Parsing the components of forgiveness: Psychological and neural mechanisms

Melike M. Fourie[1*], Ruud Hortensius[2] & Jean Decety[3]

[1] Studies in Historical Trauma and Transformation, Stellenbosch University, Stellenbosch, South Africa

[2]Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, United Kingdom

[3]Department of Psychology and Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, United States

Accepted version of Fourie, M. M., Hortensius, R. and Decety, J. (2020) Parsing the components of forgiveness: Psychological and neural mechanisms. Neuroscience and Biobehavioral Reviews, 112, pp. 437-451. (doi: 10.1016/j.neubiorev.2020.02.020) (PMID:32088347)

*Corresponding author:
Studies in Historical Trauma and Transformation
Stellenbosch University
Stellenbosch, 7600
South Africa
E-mail: melikef@sun.ac.za

ABSTRACT

Forgiveness—a shift in motivation away from retaliation and avoidance towards increased goodwill for the perceived wrongdoer—plays a vital role in restoring social relationships, and positively impacts personal wellbeing and society at large. Parsing the psychological and neurobiological mechanisms of forgiveness contributes theoretical clarity, yet has remained an outstanding challenge because of conceptual and methodological difficulties in the field. Here, we critically examine the neuroscientific evidence in support of a theoretical framework which accounts for the proximate mechanisms underlying forgiveness. Specifically, we integrate empirical evidence from social psychology and neuroscience to propose that forgiveness relies on three distinct and interacting psychological macro-components: cognitive control, perspective taking, and social valuation. The implication of the lateral prefrontal cortex, temporoparietal junction, and ventromedial prefrontal cortex, respectively, is discussed in the brain networks subserving these distinct component processes. Finally, we outline some caveats that limit the translational value of existing social neuroscience research and provide directions for future research to advance the field of forgiveness.

Keywords: forgiveness; fMRI; cognitive control; perspective taking; social valuation; cognitive neuroscience

1. INTRODUCTION

Forgiveness can be a powerful means to heal relationships and restore personal well-being and health after a transgression (Bono et al., 2008; Toussaint et al., 2015). In addition, it is a significant mediator of social change and reconciliation in society at large (Gobodo-Madikizela, 2015), helping to transform conflicts worldwide from Northern Ireland to Sierra Leone. The post-apartheid South African Truth and Reconciliation Commission (TRC) is a case in point—with its focus on restorative rather than retributive justice, on forgiveness rather than vengeance, it has been argued to play a critical role in promoting national reconciliation amid gross civil discord (Boraine et al., 1997). Greater forgiveness amongst victims of human rights abuses is also associated with reduced anger and improved mental health (Kaminer et al., 2001; Stein et al., 2008). Yet, despite its potential benefits, forgiving can be costly in terms of self-interest (Exline and Baumeister, 2000). Thus, any theory of forgiveness needs to factor in how computations regarding potential future gains of interaction with the transgressor versus likelihood of future harms affect decision-making.

From an evolutionary perspective, forgiveness might have evolved as a second-order adaptation to revenge to deal with exploitation by conspecifics (McCullough, 2008; McCullough et al., 2013). Some non-human primate species engage in post-conflict affiliation between former opponents of a fight and bystanders (Aureli et al., 1997; de Waal and Ren, 1988). This behavior, referred to as “reconciliation” or “appeasement,” has been posited to play a role in restoring valuable relationships. However, it is not clear that high cognitive processes underlie such social behavior in non-human primates. One study, which combined computational modeling and empirical data, investigated the minimum cognitive requirements for post-conflict affiliation in monkeys (Puga-Gonzalez et al., 2014). The authors found evidence for four categories of post-conflict affiliation in the model and in the empirical data, and explained how these patterns of behavior emerge from the combination of a weak hierarchy, social facilitation, risk-sensitive aggression, interactions with partners close-by, and grooming as tension-reduction mechanism.

Because revenge uses retaliation to deter future exploitation by the wrongdoer, it often comes at a personal (e.g., feelings of anger and resentment, rumination) and social (e.g., jeopardized future gains and/or escalating cycles of counter-retaliations) cost (Carlsmith et al., 2008; Noor, 2016). By comparison, forgiveness presents the individual with an alternative strategy to secure personal benefits—by inhibiting revenge, decreasing avoidant motivations, and facilitating reconciliatory behavior, it increases the individual's social fitness. Notably, the goal of such shifts in interpersonal motivation is to secure the long-term advantages of continued cooperative interaction with the transgressor (such as resources and coalitional support), provided that future exploitation does not recur (McCullough et al., 2013; Petersen et al., 2012).

Because forgiveness plays a quintessential role in social interaction and facilitates conflict resolution and cooperation within societies, a better understanding of its psychological mechanisms and their neural underpinnings is important not only to provide theoretical clarity, but also to inform its therapeutic uses. Cognitive neuroscientific investigations of forgiveness are fairly recent, however. For years, longstanding definitional controversies and lack of empirical integration has characterized the field of forgiveness (Fehr et al., 2010; Miceli and Castelfranchi, 2011; Riek and Mania, 2012; Worthington et al., 2007). Indeed, some scholars have argued that forgiveness is undertheorized (McCullough, 2008; McCullough et al., 2013). Without a clear understanding of forgiveness at the psychological level of analysis, one cannot begin to elucidate its functional architecture at the neural level (Gillihan and Farah, 2005).

The understanding of forgiveness as a construct amenable to scientific enquiry has gained considerable traction in recent years. We believe the time is now ripe for a multi-level synthesis that integrates psychological and neurobiological accounts (Krakauer et al., 2017). Identifying brain mechanisms and networks that lie at the core of forgiveness can advance the field in meaningful ways. Notably, it can contribute to discerning the underlying information-processing mechanisms, thus informing theoretical models of forgiveness. For example, mapping brain activity and connectivity at different time points and during different tasks would allow separate processes and different stages underlying forgiveness to be distinguished. In this regard, fMRI is uniquely poised to identify and track complex internal states in real-time (Huettel, 2015).

Furthermore, a better understanding of the neural mechanisms involved in forgiveness would allow vigorous testing of psychological hypotheses, which is not possible when using solely behavioral measures (Amodio, 2010). Lastly, predictive markers of forgiveness processes (such as patterns of activity analyzed with multivariate approaches) can be used for subtyping/diagnosing of individuals, and as potential biological targets for intervention, such as biofeedback-based training to enhance forgiveness (e.g., Moll et al., 2014). Such neuroscientific approaches can also be used to predict forgiveness in everyday life. For example, a recent study demonstrated that machine-learning regression techniques can distinguish between self-centered distress and other-centered empathic concern when participants listened to biographies describing stories of human suffering, and that only the latter activation patterns predict trial-by-trial donation amounts (Ashar et al., 2017).

Here we review the small but steadily growing social neuroscience literature that examined the neural underpinnings of forgiveness. As a point of departure, we discuss theory and research from social psychology to demonstrate the multifaceted nature of forgiveness (Frank and Badre, 2015), arguing that three distinct but interacting psychological macro-component processes, namely cognitive control, perspective taking, and social valuation, can be distinguished. The significance of social valuation, in particular, as the process by which potential future gains versus losses are calculated, extends Billingsley and Losin's (2017) comprehensive review. The primary thesis of that review, drawing on evolutionary psychology, is that the forgiveness system is interconnected with, and inhibits the phylogenetically older revenge/reward system. By the current account, forgiveness also necessitates a dynamic interplay between neocortical component processes to allow for flexibility in adaptively responding to transgressions. For example, recent research suggests that the cognitive control and social valuation systems are functionally interconnected, resulting in context-dependent valuation of choices (Hare et al., 2009; Rudorf and Hare, 2014). Moreover, changes in perspective taking (rather than cognitive control) can underlie reduced retaliation motivation in certain contexts (Baumgartner et al., 2013).

Below we first review the underlying neural architecture that supports each of the macro-component processes, as well as the reasons for the presence/absence thereof in current neuroimaging work on forgiveness. This qualitative analysis is followed by an exploratory meta-analysis of activation maps relevant to forgiveness. We then outline a provisional neurocognitive framework articulating the way forward with neuroimaging research—in the process highlighting caveats of previous work and providing potential directions for future research to advance the neuroscientific investigation of forgiveness.

2. PSYCHOLOGICAL COMPONENT PROCESSES

Forgiveness literally means letting go of something (refraining from retaliation), and offering an altruistic or undeserved gift (acting prosocially), despite the offender's hurtful behavior (Worthington et al., 2000). A third attribute of forgiveness is its temporal unfolding—it usually takes time to shift from an initial negative/unforgiving response to a more positive/forgiving response (McCullough et al., 2003).

Apparent in the above description, is that forgiveness necessitates important internal motivational changes. To forgive, one typically needs to overcome strong negative emotions, ruminative thoughts, or even vengeful impulses to punish the transgressor, and instead cultivate more positive feelings and concern for that person (Beyens et al., 2015; McCullough et al., 1997; Worthington et al., 2007). This kind of goal-directed, effortful emotion regulation and inhibition may be regarded as a function of cognitive control (Ochsner and Gross, 2005).

Various lines of indirect evidence support the link between forgiveness and cognitive control. Notably, superior cognitive control abilities have consistently been associated with reduced occurrence of anger and aggression (Denson et al., 2012). For instance, people who score higher compared to those who score lower on dispositional measures of cognitive control appear to be less likely to aggress against wrongdoers (Chester et al., 2014), and are more likely to accommodate a partner who has transgressed and/or inhibit destructive impulses toward that partner (Finkel and Campbell, 2001). One study demonstrated a more direct association between cognitive control processes and one's propensity to forgive (Pronk et al., 2010). In a series of four studies (including longitudinal data), the authors found that individual differences in cognitive control predicted both dispositional and actual forgiveness. Importantly, their data suggest that cognitive control facilitates forgiveness by decreasing ruminative thoughts, particularly for severe offences. Enhanced recruitment of cognitive control is furthermore considered to inhibit socially inappropriate retaliatory aggression (Wilkowski and Robinson, 2010). For example, greater cognitive control predicts forgiveness of provocations and subsequent reductions in anger and aggression (Wilkowski et al., 2010).

Intriguingly, recent evidence indicates that the inhibitory control mechanism involved in forgiveness also facilitates motivated or intentional forgetting, which prevents unwanted memories from re-entering conscious awareness (Noreen et al., 2014). Hence, forgiven offenses may result in less rumination and greater psychological distance from the event compared to unforgiven offenses (McCullough et al., 2007). While it is beyond the scope of the present review to elaborate on forgiveness's association with better health outcomes, the stress-reducing role of cognitive control is likely central in this regard (Gabrys et al., 2018; Marks et al., 2013). Indeed, forgiveness therapy's focus is foremost on clients overcoming emotions of resentment and bitterness over betrayals in the process of granting forgiveness (Enright and Fitzgibbons, 2015).

A second psychological component of forgiveness is perspective taking, imagining how someone is affected by his or her situation without confusion between the feelings experienced by the self versus feelings experienced by the other person (Ruby and Decety, 2004). A substantial body of evidence documents the effectiveness of perspective taking as a powerful means to elicit empathy and concern for others (Batson and Ahmad, 2009; Decety and Jackson, 2004; Todd and Galinsky, 2014). Recent evidence also points to its importance in explaining individual differences in justice sensitivity for others (Decety and Yoder, 2016). Notably, perspective taking seems crucial for forgiveness, because it involves temporarily suspending one's own point-of-view and feelings in an attempt to adopt and understand those of the wrongdoer. Whereas much previous work has focused broadly on empathy as a determinant of forgiveness (Macaskill et al., 2002; McCullough et al., 1997; Sandage and Worthington, 2010), there are two paths to share and understand another's emotional state: an affective sharing mechanism, and a more cognitively effortful process relying on mentalising capacities (Shamay-Tsoory et al., 2009; Zaki and Ochsner, 2012). We believe it is this latter, cognitive-driven perspective taking that plays a major role in bringing about changes in the way we see a transgressor, which, in turn, fosters concern (Decety and Cowell, 2015). This is particularly true in the absence of an apology or perceived remorse, when affective empathy may be less consequential (Davis and Gold, 2011).

Forgiveness necessitates some contextualization or reframing to understand the offender's intentions and behavior (North, 1998). For example, people are much more likely to excuse (forgive) a harmful act that was committed accidentally than one that was committed intentionally (Cushman, 2008). Enhanced perspective taking might also lead one to consider the circumstances that led to the offender's behavior, or to reflect upon attributes that are shared by oneself and them (Davis et al., 1996; Galinsky et al., 2005). Previous studies have demonstrated that enhanced perspective taking, both in terms of reflecting upon one's own previous transgressions, as well as adopting the transgressor's perspective, facilitates forgiveness (Exline et al., 2008; Steindl and Jonas, 2012; Takaku, 2001). In addition, greater disposition in perspective taking has been associated with lower incidence of punishment behavior and higher incidence of forgiveness toward transgressors (Will et al., 2015).

