Biological Psychology xxx (xxxx) 108713

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Blame and Praise cross-culturally: An fMRI investigation into causal attribution and moral judgment

BoKyung Park *, Valerie Rae Smith

The University of Texas at Dallas, USA

ARTICLE INFO

Keywords:
Culture
Causal attribution
Moral judgment
DmPFC

ABSTRACT

People from independent cultures are more likely to causally explain others' behaviors by their disposition [vs. situation] compared to those from interdependent cultures. However, few studies have directly examined how these differences in attribution shape individuals' moral judgment, nor the underlying neural mechanisms of this process. Aiming to address these questions, in the scanner, participants rated the blameworthiness or praiseworthiness of protagonists who did either a negative or positive behavior, respectively. These behaviors were pretested and found to be perceived as dispositionally or situationally caused to different extents on average. Regardless of their self-construal, participants showed enhanced dorsomedial prefrontal cortex (dmPFC) activity in response to the behaviors that were evaluated as more situationally caused on average. Importantly, relatively independent participants reduced their blame for the behaviors that they showed greater dmPFC activity to. Relatively interdependent participants reduced blame for the behaviors that they themselves inferred more situational causes for, but dmPFC activity did not explain their blame. These findings suggest that while dmPFC might support relatively independent participants' effortful consideration of situational contributors to a behavior to make moral judgments, relatively interdependent participants might engage in this process automatically and relied less on dmPFC recruitment.

When judging others' behavior as right or wrong, people differentially rely on either dispositional (i.e., internal characteristics of a target individual) or situational (i.e., outside or contextual forces) factors surrounding the behavior. Accumulated evidence suggests that people's focus on dispositional vs. situational factors can be shaped by their cultural background. However, the neural underpinnings of how these differences are reflected in people's moral judgment has been understudied. In this research, we aim to address this gap in the literature, focusing on the role of brain regions involved in theory-of-mind (ToM), or thinking about others' minds.

1. Attribution, moral judgment, and culture

Suppose you have dinner plans with your friend, Kelly. You rearranged your schedule to have this dinner and left work early. You were parking at the restaurant when you received a text from Kelly, saying that she cannot make it to today's dinner. How much would you blame her for this?

When making moral judgments, the causal responsibility of the protagonist—the extent to which responsibility for the behavior can be attributed to them—is critical (Cushman, 2008). Situational contributors of a behavior importantly reduce the causal responsibility of the protagonist (Shultz et al., 1981) and make the behaviors morally less charged (Heider, 1958; Weiner, 1995). Kelly would be more blameworthy when her unique, internal characteristic was solely responsible for her behavior (e.g., she canceled plans at the last minute because she is unreliable), compared to when there were other external/situational contributors (e.g., she canceled plans due to a last-minute change in her work shift).

Prior studies, however, have found that people often overrepresent the contributions of internal and dispositional characteristics to a behavior, while underestimating the situational influence (Jones & Harris, 1967; Nisbett et al., 1973; Ross, 1977)—a tendency that persists when observing moral vs. immoral behaviors. For example, people perceive a thief who stole something due to external pressure as still morally wrong and would remain dishonest in a new situation (Reeder & Spores, 1983). Because others' moral traits fundamentally define future interactions with them, people are motivated to learn if others are morally good or bad (Uhlmann et al., 2015), speculating their mental states to be in line with their inferred moral traits (Alicke & Zell, 2009;

* Correspondence to: Department of Psychology, The University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080, USA.
E-mail address: bokyung.park@utdallas.edu (B. Park).

https://doi.org/10.1016/j.biopsycho.2023.108713
Received 23 May 2023; Received in revised form 11 October 2023; Accepted 11 October 2023
0301-0511/© 20XX

Kliemann et al., 2008; Knobe, 2010; Siegel et al., 2017). In other words, dispositionism is prevalent especially regarding the judgment of moral behaviors.

But how "prevalent" is it? Importantly, researchers found the above-stated dispositionism varies as a function of individuals' cultural background, especially dependent on their culturally-shaped self-construal. Independent self-construal, largely endorsed in European American cultures (Markus & Kitayama, 1991), views the self as independent and autonomous. Since the self governs one's actions, free from the influences of others or the surrounding situations (Markus & Kitayama, 2003; Mesquita & Markus, 2004), protagonists' dispositions should be responsible for their behavior (Feinberg et al., 2019). On the other hand, the interdependent self, predominantly endorsed in East Asian cultures, views the self as stemming from concerted interactions between internal states and external factors, such as relationships, social roles, and the specific context of the event (Markus & Kitayama, 1991, 2003; Mesquita & Markus, 2004). The protagonists' dispositions cannot be solely responsible for their actions, and situational factors should come into play. Thus, situational consideration is essential in judging others' behaviors.

Mirroring these differences, prior works found that people from individualistic and independent backgrounds make more dispositional inferences than situational, and attribute behaviors to dispositions more than situations, compared to those from collectivistic and interdependent backgrounds (Choi et al., 2003; Lee et al., 2017; Miller, 1984; Miller & Luthar, 1989; Nisbett et al., 2001). For example, when causally explaining mass murders, American news reports focused more on the disposition of the murderer (e.g., bad temper) while Chinese news reports provided more speculation about situational and contextual backgrounds (e.g., societal pressure) (Morris & Peng, 1994). Similarly, Indians used more situation-focused terms and less disposition-focused terms than Americans to describe morally good or bad behaviors (Miller & Luthar, 1989), as well as discounted the accountability of the protagonist when enough contextual information was available (Bersoff & Miller, 1993).

Despite the large volume of evidence on the cultural shaping of attribution, little research has directly connected this to moral judgment, nor to the underlying neural mechanisms (Park et al., 2022). A few exceptional studies reported that Chinese participants perceived financial fraudsters who were under strong situational pressure as less immoral compared to Americans (Wong-On-Wing & Lui, 2007, 2013). Additionally, when punishing a protagonist's harmful actions, their degree of autonomy was more important for Americans than for Chinese (Feinberg et al., 2019). One thing to note is that most of these studies provided strong situational causes for the protagonists' behaviors (e.g., orders from someone higher in the hierarchy, pressure from their group, etc.; see also Feinberg et al., 2019, Study 3). This corroborates the findings that in order to incorporate situational contributors in moral judgments, the situations might need to be salient, even for people from collectivistic and interdependent cultures (Di et al., 2021; Masuda & Kitayama, 2004; Miyamoto & Kitayama, 2002; Norenzayan et al., 2002). For example, prior research found that without experiencing the same situational constraints as the protagonists, East Asians made similar attributions to their American counterparts, focusing on dispositional contributors (Choi & Nisbett, 1998). Yet, people from different cultures may still consider dispositional vs. situational factors differently when making moral judgments, but in a rather subtle way in the absence of salient situational cues. While traditional behavioral methods may not be optimal to detect this, a neural approach can provide insights into this nuanced process.

Another thing to note is that the aforementioned body of research focused on the judgment of negative behaviors (Feinberg et al., 2019; Wong-On-Wing & Lui, 2007, 2013). While the causal responsibility of the protagonist is critical in blaming them (Cushman, 2008), judgments of praise rely less on causal attribution of the behavior (Bostyn & Roets, 2016). People also engage in causal attribution less for positive behaviors than for negative behaviors (Bohner et al., 1988), as well as find less intentionality in positive behaviors (Hindriks, 2008; Knobe, 2003; Pizarro et al., 2003). Given the reduced importance of causal attribution for judging positive behaviors, cultural differences in the role of causal attribution in moral judgment may be less apparent in judging positive behaviors. By using a neuroimaging approach, we investigated whether ToM-associated brain regions explain how people consider dispositional vs. situational causes when judging others' morality, based on their negative vs. positive behaviors.

2. Attribution and moral judgment are supported by ToM regions

Brain regions involved in ToM—the bilateral posterior superior temporal sulcus (pSTS), precuneus, bilateral temporo-parietal junction (TPJ), and dorsomedial prefrontal cortex (dmPFC)—support inferences of others' goals, beliefs, and traits (Ma et al., 2011; Mitchell et al., 2006; Saxe & Kanwisher, 2003; Van Overwalle, 2009; Van Overwalle & Baetens, 2009; Van Overwalle & Vandekerckhove, 2013). Since considering others' mental states and traits is critical in judging their behaviors, these regions are similarly recruited when people make judgments of morality (Buckholtz & Marois, 2012; Greene et al., 2004; Young & Koenigs, 2007). Particularly, prior research reported that the dmPFC is recruited in response to a variety of moral judgment tasks, including those associated with viewing moral transgressions [vs. neutral behaviors] (Parkinson et al., 2011; see also Jenkins & Mitchell, 2010), viewing behaviors that harmed another person [vs. did not harm] (Young & Koenigs, 2007), learning of extreme positive and negative behaviors associated with faces (Baron et al., 2011), the perceived praiseworthiness of others' behaviors, especially for participants with high justice sensitivity (Yoder & Decety, 2014), and the reduction of punishment considering sympathetic circumstances (e.g., murdering one's husband after suffering from domestic violence from him; Yamada et al., 2012). Given that dmPFC is associated with the effortful process of (Lieberman, 2022; Meyer et al., 2015) and abstract construal of (Baetens et al., 2014) social information, these findings suggest that people recruit dmPFC in order to effortfully construe the complex nature of social events, ultimately feeding their moral judgment.

