Functional Specificity and Sex Differences in the Neural Circuits Supporting the Inhibition of Automatic Imitation

Kohinoor M. Darda*, Emily E. Butler*, and Richard Ramsey

Abstract

■ Humans show an involuntary tendency to copy other people's actions. Although automatic imitation builds rapport and affiliation between individuals, we do not copy actions indiscriminately. Instead, copying behaviors are guided by a selection mechanism, which inhibits some actions and prioritizes others. To date, the neural underpinnings of the inhibition of automatic imitation and differences between the sexes in imitation control are not well understood. Previous studies involved small sample sizes and low statistical power, which produced mixed findings regarding the involvement of domain-general and domain-specific neural architectures. Here, we used data from Experiment 1 ($N = 28$) to perform a power analysis to determine the sample size required for Experiment 2 ($N = 50$; 80% power). Using independent functional localizers and an analysis pipeline that bolsters sensitivity, during imitation control we show clear engagement of the multiple-demand network (domain-general), but no sensitivity in the theory-of-mind network (domain-specific). Weaker effects were observed with regard to sex differences, suggesting that there are more similarities than differences between the sexes in terms of the neural systems engaged during imitation control. In summary, neurocognitive models of imitation require revision to reflect that the inhibition of imitation relies to a greater extent on a domain-general selection system rather than a domain-specific system that supports social cognition. ■

INTRODUCTION

Human social interactions are guided by nonverbal cues, such as copying behaviors. In the last two decades, much research has investigated the involuntary tendency to copy other's actions—a phenomenon known as automatic imitation (Heyes, 2011). Automatic imitation is thought to be beneficial in social situations because it develops affiliative attitudes, better cooperation, and feelings of closeness between interacting partners (Chartrand & Lakin, 2013). Prior neuroscience research has shown that imitation is supported by the mirror neuron system, a neural network engaged in perceiving and performing actions (Iacoboni, 2009; Rizzolatti & Craighero, 2004; Iacoboni et al., 1999). Imitation, however, is unlikely to rely on a single cognitive or brain system (Southgate & Hamilton, 2008). For example, in many circumstances, imitation is maladaptive and requires inhibition (Cross & Iacoboni, 2014; Cross, Torrisi, Losin, & Iacoboni, 2013; van Schie, van Waterschoot, & Bekkering, 2008; Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007). In such situations, a selection mechanism is required to suppress the tendency to imitate and prioritize alternative actions (Brass, Ruby, & Spengler, 2009). To date, studies investigating the neural mechanisms of imitation control

have been limited by small sample sizes and low statistical power, which has produced mixed findings (Table 1). Furthermore, no neuroscience research has investigated how individual differences such as sex modulate imitation control, even though behavioral research has shown that imitative tendencies vary as a function of sex (Butler, Ward, & Ramsey, 2015; Sonnby-Borgström, Jönsson, & Svensson, 2008; Dimberg & Lundquist, 1990). Across two fMRI experiments, which had higher statistical power and functional sensitivity than prior studies, we investigated the extent to which imitation inhibition relies on a domain-specific or domain-general neural network, which varies its response as a function of sex.

Much like cognitive science in general (Kanwisher, 2010; Hirschfeld & Gelman, 1994), inhibitory control research has focused on a neat division between domain-general and domain-specific mental operations. Domain-general inhibitory systems, which operate across multiple tasks, have been identified in dorsal frontoparietal cortices (Aron, Robbins, & Poldrack, 2014; Hazeltine, Poldrack, & Gabrieli, 2007; Nee, Wager, & Jonides, 2007; Wager et al., 2005; Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002). This brain circuit has been labeled the multiple demand (MD) network because of its engagement in a diversity of mental operations (Duncan, 2010). By contrast, evidence from fMRI, neurostimulation, and neuropsychological patient studies has suggested that a domain-specific

Bangor University
*Contributed equally to this work.

circuit in an anterior portion of medial prefrontal cortex (mPFC) and right temporo-parietal junction (rTPJ) operates during the inhibition of imitation (Bardi, Gheza, & Brass, 2017; Santiesteban, Banissy, Catmur, & Bird, 2012, 2015; Sowden & Catmur, 2015; Hogeveen et al., 2014; Klapper, Ramsey, Wigboldus, & Cross, 2014; Wang, Ramsey, & Hamilton, 2011; Spengler, von Cramon, & Brass, 2009, 2010; Brass, Derrfuss, Matthes-von Cramon, & von Cramon, 2003; Brass, Zysset, & von Cramon, 2001). Beyond the control of imitation, mPFC and rTPJ have been consistently implicated in a variety of social cognition functions, which require distinguishing between self and other, as well as reasoning about other people's mental states (theory of mind [ToM]; Van Overwalle, 2009; Amodio & Frith, 2006; Saxe & Kanwisher, 2003; Frith & Frith, 1999). These results led to theorizing that a key neural circuit for social cognition also regulates imitative tendencies (Brass et al., 2009).

Although theories of imitation control have been developed that are based on functioning of the ToM network, evidence from fMRI studies that used an RT measure of imitation inhibition have not provided consistent support for the involvement of a domain-specific neural network (Table 1). The RT measure of imitation involves making finger movements while simultaneously watching compatible or incompatible finger movements (Brass, Bekkering, Wohlschlager, & Prinz, 2000; Stürmer et al., 2000). The dif-ference between RTs in these two conditions (i.e., the general compatibility effect) has been argued to index imitative control, as greater cognitive resources are required to inhibit movements that are incompatible to one's own responses (Heyes, 2011; Brass & Heyes, 2005). Approximately half of the fMRI studies using this paradigm failed to find engagement of rTPJ and anterior mPFC. In addition, a number of studies showed engagement of regions associated with the MD network, including dorsal fronto-parietal cortex, supplementary motor area (SMA) and anterior insula (Marsh, Bird, & Catmur, 2016; Cross & Iacoboni, 2013; Mengotti, Corradi-Dell'Acqua, & Rumiati, 2012; Crescentini, Mengotti, Grecucci, & Rumiati, 2011; Bien, Roebroeck, Goebel, & Sack, 2009). Moreover, the most common measure of imitation interference is confounded by spatial compatibility or the tendency to respond faster to a stimulus when it is on the same side of space as the response (e.g., Simon, 1969). To measure imitation interference independent of spatial compatibility effects, spatial and imitative processes need to be dissociated (Gowen, Bolton, & Poliakoff, 2016; Marsh et al., 2016; Boyer, Longo, & Bertenthal, 2012; Cooper, Catmur, & Heyes, 2012; Catmur & Heyes, 2011; Wiggett, Hudson, Tipper, & Downing, 2011; Bertenthal, Longo, & Kosobud, 2006). Therefore, the extent to which imitation inhibition relies on domain-specific and domain-general architectures remains unclear. Indeed, no

Table 1. fMRI Studies Investigating Imitation Control Using Modified Versions of the Imitation Inhibition Task

| | Sample (Male:Female) | Dissociation of Imitative and Spatial Processes | Analysis | | Brain Networks ToM | | |
			ROI	Whole-brain	mPFC	rTPJ	MD
Brass et al., 2001	10 (4:6)			✓	✓		✓
Brass et al., 2005	20 (8:12)			✓	✓	✓	✓
Brass et al., 2009	20[a]		✓		✓	✓	
Spengler et al., 2009	18 (9:9)		✓	✓	✓	✓	✓
Bien et al., 2009	15 (5:10)	✓		✓			✓
Crescentini et al., 2011	19 (9:10)	✓		✓			✓
Cross & Iacoboni, 2013	24 (12:12)	✓		✓			✓
Mengotti et al., 2012	22 (10:12)	✓		✓			✓
Cross et al., 2013	25 (5:15)	✓		✓	✓		✓
Klapper et al., 2014	19 (2:17)		✓		✓[b]		
Marsh et al., 2016	24 (7:17)	✓		✓			✓
Wang et al., 2011	20 (5:15)			✓	✓	✓	✓

Evidence that the engagement of mPFC and rTPJ is inconsistent across fMRI studies that investigated imitation control using modified versions of the imitation inhibition task. For all studies, engagement of mPFC or rTPJ is reported only for contrasts that test for inhibiting the urge to automatically imitate. Engagement of the MD network is reported only for whole-brain analyses. Except for Wang et al. (2011), which used hand movements, all other tasks used modified versions of the imitation inhibition tasks involving finger movements (Brass et al., 2000). For a more detailed version of this table, see Supplementary Table S6.

[a] Number of male and female individuals not mentioned.

[b] mPFC showed engagement only at $p < .005$, uncorrected.

research to date has dissociated spatial from imitative processes and used a functional ROI (fROI) approach (Fedorenko, Duncan, & Kanwisher, 2013; Kanwisher, 2010). Using a fROI approach enables investigation of how functionally defined brain circuits, such as the MD and ToM networks, operate during the control of imitation.

A further area of imitation research that has received little attention is the extent to which imitative control varies across individuals, especially between the sexes. It has been argued that imitation is modulated by stable individual differences, such as empathy (Chartrand & Lakin, 2013) and sex (Butler et al., 2015; Sonnby-Borgström et al., 2008). Although it has been suggested that women excel across a range of social processes compared with men (Baron-Cohen, 2002), only a limited number of studies have investigated sex differences in social cognition, and the results are often mixed, do not replicate, or are specific to very select contexts or samples (Hyde, 2014; Miller & Halpern, 2014). Furthermore, studies of sex differences in social cognition have mainly focused on emotional expression perception and mental state reasoning with little emphasis placed on imitation (Krach et al., 2009; Russell, Tchanturia, Rahman, & Schmidt, 2007; Rahman, Wilson, & Abrahams, 2004; Campbell et al., 2002; Thayer & Johnsen, 2000).

A recent study that used an RT measure of imitation inhibition (Brass et al., 2000) showed that female individuals showed a greater level of interference than male individuals (Butler et al., 2015). It is possible that this sex difference in imitation control may be mediated by empathy—female individuals have been shown to be more empathetic compared with male individuals (Christov-Moore et al., 2014; Baron-Cohen & Wheelwright, 2004). However, even though empathy has been associated with different types of imitation paradigms (Müller, Leeuwen, Baaren, Bekkering, & Dijksterhuis, 2013; Sonnby-Borgström, Jönsson, & Svensson, 2003; Sonnby-Borgström, 2002; Chartrand & Bargh, 1999), the evidence to date suggests that there is no link between imitation, as measured by RTs, and empathy (Genschow et al., 2017; Butler et al., 2015). In addition, in the study by Butler and colleagues (2015), it is unclear whether sex modulates the tendency to automatically imitate or the tendency to automatically respond in the same spatial location to the observed action. The former indicates a sex difference that is specifically tied to imitation control, whereas the latter might indicate a sex difference in processes associated with resolving spatial conflict. More recent work also showed a greater interference effect for female individuals compared with male individuals (Genschow et al., 2017), as well as greater error rates for predominantly female samples than male samples (Cracco et al., 2018). The imitation task used by Genschow and colleagues (2017) was controlled for left–right spatial compatibility by presenting the stimulus hand orthogonal to the response. Even though this shows that the sex difference remains when spatial compatibility is reduced, it does not rule out the possibility of orthogonal spatial compatibility (Weeks &
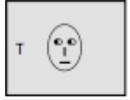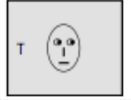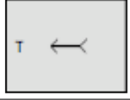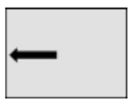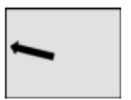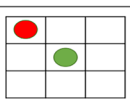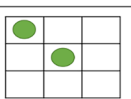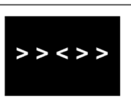
Proctor, 1990). More generally, sex differences have been found on a wide range of inhibitory control tasks, including flanker, gaze cueing, arrow cueing, oddball, and Simon tasks, wherein female individuals have been shown to require more cognitive resources than male individuals to inhibit automatic response tendencies (Figure 1; Stoet, 2010, 2017; Clayson, Clawson, & Larson, 2011; Rubia, Hyde, Halari, Giampietro, & Smith, 2010; Bayliss, di Pellegrino, & Tipper, 2005). It is possible, therefore, that a domain-general system may underpin the sex differences observed across these tasks, including during imitation control, but no research to date has directly investigated this proposal.

