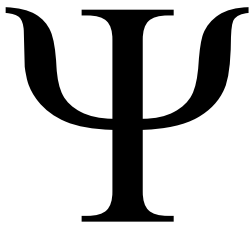
DET PSYKOLOGISKE FAKULTET



Effective Connectivity in the Theory of Mind Network in Typically Developing Individuals and Individuals with Autism Spectrum Disorder

HOVEDOPPGAVE

profesjonsstudiet i psykologi

Anna Edit Ring

Kristine Fagerheim Sættem

Høst 2020

Supervisor: Prof. Dr. Karsten Specht

Department of Biological and Medical Psychology

University of Bergen

Preface

The datasets included in this study were obtained from the Human Connectome Project and the second iteration of the Autism Brain Imaging Exchange. Data from the Human Connectome Project was provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The Human Connectome Project dataset was preprocessed by the Human Connectome Project.

The second iteration of the Autism Brain Imaging Data Exchange was funded by the NIMH 5R21MH107045. This dataset was preprocessed by our supervisor Prof. Dr. Karsten Specht, at the Department of Biological and Medical Psychology at the University of Bergen. Specht also conducted the Volume-based Statistical Analysis and the Dynamic Causal Modelling analysis. The statistical analyses to further investigate the effective connectivity strengths in the Dynamic Causal Modelling were conducted by us.

The thesis is written according to the American Psychological Association 6th guidelines. Figures and tables are presented in the text.

We want to thank our supervisor Karsten Specht for excellent, corona-friendly supervision over Teams. We would also like to thank the Re:State Research Group at the Department of Biological and Medical Psychology at the University of Bergen for including us in their group.

Abstract

This study aims to examine differences in effective connectivity (EC) between brain regions involved in Theory of Mind (ToM) during task functional magnetic resonance imaging (T-fMRI) in typically developing (TD) adults and during rest (R-fMRI) in TD adults and adults with Autism Spectrum Disorder (ASD). Data were obtained from the Human Connectome Project (HCP) and the Autism Brain Imaging Data Exchange (ABIDE II). The HCP dataset contained T-fMRI data (Frith-Happé Animations) from 50 TD adults (27 females, 23 males, age range: 22-35). The ABIDE II dataset contained R-fMRI data from 35 TD adults (15 females, 20 males, M_{age} = 24 years, age range: 18-30 years) and 24 adults with ASD (four females, 20 males, M_{age} = 22 years, age range: 18-31 years). Dynamic Causal Modelling (DCM) was used to calculate EC in a network of regions activated by the task. Spectral DCM was used to analyze R-fMRI data based on the same regions. Results revealed significant differences in EC when comparing ABIDE TD to HCP TD (eta squared effect size (η^2) ranged from .09 to .72). Fewer differences were found in EC when comparing ABIDE ASD and HCP TD (η^2 ranged from .10 to .65) against ABIDE TD and HCP TD. No significant differences in EC were found when comparing ABIDE TD to ABIDE ASD. Overall, we cannot conclude that differences in EC were significant between the chosen regions involved in ToM when comparing R-fMRI in TD adults and adults with ASD.

Keywords: theory of mind, autism spectrum disorder, functional magnetic resonance imaging, effective connectivity, functional connectivity, dynamic causal modelling, human connectome project, autism brain imaging data exchange

