Title: Differential Recruitment of Theory of Mind Brain Network across Three Tasks: An Independent Component Analysis

Authors: Melissa D. Thye, Carla J. Ammons, Donna L. Murdaugh, Rajesh K. Kana

PII: S0166-4328(17)31809-0
DOI: https://doi.org/10.1016/j.bbr.2018.03.041
Reference: BBR 11363

To appear in: *Behavioural Brain Research*

Received date: 6-11-2017
Revised date: 5-3-2018
Accepted date: 26-3-2018

Please cite this article as: Thye MD, Ammons CJ, Murdaugh DL, Kana RK, Differential Recruitment of Theory of Mind Brain Network across Three Tasks: An Independent Component Analysis, *Behavioural Brain Research* (2010), https://doi.org/10.1016/j.bbr.2018.03.041

Differential Recruitment of Theory of Mind Brain Network across Three Tasks: An Independent Component Analysis

Melissa D. Thye[1], Carla J. Ammons[1], Donna L. Murdaugh[2], & Rajesh K. Kana[1]

[1]Department of Psychology, University of Alabama at Birmingham, Birmingham, AL USA

[2]Department of Pediatrics, University of Alabama at Birmingham, Birmingham, AL USA

Corresponding author:
Rajesh K. Kana
Department of Psychology
UAB Civitan International Research Center, CIRC 235G
1719 6th Avenue South, Birmingham, AL 35233
Phone: (205) 934-3171
Fax: (205) 975 6330
E-mail: rkana@uab.edu

Highlights

- Independent component analysis of three theory-of-mind fMRI tasks in healthy adults
- Reading the Mind in the Eyes, Reading the Mind in Voice, and Causal Attribution
- Convergence of core ToM regions seen across the networks identified for each task
- ToM network differences seen for causality task compared to other tasks
- Different networks are associated with mentalizing about emotions versus actions

Abstract

Social neuroscience research has focused on an identified network of brain regions primarily associated with processing Theory of Mind (ToM). However, ToM is a broad cognitive process, which encompasses several sub-processes, such as mental state detection and intentional attribution, and the connectivity of brain regions underlying the broader ToM network in response to paradigms assessing these sub-processes requires further characterization. Standard fMRI analyses which focus only on brain activity cannot capture information about ToM processing at a network level. An alternative method, independent component analysis (ICA), is a data-driven technique used to isolate intrinsic connectivity networks, and this approach provides insight into network-level regional recruitment. In this fMRI study, three complementary, but distinct ToM tasks assessing mental state detection (e.g. RMIE: Reading the Mind in the Eyes; RMIV: Reading the Mind in the Voice) and intentional attribution (Causality task) were each analyzed using ICA in order to separately characterize the recruitment and functional connectivity of core nodes in the ToM network in response to the sub-processes of ToM. Based on visual comparison of the derived networks for each task, the spatiotemporal network patterns were similar between the RMIE and RMIV tasks, which elicited mentalizing about the mental states of others, and these networks differed from the network derived for the Causality task, which elicited mentalizing about goal-directed actions. The medial prefrontal cortex, precuneus, and right inferior frontal gyrus were seen in the components with the highest correlation with the task condition for each of the tasks highlighting the role of these regions in general ToM processing. Using a data-driven approach, the current study captured the differences in task-related brain response to ToM in three distinct ToM paradigms. The findings of this study further elucidate the neural mechanisms associated with mental state detection and causal attribution, which represent possible sub-processes of the complex construct of ToM processing.

Keywords: RMIE, RMIV, Intentional Causality, Theory of Mind, Independent Component Analysis

Introduction

Understanding and decoding the emotions, intentions, and beliefs of others is a core component of human interpersonal interaction. The cognitive mechanism that underlies this ability, Theory of Mind (ToM), develops throughout childhood [1]. Mastery of ToM is achieved by adulthood in typical development and is believed to be functionally coordinated by a network of brain regions including anterior cingulate cortex (ACC), posterior cingulate cortex (PCC), medial prefrontal cortex (MPFC), precuneus, inferior frontal gyrus (IFG), superior temporal sulcus (STS), and temporoparietal junction (TPJ) [2–4]. Although these regions are consistently reported to be active in neuroimaging studies of ToM processing, there is still a need to elucidate the network-level roles of these regions within ToM paradigms. Previous evidence has highlighted the differential recruitment of regions such as MPFC and TPJ in response to ToM tasks [4,5]. For instance, MPFC may play a broader role in ToM, with MPFC recruitment noted across various ToM tasks [6]. Conversely, activity in other areas such as TPJ has been provided as evidence of specialization within the ToM network due to its selective response to only certain ToM paradigms [3]. This presents a significant gap in the literature concerning the specialization of nodes within the ToM network.

Although studying the localization of each region in isolation can be informative in terms of functional specialization, examining the connections among regions provides a more comprehensive network-level understanding of neural processing [7]. An innovative data-driven approach to studying functional connectivity in neuroimaging datasets is independent component analysis (ICA) [8]. This method is uniquely suited to isolating the multivariate signals inherent to fMRI data. Additionally, ICA analyzes the signal from each voxel as an independent function from all other voxels and can recover signals that would be excluded or inappropriately modeled if a Gaussian distribution is assumed [9]. The application of ICA to functional neuroimaging data has several advantages. First, ICA examines the correspondence of time series across different brain regions, and is thus a proxy of functional connectivity [10]. This type of approach provides more information about distributed, network-level processing compared to the traditional localist approach to studying regional activation in isolation. Second, due to the data-driven, exploratory nature of the analysis, the identification of any components that contain regions with covarying time series in response to the task manipulation provides compelling evidence of functional synchrony among regions within a network. This is a more rigorous approach to studying the ToM network as the components are derived in a bottom-up process rather than being imposed by the researcher. As a result, the risk of confirmatory bias is eliminated. Third, ICA is an exploratory, whole-brain analysis not constrained to *a priori* regions of interest. This is well-suited to studying areas outside of the known ToM network that may play an ancillary role in the network. This allows for an exploration of within and between network dynamics involving regions not classically considered integral to ToM processing. Fourth, no manipulations are made to the data to distort the findings of the analysis; the data that is derived from the ICA is reported. Thus, there is relatively more transparency in the overall analysis and the results reported. Fifth, the ICA method isolates components associated with head motion, and eliminates this artefactual data from the resulting networks. This analytic method has been applied previously to neuroimaging data to characterize ToM processing in autism [11] and to define core brain networks for follow-up analyses [12]. This approach, therefore, is ideal for investigating the gap in the literature regarding network connectivity associated with different sub-processes of ToM.

