsolateral prefrontal cortex	R	8/9	33	21	57	3.85	109

The functional-connectivity analysis was conducted using a PPI between Opportunity and No-Opportunity conditions collapsing across Self and Donation trials with the ventromedial prefrontal cortex [3 57 −6] as a seed. Overall Dishonesty is defined by total self-reported % accuracy across both Opportunity-Self and Opportunity-Donation trials. The results were based on a whole-brain regression analyses with ranked Overall Dishonesty as a covariate [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. Significant positive t-scores reflect positive associations. BA, Brodmann areas.

It is important to link our findings with different theories of dishonesty. According to traditional economists, people decide whether to behave dishonestly using a trade-off between potential economic benefits and costs (Becker, 2000). However, in our task, there were no economic costs of being dishonest. Even so, most participants did not maximize their potential earnings by over-reporting their performance in every trial. This pattern is commonly found in various psychological experiments (Mazar & Ariely, 2006; Shalvi et al., 2015). This suggests that, in making dishonest decisions, people weigh economic benefits not only against economic costs, but also against psychological costs (e.g., being dishonest may pose a threat to their moral self-image) (Barkan et al., 2012; Shalvi et al., 2015). Previous cognitive-neuroscience research has proposed that psychological costs are represented in the brain by executive-control signals (Abe & Greene, 2014; Ding et al., 2013; Greene & Paxton, 2009; Hu et al., 2015; Lisofsky et al., 2014). Consistent with this idea, we found enhanced activity in several key areas in the executive-control network, such as the dlPFC, dmPFC, mmPFC, and IPL, when there was an opportunity to over-report accuracy. More importantly, we further highlighted the role of reward and valuation processing that has been largely neglected in the literature (Abe & Greene, 2014; Hu et al., 2015; Mazar & Ariely, 2006). Specifically, we argue that the relationship between Overall Dishonesty and the functional connectivity between the vmPFC and dlPFC may reflect the weighing between economic benefits and psychological costs. As discussed earlier, the values of the both self-serving and prosocial benefits from being dishonest was traced by the vmPFC activity, similar to that of other decision-making domains (Chib et al., 2009). Therefore, this may address an important question of whether "decisions about honesty are like every other decisions that individuals face" (Mazar & Ariely, 2006).

On the other hand, given an opportunity to be dishonest, participants with higher Self-Serving Dishonesty exhibited a stronger activity

TABLE 4 Neural activity of the Opportunity-Self vs. Opportunity-Donation, Self vs. Donation, No-Opportunity-Self vs. No-Opportunity-Donation effects as a function of ranked Self-Serving Dishonesty

Contrast	Region	R/L/M	BA	MNI coordinates			t-score	Voxels
				x	y	z		
Self vs. Donation contrasts in Opportunity trialswith ranked Self-Serving Dishonesty as a covariate								
Opportunity-Self > Opportunity-Donation	Dorsal striatum[2]	R		9	12	12	4.77	319
	Ventral striatum[2]	R		18	15	−12	3.99	
	Dorsal striatum[3]	L		−15	27	12	3.84	105
	Ventral striatum[3]	L		−21	24	−3	3.66	
Opportunity-Donation > Opportunity-Self	No supra-threshold clusters were found							
Self vs. Donation contrasts collapsing across Opportunity and No-Opportunity trials with ranked Self-Serving Dishonesty as a covariate								
Self > Donation	Dorsal striatum[1]	R		9	9	15	4.26	162
	Ventral striatum[1]	R		18	12	−15	3.16	
Donation > Self	No supra-threshold clusters were found							
Self vs. Donation contrasts in No-Opportunity trials with ranked Self-Serving Dishonesty as a covariate								
No-Opportunity-Self > No-Opportunity-Donation	No supra-threshold clusters were found							
No-Opportunity-Donation > No-Opportunity-Self	No supra-threshold clusters were found							

Self-Serving Dishonesty is defined by self-reported % accuracy during Opportunity-Self trials minus % accuracy during Opportunity-Donation trials. The results were based on whole-brain regression analyses with ranked Self-Serving Dishonesty as a covariate [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. BA, Brodmann areas. Superscripted numbers denote that the regions are from the same cluster.