Across two fMRI experiments, the current study investigated functional specificity and sex differences in imitation control. Several aspects of the experimental design provide grounds to extend current understanding in meaningful and concrete ways. First, this is the first study to use independent functional localizers to identify MD and ToM networks in single subjects and directly test the involvement of these networks in imitation control. By doing so, we can directly test hypotheses regarding the role of functionally defined neural circuits (i.e., MD and ToM networks) and therefore minimize the reliance on reverse inference to infer cognitive function based on anatomical localization (Poldrack, 2006). Second, we used data from Experiment 1 to perform a power analysis to determine the sample size required to achieve a desired level of power in Experiment 2. Given the inconsistent findings in prior studies, which had relatively small sample sizes, this multiexperiment approach made sure that our key experiment had over 80% power to detect expected effect sizes. Third, to avoid spatial compatibility confounds, in Experiment 2, we used a modified version of the imitation inhibition paradigm that allowed for an independent measure of spatial and imitative compatibility (Catmur & Heyes, 2011). If the inhibition of automatic imitation relies on a domain-specific neural architecture that is associated with social cognition, as proposed by Brass and colleagues (2009), mPFC and rTPJ would be engaged in imitative control. In contrast, engagement of the MD network would suggest that domain-general processes subserve imitation control. Furthermore, the sex difference found previously (Butler et al., 2015) may be supported by differences in ToM or MD networks.

METHODS

Overview of the Experimental Approach

Experiment 1 used a group-level whole-brain analysis, which provided the basis for power analyses that set up Experiment 2 as the critical experiment with high statistical power (80%). In Experiment 2, to increase sensitivity and functional resolution, we used independent localizers to identify key functional circuits (i.e., MD and ToM networks), and analyses were performed in single subjects to precisely quantify the consistency of

	Task	Sample (Male: Female)	Trials per condition	Sex Difference – Interference	Task Requirements	Conditions	
						Compatible	Incompatible
Bayliss et al. (2005)	Gaze-cueing	80 (40:40)	144	Female > Male	Respond to letter cue		
Bayliss et al. (2005)	Arrow-cueing	40 (20:20)	144	Female > Male	Respond to letter cue		
Rubia et al. (2010)	Oddball	63 (38:25)	160 congruent, 24 oddball	Female > Male	Respond to arrow direction		
Stoet (2010)	Flanker	80 (40:40)	120	Female > Male	Respond to central circle		
Clayson et al. (2011)	Flanker	114 (60:54)	450	Female > Male	Respond to central arrow		
Butler et al. (2015)	Automatic Imitation	230 (97:133)	~30	Female > Male	Respond to number cue		
Stoet (2017)	Simon task	418 (236:182)	51	Female > Male	Respond to arrow direction		

Figure 1. Sex differences in inhibitory control tasks. Female individuals experience greater interference than male individuals in multiple inhibitory control tasks. Images are produced based on figures and description in each experiment apart from Butler et al. (2015), which are the actual images used. Also, in Rubia et al. (2010), the sex difference showed increased interference by the oddball trials rather than the incongruent trials, and this is what is represented by the images. Finally, in Butler et al. (2015), participants completed 60 trials that were 30 ± 2 trials per compatible and incompatible condition.

network engagement across individuals (Nieto-Castañón & Fedorenko, 2012; Kanwisher, 2010). Group-level analyses require responses across individuals to overlap in individual voxels. In contrast, the fROI approach allows identification of corresponding functional regions without the requirement of exact voxel overlap across individuals. Therefore, the same voxels need not be active across individuals, as long as voxels within a functionally defined ROI are consistently active across individuals. Consequently, group-level analyses may underestimate functional specificity, whereas fROI analyses can show increased sensitivity (Nieto-Castañón & Fedorenko, 2012). In addition, because of a constrained search volume, fROI analyses typically have higher statistical power than whole-brain analyses (Fedorenko, Hsieh, Nieto-Castañón, Whitfield-Gabrieli, & Kanwisher, 2010; Saxe, Brett, & Kanwisher, 2006).

Experiment 1

Participants

Twenty-eight participants (M_{age} = 23.96, SD_{age} = 5.52; 14 women) participated for monetary compensation of £15. Participants gave informed consent in line with the guidelines set by the Research Ethics and Governance Committee of the School of Psychology at Bangor University, were right-handed, had normal or corrected-to-normal vision, and reported no history of neurological damage.

Design and Procedure

All participants performed the imitation task inside the scanner. The participants also did four additional tasks in the same scanning session as part of another experiment. The scanning session started with the imitation task, followed by a run of a face perception task, a flanker task (Erikson & Erikson, 1974), another run of the face perception task, a dynamic face localizer (Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011), and a ToM localizer (Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011). The order of the tasks was counterbalanced across participants such that, of the 28 participants, 14 participants did the imitation task first, and 14 participants did the flanker task first, with the order of the other tasks remaining the same.

The imitation inhibition task. The imitation task was based on a stimulus–response compatibility paradigm developed by Brass et al. (2000) consisting of observation and execution of finger-lifting movements during fMRI scanning. Before the task, participants were instructed to hold down the "blue" and "yellow" buttons on the response box with their index and middle fingers of the right hand, respectively, A number cue (either "1" or "2") was presented to participants, and they were asked to lift their index finger on presentation of the number "1" and the middle finger for the number "2." Simultaneously, they also viewed an image of an index or middle finger lift of a left hand viewed from the third-person perspective, such that the fingers extended toward the participants. Thus, there were four trial types in an event-related design that led to two conditions—participants performing the same (congruent) or different (incongruent) finger movement to the observed hand image.

Each trial started with a fixation cross (500 msec) followed by a neutral hand (for a random ISI of 500, 700, or 1000 msec) and a hand image with an index/middle finger lift, which stayed onscreen for 2000 msec, irrespective of when the participant made the response. Sequencing the hand images in such a way led to the appearance of apparent motion of the finger. After 2000 msec, the next trial started immediately with a fixation cross (500 msec). To separately model the influence of individual events in an event-related design, the four trial types were pseudorandomized, such that each trial type was preceded by each other trial type and by itself an equal number of times (Wager & Nichols, 2003; Josephs & Henson, 1999). There were 17 trials in each block. The first trial was used to set up the randomization sequence but excluded from the analysis as it was not preceded by any other trial. The remaining 16 trials within a block were analyzed and consisted of eight trials per condition. Each run consisted of five blocks separated by a 3-, 4-, or 5-sec fixation cross. All participants completed one run of the imitation task. Thus, there were 80 trials of interest (40 congruent and 40 incongruent).

Behavioral Data Analysis

RT on the imitation inhibition task was measured as the time from number cue onset to when participants made a response. To ensure participants were engaging correctly with the task, participants who had less than 80% accuracy were removed. In addition, RTs more than 3 *SD*s away from the mean were excluded from the analyses. Furthermore, trials on which participants made an "error" were excluded from the analyses. Errors included an incorrect response, no response, a response after 2000 msec, and pressing an invalid key. The general compatibility effect was calculated as the RT difference between incompatible and compatible trials. A one-sample *t* test was performed to verify the presence of a general compatibility effect. A one-tailed independent sample *t* test was performed to determine if the compatibility effect was greater for female than male individuals. Mean differences, 95% confidence

intervals (CI), and Cohen's *d* (Cohen, 1992) are reported for all effects of interest. For the one-sample *t* test, Cohen's d_z was calculated as mean difference divided by the standard deviation of the sample (Lakens, 2013). The 95% CI is reported for the lower bound for a one-tailed *t* test. For the independent samples *t* test, Cohen's *d* was calculated as mean difference between the two groups divided by the pooled standard deviation (Cohen, 1992).

fMRI Data Analysis

Data acquisition. Participants were placed supine in a 3-T Philips MRI scanner using a SENSE 32-channel phased array coil. They were requested to avoid head motion during the scanning session and were presented stimuli on a computer screen placed behind the scanner made visible by a mirror attached to the head coil. Responses on the task were recorded with the help of a button box that recorded RTs. Thirty-five axial slices were acquired in an ascending order using a T2*-weighted EPI sequence. The reference slice for slice time correction was the slice acquired in the middle of the sequence (Slice 17). Parameters are as follows: voxel size = 3 × 3 × 4 mm, repetition time = 2000 msec, echo time = 30 msec, flip angle = 90°, slice thickness = 4 mm, slice gap = 0.8 mm, field of view = $230 \times 230 \times 167$ mm^3. One hundred seventy-four volumes were collected for the imitation task.

Four dummy scans collected at the beginning of each run of the task were not included in any analyses. A high-resolution T1-weighted anatomical image was also collected with the following parameters: repetition time = 12 msec, echo time = 3.5 msec, flip angle = 8°, number of axial slices = 170, voxel size = 1 mm^3, field of view = $250 \times 250 \times 170$ mm^3.

Data preprocessing and general linear model. Functional images were preprocessed in SPM-8. Data were realigned, unwarped, and corrected for slice timing. Data were normalized to the Montreal Neurological Institute (MNI) template with a resolution of 3 mm^3, and images were spatially smoothed (8 mm).

For the imitation task, a design matrix was fit for each participant with three regressors: one each for the correct trials of the two conditions, and one for the "new" trials (i.e., the first trial of each block). The new trials were not used in any further analyses. Stimulus onsets were time-locked to the presentation of the number cue with a duration of 0 sec and convolved with the standard hemodynamic response function.

Whole-brain analyses. Contrast images (incompatible > compatible) were calculated at the single-subject level for the imitation inhibition task to identify regions of the brain showing a compatibility effect. Group-level contrast images were created from these single-subject contrast images to identify regions that were consistently engaged for the compatibility effect across the sample using one-sample

t tests. To identify a neural signature of the sex difference in imitation inhibition, a Sex × Compatibility ANOVA was computed (female [incompatible > compatible] > male [incompatible > compatible]) as female individuals have been shown to have a higher compatibility effect than male individuals in the imitation task (Butler et al., 2015). For all analyses, contrast images were taken to the group level and thresholded using a voxel-level threshold of $p < .001$ and a voxel extent of 10 voxels. Correction for multiple comparisons was performed at the cluster level (Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994), with clusters that survive correction for multiple corrections using a family-wise error correction ($p < .05$; shown in bold font in Table 2A and B; see Results). This restricts the likelihood of false positives (Eklund et al., 2016). Clusters of activity were identified with the SPM Anatomy toolbox (Eickhoff et al., 2005).

Experiment 2

Participants

Fifty-five participants ($M_{age} = 22.04$, $SD_{age} = 3.70$; 27 women) were recruited from the Bangor community and were either reimbursed with £15 or three course credits for their participation. Informed consent was obtained in line with the guidelines set by the Research Ethics and Governance Committee of the School of Psychology at Bangor University. All participants were right-handed, did not have dyslexia or dyspraxia, were not on any medication, did not report neurological damage, and had normal or corrected-to-normal vision. The sample size was determined by a power analysis based on Experiment 1 data (see Results).

