doi:10.1093/scan/nss139 SCAN (2014) 9, 297–306

Deconstructing the brain's moral network: dissociable functionality between the temporoparietal junction and ventro-medial prefrontal cortex

Oriel FeldmanHall,[1,2] Dean Mobbs,[1] and Tim Dalgleish[1]

[1]Medical Research Council, Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK and [2]Cambridge University, Cambridge CB2 1TP, UK

Research has illustrated that the brain regions implicated in moral cognition comprise a robust and broadly distributed network. However, understanding how these brain regions interact and give rise to the complex interplay of cognitive processes underpinning human moral cognition is still in its infancy. We used functional magnetic resonance imaging to examine patterns of activation for 'difficult' and 'easy' moral decisions relative to matched non-moral comparators. This revealed an activation pattern consistent with a relative functional double dissociation between the temporoparietal junction (TPJ) and ventro-medial prefrontal cortex (vmPFC). Difficult moral decisions activated bilateral TPJ and deactivated the vmPFC and OFC. In contrast, easy moral decisions revealed patterns of activation in the vmPFC and deactivation in bilateral TPJ and dorsolateral PFC. Together these results suggest that moral cognition is a dynamic process implemented by a distributed network that involves interacting, yet functionally dissociable networks.

Keywords: fMRI; moral; TPJ; vmPFC

INTRODUCTION

Over the past decade, neuroscientists exploring moral cognition have used brain imaging data to map a 'moral network' within the brain (Young and Dungan, 2011). This network encompasses circuits implicated in social, emotional and executive processes. For example, moral emotions appear to activate the limbic system (Shin *et al.*, 2000) and temporal poles (Decety *et al.*, 2011), while reasoned moral judgments reliably engage fronto-cortical areas (Berthoz *et al.*, 2002; Heekeren *et al.*, 2003; Kedia *et al.*, 2008; Harenski *et al.*, 2010). The distributed nature of the network reflects the fact that prototypical moral challenges recruit a broad spectrum of cognitive processes: inferring people's intentions, integrating social norms, computing goal-directed actions, identifying with others and displaying empathic behavior (Moll *et al.*, 2008).

The initial focus within the research field was to explore whether moral decisions have a specific neural signature. This reflected the early dominance of neurocognitive models which argued for the unique properties of moral deliberation. One such theory endorsed the idea that we are endowed with an innate human moral faculty: our moral judgments are mediated by an unconscious mechanism which evaluates good *vs* bad (Hauser, 2006). Another theory suggested that moral choices are driven by intuitive emotions: in other words, we feel our way through knowing what is right and wrong (Haidt, 2001). However, as the imaging data accumulated, the theoretical emphasis shifted toward the view that the psychological processes underlying moral choices recruit socio-emotional and cognitive processes that are domain general (Moll *et al.*, 2005). As opposed to a unique moral faculty, the evidence reflected the fact that moral choices reliably engage a delineated neural network which is also observed within the non-moral domain (Young and Dungan, 2011). In line with this view, one theory postulates that emotional processes and reason work in competition: controlled processes of cognition and automatic processes of emotion vie with each other to 'work out' a moral judgment (Greene *et al.*, 2001). An alternative model suggests that reason and emotion do not act as competitive systems, but instead interact in a continuously integrated and parallel fashion (Moll *et al.*, 2008).

Reflecting this theoretical shift, more recent research efforts have used experimental probes to fractionate the moral network into constituent parts and illustrate relative dissociations. That is, distinct regions of the broad moral network are responsible for different putative components of moral cognition, and this likely mirrors domain-general processing distinctions. For example, there is now a compelling body of evidence that the anterior cingulate cortex (ACC) underpins processes of error detection and conflict monitoring across multiple cognitive contexts. This knowledge has been fruitfully applied to the moral domain in work showing that high-conflict moral dilemmas—when compared with low-conflict moral dilemmas—recruit the ACC (Greene *et al.*, 2004). Similarly, the temporoparietal junction (TPJ) seems to subserve the general capacity to think about another's perspective in socially contextualized situations and is reliably activated when participants deliberate over moral dilemmas where the ability to appreciate the interpersonal impact of a decision is paramount (Young *et al.*, 2007, 2011; Young and Saxe, 2009). This approach has also proved productive in elucidating the role of the ventro-medial prefrontal cortex (vmPFC) in coding socio-emotional knowledge, such as stereotypes (Gozzi *et al.*, 2009) and moral emotions—such as pride (Tangney *et al.*, 2007), embarrassment (Zahn *et al.*, 2009) and guilt (Moll *et al.*, 2011). Likewise, the dorsolateral PFC (dlPFC) appears to underpin cognitive control, reasoned thinking (Mansouri *et al.*, 2009), abstract moral principles (Moll *et al.*, 2002) and sensitivity to unfairness (Sanfey *et al.*, 2003). Finally, a similar rationale has informed research controlling for cognitive load (Greene *et al.*, 2008), semantic content (Takahashi *et al.*, 2004), emotional arousal and regulation (Moll and de Oliveira-Souza, 2007; Decety *et al.*, 2011), probability (Shenhav and Greene, 2010), intent (Berthoz *et al.*, 2002; Young and Saxe, 2011) and harm (Kedia *et al.*, 2008), in each case revealing distinct patterns of neural activation within the broader moral network.

Although this broad approach of deconstructing the moral network has clearly been very productive, it rests on an important assumption: that we can experimentally isolate different components of the moral

Received 17 July 2012; Accepted 24 November 2012
Advance Access publication 15 January 2013

Correspondence should be addressed to Oriel FeldmanHall, Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK. This research was supported by the Medical Research Council Cognition and Brain Sciences Unit. E-mail: Oriel.FeldmanHall@mrc-cbu.cam.ac.uk

© The Author (2013). Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

network in the brain by varying the relevant processing parameters (conflict, harm, intent and emotion) while keeping others constant (Christensen and Gomila, 2012). Another possibility of course is that varying any given parameter of a moral decision has effects on how other involved parameters operate. In other words, components of the moral network may be fundamentally interactive.

This study investigated this issue by building on prior research examining the neural substrates of high-conflict (difficult) *vs* low-conflict (easy) moral decisions (Greene *et al.*, 2004). Consider for example the following two moral scenarios. First, while hiding with your family during wartime your baby starts to cry; would you suffocate your crying baby in order to save the rest of your family from being discovered and killed by soldiers? Second, you are out with your family when you come across a child who has clearly been assaulted and is lying by the side of the road crying; do you assist them and call for help? Both of these decisions involve processing of 'right' and 'wrong' in terms of socially constructed moral rules. Both also have emotionally laden consequences and require processing of others' points of view (theory of mind). However, the first decision feels much more difficult than the second, involves a greater degree of mental conflict, will elicit more deliberation and will be met with less unanimity as to the 'correct' choice (Greene *et al.*, 2004). Together, these two scenarios clearly represent the ends of a moral continuum and offer a powerful illustration of the extent to which moral decisions can engage us in very discrepant ways.

The key question is exactly how patterns of neural activation in the moral network might differ when processing these varied classes of moral challenge. One possibility is that network activation will only differ as a function of the different cognitive parameters recruited (i.e. conflict resolution, engagement of systems involved in deliberative reasoning). If this were the case, difficult moral decisions may only differ from easy moral decisions in their recruitment of the dlPFC and ACC (Greene *et al.*, 2004). However, another possibility is that varying decision difficulty will have interactive effects on the recruitment of other components of the moral network. In other words, both classes of moral choice might require significant and broadly comparable appreciation of how the people involved will be affected by any choice that is made (i.e. theory of mind). If this were the case, mPFC and TPJ—regions known to be associated with perspective taking—may be recruited for both difficult and easy decisions. Such a finding would suggest that a shared cognitive process underlies a broad spectrum of moral challenges. However, it is also plausible that easy moral decisions solely rely on automatic and reflexive processing—which is often associated with limbic activation (Moll *et al.*, 2005). A further possibility is that the interplay and interactive effect of these various cognitive processes may engage some regions while disengaging others. For example, an easier moral decision may elicit less activation (or even deactivation) in the dlPFC simply because any dlPFC engagement would be redundant, or even a source of interference, when choices are reflexive and automatic.