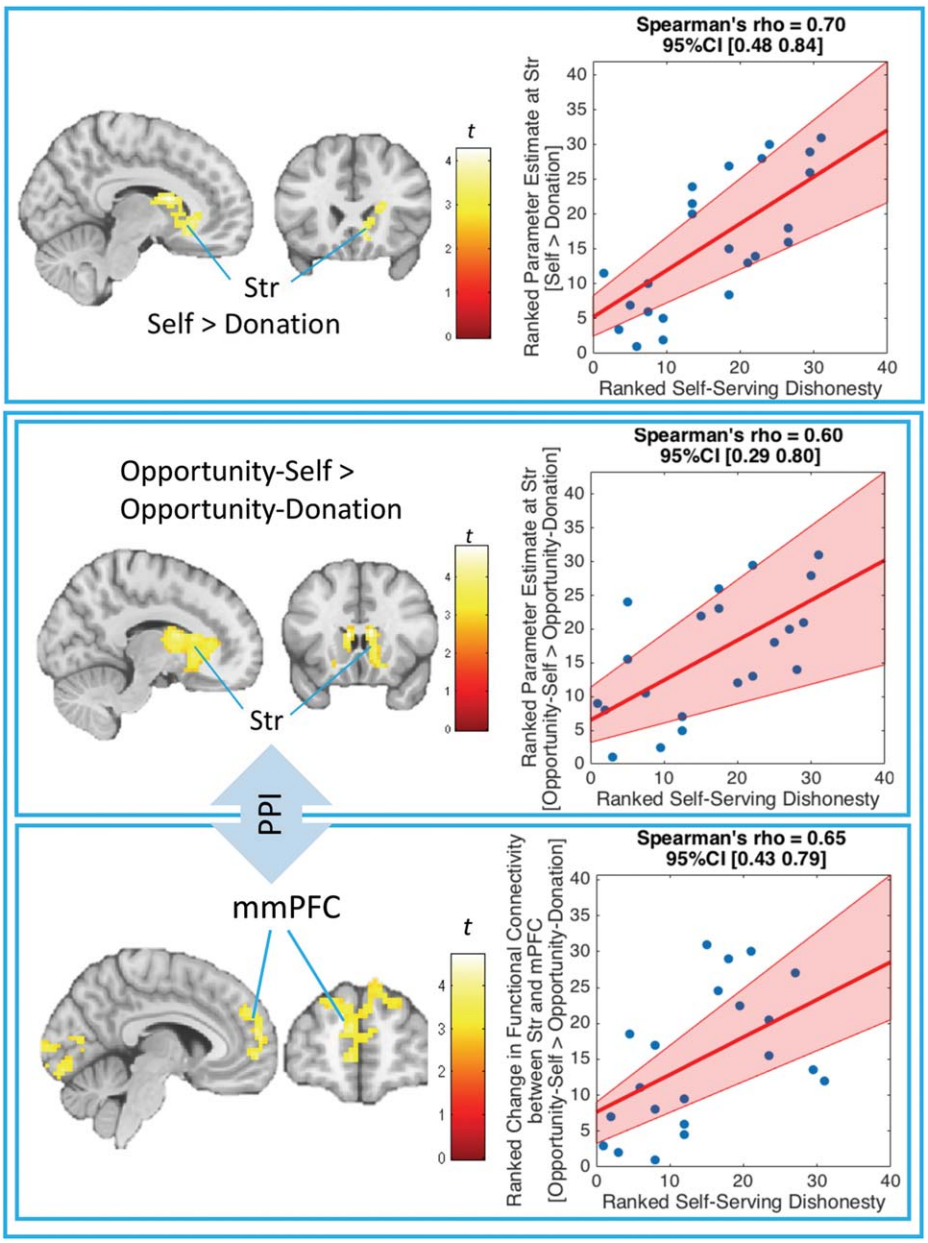

FIGURE 4 Neural activity of the Self vs. Donation effects as a function of ranked Self-Serving Dishonesty. Self-Serving Dishonesty is defined by self-reported % accuracy during Opportunity-Self trials minus % accuracy during Opportunity-Donation trials. The top section shows a positive relationship between Self-Serving Dishonesty and neural activity in the Striatum (Str) when participants evaluated the coin-flip outcome for themselves compared with for donation [Self > Donation]. The bottom section shows that this relationship also applied to situations when people had a chance to make dis/honest decisions [Opportunity-Self > Opportunity-Donation]. Note that there was no significant relationship between Self-Serving Dishonesty and neural activity in the Str when there was no chance [No-Opportunity-Self- > No-Opportunity-Donation]. The bottom section also displays a positive relationship between Self-Serving Dishonesty and the functional connectivity between the Str and middle-medial prefrontal cortex (mmPFC). The functional-connectivity analysis was conducted using a PPI between Opportunity-Self and Opportunity-Donation conditions with the dorsal part of the Str [−15 27 12] as a seed. The images were based on whole-brain regression analyses [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. The pink shaded area in the rank-transformed scatterplot (higher value = higher rank) represents bootstrapped 95% CIs around the linear regression line (Pernet et al., 2013) [Color figure can be viewed at wileyonlinelibrary.com]