Design and Procedure

Each participant performed three tasks inside the scanner—the automatic imitation task, a ToM network localizer task, and an MD network localizer task. The order of the tasks was as follows: two runs of the MD network localizer task were interspersed between three runs of the imitation task to offset boredom. This was followed by two runs of the ToM network localizer task. The ToM task was always presented at the end to reduce the likelihood that belief reasoning during the ToM task would influence performance in the imitation task. The order was the same for all participants. Participants also completed a 50-item International Personality Item Pool questionnaire (Donnellan, Oswald, Baird, & Lucas, 2006; Goldberg, 1992; unrelated to the current study) and a stimulus rating form where they were asked to rate the hand stimulus from the imitation task as either male, female, or neutral. The entire session lasted approximately 1.5 hr, with 60 min inside the scanner. All stimulus presentation was coded in MATLAB 2015b and presented with PsychToolBox 3.0.6.

The imitation inhibition task. The automatic imitation task was similar to the one used in Experiment 1, but

with two changes. First, we used a different hand stimulus, which was rated as sex-neutral by observers. The sex of the hand was an important consideration to minimize the possibility of an own-sex bias while exploring sex differences in imitation inhibition. As such, we conducted pilot work that asked observers to evaluate a range of hand stimuli in terms of masculinity and femininity, and we selected the most sex-neutral stimulus (see Supplementary Information, Development of Stimuli). We only used one hand stimulus to simplify the design space. Although using one sex-neutral hand stimulus provided greater experimental control, it may have harmed our ability to study or elicit sex differences. Future work could probe this further by varying the sex of the stimulus and/or by using more sex-typical stimuli.

The second change that we made was to calculate an imitative compatibility effect independent of spatial compatibility (Catmur & Heyes, 2011). To do so, participants viewed an image of an index or middle finger lift of either a right or left hand but always responded with their right hand. Using right- and left-hand images produced eight trial types and four main conditions of interest (see Figure 2A). For example, when cued to lift their index finger while observing a left-hand index finger lift, the observed movement is both imitatively compatible (same finger), as well as spatially compatible (same side of space to the executed movement). In contrast, when observing a right-hand index finger lift, the participant's response is imitatively compatible (same finger), but it is not on the same side of space (they are spatially incompatible). Thus, participants performed the same (imitatively compatible) or different (imitative incompatible) finger movement on the same (spatially compatible) or different (spatially incompatible) side of space to the observed finger movement, giving rise to the following four conditions:

1. Imitatively and spatially compatible
2. Imitatively and spatially incompatible
3. Imitatively compatible and spatially incompatible
4. Imitatively incompatible and spatially compatible

Sequencing information and pseudorandomization was the same as Experiment 1. There were 65 trials in each block. The first trial was used to set up the randomization sequence but excluded from the analysis as it was not preceded by any other trial. The remaining 64 trials were analyzed, consisting of 16 trials per condition. Each run consisted of two blocks separated by a 3-sec fixation cross. All participants completed three runs of the imitation task. In total, there were 384 trials of interest, 96 per condition. Experiment 2, therefore, had more than twice the number of trials per condition than Experiment 1.

Localizer tasks

The MD network localizer. To identify regions of the MD network, a verbal working memory (WM) task was used (Fedorenko, Behr, & Kanswisher, 2011). Participants

Table 2. General Compatibility Effect and Sex × Compatibility Interaction for the Imitation Inhibition Task (Experiment 1)

Region	Cluster Size	p FWE Corr	t	MNI Coordinates		
				x	y	z
(A) General Compatibility Effect (Incompatible > Compatible)						
L IPL extending into SPL and superior frontal gyrus	**986**	**<.001**	**8.40**	**−39**	**−40**	**43**
			6.50	**−36**	**−37**	**70**
			6.38	**−27**	**−7**	**70**
L cerebellum	**150**	**.001**	**5.79**	**−21**	**−55**	**−41**
			4.95	**−30**	**−55**	**−35**
			4.72	**−9**	**−70**	**−44**
R cerebellum	**198**	**<.001**	**5.71**	**21**	**−58**	**−44**
			5.12	**45**	**−46**	**−32**
			4.32	**39**	**−55**	**−23**
R PrecG extending across superior frontal gyrus and MFG	**183**	**<.001**	**5.12**	**27**	**−1**	**70**
			5.04	**42**	**2**	**58**
			4.39	**39**	**−10**	**61**
R postcentral gyrus extending into SPL and IPL	**481**	**<.001**	**5.18**	**33**	**−40**	**73**
			4.55	**42**	**−40**	**67**
			4.50	**48**	**−34**	**37**
R posterior middle temporal gyrus	41	.179	4.61	66	−46	1
			4.13	57	−40	−8
			3.63	60	−43	13
L insula	24	.458	4.67	−36	17	−2
R posterior medial frontal cortex	20	.564	4.79	3	−4	73
L posterior medial frontal cortex	55	.083	4.40	−3	−1	52
			3.82	−6	11	52
R pallidum extending into thalamus	11	.834	4.14	21	−7	−2
			3.80	15	−6	7
L paracentral lobule	11	.834	3.89	−12	−19	79
R middle cingulate cortex	20	.564	3.85	9	14	43
			3.78	6	8	49
			3.67	6	17	52
(B) Sex × Compatibility [Female (Incompatible > Compatible) > Male (Incompatible > Compatible)]						
L SPL extending into postcentral gyrus	**93**	**.011**	**4.98**	**−21**	**−37**	**70**
			4.80	**−30**	**−19**	**46**
			4.60	**−24**	**−31**	**52**
L cerebellum	16	.679	4.34	−24	−55	−38
			4.04	−21	−55	−29

Regions surviving a voxel-level threshold of *p* < .001, and 10 voxels are reported for the (A) general compatibility effect and (B) Sex × Compatibility interaction for the imitation inhibition task. Subclusters at least 8 mm from the main peak are listed. **Bold** font indicates clusters that survive correction for multiple corrections using a family-wise error (FWE) correction (*p* < .05). MNI = Montreal Neurological Institute; SPL = superior parietal lobule; IPL = inferior parietal lobule; PrecG = precentral gyrus; MFG = middle frontal gyrus; L = left hemisphere; R = right hemisphere.

Figure 2. Stimuli for the imitation inhibition and functional localizer tasks. Stimuli and trial design for the imitation inhibition task (A), the Multiple Demand (MD) network localizer task (B), and the Theory-of-Mind network localizer task (C). For the automatic imitation task, spatial compatibility was calculated as spatial incompatible (2+3) minus (1+4), and imitative compatibility was calculated as imitatively incompatible (2+4) minus imitative compatible (1+3) trials.
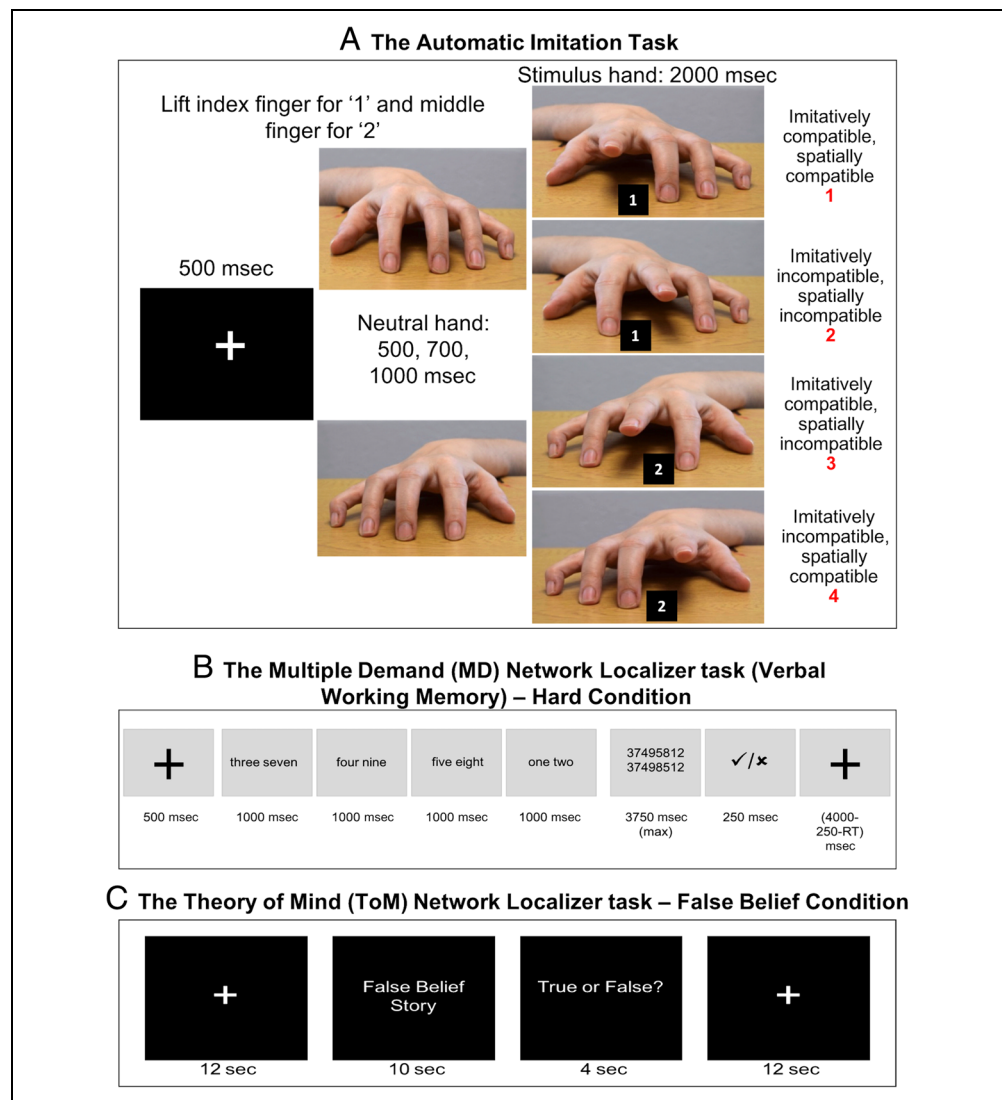
were asked to remember the sequence in which either four (easy condition) or eight (hard condition) digit sequences were presented on screen (see Figure 2B). After each trial, participants had to choose between two digit sequences presented numerically, one of which matched the sequence in which the digits were presented as words. Feedback was provided as to whether they answered correctly or incorrectly. The hard > easy contrast has been found to robustly activate regions of the MD network (Fedorenko et al., 2011, 2013). Each run consisted of 10 experimental blocks (each 34 sec long) and 6 fixation blocks (each 16 sec long). The total run lasted for 436 sec. Each participant completed two runs of the WM task.

The ToM localizer. To localize brain regions involved in mental state reasoning, we used a paradigm developed by Dodell-Feder and colleagues (2011; saxelab.mit.edu/superloc.php). This localizer task (see Figure 2C) includes 20 stories, each describing a false representation. Ten stories included out-of-date beliefs (the false belief condition), and the other 10 included out-of-date physical repre-

sentations (photographs/maps; the false photograph condition). The false belief > false photograph contrast has been shown in prior work to robustly activate regions involved in mentalizing (Dufour et al., 2013). All trials consisted of a story (10 sec), followed by a true or false question (4 sec). Each story was separated by a 12-sec rest period. The order of the stories and conditions was the same for all participants. Each participant completed two runs of this task, with five trials per condition presented in each run.

Behavioral Data Analysis

RT and accuracy were recorded in the same way as Experiment 1. Compatibility effects were calculated as follows: spatial compatibility = spatially incongruent trials − spatially congruent trials; imitative compatibility = imitatively incongruent trials − imitatively congruent trials. Behavioral data were analyzed in the same fashion as Experiment 1, only separately for imitative and spatial compatibility effects. The main aim of the experiment was to test for the presence of imitative and spatial compatibility effects,

as well as for differences between the sexes (female > male). Hence, we used a one-sample *t* test to verify the presence of spatial and imitative compatibility effects and a one-tailed independent samples *t* test to test whether female individuals showed a higher spatial/imitative compatibility effect than male individuals.

fMRI Data Analysis

Data acquisition. Data acquisition procedures were the same as Experiment 1. There were 249 volumes collected for the imitation task, 219 for the MD network localizer, and 136 for the ToM localizer for each run.