We sought to investigate these various possibilities using functional magnetic resonance imaging (fMRI) while participants negotiated difficult *vs* easy moral decisions. Critically, we also included matched difficult and easy non-moral decision conditions. This allowed us to evaluate not only differences within the moral domain as a function of decision difficulty but also to investigate whether manipulation of 'difficulty' changes the pattern of activation in other regions of the moral network—relative to activation patterns for comparable non-moral choices. In other words, does moral cognition make flexible use of different regions of the moral network as a function of the demands of the moral challenge?

MATERIALS AND METHODS

Subjects

Overall, 89 subjects participated in the research reported here. Fifty-one subjects assisted us in rating the scenarios (mean age 29.6 years and s.d. ±7.2; 30 females). Thirty-eight subjects (all right handed, mean age 24.6 years and s.d. ±3.8; 22 females) participated in the main experiment and underwent fMRI. Three additional subjects were excluded from fMRI analyses due to errors in acquiring scanning images. Subjects were compensated for their time and travel. All subjects were right-handed, had normal or corrected vision and were screened to ensure no history of psychiatric or neurological problems. All subjects gave informed consent, and the study was approved by the University of Cambridge, Department of Psychology Research Ethics Committee.

Experimental procedures

Moral scenarios

In an initial stage of materials development, we created four categories of scenario for use in the imaging study: Difficult Moral Scenarios; Easy Moral Scenarios; Difficult Non-Moral Scenarios and Easy Non-Moral Scenarios. To achieve this, subjects ($N = 51$) were presented with a set of 65 moral and non-moral scenarios and asked which action they thought they would take in the depicted situation (a binary decision), how comfortable they were with their choice (on a five-point Likert scale, ranging from 'very comfortable' to 'not at all comfortable'), and how difficult the choice was (on a five-point Likert scale, ranging from 'very difficult' to 'not at all difficult'). This initial stimulus pool included a selection of 15 widely used scenarios from the extant literature (Greene *et al.*, 2001; Valdesolo and DeSteno, 2006; Crockett *et al.*, 2010; Kahane *et al.*, 2012; Tassy *et al.*, 2012) as well as 50 additional scenarios describing more everyday moral dilemmas that we created ourselves. These additional 50 scenarios were included because many of the scenarios in the existing literature describe extreme and unfamiliar situations (e.g. deciding whether to cut off a child's arm to negotiate with a terrorist). Our aim was for these additional scenarios to be more relevant to subjects' backgrounds and understanding of established social norms and moral rules (Sunstein, 2005). The additional scenarios mirrored the style and form of the scenarios sourced from the literature, however they differed in content. In particular, we over-sampled moral scenarios for which we anticipated subjects would rate the decision as very easy to make (e.g. would you pay $10 to save your child's life?), as this category is vastly under-represented in the existing literature. These scenarios were intended as a match for non-moral scenarios that we assumed subjects would classify as eliciting 'easy' decisions [e.g. would you forgo using walnuts in a recipe if you do not like walnuts? (Greene *et al.*, 2001)]—a category of scenarios that is routinely used in the existing literature as control stimuli.

Categorization of scenarios as moral *vs* non-moral was carried out by the research team prior to this rating exercise. To achieve this, we applied the definition employed by Moll *et al.*, (2008), which states that moral cognition altruistically motivates social behavior. In other words, choices, which can either negatively or positively affect others in significant ways, were classified as reflecting moral issues. Independent unanimous classification by the three authors was required before assigning scenarios to the moral *vs* non-moral category. In reality, there was unanimous agreement for every scenario rated.

We used the participants' ratings to operationalize the concepts of 'easy' and 'difficult'. First, we examined participants' actual yes/no decisions in response to the scenarios. We defined difficult scenarios as those where there was little consensus about what the 'correct' decision should be and retained only those where the subjects were more or less evenly split as to what to do (scenarios where the mean

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

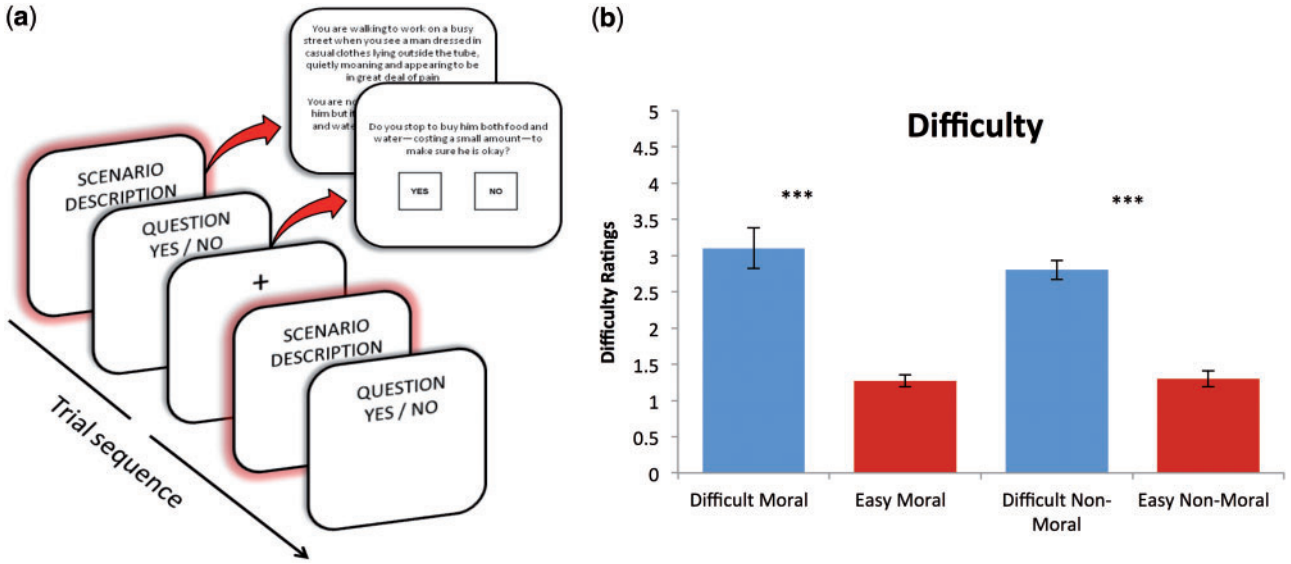

Fig. 1 (**a**) Experimental design. Subjects were presented with each scenario over two screens, the first describing the scenario and the second posing a question about their response to it. Subjects were required to select yes or no to make a choice. A fixation cross was presented for 2 s at the start of each trial. (**b**) Difficulty ratings from the subjects completing the fMRI study revealed that the categories Difficult/Easy and Moral/Non-Moral were controlled and matched across condition as rated on a five-point Likert scale.

proportion of responses was between 0.45 and 0.55 on the binary choice). In contrast, we defined easy scenarios as those where there was a strong consensus (either >0.80 or <0.20).

For these retained scenarios, we then examined participants' actual difficulty ratings. Scenarios that consistently (≥80% of the time) received high ratings of 'difficulty' (four or five on our five-point scale) or high ratings of 'easy' (one or two on the scale) were categorized as Difficult or Easy scenarios, respectively. This gave us 24 scenarios in the final set, 6 in each of our four categories (difficulty scores for each category: DM mean 3.2, s.d. ±0.71; DNM 2.9, s.d. ±0.70; EM 1.2, s.d. ±0.28; ENM mean 1.3, s.d. ±0.35). Of these 24, 6 came from the stimulus set drawn from the existing literature (Greene *et al.*, 2001) and a further 18 came from our supplementary set.