Finally, forgiveness entails a third process, social valuation, that critically affects the decision to forgive. Following an offense or social norm violation, one has the decision to forgive or punish the offender (McCullough et al., 1997). Social valuation can be described as the process whereby social information, including the outcomes of the actions of the wrongdoer (as well as the self, in the case of unforgiveness), is assigned value and hence assessed for its forgivability and appropriateness. For example, victims are more forgiving following costly compared to non-costly apologies, as the former is perceived to be more sincere, thereby reducing the risk of future exploitation (Ohtsubo and Watanabe, 2009; Ohtsubo et al., 2012).

Previous studies have found that various situational factors affect a person's decision to forgive, from the attitude, relationship value and exploitation risk of the perpetrator, to the severity of the transgression, and the presence and nature of an apology and/or repentance (Bennett and Earwaker, 1994; Berndsen et al., 2015; Burnette et al., 2012; Fehr et al., 2010; Hayashi et al., 2010; Ohbuchi et al., 1989). Social valuation is the computational process whereby the motivational significance of these factors is weighed and assessed, influencing the decision to ultimately forgive. It should be noted that clinical psychology literature distinguishes between decisional and emotional forgiveness (Worthington et al., 2007). Whereas the former involves reducing unforgiveness and controlling one's behavior, the latter type of forgiveness is more multifaceted, involving also setting aside resentment-based emotions and cultivating more positive, other-oriented emotions. In terms of the three psychological components described here, decisional forgiveness might thus rely strongly on processes of social valuation (i.e., cost-benefit analysis), whereas emotional (true) forgiveness might require the interaction of all three component processes.

Of interest is that there appears to be substantial overlap between judgments concerning the morality (rightness vs. wrongness) and forgivability of an action, such that the former is thought to influence the latter (Farrow and Woodruff, 2005). Although these processes are deeply entwined, they should not be equated. That is, forgivability judgments, particularly in close interpersonal relationships, are less objective than moral judgments, which are guided by more cross-culturally invariant moral standards (Tangney et al., 2007; Wohl and Reeder, 2004). For example, deciding to forgive a partner who made you feel excluded/unvalued at a social event is unlikely to involve the same moral judgment processes typically evoked by a moral dilemma. Moreover, one might decide to forgive someone despite the moral wrongness of his/her behavior—this is where perspective taking is paramount (Rogé and Mullet, 2011; Young and Saxe, 2009).

Taken together, forgiveness requires important shifts in motivation and emotion toward the wrongdoer (cognitive control), understanding the wrongdoer's intentions and emotional state (perspective taking), as well as judgments concerning the appropriateness or value of forgiveness in the specific context (social valuation).

3. NEURAL UNDERPINNINGS

As proposed above, the decision to forgive encompasses at least three psychological macro-processes that are supported by distinct brain networks involved in social cognition. The aim of the present review is to determine whether these proposed constructs could also be distinguished in neuroimaging work examining forgiveness to begin specifying a framework for future research. Below we review brain regions and networks supporting these component processes (Table 1, Fig. 1), followed by a qualitative discussion of the presence/absence thereof in previous neuroimaging work on forgiveness. In addition, to provide preliminary support for our framework, we conducted an exploratory quantitative meta-analysis incorporating a subset of neuroimaging studies on forgiveness that met inclusion criteria (see Supplementary Material). Because the number of studies included in this analysis is small (*N* = 8), with significant variation in the methodologies employed and contrasts performed (David et al., 2013), these results should be interpreted with caution.

Table 1 Forgiveness Component Processes

Psychological Macro-Process	**Description**	**Key Brain Areas**
Cognitive control	Emotion regulation Cognitive conflict Countering response tendencies Reappraisal processes	dlPFC, vlPFC, dACC
Perspective taking	Mentalizing Cognitive empathy Mindreading Third-person perspective	TPJ, mPFC, precuneus, PCC
Social valuation	Social decision making Cost/benefit analysis Relational and socio-moral constraints Value computations	vmPFC/OFC

Dorsal anterior cingulate cortex (dACC), dorsolateral prefrontal cortex (dlPFC), medial prefrontal cortex (mPFC), orbitofrontal cortex (OFC), posterior cingulate cortex (PCC), temporoparietal junction (TPJ), ventrolateral prefrontal cortex (vlPFC), ventromedial prefrontal cortex (vmPFC).

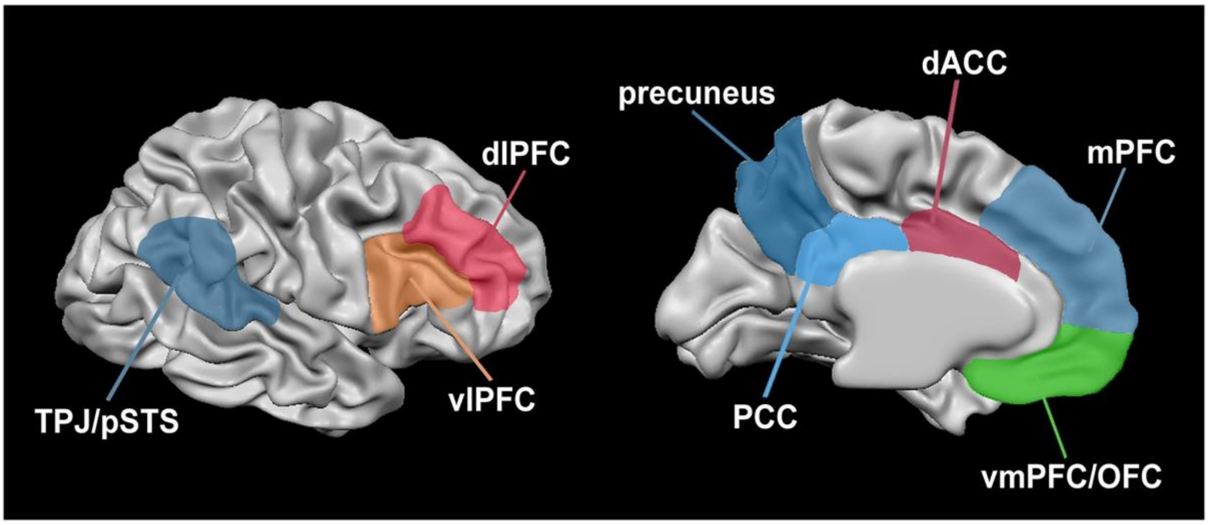

Fig. 1. Network of interconnected regions implicated in forgiveness. Regions involved in cognitive control are highlighted in red/orange; regions involved in perspective taking are highlighted in blue; and the region involved in social valuation is highlighted in green. dACC = dorsal anterior cingulate cortex; dlPFC = dorsolateral prefrontal cortex; mPFC = medial prefrontal cortex; OFC = orbitofrontal cortex; PCC = posterior cingulate cortex; pSTS = posterior superior temporal sulcus; TPJ = temporoparietal junction; vlPFC = ventrolateral prefrontal cortex; vmPFC = ventromedial prefrontal cortex.

3.1 Systematic literature search

A systematic literature search was conducted using the following strategy: First, we performed standard key-word searches in the databases PubMed, ISI Web of Science, and PsychInfo for studies published until Dec 2018. Our search terms included one of the key-words 'neuroimaging' OR 'functional magnetic resonance imaging (fMRI)' OR 'positron emission tomography (PET)' OR 'voxel-based morphometry (VBM)' AND 'forgiveness'. Twenty studies were identified in this way, of which 14 were original studies that matched our criteria (see below).[1] Second, we updated the literature sample by reviewing the reference lists of relevant articles found in step one, as well as several review articles (Billingsley and Losin, 2017; Farrow and Woodruff, 2005; Fatfouta et al., 2013), which yielded 1 more study. Document types were limited to peer-reviewed journal articles, thus conference abstracts or presentations were excluded.

We constrained our review to only include studies if they reported the direct association between actual forgiveness, forgiveness judgments, or the tendency to forgive, and brain structure and/or function. Specifically, studies were included if forgiveness was inferred based on participant self-report data or observational techniques, or if forgiveness was inferred by the authors based on theoretical reasoning. Finally, studies were excluded if the authors did not explicitly investigate forgiveness or conducted analyses to do so (e.g., studies investigating economic decision-making or moral judgment processes more broadly). To keep the scope of this mini review as comprehensive as possible, we have included studies that measured responses from either or both healthy adult and clinical populations, as well as adolescents. In addition, because of the exploratory nature of the review, we have included the following: studies whose primary analyses included voxel-based morphometry (VBM) or

[1] The full texts of all retrieved studies were reviewed, unless the abstract indicated that it was not original research.

resting state analysis, studies whose main findings relied on region of interest analyses (ROI), and studies that reported associations between forgiveness and functional connectivity analysis.

Because of discrepancies and inconsistencies in the manner that similar brain regions were labeled across studies, regions have been checked and relabeled according to our areas of interest presented in Table 1 to present a more unified scheme. For example, significant activation reported in the inferior parietal lobe or angular gyrus was relabeled to temporoparietal junction (TPJ) if deemed appropriate after inspecting peak voxel locations (Schurz et al., 2014).

3.2 Methodologies of selected studies

Our literature search identified 15 neuroimaging studies on forgiveness published between 2001 and 2018: whereas only 4 neuroscientific studies on forgiveness were published in the 12 years spanning 2001–2012, 11 neuroscientific studies on forgiveness have been published in the last 7 years, suggesting growing scientific interest in the topic. A diverse array of methods was employed in the studies retrieved.

A number of studies explored forgivability judgments in response to scenario-based descriptions (Farrow et al., 2005, 2001; Hayashi et al., 2010; Patil et al., 2017; Young and Saxe, 2009). Three studies explored imagined or direct forgiveness: one looked at re-appraisal driven forgiveness in response to hypothetical, personally-distressing events (Ricciardi et al., 2013), whereas the other two examined active forgiveness following an apology (or no apology) in response to an ambiguous offense (Strang et al., 2014) or hypothetical transgression (Ohtsubo et al., 2018). Five studies used economic games to measure forgiveness. In these studies, forgiveness was operationalized as the acceptance of unfair offers from close others or strangers during an Ultimatum Game (Fatfouta et al., 2016);

or as sharing equally in modified Dictator Games, and thus refraining from punishing people who had previously excluded them socially (Will et al., 2015, 2016),[2] or had treated them unfairly during an Ultimatum Game (Brüne et al., 2013). One study looked at the role of forgiveness following financial compensation for unequal resource sharing (during a Dictator Game) in restoring trust (Haesevoets et al., 2018). Finally, two studies explored associations between individual differences in the tendency to forgive and (i) resting state brain activity (Li and Lu, 2017) and (ii) neuroanatomical differences in gray and white matter volume using VBM (Li et al., 2017).

Table 2 summarizes the main characteristics of each study in terms of imaging modality, population, methodology, and results. Based on the methodologies described above, we have also categorized studies into task groups that made use of relatively similar stimuli and task instructions. Because our primary interest concerned structural areas and neural activation responses associated with forgiveness, we only report on these results for each study. Authors MF and RH independently extracted data that were subsequently cross-matched to ensure consistency and accuracy.

3.3 Cognitive control

Regions of the prefrontal cortex, including the dorsolateral prefrontal cortex (dlPFC), ventrolateral prefrontal cortex (vlPFC), and dorsal anterior cingulate cortex (dACC) have long been implicated in cognitive control (Egner and Hirsch, 2005; Miller and Cohen, 2001; Ridderinkhof et al., 2004). These areas are part of a circuit involved in top-down regulatory control which mediates emotion regulation in a goal-directed manner: modification of existing or initiating new emotional responses. Cognitive reappraisal, the most commonly

[2] Another study used a similar paradigm, but in their neuroimaging analyses of responses to excluders during the modified Dictator Game did not distinguish between equal sharing (forgiveness) and unequal sharing (punishment), and were thus not included in this review (Moor et al., 2012).

employed strategy, for example, involves thinking about emotionally charged situations or stimuli in a way that lessens the emotional impact thereof (Ochsner and Gross, 2008). It has been argued that two types of control processes can be distinguished because of differences in functional connectivity patterns to emotion-generative brain regions, such as the amygdala (Aron et al., 2007; Ochsner and Gross, 2005; Ochsner et al., 2012; Wager et al., 2008). Accordingly, the first type of control process involves regions of the vlPFC which, because of its direct functional connectivity to subcortical emotion systems, might be involved in directly altering emotional associations (e.g., reversal learning) and response inhibition. By contrast, dorsolateral and posterior portions of the PFC, implicated in working memory and selective attention, support explicit reappraisal of situations and thus reflect a more general, indirect, mechanism to alter emotional associations.

The dACC'sfunction in cognitive control appears to involve conflict detection and performance monitoring (Cole et al., 2009). The dACC has consistently been implicated in monitoring response tendencies for competition, in overriding prepotent responses, and in signaling the need for enhanced cognitive control within the dlPFC and related prefrontal control regions in conflict situations (Botvinick et al., 2004; Gabay et al., 2014; Kerns et al., 2004). The dACC may also signal internal conflict, for example, in response to the undesired activation of racial stereotypes (Fourie et al., 2014), or when one acts in a prosocial manner toward wrongdoers (Moor et al., 2012).