Data preprocessing and general linear model. All MRI data were preprocessed in SPM-8. Data were realigned, unwarped, and corrected for slice timing. Data were normalized to the MNI template with a resolution of 3 mm^3. Normalizing to a common space instead of the individual's native anatomical space allows for comparisons with previous studies (relying on the common space) and is preferred when definition of fROIs is based on group-constrained functional data (Nieto-Castañón & Fedorenko, 2012). Images were spatially smoothed (8 mm).

For the imitation task, a design matrix was fit for each participant with five regressors: one each for the correct trials of the four conditions and one for "new" trials (i.e., the first trial of each block). Stimulus onsets were time-locked to the presentation of the number cue with a duration of 0 sec and convolved with the standard hemodynamic response function. Contrast images were calculated for each individual participant to identify regions of the brain showing a spatial (spatially incompatible > spatially compatible) or imitative (imitatively incompatible > imitatively compatible) compatibility effect.

For the localizer tasks, the design matrix consisted of regressors for each experimental condition ("Belief" and "Photo" for the ToM localizer and "Hard" and "Easy" for

the MD localizer). The onset and duration of each condition was specified and convolved with the standard hemodynamic response function. Contrast images were then calculated for each individual subject to identify regions that responded to cognitive demand (hard > easy) and mentalizing (belief > photo).

Definition of group-constrained subject-specific analyses. For the group-constrained subject-specific (GSS) analyses, the spm_ss toolbox was used, which runs in SPM using MATLAB (web.mit.edu/evelina9/www/funcloc.html). The GSS approach developed by Fedorenko et al. (2010) and Julian, Fedorenko, Webster, and Kanwisher (2012) was used to define fROIs for each participant. These fROIs were defined using (1) each individual's activation map for the localizer tasks and (2) group-constraints or masks. These masks refer to a set of "parcels," which demarcate areas in the brain where prior work has been shown to exhibit activity for the localizer contrasts.

Two sets of fROIs were defined (Figure 3): MD network fROIs that have been known to exhibit activity for a variety of cognitive control tasks (Fedorenko et al., 2013; Duncan, 2010) and ToM network fROIs that support mentalizing and have been specifically implicated for imitation inhibition (Brass et al., 2009; Saxe & Kanwisher, 2003). For the ToM network, four parcels were derived from a group-level map from 462 participants for the false belief > false photograph contrast (Dufour et al., 2013). These regions included the dorsal, middle, and ventral mPFC (DMPFC, MMPFC, and VMPFC, respectively) and the rTPJ. For the MD network, we used 16 parcels derived from a set of functional parcels created by Idan Blank based on a probabilistic overlap map from 197 participants (available at https://evlab.mit.edu/funcloc/download-parcels). These included areas in bilateral superior and inferior parietal lobules (SPL and IPL, respectively), inferior parietal sulcus (IPS), inferior and middle frontal gyri (IFG, MFG), precentral gyrus (PrecG),
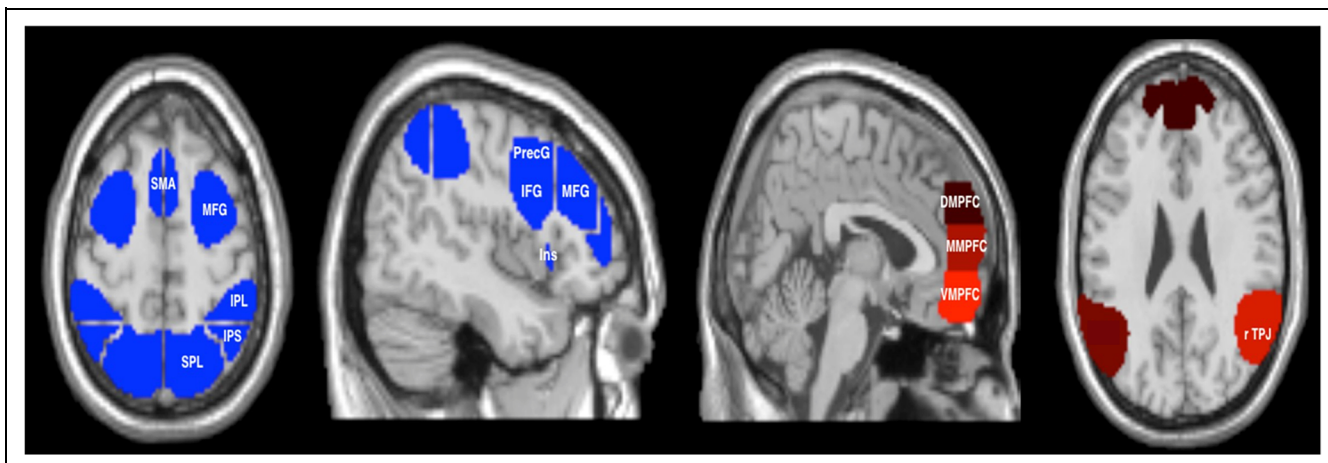
Figure 3. Graphical representation of the parcels used to define the MD and ToM network fROIs. The MD network consisted in 16 parcels, and the ToM network included 4 parcels.

insula, and the SMA. These areas were chosen for two reasons: (1) they were part of the MD network (Fedorenko et al., 2013) and (2) they have been shown to respond in prior work to the specific type of interference control of relevance to the current study (Marsh et al., 2016; see also Experiment 1).

For each individual, these masks were used to constrain the selection of subject-specific fROIs. For each individual, for the ToM network mask, the belief > photo contrast was used, and the top 10% of voxels (based on *t* values) within each parcel were defined as that individual's fROI. Similarly, for the MD network mask, each individual's top 10% of voxels (based on *t* values) in the hard > easy contrast were defined as that individual's fROI. Using the top 10% of voxels, rather than a fixed threshold (e.g., all voxels with $p < .001$), ensures a constant size of each fROI across individuals (Blank, Kanwisher, & Fedorenko, 2014). We also ran the analyses using a fixed threshold ($p < .001$, uncorrected) and found the same pattern of results (see Supplementary Tables S1.1 and S1.2). All analyses reported below are based on the top 10% of voxels that were generated from the localizer data fROIs. Percent signal change values were extracted from all fROIs. For the main analysis, all runs of the localizer tasks were used to define fROIs in each individual. Responses in these fROIs were estimated for spatial and imitative compatibility effects.

In a supplementary analysis, responses to the localizer contrasts were also estimated to ensure that all the fROIs showed the expected response with respect to the localizer contrasts. This supplementary analysis ensured that the ToM network showed a robust belief > photo and the MD network showed a robust hard > easy effect. For these localizer analyses, an across-run cross-validation approach was used (Nieto-Castañón & Fedorenko, 2012) to ensure that data used for defining fROIs were independent of data used for estimating response (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009).

As implemented in GSS, statistical tests were performed on the percent signal change values using standard Student's *t* tests. One-sample *t* tests were performed to investigate the response of the MD and ToM network fROIs to spatial and imitative compatibility effects. Based on prior behavioral findings, which showed greater RT interference for female than male individuals during imitation inhibition (Butler et al., 2015), we expected to observe sex differences in those regions that also show simple compatibility effects. That is, we expected brain regions that were generally involved in spatial and/or imitative control to show sex differences. As such, we only investigated sex difference in those fROIs that showed spatial or imitative compatibility effects. To do so, one-tailed independent samples *t* tests were performed that tested for greater engagement for female than male individuals. False discovery rate (FDR) multiple-comparison correction ($p < .05$) was used to correct for the number of fROIs in each functional network.

RESULTS

Experiment 1

Behavioral Results

A one-sample *t* test confirmed a general compatibility effect (mean = 80.02, *SE* = 8.19), $t(27) = 9.77$, $p \leq .001$, 95% CI (63.22, 96.82), Cohen's $d_z = 1.85$. A one-tailed independent samples *t* test showed no differences between male individuals (mean = 70.94, *SE* = 13.30) and female individuals (mean = 89.10, *SE* = 9.43), $t(26) = 1.114$, $p = .138$, 95% mean difference = 18.16, CI (−9.64), Cohen's $d = 0.42$. All participants had >80% accuracy; hence, all were included in the analysis. Trials on which participants made an incorrect response (0.95%) did not make a response or responded after 2000 msec (0.52%) or pressed an invalid key or responded too fast (0.09%) were excluded from the analyses.
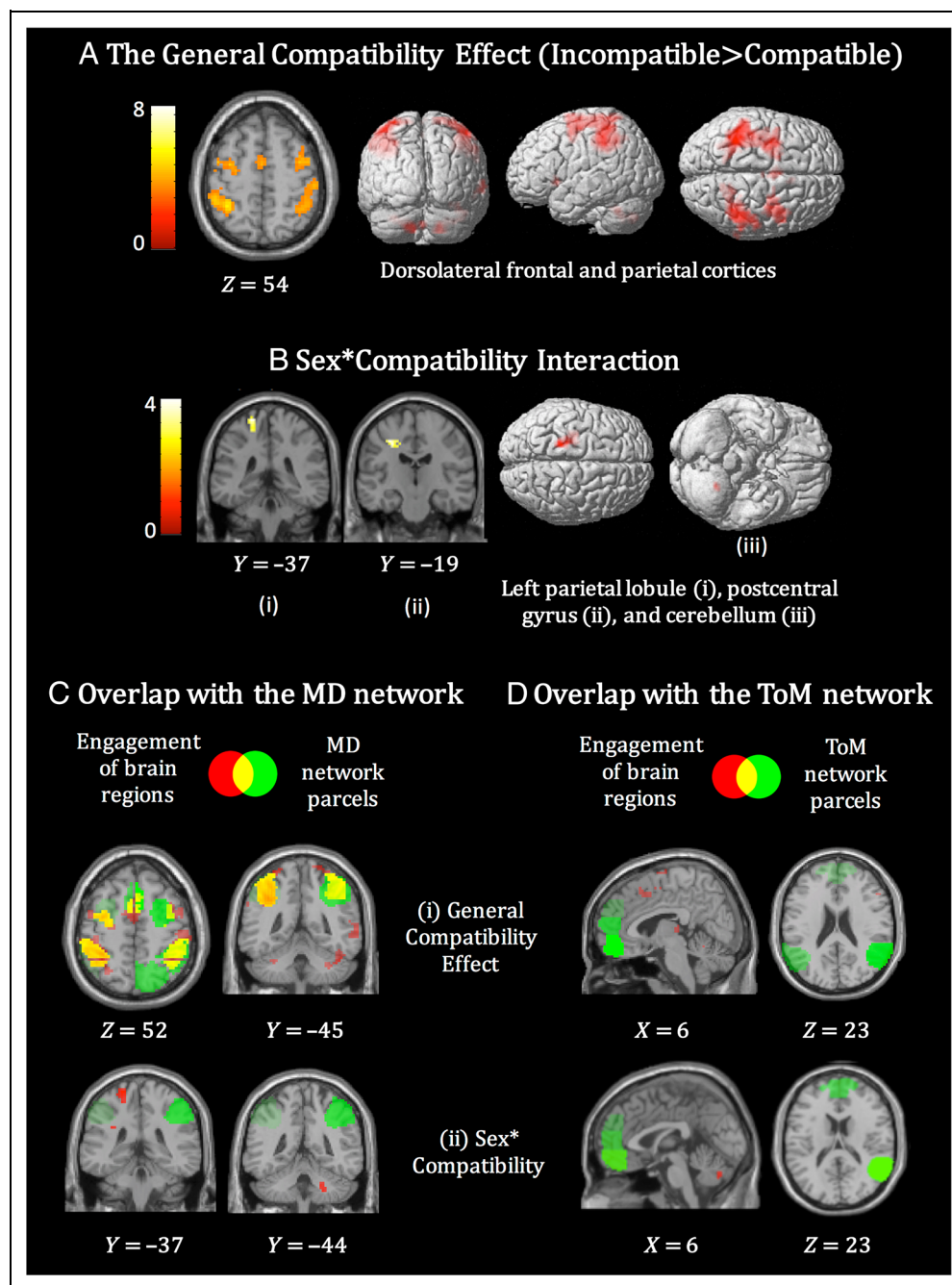
fMRI Results

In a whole-brain analysis, compatibility effects (general incompatible > general compatible) were observed in dorsomedial frontal cortex and bilaterally in dorsolateral frontal and parietal cortices (Figure 4A; Table 2A). A small volume correction (SVC) using MD and ToM network parcels was performed to restrict the search area to ToM and MD networks. Using the MD network SVC, results showed widespread activation of frontal and parietal regions, which survived correction for multiple comparisons (Figure 4A, Ci). In contrast, using the ToM network SVC, no clusters survived correction for multiple comparison, and only rTPJ showed a compatibility effect at more lenient threshold ($p < .001$, uncorrected; see Supplementary Tables S2.1 and S2.2). Anterior mPFC did not show the general compatibility effect even at this more lenient threshold.