Interestingly, prior studies that primarily recruited participants from independent cultural contexts (i.e., the US and Western Europe) reported that dmPFC activity tracked greater consideration of situational information as well. In one study (Kestemont et al., 2013), participants judged the causes of others' behaviors that could be attributed to their disposition (e.g., Jun gives a bouquet at arrival [Jun is romantic]) or to the situation (e.g., Gabril changes the ink [the ink holder was empty]). Bilateral TPJ and dmPFC were recruited more when they made situational [vs. dispositional] attributions. These regions were also recruited when participants were processing others' behaviors (e.g., "Tom left the restaurant in a hurry without tipping the waitress") and subsequent situational explanations of the behaviors (e.g., "Tom's baby was screaming"); dmPFC further tracked how much participants endorsed the situational explanation (Brosch et al., 2013). These findings suggest that greater engagement of dmPFC, indexing enhanced ToM and/or effortful construal of the social event, may be needed for individualistic and independent participants to engage in situational attribution which they do not often make. Whether these regions similarly support collectivistic and interdependent individuals' attribution and moral judgment, though, has not yet been examined.

3. The current study

In the current study, first, we explored if people with relatively independent vs. relatively interdependent self-construal use dispositional vs. situational attribution differently when making moral judgments. Although a few previous studies (Feinberg et al., 2019; Wong-On-Wing

& Lui, 2007, 2013) demonstrated that people from independent [vs. interdependent] backgrounds consider the protagonist's agency [vs. salient situation] more when blaming others, these studies 1) provided salient situational information to participants, which may not always be apparent in the real world, and 2) only examined blame for negative behaviors. We will explore if people use dispositional vs. situational causes any differently depending on their self-construal, even in the absence of salient situational cues. We will also test if this self-construal influence varies as a function of the nature of the behavior, i.e., blameworthy negative behaviors and praiseworthy positive behaviors.

Second, we examined if ToM regions—especially dmPFC—can account for any of the cultural shaping of attribution and moral judgment. In a recent theoretical review (Park et al., 2022), we predicted that when people from individualistic and independent [vs. collectivistic and interdependent] backgrounds situationally explain others' behaviors, they would recruit ToM regions to a greater extent [vs. lesser extent], mirroring their effortful processing of social events to mitigate their blame. In contrast, when people from collectivistic and interdependent [vs. individualistic and independent] backgrounds dispositionally explain others' behaviors, they would recruit ToM regions to a greater extent [vs. lesser extent], effortfully construing the behavior as reflecting the protagonist's disposition and enhancing their blame. Elaborating on these predictions, we hypothesized that H1) people with relatively independent [vs. relatively interdependent] self-construal would recruit ToM regions, especially dmPFC, to a greater extent in response to the behaviors that were more situationally [vs. dispositionally] explained, and H2) increased dmPFC activity of those with relatively independent [vs. relatively interdependent] self-construal would predict mitigated [vs. magnified] blame. Due to the reduced importance of causal attribution in judging positive behaviors, we did not expect distinctive cultural differences in the role of dmPFC in participants' praise.

4. Methods

4.1. Participants

We recruited 41 healthy, right-handed European American and East Asian American college students from the Dallas-Fort Worth area. European American participants were required to be born and raised in the US, have parents who were born and raised in the US, and have grandparents who were born and raised in the US or Western Europe (e.g., Germany). East Asian American participants were required to be born and raised in the US, and have parents and grandparents who were born and raised in an East Asian nation (e.g., China, Japan, Korea, Singapore, Taiwan, or Vietnam). One participant was excluded due to excessive head movement (> 2.5 mm), leaving 40 participants in the final analyses (24 females, 16 males; gender identification based on participants' response to "What is your gender? Male, Female, Other"; age M = 20.15 [range 18 – 26], S.D. = 1.79). This final sample size corresponds to or exceeds the sample size of similar studies contrasting two cultural groups (Adams et al., 2010, Han et al., 2011; Park et al., 2016, 2017). However, due to the lack of cultural differences in self-construal (1: Extremely disagree – 7: Extremely agree; Independence: European American M = 4.64, S.E. =.18, East Asian American M = 4.58, S.E. =.12 [range 3.13 – 6.07]; Interdependence: European American M = 4.71, S.E. =.19, East Asian American M = 4.91, S.E. =.14 [range 3.07 – 6.20], *ps* >.38), we collapsed across cultural groups and used self-construal scores as our main predictors in the subsequent analyses. The effect of self-construal persisted after we controlled for participants' cultural groups (Supplementary Section 1). Since the findings were not qualified as a function of participant gender, we dropped this factor from the final analyses. All procedures were approved by the Institutional Review Board at the University of Texas at Dallas. Informed consent was obtained from all participants in advance of any experimental procedure.

4.2. Stimuli

We retrieved a total of 96 different behavioral descriptions of 96 different protagonists from Kim et al., 2021, half negative (e.g., "Emily stole and used her cousin's credit card") and the other half positive (e.g., "Angela covered an overnight shift for her sick coworker"). The findings of the current study persisted after we controlled for perceived frequency, moral relevance, and emotional arousal of the behaviors, and the trustworthiness and intelligence of the protagonists, that were pre-tested in the prior study (Kim et al., 2021) (Supplementary Section 2).

To measure how much dispositional vs. situational cause people perceived in each behavior on average, each behavioral description was presented to 24–26 raters on Prolific (https://www.prolific.co/), for a total of 2400 raters (age M = 40.70, S.D. = 13.24; 44.2% female, 54.6% male, 1.2% other; 74.6% White American, 9.2% Black American, 5.2% Latino American, 6.9% Asian/Asian American, 0.3% Arab American, 0.4% Native American, 3.0% Multiracial, 0.2% Other, 0.2% no responses). The gender of the protagonists matched the Prolific raters' gender. The Prolific raters evaluated how much the behavior was primarily caused by the disposition of the protagonist, such as the protagonist's personality and/or internal desires, and how much the behavior was primarily caused by the situation, such as the protagonist's current circumstance and/or the influence of other people, using two separate 7-point likert scales ("To what extent is this behavior primarily caused by the person? (for example [Name's] personality and/or internal desires)?" 1: Not at all – 7: Entirely the person; "To what extent is this behavior primarily caused by the situation? (for example [Name's] current circumstance and/or the influence of other people)?" 1: Not at all – 7: Entirely the situation; Supplementary Section 3, Table S1). We averaged Prolific raters' ratings of dispositional attribution and situational attribution per each behavior, creating "Average Dispositional Attribution" and "Average Situational Attribution" indexes, respectively.[1]

4.3. Moral judgment task

We created the "Moral judgment task" to measure fMRI participants' evaluation of the 96 protagonists who did the 96 negative or positive behaviors (Fig. 1A). Each behavior was presented for 4 s, followed by the prompt to evaluate the blameworthiness of the protagonist who did the negative behavior ("How blameworthy is [Name]?") or the praiseworthiness of the protagonist who did the positive behavior ("How praiseworthy is [Name]?") using a 7-point sliding bar (1: Not at all – 4: Moderately – 7: Extremely) for 4 s. Trials were divided by a jittered fixation (4–8 s, Mean = 6 s). The gender of the protagonists matched the participants' gender. Behaviors were presented in a randomized order, unique to each participant.

4.4. Procedure

After providing informed consent, participants were told that they would do the Moral judgment task, in which they would read about others' negative or positive behaviors and evaluate the blameworthiness or praiseworthiness of them. Afterwards, they were informed of the differences between dispositional and situational attribution, and rated if the

[1] Average Dispositional Attribution and Average Situational Attribution scores were inversely correlated ($r = -.61$, $p < .001$), consistent with prior research (Guimond et al., 1989; Houston, 1990; Howard, 1987; Reeder et al., 2004; see also Miller et al., 1981). The findings persisted when only Average Dispositional Attribution and only Average Situational Attribution were entered in the model, separately (Supplementary Section 4, Tables S2, S3).