The aim of the current study is to capture spatiotemporal patterns of connectivity within three ToM paradigms which involve stimuli that target different aspects of ToM processing. The first task, *Reading the Mind in the Eyes* (RMIE) [13] is a widely used ToM task that relies on visual presentation of a static photo of the eye region of the face to examine mental state detection. Similarly, the second task, *Reading the Mind in the Voice* (RMIV) [14], elicits mental state detection but with auditory stimuli. The final task, *Causality* (intentional and physical causal attribution), provides participants with social contexts in the form of visual comic strip vignettes and forces participants to choose an ending for the vignette in order to elicit intentional attribution of the protagonists [15]. In this study, ICA was performed for each paradigm separately and the top task-related components were examined in order to provide information concerning the networks of brain regions associated with each task. We hypothesized that RMIE and RMIV would have components comprised of similar regions due to the common cognitive process of mental state detection addressed in both tasks. We also expected that the components identified in the Causality task would be comprised of a greater number of functionally synchronized regions within the ToM network due to the contextual nature of the task which demands integration of social information across situations.

Overall, the results of this study capture the relationships among core regions within the broader ToM network during two ToM processes: mental state detection and intentional attribution. Furthermore, the data-driven nature of the analytic approach adopted in the current study provides an avenue to derive networks of regions without constraining the analysis with a model. Thus, the current study uses a bottom-up approach to investigate and characterize spatiotemporal patterns of connectivity within tasks assessing mental state detection and intentional attribution, which represent possible sub-processes of the complex construct of ToM processing.

Methods

2.1 Participants

Twenty-four healthy undergraduate students recruited from the *Introduction to Psychology (PY 101)* course at our university participated in this fMRI study and results from different analyses using this sample of participants have been previously published [16]. Due to excessive head motion, one participant was excluded from the analysis. Four participants completed the RMIE and RMIV tasks, but did not complete the Causality task. Thus, the final sample included 23 participants for the RMIE and RMIV tasks and 19 participants for the Causality task. Participants completed the following pre-screening measures: Edinburgh Handedness Inventory [17] to measure handedness, Empathizing Quotient (EQ) [18] questionnaire to measure empathy, and Wechsler Abbreviated Scale of Intelligence (WASI) [19] to measure IQ (Table 1). All participants were right handed. Exclusion criteria included any contraindications that would interfere with fMRI safety. All participants gave written informed consent which was approved by the Institutional Review Board of our institution.

Table 1

Participant Demographic Information

		RMIE/ RMIV			Causality	
Variable	*N*	*Mean (±SD)*	*Range*	*N*	*Mean (±SD)*	*Range*
Age	23	19.96 (± 1.30)	19-24	19	20.08 (± 1.41)	19-24
PIQ	23	110 (± 9.89)	91-126	19	108.53 (± 10.02)	91-126
VIQ	23	112.7 (± 19.75)	97-192	19	113.59 (± 21.14)	98-192
FSIQ	23	110.35 (± 7.46)	96-128	19	109.59 (± 7.52)	96-128
EQ	23	43.61 (± 11.97)	16-73	19	44.05 (± 12.83)	16-73
Gender (M:F)	23	7:16		19	6:13	

Note. N, number of participants; RMIE, Reading the Mind in the Eyes; RMIV, Reading the Mind in the Voice; PIQ, Performance Intelligence Quotient; VIQ, Verbal Intelligence Quotient; FSIQ, Full Scale Intelligence Quotient; EQ, Empathizing Quotient; M, Male; F, Female; SD, standard deviation of the mean. Scores on the EQ range from 0 to 80 with higher scores indicating greater empathizing.

2.2 Experimental Paradigm

In the MRI scanner, participants completed the RMIE task, the RMIV task, and the Causality task that was used in two of our previous studies [11,20]. Participants completed all 3 tasks in the same scanning session. The experimental design and example stimuli for each task are shown in Figure 1. Both RMIE and RMIV had the same experimental conditions: gender identification, mental state (e.g. emotion) identification, and a fixation block. For RMIE, participants were visually presented with a static image of the eye region photograph of a face and asked to either identify the gender of the person (gender condition) or the mental state expressed by the person in the stimuli (emotion condition). For RMIV, participants heard an utterance and were asked to identify the gender of the speaker (gender condition) or the mental state or emotion conveyed by the speaker (emotion condition). The utterance was an auditory stimuli, but the response options (gender or mental state words, depending on the condition) were presented visually. The stimuli for the RMIE and RMIV tasks were adapted from previous research [13,14] and, with the exception of the gender options, none of the stimuli or response options were the same across the two tasks. For the Causality task, participants were shown comic strip vignettes with three images of objects or characters and instructed to identify a logical ending to the scenario from three choices. In the physical causality condition, the participants were presented with cartoon images depicting a physical action, whereas in the intentional causality condition, participants were presented with three images depicting a goal-oriented and socially interactive behavior. All stimuli for the Causality task were visually presented.

During the RMIE task, participants were presented with a total of six gender blocks, each containing six stimuli presented for five seconds each, six emotion blocks, each containing six stimuli which were presented for seven seconds each, and six twenty-four second fixation blocks. Gender blocks were thirty seconds long and emotion blocks were forty-two seconds long. The entire task lasted 10 minutes. During the RMIV task, participants were presented with a total of six gender blocks each containing five stimuli which were presented for five seconds each, six emotion blocks each containing five stimuli which were presented for seven seconds each, and

six twenty-four second fixation blocks. The duration of the gender blocks and the emotion blocks were twenty-five seconds and thirty-five seconds respectively. The entire task lasted approximately 8 minutes. As depicted in Figure 1, four response options were provided for the emotion blocks for both RMIE and RMIV, and two response options (e.g. male or female) were provided for the gender blocks. The task design was adapted from the revised version of the RMIE task which included additional mental state response options to reduce the likelihood of selecting the correct answer by chance, a previous criticism of the task [13]. Due to the inability to increase the number of gender options to correspond to this revision of the emotion condition, the stimuli were presented for a shorter duration during the gender condition compared to the emotion condition. This design allowed for capturing the hemodynamic response function for both processes without modeling unrelated cognitive processes in the gender condition.