Sammendrag

Formålet med denne studien er å undersøke forskjeller i effektiv konnektivitet (EK) mellom hjerneområder involvert i Theory of Mind (ToM) under oppgavebasert funksjonell magnetresonans tomografi (T-fMRI) og under hvile (R-fMRI) i typisk utviklede (TU) voksne og voksne med autismespekterforstyrrelser (ASF). Data ble hentet fra Human Connectome Project (HCP) og Autism Brain Imaging Data Exchange (ABIDE II). Datasettet fra HCP inneholdt T-fMRI (Frith-Happé Animations) fra 50 TU voksne (27 kvinner, 23 menn, aldersspenn: 22-35). Datasettet fra ABIDE II inneholdt R-fMRI data fra 35 TU voksne (15 kvinner, 20 menn, *M* = 24 år, aldersspenn: 18-30) og 24 voksne med ASF (fire kvinner, 20 menn, *M* = 22 år, aldersspenn: 18-31). Dynamisk kausal modellering ble anvendt for å kalkulere EK i et nettverk av hjerneregioner aktivert av oppgaven. Spektral dynamisk kausal modellering ble anvendt for å analysere R-fMRI data basert på de samme områdene. Resultatene viste signifikante forskjeller i EK ved sammenligning av ABIDE TU og HCP TU (eta kvadrert effektstørrelse (η^2) varierte fra .09 til .72). Færre signifikante forskjeller ble funnet ved sammenligning av ABIDE ASF og HCP TU (η^2 varierte fra .10 til .65) mot ABIDE TU og HCP TU. Ingen signifikante forskjeller ble funnet ved sammenligning av ABIDE TU og ABIDE ASF. Overordnet kan vi ikke konkludere med at forskjeller i EK var signifikante mellom de valgte områdene involvert i ToM ved sammenligning av R-fMRI i TU voksne og voksne med ASF.

Nøkkelord: theory of mind, autismespekterforstyrrelser, funksjonell magnetresonans tomografi, effektiv konnektivitet, funksjonell konnektivitet, dynamisk kausal modellering, human connectome project, autism brain imaging data exchange

Abbreviations

ABIDE = the Autism Brain Imaging Data Exchange; **ABIDE I** = the first iteration of the Autism Brain Imaging Data Exchange; **ABIDE II** = the second iteration of the Autism Brain Imaging Data Exchange; **ABIDEII-ONRC_2** = the ABIDE II dataset provided by Olin Neuropsychiatry Research Center; **ADHD** = Attention-Deficit/Hyperactivity Disorder; **ADI-R** = the Autism Diagnostic Interview-Revised; **ADOS-G** = the Autism Diagnostic Observation Schedule-Generic; **ALE** = activation likelihood estimation; **AMY** = amygdala; **AnG** = angular gyrus; **APA** = American Psychiatric Association; **ASD** = Autism Spectrum Disorder; **BMS** = Bayesian Model Selection; **BOLD** = blood oxygenation level dependent; **BW** = bandwidth; **DCM** = dynamic causal modelling; **dMRI** = diffusion magnetic resonance imaging; **DMN** = default mode network; **DSM-5** = Diagnostic and Statistical Manual of Mental Disorders, fifth edition; **EC** = effective connectivity; **EEG** = electroencephalography; **EPI** = echo-planar imaging; **FA** = flip angle; **FC** = functional connectivity; **FCP** = the 1000 Functional Connectomes Project; **fMRI** = functional magnetic resonance imaging; **FOV** = field-of-view; **FWE** = family-wise error; **FWHM** = full-width-at-half-maximum; **HCP** = Human Connectome Project; **Hz** = hertz; **IFG** = inferior frontal gyrus; **INDI** = International Neuroimaging Data-Sharing Initiative; **IOG** = inferior occipital gyrus; **IRB** = Institutional Review Board; **ITG** = inferior temporal gyrus; **KSADS-PL** = the Schedule of Affective Disorders and Schizophrenia for Children-Present and Lifetime Version; ***M*** = mean; ***Mdn*** = median; **MEG** = magnetoencephalography; **MFG** = medial frontal gyrus; **MNI** = Montreal Neurological Institute; **mPFC** = medial prefrontal cortex; **MT** = middle temporal visual area; **MTG** = middle temporal gyrus; **mT/m** = Millitesla Per Meter; **pCC** = posterior cingulate cortex; **PEB** = Parametric Empirical Bayes; **PreC** = precuneus; **pSTG** = posterior superior temporal gyrus; **pSTS** = posterior superior temporal sulcus; **REC** = Regional Committees for Medical and Health Research Ethics; **R-fMRI** = resting state functional magnetic resonance