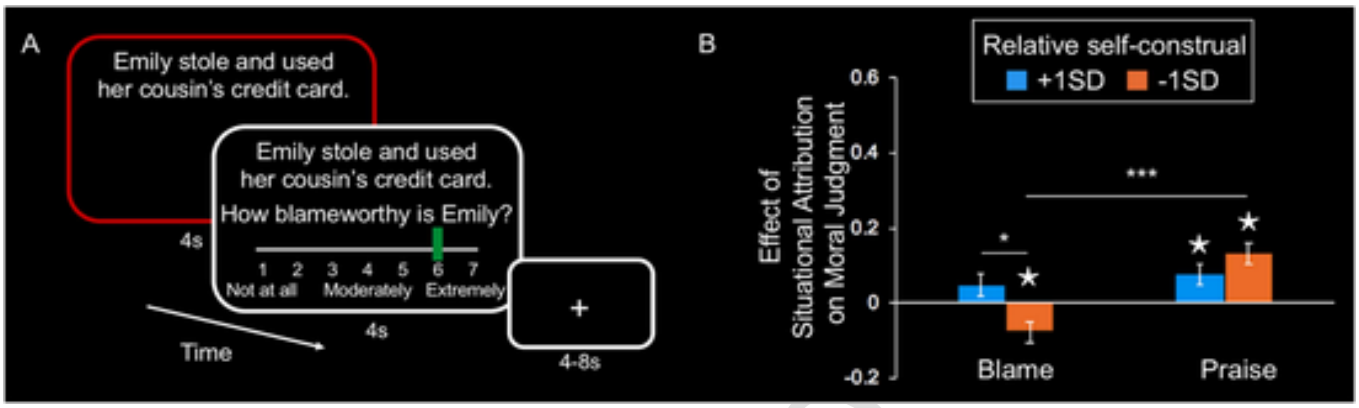

Fig. 1. A) Moral judgment task. Participants first read each protagonist's negative or positive behavior for 4 s, then evaluated their blameworthiness (for negative protagonist) or praiseworthiness (for positive protagonist) for another 4 s, followed by a jittered fixation (4–8 s). Functional brain imaging analyses focused on the time participants first read the behaviors, marked in red. B) Effect of participants' situational attribution on their moral judgments. Relatively interdependent (relative self-construal −1 SD) participants mitigated blame for protagonists whose behaviors they saw as more situationally caused. Greater situational attribution was associated with increased praise overall. Relative self-construal: Independence – Interdependence. * $p < .05$, * ** $p < .001$, ⋆ Effects that do not contain 0 in 95% Confidence Interval.

explanations of two exemplar behaviors were dispositional or situational as a comprehension check (e.g., Kelly canceled plans with a friend at the last moment due to the last-minute change in her work shift; Supplementary Section 5). Participants then practiced four trials of the Moral judgment task. In this practice as well as in the real task, unlike the two examples, participants viewed only the behaviors without any explanations (e.g., Kelly canceled plans with a friend at the last moment). The behaviors participants saw during the practice were not presented during the real task.

Participants then moved to the scanner, underwent a T1-weighted anatomical scan, and then completed 96 trials of the Moral judgment task divided into four runs (each run with 24 trials; total time = 23 min 44 s) while their functional brain activities were acquired. Afterwards, participants completed two runs of a ToM localizer task (10 trials in each run; total time = 9 min 4 s) (Dodell-Feder et al., 2011). On each trial, participants read a story about social protagonists ("belief" condition) or physical objects ("photo" condition) (10 s) and judged if a statement was true or false based on the story (4 s), followed by a fixation (12 s). In the belief condition, participants had to infer the protagonist's mental state to judge the statement (e.g., "Lisa now believes that Jacob is sleeping"). In the photo condition, participants had to think about the physical attributes of an object (e.g., "Today the color of the blouse is white"). One participant did not complete this task due to a computer malfunction.

After the ToM localizer task, participants moved to a behavioral room and viewed the 96 behaviors from the Moral judgment task again. Participants estimated how much each behavior was caused by dispositional or situational factors using two separate 7-point likert scales ("To what extent is this behavior primarily caused by the person [the situation]?" 1: Not at all – 7: Extremely). Finally, they completed the self-construal scale (Singelis, 1994), demographics questionnaire, and were thanked and compensated (see Supplementary Section 6 for other measures administered for a different project).

4.5. FMRI acquisition and analyses

We used a 3 T Siemens scanner outfitted with a 32-channel head coil at the Sammons BrainHealth Imaging Center at the University of Texas at Dallas. Sixty 2.5 × 2.5 × 2.5 mm slices of gradient echo T2 * -weighted echo-planar images provided whole brain coverage [time repetition (TR) = 2 s, echo time (TE) = 25 ms, flip angle = 77°] for functional scans. High-resolution [1.0 × 1.0 × 1.0 mm; TR = 2.30 s, TE = 2.26 ms] anatomical scans were additionally acquired while participants were looking at a blank screen.

We analyzed brain data using Analysis of Functional Neural Images (AFNI_20.3.00 version) software (Cox, 1996). Per each run of functional scans, the initial six scans were removed to compensate for magnet stabilization. All other images were slice-timing corrected (using the first slice as reference), deobliqued, motion corrected (using Fourier interpolation; third volume as a reference), spatially smoothed (using a 3D isotropic Gaussian kernel of a 8 mm full width at half maximum), normalized by the average activity over each run to generate percent signal change (PSC), high-pass filtered (removing frequencies < 0.01 Hz), and concatenated across runs within each Moral judgment task and ToM localizer task (following procedures described in Wu et al., 2014).

4.5.1. Whole-brain analyses

For the functional scans acquired during the Moral judgment task, we conducted the analyses on the time points when participants first looked at the behaviors, before being prompted to evaluate the behaviors (Fig. 1A). During this period, participants read the behaviors and processed them, but were not yet able to view the scale or indicate their judgment. This allowed us to control for motor responses associated with moving the index on the scale, as well as capture the spontaneous processing of others' behaviors not explicitly under the prompt of blaming or praising the behavior (Genevsky et al., 2013; Park et al., 2017).

We examined whether participants showed any distinctive neural activities in response to the behaviors that were on average rated as more dispositionally driven vs. situationally driven. In a general linear model (GLM, ordinary least-squares regression) including fourteen regressors, eight regressors of no interest were included to minimize the influence of physiological confounds: six modeling head movement, one sampling white matter activity, and one sampling cerebrospinal fluid activity (Chang & Glover, 2009). Among the six regressors of interest, the first regressor marked the time points of each trial when participants read the behavior. Other regressors marked: (1) the Valence of the behavior, contrasting the negative [− 1] and positive [+ 1] behaviors, (2) Average Dispositional Attribution ratings provided by Prolific raters, (3) Average Situational Attribution ratings provided by Prolific raters, (4) the Valence X Average Dispositional Attribution interaction, and (5) the Valence X Average Situational Attribution interaction. Before they were submitted to the model, regressors of interest were convolved with a canonical gamma variate hemodynamic delay (Cohen,

1997). Linear regression t-statistic maps were converted to Z-scores, co-registered with structural maps, spatially normalized by warping to Montreal Neurological Institute space (MNI; linear to colin27T1_seg template), and resampled as 2.5 mm cubic voxels.

To examine if participants' self-construal influenced their neural responses to each of these regressors, we created participants' independence (e.g., "My personal identity, independent of others, is very important to me"; Cronbach's α =.69) and interdependence (e.g., "It is important to me to respect decisions made by the group"; Cronbach's α =.78) scores by averaging the corresponding subscales from the self-construal scale (Singelis, 1994). We subtracted each participant's interdependence score from their independence score, generating a "relative self-construal" index (ranges between −2.47 and 1.93) per each participant, which was submitted as a covariate in the one-sample t-test on the regressors of interest. All findings persisted when we entered participants' independence and interdependence separately in the model (Supplementary Section 7; Table S4). These group maps were initially voxelwise thresholded (at $p < .005$) and then cluster thresholded (cluster size > 142 continuous 2.5 mm cubic voxels) to yield corrected maps for detecting whole-brain activity ($p < 0.05$ corrected), derived with 10000 Monte Carlo iterations using AFNI program 3dClustSim (Supplementary Section 4).