During the Causality task, the first three images in the incomplete vignette were presented for five seconds and the three choices to complete the vignette were presented for six seconds. The entire epoch lasted for eleven seconds. There were a total of twenty stimuli presented to participants (ten intentional causality and ten physical causality stimuli). Including the five twenty-four second fixation blocks, the entire task lasted approximately 8 minutes. RMIE and RMIV were modeled as blocked design, and the Causality task was modeled as event-related due to the longer stimulus presentation time and the presentation of discrete trials. All participants were shown the same stimuli for each task, but the order in which stimuli were presented within each run was counterbalanced across participants. Stimuli were displayed in the scanner using the E-prime software 1.2 (Psychology Software Tools, Pittsburgh, PA, USA), and participant responses in the scanner were recorded using a fiber optic button box response system with the right hand. Participants were able to practice all of the tasks prior to the experiment.

Insert Figure 1 here

2.3 Data Acquisition

Functional and structural imaging data were acquired using a 3T Siemens Allegra head-only scanner located in the Civitan International Research Center at the University of Alabama at Birmingham. Structural images were acquired using high-resolution T1-weighted anatomical scans with a 160 slice 3D Magnetization Prepared Repaid Acquisition Gradient Echo (MPRAGE) volume scan (TR = 1000ms, TE = 3.34ms; flip angle = 12°; field of view (FOV) = 25.6 cm; matrix size = 256 × 256; slice thickness = 1 mm). T2*-weighted functional images were acquired using a single-shot gradient-recalled echo-planar pulse sequence (TR = 1000ms, TE = 30ms, flip angle = 60°, FOV = 24 × 24 cm, matrix = 64 × 64). Seventeen adjacent oblique-axial slices were acquired with parallel reference to the anterior commissure-posterior commissure (AC-PC) line in a single cycle of scanning in an interleaved sequence, with a 5-mm slice thickness and 1-mm slice gap resulting in an in-plane resolution of $3.75 \times 3.75 \times 5\ mm^3$.

2.4 Data Analysis

2.4.1 Pre-processing. Functional images were preprocessed using a combination of the functional connectivity toolbox (CONN) [21] and Statistical Parametric Mapping (SPM) 12 (Wellcome Department of Cognitive Neurology, London, UK). Preprocessing of functional data was completed in CONN. All images were slice-time corrected, motion corrected by registering

each functional volume to the middle time point, normalized to a 3mm-isotropic MNI152 template, and smoothed using an 8mm Gaussian kernel. Statistical analyses on the individual and group level data were conducted using SPM 12's implementation of the general linear model [22].

The impact of head motion on functional connectivity analyses has been widely documented [23,24]. In order to address potential confounds and control for motion artifacts, the Euclidean distance from the six rigid-body motion parameters for two consecutive time points was calculated and any deviation greater than .2mm were labeled as excessive head motion. For the excluded participant, the number of volumes exceeding the movement threshold was greater than 20% of the total number of volumes. For all of the remaining participants, less than 10% of the total number of volumes exceeded the movement threshold.

2.4.2 Independent Component Analysis. ICA of functional images was implemented in the Group ICA of fMRI Toolbox (ICA GIFT v4.0a) to establish networks of spatially distinct, temporally correlated regions associated with a time course of interest [25]. Separate ICAs were conducted for the RMIE, RMIV, and Causality tasks. Each ICA was implemented in GIFT using a multi-subject principal component analysis (PCA) back-reconstruction with ICASSO, a toolbox which measures stability of the components, and followed up with one-sample t-tests of resulting components in SPM12.

Briefly, the following steps were performed in GIFT for each ICA: 1) the data were reduced through two stages of PCA: single subject PCA with temporal concatenation followed by group level PCA; 2) the number of components was estimated from the fMRI data of all subjects using minimum descriptive length (MDL) criteria. A total of 30 independent components were estimated at this stage for RMIE, 27 for RMIV and 29 for the Causality task; 3) spatial ICA with Infomax was run on the reduced data and independent spatial component maps and their accompanying time courses were constructed for each subject and the group average; 4) stability analysis using ICASSO estimated the reliability of components over 10 iterations of the ICA algorithm and unstable components were eliminated; 5) artefactual components (e.g. those produced by head motion or cardiorespiratory effects) were removed based on visual inspection following guidelines outlined previously [26]; and 6) the ICA was repeated without the eliminate components to improve source detection. This resulted in a final set of independent components for each task. For RMIE, 30 components were initially estimated, of which 7 were eliminated (0 unstable, 7 artefactual), and a final set of 23 components were constructed. For RMIV, 27 components were initially estimated, 5 were eliminated (0 unstable, 5 artefactual), and a final set of 22 components were constructed. For causality, 29 components were initially estimated, 7 were eliminated (0 unstable, 7 artefactual), and a final set of 22 components were constructed.

After the implementation of the ICA, the resulting components were temporally sorted based on the correspondence between the component time series and a design matrix. The design matrices for each task were generated in SPM12 by modeling the onsets and durations of the blocks or events within each task, and convolving this design matrix with a canonical hemodynamic response function. Using GIFT's temporal sorting feature, these components were ranked based on their correlation with task conditions. A correlation threshold of .10 was used to determine which components were sufficiently correlated with the task manipulation to be reported and discussed. The justification for this threshold was two-fold. First, the cognitive process of interest, ToM processing, is a complex, higher-order cognitive skill which displays distributed connectivity across a network of regions. Given the diffuse nature of ToM

processing, it is expected that the correlation coefficients obtained for a study of ToM processing would be relatively lower than that have been previously reported for sensory or motor ICA studies. In fact, a correlation cut-off of .10 is relatively high based on a previously published ICA study of social cognition which used a cut-off of .06 [11]. Second, a lower threshold promotes greater transparency in reporting the results by allowing for several top components to be reported and discussed. Although greater emphasis is placed on those components which showed the highest correlation with the task, the exploratory, data-driven nature of the analysis necessitates a comprehensive overview of the data and the derived components. The selection of a more stringent threshold may result in selectively removing components which may be informative for understanding weaker within or between-network dynamics. Correlations between the group-level components and the design matrices ranged from .02 to .33.