imaging; **ROI** = region of interest; **SBRef** = single-band reference; **SCID-I RV** = Structured Clinical Interview for DSM-IV Axis I Disorders-Research Version; **SCQ** = Social Communication Questionnaire; **SD** = standard deviation; **SFG** = superior frontal gyrus; **SFS** = superior frontal sulcus; **SMA** = supplementary motor area; **SMG** = supramarginal gyrus; **SNR** = signal to noise ratio; **spDCM** = spectral DCM; **SSAGA** = Semi-Structured Assessment for the Genetics of Alcoholism; **STG** = superior temporal gyrus; **STS** = superior temporal sulcus; ***T*** = *T*-test; **TD** = typically developing; **TE** = echo time; **T-fMRI** = task evoked fMRI; **TG** = temporal gyrus; **ToM** = Theory of Mind; **TP** = temporal pole; **TPJ** = temporo-parietal junction; **TR** = repetition time; ***U*** = Mann-Whitney U; **V5** = middle temporal visual area; **vmPFC** = ventromedial prefrontal cortex; **VOE** = violation-of-expectation; **WAIS-III** = Wechsler Adult Intelligence Scale - Third Edition; ***z*** = Standardized Test Statistics; **3T** = 3 Tesla; η^2 = eta squared effect size

Table of Contents

Theory of Mind

The ability to take another person's point of view, and to explain and predict the person's behavior as determined by their desires, attitudes, intentions, emotions, goals, and beliefs, is essential for making social interaction and communication effective (Gallagher & Frith, 2003; Kana et al., 2015; Rubio-Fernández & Geurts, 2013). This cognitive mechanism, referred to as "theory of mind" (ToM), "mentalizing", "mindreading", "intentional stance", and "social intelligence", enables us to understand false beliefs about the world, to deceive and understand deception, to read body language, to empathize, and to predict and explain other people's behavior (Frith & Frith, 2003; Gallagher & Frith, 2003; Kana et al., 2015; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Sidera, Perpiñà, Serrano, & Rostan, 2018). Imagining a person without the ability to have a ToM, highlights its significance for social interaction and communication: the person would have trouble understanding and adapting to other people's behavior, emotions, and mental states, which in turn would lead to difficulties regulating own behavior, misunderstandings, and in general difficulties interacting with other people in socially normative ways (Brewer, Young, & Barnett, 2017; Byom & Mutlu, 2013; Rubio-Fernández & Geurts, 2013; Sidera, Perpiñà, Serrano, & Rostan, 2018).

ToM is a topic of interest within several disciplines. Researchers within developmental psychology have long been interested in uncovering how and when typical, as well as atypical, ToM development takes place, whereas neuroscientific researchers are interested in uncovering which brain regions and which cognitive processes and mechanisms are involved (Mahy, Moses, & Pfeifer, 2014; Mar, 2011; Schurz & Perner, 2015). Acknowledging the importance of ToM for social functioning, researchers within clinical psychology are interested in studying clinical populations in which social interaction and

communication is difficult, such as in individuals with autism spectrum disorder (Baron-Cohen, Leslie, & Frith, 1985; Byom & Mutlu, 2013).

Autism Spectrum Disorder

Autism spectrum disorder (ASD) refers to a spectrum of disorders sharing similar developmental characteristics, including classic autism or Kanner's Syndrome, Childhood Disintegrative Disorder, Asperger's Syndrome, Rett Syndrome, and Pervasive Developmental Disorder-Not Otherwise Specified (Kana, Libero, & Moore, 2011; Lai, Lombardo, & Baron-Cohen, 2014).