We then carried out a number of additional checks of potential between-category differences that we felt might drive behavioral and neural responses in our study. Consequently, we had a subset of the subjects ($n = 15$) rate each scenario on four further dimensions, all on five-point Likert scales. These comprised: (i) How much effort is required to complete the action resulting from your decision?; (ii) How much effort is required to weigh up each aspect/component of this scenario?; (iii) How many aspects/components did you consider when making your decision? and (iv) How emotionally involving is this scenario?

We wanted to ensure that the two sets of Difficult scenarios were rated as more effortful and complex (ratings, 1, 2 and 3) than the two sets of Easy scenarios, but that there were no differences on these ratings within the Difficult and Easy pairings. The data showed that this was the case [main effects of difficulty for the ratings 1, 2 and 3 (Fs > 49.74, Ps < 0.000), but no effects of difficulty within the pairings]. We also wanted to verify that the two sets of Moral scenarios were rated as more emotive (as we would predict) than the two sets of Non-Moral scenarios (as was the case, $t = -13.37$; $P < 0.001$; paired samples t-test, two-tailed), but that there were no differences within either the Moral or Non-Moral pairings (paired ts < 0.18) importantly illustrating that the difficult and easy scenarios in the moral and non-moral domains were matched on how emotionally involving they were. Finally, we ensured that the stimuli were matched for word length across categories [($F(3,20) = 0.51$, $P = 0.68$); DM word count (mean 86.3, s.d. ±25.3); EM word count (mean 92.0, s.d. ±20.1); DNM word count (mean 90.2, s.d. ±18.6) and ENM word count (mean 79.3, s.d. ±9.7)].

Functional MRI procedure

Within the scanner, subjects were presented with the 24 written scenarios. We structured our task using an event-related design, which closely mimicked past fMRI designs within this literature (Greene *et al.*, 2001). Scenarios were randomly presented in a series of four blocks with six trials (scenarios) per block. Each trial was presented as text through a series of two screens, the first of which described the short scenario and the second of which asked whether the subject would do the relevant action, requiring a yes/no button press (Figure 1a). Subjects read each scenario and question at their own pace (up to 25 s for the scenario and 15 s to make their choice) and pressed a button to advance through the screens. Between each trial, a fixation cross was displayed for 2 s. At the end of each block, there was an inter-block-interval (IBI) of 16 s to allow the hemodynamic response function to return to baseline. Baseline was defined as the mean signal across the last four images of this 16 s IBI. Neural activity was measured using the floating window method (Greene *et al.*, 2001). This method isolates the decision phase by including the time around the decision—8 s before the response, 1 s during the response and 6 s following the response—for a total of 15 s of recorded activity for every response. The rationale for using the floating window approach is to not only account for the 4–6 s delay following a psychological event in the hemodynamic response but also to create a flexible analysis structure for a complex, self-paced task.

Imaging acquisition

MRI scanning was conducted at the Medical Research Council Cognition and Brain Sciences Unit on a Siemens 3-Tesla Tim Trio MRI scanner by using a head coil gradient set. Whole-brain data were acquired with echoplanar T2* weighted imaging, sensitive to BOLD signal contrast (48 sagittal slices, 3 mm-thickness; TR = 2400 ms; TE = 30 ms; flip angle = 78° and FOV 192 mm). To provide for equilibration effects, the first 8 vol were discarded. T1

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

weighted structural images were acquired at a resolution of $1 \times 1 \times 1$ mm.

Imaging processing

Statistical parametric mapping software (SPM5: www.fil.ion.ucl.ac.uk/spm/software/spm5/) was used to analyze all data. Preprocessing of fMRI data included spatial realignment, coregistration, normalization and smoothing. The first eight scans were discarded as dummy scans. To control for motion, all functional volumes were realigned to the mean volume. Images were spatially normalized to standard space using the Montreal Neurological Institute (MNI) template with a voxel size of $3 \times 3 \times 3$ mm and smoothed using a Gaussian kernel with an isotropic full width at half maximum of 8 mm. Additionally, high-pass temporal filtering with a cut-off of 128 s was applied to remove low-frequency drifts in signal.

Data analysis

After preprocessing, statistical analysis was performed using the general linear model. Activated voxels were identified using an event-related statistical model representing each of the response events, convolved with a canonical hemodynamic response function and mean corrected. Six head-motion parameters defined by the realignment were added to the model as regressors of no interest. Analysis was carried out to establish each participant's voxel-wise activation when subjects made their response regarding each scenario (the aforementioned fixed 15 s floating window approach). For each subject, contrast images were calculated for each of the four scenario categories. These first level contrasts were then aggregated into second level full factorial analyses of variance (ANOVAs) in order to compute group statistics.

We report activity at $P < 0.001$ uncorrected for multiple spatial comparisons across the whole brain, and $P < 0.05$ family wise error (FWE) corrected for the following *a priori* regions of interest (ROIs; attained by independent coordinates): TPJ, ACC, dlPFC and vmPFC, reflecting the 'moral network' (coordinates listed in tables). Coordinates were taken from previous related studies.

RESULTS

Manipulation check: behavioral data

To validate our *a priori* allocation of scenarios to the Easy and Difficult categories based on participants' ratings, we administered a post-scan questionnaire to assess how difficult the fMRI subjects reported finding the scenarios using the same five-point Likert scale of difficulty. A repeated measures ANOVA with two within-subjects factors: Difficulty (difficult and easy) and Morality (moral and non-moral) confirmed the expected main effect of difficulty ($F(1,36) = 287.27$, $P < 0.001$), with Difficult scenarios rated as more difficult than Easy scenarios (Figure 1b). As anticipated, the main effect of morality and the morality by difficulty interaction were not significant, indicating that there was no support for self-reported differences in difficulty between moral and non-moral scenarios and no support for any differential discrepancy between difficult *vs* easy scenarios in the moral compared with non-moral domains ($Fs < 2.62$, $Ps > 0.13$).

As a further validation of our *a priori* categorization of scenarios as Difficult or Easy, we also examined response patterns for each of the different categories. Subjects had near perfect agreement in their responses for Easy decisions (98% of the subjects responded in the same manner). However, for Difficult scenarios, there was little consensus in response selection (only 57% of the subjects responded in the same manner). A repeated measures ANOVA exploring reaction times (Greene *et al.*, 2004) offered further support for this Difficult–Easy distinction, as Difficult scenarios (mean 4.0 s, s.d. ± 1.6) took significantly longer to respond to than Easy scenarios (mean 3.1 s, s.d. ± 1.1) ($F(1,36) = 24.34$, $P < 0.000$). Interestingly, moral scenarios (mean 3.65 s, s.d. ± 0.14) also took slightly longer to respond to relative to non-moral scenarios (mean 3.43 s, s.d. ± 0.15), likely reflecting their higher emotional impact ($F(1,36) = 5.35$, $P = 0.027$). There was therefore also a significant Difficulty by morality interaction ($F(1,36) = 143.14$, $P < 0.000$), reflecting the fact that the moral–difficult scenarios took the longest to respond to.

IMAGING RESULTS

We contrasted neural activation associated with making a decision for each of the four categories against one another: Easy Moral, Difficult Moral, Difficult Non-Moral and Easy Non-Moral. To explore potential interactions among the four conditions and to verify that overall the current scenarios elicited activations consistent with the moral network described in the literature (Moll, Zahn *et al.*, 2005), we ran a full factorial Morality × Difficulty ANOVA (Morality × Difficulty interaction). A whole-brain analysis of the interaction term (thresholded at $P = 0.001$ uncorrected) revealed a robust network of areas including bilateral TPJ, mid temporal poles, vmPFC, dACC and dlPFC (Figure 2; a full list of coordinates can be found in Table 1). We then examined *a priori* ROIs (Greene *et al.*, 2001; Young and Saxe, 2009) (thresholded at FWE $P = 0.05$) to determine if this network specifically overlapped with the regions delineated within the literature. As expected, the vmPFC, ACC and bilateral TPJ ROIs revealed significant activation for the interaction term. The interaction term qualified significant main effects of Morality and Difficulty. Although these activations are suprasumed by the interaction, for completeness, we report then in Tables 2 and 3.