Finally, individual subject component maps were entered into a one-sample *t*-test in SPM12 to create task-specific component maps for each of the top-ranked components. A Monte Carlo simulation was run on the average grey matter mask in Analysis of Functional NeuroImages (AFNI) [27] using 3dClustSim and yielded a threshold of 74 contiguous voxels at an uncorrected *p* value of .001 (at a family wise error (FWE) threshold of .05), and this threshold was applied to all of the results. Simulations such as Monte Carlo can identify the cluster size thresholds that optimize the balance between Type I and Type II error [28].

Results

3.1 Reading the Mind in the Eyes

The top-ranked components identified in the ToM conditions were all unique and distinct from the components identified in the control conditions in each of these tasks (Figures 2-4; see Supplementary Materials for coordinate tables and control condition results). For RMIE, component A (R^2 = .23) was comprised primarily of visual associative areas such as the calcarine sulcus, lingual gyrus, left superior parietal lobule (LSPL), and thalamus as well as areas implicated in emotion processing and empathy such as the cingulate gyrus [29], and supramarginal gyrus [30]. Component B for RMIE (R^2 = .22) included similar areas, such as the supramarginal gyrus and cingulate gyrus as well as additional areas such as middle frontal gyrus (MFG), insula [31], left inferior parietal lobule (LIPL) [32], right inferior frontal gyrus (RIFG), precuneus, right superior frontal gyrus (RSFG), and medial prefrontal cortex (MPFC). Additional areas were also seen in this component such as, LSPL, cuneus, left superior occipital gyrus, and right calcarine gyrus. Component C for RMIE (R^2 = .19) contained many of the same areas seen in the previous components, but the overall network structure differed. Specifically, there was more MPFC involvement extending anteriorly, greater IFG recruitment, and caudate nucleus involvement. Component D (R^2 = .17) was comprised of classic mentalizing regions such as ACC and IPL. Additional regions not previously identified in the other components were also present, such as an anterior portion of the cerebellum and putamen. Component E (R^2 = .14) was highly similar to the other identified networks with the exception of diffuse middle temporal gyrus involvement. The other regions identified within this component were also captured in the other RMIE networks such as IFG, MPFC, SFG, and insula. Component F (R^2 = .13) was predominately comprised of middle frontal gyrus and cingulate with additional recruitment of cerebellum and lingual gyrus. Component G (R^2 = .11) was comprised of a network of ToM regions including: ACC, bilateral IFG, SFG, precuneus, and left and right fusiform gyrus. The final RMIE network, component H (R^2 = .11), contained a large region encompassing parts of the ACC, supramarginal gyrus, left IFG, and right insula. In addition, there was a small superior

parietal region identified within this network (Figure 2; Supplemental Table 1). All remaining components fell below the .10 threshold.

3.2 Reading the Mind in the Voice

In the mental state condition of the RMIV task, the top component A (R^2 = .33) consisted of regions including left middle temporal gyrus (LMTG), cuneus, superior occipital gyrus, lingual gyrus, left superior temporal gyrus (LSTG), and bilateral SPL as well as areas associated with ToM processing such as bilateral IFG, right insula, MPFC, left precuneus, LIPL, and right supramarginal gyrus. Component B (R^2 = .21) included both auditory and visual processing areas such as right calcarine gyrus, occipital gyrus, RSPL. Additional regions such as RIFG, cingulate gyrus, and right insula were also identified in this component. Component C (R^2 = .17) was comprised of distributed connections among right supramarginal gyrus, ACC, STG, and MPFC. The caudate nucleus, insula, and a small portion of the cerebellum were also recruited in this network. The final network for RMIV, component D (R^2 = .15), contained right medial orbitofrontal cortex, right SFG and MPFC, right angular gyrus, caudate nucleus, putamen, right IFG, and a small anterior portion of the cerebellum (Figure 3; Supplemental Table 2). All remaining components fell below the .10 threshold.

3.3 Causality

One component was identified for the intentional causal attribution condition which exceeded the .10 threshold. Component A for the Causality task (R^2 = .14) was comprised of ToM regions such as left MFG, MPFC, precentral gyrus, and RSPL. Other areas associated with sensory processing were also seen in component A such as the LSTG, bilateral MTG, and cerebellum. All of the remaining components associated with the intentional causality condition fell below .10 indicating lower task correlation (Figure 4; Supplemental Table 3).

Insert Figures 2-4 here

Discussion

The core brain networks showing the greatest temporal correlation with the ToM condition in three distinct tasks were isolated using a data-driven approach that was blind to *a priori* characterizations of the ToM network. The top task related components for each ToM task contained sensory processing areas as well as areas involved in mentalizing. The most notable finding in the present study is the recruitment of ToM network regions within each task, with some specificity in regional participation emerging in the emotion condition for both RMIE and RMIV.

Although all three tasks required participants to engage in mentalizing, the type of mentalizing required to successfully complete the trials differed across tasks. The RMIE and RMIV tasks involved understanding the mental states of others (e.g. through eyes or voice) while the Causality task invoked action and intention understanding and prediction of behavior based on intention. Understanding and decoding the mental states, beliefs, morals, and goals of others is a distinct mentalizing process from action understanding or prediction. This type of mentalizing is associated with a distinct network of brain regions including MPFC, precuneus, and TPJ [33]. One complementary theory which aligns with this type of mentalizing, *the structured-event-complex framework*, suggests that the prefrontal cortex is largely involved in understanding goal-directed social events and decoding emotional information from social

sequences [34,35]. The second type of mentalizing which was elicited by the Causality task, involves decoding goal-oriented biological motion, understanding actions, and making predictions based on observed tasks [33]. The premotor cortex and the anterior intraparietal sulcus are consistently reported as components of this type of mentalizing in both animal and human research. Thus, although all of the tasks underlie mentalizing, the type of mentalizing elicited differed across the tasks.