Together with the dACC, prefrontal cognitive control areas are thus crucial in countering one's own response tendencies and in using cognitive strategies to regulate emotions (e.g., through reappraisal processes). For example, lateral prefrontal areas are activated when people overcome a selfish impulse (Steinbeis et al., 2012), when they regulate racial bias (Richeson et al., 2003), when they reappraise an emotive situation in a positive manner (Drabant et al., 2009), and when they regulate strong negative affect (Sebastian et al., 2011). Importantly, enhanced recruitment of cognitive control is also pivotal when dealing with a transgression and experiencing conflicting desires (e.g., emotional "punish" vs cognitive "forgive"). Arguably the most direct evidence supporting the importance of cognitive control for forgiveness decisions comes from a recent study where cognitive control was manipulated in real time through inhibitory continuous theta-burst stimulation (cTBS) of the dlPFC (Maier et al., 2018). Following cTBS (versus placebo), participants displayed significantly more revenge than forgiveness behavior in a dictator game against previously unfair opponents.

In the current review, seven functional neuroimaging studies found increased activation in the dlPFC, vlPFC and/or dACC to be associated with forgiveness (Table 2).[3] In addition, a VBM study conducted by Li and colleagues (2017) reported a significant positive correlation between participants' dispositional tendency to forgive and gray matter volume in the dlPFC. The authors argued that this local increase in gray matter may facilitate regulation of prepotent responses to retaliate against wrongdoers in those with higher trait forgiveness. Consistent with this interpretation, Will and colleagues (2016) found that chronically rejected compared to stable adolescents require enhanced recruitment of the lateral PFC during forgiveness, as they may suffer greater difficulties to control retaliatory responses than stable adolescents.

Consistent with the reasoning above, most studies that involved the generation of strong negative affect as a result of personal harm [e.g., social exclusion (Will et al., 2015, 2016), being treated unfairly (Brüne et al., 2013; Fatfouta et al., 2016; Haesevoets et al., 2018), or suffering a personally hurtful, albeit hypothetical, event (Ohtsubo et al., 2018;

[3] The study by Fatfouta et al. (2016) detected dlPFC activity during unfair offers, however, it is not possible to determine how often these unfair offers were accepted (and thus forgiveness presumably occurred).

Ricciardi et al., 2013)] were associated with significant dlPFC activation.[4] In each case, it could be argued that a prepotent response to retaliate had to be controlled in order to forgive. By contrast, those studies that involved forgivability judgments of scenario-based vignettes unrelated to the self (e.g., Farrow et al., 2001; Young and Saxe, 2009), or forgiving of an ambiguous offense (Strang et al., 2014) typically did not elicit dlPFC activity. Interestingly, a recent study by Fatfouta et al. (2016), found reduced functional connectivity between the medial prefrontal cortex (mPFC) and dACC to be associated with increased acceptance of unfair offers from close others (forgiveness). Although the dACC may thus be critical in alerting to conflicting tendencies, these results suggest that less available information about conflict is more conducive to forgiving.

[4] Previous imaging studies of the Ultimatum Game also observed heightened dlPFC/vlPFC compared to insula activation when unfair monetary offers were accepted, presumably an indication that prepotent emotional responses had to be controlled to resist unfairness (Sanfey et al., 2003; Tabibnia et al., 2008).

Table 2

Neuroimaging Studies Involving Forgiveness

Study	Imaging Modality and Population	Constructs Investigated	Behavioral Task	Cognitive Control			Perspective Taking				Social Valuation	Other Regions
				dlPFC	**vlPFC**	**dACC**	**mPFC**	**TPJ/ pSTS**	**PCU**	**PCC**	**vmPFC / OFC**	
Social Scenarios Involving Others												
(Farrow et al., 2001)	fMRI 10 healthy individuals	Forgivability judgments	Participants selected the more forgivable explanation for the situation described.						V	V	V	superior frontal gyrus
(Farrow et al., 2005)	fMRI 13 patients with PTSD	Forgivability judgments	Participants selected the more forgivable explanation for the situation described (both pre and post cognitive behavioral therapy).					V	V	V	V (frontal pole)	L middle frontal gyrus, posterior MTG
(Young & Saxe, 2009)	fMRI 15 healthy individuals	Moral judgments	Participants judged the blameworthiness of agents on a 4-point scale. (outcome: harm vs. no harm; belief of agent: negative vs. neutral).					V[a]			V[a]	
(Hayashi et al., 2010)	PET 12 healthy individuals	Forgivability judgments	Participants judged the forgivability of each scenario on a 4-point scale. (transgression: serious vs. minor; perpetrator: honest vs. dishonest)								V	

(Patil et al., 2017)	VBM and fMRI 50 healthy individuals	Moral judgments	Participants judged the blameworthiness of agents (outcome: negative vs. neutral; belief of agent: negative vs. neutral)					V (L STS)		
Economic Decision-Making										
(Brüne et al., 2013)	fMRI 29 healthy individuals	Treatment of opponent	Ultimatum Game followed by Dictator Game Participants accepted or rejected unfamiliar proposer’s fair or unfair offers, and then proposed fair or unfair offers to the previous opponent.	V[b]						
(Will et al., 2015)	fMRI 26 healthy individuals	Treatment of opponent	Cyberball exclusion game followed by Dictator Game Participants proposed fair or unfair offers to opponents who had previously either excluded or included them.	V	V	V	V	V		bilateral insula
(Will et al., 2016)	fMRI 25 healthy adolescents and 18 chronically peer rejected adolescents	Treatment of opponent	Cyberball exclusion game followed by Dictator Game Participants proposed fair or unfair offers to opponents who had previously either excluded or included them.	V[c] (lateral PFC)		V	V			dorsal striatum

(Fatfouta et al., 2016)	fMRI 23 healthy individuals in romantic relationships	Treatment of opponent	Ultimatum Game Participants accepted or rejected fair or unfair offers by their romantic partner of unknown others.			V[d]	V					right insula, OCC
(Haesevoets et al., 2018)	fMRI 27 healthy individuals	Receipt of financial compensation	Dictator Game followed by financial compensation (or not) that restored equality Participants rated trust for the allocator on a 4-point scale.		V			V				superior frontal gyrus, bilateral insula, parietal cortex
Imagined or Direct Forgiveness												
(Ricciardi et al., 2013)	fMRI 10 healthy individuals	Reappraisal-driven forgiveness	Participants read narrative scenarios of emotionally hurtful personal events, followed by the indication to forgive or harbor a grudge toward the imagined offender	V		V		V	V	V	V (frontal pole)	bilateral MTG, OCC
(Strang et al., 2014)	fMRI 32 healthy individuals	Active forgiveness	Participants decided to forgive/not forgive the other player for a wrong response following an apology/no apology in a monetary game. The intention of the other player was not known.					V[b]				

(Ohtsubo et al., 2018)	fMRI 37 healthy individuals	Active forgiveness	Participants read scenarios of a friend committing mild interpersonal transgressions, followed by a costly/non-costly apology (or non-apology). Willingness to forgive was assessed on Visual Analog Scale (VAS) sliders.	V	V	V	V	V	MTG
Individual Differences									
(Li et al., 2017)	VBM 194 healthy individuals	Dispositional forgiveness	Tendency to forgive scale (associations with gray/white matter volume)	V					R insula, IFG
(Li & Lu, 2017)	Resting-state fMRI 178 healthy individuals	Dispositional forgiveness	Tendency to forgive scale (associations with spontaneous brain activity)		V		V		Parietal cortex

Note. Results are reported for the forgiveness measures only. Dorsal anterior cingulate cortex (dACC), dorsolateral prefrontal cortex (dlPFC); inferior frontal gyrus (IFG), medial prefrontal cortex (mPFC), middle temporal gyrus (MTG), occipital cortex (OCC), orbitofrontal cortex (OFC), posterior cingulate cortex (PCC), precuneus (PCU), superior temporal sulcus (STS), temporoparietal junction (TPJ), voxel-based morphometry (VBM), ventrolateral prefrontal cortex (vlPFC), ventromedial prefrontal cortex (vmPFC).
[a]Young and Saxe report only correlations between participants' judgements of blame and brain activity.
[b]Region of interest (ROI) analysis.
[c]Chronically rejected compared to stably highly accepted adolescents showed higher activity in this region when refraining from punishment.
[d]Functional connectivity analysis.

3.4 Perspective taking

The construct of perspective taking largely overlaps with theory of mind (ToM), the ability to explain, predict, and interpret behavior by attributing mental states such as desires, beliefs, intentions and emotions to oneself and to other people (Decety and Svetlova, 2012). The mentalizing system typically engages a neural network that includes the medial prefrontal cortex (mPFC), the temporoparietal junction (TPJ) and posterior superior temporal sulcus (pSTS), and medial parts of the parietal cortex, including the precuneus and posterior cingulate cortex (PCC) (Ciaramidaro et al., 2007; Koster-Hale and Saxe, 2013; Mitchell, 2009).

During the course of our daily routines, we seamlessly and continually attribute invisible internal states to others (real or fictitious), as such inferences underpin all social interaction (Saxe, 2006). Perspective taking, however, represents a more effortful, extended process whereby we actively try to imagine how another person thinks and feels given his/her situation (i.e., an imagine-other perspective), without self-other confusion (Lamm et al., 2007; van der Heiden et al., 2013). In this sense, perspective taking involves more than simply making mental inferences, it involves an "empathic attentional set" (Barrett-Lennard, 1981), whereby one is simultaneously sensitive to the thoughts and feelings of another and conscious of how this conception affects the self (Batson and Ahmad, 2009).

Consistent with the above reasoning, various lines of evidence suggest that adopting another's perspective is cognitively demanding, and hence requires higher demands on executive resources to be met. For example, in addition to brain areas involved in ToM, various studies investigating third-person perspective taking have found increased activation in prefrontal areas associated with executive attention, working memory, and inhibition (including the inferior frontal gyrus, dmPFC, and frontopolar cortex) (Lamm et al., 2010; Ruby and Decety, 2003, 2004; van der Heiden et al., 2013). Whereas controlled attention is required to activate relevant representations of other persons, inhibition of egocentric thoughts may facilitate cognitive flexibility, which is necessary to consider ideas and response options different to our own (Ruby and Decety, 2003; Samson et al., 2015).

By far the most consistent area activated by third-person perspective taking and mentalizing tasks, involves the TPJ (Cheng et al., 2010; Jackson et al., 2006; Lamm et al., 2007; Vistoli et al., 2016). Whereas the mPFC is implicated in reasoning about a person's stable psychological properties across time, including their enduring personality traits or social value, the TPJ appears to be more specifically involved in reasoning about another person's transitory mental states, such as specific goals, intentions, and desires (Schurz et al., 2014; Van Overwalle, 2009). In fact, because of the convergence of several basic cognitive (such as attention, memory and language) and social processing streams within the TPJ, it has been argued that this region serves a unique higher-order role in the creation of a social context for behavior (Carter and Huettel, 2013). Whatever the more domain-general computational mechanism contributed by the TPJ (Decety and Lamm, 2007), it appears that both the affective and cognitive understanding of others (Kanske et al., 2015), and the ability to distinguish between self and others (Decety and Grèzes, 2006), rely critically on processes subserved by this area.

The medial posterior areas involved in ToM appear to be instrumental in representing one's own self as a means to understand others. For example, the PCC seems to support internally directed thought (Leech et al., 2011), and the precuneus has been associated with episodic memory retrieval, self-related mental representations, and first-person perspective taking (Cavanna and Trimble, 2006). Converging evidence also suggests the precuneus contributes visuospatial mental imagery to represent the perspective of another person (Schurz et al., 2014).

Given the importance of understanding the wrongdoer's behavior and intentions for forgiveness to take place, it is not surprising that 11 out of 13 functional neuroimaging studies on forgiveness have found activation in areas associated with perspective taking (including the TPJ or STS, mPFC, precuneus, and PCC), regardless of the experimental paradigm employed (Table 2). Consistently, resting-state brain activity variation in mentalizing regions were associated with individual differences in the tendency to forgive (Li and Lu, 2017). While third-person perspective taking was not manipulated explicitly in these paradigms, each task involved judgments or decisions regarding an act of wrongdoing by a known or unknown other, so that perspective taking was required implicitly. In this regard, recent evidence suggests that implicit and explicit inferences regarding the contents of another's mind are subserved by a shared neural network involving core ToM areas (Van Overwalle and Vandekerckhove, 2013).

Young and Saxe (2009) observed that during moral judgments of accidental harms (i.e., unintentional harm on the basis of a false belief), participants with higher activation in the rTPJ were more likely to clear agents from blame, thus relying on information regarding the intent of the wrongdoer. Both studies that directly compared forgiving to unforgiving responses also detected significant activation in the rTPJ (Ricciardi et al., 2013; Strang et al., 2014). Ricciardi and colleagues (2013) reported that during forgiveness, the strength of the connection between the precuneus and inferior parietal lobule significantly correlated with participants' subjective relief. The authors argued that perspective taking may thus play a role in inducing positive affective states associated with forgiveness.