As this initial full factorial analysis identified brain areas differing in activity as a function of the interaction of the Morality and Difficulty factors (the TPJ, dACC and vmPFC), our next aim was to deconstruct these interactions to examine functionality within those regions for Difficult and Easy Moral decisions relative to the matched Non-Moral comparison conditions.

First, in order to understand which areas are differentially more activated for difficult moral decisions, we compared Difficult Moral with Difficult Non-Moral scenarios (DM > DN) at the whole-brain level. This revealed a network starting at the TPJ and extending the length of the temporal lobe into the temporal pole (Figure 3a and Table 4). These findings demonstrate that difficult moral choices activate a network within the temporal lobe—areas implicated in theory of mind (Young and Saxe, 2009), attentional switching (Tassy *et al.*, 2012), higher order social concepts (Moll *et al.*, 2008) and the understanding of social cues (Van Overwalle, 2009).

To reveal brain regions demonstrating relative *decreases* in activity for difficult moral decisions, Difficult Non-Moral scenarios were contrasted with Difficult Moral scenarios (DN > DM), revealing vmPFC and bilateral orbital frontal cortex (OFC) deactivation (Figure 3a and Table 5). Thus, regions often associated with the moral network were found to be relatively less activated during difficult moral (*vs* non-moral) decisions once the difficulty of the scenario was controlled for.

Using a similar rationale, we compared Easy Moral decisions with Easy Non-Moral decisions (EM > EN), revealing activation of the vmPFC—an area known to integrate emotion into decision making and planning (Moretto *et al.*, 2010). Research has also shown that patients suffering damage to the vmPFC exhibit poor practical judgment (Raine and Yang, 2006; Blair, 2008). Interestingly, there was a pattern of TPJ and dlPFC relative deactivation for Easy Moral decisions (EN > EM) (Figure 3b and Tables 6 and 7).

Taken together, these patterns of activation and deactivation highlight that difficult moral decisions appear to differentially recruit the

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

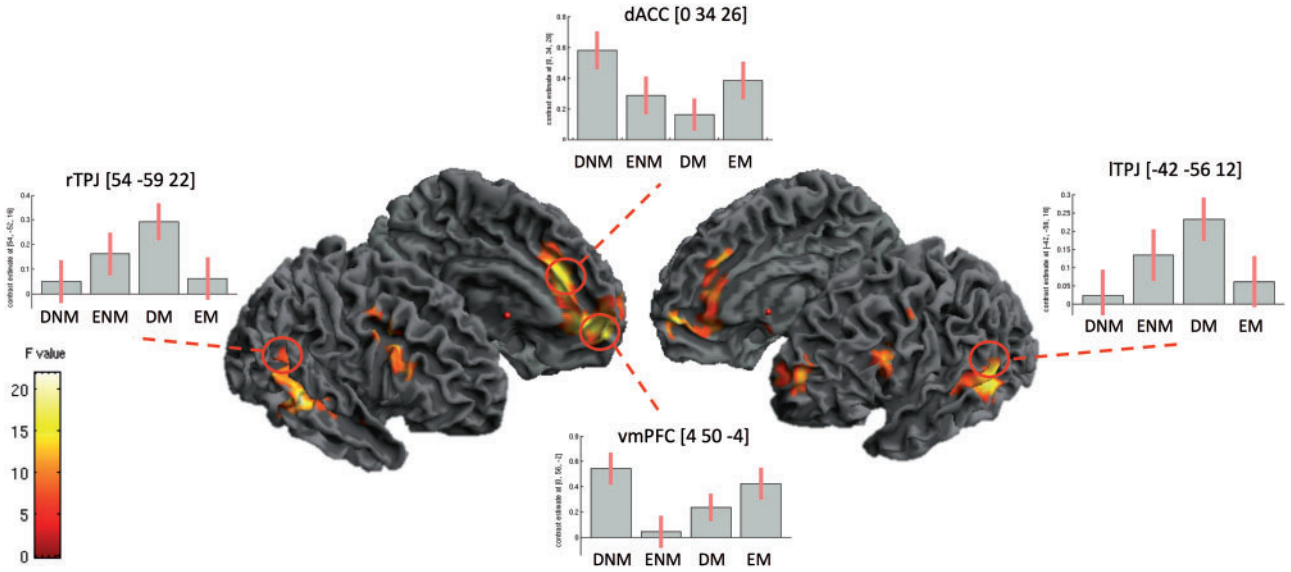

Fig. 2 *F*-test examining the interaction of the factors Morality and Difficulty. This contrast reveals activation of the moral network traditionally described in the literature, consisting of the TPJ (bilaterally), vmPFC, dlPFC and dACC. The red circles indicate the location of the regions used in the ROI analysis (taken from a priori coordinates), all thresholded at $P < 0.05$ FWE.

Table 1 ANOVA *F*-test interaction Morality × Difficulty

Region	Peak MNI coordinates			*F*-statistic/*z*-value
Medial OFC	−0	56	−2	21.89/4.36
Left ACC	−10	42	−4	17.95/3.95
Left dlPFC	−24	52	10	14.13/3.49
Right TPJ	56	−40	−4	20.17/4.19
Right TPJ	58	−52	14	13.73/3.43
Left TPJ	−56	−52	−2	16.67/3.80
Left TPJ	−50	−52	−12	14.23/3.50
Left ACC	−6	28	30	18.30/3.98
Right mid frontal gyrus	38	12	30	15.32/3.64
Left precentral gyrus	−52	−2	48	13.75/3.44
Right precentral gyrus	46	8	36	11.54/3.71
A priori ROIs	**MNI coordinates**			***F*-statistic/*z*-value**
[a]ACC	0	34	26	18.30/3.98
[a]Middle frontal gyrus	−28	49	7	14.13/3.49
[b]Right TPJ	54	−59	22	12.44/3.36
[b]Right TPJ	54	−52	16	13.73/3.44
[b]Right TPJ	52	−54	22	13.04/3.34
[b]Left TPJ	−52	−58	20	11.14/3.07
[b]vmPFC	2	58	17	11.57/3.13
[b]vmPFC	2	62	16	12.56/3.28
[b]vmPFC	2	50	−10	21.61/4.33
[b]vmPFC	4	50	−4	21.89/4.36

Notes: We used *a priori* coordinates to define ROI in our analysis. All ROIs were selected on the basis of independent coordinates using a sphere of 10 mm and corrected at $P < 0.05$ FWE and were attained through MarsBaRs. Peak voxels are presented in the tables at $P < 0.001$ uncorrected and all images are shown at $P < 0.005$ uncorrected. Cluster size was defined by a minimum of 10 contiguous voxels. All coordinates are in MNI Space. ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Greene *et al.* (2004) and [b]Young and Saxe (2009).

Table 2 Main effect of Difficulty (DM + DN > EM + EN)

Region	Peak MNI coordinates			*z*-value
vmPFC	−4	55	12	3.10

See footnote of Table 1 for more information.

Table 3 Main effect of Morality (DM + DN > EM + EN)

Region	Peak MNI coordinates			*z*-value
TPJ	−44	−78	34	3.82

See footnote of Table 1 for more information.