4.5.2. Volume-of-interest analyses

For Volume-of-Interest (VOI) analyses, we located each participant's dmPFC coordinates from the ToM localizer task, by subtracting their brain activities during the photo condition from those during the belief condition (Supplementary Section 8A). We created a spherical VOI mask (5 mm radius) centered on dmPFC coordinates unique to each participant. For the one participant who did not complete the ToM task, and for the participants who did not show dmPFC activations during the ToM task, we averaged belief vs. photo contrast maps across all participants, located the group dmPFC coordinates (MNI coordinates dmPFC: [8, 50, 24]; Table S5), and created a VOI mask centered on it. We extracted PSC data averaged across all voxels within the masks, during the phase that participants first read the behavioral description of each protagonist (Fig. 1A). Sampling was lagged by 4 s because of the hemodynamic lag to peak (Knutson et al., 2007). PSC outliers exceeding 4 SD from the mean activity were excluded from further analyses (Stallen et al., 2021), although findings persist without excluding any outliers.

5. Results

5.1. Self-construal modulated the effect of situational attribution on moral judgment, but not the effect of dispositional attribution

We explored if participants' self-construal modulated the degrees to which they made dispositional and situational attributions about the behaviors. We conducted a linear mixed-effect regression using R package nlme (Pinheiro et al., 2019) on the dispositional and situational attribution participants made about the behaviors after the scan, separately, including individual participants as random effects. Participants' relative self-construal, the valence (negative, positive) of the protagonists' behaviors, and the interaction between relative self-construal and valence of the behaviors were entered as fixed effects. We also controlled for participants' situational attribution ratings in the model that predicted their dispositional attribution, and controlled for participants' dispositional attribution ratings in the model that predicted their situational attribution. In all mixed-effect regression models including these, findings persisted when we replaced relative self-construal with relatively independent and relatively interdependent median-split subgroups (Supplementary Section 9). In both models, there were no main effects or interaction effects involving relative self-construal (*ps* >.12; Table S6), consistent with previous literature that people from interdependent cultures make similar attributions to their independent counterparts in the absence of salient situational cues (Choi & Nisbett, 1998; Di et al., 2021; Masuda & Kitayama, 2004; Miyamoto & Kitayama, 2002; Norenzayan et al., 2002).

We further examined how participants considered their dispositional and situational attributions in moral judgment. We conducted another linear mixed-effect regression on participants' trial-by-trial moral judgment ratings, including individual participants as random effects. Participants' relative self-construal scores, the valence of the protagonists' behaviors, and the dispositional and situational attribution ratings they made after the scan, were entered as fixed effects. The interactions between relative self-construal, dispositional attribution ratings, and valence, and relative self-construal, situational attribution ratings, and valence were also entered.

First, replicating previous literature (Heider, 1958; Shultz et al., 1981; Weiner, 1995), we found a significant main effect of dispositional attribution (B =.42, S.E. =.02, t = 21.72, $p < .001$). The more dispositional attribution participants made about a behavior, the more they blamed or praised that behavior. This effect was modulated by a significant Valence X Dispositional attribution interaction (B = −.05, S.E. =.02, t = −2.85, p = .004), indicating that greater dispositional attribution increased participants' blame (B =.47, S.E. =.03, 95% CI = [.42, .52], $p < .001$) more than their praise (B =.37, S.E. =.02, 95% CI = [.32, .42], $p < .001$). These effects were not modulated by relative self-construal (*ps* >.15; Table S7).

Second, contrary to previous literature (Heider, 1958; Shultz et al., 1981; Weiner, 1995), greater situational attribution was associated with enhanced moral judgment (B =.05, S.E. =.01, t = 3.20, p = .001). This main effect was qualified by a Valence X Situational attribution interaction (B =.05, S.E. =.01, t = 3.74, $p < .001$), indicating that greater situational attribution was associated with greater praise (B =.10, S.E. =.02, 95% CI = [.07, .14], $p < .001$) but not with blame (B = −.02, S.E. =.02, 95% CI = [−.06, .03], p = .465). Importantly, this effect was again qualified by a significant Relative self-construal X Valence X Situational attribution interaction (B = −.04, S.E. =.01, t = −3.54, $p < .001$; Fig. 1B). While situational attribution did not impact relatively independent participants' (relative self-construal +1 SD) blame (B =.05, S.E. =.03, 95% CI = [−.01, .11], p = .108), relatively interdependent participants (relative self-construal −1 SD) blamed the protagonist less for the behaviors that they found more situational causes for (B = −.08, S.E. =.03, 95% CI = [−.13, −.02], p = .005), significantly different from relatively independent participants (p = .010). Situational attribution increased the praise of relatively independent (B =.08, S.E. =.03, 95% CI = [.02, .13], p = .007) and relatively interdependent (B =.13, S.E. =.03, 95% CI = [.08, .19], $p < .001$) participants to the same degrees (p = .519). These findings suggest that while people consider dispositional reasons similarly in general, how much situational attribution is considered while making moral judgments may vary as a function of their self-construal.

5.2. Participants recruited dmPFC in response to situationally explained behaviors

We predicted that relatively independent participants would show enhanced dmPFC activity in response to the behaviors that were more situationally explained, while relatively interdependent participants would show enhanced dmPFC activity in response to the behaviors that were more dispositionally explained (H1). To test these hypotheses, we examined participants' brain activities in response to the average dispositional and situational attribution ratings for each behavior made by Prolific raters. Contrary to our predictions, participants' self-construal did not significantly modulate their brain activities in response to the behaviors. Instead, whole-brain analysis revealed that participants overall showed enhanced dmPFC activity in response to the behaviors

that Prolific raters rated as driven by more situational causes, conceptually replicating previous studies (Brosch et al., 2013; Kestemont et al., 2013) (Table 1; Fig. 2A).[2]

5.3. Increased dmPFC activity explained relatively independent participants' reduced blame

We predicted that increased dmPFC activity of relatively independent participants would predict mitigated blame, while increased dmPFC activity of relatively interdependent participants would predict enhanced blame (H2). To test this hypothesis, we conducted a linear mixed-effect regression using R package nlme (Pinheiro et al., 2019) on participants' moral judgment ratings, including individual participants as random effects. Participants' relative self-construal scores, the valence of the behaviors, and dmPFC PSC, and the interactions between these variables were entered as fixed effects.

There were significant main effects of relative self-construal (B = −.24, S.E. =.09, t = −2.61, p = .013) and valence of the behaviors (B = −.12, S.E. =.02, t = −5.74, p < .001). More interdependent participants blamed or praised the protagonists to a greater extent. Also, participants overall blamed the protagonists more than they praised the protagonists.

Importantly, these main effects were qualified by a significant Relative self-construal X Valence X dmPFC PSC interaction (B =.14, S.E. =.05, t = 2.64, p = .008). Partially supporting our hypotheses, for relatively independent participants (relative self-construal +1 SD), increased dmPFC activity was associated with reduced blaming (Effect = −.24, S.E. =.11, 95% CI = [−.46, −.03], p = .027). The activity of dmPFC was not associated with relatively interdependent participants' (relative self-construal −1 SD) blame (Effect =.002, S.E. =.11, 95% CI = [−.22, .23], p = .985), nor participants' praise, regardless of their self-construal (relatively independent: Effect =.13, S.E. =.10, 95% CI = [−.06, .33], p = .185; relatively interdependent: Effect = −.19, S.E. =.11, 95% CI = [−.41, .03], p = .089; Table S8). Other regions that were previously reported to be implicated in moral judgments, such as right TPJ (rTPJ; Park et al., 2022; Young & Koenigs, 2007, 2010, 2011; Young & Saxe, 2009), did not modify participants' moral judgments as a function of their self-construal (Supplementary Section 11). Together, these results indicate that dmPFC activity was uniquely associated with relatively independent participants' mitigated blame.

6. Discussion

This research presents a first attempt to investigate how cultural differences in attribution style can be reflected in individuals' moral judgments of others' behaviors. While participants' self-construal did not modify how they dispositionally or situationally explained others' behaviors, those with relatively interdependent self-construal were less likely to blame protagonists for the behaviors that they found more situational explanations for, compared to those with relatively independent self-construal. Also, consistent with previous literature (Brosch et al., 2013; Kestemont et al., 2013), participants showed increased dmPFC activity in response to the behaviors that were more situationally explained on average. Importantly, the degrees to which dmPFC activity was associated with participants' moral judgments varied as a function of their self-construal. When relatively independent participants showed greater dmPFC activity in response to a negative behavior, they mitigated their blame for the protagonist who did the behavior.

Table 1
Significant activation foci from whole-brain analyses during Moral judgment task.