The act of mentalizing may be performed in different ways and has historically been studied using either *the simulation theory* or *the theory theory* perspective. Simulation theory is an embodied approach which suggests that humans internally simulate what another person is feeling, thinking, or doing in order to arrive at an accurate prediction of the mental or physical state of others [36]. This theoretical approach aligns with a mirroring mechanism where neural responses are identical when engaging or when observing an action. The alternative, the theory theory perspective, argues that individuals do not fully simulate the mental states of others, and instead develop a rough theory of others' minds and then make inferences from that internal representation of others' minds. This approach aligns with mentalizing about the mental states that are not grounded in biological motion. There are mixed findings supporting both theories [37] suggesting that each perspective is elicited depending on the nature of the stimuli and the cognitive task. For instance, in studies of motor understanding and prediction, there is activation in premotor cortex which suggests that primates are fully simulating the motor action simply by observing another primate engaged in the action [38–40]. Other studies of emotional understanding have reported activation in paracingulate cortex and STS which is more consistent with the theory theory perspective [41]. Thus, different aspects of ToM processing may rely on different internal strategies to successfully describe, understand, and predict the intentions or mental states of others, and these strategies may broadly align with different types mentalizing.

The results across the three tasks converged around several regions which emerged within the top-ranked components for each task. A number of areas involved in emotion processing were identified in both RMIE and RMIV including the precuneus, IFG, superior temporal sulcus extending into inferior parietal lobule, and MPFC which are consistently reported ToM and mentalizing regions [2,4]. The most commonly reported sites of activation in previous fMRI studies of RMIE have been the IFG and the posterior STS (pSTS) [42–44]. Activity in the IFG has been linked to performance on the RMIE task specifically [45] with possible explanations of the role of this region in semantic memory [45], a mirroring mechanism [46], and mental state detection [47]. In addition, the IFG plays a dual role in both language processing and mentalizing, neural systems with overlapping neuroanatomy. The overlap of these functional systems, observed in both monkey and human research, has been used as evidence to suggest that language emerges from an intact imitation system [39,46,48–50]. In neuroimaging meta-analyses, the MPFC emerges across ToM paradigms as a consistent area of activation highlighting its role in mentalizing broadly regardless of task specific demands [4,5]. Across all three tasks, MPFC was identified in the top-ranked components suggesting the important role this region plays in facilitating ToM-based inferences.

Although RMIV task is similar in nature to RMIE, no studies have used neuroimaging methods to study network-level differences across these tasks. The network identified in Component A for RMIE can be characterized predominately as a visual network with projections to memory processing areas. For instance, the hippocampus has direct connections with other structures within the limbic cortex such as amygdala and the reciprocal connections between these regions support memory and encoding of emotional stimuli [51]. Additionally, the visual

associative regions within the network may project to the limbic cortex to aid in the retrieval and processing of the emotional information based on previous experiences. Thus, within the derived network, the hippocampus may serve as the hub receiving inputs from the other regions within the network to draw from memory in order to label the emotions depicted in the task stimuli. This component for the RMIE task closely resembled component B identified for the RMIV task. The remaining components for RMIE more closely resemble the known ToM and mentalizing networks. For RMIV, component A may be reflective of a ToM network that is visually similar to component C for RMIE, although some of the regions that comprise the network differ from the ToM network identified for the RMIE task. Interestingly, the role of the limbic cortex is also seen in the recruitment of the thalamus as the primary limbic region involved in decoding mental state which aligns with previous findings regarding the function of the thalamus in emotion processing [52]. The predominately sensory processing network identified in Component B for RMIV more closely resembles the first network derived for the RMIE task with additional auditory processing areas that were not seen in the RMIE networks. The top components for RMIE and RMIV also contained the caudate nucleus, which has been found to play a role in emotion recognition [53]. The remaining components identified for the RMIV task were visually similar to components identified for RMIE. RMIV component C, for instance, displayed a similar network structure to RMIE component F. Both networks heavily recruit MFG and ACC. The final component identified for RMIE (component D) was similar to RMIE component D; both networks include SFG and IPL extending to angular gyrus. Thus, there was a high level of correspondence between the components identified for RMIE and RMIV. This is likely due to the fact that both tasks require participants to infer mental states and are thus eliciting the same underlying ToM sub-process.

As expected, there were differences in sensory processing between RMIE and RMIV with visual associative areas seen in both tasks, such as the calcarine sulcus, lingual gyrus, and superior occipital gyrus and auditory processing areas such as left middle and superior temporal gyrus seen in components identified for RMIV. The top-ranked components identified for both RMIE and RMIV contained the right supramarginal gyrus which has been implicated in empathetic understanding of others' emotions when they are incongruent with one's own emotions [30], the cingulate gyrus which has been implicated in social cognition in general and emotion processing in particular [29], and the insula, involved in empathizing [54]. It is interesting to note that these regions were not seen in the top-ranked component for the Causality task, further highlighting how the mentalizing involved in detecting mental states in RMIE and RMIV differed from the action understanding and prediction inherent to the Causality task. In this way, the RMIE and RMIV tasks diverge from the Causality task in the explicit labeling of mental state and the subsequent neural correlates associated with this social cognitive process.

As previously mentioned, the networks derived for the Causality task differed from the networks identified for the RMIE and RMIV tasks. This is attributed to the different mentalizing process elicited by the Causality task. Specifically, the intentional condition in the Causality task required participants to engage in mentalizing about goal-directed actions, which engages different brain networks than mentalizing about non-goal-directed actions, presumably because an understanding of intention is required to predict the response [33]. Importantly, the physical causality trials were non-goal directed, instead showing a mechanical event such as a glass falling off of a table. Understanding and predicting actions can also be achieved by the mirroring system, which is implicated in observed goal-oriented biological motion processing and prediction [33]. Animal studies of action understanding have consistently highlighted neural