The Sex × Compatibility interaction revealed clusters in left SPL extending into postcentral gyrus and a further cluster in the cerebellum (Figure 4B; Table 2B). No clusters emerged following an SVC analysis using the MD and ToM network masks, which demonstrate that the clusters emerging from the Sex × Compatibility interaction do not overlap with the MD or ToM networks (see Supplementary Tables S2.1 and S2.2; Figure 4Cii and Dii).

Power Analysis

We set up Experiment 1 to estimate the appropriate sample size for our critical experiment (Experiment 2). To this end, a power analysis was performed using the fMRIpower software package (fMRIpower.org; Mumford & Nichols, 2008). We performed the power analysis as follows: First, a whole-brain map of the imitation task general compatibility effect (incompatible > compatible) from Experiment 1 was entered into fMRIpower. Next, two ROIs were identified: the MD network (Duncan, 2010) and the ToM network (Saxe & Kanwisher,

Figure 4. General compatibility effect and Sex × Compatibility interaction in the imitation inhibition task (Experiment 1). (A) Results for the general compatibility effect (incompatible > compatible). Clusters emerged in the dorsal frontoparietal cortices. (B) Results for the Sex × Compatibility interaction (defined as [female (compatibility effect) > male (compatibility effect)]). Clusters emerged in the left superior parietal cortex extending into the postcentral gyrus. The MD network parcels were overlapped with the general compatibility effect (C(i)) and the Sex × Compatibility interaction (C(ii)). An overlap was found between MD parcels and regions engaged by the general compatibility effect. The ToM network parcels were also overlapped with the general compatibility effect (D(i)) and Sex × Compatibility interaction (D(ii)). No overlapping regions were found for the ToM network and regions engaged for the general compatibility effect or for the Sex × Compatibility interaction. Voxel-wise threshold used for all images was $p < .001$, $k = 10$. For a complete set of results, see Table 2 and Supplementary Tables S2.1 and S2.2.

A The General Compatibility Effect (Incompatible>Compatible)

$Z = 54$

Dorsolateral frontal and parietal cortices

B Sex*Compatibility Interaction

$Y = -37$ (i) $Y = -19$ (ii)   (iii)

Left parietal lobule (i), postcentral gyrus (ii), and cerebellum (iii)

C Overlap with the MD network

Engagement of brain regions   MD network parcels

$Z = 52$ $Y = -45$ (i) General Compatibility Effect

$Y = -37$ $Y = -44$ (ii) Sex* Compatibility

D Overlap with the ToM network

Engagement of brain regions   ToM network parcels

$X = 6$ $Z = 23$ (i) General Compatibility Effect

$X = 6$ $Z = 23$ (ii) Sex* Compatibility

2003). The MD and ToM network masks used were the same as in Experiment 2 (see Methods). As recommended, we corrected the alpha value by the number of ROIs (0.05/2 = 0.025) before performing power analyses (Mumford, 2012).

Results from these power analyses showed that testing 50 participants in Experiment 2 would provide 80% power to detect effects as large as (or larger than) the average effect size that was observed across all nodes in the MD network in Experiment 1 (Cohen's $d = 0.4$, mean signal change = 0.23, $SD = 0.58$). We did not have the same level of power to detect smaller effects than these, such as those observed in the ToM network in Experiment 1.

Indeed, the effects in the ToM network in Experiment 1 were so small that we would have needed an impractically large sample size to achieve 80% power. As such, in Experiment 2, we decided to test participants until we had 50 usable data sets.

Design differences between Experiments 1 and 2 are worth considering when interpreting these power calculations because we may be underestimating the power of our design in Experiment 2. The toolbox used to run power calculations (fmripower.org) can only estimate power for a future experiment with the same design as the current data set (Mumford & Nichols, 2008). However, the designs of Experiments 1 and 2 differed in

Table 3. Responses in Each MD Network fROI for Spatial and Imitative Compatibility

ROI	ROI Size	Intersubject Overlap	Average ROI Mask Size (Voxels)	Spatial Compatibility			Imitative Compatibility		
				t	p	p-FDR	t	p	p-FDR
L_SPL	1173	1	117	2.00	.026	**.028**	1.13	.131	.191
L_IPS	287	1	28	2.00	.026	**.028**	1.96	.028	.089
L_IPL	641	1	64	2.72	.005	**.019**	2.05	.023	.089
L_MFG	536	1	53	2.16	.018	**.028**	0.53	.301	.324
L_PrecG	338	1	33	2.17	.018	**.028**	0.91	.184	.227
L_IFG	181	1	18	1.83	.040	**.037**	1.53	.066	.118
L_Insula	197	1	19	2.78	.004	**.019**	0.52	.304	.324
L_SMA	294	1	29	2.52	.008	**.020**	0.39	.349	.349
R_SPL	1181	1	118	2.30	.013	**.026**	1.56	.062	.118
R_IPS	227	1	22	2.03	.024	**.028**	2.30	.013	.069
R_IPL	599	1	59	2.65	.005	**.019**	2.50	.008	.063
R_MFG	535	1	53	3.57	<.001	**.006**	1.55	.064	.118
R_PrecG	269	1	26	2.43	.009	**.021**	1.59	.060	.118
R_IFG	265	1	26	2.61	.006	**.019**	2.53	.007	.063
R_Insula	184	1	18	2.09	.021	**.028**	1.24	.120	.175
R_SMA	328	1	32	2.30	.022	**.028**	1.06	.148	.198

For each individual, for the MD network mask, the hard > easy contrast was used, and the top 10% of voxels (based on *t* values) within each parcel were defined as that individual's fROI. Uncorrected *p* values as well as FDR-corrected *p* values are reported. Cells in **bold** are fROIs that survive correction for multiple comparisons (*p* < .05, FDR-corrected).

two ways. First, Experiment 1 measured a general compatibility effect, whereas in Experiment 2, we broke this effect down into spatial and imitative compatibility effects. Second, Experiment 2 had more than double the amount of trials per condition as Experiment 1. Therefore, the primary contrast used to determine power was not identical to the contrast used in Experiment 2, but due to a greater number of trials per condition to estimate the effects of interest, we may underestimate power in Experiment 2. Given the lack of sex differences in Experiment 1 in our ROIs, we did not have sufficient power to convincingly investigate neural differences between male and female individuals in Experiment 2. However, given our a priori predictions regarding sex, we continue to report sex difference analyses throughout the article.

Experiment 2

Behavioral Results

The hand stimulus used in Experiment 2 for the imitation inhibition task was perceived as "neutral" by most participants (mean$_{rating}$ = 5.20, SD_{rating} = 2.04; rated on a scale of 1–9, where 1 = *most masculine*, 5 = *neutral*, and 9 = *most feminine*). To ensure participants were engaging

correctly with the task, runs on which participants had less than 80% accuracy (two runs of one participant) were removed. In addition, RTs more than 3 *SD*s away from the mean (two runs of one participant and one run of another participant) were excluded from the analyses. Furthermore, trials on which participants made an incorrect response (1.52%), did not make a response or responded after 2000 msec (0.61%), or pressed an invalid key (0.22%) were also excluded from the analyses. Figure 5 shows the imitative and spatial compatibility effects for both the sexes. For RT data, see Supplementary Table S3.

Spatial compatibility. A one-sample *t* test confirmed a spatial compatibility effect (mean = 41.94, *SE* = 2.87), *t*(54) = 14.618, *p* ≤ .001, 95% CI (36.19, 47.69), Cohen's d_z = 1.97. A one-tailed independent samples *t* test evidenced a greater spatial interference effect for female individuals (mean = 50.98, *SE* = 3.67) as compared with male individuals (mean = 33.20, *SE* = 3.75), *t*(53) = −3.38, *p* < .001, mean difference = 17.76, 95% CI (8.91); Cohen's *d* = 0.91.

Imitative compatibility. A one-sample *t* test showed a significant imitative compatibility effect (mean = 15.37, *SE* = 2.86), *t*(54) = 5.37, *p* < .001, 95% CI (9.63, 21.11), Cohen's d_z = 0.72. There was no significant difference
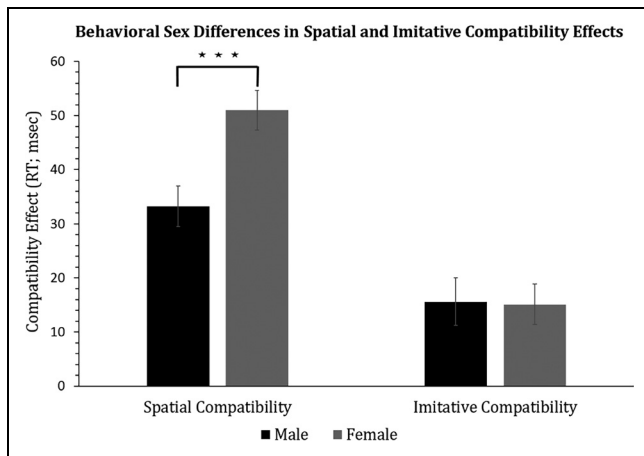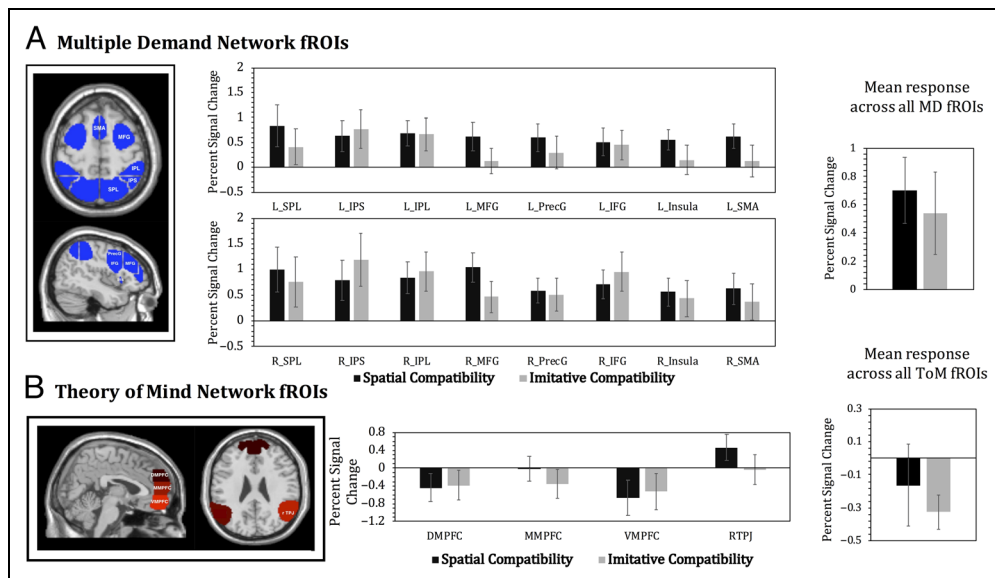
Figure 5. Behavioral sex differences in imitative and spatial compatibility effects. The spatial and imitative compatibility effects (RTs) in male and female individuals displayed in milliseconds. Error bars denote standard error of mean.

between male individuals (mean = 15.62, SE = 4.39) and female individuals (mean = 15.11, SE = 3.73), $t(53) = 0.09$, $p = .465$, mean difference = -0.51, 95% CI (-10.18), Cohen's $d = 0.02$.

fMRI Results

Five participants were excluded from the fMRI analyses due to lower than 80% accuracy in two runs of the imitation task and the MD network localizer task ($n = 1$) and excessive head motion ($n = 4$; displacement > 4 mm) in all runs of the imitation task and/or all runs of either of the localizer tasks. Thus, the final sample consisted of 50 participants ($M_{age} = 22.26$, $SD_{age} = 3.71$; 24 female). From these 50 participants, two sessions of the imitation task were also excluded for one participant due to excessive head motion

and one participant's data for one session of the imitation task could not be used because the data file was corrupted.