Region	x	y	z	Peak Z	Voxels
Each time point					
L Lingual Gyrus	-14	-93	-13	5.66	487
R Lingual Gyrus	14	-91	-12	5.84	313
Average Dispositional Attribution					
L Lingual Gyrus	-16	-97	-7	-5.72	607
R Lingual Gyrus	14	-93	-13	-5.56	365
R Cuneus	14	-82	13	3.94	189
Valence X Average Dispositional Attribution					
R Superior Frontal Gyrus	21	62	24	4.11	160
R Middle Frontal Gyrus	29	12	45	3.55	150
Average Situational Attribution					
L dmPFC	-11	55	15	3.51	196
Valence X Average Situational Attribution					
R Cuneus	19	-97	-4	4.23	234

Note. R: Right, L: Left, dmPFC: dorsomedial prefrontal cortex, region of interest in bold; p < .005, cluster size > 142 continuous, face touching voxels, corrected p < .05, MNI coordinates

These findings suggest that relatively independent participants needed to recruit dmPFC when judging the morality of behaviors that are explained situationally, ultimately reducing their blame. Supporting this notion, although statistically marginal, relatively independent participants made greater situational attribution after the scan for the negative behaviors that they previously showed increased dmPFC activity to (Supplementary Section 10). Prior research has also linked enhanced dmPFC activity with effortful social processing (Lieberman, 2022; Meyer et al., 2015), suggesting that for relatively independent participants, the incorporation of situational contributors might be effortful and need to be supported by dmPFC recruitment. Indeed, it has been reported that situational attribution is more effortful for more independent people (Buchtel & Norenzayan, 2009); for example, a prior work found that American participants failed in considering the situational contributors to others' behaviors when they had limited cognitive resources, while Hong Kong Chinese could successfully factor in these contributors, regardless of their available cognitive resources (Knowles et al., 2001).

In contrast, relatively interdependent participants' dmPFC activity was not associated with their moral judgment, contrary to our prediction that dmPFC would support their effortful processing of dispositional contributors and subsequent blame. One possibility is that incorporating dispositional contributors, as well as situational contributors, into moral judgment might not be effortful, but rather automatic, and not need to be supported by dmPFC, for these participants. Supporting this notion, a body of previous research suggested that without salient situational information—just like in the current research—people from interdependent cultures also focus on the dispositional contributors to others' behaviors (Choi & Nisbett, 1998; Di et al., 2021; Masuda & Kitayama, 2004; Miyamoto & Kitayama, 2002; Norenzayan et al., 2002). Moreover, even when participants had limited cognitive resources, both Americans and Hong Kong Chinese could successfully make dispositional attributions, suggesting that dispositional attribution is automatic for both cultural groups (Knowles et al., 2001).

Another possibility would be that above and beyond thinking about others' minds and causal contributors tracked by dmPFC, relatively interdependent participants might search for other information to make their moral judgments; such as how much social harmony the given behavior interrupted or enhanced as a result (Feinberg et al., 2019) or

[2] When we replaced the Average Dispositional Attribution and Average Situational Attribution provided by Prolific raters with fMRI participants' dispositional and situational attribution ratings, whole-brain analysis did not reveal any significant activations (Supplementary Section 8B). We speculate that since participants made their attribution ratings when they viewed the behaviors for the second time outside of the scanner, after judging the morality of the protagonists, the attribution ratings might reflect their moral judgments more and the characterization of dispositional and situational accounts less compared to the Prolific raters' ratings. Still, analyses with dmPFC PSC showed a marginal association between dmPFC activity and relatively independent participants' greater situational attribution for the negative behaviors (Supplementary Section 10).

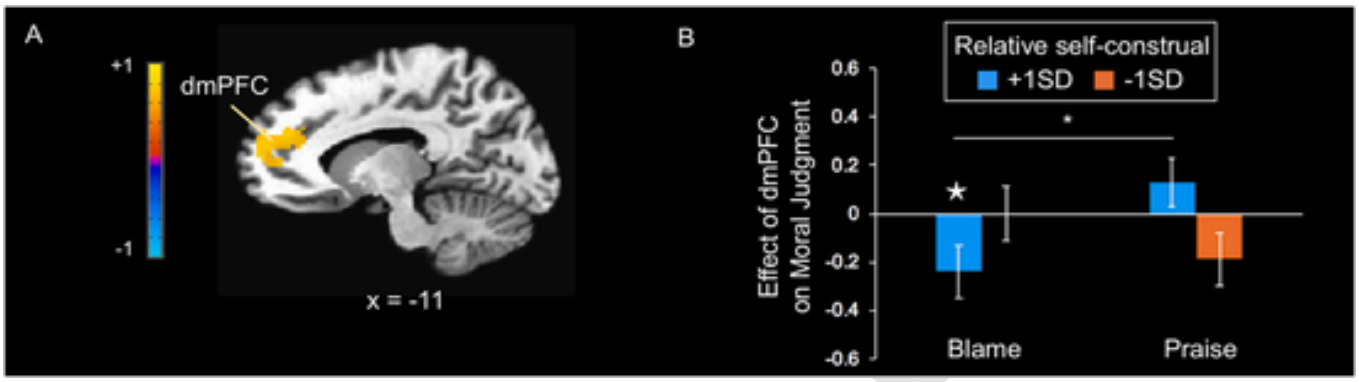

Fig. 2. A) Dorsomedial prefrontal cortex (dmPFC) was activated in response to behaviors that were more situationally attributed on average. $P_{FWE} < .05$, whole-brain corrected. B) Increased dmPFC activity accounted for reduced blame of relatively independent participants (relative self-construal +1 SD). Relative self-construal: Independence – Interdependence. * $p < .05$, ⋆ Effects that do not contain 0 in 95% Confidence Interval.

whether the behavior was consistent with social norms (Mu et al., 2015). Brain regions associated with detecting the outcome of behaviors and integrating information from various sources, such as ventromedial prefrontal cortex (Chib et al., 2009; Ciaramelli et al., 2012; Cooper et al., 2010; De Martino et al., 2013; Glimcher & Fehr, 2013; Park et al., 2022), or regions associated with the detection of norm violation, such as ventrolateral orbitofrontal cortex and insula (Bellucci et al., 2018; Berthoz et al., 2002), may account for relatively interdependent participants' moral judgments better.

RTPJ, another ToM-implicated region that has been reported to support moral judgments (Park et al., 2022, Young et al., 2007, 2010, 2011; Young & Saxe, 2009), indeed showed enhanced activity preceding participants' greater blame of negative behaviors (Supplementary Section 11). Yet, this association was not qualified as a function of participants' self-construal. Given rTPJ's role in encoding the unexpectedness of others' behaviors (Koster-Hale & Saxe, 2013; Park et al., 2021), rTPJ activity might mirror participants' immediate "surprise", evoked by normatively unexpected negative behaviors, and the greater blame that followed. On the other hand, the role of dmPFC in effortful social cognitive processes (Lieberman, 2022; Meyer et al., 2015) may suggest that the influence of self-construal on moral judgment is channeled through rather effortful processing of social events, and is manifested in relatively independent individuals' incorporation of situational explanations of the events.

As predicted, self-construal only shaped how participants blamed the protagonists, but not how they praised the protagonists. Specifically, relatively interdependent participants mitigated their blame for the behaviors that they attributed to more situational causes, suggesting that they successfully factored in situational contributors to the behaviors in their judgment. Interestingly, the more situational causes participants inferred about a behavior, the more they praised the protagonist, regardless of their self-construal. While causal reasoning is less important for the judgment of praise than for the judgment of blame (Bohner et al., 1988; Bostyn & Roets, 2016; Pizarro et al., 2003), people may value when others adaptively go along with the situation to produce a positive behavior, praising the behavior more as well as perceiving the behavior as situationally caused. Further investigation into the effect of situational attribution on blame vs. praise would be required.

These findings importantly broaden cultural psychology by presenting initial evidence about the neural underpinnings of subtle cultural differences in attribution and moral judgment. Without salient situational information, people make similar attributions regardless of their cultural backgrounds (Choi & Nisbett, 1998; Di et al., 2021; Masuda & Kitayama, 2004; Miyamoto & Kitayama, 2002; Norenzayan et al., 2002). However, how they incorporate these attributions into their moral judgment is shaped by culture, which is reflected in their neural responses. This study also expands causal attribution and moral judgment literature, suggesting that under certain conditions, situational consideration may not always be associated with mitigated moral judgment, but rather magnified judgment; the more situational causes participants identified in the behavior, the more they praised the protagonist. Further, this study strengthens the prior body of work that dmPFC may account for situational consideration (Brosch et al., 2013; Kestemont et al., 2013), above and beyond traditionally studied dmPFC functions, such as inferring the mental states of others. Importantly, even though overall participants' dmPFC activity scaled with the average situational attribution about the behaviors, the degrees to which the dmPFC activity contributed to participants' moral judgments varied by their self-construal.