Table 4 Difficult Moral > Difficult Non-Moral (DM > DN)

Region	Peak MNI coordinates			*z*-value
Right mid temporal lobe	56	−2	−14	4.04
Right TPJ	56	−52	14	3.55
Left TPJ	−40	−58	16	3.74
Right mid temporal lobe	50	−16	−14	3.52
Left mid temporal lobe	−64	−56	10	3.61
Left post central gyrus	−54	−6	46	3.17
A priori ROIs	**MNI coordinates**			***t*-Statistic**
[a]Left TPJ	−58	−66	22	2.84
[a]Right TPJ	54	−52	16	3.64
[a]Right TPJ	54	−59	22	3.56

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Young and Saxe (2009). See footnote of Table 1 for more information.

TPJ and deactivate the vmPFC, while easy moral decisions appear to differentially deactivate the TPJ and activate the vmPFC, relative to the appropriate non-moral controls. These findings therefore suggest a degree of relative functional dissociation between the TPJ and vmPFC for moral decision making. The TPJ was selectively more engaged for difficult moral decisions, while in contrast, the vmPFC was selectively more activated for easy moral decisions, suggesting that these regions have different functional roles in the moral network.

To identify whether this activation and deactivation pattern associated with making difficult moral decisions overlapped with the network showing the reverse pattern implicated in making easy moral decisions, we performed a conjunction analysis. We first applied a conjunction to the contrasts Difficult Moral > Difficult Non-Moral

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

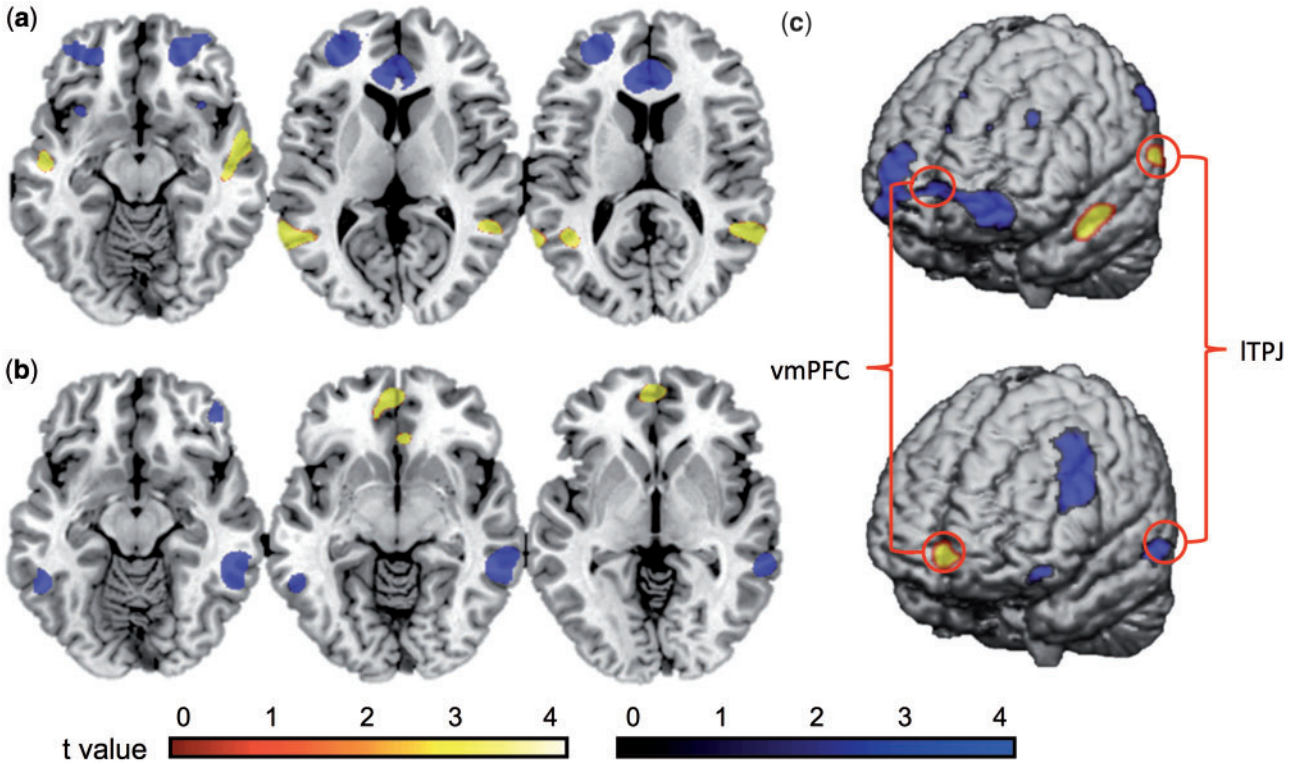

Fig. 3 (**a**) Whole-brain images for the contrast Difficult Moral > Difficult Non-Moral scenarios. The TPJ was activated (shown in yellow) while the vmPFC and bilateral OFC were deactivated (shown in blue: computed as Difficult Non-Moral > Difficult Moral). (**b**) Whole-brain images for contrast Easy Moral > Easy Non-Moral scenarios. The vmPFC was activated (shown in yellow) while the TPJ and dlPFC were deactivated (shown in blue: computed as Easy Non-Moral > Easy Moral scenarios). (**c**) A priori ROIs (indicated by red circles, corrected at FWE $P < 0.05$, are shown for the conjunction analysis of contrasts illustrated in Figure 3a and b (vmPFC [-2 54 -4] and TPJ [-52 -46 4]).

Table 5 Difficult Non-Moral > Difficult Moral (DN > DM)

Region	Peak MNI coordinates			*z*-value
MCC	0	28	34	4.66
vmPFC	0	54	2	3.37
Right OFC	22	46	−12	3.98
Left OFC	−26	48	−12	4.01
Left anterior insula	−32	16	−10	3.37
Right anterior insula	36	18	−10	3.24
A priori ROIs	**MNI coordinates**			***t*-statistic**
[a]ACC	0	34	26	4.84
[a]Middle frontal gyrus	−28	49	7	4.20
[b]vmPFC	2	50	−10	3.47
[b]vmPFC	4	50	−4	3.76

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Greene *et al.* (2004) and [b]Young and Saxe (2009). See footnote of Table 1 for more information.

Table 6 Easy Moral > Easy Non-Moral (EM > EN)

Region	Peak MNI coordinates			*z*-value
vmPFC	−2	54	−4	3.64
vmPFC	−12	46	6	3.19
ACC	6	30	−6	3.32
PCC	−2	60	26	3.00
A priori ROIs	**MNI coordinates**			***t*-statistic**
[a]vmPFC	2	50	−10	3.73

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Young and Saxe (2009). See footnote of Table 1 for more information.

(DM > DN) and Easy Non-Moral > Easy Moral (EN > EM) to clarify whether the TPJ activation associated with the former and the TPJ deactivation associated with the latter were occurring within the same region. A whole-brain analysis revealed bilateral TPJ activation, however, when *a priori* (Berthoz *et al.*, 2002) ROIs were applied, only the LTPJ survived SVC correction at $P < 0.05$ FWE (Figure 3c and Table 8). We also ran a conjunction analysis for Easy Moral > Easy Non-Moral (EM > EN) and Difficult Non-Moral > Difficult Moral (DN > DM) to determine whether the vmPFC activations and deactivations found in the original set of contrasts shared a common network. We found robust activity within the vmPFC region both at a whole-brain uncorrected level and when *a priori* (Young and Saxe, 2009) ROIs were applied (Figure 3c and Table 9).

We next investigated whether difficult moral decisions exhibited a neural signature that is distinct to easy moral decisions for our scenarios. By directly comparing Difficult Moral to Easy Moral decisions (DM > EM), bilateral TPJ as well as the right temporal pole were activated specifically for Difficult Moral decisions (Figure 4a and Table 10). A direct contrast of Easy Moral compared with Difficult Moral (EM > DM) revealed a network comprised of the Left OFC (extending into the superior frontal gyrus), vmPFC and middle cingulate (Figure 4b and Table 11). Interestingly, these results diverge from past findings which indicated that the dlPFC and ACC underpin difficult moral decisions (relative to easy moral decisions), while the TPJ and middle temporal gyrus code for easy moral decisions (relative to difficult moral decisions) (Greene *et al.*, 2004). One explanation for these differential findings may be that in our task, we independently categorized scenarios as difficult *vs* easy prior to scanning, instead of using each participant's response latencies as a metric of the difficulty of the moral dilemma (Greene *et al.*, 2004).