in the Str when making (dis)honest decisions for self-serving (compared with prosocial) benefits. In other words, Str activity for these individuals was differentially enhanced toward self-serving benefits. It is important to note that this enhanced activity in the Str as a function of Self-Serving Dishonesty could not simply be explained by higher earnings to the self. In our task, people who obtained higher earnings to themselves can either be higher or lower in Self-Serving Dishonesty depending on whether they decided to over-report accuracy (1) selectively for themselves (i.e., higher in Self-Serving Dishonesty) or (2) indiscriminatingly for both themselves and donations (i.e., lower in Self-Serving

TABLE 5 Functional connectivity of the Opportunity-Self > Opportunity-Donation contrasts as a function of ranked Self-Serving Dishonesty

Contrast	Region	R/L/M	BA	MNI coordinates			t-score	Voxels
				x	y	z		
Opportunity-self > Opportunity-Donation	Occipital lobe	R	19	24	−93	24	4.20	614
	Middle-medial prefrontal cortex	L	9	−6	57	18	3.86	312
	Sensorimotor cortex	L	6	−48	−12	51	3.64	122

The functional-connectivity analysis was conducted using a PPI between Opportunity-Self and Opportunity-Donation conditions with the dorsal striatum [−15 27 12] as a seed. Self-Serving Dishonesty is defined by self-reported % accuracy during Opportunity-Self trials minus % accuracy during Opportunity-Donation trials. The results were based on a whole-brain regression analyses with ranked Self-Serving Dishonesty as a covariate [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. Significant positive t-scores reflect positive associations. BA, Brodmann areas.

Dishonesty). Researchers have observed a similar pattern of Str activity when examining vicarious reward-processing (Braams et al., 2014; Mobbs et al., 2009; Ruff & Fehr, 2014). For instance, the reaction of the Str when observing others gain money depends on whether participants like the observed other (Braams et al., 2014). The enhancement of Str activity is stronger for liked others than for disliked others. Thus, in our study, the relatively stronger Str activity toward self-serving benefits among participants with higher Self-Serving Dishonesty may reflect that, for these participants, their self-serving benefits were much more rewarding than their prosocial benefits.

We also found a stronger functional connectivity between the Str and mmPFC among people with higher Self-Serving Dishonesty. This was consistent with a framework that, when making social-related decisions, basic reward-processing (Str) is interconnected with "higher cognitive-processing" in the prefrontal cortex (mmPFC) (Ruff & Fehr, 2014). In a recent study that investigated altruistic behaviors (Hu et al., 2017), activity in the mmPFC is shown to trace the potential risk to the self (as opposed to the need of others) when deciding whether to help another person. Based on this role of the mmPFC in evaluating self-risk, we speculated the processes that may underlie a strategy to be dishonest in the Self trials, but honest in the Donation trials, among people with higher Self-Serving Dishonesty. In particular, these high self-serving-biased individuals may implement this strategy not only because, for them, self-serving benefits were much more rewarding than their prosocial benefits (reflected by differentially enhanced Str activity toward self-serving benefits), but also because they may try to avoid the risk of being judged as dishonest. By implementing the strategy, they were able to earn a higher amount of reward, while keeping the total frequency of dishonest decisions throughout the task low, compared with those who decided to be dishonest in both the Self and Donation trials. Future studies are needed to test this conjecture regarding the role of the mmPFC more systematically, perhaps by directly manipulating self-risk in the dishonest decision-making context.