There are some limitations in this research that raise future research questions. First, we did not find support for our prediction that relatively interdependent participants would recruit the ToM network to a greater extent to make dispositional attributions (Park et al., 2022). Potentially this limitation is because all our participants were Americans, born and raised in the US. A clearer cultural distinction may be found if the same study is conducted again with European Americans born and raised in the US, and East Asians born and raised in East Asia. Additionally, due to the nature of being morally relevant, the behaviors—especially negative behaviors—were rated as higher on dispositional attribution than situational attribution by Prolific raters (Supplementary Section 3). These highly dispositional behaviors might yield a ceiling effect, driving all participants to effortlessly consider the dispositional causes of the behaviors in their moral judgments. Future studies may create behaviors that are equally attributable to disposition and situation, and see if individuals' incorporation of dispositional causes can vary as a function of their cultural backgrounds. Further, the behavioral descriptions used in the current study were decontextualized, only a single sentence without any explanation about the protagonists' mind or surrounding situations (e.g., Kelly canceled plans with a friend at the last moment). Participants were also not explicitly prompted to think of the dispositional and situational causes of the behaviors in the scanner, but only provided these ratings afterwards. While the current paradigm provided us an opportunity to examine participants' spontaneous moral judgments, future research may explicitly ask people to think about the dispositional and situational causes of the behavior in the scanner. Second, even though participants praised more situationally driven positive behaviors regardless of their self-construal, the specific reason behind the praise may be different. For example, the maintenance of positive self-view, which could be acquired by being responsive to situational concerns, might be more praiseworthy for relatively independent participants (Scholer &

Higgins, 2010). On the other hand, the increased social harmony which could be achieved by responding to situational concerns might be more praiseworthy for relatively interdependent participants (Feinberg et al., 2019). Third, this research did not examine the cultural differences in how people construe morality. For example, while mature moral agents in independent cultures are expected to make rational decisions and be responsible for their actions, they are expected to respect social harmony and reconcile collectivistic opinions in interdependent cultures (Dien, 1982). These fundamentally different perceptions of moral agents may drive people to judge the same behaviors differently, which should be further studied.

In summary, participants overall showed enhanced dmPFC activity in response to behaviors that were more situationally explained on average. Relatively independent participants' increased dmPFC activity was associated with their reduced blame, suggesting that they might need to effortfully incorporate situational causes of the behavior to mitigate their blame. Relatively interdependent participants' dmPFC activity was not associated with their moral judgment, consistent with the prior work suggesting that both dispositional and situational consideration is effortless and automatic for interdependent people. Together, these findings provide the first neural evidence showing the nuanced cultural shaping of causal attribution and moral judgment.

Declaration of Generative AI and AI-assisted technologies in the writing process

No generative AI or AI-assisted technologies were used in writing this paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The stimuli, data, and code for analyses presented in this study are available in the Open Science Framework repository at [https://osf.io/bt9xa/?view_only=7caff660c8e54385a9a26f21bf5019a7].

Acknowledgment

We thank U. Hsieh, A. Alhaque, and O. Drake for assistance with running participants, members of SPRING Lab for feedback on earlier versions of this manuscript, M. Kim for sharing the stimuli, and the staff and faculty of the Sammons BrainHealth Imaging Center at UT Dallas for their assistance with collecting the fMRI data. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.biopsycho.2023.108713.

References

Adams, R., Rule, N., Franklin, R., Wang, E., Stevenson, M., Yoshikawa, S., et al. (2010). *Cross- cultural Reading the Mind in the Eyes: An fMRI Investigation. Journal of Cognitive Neuroscience*, *22*(1), 97–108. https://doi.org/10.1162/jocn.2009.21187.

Alicke, M.D., & Zell, E. (2009). *Social attractiveness and blame. Journal of Applied Social Psychology*, *39*(9), 2089–2105. https://doi.org/10.1111/j.1559-1816.2009.00517.x.

Baetens, K., Ma, N., Steen, J., & Van Overwalle, F. (2014). *Involvement of the mentalizing network in social and non-social high construal. Social Cognitive and Affective Neuroscience*, *9*(6), 817–824. https://doi.org/10.1093/scan/nst048.

Baron, S.G., Gobbini, M.I., Engell, A.D., & Todorov, A. (2011). *Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. Social Cognitive and Affective Neuroscience*, *6*(5), 572–581. https://doi.org/10.1093/scan/nsq086.

Bellucci, G., Feng, C., Camilleri, J., Eickhoff, S.B., & Krueger, F. (2018). *The role of the anterior insula in social norm compliance and enforcement: Evidence from coordinate-based and functional connectivity meta-analyses. Neuroscience & Biobehavioral Reviews*, *92*, 378–389. https://doi.org/10.1016/j.neubiorev.2018.06.024.

Bersoff, D.M., & Miller, J.G. (1993). *Culture, context, and the development of moral accountability judgments. Developmental Psychology*, *29*(4), 664–676. https://doi.org/10.1037/0012-1649.29.4.664.

Berthoz, S., Armony, J.L., Blair, R.J.R., & Dolan, R.J. (2002). *An fMRI study of intentional and unintentional (embarrassing) violations of social norms. Brain*, *125*(8), 1696–1708. https://doi.org/10.1093/brain/awf190.

Bohner, G., Bless, H., Schwarz, N., & Strack, F. (1988). *What triggers causal attributions? The impact of valence and subjective probability. European Journal of Social Psychology*, *18*(4), 335–345. https://doi.org/10.1002/ejsp.2420180404.

Bostyn, D.H., & Roets, A. (2016). *The morality of action: The asymmetry between judgments of praise and blame in the action–omission effect. Journal of Experimental Social Psychology*, *63*, 19–25. https://doi.org/10.1016/j.jesp.2015.11.005.

Brosch, T., Schiller, D., Mojdehbakhsh, R., Uleman, J.S., & Phelps, E.A. (2013). *Neural mechanisms underlying the integration of situational information into attribution outcomes. Social Cognitive and Affective Neuroscience*, *8*, 640–646. https://doi.org/10.1093/scan/nst019.

Buchtel, E.E., & Norenzayan, A. (2009). Thinking across cultures: Implications for dual processes. In J.St.B.T., Evans, & K., Frankish (Eds.), *Two Minds: Dual Processes and Beyond* (pp. 217–238). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230167.003.0010.

Buckholtz, J., & Marois, R. (2012). *The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. Nature Neuroscience*, *15*, 655–661. https://doi.org/10.1038/nn.3087.

Chang, C., & Glover, G.H. (2009). *Relationship between respiration, end-tidal CO_2, and BOLD signals in resting-state fMRI. Neuroimage*, *47*(4), 1381–1393. https://doi.org/10.1016/j.neuroimage.2009.04.048.

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. https://doi.org/10.1523/JNEUROSCI.2575-09.2009.

Choi, I., & Nisbett, R.E. (1998). *Situational Salience and Cultural Differences in the Correspondence Bias and Actor-Observer Bias. Personality and Social Psychology Bulletin*, *24*, 949–960. https://doi.org/10.1177/0146167298249003.

Choi, I., Dalal, R., Kim-Prieto, C., & Park, H. (2003). *Culture and Judgment of Causal Relevance. Journal of Personality and Social Psychology*, *84*, 46–59. https://doi.org/10.1037/0022-3514.84.1.46.

Ciaramelli, E., Braghittoni, D., & di Pellegrino, G. (2012). *It is the outcome that counts! Damage to the ventromedial prefrontal cortex disrupts the integration of outcome and belief information for moral judgment. Journal of the International Neuropsychological Society*, *18*(6), 962–971. https://doi.org/10.1017/S1355617712000690.

Cohen, M.S. (1997). *Parametric analysis of fMRI data using linear systems methods. NeuroImage*, *6*(2), 93–103. https://doi.org/10.1006/nimg.1997.0278.

Cooper, J.C., Kreps, T.A., Wiebe, T., Pirkl, T., & Knutson, B. (2010). *When giving is good: Ventromedial prefrontal cortex activation for others' intentions. Neuron*, *67*, 511–521. https://doi.org/10.1016/j.neuron.2010.06.030.

Cox, R.W. (1996). *AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014.

Cushman, F. (2008). *Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. Cognition*, *108*, 353–380. https://doi.org/10.1016/j.cognition.2008.03.006.

De Martino, B., O'Doherty, J. p, Ray, D., Bossaerts, P., & Camerer, C. (2013). *In the mind of the market: Theory of mind biases value computation during financial bubbles. Neuron*, *79*, 1222–1231. https://doi.org/10.1016/j.neuron.2013.07.003.

Di, M., Wang, X., Zhao, J., Feng, W., Zhao, J., Nguyen, K., et al. (2021). *Correspondence bias of Chinese and American undergraduates: similar initial trait attributions but different situational corrections. The Journal of General Psychology*, *148*, 105–123. https://doi.org/10.1080/00221309.2019.1703632.