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

Table 7 Easy Non-Moral > Easy Moral (EN > EM)

Region	Peak MNI coordinates			*z*-value
Right TPJ	54	−44	−14	4.55
Left TPJ	−52	50	−14	3.80
Right dlPFC	46	12	50	3.87
Right dlPFC	52	16	28	3.43
A priori ROIs	**MNI coordinates**			***t*-statistic**
[a]Left TPJ	−51	−46	4	3.17

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Berthoz *et al*. (2002). See footnote of Table 1 for more information.

Table 8 Conjunction Difficult Moral > Difficult Non-Moral (DM > DN) + Easy Non-Moral > Easy Moral (EN > EM)

Region	Peak MNI coordinates			*z*-value
Right TPJ	56	42	0	2.80
Left TPJ	−56	−54	−2	2.79
A priori ROIs	**MNI coordinates**			***t*-statistic**
[a]Left TPJ	−52	−46	4	2.83

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Berthoz *et al*. (2002). See footnote of Table 1 for more information.

Table 9 Conjunction Easy Moral > Easy Non-Moral (EM > EN) + Difficult Non-Moral > Difficult Moral (DN > DM)

Region	Peak MNI coordinates			*z*-value
vmPFC	0	56	0	3.27
A priori ROIs	**MNI coordinates**			***t*-Statistic**
[a]vmPFC	4	50	−4	3.37

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Young and Saxe (2009). See footnote of Table 1 for more information.

DISCUSSION

The aim of the study reported here was to examine how the brain processes various classes of moral choices and to ascertain whether specific and potentially dissociable functionality can be mapped within the brain's moral network. Our behavioral findings confirmed that difficult moral decisions require longer response times, elicit little consensus over the appropriate response and engender high ratings of discomfort. In contrast, easy moral and non-moral dilemmas were answered quickly, elicited near perfect agreement for responses and created minimal discomfort. These differential behavioral profiles had distinct neural signatures within the moral network: relative to the appropriate non-moral comparison conditions, difficult moral dilemmas selectively engaged the bilateral TPJ but deactivated the vmPFC, while easy moral dilemmas revealed the reverse finding—greater vmPFC activation and less engagement of the TPJ. These results suggest a degree of functional dissociation between the TPJ and vmPFC for moral decisions and indicate that these cortical regions have distinct roles. Together, our findings support the notion that, rather than comprising a single mental operation, moral cognition makes flexible use of different regions as a function of the particular demands of the moral dilemma.

Our neurobiological results show consistency with the existing research on moral reasoning (Moll *et al*., 2008) which identifies both the TPJ and vmPFC as integral players in social cognition (Van Overwalle, 2009; Janowski *et al*., 2013). The vmPFC has largely been associated with higher ordered deliberation (Harenski *et al*., 2010), morally salient contexts (Moll *et al*., 2008) and emotionally engaging experiences (Greene *et al*., 2001). Clinical data have further confirmed these findings: patients with fronto-temporal dementia (FTD)—deterioration of the PFC—exhibit blunted emotional responses and diminished empathy when responding to moral dilemmas (Mendez *et al*., 2005). Additionally, lesions within the vmPFC produce a similar set of behaviors (Anderson *et al*., 1999). Unlike healthy controls, vmPFC patients consistently endorse the utilitarian response when presented with high-conflict moral dilemmas, despite the fact that such a response often has an emotionally aversive consequence (Koenigs *et al*., 2007). This clinical population is unable to access information that indicates a decision might be emotionally distressing, and they therefore rely on explicit norms that maximize aggregate welfare. This signifies that the vmPFC likely plays a role in generating pro-social sentiments such as compassion, guilt, harm aversion and interpersonal attachment (Moll *et al*., 2008).

In the experiment presented here, differential activity was observed within the vmPFC in response to easy moral dilemmas, suggesting that when a moral dilemma has a clear, obvious and automatic choice (e.g. pay $10 to save your child's life), this region supports a neural representation of the most motivationally compelling and 'morally guided' option. In other words, the vmPFC appears sensitive to a decision that has a low cost and high benefit result. This converges with the evidence that this area is critical for the experience of pro-social sentiments (Moll *et al*., 2008) and fits with the extant research demonstrating a strong association between the subjective value of reward and vmPFC activity (Hare *et al*., 2010). Because our moral scenarios were matched for emotional engagement, it seems unlikely that the vmPFC is only coding for the emotional component of the moral challenge. We speculated that when presented with an easy moral dilemma, the vmPFC may also be coding for both the subjective reward value and the pro-social nature of making a decision which produces a highly positive outcome.

Interestingly, when a moral dilemma is relatively more difficult, less activation within the vmPFC was observed. The nature of these more difficult moral scenarios is that there is no salient or motivationally compelling 'correct' choice. The options available to subjects elicit no explicit morally guided choice and are instead unpleasant and often even aversive (indicated by subjects' discomfort ratings). As a result, subjects understandably appear to be more reflective in their decision making, employing effortful deliberation (longer response latencies) during which they may be creating extended mental simulations of each available option (Evans, 2008). Thus, if the vmPFC is specifically coding the obvious and easy pro-social choice, then it is reasonable to assume that when there is no clear morally guided option, the vmPFC is relatively disengaged. This may be due to simple efficiency—suppression of activity in one region facilitates activity in another region. For example, any activity in the vmPFC might represent a misleading signal that there is a pro-social choice when there is not. In fact, patients with vmPFC lesions lack the requisite engagement of this region, and as a result, show behavioral abnormalities when presented with high-conflict moral dilemmas (Koenigs *et al*., 2007).

In contrast to easy moral dilemmas, difficult moral dilemmas showed relatively increased activity in the TPJ, extending down