It should be noted that both studies where significant activation in perspective taking regions were not observed were limited in terms of their analyses to do so: Brüne and colleagues' (2013) fMRI analysis focused only on the dlPFC as a region of interest, whereas

Hayashi and colleagues (2010) did not separate in their analyses those scenarios for which forgiveness judgments were high from those for which forgiveness judgments were low.

3.5 Social valuation

In interpersonal relationships, people respond to wrongdoing by unwittingly calculating the potential future harm versus reward value inherent in the relationship with the harmdoer (Burnette et al., 2012; McCullough et al., 2013). This implicit value-tagging influences the decision to act in a retaliatory or reconciliatory fashion, depending on the costs that either choice incurs. For example, retribution might mean losing future benefits from a previous ally, whereas forgiveness might result in future exploitation. The neural architecture involved in this decision-making process likely relies fundamentally on the ventromedial prefrontal cortex (vmPFC)—an area that includes the anterior PFC, the medial sector of the orbitofrontal cortex (OFC), and the subgenual ACC (Rudebeck et al., 2008).

The vmPFC has consistently been implicated in studies investigating moral judgments and social decision-making (D'Argembeau, 2013; Moll and Schulkin, 2009; Young and Koenigs, 2007). It appears to play a particularly important role when the stakes are uncertain, that is, when the information available is insufficient to make decisions with certainty (Elliott et al., 2000). In such situations, vmPFC activation might reflect a course of action whereby the potential reward value of one's response is considered (Ruff and Fehr, 2014). In the context of forgiveness, this might mean taking into account, for example, the perceived association value of the perpetrator (Petersen et al., 2012; see also "welfare trade-off ratio", Tooby et al., 2008). This formulation is consistent with more recent accounts suggesting that the vmPFC is a core area encoding the subjective value of social and non-social stimuli in a context and goal-dependent manner (Bartra et al., 2013; Levy and Glimcher, 2012; McNamee et al., 2013). Of importance for its proposed role in forgiveness, is that the vmPFC is thus concerned with representing the value of decisions, thereby guiding behavior in terms of the reward-value of potential future outcomes (Amodio and Frith, 2006; Schoenbaum et al., 2011).

Evidence provided by neurological lesion studies dovetail with these functional neuroimaging findings, emphasising the importance of a functionally intact vmPFC for uncompromised social reasoning (Ciaramelli et al., 2007; Koenigs et al., 2007). Damage of this region, for example, is consistently associated with utilitarian choices in high-conflict, emotionally aversive, moral dilemmas (Young and Dungan, 2012). Specifically, the vmPFC/OFC appear to signal inappropriate social behavior, such that dysfunction in this area has been associated with reduced sensitivity to social norms, impaired ability to alter behavior in response to socially aversive cues, and socially unacceptable behavior in general (Beer et al., 2006; Blair and Cipolotti, 2000; Saver and Damasio, 1991). Functional suppression of the vmPFC might thus be necessary to act upon socially unacceptable impulses, including negative emotive or potentially aggressive behavior associated with unforgiveness (Worthington et al., 2007).

Scrutinising neuroimaging forgiveness studies to date, it appears only those studies involving judgment of a transgression or apology, and deciding on the consequent appropriateness of forgiveness resulted in significant activation in the vmPFC (Farrow et al., 2001, 2005; Hayashi et al., 2010; Ohtsubo et al., 2018; Young and Saxe, 2009). In each of these studies, participants were required to evaluate the specific social context from one situation to the next in order to assign blame to the perpetrator, judge the forgivability of the action (given the context) or the sincerity of the apology, or choose the most forgivable explanation for the event in question. In two of these studies, activity of the vmPFC was increased in the context of harmful mental states of the perpetrator: malicious desires to do harm intentionally (Young and Saxe, 2009), and dishonesty or deception (Hayashi et al., 2010). It should be noted that in the study by Ricciardi and colleagues (2013), which looked at neural activation in response to social scenarios concerning the self, vmPFC activation was observed for both forgiveness and unforgiveness (specifically the anterior part). Contrasting forgiveness against unforgiveness might thus have resulted in canceling out activity in this area. Consistent with the notion that the vmPFC is involved in computing the subjective value of a decision, it makes sense that this area would also be involved in unforgiveness.

Of significance, is that studies that involved forgiveness decisions in the context of economic games did not detect significant vmPFC activation. Those studies typically provided participants with limited or no social information in paradigms where they simply had to make offers or respond to others' fair or unfair offers in terms of personal financial gain (Brüne et al., 2013; Fatfouta et al., 2016; Haesevoets et al., 2018; Will et al., 2015, 2016). Perhaps information about the other player's socio-economic status (e.g., having recently lost a job), social group status (e.g., belonging to the same or a rival group) or personal attributes (e.g., being a bully) would make the decision to propose or accept fair/unfair offers more uncertain or complex, and hence engage the vmPFC more strongly. In this regard, the vmPFC has been shown to support flexible, value-based decisions across multiple domains (Hackel et al., 2017; Zaki et al., 2011), but may be less involved when behavior conforms to normative social principles that are stable from one trial to the next (Ruff and Fehr, 2014). More research is needed to tease apart these possibilities, and to determine whether economic games tap into forgiveness processes that are representative of those in real life situations.

3.6 Exploratory Meta-analysis

To statistically verify concurrence across previous work on forgiveness and provide preliminary support for our theoretical framework, we conducted a coordinate-based meta-analysis to reveal the regions with the highest likelihood of activation. Specifically, we employed a random-effects activation likelihood estimation (ALE) algorithm implemented in GingerALE (Eickhoff et al., 2012) (see Supplementary Material). While the contrasts across studies varied considerably (ranging across, for example, forgivability judgments, refraining from punishment, acceptance of unfair offers, and active forgiveness), coordinates of relevant foci were extracted from those contrasts that best represented forgiveness processing. It should further be noted that from the 8 studies included in this analysis, 50 % employed economic decision-making paradigms, hence affording greater weight to this task group than to paradigms employing social scenarios or direct forgiveness.

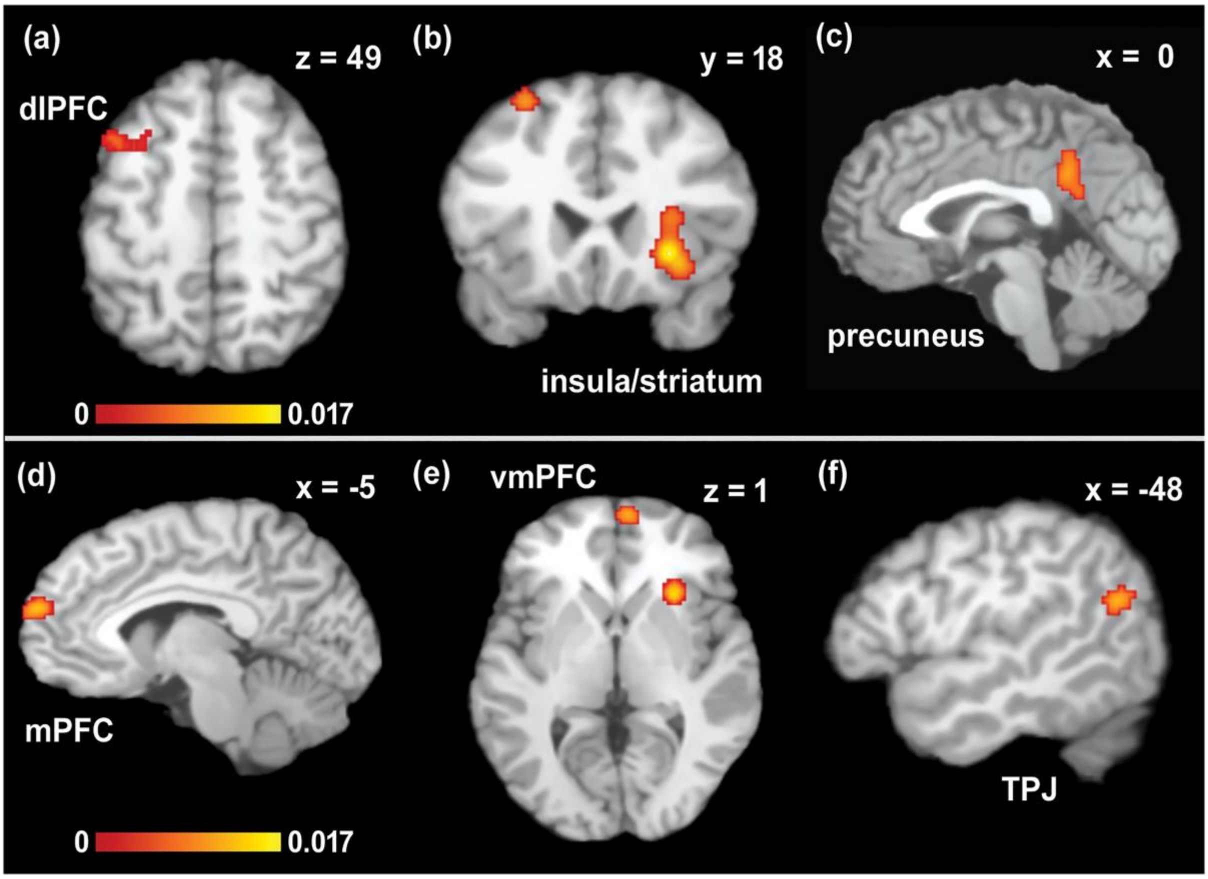

Figure 2. Results of the ALE meta-analysis of fMRI forgiveness studies. For (a), (b), and (c): cFWE, P < 0.05, and uncorrected cluster-defining, p < 0.01. For (d), (e), and (f): p < 0.005 (uncorrected), min cluster size 350 mm^3. The legend represents activation likelihood estimation (ALE) values as described in Table S1. dlPFC = dorsolateral prefrontal cortex; mPFC = medial prefrontal cortex; TPJ = temporoparietal junction; vmPFC = ventromedial prefrontal cortex.

The meta-analysis of fMRI forgiveness studies resulted in three activation clusters: (a) left dlPFC (centered at x = -34, y = 13, z = 49) with three peaks, (b) right anterior insula extending to inferior frontal gyrus (IFG) and striatum (centered at x = 29, y = 18, z = 0) with two peaks, and (c) precuneus extending to PCC (centered at x = 2, y = -50, z = 30) with three peaks (see Fig. 2**(a–c)** and **Table S1**). In addition to these clusters, at a more liberal threshold of $p < 0.005$ (uncorrected), we also observed the following clusters: (d) mPFC (centered at x = -5, y = 58, z = 20), (e) vmPFC (centered at x = 6, y = 57, z = 1), and (f) left TPJ (centered at x = -49, y = -65, z = 20) (see Fig. 2**(d–f)** and **Table S1**). These results, though tentative, thus corroborate engagement of all three proposed component processes in forgiveness studies to date.

4. DISCUSSION

In the current paper, we reviewed the growing body of literature into the neural architecture of forgiveness to begin conceptualizing a framework of its component processes. Consistent with behavioral work in cognitive and social psychology, we found support from functional neuroimaging studies of forgiveness for three distinct psychological constructs, namely cognitive control, perspective taking and social valuation. This parcellation of forgiveness component processes was furthermore supported by results from an exploratory ALE meta-analysis. Whereas previous work has made reference to the first two constructs (e.g., Billingsley and Losin, 2017; Brüne et al., 2013; Strang et al., 2014; Will et al., 2015), a unique contribution of the present review is highlighting the importance and neural architecture of social valuation in the decision to forgive. In addition, our review captures the diversity in tasks used to operationalise forgiveness. Here, our analysis sheds light on important associations between the extent to which previous methodologies weighed in on proposed forgiveness component processes and accompanied neural activation patterns. For the field to move forward, it would be key to develop paradigms that tap into all aspects of forgiveness, while having sound hypotheses of the underlying subprocesses and related brain substrates.

4.1 A neuroimaging framework of forgiveness

Behavioral studies converge on three dissociable, but interacting, components that are essential for forgiveness: cognitive control, perspective taking, and social valuation. Fortunately, neuroimaging research has established the validity of the neural indicators for these psychological constructs. In particular, we have highlighted the importance of the dlPFC, vlPFC, and dACC in cognitive control; the TPJ/pSTS, mPFC, precuneus and PCC (in addition to areas involved in controlled attention and inhibition processes) in perspective taking; and the vmPFC/OFC in social valuation; and as such, we have focused our review on the presence/absence of these regions in previous neuroimaging work on forgiveness.