Localizer Tasks

All fROIs showed the predicted responses to the localizer contrasts (as estimated using data not used for defining ROIs; see Methods). All the MD network fROIs showed a robust hard > easy effect ($ts > 9.13$, $ps < .0001$), and ToM network fROIs showed a robust belief > photo effect ($ts > 5.70$, $ps < .0001$). For responses for each individual fROI separately, see Supplementary Tables S4.1 (MD) and S4.2 (ToM).
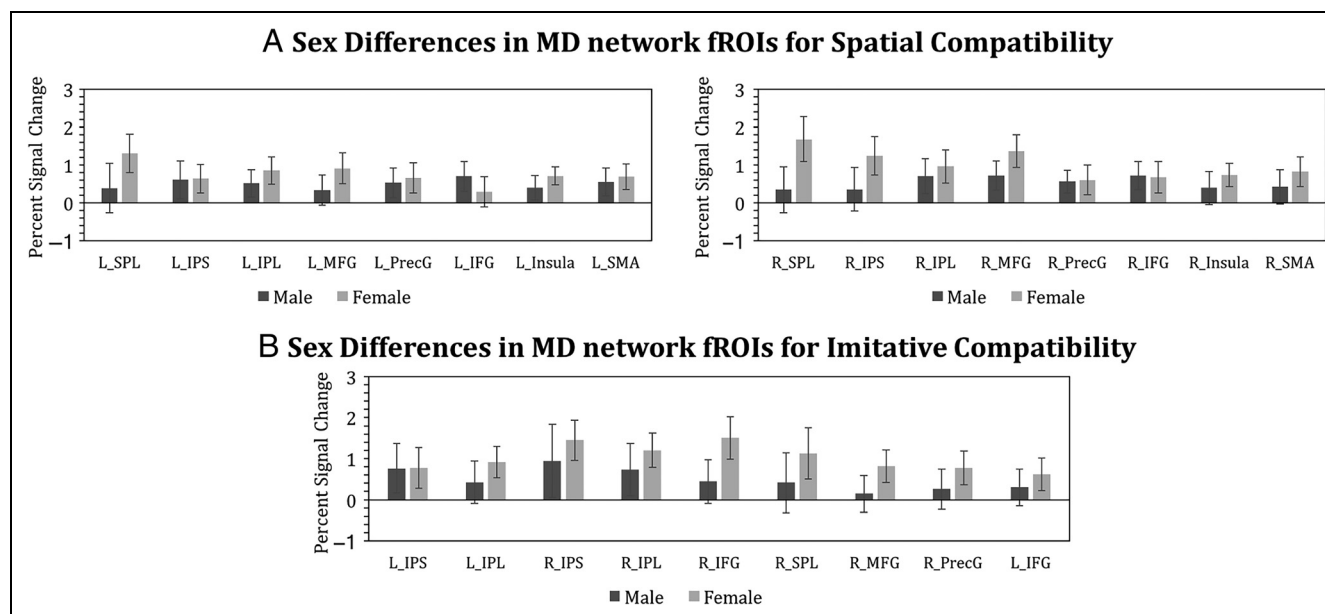
The Automatic Imitation Task

GSS analyses. Figure 6 shows the mean percent signal change for each fROI in the MD and ToM networks for spatial (spatial incompatible > spatial compatible) and imitative compatibility (imitative incompatible > imitative compatible) effects.

MD network fROIs

Spatial compatibility. All 16 fROIs of the MD network showed a spatial compatibility effect ($ts > 1.8$, $ps < .04$; Figure 6A, Table 3), which survived correction for multiple comparisons ($p < .05$, FDR-corrected). The mean percent signal change across the MD network for spatial compatibility was 0.70, $SD = 1.66$, Cohen's $d = 0.42$. No significant differences were found between male and female individuals in percent signal change values in any of the fROIs ($ts < 1.6$, $ps > .1$), except right SPL which approached significance ($p = .062$; Figure 7A).

Imitative compatibility. None of the 16 MD network fROIs showed an imitative compatibility effect, which

Figure 6. The parcels used to define individual fROIs and the responses to spatial and imitative compatibility effects in the MD (A) and ToM (B) network fROIs are shown. Error bars denote standard error of mean. All MD network fROIs were sensitive to spatial compatibility effects (FDR-corrected, $p < .05$). Bilateral IPL, bilateral IPS, and the right IFG showed a significant response for imitative compatibility effects, but at an uncorrected threshold of $p < .001$. No ToM network fROIs showed engagement for either spatial or imitative compatibility effects.

Figure 7. Responses to spatial (A) and imitative compatibility (B) effects separately for male and female participants in the MD network. Error bars denote standard error of mean. None of the fROIs showed a sex difference either in imitative or spatial compatibility that survived correction for multiple comparisons.

survived correction for multiple comparisons (all $ps > .05$, FDR-corrected). Five MD network fROIs showed an imitative compatibility effect at an uncorrected threshold ($ts > 1.95$, $ps < .05$). These fROIs include bilateral IPL, bilateral IPS, and the right IFG (Figure 6A, Table 3). Four further fROIs showed an imitative compatibility effect that approached significance, which included left IFG ($p = .07$), right SPL, right MFG, and right PrecG ($p = .06$). The mean percent signal change across the MD network for imitative compatibility was 0.54, $SD = 2.06$, Cohen's $d = 0.26$. There was no significant difference between male and female individuals in any of these fROIs ($ts < 1.5$, $ps > .08$; see Figure 7B).

ToM network fROIs. None of the ToM network fROIs showed imitative ($ts < 1.3$, $ps > .50$) or spatial ($ts < 1.6$, $ps > .06$) compatibility effects, even at an uncorrected significance threshold (Figure 6B, Table 4). rTPJ showed a spatial compatibility effect that approached

significance ($p = .065$). The mean percent signal change across the ToM network for spatial compatibility was -0.16, $SD = 1.88$, Cohen's $d = -0.08$, and the mean percent signal change across the ToM network for imitative compatibility was -0.32, $SD = 2.02$, Cohen's $d = -0.16$.

Whole-brain analyses. For completeness and for use in future meta-analyses, we also computed group-level whole-brain analyses separately for general, spatial, and imitative compatibility effects, as well as for Sex × Compatibility interactions (see Supplementary Table S4).

Open Science

Data for Experiments 1 and 2 are freely available online including behavioral and fROI data (osf.io/45x6z), as well as whole-brain t maps (https://neurovault.org/collections/3218).

Table 4. Responses in Each ToM Network fROI for Spatial and Imitative Compatibility

ROI	ROI Size	Intersubject Overlap	Average ROI Mask Size (Voxels)	Spatial Compatibility			Imitative Compatibility		
				t	p	p-FDR	t	p	p-FDR
DMPFC	576	1	57	-1.38	.913	.951	-1.167	.876	.898
MMPFC	494	1	49	-0.043	.517	.951	-1.081	.857	.898
VMPFC	382	1	38	-1.690	.951	.951	-1.286	.898	.898
rTPJ	1018	1	101	1.543	.065	.258	-0.106	.542	.898

For each individual, for the ToM network mask, the Belief > Photo contrast was used, and the top 10% of voxels (based on t values) within each parcel were defined as that individual's fROI. Uncorrected p values as well as FDR-corrected p values are reported.

DISCUSSION

The current study provides the most robust neuro-imaging evidence to date for a lack of functional specificity in the neural circuits supporting the inhibition of automatic imitation. With higher statistical power and functional sensitivity than prior studies, across two experiments the results demonstrate that imitation inhibition engages a domain-general neural network as opposed to a brain network that supports social cognition. As such, models of imitation control need updating to include an increased role for domain-general processes and a reduced or altered role for domain-specific processes. Furthermore, in terms of behavior, female individuals showed a higher spatial but not imitative compatibility effect than male individuals. However, there was no sex difference in the neural mechanisms underlying spatial or imitation control, which suggests that further exploration of sex differences in inhibitory control is required.

Functional Specificity in Imitation Inhibition

Our findings show that brain regions that are engaged in a verbal WM task, which are associated with the operation of the MD network (Fedorenko et al., 2013; Duncan, 2010), are also engaged during spatial and imitative conflict resolution. These results support the involvement of a domain-general cognitive and neural system during the control of imitation. By contrast, brain regions that are engaged in a belief reasoning task, which are associated with the operation of the ToM network (Van Overwalle, 2009; Saxe & Kanwisher, 2003; Frith & Frith, 1999), show no engagement during the inhibition of imitation. As such, we provide no evidence for domain specificity in cognitive and neural systems that control imitation.

Brass and colleagues (2009) proposed that, in the context of imitation control, rTPJ is involved in self–other distinction, and mPFC enforces the self-generated action over the observed action. Our findings are inconsistent with the hypothesis that a specific neural system related to social cognition is engaged in the inhibition of automatic imitative tendencies. mPFC and rTPJ have both been implicated in imitation inhibition by some studies (Wang et al., 2011; Brass et al., 2009; Spengler et al., 2009; Brass, Derrfuss, & von Cramon, 2005). In contrast, other studies found engagement of mPFC only (Cross et al., 2013; Brass et al., 2001) or of domain-general regions rather than mPFC and rTPJ (Marsh et al., 2016; Cross & Iacoboni, 2013; Crescentini et al., 2011; Bien et al., 2009). In both experiments in the current study, we had larger sample sizes than prior experiments, and in Experiment 2, we had sufficient statistical power to be confident in detecting effects as large as previously observed in mPFC and rTPJ, should they exist. Taken together with prior findings (Table 1), we suggest that, during the inhibition of imitation, the consistency of mPFC and rTPJ engagement across individuals is relatively low, whereas the consistency of MD network engagement across individuals is relatively high.