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

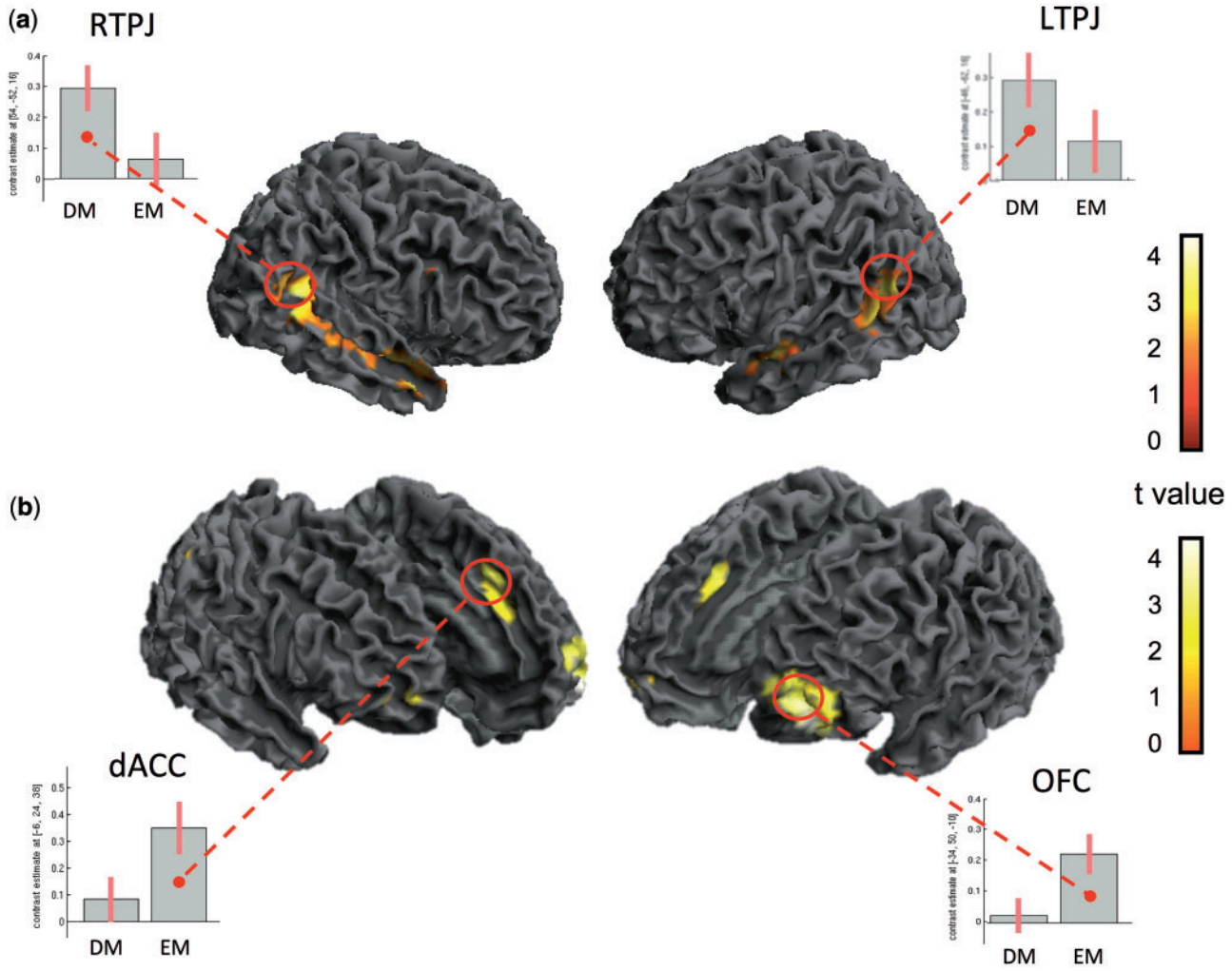

Fig. 4 (**a**) Whole-brain images for the contrast Difficult Moral > Easy Moral scenarios. Bilateral TPJ regions were activated and a priori ROIs were applied to these areas. Parameter estimates of the beta values indicate that the TPJ regions activate significantly more for Difficult Moral decisions than for Easy Moral decisions (**b**) Whole-brain images for the contrast Easy Moral > Difficult Moral scenarios reveal significant dACC and OFC activation. A priori ROIs were applied and parameter estimates of the beta values revealed that the dACC and OFC activate significantly more for Easy Moral decisions than for Difficult Moral decisions.

Table 10 Difficult Moral > Easy Moral (DM > EM)

Region	Peak MNI coordinates			*z*-value
Right TPJ	62	−54	14	3.55
Left TPJ	−38	−60	18	3.26
Right temporal pole	56	0	−18	3.26
A priori ROIs	**MNI coordinates**			***t*-statistic**
[a]Right TPJ	54	−52	16	3.63
[a]Left TPJ	−46	−62	25	3.32

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Young and Saxe (2009). See footnote of Table 1 for more information.

Table 11 Easy Moral > Difficult Moral (EM > DM)

Region	Peak MNI coordinates			*z*-value
Left OFC	−34	50	−10	3.75
Right OFC	30	62	−4	3.00
Left superior frontal gyrus	−20	54	6	3.47
MCC	−6	24	38	3.41
A priori ROIs	**MNI coordinates**			***t*-statistic**
[a]ACC	0	34	26	3.24
[a]Middle frontal gyrus	−28	49	7	3.59

ROIs, regions of interest corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Greene *et al.* (2004). See footnote of Table 1 for more information.

through the temporal poles. This activation pattern fits well with the fMRI documentation that the TPJ is integral in processing a diverse spectrum of social cognitive abilities such as empathy, theory of mind (Young and Saxe, 2009), agency and more basic processes such as attentional switching (Decety and Lamm, 2007). Converging evidence from clinical work has further implicated the TPJ in both mentalizing about the states of another, as well as attentional and spatial orientation (unilateral spatial neglect) (Mesulam, 1981). For example, during theory of mind tasks, subjects with autism either demonstrate abnormal TPJ activity (Baron-Cohen *et al.*, 1999) or fail to activate the TPJ altogether (Castelli *et al.*, 2002). Similar atypical TPJ activation was also found in autistic subjects who completed an attentional resource distribution task (Gomot *et al.*, 2006) and demonstrated difficulty in

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

Table 12 Difficult Non-Moral > Easy Non-Moral (DN > EN)

Region	Peak MNI coordinates			*z*-value
Mmfg	−6	54	0	4.57
Right ACC	6	30	−8	3.91
Right mOFC	0	38	−10	3.51
Ventral striatum (?)	0	2	−8	3.75
PCC	0	−56	32	3.42
A priori ROIs	**MNI coordinates**			***t*-statistic**
[a]ACC	0	34	26	3.26
[b]PCC	0	61	35	3.49
[b]mMPFC	2	58	17	4.13
[b]vMPFC	2	50	−10	4.70

ROIs, regions of interest SVC corrected at $P < 0.05$ FWE using a priori independent coordinates from previous studies: [a]Greene *et al*. (2004) and [b]Saxe (2009). See footnote of Table 1 for more information.

processing novel stimuli. Together, this research indicates that the TPJ seems to play a critical role in comparing and assessing socially salient stimuli (Decety and Lamm, 2007).

Based on these findings, we reasoned that more difficult moral decisions—which are not associated with normatively 'correct' choices—may rely more on reflective cognitive systems partly localized within the TPJ. Our behavioral data indicate that the major difference between difficult and easy dilemmas is not only the number of elements one must evaluate in order to make a decision but how much effort is required to do so. Thus, we speculated that the TPJ may process difficult dilemmas in two stages: the TPJ first subsumes the allocation of attentional resources to attend to the numerous socially relevant stimuli and is then critically implicated in the assessment of these stimuli to select the most compelling option. In short, the TPJ could be involved in attending to, shifting between, and then weighing up the salient nature of a difficult moral dilemma.

However, this neural result is not found when difficult and easy non-moral decisions are compared with one another (Table 12), which suggests that there is something specific about difficult moral decisions which engage the TPJ. What then distinguishes moral cognition from other forms of socially relevant decisions? While social interaction affects others, moral decisions are distinctive in that they can altruistically motivate interpersonal behavior (Moll *et al*., 2008). Accordingly, stimuli that are highly relevant and attentionally demanding—social cues, norms and taboos—necessitate processing according to their level of significance. This would mean that moral phenomena specifically require increased attentional resources because they are more consequential than non-moral phenomena. Thus, difficult decisions made within the moral domain are considerably more relevant and meaningful than difficult decisions made outside the moral domain. Hence, the TPJ appears to subserve the attention-oriented comparison of highly salient and meaningful moral stimuli.

Together, our results suggest that moral cognition emerges from the integration and coordination of disparate neural systems. This account extends the current moral cognitive framework by illustrating that not only do the TPJ and vmPFC have specific and differential roles but that they also operate within a flexible and competitive neural system. Dilemmas with a clearly guided moral choice require minimal processing of social information, and as a result, entail little cognitive demand. In contrast, moral dilemmas with ambiguously unfavorable outcomes demand greater deliberation and seemingly depend on an explicitly reflective system (Evans, 2008). The fact that the relationship between the TPJ and vmPFC appears to function within a dynamic equilibrium—when the TPJ is more engaged the vmPFC is less engaged, and vice versa—implies that moral decision making relies on a system of neural reallocation or mutual inhibition. Portions of the vmPFC and TPJ are specifically connected (Price and Drevets, 2010), and work has illustrated spontaneous correlations of activity between the TPJ and vmPFC (Burnett and Blakemore, 2009; Mars *et al*., 2012). Although speculative, such evidence of TPJ-vmPFC functional connectivity supports the idea that these regions may work together to encode moral choices. Interestingly, an experiment where the TPJ was transiently disrupted caused subjects to judge attempted harms as more morally permissible (Young *et al*., 2010). This suggests that when the TPJ 'turns off', neural resources may re-allocate to the vmPFC (where pro-social judgments may be generated). Such a mutual inhibitory process would mean that differential moral behavior competes for neural resources and thus rely on discrete and dissociable systems. Although beyond the scope of this research, it is possible that information processing taking place in these two classes of moral dilemmas act in direct opposition.