Because the neuroimaging of forgiveness is still in its infancy, with relatively few published studies employing diverse methodologies, our review was exploratory in nature. Nevertheless, we were able to draw several important inferences from our data set. Notably, we found considerable evidence in the studies reviewed for the involvement of brain areas associated with our constructs of interest. Indeed, no other activations were detected with regular frequency across studies. An interesting exception, however, was the notable presence of activation in the insula/striatum in studies that also activated cognitive control areas—an observation that was corroborated by the exploratory meta-analysis. The anterior insula is implicated in the brain's putative revenge/reward system as an area involved in instantiating aversive emotional states and punishment in response to exploitation (Rilling and Sanfey, 2011). Significant insula activity is thus in line with the proposal that cognitive control is most directly associated with activity of the revenge/reward system (Billingsley and Losin, 2017). The most important finding, however, was that forgiveness studies to date appeared to fractionate neatly into their respective task groups. That is, studies with comparable stimuli, instructions, and control conditions tended to activate common sets of brain regions.

Confirming the importance of perspective taking for forgiveness to take place, we found that 85 % of previous functional imaging studies on forgiveness detected activation in one or more areas associated with perspective taking. Among these regions, the TPJ was activated most consistently, in line with its integral role in both the affective and cognitive understanding of others' thoughts, behavior, and intentions (Ciaramidaro et al., 2007; Kanske et al., 2015). More specifically related to forgiveness, a recent study suggests that the right TPJ plays a causal role in processing the mitigating circumstances of a harmful act, resulting in reduced moral blame of the transgressor (Leloup et al., 2016). It should be noted that those studies with no activation in regions supporting perspective taking were limited in terms of the analyses they carried out to detect such areas. Perspective taking thus emerged as the construct most reliably observed across task groups.

By comparison, areas associated with cognitive control were most consistently activated in studies using economic gaming paradigms. In particular, the dlPFC (in concert with the dACC, which serves important monitoring functions) was readily activated, and might reflect the explicit reappraisal of situations to counter one's own response tendencies, e.g., to propose fair offers to previously unfair opponents (Ochsner et al., 2012).

We have argued, however, that it is not the game theory paradigm as such, but rather the elicitation of strong negative affect in response to personal harm (e.g., social exclusion or unfair treatment), that resulted in activation in areas associated with cognitive control, presumably to curb a tit-for-tat retaliatory response. For the same plausible reason, dlPFC activation was *not* observed in studies using scenario-based forgiveness judgments from a third-person perspective, i.e., the scenarios did not concern participants personally and thus no harm was felt. The studies conducted by Ricciardi et al. (2013) and Ohtsubo et al. (2018)

marks two important exceptions to the rule, however—here we discerned brain activation associated with all three of our constructs of interest. Keeping our provisional framework in mind, it is possible to speculate why: Participants in these studies were instructed to imagine themselves in emotionally hurtful events by valued/close others, so that the narrative scenarios represented *personal* harm.

Finally, only studies involving forgiveness judgments in contextual social scenarios were associated with significant vmPFC/OFC activation. Forgiveness is exquisitely context-dependent, with cost-benefit computations regarding potential future interaction with the transgressor affecting the decision to forgive (Burnette et al., 2012). Because the vmPFC appears critical in assigning the current and future value we place on something when making decisions (Hutcherson et al., 2015; Schoenbaum et al., 2011), it is likely recruited during forgiveness judgements to perform such a cost-benefit analysis based on dynamic integration of situational factors. Importantly, the scenario-based stories conveyed contextual information (e.g., the intent, honesty, or blameworthiness of the perpetrator, or the costliness of the apology), which participants factored into their decisions to forgive.

By comparison, studies that employed economic games (where significant vmPFC activation was not observed), were almost devoid of any social context. For example, participants were naive about the wrongdoer's personal characteristics or intent. Hence, forgiveness in these situations—operationalized as accepting unfair offers from others, or proposing fair offers to previously unfair opponents (e.g., Fatfouta et al., 2016; Will et al., 2016)—perhaps relied on other factors, such as the participant's inherent altruistic tendencies or normative social principles, rather than on evaluating the subjective value of the particular social context. In fact, it is not clear that forgiveness could be inferred from these observed responses in the first place. As has been argued elsewhere, accepting an unfair offer might also meet the cognitive goal of maximizing one's own monetary gains (Sanfey et al., 2003), and proposing a fair offer to a previously hurtful opponent might not necessarily reflect an attempt at reaffiliation (Will et al., 2015). While these studies surely measured some aspects of forgiveness, a complete manipulation of the different components of forgiveness, as it typically occurs in real life, is lacking.

While the use of simplified models may have been a good place to start investigating a complex construct like forgiveness, it would be critical to develop paradigms assessing forgiveness in terms of how it most often operates outside the laboratory (Zaki and Ochsner, 2009). Neuroimaging research that focuses only on some aspects or component processes of forgiveness, while remaining indifferent about the role of others, runs the risk of overlooking deeper insights about the neurobiological structure and mechanisms underlying forgiveness. This is because investigating some aspects of forgiveness in isolation might not tell us how they interact during complex social information processing embedded in real-world settings (Krakauer et al., 2017; Zaki and Ochsner, 2012). Like most other complex psychological phenomena, forgiveness is likely to be greater than the sum of its parts (i.e., emergence).

The relative lack of coherence in previous neuroimaging work on forgiveness furthermore constrained our ALE meta-analysis: While an approach based on separate task groups with comparable stimulus-materials and instructions would have been advantageous in this situation for a more fine-grained analysis (see e.g., Schurz et al., 2014), we were restricted to the use of a pooled ALE meta-analysis. Specifically, the number of studies in each task group with sufficient whole-brain data (n = 2–4) was well below the minimum number of original studies required (n = 8–20) to perform valid ALE analysis with sufficient statistical power (David et al., 2013; Wagner et al., 2014). Consequently, our results were skewed disproportionately by studies employing economic gaming paradigms. Nevertheless, we believe the meta-analysis results, in combination with our qualitative review, offer a valuable synthesis of findings to date.

In summary, we have proposed a framework for discerning the underlying specific cognitive mechanisms involved in forgiveness that may guide future cognitive and social neuroimaging research. Because, despite our hypothesis-driven analysis, we relied on reverse inference to infer psychological processes from observed patterns of brain activation, the inductive validity of these inferences may be questioned. However, it is possible to recast these inferences in likelihoodist terms (i.e., deciding which of two competing hypotheses is best supported by the data), which has been proposed to circumvent the issues associated with reverse inference and provide genuine evidence for psychological hypotheses (Machery, 2014). These concerns notwithstanding, the present framework may facilitate more robust testing of hypotheses about how forgiveness component processes interact in real life, and how contextual and individual difference factors create variance in these systems, given the highly context-dependent nature of forgiveness.

4.2 Directions for future research

In what follows, we discuss a number of conceptual and methodological limitations characteristic of previous work, and how future research might move the field forward.

First, we concur with scholars who have argued that neuroimaging studies to date are not representative of forgiveness in real-life encounters, and therefore possess little ecological validity (Billingsley and Losin, 2017; Fatfouta et al., 2013). Forgiveness paradigms have typically involved artificial or decontextualized stimuli, thus differing qualitatively from real-world experiences. Moreover, for the largest part, previous studies have focused on wrongdoers that are unfamiliar, anonymous, or hypothetical whom participants do not meet or expect to meet again in future; instead of known others with whom participants are likely to affiliate with. And because no future contact with the wrongdoer is anticipated, these studies typically fail to involve consequences for either party, limiting their applicability to real-life situations.

More pertinent to justify the use of known others in forgiveness work, however, is that the 'perceived likelihood of affiliation' and 'value' of the social relationship appear defining when it comes to efforts to re-establish social connection, and therefore forgiveness (Maner et al., 2007; McCullough et al., 2010). In fact, it has been argued that forgiveness of strangers or people with whom one does not expect continuing contact with is fundamentally different from forgiving a loved one: whereas forgiving a stranger involves reducing *unforgiveness* and may best be described as *decisional forgiveness*, genuinely forgiving someone in a close relationship involves a more multifaceted change in cognition, emotion, and motivation, termed *emotional forgiveness* (Worthington et al., 2007).

To circumvent the issues described above, we believe more neuroimaging work on forgiveness should enter the personal realm, that is, forgiving familiar (in addition to unfamiliar) others whom one is affiliated to. The use of familiar others in paradigms tailored for each individual participant has been employed successfully in previous imaging studies investigating complex constructs like love (Bartels and Zeki, 2004) or social ostracism (Beeney et al., 2011). Moreover, results of such studies suggest that emotional closeness is a significant modulator (both qualitatively and quantitatively) of neural activation patterns. For example, when imagining a loved one in pain, greater relationship intimacy was associated with deactivation in the right TPJ and increased response in the insula and ACC (Cheng et al., 2010). When it comes to forgiveness, it remains to be seen how relationship closeness influences perspective taking, cognitive control and social valuation. Initial data suggest that functional connectivity between perspective taking and cognitive control areas are modulated when forgiving a partner (Fatfouta et al., 2016). Another important consideration for future research is how *expectations* about future interactions with wrongdoers might influence forgiveness component processes.

Second, as elaborated in the previous section, the contribution of neuroimaging studies to the understanding of forgiveness is often tempered by questions around construct validity (Cook and Campbell, 1979). In this regard, while studies may profess to measure forgiveness, this is often not assessed directly, but inferred through brain activation (Haesevoets et al., 2018) or behavioral responses (Brüne et al., 2013). In the context of game theory paradigms, for example, implicit forgiveness differs from mere acceptance of unfair offers (Fatfouta et al., 2013). Moreover, when forgiveness *is* assessed explicitly, social desirability and self-presentation issues may make it difficult to distinguish between hollow (behavioral) and true (both behavioral and internal) forgiveness (Baumeister et al., 1998), or between pseudo (false) and authentic forgiveness (Enright, 2001).

Here a fruitful direction for future research might be the use of autobiographical recall—something that has not previously been employed to parse the underlying neurocognitive components of forgiveness. Several behavioral investigations have successfully employed autobiographical recall to unravel certain aspects of forgiveness in the past (e.g., McCullough et al., 2006; Wallace et al., 2008). Zechmeister and Romero (2002), for example, were able to disentangle complete (true) versus incomplete (false) forgiveness using an autobiographical recall paradigm.

Autobiographical recall paradigms are commonly used in memory research, and can tap into neural processes that are difficult to study using exogenous stimuli (Cabeza and St Jacques, 2007). For example, it allows one to study people's reactions (cognitive and emotional) to aversive memories that are impossible, unethical or dangerous to recreate in the laboratory. In a similar fashion, in forgiveness research, the advantage of drawing on personally relevant experiences is that it allows for the subjective re-experience of an intrapsychic event, which may, in turn, result in neural activation similar to that of the original event (Buchanan, 2007). The added advantage of this approach, as suggested above, is that it may permit examination of forgiveness in responses to more serious offenses than is typically the case in laboratory work. An extreme example would be autobiographical recall of forgiveness experiences in the context of gross human rights violations, such as the Rwandan genocide or South African anti-apartheid struggle (Boraine et al., 1997; Staub, 2005). Neuroimaging work of this nature would provide a crucial juxtaposition to laboratory-based paradigms of forgiveness.

A third limitation of neuroimaging studies, and admittedly perhaps the most challenging, is that they have generally failed to take into account the temporal unfolding of forgiveness (Worthington et al., 2007). That is, neuroimaging studies of forgiveness have been almost entirely limited to time and space, focusing on participants' immediate reactions following a transgression (for an exception, see Farrow and Woodruff (2005)).

Forgiveness is unlikely to be an all-or nothing experience, but rather the outcome of a gradual information processing unfolding in time (Arendt, 1958). It has been described as a transition in one's internal motivational orientation toward the transgressor—a process of 'working through' the pain (Gobodo-Madikizela, 2018; McCullough et al., 2003). Some have argued it begins with the decision to forgive and is only complete once all feelings of anger or resentment toward the wrongdoer are set aside (Wilkowski et al., 2010). We suggest that in this gap ('calculation time'), processes of cognitive control, perspective taking, and social valuation are critically at work, and may be facilitated by spaces that encourage intersubjective engagement with the transgressor (Gobodo-Madikizela, 2015). However, it should be emphasized, that unlike revenge, which is more reactive, forgiveness can never quite be predicted or forced, but remains unconditioned by the behavior that provoked it (Arendt, 1958).

One way to measure the temporal unfolding of forgiveness (and its component processes), is through use of longitudinal neuroimaging paradigms paired with behavioral assessment. Such an approach has been used successfully in behavioral work (McCullough et al., 2014; Pronk et al., 2010; Riek et al., 2013), and would allow researchers to measure and compare behavioral and phenomenological changes over time with changes in neural activation. A recent neuroimaging study has made effective use of such an approach to probe the long-lasting effects of reappraisal, as an emotion regulation strategy, on brain activity (Denny et al., 2015). Longitudinal designs might also be particularly suitable for the study of forgiveness in specific population groups, for example, victims of crime. By assessing individuals and their responses to perpetrators over time using comprehensive measures (e.g., interviews, behavioral assessment, and neuroimaging), one might be able to disentangle different aspects of forgiveness and observe their temporal unfolding. The latter approach stresses the importance of interdisciplinary teams to strengthen theoretical and conceptual frameworks of forgiveness and elucidate phenomenological changes.