ASD is a neurodevelopmental condition with a global population prevalence of around 1% (Lai, Lombardo, & Baron-Cohen, 2014). It is possible to diagnose ASD from the age of 18 months (Dawson et al., 2012). However, females are often diagnosed later than males, and more males than females are diagnosed with the disorder in general (Green, Travers, Howe, & McDougle, 2019). Although the reason for this is not clear, it can be due to sex differences in genes and hormones, or due to differences in how the symptoms present themselves in males and females (Halladay et al., 2015; Werling & Geschwind, 2013).

Thus, ASD is a heterogeneous condition (Hull et al., 2017). Whereas some individuals with ASD never develop functional speech, others develop speech late in childhood but continue struggling using language in socially appropriate ways (Kana, Libero, & Moore, 2011). Many individuals with ASD have additional problems such as intellectual disabilities, executive dysfunctions, epilepsy, Attention-Deficit/Hyperactivity Disorder (ADHD), anxiety disorders, and more (Happé, Booth, Charlton, & Hughes, 2006; Mannion, Leader, & Healy, 2013). However, restricted repetitive and stereotyped patterns of behavior, interests and activities, as well as difficulties with social communication and interaction, are central features and criteria of the disorder (American Psychiatric Association [APA], 2013, p. 50; Kana et al., 2015). According to the fifth edition of the Diagnostic and Statistical Manual of

Mental Disorders (DSM-5), these features must cause clinically significant impairments in important areas of functioning, and they must present themselves early in development (APA, 2013, p. 50).

Typical Development of Theory of Mind

The original paradigm created to test for the presence of ToM in typically developing (TD) individuals was first described by Wimmer and Perner in 1983. In a series of experiments, they presented children with a story character named Maxi. Maxi has a chocolate, in which he puts in location x. Then, Maxi's mother enters the room and moves the chocolate from location x to location y without Maxi knowing. To pass the test and demonstrate the possession of a ToM mechanism, the child has to indicate where Maxi will look for the chocolate when he returns (location x), and thus, prove that he or she is aware that another individual's belief or representation about the world can contrast with reality (false belief), and that other people can perceive the world from a different viewpoint than the child itself (Gallagher & Frith, 2003; Wimmer & Perner, 1983). Wimmer & Perner (1983) concluded that TD individuals have this knowledge between the age of four and six years.

Today, almost 40 years after their pioneering study, several studies on ToM have been conducted and several paradigms have been developed to test for the presence of ToM. Research generally shows that TD individuals pass verbal first-order false belief tasks (i.e., the ability to understand another person's false belief) from the age of four, and verbal second-order false belief tasks (i.e., the ability to understand a person's false belief about another person's belief) from the age of six (Arslan, Taatgen, & Verbrugge, 2017; Bauminger-Zviely, 2013; Mar, 2011; Perner & Wimmer, 1985; Rubio-Fernández & Geurts, 2013).

However, there is also evidence that ToM can be present even earlier. For example, it has been argued that when TD toddlers at approximately 18 months of age start to play

pretense, it can be interpreted as evidence that they have developed the ability to mentalize (Frith & Frith, 2003; Leslie, 1987). Pretend play requires the capacity to share attention with another person toward a third object or event (joint attention), which in turn requires the ability to mentalize (Akhtar & Gernsbacher, 2007; Frith & Frith, 2003; Frith & Happé, 1994; Mundy & Newell, 2007). Furthermore, when using the violation-of-expectation (VOE) task, a spontaneous-response task, it has been proven that TD toddlers at the age of 15 months can attribute to an agent a false belief about the hidden location of a toy when using the length of looking time as an implicit measure of ToM (Baillargeon, Scott, & He, 2010; Onishi & Baillargeon, 2005).