SUPPLEMENTARY DATA

Supplementary data are available at *SCAN* online.

REFERENCES

Anderson, S.W., Bechara, A., Damasio, H., Tranel, D., Damasio, A.R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, *2*(11), 1032–7.

Baron-Cohen, S., Ring, H.A., Wheelwright, S., et al. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *The European Journal of Neuroscience*, *11*(6), 1891–8.

Berthoz, S., Armony, J.L., Blair, R.J., Dolan, R.J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain: A Journal of Neurology*, *125*(Pt 8), 1696–708.

Blair, R.J. (2008). The amygdala and ventromedial prefrontal cortex: functional contributions and dysfunction in psychopathy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1503), 2557–65.

Burnett, S., Blakemore, S.J. (2009). Functional connectivity during a social emotion task in adolescents and in adults. *The European Journal of Neuroscience*, *29*(6), 1294–301.

Castelli, F., Frith, C., Happé, F., Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain: A Journal of Neurology*, *125*(Pt 8), 1839–49.

Christensen, J.F., Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: a principled review. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1249–64.

Crockett, M.J., Clark, L., Hauser, M.D., Robbins, T.W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(40), 17433–8.

Decety, J., Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *13*(6), 580–93.

Decety, J., Michalska, K.J., Kinzler, K.D. (2011). The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. *Cerebral Cortex*, *22*(1), 209–20.

Evans, J.S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–78.

Fliessbach, K., Weber, B., Trautner, P., et al. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, *318*(5854), 1305–8.

Gomot, M., Bernard, F.A., Davis, M.H., et al. (2006). Change detection in children with autism: an auditory event-related fMRI study. *NeuroImage*, *29*(2), 475–84.

Gozzi, M., Raymont, V., Solomon, J., Koenigs, M., Grafman, J. (2009). Dissociable effects of prefrontal and anterior temporal cortical lesions on stereotypical gender attitudes. *Neuropsychologia*, *47*(10), 2125–32.

Greene, J.D., Morelli, S.A., Lowenberg, K., Nystrom, L.E., Cohen, J.D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*(3), 1144–54.

Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389–400.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–8.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–34.

Hare, T.A., Camerer, C.F., Knoepfle, D.T., Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(2), 583–90.

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015

Harenski, C.L., Antonenko, O., Shane, M.S., Kiehl, K.A. (2010). A functional imaging investigation of moral deliberation and moral intuition. *NeuroImage*, *49*(3), 2707–16.
Hauser, M.D. (2006). The liver and the moral organ. *Social Cognitive and Affective Neuroscience*, *1*(3), 214–20.
Heekeren, H.R., Wartenburger, I., Schmidt, H., Schwintowski, H.P., Villringer, A. (2003). An fMRI study of simple ethical decision-making. *Neuroreport*, *14*(9), 1215–9.
Janowski, V., Camerer, C., Rangel, A. (2013). Empathic choice involves vmPFC value signals that are modulated by social processing implemented in IPL. *Social Cognitive and Affective Neuroscience*, *8*(2), 201–8.
Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., Tracey, I. (2012). The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience*, *7*(4), 393–402.
Kedia, G., Berthoz, S., Wessa, M., Hilton, D., Martinot, J.L. (2008). An agent harms a victim: a functional magnetic resonance imaging study on specific moral emotions. *Journal of Cognitive Neuroscience*, *20*(10), 1788–98.
Koenigs, M., Young, L., Adolphs, R., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*(7138), 908–11.
Mansouri, F.A., Tanaka, K., Buckley, M.J. (2009). Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nature Reviews. Neuroscience*, *10*(2), 141–52.
Mars, R.B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., Rushworth, M.F. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. *Cerebral cortex*, *22*(8), 1894–1903.
Mendez, M.F., Anderson, E., Shapira, J.S. (2005). An investigation of moral judgement in frontotemporal dementia. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology*, *18*(4), 193–7.
Mesulam, M.M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology*, *10*(4), 309–25.
Moll, J., de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends in Cognitive Sciences*, *11*(8), 319–21.
Moll, J., de Oliveira-Souza, R., Bramati, I.E., Grafman, J. (2002). Functional networks in emotional moral and nonmoral social judgments. *Neuroimage*, *16*(3 Pt 1), 696–703.
Moll, J., de Oliveira-Souza, R., Zahn, R. (2008). The neural basis of moral cognition: sentiments, concepts, and values. *Annals of the New York Academy of Sciences*, *1124*, 161–80.
Moll, J., Zahn, R., de Oliveira-Souza, R., et al. (2011). Impairment of prosocial sentiments is associated with frontopolar and septal damage in frontotemporal dementia. *NeuroImage*, *54*(2), 1735–42.
Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., Grafman, J. (2005). Opinion: the neural basis of human moral cognition. *Nature Reviews. Neuroscience*, *6*(10), 799–809.
Moretto, G., Làdavas, E., Mattioli, F., di Pellegrino, G. (2010). A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *Journal of Cognitive Neuroscience*, *22*(8), 1888–99.
Price, J.L., Drevets, W.C. (2010). Neurocircuitry of mood disorders. *Neuropsychopharmacology*, *35*(1), 192–216.
Raine, A., Yang, Y. (2006). Neural foundations to moral reasoning and antisocial behavior. *Social Cognitive and Affective Neuroscience*, *1*(3), 203–13.
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*(5626), 1755–8.
Shenhav, A., Greene, J.D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*(4), 667–77.
Shin, L.M., Dougherty, D.D., Orr, S.P., et al. (2000). Activation of anterior paralimbic structures during guilt-related script-driven imagery. *Biological Psychiatry*, *48*(1), 43–50.
Sunstein, C.R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, *28*(4), 531–541.
Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *NeuroImage*, *23*(3), 967–74.
Tangney, J.P., Stuewig, J., Mashek, D.J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*, 345–72.
Tassy, S., Oullier, O., Duclos, Y., et al. (2012). Disrupting the right prefrontal cortex alters moral judgement. *Social Cognitive and Affective Neuroscience*, *7*(3), 282–8.
Valdesolo, P., DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science*, *17*(6), 476–7.
Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, *30*(3), 829–58.
Vrticka, P., Andersson, F., Grandjean, D., Sander, D., Vuilleumier, P. (2008). Individual attachment style modulates human amygdala and striatum activation during social appraisal. *PLoS One*, *3*(8), e2868.
Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(15), 6753–8.
Young, L., Cushman, F., Hauser, M., Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(20), 8235–40.
Young, L., Dungan, J. (2011). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, *7*(1), 1–10.
Young, L., Saxe, R. (2009). An FMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, *21*(7), 1396–405.
Young, L., Saxe, R. (2011). When ignorance is no excuse: different roles for intent across moral domains. *Cognition*, *120*(2), 202–14.
Young, L., Scholz, J., Saxe, R. (2011). Neural evidence for "intuitive prosecution": the use of mental state information for negative moral verdicts. *Social Neuroscience*, *6*(3), 302–15.
Zahn, R., Moll, J., Paiva, M., et al. (2009). The neural basis of human social values: evidence from functional MRI. *Cerebral cortex*, *19*(2), 276–83.

Downloaded from http://scan.oxfordjournals.org/ at Columbia University Health Sciences Library on July 30, 2015