It is important that we discuss our findings in light of two very recent cognitive neuroscience studies that also investigated the modulatory roles of social-related goals on dishonest dicision-making (Cui et al., 2018; Yin et al., 2017). One EEG study used the Coin-Guessing task (Greene & Paxton, 2009) and found a higher propensity to be dishonest for self-serving, compared with for prosocial, benefits (Cui et al., 2018), while the other fMRI study used the Sender-Receiver Game (Gneezy, 2005) and found the opposite behavioral pattern (Yin et al., 2017). Because we used the same task as Cui et al.' (2018) study, and found replicating behavioral pattern with theirs, we attribute the discrepancy between ours and Yin et al.' (2017) to the difference in study design. As argued by Cui et al. (2018), deciding to be dishonest in the Sender-Receiver Game involves a stronger concern about self-image and reputation. Specifically, participants in this game need to record their dishonesty, knowing that the experimenters would be aware of them being dishonest with another person at that moment. Thus, in this context, being dishonest for prosocial benefits may alleviate the threat to participants' own self-image and reputation, making them more likely to lie for prosocial, than for self-serving, benefits. The Coin-Guessing task, on the other hand, eases this concern about self-image and reputation by allowing participants to make their dishonest decisions in private, rendering the dishonesty unnoticeable by the experiments at the time of decisions. Thus, the Coin-Guessing task is more akin to real-life situations when people have opportunities to be dishonest for some kind of benefits, knowing that others are not aware of their dishonesty. This design may allow people to be dishonest if they are motivated to do so and may "show the real power of self-interest drive" (Cui et al., 2018). The difference in a paradigm used may also explain the discrepancy in neural activity found Yin et al.' (2017) and ours. When having an opportunity to be dishonest for self-serving benefits (compared with for prosocial benefits), participants in their study with higher self-serving dishonesty showed a stronger activity in the anterior insula (AI) whereas those in our study showed a stronger activity in the Str. Given the roles of the AI in interoception and self-awareness (Critchley, Wiens, Rotshtein, Öhman, & Dolan, 2004), Yin et al.' (2017) data seem to suggest that being dishonest for self-serving benefits in the Sender-Receiver Game involves a stronger concern about self-image. After easing this concern as done in the Coin-Guessing task (Cui et al., 2018), we found that reward-processing reflected by Str activity (Diekhof et al., 2012) was able to explain individual variability in self-serving dishonesty. Altogether, this discrepancy in neural activity may suggest that being selfishly dishonest when the self-image is at stake involves interception-related processes whereas being selfishly dishonest when the self-image is more protected involves reward-related processes. Thus, combining the results from the two tasks provides a clearer picture of the extent to which social-related goals modulate neural cognitive processes of dishonest decision-making.

TABLE 6 Neural activity during the coin-guessing task across participants regardless of the level of dishonesty

Contrast	Region	R/L/M	BA	MNI coordinates			*t*-score	Voxels
				x	*y*	*z*		
Opportunity > No-Opportunity	Inferior parietal lobe	R	40	57	−63	39	8.85	268
	Inferior parietal lobe	L	40	−60	−60	30	7.79	498
	Middle-medial prefrontal cortex[1]	L	10	−6	51	21	6.48	1,063
	Dorsomedial prefrontal cortex[1]	L	10	−6	39	57	4.34	
	Dorsolateral prefrontal cortex[1]	L	9	−39	18	51	4.47	
	Dorsomedial prefrontal cortex	R	10	12	39	54	5.14	236
	Temporal lobe	L	20	−63	−15	−18	4.98	140
No-Opportunity > Opportunity	Occipital cortex and cerebellum	R	19	9	−60	−12	5.51	1,687
Opportunity-Self > Opportunity-Donation	Dorsomedial prefrontal cortex	R	10	3	42	42	5.09	206
	Inferior parietal lobe	R		45	−48	36	4.04	134
Opportunity-Donation > Opportunity-Self	Occipital cortex	R	18	9	−84	−6	8.69	3,431
	Premotor cortex	L	6	−60	3	12	6.03	504
	Premotor cortex	R	6	66	3	15	5.91	471
	Anterior cingulate cortex and dorsomedial prefrontal cortex	R	10	9	6	42	4.62	658
Self > Donation	Dorsal striatum	R		9	12	3	5.04	183
	Anterior cingulate cortex and dorsomedial prefrontal cortex	L	9	−3	39	24	5.23	571
	Anterior insula	R	16	42	24	0	4.85	146
	Inferior parietal lobe	R	40	51	−48	39	4.57	164
Donation > Self	Occipital cortex	L	18	−9	−84	−9	10.07	5,531
	Ventromedial prefrontal cortex	M	11	0	33	−21	5.50	276
	Premotor cortex	R	6	51	−33	63	5.26	1,358
	Premotor cortex	L	6	−51	−15	57	4.35	258
No-Opportunity-Self > No-Opportunity-Donation	Dorsal striatum (caudate)	R	10	9	12	3	5.00	116
	Dorsomedial prefrontal cortex and anterior cingulate cortex	R		3	33	33	4.71	352
No-Opportunity-Donation > No-Opportunity-Self	Occipital cortex[2]	L	18	−9	−84	−9	8.13	1,912
	Posterior cingulate cortex[2]	L	31	−6	−54	21	5.17	
	Angular gyrus	L	39	−39	−63	30	5.93	239
	Occipito-temporal gyrus	L	37	−36	−42	−27	5.81	688
	Ventromedial prefrontal cortex	L	11	−9	42	−12	5.76	196
	Occipital cortex and cerebellum	R	18	15	−105	9	4.98	212
(Opportunity-Self > Opportunity-Donation) > (No-Opportunity-Self > No-Opportunity-Donation)	No supra-threshold clusters were found							
(No-Opportunity-Self > No-Opportunity-Donation) > (Opportunity-Self > Opportunity-Donation)	No supra-threshold clusters were found							