responses for when actions are observed or predicted from observed actions [38,50,55]. Some of this research has been extended to human neuroimaging studies which have isolated the left STS, left IPL, and left IFG in goal-directed action understanding and prediction [46,50]. Neither the RMIE nor the RMIV task required participants to mentalize about goals or make a prediction based on presumed goals, so the differences between the obtained networks for the other two tasks and the Causality task is likely due to the fact that a different sub-process of ToM was required to complete the Causality task. Interestingly, the correlation coefficients of the intrinsic networks identified using ICA were higher for both the RMIE and RMIV tasks when compared to the Causality task. This resulted in fewer identified components for the Causality task. There are a number of possible explanations for this finding. First, this suggests that an ICA approach may be less efficient in capturing transient network dynamics associated with some cognitive processes. This argument is further supported by the low task correlation coefficients for the causality control condition component (see Supplemental Materials). This suggests that the action understanding and intentional attribution inherent to the Causality task are higher-order cognitive processes with transient, widespread network level dynamics that vary across the time window imposed in a task-based study. Thus, a model-based approach would likely have missed the network-level fluctuations that were uncovered by ICA. Second, the RMIE and RMIV tasks may strongly elicit one type of mentalizing (e.g. mental state detection) whereas the Causality task may require a number of cognitive skills (e.g. mentalizing about actions, mirroring, intentional attribution, causal reasoning). The causality stimuli presented to participants are multifaceted and provide more detail and context compared to the static image or brief audio clip presented during the RMIE and RMIV tasks respectively. Therefore, participants may be shifting their attentional and cognitive mechanisms throughout the duration of the task resulting in weak network correlation with the overall task manipulation. The ICA approach adopted here allowed for a closer examination of these nuanced fluctuations.

Differential network connectivity was seen predominantly in the top component identified for the Causality task. The role of the cerebellum across all of these tasks may be reflective of the functional connections between the cerebellum and other cortical networks [56] or may underscore the role of the cerebellum in social cognition. Specifically, the cerebellum, through its extensive cortical projections, is involved in inferring intentions and extracting social information from actions [57] which may be more pronounced in the causality paradigm compared to either RMIE or RMIV. The cerebellum is also involved in simulating goal-directed actions which directly involves understanding and predicting future actions, processes that are inherent to mentalizing [58,59]. Activation in the STG and MPFC has been previously reported in response to this Causality task in a positron emission tomography (PET) study [15]. There was more widespread right superior parietal lobe involvement in the causality component when visually compared to the components identified for RMIE or RMIV. This region plays a role in social cognition through the understanding of other's movements and actions [60] which aligns with the cognitive requirements of the task (e.g. providing the next action for a person completing a goal-oriented task), and is also implicated in the mirroring system. Upon visual inspection, the Causality component did not resemble any single component identified for RMIE or RMIV, although many of the same regions were present in some of the RMIE and RMIV components. This may reflect the differences in the type of mentalizing elicited by the task.

Conclusions

Using a data-driven statistical approach, the findings of this study illustrate the spatiotemporal patterns of functional connectivity within three distinct, but complementary ToM tasks. Despite the expected divergence in sensory processing mechanisms, the RMIE and RMIV results suggest that there is stability within the cognitive domain of mental state detection which may represent a sub-process of ToM, and more specifically mentalizing. The higher degree of correspondence between the identified components for RMIE and RMIV suggests that mentalizing involving decoding and simulating the mental states of others relies on different neural networks than mentalizing about actions and making predictions from goal-directed actions in the Causality task. Although the tasks were not statistically compared with each other, there is evidence of similarities as well as core differences across these tasks. These findings provide insight into network-level variation in ToM regional involvement in response to different task manipulations. The current study further underscores the need to study the sub-processes of ToM, such as mental state attribution and intentional causality, and expand beyond traditional activation analyses by looking at network level connectivity.

Funding: This research was funded by the UAB Department of Psychology faculty funds.

References

[1] H.M. Wellman, D. Cross, J. Watson, Meta-analysis of theory-of-mind development: The truth about false belief., Child Dev. 72 (2001) 655–684. doi:10.1111/1467-8624.00304.

[2] C.E.V. Mahy, L.J. Moses, J.H. Pfeifer, How and where: Theory-of-mind in the brain, Dev. Cogn. Neurosci. 9 (2014) 68–81. doi:10.1016/j.dcn.2014.01.002.

[3] R. Saxe, L.J. Powell, It's the thought that counts: Specific brain regions for one component of theory of mind., Psychol. Sci. 17 (2006) 692–9. doi:10.1111/j.1467-9280.2006.01768.x.

[4] M. Schurz, J. Radua, M. Aichhorn, F. Richlan, J. Perner, Fractionating theory of mind: A meta-analysis of functional brain imaging studies, Neurosci. Biobehav. Rev. 42 (2014) 9–34. doi:10.1016/j.neubiorev.2014.01.009.

[5] S.J. Carrington, A.J. Bailey, Are there theory of mind regions in the brain? A review of the neuroimaging literature, Hum. Brain Mapp. 30 (2009) 2313–2335. doi:10.1002/hbm.20671.

[6] J.P. Mitchell, M.R. Banaji, C.N. Macrae, General and specific contributions of the medial prefrontal cortex to knowledge about mental states., Neuroimage. 28 (2005) 757–62. doi:10.1016/j.neuroimage.2005.03.011.

[7] S.M. Smith, The future of FMRI connectivity, Neuroimage. 62 (2012) 1257–1266. doi:10.1016/j.neuroimage.2012.01.022.

[8] V.D. Calhoun, T. Adali, G.D. Pearlson, J.J. Pekar, Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms., Hum. Brain Mapp. 13 (2001) 43–53. http://www.ncbi.nlm.nih.gov/pubmed/11284046.

[9] V.D. Calhoun, G. Pearlson, T. Adali, Independent component analysis applied to fMRI Data : A generative model for validating results, Signal Processing. (2004) 281–291.

[10] M.P. van den Heuvel, H.E. Hulshoff Pol, Exploring the brain network: A review on resting-state fMRI functional connectivity, Eur. Neuropsychopharmacol. 20 (2010) 519–534. doi:10.1016/j.euroneuro.2010.03.008.

[11] D.L. Murdaugh, K.D. Nadendla, R.K. Kana, Differential role of temporoparietal junction and medial prefrontal cortex in causal inference in autism: An independent component analysis, Neurosci. Lett. 568 (2014) 50–55. doi:10.1016/J.NEULET.2014.03.051.

[12] W.R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, M.D. Greicius, Decoding subject-driven cognitive states with whole-brain connectivity patterns., Cereb. Cortex. 22 (2012) 158–65. doi:10.1093/cercor/bhr099.

[13] S. Baron-Cohen, S. Wheelwright, J. Hill, Y. Raste, I. Plumb, The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism., J. Child Psychol. Psychiatry. 42 (2001) 241–51. doi:10.1111/1469-7610.00715.