Because longitudinal neuroimaging designs require extensive planning and resources, test-retest paradigms might be employed effectively to collect behavioral and/or neuroimaging data in only two sessions. Imaging paradigms that track the unfolding of forgiveness processes in a single session might also be developed. For example, a ‘floating window’ technique could be employed to allow participants to deal with a complex affective process at their own pace (Ricciardi et al., 2013). In this method, the data are analyzed based on the response of the participant, who indicates the occurrence of the desired internal state—in this case, forgiveness.

Beyond the directions described above, we briefly note a few additional avenues of enquiry to advance the field. In particular, despite the long-term association between forgiving and forgetting in popular culture, our review of the literature suggests that empirical studies investigating this association at the neural level are lacking. While forgetting is not necessary or perhaps even useful for forgiveness, it may serve as an adaptive coping strategy to resolve hurt and anger associated with transgressions for some individuals (Cosgrove and Konstam, 2008). As with the ability to forgive, motivated forgetting is facilitated by lateral prefrontal areas involved in cognitive control (Anderson and Hanslmayr, 2014), and may thus present a mechanism whereby unwanted thoughts are removed from conscious awareness (Noreen et al., 2014). Future research should also examine the relationship between social group status and forgiveness processes in more detail, and how these align with our understanding of forgiveness at the interpersonal level (Baumgartner et al., 2012; Noor, 2016). Finally, different cultures are likely to differ in their valuation of social principles (Ruff and Fehr, 2014), and their expectations and understanding of forgiveness more broadly (Forster, 2018). Neuroimaging studies may elucidate how such cultural differences are instantiated at the neural level.

If the cognitive neuroscience of forgiveness is to provide valuable information to our understanding of forgiveness in real-world settings, we need to integrate naturalistic approaches. While the use of simplified, controlled stimuli is necessary to inform early models of any complex cognitive process, the translational value of neuroscience would be undermined if researchers rely on overly simplified models for too long (Zaki and Ochsner, 2012). Notably, data obtained using laboratory stimuli may differ qualitatively from the natural social environment in terms of information processing. Whereas laboratory paradigms might rely primarily on overt, cost/benefit reasoning, forgiveness in real-life encounters might be more implicit or elusive, and more time-consuming—materializing as ‘the emergence of the unexpected’ (Gobodo-Madikizela, 2018). In this latter sense, traditional neuroimaging techniques may be some way off in capturing fully the enigmatic complexity of forgiveness.

5. CONCLUSIONS

Understanding the neural and cognitive mechanisms involved in forgiving is a new area of inquiry—still fraught with challenges in eliciting and measuring *genuine* forgiveness within a controlled experimental environment. As a result, most neuroimaging research to date have homed in on isolated components of forgiveness, leaving unclear how these are put together or interact in more natural settings.

Here, drawing on behavioral work in cognitive and social psychology, we have construed a preliminary theoretical framework that may guide future neuroimaging analysis. In particular, we have proposed that forgiveness involves the dynamic interplay between three macro-component processes: cognitive control (contributed by the lateral PFC and dACC), perspective taking (contributed by the TPJ/pSTS, dmPFC, precuneus, and PCC), and social valuation (contributed by the vmPFC), which unfold over time in a highly context-dependent manner.

The framework presented here, and supported by our review of the literature, may facilitate the examination of hypotheses about the respective contribution of contextual, social, and individual differences to the variation in the ability to forgive. Future work may employ this framework to understand how certain aspects of forgiveness relate to important downstream social behaviors (e.g., prosocial affiliation with the wrongdoer), by relating known neural indicators of component processes to validated behavioral indices. Work of this nature will facilitate the bidirectional exchange of ideas between behavioral and neuroimaging research, where, at present, there is a lack of cross-talk. We believe the field will be advanced meaningfully by neuroimaging paradigms that examine the full extent of forgiveness using personally relevant stimuli, and through implementation of autobiographical and longitudinal designs in interdisciplinary contexts that may be especially suited to investigate its complexity.

ACKNOWLEDGEMENTS
MF was supported by a Blue Skies Research grant on Histories of Trauma from the National Research Foundation (NRF) of South Africa, as well as by Stellenbosch University. RH was supported by a Van der Gaag Grant of the Royal Netherlands Academy of Arts and Sciences, and a Professors' Fund grant of the University Fund Limburg. No conflicts of interest exist for any author regarding this manuscript.

REFERENCES

Amodio, D. M. (2010). Can neuroscience advance social psychological theory? Social neuroscience for the behavioral social psychologist. *Social Cognition, 28*, 695-716.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*, 268-277.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences, 18*, 279-292.

Arendt, H. (1958). *The Human Condition*. Chicago: University of Chicago Press.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *The Journal of Neuroscience, 27,* 3743-3752.

Ashar, Y. K., Andrews-Hanna, J. R., Dimidjian, S., & Wager, T. D. (2017). Empathic care and distress: Predictive brain markers and dissociable brain systems. *Neuron, 94,* 1263-1273.e1264.

Aureli, F., Das, M., & Veenema, H. C. (1997). Differential kinship effect on reconciliation in three species of macaques (Macaca fascicularis, M. fuscata, and M. sylvanus). *Journal of Comparative Psychology, 111*, 91-99.

Barrett-Lennard, G. T. (1981). The empathy cycle: Refinement of a nuclear concept. *Journal of Counseling Psychology*, *28*, 91-100.

Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *Neuroimage, 21*, 1155-1166.

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.

Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review, 3*, 141-177.

Baumeister, R. F., Exline, J. J., & Sommer, K. L. (1998). The victim role, grudge theory, and two dimensions of forgiveness. In E. L. Worthington (Ed.), *Dimensions of Forgiveness: Psychological Research & Theological Perspectives* (pp. 79-104). Philadelphia, PA: Templeton Foundation Press.

Baumgartner, T., Götte, L., Gügler, R., & Fehr, E. (2012). The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Human Brain Mapping, 33*, 1452-1469.

Baumgartner, T., Schiller, B., Rieskamp, J., Gianotti, L. R. R., & Knoch, D. (2013). Diminishing parochialism in intergroup conflict by disrupting the right temporo-parietal junction. *Social Cognitive and Affective Neuroscience, 9*, 653-660.

Beeney, J. E., Franklin, R. G., Jr., Levy, K. N., & Adams, R. B., Jr. (2011). I feel your pain: Emotional closeness modulates neural responses to empathically experienced rejection. *Social Neuroscience, 6*, 369-376.

Beer, J. S., John, O. P., Scabini, D., & Knight, R. T. (2006). Orbitofrontal cortex and social behavior: Integrating self-monitoring and emotion-cognition interactions. *Journal of Cognitive Neuroscience, 18*, 871-879.

Bennett, M., & Earwaker, D. (1994). Victims' responses to apologies: The effects of offender responsibility and offense severity. *The Journal of Social Psychology, 134*, 457-464.

Berndsen, M., Hornsey, M. J., & Wohl, M. J. A. (2015). The impact of a victim-focused apology on forgiveness in an intergroup context. *Group Processes and Intergroup Relations, 18*, 726-739.

Beyens, U., Yu, H., Han, T., Zhang, L., & Zhou, X. (2015). The strength of a remorseful heart: Psychological and neural basis of how apology emolliates reactive aggression and promotes forgiveness. *Frontiers in Psychology, 6*, 67-77.

Billingsley, J., & Losin, E. A. R. (2017). The neural systems of forgiveness: An evolutionary psychological perspective. *Frontiers in Psychology, 8*, 737.

Blair, R. J., & Cipolotti, L. (2000). Impaired social response reversal. A case of 'acquired sociopathy'. *Brain, 123* 1122-1141.

Bono, G., McCullough, M. E., & Root, L. M. (2008). Forgiveness, feeling connected to others, and well-being: Two longitudinal studies. *Personality and Social Psychology Bulletin, 34*, 182-195.

Boraine, A., Levy, J., & Scheffer, R. (1997). *Dealing with the past: Truth and Reconciliation in South Africa* (2nd ed.). Cape Town: IDASA.

Botvinick, M., Cohen, J. D., & Carter, C. S. (2004). Conflict-monitoring and anterior cingulate cingulate cortex: An update. *Trends in Cognitive Sciences, 8*, 539-546.

Brüne, M., Juckel, G., & Enzi, B. (2013). "An eye for an eye"? Neural correlates of retribution and forgiveness. *PLoS One, 8*, e73519.

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychological Bulletin, 133*, 761-779.

Burnette, J. L., McCullough, M. E., Van Tongeren, D. R., & Davis, D. E. (2012). Forgiveness results from integrating information about relationship value and exploitation risk. *Personality and Social Psychology Bulletin, 38*, 345-356.

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences*, *11*, 219-227.

Carlsmith, K. M., Wilson, T. D., & Gilbert, D. T. (2008). The paradoxical consequences of revenge. *Journal of Personality and Social Psychology, 95*, 1316-1324.

Carter, R. M., & Huettel, S. A. (2013). A nexus model of the temporal-parietal junction. *Trends in Cognitive Sciences, 17*, 328-336.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain, 129*, 564-583.

Cheng, Y., Chen, C., Lin, C. P., Chou, K. H., & Decety, J. (2010). Love hurts: An fMRI study. *NeuroImage, 51*, 923-929.

Chester, D. S., Eisenberger, N. I., Pond, R. S., Jr., Richman, S. B., Bushman, B. J., & Dewall, C. N. (2014). The interactive effect of social pain and executive functioning on aggression: An fMRI experiment. *Social Cognitive and Affective Neuroscience, 9*, 699-704.

Ciaramelli, E., Muccioli, M., Ladavas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience, 2*, 84-92.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., et al. (2007). The intentional network: How the brain reads varieties of intentions. *Neuropsychologia, 45*, 3105-3113.

Cole, M. W., Yeung, N., Freiwald, W. A., & Botvinick, M. (2009). Cingulate cortex: Diverging data from humans and monkeys. *Trends in Neurosciences, 32*, 566-574.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houhton-Mifflin.

Cosgrove, L., & Konstam, V. (2008). Forgiveness and forgetting: Clinical implications for mental health counselors. *Journal of Mental Health Counseling, 30*, 1-13.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*, 353-380.

D'Argembeau, A. (2013). On the role of the ventromedial prefrontal cortex in self-processing: The valuation hypothesis. *Frontiers in Human Neuroscience, 7*, 372.

David, S. P., Ware, J. J., Chu, I. M., Loftus, P. D., Fusar-Poli, P., Radua, J., et al. (2013). Potential reporting bias in fMRI studies of the brain. *PLoS ONE, 8*, e70104.

Davis, J. R., & Gold, G. J. (2011). An examination of emotional empathy, attributions of stability, and the link between perceived remorse and forgiveness. *Personality and Individual Differences, 50*, 392-397.

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology, 70*, 713-726.

de Waal, F. B. M., & Ren, R. (1988). Comparison of the reconciliation behavior of stumptail and rhesus macaques. *Ethology, 78*, 129-142.

Decety, J., & Cowell, J. M. (2015). Empathy, justice, and moral behavior. *AJOB Neuroscience, 6*, 3-14.

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research, 1079*, 4-14.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*, 71-100.

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *Neuroscientist, 13,* 580-593.

Decety, J., & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience, 2*, 1-24.

Decety, J., & Yoder, K. J. (2016). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social Neuroscience, 11*, 1-14.

Denny, B. T., Inhoff, M. C., Zerubavel, N., Davachi, L., & Ochsner, K. N. (2015). Getting over it: Long-lasting effects of emotion regulation on amygdala response. *Psychological Science, 26,* 1377-1388.

Denson, T. F., DeWall, C. N., & Finkel, E. J. (2012). Self-Control and Aggression. *Current Directions in Psychological Science, 21*, 20-25.

Drabant, E. M., McRae, K., Manuck, S. B., Hariri, A. R., & Gross, J. J. (2009). Individual differences in typical reappraisal use predict amygdala and prefrontal responses. *Biological Psychiatry, 65*, 367-373.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8,* 1784.

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage, 59*, 2349-2361.

Elliott, R., Dolan, R. J., & Frith, C. D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: Evidence from human neuroimaging studies. *Cerebral Cortex, 10,* 308-317.

Enright, R. D. (2001). *Forgiveness is a Choice: A Step-by-step Process for Resolving Anger and Restoring Hope*. Washington, DC.: American Psychological Association.

Enright, R. D., & Fitzgibbons, R. P. (2015). *Forgiveness therapy: An empirical guide for resolving anger and restoring hope*. Washington, DC: American Psychological Association.

Exline, J. J., & Baumeister, R. F. (2000). Expressing forgiveness and repentance: Benefits and barriers. In M. E. McCullough, K. Pargament & C. Thoresen (Eds.), *Forgiveness: Theory, Research, and Practice* (pp. 133-155). New York: Guilford.