The results were based on whole-brain one-sample *t*-test analyses [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. BA, Brodmann areas. Superscripted numbers denote that the regions are from the same cluster.

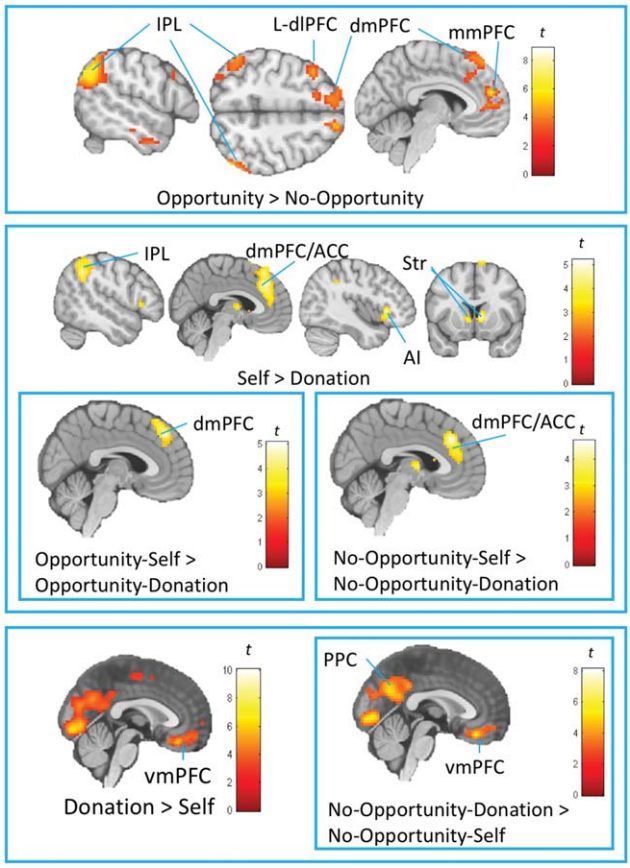

FIGURE 5 Neural activity during the coin-guessing task across participants. The images were based on whole-brain one-sample *t*-test analyses [Cluster-forming threshold at $p < .005$, cluster-wise corrected (pFWE $< .05$)]. IPL, inferior parietal lobe; L-dlPFC, left dorsolateral prefrontal cortex; dmPFC, dorsomedial prefrontal cortex; mmPFC, middle-medial prefrontal cortex; vmPFC, ventromedial prefrontal cortex; ACC, anterior cingulate cortex; AI, anterior insula; Str, striatum; PPC, posterior cingulate cortex [Color figure can be viewed at wileyonlinelibrary.com]