[14] Rutherford, S. Baron-Cohen, S. Wheelwright, Reading the mind in the voice: A study with

normal adults and adults with asperger syndrome and high functioning autism, J. Autism Dev. Disord. 32 (2002) 189–194.

[15] E. Brunet, Y. Sarfati, M.C. Hardy-Baylé, J. Decety, A PET investigation of the attribution of intentions with a nonverbal task., Neuroimage. 11 (2000) 157–166. doi:10.1006/nimg.1999.0525.

[16] M.D. Thye, D.L. Murdaugh, R.K. Kana, Brain mechanisms underlying reading the mind from eyes, voice, and actions, Neuroscience. 374 (2018) 172–186. doi:10.1016/J.NEUROSCIENCE.2018.01.045.

[17] R.C. Oldfield, The assessment and analysis of handedness: The Edinburgh inventory., Neuropsychologia. 9 (1971) 97–113.

[18] E.J. Lawrence, P. Shaw, D. Baker, S. Baron-Cohen, a S. David, Measuring empathy: Reliability and validity of the Empathy Quotient., Psychol. Med. 34 (2004) 911–919. doi:10.1017/S0033291703001624.

[19] D. Wechsler, Wechsler Abbreviated Scale of Intelligence (WASI), The Psychological Corporation, San Antonio, TX, 1999.

[20] R.K. Kana, L.E. Libero, C.P. Hu, H.D. Deshpande, J.S. Colburn, Functional brain networks and white matter underlying theory-of-mind in autism, Soc. Cogn. Affect. Neurosci. 9 (2014) 98–105. doi:10.1093/scan/nss106.

[21] S. Whitfield-Gabrieli, A. Nieto-Castanon, Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks, Brain Connect. 2 (2012) 125–141. doi:10.1089/brain.2012.0073.

[22] K.J. Friston, A.P. Holmes, K.J. Worsley, J.-P. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: A general linear approach, Hum. Brain Mapp. 2 (1995).

[23] T.D. Satterthwaite, D.H. Wolf, J. Loughead, K. Ruparel, M.A. Elliott, H. Hakonarson, R.C. Gur, R.E. Gur, Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth, Neuroimage. 60 (2012) 623–632. doi:10.1016/j.neuroimage.2011.12.063.

[24] K.R.A. Van Dijk, M.R. Sabuncu, R.L. Buckner, The influence of head motion on intrinsic functional connectivity MRI, Neuroimage. 59 (2012) 431–438. doi:10.1016/j.neuroimage.2011.07.044.

[25] Calhoun, T. Adali, G.D. Pearlson, J.J. Pekar, A method for making group inferences from functional MRI data using independent component analysis, Hum. Brain Mapp. 14 (2001) 140–151. doi:10.1002/hbm.1048.

[26] L. Griffanti, G. Douaud, J. Bijsterbosch, S. Evangelisti, F. Alfaro-Almagro, M.F. Glasser, E.P. Duff, S. Fitzgibbon, R. Westphal, D. Carone, C.F. Beckmann, S.M. Smith, Hand classification of fMRI ICA noise components, Neuroimage. 154 (2017) 188–205. doi:10.1016/J.NEUROIMAGE.2016.12.036.

[27] R.W. Cox, AFNI: Software for analysis and visualization of functional magnetic

resonance neuroimages, Comput. Biomed. Res. 29 (1996) 162–173. doi:10.1006/cbmr.1996.0014.

[28] M.D. Lieberman, W.A. Cunningham, Type I and Type II error concerns in fMRI research: Re-balancing the scale., Soc. Cogn. Affect. Neurosci. 4 (2009) 423–8. doi:10.1093/scan/nsp052.

[29] K.A. Hadland, M.F.S. Rushworth, D. Gaffan, R.E. Passingham, The effect of cingulate lesions on social behaviour and emotion., Neuropsychologia. 41 (2003) 919–31.

[30] G. Silani, C. Lamm, C.C. Ruff, T. Singer, Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments, J. Neurosci. 33 (2013) 15466–15476.

[31] T. Singer, The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research., Neurosci. Biobehav. Rev. 30 (2006) 855–63. doi:10.1016/j.neubiorev.2006.06.011.

[32] J. Radua, M.L. Phillips, T. Russell, N. Lawrence, N. Marshall, S. Kalidindi, W. El-Hage, C. McDonald, V. Giampietro, M.J. Brammer, A.S. David, S.A. Surguladze, Neural response to specific components of fearful faces in healthy and schizophrenic adults, Neuroimage. 49 (2010) 939–946. doi:10.1016/j.neuroimage.2009.08.030.

[33] F. Van Overwalle, K. Baetens, Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis, Neuroimage. 48 (2009) 564–584. doi:10.1016/J.NEUROIMAGE.2009.06.009.

[34] J. Moll, R. Zahn, R. de Oliveira-Souza, F. Krueger, J. Grafman, The neural basis of human moral cognition, Nat. Rev. Neurosci. 6 (2005) 799–809. doi:10.1038/nrn1768.

[35] C.E. Forbes, J. Grafman, The role of the human prefrontal cortex in social cognition and moral judgment, Annu. Rev. Neurosci. 33 (2010) 299–324. doi:10.1146/annurev-neuro-060909-153230.

[36] K. Shanton, A. Goldman, Simulation theory, Wiley Interdiscip. Rev. Cogn. Sci. 1 (2010) 527–538. doi:10.1002/wcs.33.

[37] M. Coltheart, What has functional neuroimaging told us about the mind (so far)?, Cortex. 42 (2006) 323–31.

[38] J. Grèzes, J.L. Armony, J. Rowe, R.E. Passingham, Activations related to "mirror" and "canonical" neurons in the human brain: An fMRI study., Neuroimage. 18 (2003) 928–37.

[39] E. Kohler, C. Keysers, M.A. Umiltà, L. Fogassi, V. Gallese, G. Rizzolatti, Hearing sounds, understanding actions: Action representation in mirror neurons, Science (80-.). 297 (2002) 846–848. doi:10.1126/science.1070311.

[40] G. Rizzolatti, L. Fogassi, V. Gallese, Neurophysiological mechanisms underlying the understanding and imitation of action, Nat. Rev. Neurosci. 2 (2001) 661–670. doi:10.1038/35090060.