Atypical Development of Theory of Mind

Signs of ASD often emerge during the first years of life through a process of diminishing, delayed, or atypical development of behaviors considered important in social communication (Lai, Lombardo, & Baron-Cohen, 2014; Landa & Garrett-Mayer, 2006). For example, children with ASD often pay less attention to human voices and faces, and they often engage less in joint attention and pretend play than TD children (Campbell, Leezenbaum, Mahoney, Moore, & Brownell, 2016; Lai, Lombardo, & Baron-Cohen, 2014). Furthermore, Baron-Cohen, Leslie, and Frith (1985) discovered that children with ASD failed an adopted version of the false-belief task, even when controlling for intelligence, whether the child had knowledge of which doll was which (naming question), whether the child had knowledge of the actual location of the doll (reality question), and whether the child's memory was correct regarding the previous location of the doll (memory question). More recent research supports these findings: individuals with ASD generally perform poorer on false-belief tasks compared to TD controls (Castelli, Frith, Happé, & Frith, 2002; Girli & Tekin, 2010; Happé et al., 1996).

False-belief tasks typically require the individual to reflect on a given situation and verbally explain the behavior of the story characters in retrospect ("off-line") (Gallagher & Frith, 2003). It should be noted that some individuals diagnosed with ASD can pass first-order, as well as and second-order false belief tasks, by using cognitive compensation strategies and logical reasoning (Fletcher et al., 1995; Happé et al., 1996; Zalla & Korman, 2018). These individuals are typically older, they have a greater vocabulary, and they are intellectually more functioning compared to other individuals diagnosed with the disorder (Fletcher et al., 1995; Happé et al., 1996; Zalla & Korman, 2018). However, difficulties with mentalizing have been detected in even high-functioning individuals diagnosed with ASD when they are presented with "on-line" tasks requiring immediate mentalizing (e.g., White, Coniston, Rogers, & Frith, 2011; Castelli, Frith, Happé, & Frith, 2002). One such task is the "Frith-Happé Animations Test", also referred to as "animated shapes task" (Mar, 2011) or just "social animations" (Schurz et al., 2014), in which the individual is presented with short video clips of geometric shapes moving around on a screen in a way that is intended to stimulate to the attribution of mental states by their kinetic properties alone (Castelli, Happé, Frith, & Frith, 2000; White, Coniston, Rogers, & Frith, 2011).

The Neural Basis of Theory of Mind

Advances in neuroscientific research and neuroimaging methods have made it possible to investigate which brain regions and which processes are involved in ToM (Mar, 2011; Schurz & Perner, 2015). In a meta-analysis of neuroimaging studies on ToM published between 1995 and 2010 involving clinical as well as non-clinical adult participants, Mar (2011) categorized studies into story-based (20 samples, $N = 274$) (e.g., false-belief stories) and nonstory-based (43 samples, $N = 623$) (e.g., animated shapes tasks) and looked at meta-analytic overlaps between them. A meta-analysis was also conducted for narrative comprehension studies (stories resembling those we typically read in our daily lives, but not

entailing ToM) (23 samples, $N = 355$). The author used an activation likelihood estimation (ALE) approach, which is a quantitative meta-analytic method used to determine the convergence of foci reported in neuroimaging studies previously published (Eickhoff, Bzdok, Laird, Kurth, & Fox, 2012; Mar, 2011).

ALE results for story-based ToM studies revealed that the cluster with the highest likelihood to be involved in ToM was located at the right angular gyrus (AnG), the right posterior superior temporal sulcus (pSTS), the right temporo-parietal junction (TPJ), and the right posterior superior temporal gyrus (pSTG). A similar activation appeared on the left, but of lower likelihood. The cluster with the second-highest likelihood to be involved in ToM was located in the medial prefrontal cortex (mPFC), and next the precuneus (PreC), extending to the posterior cingulate cortex (pCC). In addition, clusters were found in the superior temporal sulcus (STS) bilaterally, the left temporal pole (TP), the left amygdala (AMY), and the left superior frontal gyrus (SFG).