It should be noted, however, that the Coin-Guessing task (Greene & Paxton, 2009) when used with fMRI has certain limitations. Unlike in previous EEG studies (Cui et al., 2018; Hu et al., 2015), the BOLD activity following the onset of the coin-flip outcome may comprise of multiple neural-cognitive processes that overlapped in time due to the poor temporal resolution of fMRI. While we focused on reward and valuation processes, other processes, such as prediction error, may also involve in the BOLD activity in our study. We argue, however, that prediction error is not likely to explain the effects of Overall Dishonesty and Self-Serving Dishonesty shown here. First, it is less probable that participants who had higher Overall Dishonestly would have a stronger prediction error during Opportunity trials than during No-Opportunity trials, given that they lose more money during No-Opportunity trials. Similarly, it is less probable that participants who had higher Self-Serving Dishonesty would have a stronger prediction error during Opportunity-Self trials than during Opportunity-Donation trials, given that they reported higher accuracy during Opportunity-Self trials. Additionally, we found that Overall Dishonesty and Self-Serving Dishonesty were positively associated with stronger activity in the vmPFC and Str. If prediction error explains the behavioral patterns, one would expect Overall Dishonesty and Self-Serving Dishonesty to be explained by a stronger activity in the Anterior Cingulate Cortex (Brown & Braver, 2005) and a weaker (not stronger) activity in the Str (Diekhof et al., 2012). Future studies with different experimental designs are needed to formally investigate the involvement of prediction error.

In summary, we demonstrated that activity in the valuation system and functional connectivity between the valuation and executive control systems play an important role in social-related dishonest decision-making. Specifically, activity in two key areas of the valuation system, the vmPFC and Str, and their interactions with areas in the executive control systems were separately associated with the two processes: one concerned with overall benefits of dishonest acts and the other concerned with whether the self was the beneficiary. Therefore, we suggest that theories of dishonesty should be modified to include the reward and valuation processing as another important factor in explaining self-serving/prosocial dishonesty (Mazar & Ariely, 2006; Shalvi et al., 2015).

ACKNOWLEDGMENTS

The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript. The authors thank Jing Wen Chai, Anna Jos and Avijit Chowdhury for their help with proofreading the final version of this paper.

ORCID

Rongjun Yu http://orcid.org/0000-0003-0123-1524

REFERENCES

Abe, N., & Greene, J. D. (2014). Response to anticipated reward in the nucleus accumbens predicts behavior in an independent test of honesty. *Journal of Neuroscience*, *34*(32), 10564–10572.

Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience*, *19*(2), 287–295.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384.

Barkan, R., Ayal, S., Gino, F., & Ariely, D. (2012). The pot calling the kettle black: Distancing response to ethical dissonance. *Journal of Experimental Psychology: General*, *141*(4), 757–773.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, *76*, 412–427.

Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, *64*(5), 756–770.

Becker, G. S. (2000). Crime and punishment: An economic approach. In N. G. Fielding, A. Clarke, & R. Witt (Eds.), *The economic dimensions of crime* (pp. 13–68). London: Palgrave Macmillan UK.

Braams, B. R., Güroğlu, B., de Water, E., Meuwese, R., Koolschijn, P. C., Peper, J. S., & Crone, E. A. (2014). Reward-related neural responses are dependent on the beneficiary. *Social Cognitive and Affective Neuroscience*, *9*(7), 1030–1037.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118–1121.

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *The Journal of Neuroscience*, *29*(39), 12315–12320.

Cox, R. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173.

Cox, R., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017). FMRI clustering in AFNI: False positive rates redux. *Brain Connectivity*, *7*(3), 152.

Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2), 189–195.

Cui, F., Wu, S., Wu, H., Wang, C., Jiao, C., & Luo, Y. (2018). Altruistic and self-serving goals modulate behavioral and neural responses in deception. *Social Cognitive and Affective Neuroscience*, *13*(1), 63–71.

Diekhof, E. K., Kaps, L., Falkai, P., & Gruber, O. (2012). The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude – An activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia*, *50*(7), 1252–1266.

Ding, X. P., Gao, X., Fu, G., & Lee, K. (2013). Neural correlates of spontaneous deception: A functional near-infrared spectroscopy (fNIRS) study. *Neuropsychologia*, *51*(4), 704–712.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, *6*(3), 218–229.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210.

Garrett, N., Lazzaro, S. C., Ariely, D., & Sharot, T. (2016). The brain adapts to dishonesty. *Nature Neuroscience*, *19*(12), 1727–1732.

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, *95*(1), 384–394.

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences*, *106*(30), 12506–12511.

Hare, T. A., Camerer, C. F., Knoepfle, D. T., O'Doherty, J. P., & Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *The Journal of Neuroscience*, *30*(2), 583–590.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience*, *28*(22), 5623–5630.