[41] N. Ramnani, R.C. Miall, A system in the human brain for predicting the actions of others,

Nat. Neurosci. 7 (2004) 85–90. doi:10.1038/nn1168.

[42] R.B. Adams, N.O. Rule, R.G. Franklin, E. Wang, M.T. Stevenson, S. Yoshikawa, M. Nomura, W. Sato, K. Kveraga, N. Ambady, Cross-cultural reading the mind in the eyes: An fMRI investigation, J. Cogn. Neurosci. 22 (2010) 97–108. doi:10.1162/jocn.2009.21187.

[43] B. Gunther Moor, Z.A. Op de Macks, B. Guroglu, S.A.R.B. Rombouts, M.W. Van der Molen, E.A. Crone, Neurodevelopmental changes of reading the mind in the eyes, Soc. Cogn. Affect. Neurosci. 7 (2012) 44–52. doi:10.1093/scan/nsr020.

[44] S. Overgaauw, A.C.K. Van Duijvenvoorde, B.G. Moor, E.A. Crone, A longitudinal analysis of neural regions involved in reading the mind in the eyes, Soc. Cogn. Affect. Neurosci. 10 (2015) 619–627. doi:10.1093/scan/nsu095.

[45] O. Dal Monte, S. Schintu, M. Pardini, A. Berti, E.M. Wassermann, J. Grafman, F. Krueger, The left inferior frontal gyrus is crucial for reading the mind in the eyes: Brain lesion evidence., Cortex. 58 (2014) 9–17. doi:10.1016/j.cortex.2014.05.002.

[46] M. Iacoboni, R.P. Woods, M. Brass, H. Bekkering, J.C. Mazziotta, G. Rizzolatti, Cortical mechanisms of human imitation, Science (80-.). 286 (1999) 2526–2528. doi:10.1126/science.286.5449.2526.

[47] G. Liakakis, J. Nickel, R.J. Seitz, Diversity of the inferior frontal gyrus—A meta-analysis of neuroimaging studies, Behav. Brain Res. 225 (2011) 341–347. doi:10.1016/j.bbr.2011.06.022.

[48] M. Iacoboni, A. Brain, Understanding others: Imitation, language, empathy, in: Perspect. Imitation From Neurosci. to Soc. Sci., 2005: pp. 77–99.

[49] M. Iacoboni, Imitation, empathy, and mirror neurons, Annu. Rev. Psychol. 60 (2009) 653–670. doi:10.1146/annurev.psych.60.110707.163604.

[50] M. Arbib, The mirror system, imitation, and the evolution of language, Imitation Anim. Artifacts. (2002) 229–280.

[51] M.P. Richardson, B.A. Strange, R.J. Dolan, Encoding of emotional memories depends on amygdala and hippocampus and their interactions, Nat. Neurosci. 7 (2004) 278–285. doi:10.1038/nn1190.

[52] K. Hirayama, [Thalamus and Emotion]., Brain Nerve. 67 (2015) 1499–508. doi:10.11477/mf.1416200328.

[53] J. Kemp, M.-C. Berthel, A. Dufour, O. Després, A. Henry, I.J. Namer, M. Musacchio, F. Sellal, Caudate nucleus and social cognition: Neuropsychological and SPECT evidence from a patient with focal caudate lesion, Cortex. 49 (2013) 559–571. doi:10.1016/j.cortex.2012.01.004.

[54] T. Singer, The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research, Neurosci. Biobehav. Rev. 30 (2006) 855–863. doi:10.1016/j.neubiorev.2006.06.011.

[55] E. Oztop, M. Kawato, M.A. Arbib, Mirror neurons: Functions, mechanisms and models, Neurosci. Lett. 540 (2013) 43–55. doi:10.1016/j.neulet.2012.10.005.

[56] R.L. Buckner, F.M. Krienen, A. Castellanos, J.C. Diaz, B.T.T. Yeo, The organization of the human cerebellum estimated by intrinsic functional connectivity., J. Neurophysiol. 106 (2011) 2322–45. doi:10.1152/jn.00339.2011.

[57] F. Van Overwalle, T. D'aes, P. Mariën, Social cognition and the cerebellum: A meta-analytic connectivity analysis, Hum. Brain Mapp. 36 (2015) 5137–5154. doi:10.1002/hbm.23002.

[58] A.S. Therrien, A.J. Bastian, Cerebellar damage impairs internal predictions for sensory and motor function, Curr. Opin. Neurobiol. 33 (2015) 127–133. doi:10.1016/J.CONB.2015.03.013.

[59] O. Oullier, K.J. Jantzen, F.L. Steinberg, J.A.S. Kelso, Neural substrates of real and imagined sensorimotor coordination, Cereb. Cortex. 15 (2005) 975–985. doi:10.1093/cercor/bhh198.

[60] R. Adolphs, The neurobiology of social cognition, Curr. Opin. Neurobiol. 11 (2001) 231–239. doi:10.1016/S0959-4388(00)00202-6.

Figure Legend:

*Figure 1.*Task Design and Example Stimuli. Reading the Mind in the Eyes (*left*): the correct answers for the example stimuli pictured here are **"regretful"** (emotion condition) and **"male"** (gender condition). Reading the Mind in the Voice (*middle*): the correct choice for the emotion condition trial presented here is **"worried"**. Causality task (*right*): the correct answers for the example stimuli pictured here are **"(c)" (**intentional condition) and **"(a)"** (physical condition). s, seconds.

Figure 2. Independent components for RMIE where the temporal correlation with the emotion condition exceeded .10. L, Left; R, Right. The number below each image corresponds to the slice number. Region and coordinate information for the RMIE components are presented in Supplemental Table 1. Maps are thresholded at $p < .001$, $k = 74$.

Figure 3. Independent components for RMIV where the temporal correlation with the emotion condition exceeded .10. L, Left; R, Right. The number below each image corresponds to the slice number. Region and coordinate information for the RMIV components are presented in Supplemental Table 2. Maps are thresholded at $p < .001$, $k = 74$.

Figure 4. The independent component for the Causality task where the temporal correlation with the intentional causality condition exceeded .10. L, Left; R, Right. The number below each image corresponds to the slice number. Region and coordinate information for the Causality component are presented in Supplemental Table 3. Maps are thresholded at $p < .001$, $k = 74$.