ALE results for nonstory-based ToM studies revealed that the cluster with the highest likelihood to be involved in ToM was located in the mPFC. A cluster in the right anterior STS, proceeding toward the right TP, was next in likelihood. Similar activations were found on the left, but the likelihood was lower. Next in likelihood was a cluster in the left inferior frontal gyrus (IFG), followed by a cluster in the left posterior middle temporal gyrus (MTG), extending caudally to the pSTS and the TPJ, and then dorsally to the supramarginal gyrus (SMG). Similar activations, but of slightly lower likelihood, appeared in the right. In addition, clusters were found in the PreC, the pCC, the right AMY, the right insula, the superior frontal sulcus (SFS), the medial frontal gyrus (MFG), the fusiform gyrus, the lingual gyrus, the putamen, and the thalamus.

A conjunction analysis revealed an overlap in clusters between the two approaches in the mPFC (predominantly in the right hemisphere), the right STS, the left IFG, the pSTS, the

TPJ, the MTG, and the AnG bilaterally, as well as the PreC and the pCC. The AMY was activated on the left for story-based studies but on the right for nonstory-based studies. Furthermore, the three groups overlapped in the mPFC (predominantly in the right hemisphere), the left IFG, the pSTS, the TPJ bilaterally, and the anterior temporal regions bilaterally (aMTG).

Building on the work by Mar (2011), Schurz et al. (2014) categorized previously published studies into six different task groups with equivalent stimulus-material, instructions, and control conditions (73 samples, $N = 1241$). Only data from nonclinical samples were included. An overlap in activation was found in the mPFC and the TPJ on both sides of the hemispheres in all tasks. The cluster in the right TPJ showed the most robust overlap among all findings. However, a region of interest (ROI) analysis revealed task-related differences in activation. For example, tasks requiring processing of mental perspectives (e.g., false-belief tasks) activated more dorsal/posterior parts of the TPJ, while tasks presenting depictions of human action or behavior (e.g., social animations, rational actions, and mind in the eyes tasks) activated more ventral/anterior parts of the TPJ (the STS).

Further, research generally shows that the activity in brain regions involved in ToM is significantly lower in individuals diagnosed with ASD compared to TD individuals (Castelli, Frith, Happé, & Frith, 2002; Happé et al., 1996). For example, Castelli, Frith, Happé, and Frith (2002) presented ten TD adults and ten high-functioning adults with ASD with the Frith-Happé Animations during PET-scanning. Participants watched short video clips of triangles of different sizes that were persuading, bluffing, mocking, and surprising one another, and thus intending to evoke mental state attribution. Participants also watched geometric shapes moving in random and goal-directed order. Behavioral results indicated that the ASD group gave equally accurate descriptions of the animations moving in random and

goal-directed order as the TD group. However, they gave fewer and less accurate descriptions of the animations intended to evoke the attribution of mental states.

Further, when watching the ToM animations, increased activation was found in the mPFC, the inferior temporal gyrus (ITG) extending to anterior fusiform gyrus and the TP adjacent to the AMY, the STS, and the inferior occipital gyrus (IOG) in both groups. However, activity was significantly lower in all these regions in the ASD group, except for the IOG, which showed the same amount of activity as in the TD group.

Furthermore, studies examining the relationship between brain and behavior, indicate that symptom severity is correlated with lower brain activity (Libero & Kana, 2013).

Brain Connectivity During Task and Rest

Today there is general agreement that the attribution of mental states is a complex cognitive process requiring several brain regions to interact (Carrington & Bailey, 2009; Kana, Keller, Cherkassky, Minshew, & Just, 2009; Kana, Libero, & Moore, 2011; Müller & Fishman, 2018). Thus, equal awareness should be given to how these regions are connected to each other and forming an integrated functional "mentalizing network" (Carrington & Bailey, 2009; Lai, Lombardo, & Baron-Cohen, 2014).