Dien, D.S.F. (1982). *A Chinese perspective on Kohlberg's theory of moral development. Developmental Review*, *2*(4), 331–341. https://doi.org/10.1016/0273-2297(82)90017-X.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). *FMRI item analysis in a theory of mind task. NeuroImage*, *55*, 705–712. https://doi.org/10.1016/j.neuroimage.2010.12.040.

Feinberg, M., Fang, R., Liu, S., & Peng, K. (2019). *A World of Blame to Go Around: Cross-Cultural Determinants of Responsibility and Punishment Judgments. Personality and Social Psychology Bulletin*, *45*, 634–651. https://doi.org/10.1177/0146167218794631.

Genevsky, A., Västfjäll, D., Slovic, P., & Knutson, B. (2013). *Neural underpinnings of the identifiable victim effect: Affect shifts preferences for giving. Journal of Neuroscience*, *33*(43), 17188–17196. https://doi.org/10.1523/JNEUROSCI.2348-13.2013.

Glimcher, P.W., & Fehr, E. (2013). *Neuroeconomics: Decision Making and the Brain* (Eds.). Academic Press,. https://doi.org/10.1016/C2011-0-05512-6.

Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., & Cohen, J.D. (2004). *The neural bases of cognitive conflict and control in moral judgment. Neuron*, *44*(2), 389–400. https://doi.org/10.1016/j.neuron.2004.09.027.

Guimond, S., Begin, G., & Palmer, D.L. (1989). *Education and causal attributions: The development of" person-blame" and" system-blame" ideology. Social Psychology Quarterly*, 126–140.

Han, S., Mao, L., Qin, J., Friederici, A.D., & Ge, J. (2011). *Functional roles and cultural modulations of the medial prefrontal and parietal activity associated with causal*