Exline, J. J., Baumeister, R. F., Zell, A. L., Kraft, A. J., & Witvliet, C. V. (2008). Not so innocent: Does seeing one's own capacity for wrongdoing predict forgiveness? *Journal of Personality and Social Psychology, 94*, 495-515.

Farrow, T. F., Hunter, M. D., Wilkinson, I. D., Gouneea, C., Fawbert, D., Smith, R., et al. (2005). Quantifiable change in functional brain response to empathic and forgivability judgments with resolution of posttraumatic stress disorder. *Psychiatry Research: Neuroimaging, 140,* 45-53.

Farrow, T. F., & Woodruff, P. W. (2005). Neuroimaging of forgivability. In E. L. Worthington Jr (Ed.), *Handbook of Forgiveness* (pp. 259-272). New York: Brunner-Routledge.

Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., Tarrier, N., et al. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport, 12*, 2433-2438.

Fatfouta, R., Jacobs, A., & Merkl, A. (2013). Towards a multifaceted understanding of revenge and forgiveness. *Behavioral and Brain Sciences, 36*, 21.

Fatfouta, R., Meshi, D., Merkl, A., & Heekeren, H. R. (2016). Accepting unfairness by a significant other is associated with reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex. *Social Neuroscience, 13*, 61-73.

Fehr, R., Gelfand, M. J., & Nag, M. (2010). The road to forgiveness: A meta-analytic synthesis of its situational and dispositional correlates. *Psychological Bulletin, 136*, 894-914.

Finkel, E. J., & Campbell, W. K. (2001). Self-control and accommodation in close relationships: An interdependence analysis. *Journal of Personality and Social Psychology, 81*, 263-277.

Forster, D. (2018). Translation and a politics of forgiveness in South Africa? What black Christians believe, and white Christians do not seem to understand. *Stellenbosch Theological Journal, 4*, 77-93.

Fourie, M. M., Thomas, K. G., Amodio, D. M., Warton, C. M., & Meintjes, E. M. (2014). Neural correlates of experienced moral emotion: An fMRI investigation of emotion in response to prejudice feedback. *Social Neuroscience, 9*, 203-218.

Frank, M. J., & Badre, D. (2015). How cognitive theory guides neuroscience. *Cognition, 135*, 14-20.

Gabay, A. S., Radua, J., Kempton, M. J., & Mehta, M. A. (2014). The Ultimatum Game and the brain: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 47*, 549-558.

Gabrys, R. L., Tabri, N., Anisman, H., & Matheson, K. (2018). Cognitive control and flexibility in the context of stress and depressive symptoms: The cognitive control and flexibility questionnaire. *Frontiers in Psychology, 9*.

Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations, 8*, 109-124.

Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin, 131*, 76-97.

Gobodo-Madikizela, P. (2015). Psychological repair: The intersubjective dialogue of remorse and forgiveness in the aftermath of gross human rights violations. *Journal of the American Psychoanalytic Association, 63*, 1085-1123.

Gobodo-Madikizela, P. (2018). Forgiveness is 'the wrong word': Empathic repair and the potential for human connection in the aftermath of historical trauma. In M. Leiner & C. Shliesser (Eds.), *Alternative Approaches in Conflict Resolution* (pp. 111-123): Palgrave Macmillan.

Hackel, L. M., Zaki, J., & Van Bavel, J. J. (2017). Social identity shapes social valuation: Evidence from prosocial behavior and vicarious reward. *Social Cognitive and Affective Neuroscience, 12*, 1219-1228.

Haesevoets, T., De Cremer, D., Van Hiel, A., & Van Overwalle, F. (2018). Understanding the positive effect of financial compensation on trust after norm violations: Evidence from fMRI in favor of forgiveness. *Journal of Applied Psychology, 103*, 578-590.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science, 324*, 646-648.

Hayashi, A., Abe, N., Ueno, A., Shigemune, Y., Mori, E., Tashiro, M., et al. (2010). Neural correlates of forgiveness for moral transgressions involving deception. *Brain Research, 1332*, 90-99.

Huettel, S. (2015). Novel developments in cognitive fMRI. In U. K., U. K. & B. L. (Eds.), *fMRI: From Nuclear Spins to Brain Functions* (Vol. 30, pp. 557-579). Boston, MA: Springer.

Hutcherson, Cendri A., Bushong, B., & Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron, 87*, 451-462.

Jackson, P. L., Brunet, E., Meltzoff, A. N., & Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia, 44*, 752-761.

Kaminer, D., Stein, D. J., Mbanga, I., & Zungu-Dirwayi, N. (2001). The Truth and Reconciliation Commission in South Africa: Relation to psychiatric status and forgiveness among survivors of human rights abuses. *British Journal of Psychiatry, 178*, 373-377.

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

Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*, 1023-1026.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature, 446*, 908-911.

Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron, 79*, 836-848.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron, 93*, 480-490.

Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience, 19*, 42-58.

Lamm, C., Meltzoff, A. N., & Decety, J. (2010). How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience, 22*, 362-376.

Leech, R., Kamourieh, S., Beckmann, C. F., & Sharp, D. J. (2011). Fractionating the default mode network: Distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *The Journal of Neuroscience, 31*, 3217-3224.

Leloup, L., Miletich, D. D., Andriet, G., Vandermeeren, Y., & Samson, D. (2016). Cathodal transcranial direct current stimulation on the right temporo-parietal junction modulates the use of mitigating circumstances during moral judgments. *Frontiers in Human Neuroscience, 10*.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology, 22*, 1027-1038.

Li, H., Chen, Q., Lu, J., & Qiu, J. (2017). Brain structural bases of tendency to forgive: Evidence from a young adults sample using voxel-based morphometry. *Scientific Reports, 7*, 16856.

Li, H., & Lu, J. (2017). The neural association between tendency to forgive and spontaneous brain activity in healthy young adults. *Frontiers in Human Neuroscience, 11*.

Macaskill, A., Maltby, J., & Day, L. (2002). Forgiveness of self and others and emotional empathy. *Journal of Social Psychology, 142*, 663-665.

Machery, E. (2014). In defense of reverse inference. *The British Journal for the Philosophy of Science, 65*, 251-267.

Maier, M. J., Rosenbaum, D., Haeussinger, F. B., Brüne, M., Enzi, B., Plewnia, C., et al. (2018). Forgiveness and cognitive control – Provoking revenge via theta-burst-stimulation of the DLPFC. *NeuroImage, 183*, 769-775.

Maner, J. K., DeWall, C. N., Baumeister, R. F., & Schaller, M. (2007). Does social exclusion motivate interpersonal reconnection? Resolving the "porcupine problem". *Journal of Personality and Social Psychology*, *92*, 42-55.

Marks, M. J., Trafimow, D., Busche, L. K., & Oates, K. N. (2013). A function of forgiveness: Exploring the relationship between negative mood and forgiving. *SAGE Open, 3*, 2158244013507267.

McCullough, M. E. (2008). *Beyond Revenge: The Evolution of the Forgiveness Instinct*. San Francisco: Jossey-Bass.

McCullough, M. E., Bono, G., & Root, L. M. (2007). Rumination, emotion, and forgiveness: Three longitudinal studies. *Journal of Personality and Social Psychology*, *92*, 490-505.

McCullough, M. E., Fincham, F. D., & Tsang, J. A. (2003). Forgiveness, forbearance, and time: The temporal unfolding of transgression-related interpersonal motivations. *Journal of Personality and Social Psychology, 84*, 540-557.

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences, 36*, 1-15.

McCullough, M. E., Luna, L. R., Berry, J. W., Tabak, B. A., & Bono, G. (2010). On the form and function of forgiving: Modeling the time-forgiveness relationship and testing the valuable relationships hypothesis. *Emotion, 10,* 358-376.

McCullough, M. E., Pedersen, E. J., Tabak, B. A., & Carter, E. C. (2014). Conciliatory gestures promote forgiveness and reduce anger in humans. *Proceedings of the National Academy of Sciences, 111*, 11211-11216.

McCullough, M. E., Root, L. M., & Cohen, A. D. (2006). Writing about the benefits of an interpersonal transgression facilitates forgiveness. *Journal of Consulting and Clinical Psychology, 74*, 887-897.

McCullough, M. E., Worthington, E. L., Jr., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology, 73,* 321-336.

McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience, 16*, 479.

Miceli, M., & Castelfranchi, C. (2011). Forgiveness: A cognitive-motivational anatomy. *Journal for the Theory of Social Behaviour, 41*, 260-290.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 364*, 1309-1316.

Moll, J., & Schulkin, J. (2009). Social attachment and aversion in human moral cognition. *Neuroscience and Biobehavioral Reviews, 33*, 456-465.

Moll, J., Weingartner, J. H., Bado, P., Basilio, R., Sato, J. R., Melo, B. R., et al. (2014). Voluntary enhancement of neural signatures of affiliative emotion using fMRI neurofeedback. *PLOS ONE, 9*, e97343.

Moor, B. G., Guroglu, B., Op de Macks, Z. A., Rombouts, S. A., Van der Molen, M. W., & Crone, E. A. (2012). Social exclusion and punishment of excluders: Neural correlates and developmental trajectories. *NeuroImage, 59*, 708-717.

Noor, M. (2016). Suffering need not beget suffering: Why we forgive. *Current Opinion in Psychology, 11*, 100-104.

Noreen, S., Bierman, R. N., & MacLeod, M. D. (2014). Forgiving you is hard, but forgetting seems easy: Can forgiveness facilitate forgetting? *Psychological Science, 25*, 1295-1302.

North, J. (1998). The “ideal” of forgiveness: A philosopher’s exploration. In R. D. Enright & J. North (Eds.), *Exploring Forgiveness* (pp. 15-34). Madison: University of Wisconsin Press.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences, 9*, 242-249.

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

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: A synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences, 1251*, E1-24.

Ohbuchi, K., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology, 56*, 219-227.

Ohtsubo, Y., Matsunaga, M., Tanaka, H., Suzuki, K., Kobayashi, F., Shibata, E., et al. (2018). Costly apologies communicate conciliatory intention: An fMRI study on forgiveness in response to costly apologies. *Evolution and Human Behavior, 39*, 249-256.

Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior, 30*, 114-123.

Ohtsubo, Y., Watanabe, E., Kim, J., Kulas, J. T., Muluk, H., Nazar, G., et al. (2012). Are costly apologies universally perceived as being sincere? *Journal of Evolutionary Psychology, 10*, 187-204.

Patil, I., Calò, M., Fornasier, F., Young, L., & Silani, G. (2017). Neuroanatomical correlates of forgiving unintentional harms. *Scientific Reports, 7*, 45967.

Petersen, M. B., Sell, A., Tooby, J., & Cosmides, L. (2012). To punish or repair? Evolutionary psychology and lay intuitions about modern criminal justice. *Evolution and Human Behavior, 33*, 682-695.

Pronk, T. M., Karremans, J. C., Overbeek, G., Vermulst, A. A., & Wigboldus, D. H. J. (2010). What it takes to forgive: When and why executive functioning facilitates forgiveness. *Journal of Personality and Social Psychology*, *98*, 119-131.

Puga-Gonzalez, I., Butovskaya, M., Thierry, B., & Hemelrijk, C. K. (2014). Empathy versus parsimony in understanding post-conflict affiliation in monkeys: Model and epirical data. *PLOS ONE*, *9,* e91262.

Ricciardi, E., Rota, G., Sani, L., Gentili, C., Gaglianese, A., Guazzelli, M., et al. (2013). How the brain heals emotional wounds: The functional neuroanatomy of forgiveness. *Frontiers in Human Neuroscience, 7*, 839.

Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., et al. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience, 6*, 1323-1328.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*, 443-447.

Riek, B. M., Luna, L. M. R., & Schnabelrauch, C. A. (2013). Transgressors' guilt and shame: A longitudinal examination of forgiveness seeking. *Journal of Social and Personal Relationships, 31*, 751-772.

Riek, B. M., & Mania, E. W. (2012). The antecedents and consequences of interpersonal forgiveness: A meta-analytic review. *Personal Relationships, 19,* 304-325.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology, 62*, 23-48.

Rogé, B., & Mullet, E. (2011). Blame and forgiveness judgements among children, adolescents and adults with autism. *Autism, 15*, 702-712.

Ruby, P., & Decety, J. (2003). What you believe versus what you think they believe: A neuroimaging study of conceptual perspective-taking. *European Journal of Neuroscience, 17*, 2475-2480.

Ruby, P., & Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience, 16*, 988-999.

Rudebeck, P. H., Bannerman, D. M., & Rushworth, M. F. (2008). The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cognitive, Affective, & Behavioral Neuroscience, 8*, 485-497.

Rudorf, S., & Hare, T. A. (2014). Interactions between dorsolateral and ventromedial prefrontal cortex underlie context-dependent stimulus valuation in goal-directed choice. *The Journal of Neuroscience, 34*, 15988-15996.

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

Samson, D., Houthuys, S., & Humphreys, G. W. (2015). Self-perspective inhibition deficits cannot be explained by general executive control difficulties. *Cortex, 70*, 189-201.