Brain connectivity can be examined by looking at the physical connections between regions regarding white matter tracts and synaptic connections (structural connectivity) (Vissers, Cohen, & Geurts, 2011; Zeidman et al., 2019), by looking at statistical dependencies (i.e., correlations) between different regions (functional connectivity) (Kahan & Foltynie, 2013; Zeidman et al., 2019), or by looking at the directed causal effect one region has on another in a specific direction (effective connectivity) (Rolls et al., 2020; Zeidman et al., 2019).

Due to its relatively high spatial resolution and non-invasiveness, functional magnetic resonance imaging (fMRI) is generally considered the preferred imaging modality when

examining the functional network of the human brain (Friston, 2009; Li, Guo, Nie, Li, & Liu, 2009). Here, changes in the brains' concentration of deoxygenated hemoglobin, evoked by a task or endogenous (spontaneous) brain activity, is detected and measured (Glover, 2011).

The majority of fMRI studies examining functional connectivity (FC) in individuals diagnosed with ASD compared to TD individuals during task performance (T-fMRI) have found reduced FC between cortical regions, and especially between frontal and other cortical regions, as well as between frontal lobe regions in individuals with ASD (Vissers, Cohen, & Geurts, 2011).

For example, Kana et al. (2009) presented 12 high-functioning young adults diagnosed with autism (10 males, two females, M_{age} = 24.6 years, age range: 15.8-36.7 years) and 12 TD young males (M_{age} = 24.4 years, age range: 16.6-31.4 years) with the same Frith-Happé Animations as used by Castelli, Frith, Happé, and Frith (2002), in which the participants were presented with short video clips of triangles that were either persuading, bluffing, mocking, or surprising one another, and thus intending to evoke the attribution of mental states. Participants were also presented with triangles moving in random and in goal-directed order. Results indicated that during the attribution of mental states to the animated figures, the FC in individuals with autism was reduced between frontal (MFG, anterior paracingulate, orbital frontal gyrus) and posterior (right middle and STG) brain regions. Reduced FC was also found in the connections between the frontal lobe and other posterior regions, as well as within the occipital and parietal lobes, in individuals with autism.

It should be noted that some studies have found patterns of both reduced and increased FC in individuals with ASD during task performance. For example, Shukla, Keehn, and Müller (2010) found reduced local connectivity in frontal and parietal regions, but increased connectivity in temporal and parahippocampal regions in 26 adolescents diagnosed with ASD (25 males, one female, M_{age} = 13.7 years, age range: 9-18 years) compared to 29

TD adolescents (28 males, one female, M_{age} = 13.8 years, age range: 8-19 years), while Mizuno, Villalobos, Davies, Dahl, and Müller (2006) found evidence of reduced cortico-cortical FC, but partially increased subcortico-cortical FC, in eight young adult males diagnosed with high-functioning autism (M_{age} = 28.4 years, age range: 15-39) compared to eight TD young adult males (M_{age} = 28.1 years, age range: 21-43 years).

Over the past few years, fMRI studies examining the brain at rest (R-fMRI) have become increasingly popular (Hull et al., 2017; Specht, 2020). Here, endogenous (spontaneous) brain activity at low frequencies (0.01-0.1 Hz) is measured while the individual lies inside the scanner in the absence of any sensory or cognitive stimulus (Kristo et al., 2014; Razi et al., 2017; Smitha et al., 2017). During R-fMRI, the individual is often instructed to relax and to have their eyes shut, open, or fixating at a cross (Gotts, Ramot, Jasmin, & Martin, 2019; Specht, 2020).

Research has shown that even in the absence of cognitively demanding tasks, some brain networks increase in activity (Fox & Greicius, 2010; Raichle et al., 2001). Furthermore, some of these resting-state networks resemble those found through T-fMRI (Biswal, Yetkin, Haughton, & Hyde, 1995; Smith et al., 2009). This makes R-fMRI a promising technique for studying brain connectivity in the absence of a task, and particular in individuals who cannot follow instructions or perform tasks accurately in the scanner (Fox & Greicius, 2010).