attribution. *Neuropsychologia*, *49*, 83–91. https://doi.org/10.1016/j.neuropsychologia.2010.11.003.
Heider, F. (1958). Perceiving the other person. In F., Heider (Ed.), *The psychology of interpersonal relations* (pp. 20–58). John Wiley & Sons Inc. https://doi.org/10.1037/10628-002.
Hindriks, F. (2008). *Intentional Action and the Praise-Blame Asymmetry*. *The Philosophical Quarterly*, *58*(233), 630–641. https://doi.org/10.1111/j.1467-9213.2007.551.x.
Houston, D.A. (1990). *Empathy and the self: Cognitive and emotional influences on the evaluation of negative affect in others*. *Journal of personality and Social Psychology*, *59*(5), 859.
Howard, J.A. (1987). *The conceptualization and measurement of attributions*. *Journal of Experimental Social Psychology*, *23*(1), 32–58.
Jenkins, A.C., & Mitchell, J.P. (2010). *Mentalizing under uncertainty: dissociated neural responses to ambiguous and unambiguous mental state inferences*. *Cerebral Cortex*, *20*(2), 404–410. https://doi.org/10.1093/cercor/bhp109.
Jones, E.E., & Harris, V.A. (1967). *The attribution of attitudes*. *Journal of Experimental Social Psychology*, *3*, 1–24. https://doi.org/10.1016/0022-1031(67)90034-0.
Kestemont, J., Vandekerckhove, M., Ma, N., Van Hoeck, N., & Van Overwalle, F. (2013). *Situation and person attributions under spontaneous and intentional instructions: an fMRI study*. *Social Cognitive and Affective Neuroscience*, *8*, 481–493. https://doi.org/10.1093/scan/nss022.
Kim, M.J., Mende-Siedlecki, P., Anzellotti, S., & Young, L. (2021). *Theory of Mind Following the Violation of Strong and Weak Prior Beliefs*. *Cerebral Cortex*, *31*, 884–898. https://doi.org/10.1093/cercor/bhaa263.
Koster-Hale, J., & Saxe, R. (2013). *Theory of mind: a neural prediction problem*. *Neuron*, *79*(5), 836–848.
Kliemann, D., Young, L., Scholz, J., & Saxe, R. (2008). *The influence of prior record on moral judgement*. *Neuropsychologia*, *46*, 2949–2957. https://doi.org/10.1016/j.neuropsychologia.2008.06.010.
Knobe, J. (2003). *Intentional action and side effects in ordinary language*. *Analysis*, *63*(3), 190–194. ⟨https://www.jstor.org/stable/3329308⟩.
Knobe, J. (2010). *Person as scientist, person as moralist*. *Behavioral and Brain Sciences*, *33*, 315–365. https://doi.org/10.1017/S0140525×10000907.
Knowles, E.D., Morris, M.W., Chiu, C., & Hong, Y. (2001). *Culture and the Process of Person Perception: Evidence for Automaticity Among East Asians in Correcting for Situational Influences on Behavior*. *Personality and Social Psychology Bulletin*, *27*, 1344–1356. https://doi.org/10.1177/01461672012710010.
Knutson, B., Rick, S., Wimmer, G.E., Prelec, D., & Loewenstein, G. (2007). *Neural predictors of purchases*. *Neuron*, *53*(1), 147–156. https://doi.org/10.1016/j.neuron.2006.11.010.
Lee, H., Shimizu, Y., Masuda, T., & Uleman, J.S. (2017). *Cultural Differences in Spontaneous Trait and Situation Inferences*. *Journal of Cross-Cultural Psychology*, *48*, 627–643. https://doi.org/10.1177/0022022117699279.
Lieberman, M.D. (2022). *Seeing minds, matter, and meaning: The CEEing model of pre-reflective subjective construal*. *Psychological Review*, *129*(4), 830. https://doi.org/10.1037/rev0000362.
Ma, N., Vandekerckhove, M., & Van Overwalle, F. (2011). *Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: Spontaneous inferences activate only its core areas*. *Social Neuroscience*, *6*, 123–138. https://doi.org/10.1080/17470919.2010.485884.
Markus, H.R., & Kitayama, S. (1991). *Culture and the self: Implications for cognition, emotion, and motivation*. *Psychological Review*, *98*, 224–253. https://doi.org/10.1037/0033-295X.98.2.224.
Markus, H.R., & Kitayama, S. (2003). Models of agency: Sociocultural diversity in the construction of action (of the Nebraska symposium on motivation). In V., Murphy-Berman, & J.J., Berman (Eds.), *Cross-cultural differences in perspectives on the self: Vol. 49* (p. p.18-p74). University of Nebraska Press. (of the Nebraska symposium on motivation).
Masuda, T., & Kitayama, S. (2004). *Perceiver-induced constraint and attitude attribution in Japan and the US: A case for the cultural dependence of the correspondence bias*. *Journal of Experimental Social Psychology*, *40*, 409–416. https://doi.org/10.1016/j.jesp.2003.08.004.
Mesquita, B., & Markus, H.R. (2004). Culture and emotion: Models of agency as sources of cultural variation in emotion. In A.S.R., Manstead, N., Frijda, & A., Fischer (Eds.), *Feelings and Emotions: The Amsterdam Symposium* (pp. 341–358). Cambridge University Press. https://doi.org/10.1017/CBO9780511806582.
Meyer, M.L., Taylor, S.E., & Lieberman, M.D. (2015). *Social working memory and its distinctive link to social cognitive ability: an fMRI study*. *Social Cognitive and Affective Neuroscience*, *10*, 1338–1347. https://doi.org/10.1093/scan/nsv065.
Miller, F.D., Smith, E.R., & Uleman, J. (1981). *Measurement and interpretation of situational and dispositional attributions*. *Journal of Experimental Social Psychology*, *17*(1), 80–95.
Miller, J.G. (1984). *Culture and the development of everyday social explanation*. *Journal of Personality and Social Psychology*, *46*, 961–978. https://doi.org/10.1037/0022-3514.46.5.961.
Miller, J.G., & Luthar, S. (1989). *Issues of interpersonal responsibility and accountability: A comparison of Indians' and Americans' moral judgements*. *Social Cognition*, *7*(3), 237–261. https://doi.org/10.1521/soco.1989.7.3.237.
Mitchell, J.P., Macrae, C.N., & Banaji, M.R. (2006). *Dissociable medial prefrontal contributions to judgments of similar and dissimilar others*. *Neuron*, *50*(4), 655–663. https://doi.org/10.1016/j.neuron.2006.03.040.
Miyamoto, Y., & Kitayama, S. (2002). *Cultural Variation in Correspondence Bias: The Critical Role of Attitude Diagnosticity of Socially Constrained Behavior*. *Journal of Personality and Social Psychology*, *83*, 1239–1248. https://doi.org/10.1037/0022-3514.83.5.1239.
Morris, M.W., & Peng, K. (1994). *Culture and cause: American and Chinese attributions for social and physical events*. *Journal of Personality and Social Psychology*, *67*, 949–971. https://doi.org/10.1037/0022-3514.67.6.949.
Mu, Y., Kitayama, S., Han, S., & Gelfand, M.J. (2015). *How culture gets embrained: Cultural differences in event-related potentials of social norm violations*. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 15348–15353. https://doi.org/10.1073/pnas.1509839112.
Nisbett, R.E., Caputo, C., Legant, P., & Marecek, J. (1973). *Behavior as seen by the actor and as seen by the observer*. *Journal of Personality and Social Psychology*, *27*, 154–164. https://doi.org/10.1037/h0034779.
Nisbett, R.E., Peng, K., Choi, I., & Norenzayan, A. (2001). *Culture and systems of thought: Holistic versus analytic cognition*. *Psychological Review*, *108*, 291–310. https://doi.org/10.1037/0033-295X.108.2.291.
Norenzayan, A., Choi, I., & Nisbett, R.E. (2002). *Cultural Similarities and Differences in Social Inference: Evidence From Behavioral Predictions and Lay Theories of Behavior*. *Personality and Social Psychology Bulletin*, *28*, 109–120. https://doi.org/10.1177/0146167202281010.
Park, B., Tsai, J.L., Chim, L., Blevins, E., & Knutson, B. (2016). *Neural evidence for cultural differences in the valuation of positive facial expressions*. *Social Cognitive and Affective Neuroscience*, *11*(2), 243–252. https://doi.org/10.1093/scan/nsv113.
Park, B., Blevins, E., Knutson, B., & Tsai, J. (2017). *Neurocultural evidence that ideal affect match promotes giving*. *Social Cognitive and Affective Neuroscience*, *12*(7), 1083–1096. https://doi.org/10.1093/scan/nsx047.
Park, B., Fareri, D., Delgado, M., & Young, L. (2021). *The role of right temporoparietal junction in processing social prediction error across relationship contexts*. *Social Cognitive and Affective Neuroscience*, *16*(8), 772–781.
Park, B., Vepachedu, S., Keshava, P., & Minns, S. (2022). *Culture, theory-of-mind, and morality: How independent and interdependent minds make moral judgments*. *Biological Psychology*, 108423. https://doi.org/10.1016/j.biopsycho.2022.108423.
Parkinson, C., Sinnott-Armstrong, W., Koralus, P.E., Mendelovici, A., McGeer, V., & Wheatley, T. (2011). *Is morality unified? Evidence that distinct neural systems underlie moral judgments of harm, dishonesty, and disgust*. *Journal of Cognitive Neuroscience*, *23*(10), 3162–3180. https://doi.org/10.1162/jocn_a_00017.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2019). nlme: Linear and nonlinear mixed effects models. R package version 3.1–143. https://CRAN.R-project.org/package=nlme.
Pizarro, D., Uhlmann, E., & Salovey, P. (2003). *Asymmetry in judgments of moral blame and praise: The role of perceived metadesires*. *Psychological Science*, *14*(3), 267–272. https://doi.org/10.1111/1467-9280.03433.
Reeder, G.D., & Spores, J.M. (1983). *The attribution of morality*. *Journal of Personality and Social Psychology*, *44*, 736–745. https://doi.org/10.1037/0022-3514.44.4.736.
Reeder, G.D., Vonk, R., Ronk, M.J., Ham, J., & Lawrence, M. (2004). *Dispositional attribution: multiple inferences about motive-related traits*. *Journal of personality and Social Psychology*, *86*(4), 530.
Ross, L. (1977). *The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process*. *Advances in Experimental Social Psychology*, *10*, 173–220. https://doi.org/10.1016/S0065-2601(08)60357-3.
Saxe, R., & Kanwisher, N. (2003). *People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind."*. *NeuroImage*, *19*, 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1.
Scholer, A.A., & Higgins, E.T. (2010). Regulatory focus in a demanding world. In R.H., Hoyle (Ed.), *Handbook of personality and self-regulation* (pp. 291–314). Wiley Blackwell. https://doi.org/10.1002/9781444318111.ch13.
Shultz, T.R., Schleifer, M., & Altman, I. (1981). *Judgments of causation, responsibility, and punishment in cases of harm-doing*. *Canadian Journal of Behavioural Science*, *13*, 238–253. https://doi.org/10.1037/h0081183.
Siegel, J.Z., Crockett, M.J., & Dolan, R.J. (2017). *Inferences about moral character moderate the impact of consequences on blame and praise*. *Cognition*, *167*, 201–211. https://doi.org/10.1016/j.cognition.2017.05.004.
Singelis, T.M. (1994). *The measurement of independent and interdependent self-construals*. *Personality and Social Psychology Bulletin*, *20*(5), 580–591. https://doi.org/10.1177/0146167294205014.
Stallen, M., Borg, N., & Knutson, B. (2021). *Brain Activity Foreshadows Stock Price Dynamics*. *The Journal of Neuroscience*, *41*, 3266–3274. https://doi.org/10.1523/JNEUROSCI.1727-20.2021.
Uhlmann, E.L., Pizarro, D.A., & Diermeier, D. (2015). *A Person-Centered Approach to Moral Judgment*. *Perspectives on Psychological Science*, *10*, 72–81. https://doi.org/10.1177/1745691614556679.
Van Overwalle, F. (2009). *Social cognition and the brain: a meta-analysis*. *Human brain Mapping*, *30*(3), 829–858. https://doi.org/10.1002/hbm.20547.
Van Overwalle, F., & Baetens, K. (2009). *Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis*. *Neuroimage*, *48*(3), 564–584. https://doi.org/10.1016/j.neuroimage.2009.06.009.
Van Overwalle, F., & Vandekerckhove, M. (2013). *Implicit and explicit social mentalizing: dual processes driven by a shared neural network*. *Frontiers in Human Neuroscience*, *7*(560). https://doi.org/10.3389/fnhum.2013.00560.
Weiner, B. (1995). *Judgments of Responsibility: A Foundation for a Theory of Social Conduct*. New York: Guilford Press,.
Wong-On-Wing, B., & Lui, G. (2007). *Culture, implicit theories, and the attribution of morality*. *Behavioral Research in Accounting*, *19*, 231–246. https://doi.org/10.2308/bria.2007.19.1.231.
Wong-On-Wing, B., & Lui, G. (2013). *Beyond cultural values: An implicit theory approach to cross-cultural research in accounting ethics*. *Behavioral Research in Accounting*, *25*, 15–36. https://doi.org/10.2308/bria-10315.
Wu, C.C., Samanez-Larkin, G.R., Katovich, K., & Knutson, B. (2014). *Affective traits link to reliable neural markers of incentive anticipation*. *NeuroImage*, *84*, 279–289. https://doi.org/10.1016/j.neuroimage.2013.08.055.

Yamada, M., Camerer, C.F., Fujie, S., Kato, M., Matsuda, T., Takano, H., & Takahashi, H. (2012). *Neural circuits in the brain that are activated when mitigating criminal sentences*. *Nature Communications*, *3*(1), 759. https://doi.org/10.1038/ncomms1757.

Yoder, K.J., & Decety, J. (2014). *The good, the bad, and the just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others*. *Journal of Neuroscience*, *34*(12), 4161–4166. https://doi.org/10.1523/JNEUROSCI.4648-13.2014.

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). *The neural basis of the interaction between theory of mind and moral judgment. Proceedings of the National Academy of Sciences- PNAS*, *104*(20), 8235–8240. https://doi.org/10.1073/pnas.0701408104.

Young, L., & Koenigs, M. (2007). *Investigating emotion in moral cognition: a review of evidence from functional neuroimaging and neuropsychology*. *British Medical Bulletin*, *84*, 69–79. https://doi.org/10.1093/bmb/ldm031.

Young, L., Nichols, S., & Saxe, R. (2010). *Investigating the neural and cognitive basis of moral luck: It's not what you do but what you know*. *Review of Philosophy and Psychology*, *1*(3), 333–349. https://doi.org/10.1007/s13164-010-0027-y.

Young, L., & Saxe, R. (2009). *Innocent intentions: A correlation between forgiveness for accidental harm and neural activity*. *Neuropsychologia*, *47*(10), 2065–2072. https://doi.org/10.1016/j.neuropsychologia.2009.03.020.

Young, L., Scholz, J., & Saxe, R. (2011). *Neural evidence for "intuitive prosecution": The use of mental state information for negative moral verdicts*. *Social Neuroscience*, *6*(3), 302–315. https://doi.org/10.1080/17470919.2010.529712.