Sandage, S. J., & Worthington, E. L. (2010). Comparison of two group interventions to promote forgiveness: Empathy as a mediator of change. *Journal of Mental Health Counseling*, *32*, 35-57.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science, 300*, 1755-1758.

Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, *29*, 1241-1249.

Saxe, R. (2006). Why and how to study Theory of Mind with fMRI. *Brain Research, 1079,* 57-65.

Schoenbaum, G., Takahashi, Y., Liu, T. L., & McDannald, M. A. (2011). Does the orbitofrontal cortex signal value? *Annals of the New York Academy of Sciences, 1239*, 87-99.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews, 42*, 9-34.

Sebastian, C. L., Tan, G. C. Y., Roiser, J. P., Viding, E., Dumontheil, I., & Blakemore, S.-J. (2011). Developmental influences on the neural bases of responses to social rejection: Implications of social neuroscience for education. *NeuroImage, 57,* 686-694.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain, 132*, 617-627.

Staub, E. (2005). Constructive rather than harmful forgiveness, reconciliation, and ways to promote them after genocide and mass killing. In E. L. Worthington Jr (Ed.), *Handbook of Forgiveness* (pp. 443-459). New York, NY: Routledge.

Stein, D. J., Seedat, S., Kaminer, D., Moomal, H., Herman, A., Sonnega, J., et al. (2008). The impact of the Truth and Reconciliation Commission on psychological distress and forgiveness in South Africa. *Social Psychiatry and Psychiatric Epidemiology, 43*, 462-468.

Steinbeis, N., Bernhardt, B. C., & Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron, 73*, 1040-1051.

Steindl, C., & Jonas, E. (2012). What reasons might the other one have?—Perspective taking to reduce psychological reactance in individualists and collectivists. *Psychology (Irvine, Calif.), 3*, 1153-1160.

Strang, S., Utikal, V., Fischbacher, U., Weber, B., & Falk, A. (2014). Neural correlates of receiving an apology and active forgiveness: An fMRI study. *PLoS One, 9*, e87654.

Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science, 19,* 339-347.

Takaku, S. (2001). The effects of apology and perspective taking on interpersonal forgiveness: A dissonance-attribution model of interpersonal forgiveness*. Journal of Social Psychology, 141*, 494-508.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology, 58*, 345-372.

Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass, 8*, 374-387.

Tooby, J., Cosmides, L., Sell, A., Lieberman, D., & Sznycer, D. (2008). Internal regulatory variables and the design of human motivation: A computational and evolutionary approach. In A. J. Elliott (Ed.), *Handbook of Approach and Avoidance Motivation* (pp. 251-272). New York, NY: Psychology Press.

Toussaint, L. L., Worthington, E. L., & Williams, D. R. (2015). *Forgiveness and Health: Scientific Evidence and Theories Relating Forgiveness to Better Health*. Dordrecht: Springer.

van der Heiden, L., Scherpiet, S., Konicar, L., Birbaumer, N., & Veit, R. (2013). Inter-individual differences in successful perspective taking during pain perception mediates emotional responsiveness in self and others: An fMRI study. *NeuroImage, 65*, 387-394.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping, 30*, 829-858.

Van Overwalle, F., & Vandekerckhove, M. (2013). Implicit and explicit social mentalizing: Dual processes driven by a shared neural network. *Frontiers in Human Neuroscience, 7*.

Vistoli, D., Achim, A. M., Lavoie, M. A., & Jackson, P. L. (2016). Changes in visual perspective influence brain activity patterns during cognitive perspective-taking of other people's pain. *Neuropsychologia, 85*, 327-336.

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron, 59*, 1037-1050.

Wagner, S., Sebastian, A., Lieb, K., Tüscher, O., & Tadić, A. (2014). A coordinate-based ALE functional MRI meta-analysis of brain activation during verbal fluency tasks in healthy control subjects. *BMC Neuroscience, 15*, 19.

Wallace, H. M., Exline, J. J., & Baumeister, R. F. (2008). Interpersonal consequences of forgiveness: Does forgiveness deter or encourage repeat offenses? *Journal of Experimental Social Psychology, 44*, 453-460.

Wilkowski, B. M., & Robinson, M. D. (2010). The anatomy of anger: An integrative cognitive model of trait anger and reactive aggression. *Journal of Personality, 78*, 9-38.

Wilkowski, B. M., Robinson, M. D., & Troop-Gordon, W. (2010). How does cognitive control reduce anger and aggression? The role of conflict monitoring and forgiveness processes. *Journal of Personality and Social Psychology, 98*, 830-840.

Will, G. J., Crone, E. A., & Guroglu, B. (2015). Acting on social exclusion: Neural correlates of punishment and forgiveness of excluders. *Social Cognitive and Affectice Neuroscience, 10*, 209-218.

Will, G. J., Crone, E. A., van Lier, P. A., & Guroglu, B. (2016). Neural correlates of retaliatory and prosocial reactions to social exclusion: Associations with chronic peer rejection. *Developmental Cognitive Neuroscience, 19*, 288-297.

Wohl, M. J. A., & Reeder, G. D. (2004). When bad deeds are forgiven: Judgments of morality and forgiveness for intergroup aggression. In J. P. Morgan (Ed.), *Focus on Aggression Research* (pp. 59-74). Hauppauge, NY, US: Nova Science Publishers.

Worthington, E. L., Kurusu, T. A., Collins, W., Berry, J. W., Ripley, J. S., & Baier, S. N. (2000). Forgiveness usually takes time. *Journal of Psychology and Theology, 28*, 3-20.

Worthington, E. L., Witvliet, C. V. O., Pietrini, P., & Miller, A. J. (2007). Forgiveness, health, and well-being: A review of evidence for emotional versus decisional forgiveness, dispositional forgivingness, and reduced unforgiveness. *Journal of Behavioral Medicine, 30*, 291-302.

Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience, 7*, 1-10.

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.

Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia, 47*, 2065-2072.

Zaki, J., & Ochsner, K. (2009). The need for a cognitive neuroscience of naturalistic social cognition. *Annals of the New York Academy of Sciences*, *1167*, 16-30.

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15,* 675-680.

Zaki, J., Schirmer, J., & Mitchell, J. P. (2011). Social influence modulates the neural computation of value. *Psychological Science*, *22*, 894-900.

Zechmeister, J. S., & Romero, C. (2002). Victim and offender accounts of interpersonal conflict: Autobiographical narratives of forgiveness and unforgiveness. *Journal of Personality and Social Psychology, 82*, 675-686.

SUPPLEMENTARY MATERIAL

Coordinate-Based Meta-Analysis

Inclusion and Exclusion Criteria

All studies included in the systematic review (*N* = 15) were evaluated for their suitability to include in the subsequent meta-analysis. Included in this analysis were original fMRI studies that (1) reported whole-brain data on contrasts related to forgiveness, (2) in either Talairach (TAL) or Montreal Neurological Institute (MNI) standard stereotactic coordinate space, and (3) with a sample size of 10 or more participants. Exclusion criteria were other brain scanning modalities, such as positron emission tomography (PET) (*n* = 1) or voxel-based morphometry (VBM) (*n* = 2), studies that measured resting-state fMRI (*n* = 1), or studies where results were based only on region of interest (ROI) analysis (*n* = 3). Only 8 studies satisfied these criteria and were included in the meta-analysis: 2 studies made use of social scenarios involving others (Farrow et al., 2005; Farrow et al., 2001), 4 studies employed economic decision-making paradigms (Fatfouta et al., 2016; Haesevoets et al., 2018; Will et al., 2015; Will et al., 2016), and 2 studies investigated imagined or direct forgiveness (Ohtsubo et al., 2018; Ricciardi et al., 2013).

Data Extraction

Foci data (peak coordinates in TAL or MNI space) were extracted for the ALE analysis and read as a text file containing the study name, number of subjects for the group of foci, and coordinate data. While the contrasts across studies varied considerably (ranging across, for example, forgivability judgments, refraining from punishment, and active forgiveness), coordinates of relevant foci were extracted from those contrasts that best represented forgiveness processing. Coordinates previously reported in MNI space were converted to Talairach space using the MNI to Talaraich conversion function (icbm2tal) implemented in GingerALE (Lancaster et al., 2007). The analysis included 87 foci from the 8 published studies.

Activation Likelihood Estimation

The meta-analysis was conducted using the revised random-effects ALE algorithm implemented in the GingerALE toolbox (version 3.0.2, http://brainmap.org/ale/, Research Imaging Center of the University of Texas Health Science Institute, San Antonio, Texas) (Eickhoff et al., 2009). GingerALE employs a 3D Gaussian probability distribution around individual foci to create a model activation map for each included study. Rather than using a specified Full-Width-Half-Maximum (FWHM) of the Gaussian function to blur all foci, GngerALE uses the number of subjects in each study to calculate the FWHM for that group of foci. In our analysis, ALE maps were computed using a cluster-level family-wise error (FWE)-corrected threshold value of $P < 0.05$ to correct for multiple comparisons (Eickhoff et al., 2012), a cluster-defining threshold of $p < 0.01$ (uncorrected), and 1000 permutations. Meta-analysis results were overlaid onto an anatomical template (Colin_tlrc_1x1x1.nii, http://brainmap.org/ale/) and displayed in Talairach space using the anatomical image overlay program Mango (http://ric.uthscsa.edu/mango/). In addition, to detect nonsignificant trends in the data that might be overlooked because of the small number of included studies, we conducted a post-hoc uncorrected ALE meta-analysis with a cluster-defining threshold of $p < 0.005$ and a minimum cluster size of 350 mm^3.

Supplementary Results

Table S1

Significant Clusters That Were Activated During Forgiveness Processing

Brain regions	BA	TAL Coordinates *x*	*y*	*z*	Cluster size (mm^3)	ALE value Max ($x10^{-2}$)
cFWE, $P < 0.05$, and uncorrected cluster-defining, $p < 0.01$						
dlPFC	9	-38	14	38	1456	1.19
Middle frontal gyrus	6	-26	16	54		
Middle frontal gyrus	6	-36	10	52		
Insula extending to IFG	13/47	28	18	-2	1968	1.66
Claustrum	-	28	16	12		
Precuneus	31	2	-50	32	1512	1.16
Precuneus	31	8	-52	32		
Posterior cingulate cortex	23	-2	-54	22		
Additional clusters detected at $p < 0.005$ (uncorrected, min cluster size 350 mm^3)						
mPFC	10	-4	58	20	808	1.27
vmPFC	10	6	58	2	384	1.08
TPJ	39	-48	-64	20	424	1.14
Superior frontal gyrus	6	10	12	64	680	1.27
Superior/middle frontal gyrus	6	20	14	60		

Note. Talairach coordinates refer to the peaks within each brain region.
ALE = activation likelihood estimation; BA = Brodmann Area; mPFC = medial prefrontal cortex; dlPFC = dorsolateral prefrontal cortex; IFG = inferior frontal gyrus; TPJ = temporoparietal junction; vmPFC = ventromedial prefrontal cortex

References

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage, 59,* 2349-2361.

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping, 30*, 2907-2926.

Farrow, T. F., Hunter, M. D., Wilkinson, I. D., Gouneea, C., Fawbert, D., Smith, R., et al. (2005). Quantifiable change in functional brain response to empathic and forgivability judgments with resolution of posttraumatic stress disorder. *Psychiatry Research: Neuroimaging, 140*, 45-53.

Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., Tarrier, N., et al. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport, 12*, 2433-2438.

Fatfouta, R., Meshi, D., Merkl, A., & Heekeren, H. R. (2016). Accepting unfairness by a significant other is associated with reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex. *Social Neuroscience, 13*, 61-73.

Haesevoets, T., De Cremer, D., Van Hiel, A., & Van Overwalle, F. (2018). Understanding the positive effect of financial compensation on trust after norm violations: Evidence from fMRI in favor of forgiveness. *Journal of Applied Psychology, 103*, 578-590.

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., et al. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping, 28*, 1194-1205.

Ohtsubo, Y., Matsunaga, M., Tanaka, H., Suzuki, K., Kobayashi, F., Shibata, E., et al. (2018). Costly apologies communicate conciliatory intention: An fMRI study on forgiveness in response to costly apologies. *Evolution and Human Behavior, 39*, 249-256.

Ricciardi, E., Rota, G., Sani, L., Gentili, C., Gaglianese, A., Guazzelli, M., et al. (2013). How the brain heals emotional wounds: The functional neuroanatomy of forgiveness. *Frontiers in Human Neuroscience, 7*, 839.

Will, G. J., Crone, E. A., & Guroglu, B. (2015). Acting on social exclusion: Neural correlates of punishment and forgiveness of excluders. *Social Cognitive and Affectice Neuroscience, 10,* 209-218.

Will, G. J., Crone, E. A., van Lier, P. A., & Guroglu, B. (2016). Neural correlates of retaliatory and prosocial reactions to social exclusion: Associations with chronic peer rejection. *Developmental Cognitive Neuroscience, 19*, 288-297.