Hu, J., Li, Y., Yin, Y., Blue, P. R., Yu, H., & Zhou, X. (2017). How do self-interest and other-need interact in the brain to determine altruistic behavior?. *Neuroimage*, *157*, 598–611.

Hu, X., Pornpattananangkul, N., & Nusslock, R. (2015). Executive control- and reward-related neural processes associated with the opportunity to engage in voluntary dishonest moral decision making. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(2), 475–491.

Lisofsky, N., Kazzer, P., Heekeren, H. R., & Prehn, K. (2014). Investigating socio-cognitive processes in deception: A quantitative meta-analysis of neuroimaging studies. *Neuropsychologia*, *61*, 113–122.

Mameli, F., Sartori, G., Scarpazza, C., Zangrossi, A., Pietrini, P., Fumagalli, M., & Priori, A. (2016). Chapter 16 - Honesty A2 - Absher. In J. Cloutier (Ed.), *Neuroimaging personality, social cognition, and character* (pp. 305–322). San Diego: Academic Press.

Maréchal, M. A., Cohn, A., Ugazio, G., & Ruff, C. C. (2017). Increasing honesty in humans with noninvasive brain stimulation. *Proceedings of the National Academy of Sciences*, *114*(17), 4360–4364.

Mazar, N., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Journal of Public Policy & Marketing*, *25*(1), 117–126.

McDonald, I. (2002). Brokers get extra incentive to push funds. *Wall Street Journal* (April 08).

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage*, *61*(4), 1277–1286.

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., … Dalgleish, T. (2009). A key role for similarity in vicarious reward. *Science (New York, N.Y.)*, *324*(5929), 900.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.

Nicolle, A., Klein-Flugge, M. C., Hunt, L. T., Vlaev, I., Dolan, R. J., & Behrens, T. E. (2012). An agent independent axis for executed and modeled choice in medial prefrontal cortex. *Neuron*, *75*(6), 1114–1121.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–776.

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E. J., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: Psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, *7*(5), 604–609.

Penny, W., & Holmes, A. (2004). Random-effects analysis. In R. Frackowiak, J. Ashburner, W. Penny, S. Zeki, K. Friston, C. Frith, R. Dolan, & C. Price (Eds.), *Human brain function* (pp. 843–850). San Diego: Elsevier.

Pernet, C. R., Wilcox, R., & Rousselet, G. A. (2013). Robust correlation analyses: False positive and power validation using a new open source matlab toolbox. *Frontiers in Psychology*, *3*, 606.

Perry, R. W., & Lindell, M. K. (2003). Understanding citizen response to disasters with implications for terrorism. *Journal of Contingencies and Crisis Management*, *11*(2), 49–60.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562.

Sanford, J. (2014). Confessions of a financial advisor. *CNBC*. Retrieved from http://www.cnbc.com/2014/06/19/confessions-of-a-financial-advisorpersonal-financecommentary.html

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Shalvi, S., & De Dreu, C. K. W. (2014). Oxytocin promotes group-serving dishonesty. *Proceedings of the National Academy of Sciences*, *111*(15), 5503–5507.

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications. *Current Directions in Psychological Science*, *24*(2), 125–130.

Sun, D., Chan, C. C. H., Hu, Y., Wang, Z., & Lee, T. M. C. (2015). Neural correlates of outcome processing post dishonest choice: An fMRI and ERP study. *Neuropsychologia*, *68*, 148–157.

Tenbrunsel, A. E. (1998). Misrepresentation and expectations of misrepresentation in an ethical dilemma: The role of incentives and temptation. *The Academy of Management Journal*, *41*(3), 330–339.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science (New York, N.Y.)*, *315*(5811), 515–518.

Yin, L., Hu, Y., Dynowski, D., Li, J., & Weber, B. (2017). The good lies: Altruistic goals modulate processing of deception in the anterior insula. *Human Brain Mapping*, *38*(7), 3675–3690.

Yin, L., & Weber, B. (2016). Can beneficial ends justify lying? Neural responses to the passive reception of lies and truth-telling with beneficial and harmful monetary outcomes. *Social Cognitive and Affective Neuroscience*, *11*(3), 423–432.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Pornpattananangkul N, Zhen S, Yu R. Common and distinct neural correlates of self-serving and prosocial dishonesty. *Hum Brain Mapp*. 2018;00:1–18. https://doi.org/10.1002/hbm.24062