These results have potential implications for self–other control theories of social cognition more generally. Mostly based on imitation research, which previously suggested that mPFC and rTPJ are engaged in imitation inhibition, self–other control is thought to be a candidate mechanism for a diverse set of social functions (de Guzman, Bird, Banissy, & Catmur, 2016; Sowden & Shah, 2014; Brass et al., 2009). For example, self–other control processes have been linked to autism, empathy, and theory of mind (de Guzman et al., 2016; Sowden & Shah, 2014; Spengler et al., 2009). However, recent behavioral findings, which used larger sample sizes than prior work and meta-analytical approaches, do not support the view that the control of imitation varies as a function of social disposition as indexed by autistic-like traits and empathy (Cracco et al., 2018; Genschow et al., 2017; Butler et al., 2015). In light of these recent behavioral results, the lack of engagement of mPFC and rTPJ in the current study raises an important question about the reliance of imitation inhibition on a self–other distinction. One possibility is that, instead of a distinctly social mechanism (Bertenthal & Scheutz, 2013; Boyer et al., 2012), inhibiting imitative tendencies may involve the same cognitive processes that are used when inhibiting other nonsocial external influences (Cooper et al., 2012; Heyes, 2011).

Alternatively, the engagement of mPFC and rTPJ during self–other control processes may be more complicated than current models of social cognition suggest. Indeed, a small number of neurostimulation studies have shown that modulation to rTPJ can influence performance on RT measures of imitation (Sowden & Catmur, 2015; Hogeveen et al., 2014). In addition, mPFC and rTPJ have been found to be involved in the modulation of automatic imitation. For example, Klapper et al. (2014) found a higher response in rTPJ when an interaction partner looked human and was believed to be human compared with when neither of these animacy cues was present. Wang and colleagues (2011) demonstrated that mPFC had a top–down influence on other brain circuits during social modulation of imitation via direct gaze. These studies suggest that mPFC and/or rTPJ may have a regulatory role, be sensitive to social context, and be functionally connected to other regions during the inhibition of automatic imitation. Indeed, regions that do not show direct engagement in a cognitive process of interest have been known to have a regulatory influence on other regions that are directly engaged (Burnett & Blakemore, 2009). In line with this proposal, Cross and colleagues (2013) suggested that imitation control involves top–down regulation between a domain-general cognitive control network and a domain-specific network relevant for imitation. More generally, research from other domains of social cognition shows growing evidence for higher complexity and functional interplay within and between so-called domain-specific

and domain-general networks (Spunt & Adolphs, 2015; Baetens, Ma, Steen, & Van Overwalle, 2014; Quadflieg et al., 2011; Zaki, Hennigan, Weber, & Ochsner, 2010). These studies suggest that models including neat divisions between these networks may be an overly simplistic characterization of mental function (Michael & D'Ausilio, 2015; Barrett, 2012). Much like social cognition in general, therefore, imitation control may be best explained by interactions between component functional circuits, which themselves need not be domain-specific (Spunt & Adolphs, 2017). A crucial direction for future research is testing for more complex models of imitation, which may involve connectivity in and between regions of the MD and ToM networks.

An important point to note, however, is that any conclusions made regarding possible domain specificity of mPFC and rTPJ are based on the assumption that mPFC and rTPJ are at least partly specialized for social cognition (Brass et al., 2009). Recent evidence suggests that mPFC and rTPJ may be functionally versatile in the sense that they show general cognitive properties, which may not be specific to social cognition (Dugué, Merriam, Heeger, & Carrasco, 2017; Schurz, Tholen, Perner, Mars, Sallet, 2017; Schuwerk, Schurz, Müller, Rupprecht, & Sommer, 2017; de la Vega, Chang, Banich, Wager, & Yarkoni, 2016; Carter & Huettel, 2013; Alexander & Brown, 2011; Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). Thus, the argument that the engagement of mPFC and rTPJ in imitation inhibition may be specific to social cognition might need further validation. In addition, social cognition itself has been broken down in "bottom–up" and "top–down" domains (Zaki & Ochsner, 2012). The bottom–up domain refers to prereflective processes that are fast and stimulus driven, whereas the top–down domain maps on reflective, cognitively laborious, and flexible processes (Bohl & van den Bos, 2012). When extended to imitation control, prior research has consistently implicated regions involved in top–down control for automatic imitation (Brass et al., 2009). However, recent studies suggest that imitation control (and social cognition more broadly) relies on interactions between bottom–up and top–down processes (Christov-Moore, Conway, & Iacoboni, 2017; Cross & Iacoboni, 2014; Bohl & van den Bos, 2012). Thus, another important avenue for future research would be to investigate imitation control based on bottom–up and top–down processes and their interactions, rather than considering these processes as mutually exclusive.

Nonetheless, results from the current study remain clear: The basic imitation inhibition mechanism engages the MD network, which has been consistently associated with domain-general processes (Duncan, 2010). Given the mixed findings in prior imitation studies (Table 1) as well as in psychology and neuroscience more generally (Open Science Framework, 2015; Button et al., 2013), future fMRI research may also consider reliability and reproducibility as key concerns in imitation research and consider the possible use of fROI approaches as a means to quantify consistency across individuals.

Sex Differences in Imitation Inhibition

This study is the first to investigate sex differences in the neural mechanisms that inhibit imitation. The behavioral data demonstrated that female individuals show a greater spatial but not imitative compatibility effect than male individuals. This result extends prior behavioral research on sex differences, which did not separate spatial (or orthogonal spatial) from imitative responses in imitation control (Genschow et al., 2017; Butler et al., 2015). The result is also consistent with reports in a wide range of nonsocial inhibitory control tasks, which show similar sex differences (Figure 2; Stoet, 2010, 2017; Clayson et al., 2011; Rubia et al., 2010; Bayliss et al., 2005). All these tasks share a common feature—they require the inhibition of a response to a task-irrelevant spatial feature to enforce a task-relevant response. Taken together, this pattern of results suggests that response inhibition relating to spatial conflict differs between the sexes, rather than a process that is tied to the control of imitation. An alternative possibility is that the difference between the sexes for spatial compatibility is larger than for imitative compatibility, and we were unable to detect the imitative effect behaviorally. Future research will have to probe these possibilities further.

Given the proposed role of MD and ToM networks in imitation control, we anticipated sex differences in one or both of these networks. The neuroimaging data, however, demonstrated no sex differences in the ToM or MD networks in either experiment. Furthermore, even though regions outside our ROIs mediated the sex difference in Experiment 1, these regions were not consistently engaged differently for male and female individuals in Experiment 2. Thus, based on data across both experiments, our best estimate is that univariate analyses, which assess the magnitude of BOLD response, do not show large effects of sex in MD or ToM neural networks. This being said, there does seem to be a trend for greater engagement in the MD network for female individuals compared with male individuals for both spatial and imitative effects, but this does not survive our statistical thresholding (Figure 7). As a consequence, we are cautious to interpret this null result as we did not have the same level of statistical power to detect sex differences as we did to detect simple compatibility effects. Indeed, it remains a possibility that small univariate effects exist or that the sex difference is underpinned by more complex neural organization. Future studies that use connectivity measures (Sporns, Tononi, & Kötter, 2005) or multivoxel pattern analysis (Kriegeskorte, Mur, & Bandettini, 2008; Norman, Polyn, Detre, & Haxby, 2006) may show increased sensitivity and be better able to capture the complexity of neural organization that we are aiming to measure.

Limitations

The primary limitation of the current work is that we studied a relatively simple model of brain organization

based on univariate measures. Given the mixed evidence from prior studies regarding imitation control (Table 1), we felt it was an important step to first establish the extent to which general and specific systems were engaged in a univariate manner. By doing so, we aimed to build an appropriate foundation for future work to build upon. Moreover, as we only identified the MD and ToM networks, it is possible that neural regions outside our key networks may play a role in imitation inhibition or mediate the sex difference in spatial response inhibition or imitation control. Even though our whole-brain analyses showed no consistent effects outside our fROIs, this only shows that there was no univariate engagement of extended brain regions. We thus acknowledge that we have tested a relatively simple model of brain organization that is likely to underestimate the complexity of neural processes associated with social and cognitive mechanisms such as imitation control. As mentioned before, future work may consider interactions between general and specific systems and more complex, multivariate measures of brain organization.

A second limitation regards the functional localization approach used to identify the ToM and MD networks in Experiment 2. The validity of the fROI approach is based on assumptions about the functional processes that are engaged by the localizers used to identify fROIs. For example, different ToM localizers may engage partly nonoverlapping aspects of the ToM network (Schaafsma, Pfaff, Spunt, & Adolphs, 2015; Spunt & Adolphs, 2014). Therefore, our conclusions about the role of ToM and MD networks are limited to the type of localizer paradigms that we used in the current study. Future research that uses different functional partitions of these networks would be instructive.

A final potential limitation is that the order of tasks in Experiment 2 could have influenced our results. We ordered the tasks such that the ToM localizer was always performed at the end, but the MD task was interspersed between imitations runs in order to offset boredom. We arranged blocks in this manner because we were primarily concerned that asking people to perform a belief reasoning task would introduce a social bias to treat the person (hand image) in an artificially more social/belief reasoning manner during the imitation inhibition task. We did not share the same level of concern that performing a memory task, which we used to localize the domain-general system, would introduce a memory or "cognitive control" bias to the imitation inhibition task. However, we cannot rule out the possibility in the current experiment that the MD task influenced the way the imitation task was performed. This being said, we did get the same results in Experiment 1, when the MD task was not performed before the imitation task. As such, although possible, we find it unlikely that task order had a meaningful impact on our results in Experiment 2.

REFERENCES

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience, 14,* 1338–1344.

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

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences, 18,* 177–185.

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

Bardi, L., Gheza, D., & Brass, M. (2017). TPJ-M1 interaction in the control of shared representations: New insights from tDCS and TMS combined. *Neuroimage, 146,* 734–740.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6,* 248–254.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163–175.

Barrett, H. C. (2012). A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences, U.S.A., 109,* 10733–10740.

Bayliss, A. P., di Pellegrino, G., & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *Quarterly Journal of Experimental Psychology, 58A,* 631–650.

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 210.

Bertenthal, B. I., & Scheutz, M. (2013). In praise of a model but not its conclusions: Commentary on Cooper et al. (2012). *Cognitive Science, 37,* 631–641.

Bien, N., Roebroeck, A., Goebel, R., & Sack, A. T. (2009). The brain's intention to imitate: The neurobiology of intentional versus automatic imitation. *Cerebral Cortex, 19,* 2338–2351.

Blank, I., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology, 112,* 1105–1118.

Bohl, V., & van den Bos, W. (2012). Toward an integrative account of social cognition: Marrying theory of mind and interactionism to study the interplay of Type 1 and Type 2 processes. *Frontiers in Human Neuroscience, 6,* 274.

Boyer, T. W., Longo, M. R., & Bertenthal, B. I. (2012). Is automatic imitation a specialized form of stimulus–response compatibility? Dissociating imitative and spatial compatibilities. *Acta Psychologica, 139,* 440–448.

Brass, M., Bekkering, H., Wohlschlager, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44,* 124–143.

Brass, M., Derrfuss, J., Matthes-von Cramon, G., & von Cramon, D. Y. (2003). Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology, 17,* 265–271.

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia, 43,* 89–98.

Brass, M., & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences, 9,* 489–495.

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 364,* 2359–2367.

Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *Neuroimage, 14,* 1416–1423.

Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. E. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage, 17,* 1562–1571.

Burnett, S., & Blakemore, S. J. (2009). Functional connectivity during a social emotion task in adolescents and in adults. *European Journal of Neuroscience, 29,* 1294–1301.

Butler, E. E., Ward, R., & Ramsey, R. (2015). Investigating the relationship between stable personality characteristics and automatic imitation. *PLoS One, 10,* e0129651.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., et al. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14,* 365–376.

Campbell, R., Elgar, K., Kuntsi, J., Akers, R., Terstegge, J., Coleman, M., et al. (2002). The classification of "fear" from faces is associated with face recognition skill in females. *Neuropsychologia, 40,* 575–584.

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

Catmur, C., & Heyes, C. (2011). Time course analyses confirm independence of imitative and spatial compatibility. *Journal of Experimental Psychology. Human Perception and Performance, 37,* 409–421.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76,* 893–910.

Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioural mimicry. *Annual Review of Psychology, 64,* 285–308.

Christov-Moore, L., Conway, P., & Iacoboni, M. (2017). Deontological dilemma response tendencies and sensorimotor representations of harm to others. *Frontiers in Integrative Neuroscience, 11,* 34.

Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience & Biobehavioral Reviews, 46,* 604–627.

Clayson, P. E., Clawson, A., & Larson, M. J. (2011). Sex differences in electrophysiological indices of conflict monitoring. *Biological Psychology, 87,* 282–289.

Cohen, J. (1992). A power primer. *Quantitative Methods in Psychology, 112,* 155–159.

Cooper, R. P., Catmur, C., & Heyes, C. (2012). Are automatic imitation and spatial compatibility mediated by different processes? *Cognitive Science, 37,* 605–630.

Cracco, E., Bardi, L., Desmet, C., Rigoni, D., Radkova, I., Deschrijver, E., et al. (2018). Automatic imitation: A meta-analysis. *Psychological Bulletin.* doi:10.1037.bul0000143.

Crescentini, C., Mengotti, P., Grecucci, A., & Rumiati, R. I. (2011). The effect of observed biological and non-biological movements on action imitation: An fMRI study. *Brain Research, 1420,* 80–92.

Cross, K. A., & Iacoboni, M. (2013). Optimised neural coding? Control mechanisms in large cortical networks implemented by connectivity changes. *Human Brain Mapping, 34,* 213–225.

Cross, K. A., & Iacoboni, M. (2014). To imitate or not: Avoiding imitation involves preparatory inhibition of motor resonance. *Neuroimage, 91,* 228–236.

Cross, K. A., Torrisi, S., Losin, E. A. R., & Iacoboni, M. (2013). Controlling automatic imitative tendencies: Interactions between mirror neuron and cognitive control systems. *Neuroimage, 83,* 493–504.

de Guzman, M., Bird, G., Banissy, M. J., & Catmur, C. (2016). Self–other control processes in social cognition: From imitation to empathy. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 371,* 20150079.

de la Vega, A., Chang, L. J., Banich, M. T., Wager, T. D., & Yarkoni, T. (2016). Large-scale meta-analysis of human medial frontal cortex reveals tripartite functional organization. *Journal of Neuroscience, 36,* 6553–6562.

Dimberg, U., & Lundquist, L. O. (1990). Gender differences in facial reactions to facial expressions. *Biological Psychology, 30,* 151–159.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage, 55,* 705–712.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18,* 192–203.

Dufour, N., Redcay, E., Young, L., Mavros, P. L., Moran, J. M., Triantafyllou, C., et al. (2013). Similar brain activation during false belief tasks in a large sample of adults with and without autism. *PLoS One, 8,* e75468.

Dugué, L., Merriam, E. P., Heeger, D. J., & Carrasco, M. (2017). Specific visual subregions of TPJ mediate reorienting of spatial attention. *Cerebral Cortex, 13,* 1–16.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences, 14,* 172–179.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25,* 1325–1335.

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

Erikson, B. A., & Erikson, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16,* 143–149.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences, U.S.A., 108,* 16428–16433.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 110,* 16616–16621.

Fedorenko, E., Hsieh, P. J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations

of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology, 104,* 1177–1194.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping, 1,* 210–220.

Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science, 286,* 1692–1695.

Genschow, O., van Den Bossche, S., Cracco, E., Bardi, L., Rigoni, D., & Brass, M. (2017). Mimicry and automatic imitation are not correlated. *PLoS One, 12,* e0183784.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26–42.

Gowen, E., Bolton, E., & Poliakoff, E. (2016). Believe it or not: Moving non-biological stimuli believed to have human origin can be represented as human movement. *Cognition, 146,* 431–438.

Hazeltine, E., Poldrack, R., & Gabrieli, J. D. E. (2007). Neural activation during response competition. *Journal of Cognitive Neuroscience, 12,* 118–129.

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin, 137,* 463.

Hirschfeld, L. A., & Gelman, S. A. (1994). *Mapping the mind: Domain specificity in cognition and culture.* Cambridge, UK: Cambridge University Press.

Hogeveen, J., Obhi, S. S., Banissy, M. J., Santiesteban, I., Press, C., Catmur, C., et al. (2014). Task-dependent and distinct roles of the temporoparietal junction and inferior frontal cortex in the control of imitation. *Social Cognitive and Affective Neuroscience, 10,* 1003–1009.

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65,* 373–398.

Iacoboni, M. (2009). Neurobiology of imitation. *Current Opinion in Neurobiology, 19,* 661–665.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286,* 2526–2528.

Josephs, O., & Henson, R. N. (1999). Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 354,* 1215–1228.

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage, 60,* 2357–2364.

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 11163–11170.

Klapper, A., Ramsey, R., Wigboldus, D., & Cross, E. S. (2014). The control of automatic imitation based on bottom–up and top–down cues to animacy: Insights from brain and behavior. *Journal of Cognitive Neuroscience, 26,* 2503–2513.

Krach, S., Blumel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., et al. (2009). Are women better mindreaders? Sex differences in neural correlates of mentalizing detected with functional MRI. *BMC Neuroscience, 10,* 9.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2,* 4.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience, 12,* 535–540.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4,* 863.

Marsh, L. E., Bird, G., & Catmur, C. (2016). The imitation game: Effects of social cues on 'imitation' are domain-general in nature. *Neuroimage, 139,* 368–375.

Mengotti, P., Corradi-Dell'Acqua, C., & Rumiati, R. I. (2012). Imitation components in the human brain: An fMRI study. *Neuroimage, 59,* 1622–1630.

Michael, J., & D'Ausilio, A. (2015). Domain-specific and domain-general processes in social perception—A complementary approach. *Consciousness and Cognition, 36,* 434–437.

Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in Cognitive Sciences, 18,* 37–45.

Müller, B. C., Leeuwen, M. L., Baaren, R. B., Bekkering, H., & Dijksterhuis, A. (2013). Empathy is a beautiful thing: Empathy predicts imitation only for attractive others. *Scandinavian Journal of Psychology, 54,* 401–406.

Mumford, J. A. (2012). A power calculation guide for fMRI studies. *Social Cognitive and Affective Neuroscience, 7,* 738–742.

Mumford, J. A., & Nichols, T. E. (2008). Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. *Neuroimage, 39,* 261–268.

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective & Behavioral Neuroscience, 7,* 1–17.

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M., & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience, 10,* 817–819.

Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localisers increase sensitivity and functional resolution of multi-subject analyses. *Neuroimage, 63,* 1646–1669.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences, 10,* 424–430.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science.* doi:10.1126/science.aac4716.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage, 56,* 2356–2363.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10,* 59–63.

Quadflieg, S., Flannigan, N., Waiter, G. D., Rossion, B., Wig, G. S., Turk, D. J., et al. (2011). Stereotype-based modulation of person perception. *Neuroimage, 57,* 549–557.

Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Sex, sexual orientation, and identification of positive and negative facial affect. *Brain and Cognition, 54,* 179–185.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27,* 169–192.

Rubia, K., Hyde, Z., Halari, R., Giampietro, V., & Smith, A. (2010). Effects of age and sex on developmental neural networks of visual-spatial attention allocation. *Neuroimage, 51,* 817–827.

Russell, T. A., Tchanturia, K., Rahman, Q., & Schmidt, U. (2007). Sex differences in theory of mind: A male advantage on Happé's "cartoon" task. *Cognition and Emotion, 21,* 1554–1564.

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2012). Enhancing social ability by stimulating right temporoparietal junction. *Current Biology, 22,* 2274–2277.

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2015). Functional lateralization of temporoparietal junction—Imitation inhibition, visual perspective-taking and theory of mind. *European Journal of Neuroscience, 42,* 2527–2533.

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localisers. *Neuroimage, 30,* 1088–1096.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *Neuroimage, 19,* 1835–1842.

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences, 19,* 65–72.

Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., & Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: A review using probabilistic atlases from different imaging modalities. *Human Brain Mapping, 38,* 4788–4805.

Schuwerk, T., Schurz, M., Müller, F., Rupprecht, R., & Sommer, M. (2017). The rTPJ's overarching cognitive function in networks for attention and theory of mind. *Social Cognitive and Affective Neuroscience, 12,* 157–168.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81,* 174.

Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology, 43,* 433–443.

Sonnby-Borgström, M., Jönsson, P., & Svensson, O. (2003). Emotional empathy as related to mimicry reactions at different levels of information processing. *Journal of Nonverbal behavior, 27,* 3–23.

Sonnby-Borgström, M. A., Jönsson, P., & Svensson, O. (2008). Gender differences in facial imitation and verbally reported emotional contagion from spontaneous to emotionally regulated processing levels. *Scandinavian Journal of Psychology, 49,* 111–122.

Southgate, V., & Hamilton, A. F. (2008). Unbroken mirrors: Challenging a theory of autism. *Trends in Cognitive Sciences, 12,* 225–229.

Sowden, S., & Catmur, C. (2015). The role of the right temporoparietal junction in the control of imitation. *Cerebral Cortex, 25,* 1107–1113.

Sowden, S., & Shah, P. (2014). Self-other control: A candidate mechanism for social cognitive function. *Frontiers in Human Neuroscience, 8,* 789.

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Control of shared representations relies on key processes in mental state reasoning. *Human Brain Mapping, 30,* 3704–3718.

Spengler, S., von Cramon, D. Y., & Brass, M. (2010). Resisting motor mimicry: Control of imitation involves processes central to social cognition in patients with frontal and temporo-parietal lesions. *Social Neuroscience, 5,* 401–416.

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology, 1,* e42.

Spunt, R. P., & Adolphs, R. (2014). Validating the why/how contrast for functional MRI studies of theory of mind. *Neuroimage, 99,* 301–311.

Spunt, R. P., & Adolphs, R. (2015). Folk explanations of behavior: A specialized use of a domain-general mechanism. *Psychological Science, 26,* 724–736.

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: Insights from social neuroscience. *Nature Reviews Neuroscience, 18,* 559.

Stoet, G. (2010). Sex differences in the processing of flankers. *The Quarterly Journal of Experimental Psychology, 63,* 633–638.

Stoet, G. (2017). Sex differences in the Simon task help to interpret sex differences in selective attention. *Psychological Research, 81,* 571–581.

Thayer, J. F., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology, 41,* 243–246.

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

van Schie, H. T., van Waterschoot, B. M., & Bekkering, H. (2008). Understanding action beyond imitation: Reversed compatibility effects of action observation in imitation and joint action. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1493.

Wager, T. D., & Nichols, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *Neuroimage, 18,* 293–309.

Wager, T. D., Sylvester, C. Y. C., Lacey, S. C., Nee, D. E., Franklin, M., & Jonides, J. (2005). Common and unique components of response inhibition revealed by fMRI. *Neuroimage, 27,* 323–340.

Wang, Y., Ramsey, R., & Hamilton, A. F. (2011). The control of mimicry by eye contact is mediated by medial prefrontal cortex. *Journal of Neuroscience, 31,* 12001–12010.

Weeks, D. J., & Proctor, R. W. (1990). Salient-features coding in the translation between orthogonal stimulus and response dimensions. *Journal of Experimental Psychology: General, 119,* 355–366.

Wiggett, A. J., Hudson, M., Tipper, S. P., & Downing, P. E. (2011). Learning associations between action and perception: Effects of incompatible training on body part and spatial priming. *Brain and Cognition, 76,* 87–96.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8,* 665.

Zaki, J., Hennigan, K., Weber, J., & Ochsner, K. N. (2010). Social cognitive conflict resolution: Contributions of domain general and domain specific neural systems. *Journal of Neuroscience, 30,* 8481–8488.

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