One resting-state network that frequently has been identified and investigated is the "default mode network" (DMN) (Fox & Greicius, 2010). The DMN includes brain regions such as the mPFC, the pCC, the PreC, the TPJ, and the hippocampus, among other regions (Padmanabhan, Lynch, Schaer, & Menon, 2017; Raichle et al., 2001). These regions are active when the individual is not involved in any attention-demanding or goal-directed tasks, such as during retrieval of autobiographical memory, when visualizing the future, and when taking others' perspective (Buckner, Andrews-Hanna, & Schacter, 2008; Raichle et al.,

2001). Thus, the DMN is also referred to as the "mentalizing network" due to its involvement in social cognition, including mind-wandering, emotional processing, and mental state inference (Smitha et al., 2017).

Many R-fMRI studies have found reduced FC between regions of the DMN in individuals diagnosed with ASD (Gotts, Ramot, Jasmin, & Martin, 2019; Vissers, Cohen, & Geurts, 2011). For example, Cherkassky, Kana, Keller, & Just (2006) found that even though high-functioning individuals with autism (53 males, 4 females, M_{age} = 24 years) and TD individuals (52 males, 5 females, M_{age} = 24 years) had similar levels of activation in DMN regions during visual fixation, the FC between many of these regions was reduced in the autism group. Reduced FC was especially evident between anterior and posterior brain regions.

However, R-fMRI studies have also found patterns of both reduced and increased FC in individuals diagnosed with ASD. For example, Cheng, Rolls, Gu, Zhang, and Feng (2015) analyzed R-fMRI data from 418 individuals with autism (367 males, 51 females, M_{age} = 17.17 years) and 509 TD individuals (424 males, 85 females, M_{age} = 16.4 years) obtained from the first initiative of the Autism Brain Imaging Data Exchange (ABIDE I). In the group with autism, reduced FC was found in parts of the MTG (the right MTG and the right ITG) with the ventromedial prefrontal cortex (vmPFC), the left SFG, the PreC, and the cuneus. Reduced FC was also found between the PreC and the cuneus, the MTG, and the vmPFC. Also, the postcentral gyrus had reduced FC with the paracentral lobule. However, increased FC was found between the medial thalamus and the left MTG and the STG, as well as between the medial thalamus and the postcentral gyrus in the group with autism.

While FC describes statistical dependencies between brain regions, it does not consider the directed causal connections within the network (Kahan & Foltynie, 2013; Zeidman et al., 2019). Effective connectivity (EC) on the other hand, can provide information

about the direction of communication between different regions of the brain (Kana, Libero & Moore, 2011). One approach used to understand EC is Dynamic Causal Modelling (DCM) (Friston, 2009). DCM aims at estimating the experimental modulation of (intrinsic) self-connections and (extrinsic) connections between active regions during a specific task in a directional way (Hillebrandt, Friston, & Blakemore, 2014).

While many studies have examined FC in individuals diagnosed with ASD, relatively few studies have examined EC in individuals diagnosed with ASD (Deshpande, Libero, Sreenivasan, Deshpande, & Kana, 2013; Kana, Uddin, Kenet, Chugani, & Müller, 2014). This is unfortunate, given that our understanding of the neural basis of ASD is not only depending on knowing something about the strength of the relationship between different regions of the brain (FC), but also in which direction the information transfer (EC) (Deshpande et al., 2013; Mohammad-Rezazadeh, Frohlich, Loo, & Jeste, 2016).

In the first large-scale study examining differences in EC between brain regions in individuals diagnosed with autism compared to TD controls, Rolls et al. (2020) analyzed R-fMRI data between 94 regions in 394 individuals with autism (347 males, 47 females, M_{age} = 16.85 years) and 473 TD individuals (385 males, 88 females; M_{age} = 16.42 years) obtained from the ABIDE I. Rolls et al. (2020) found reduced EC from the MTG and other temporal cortical regions to the PreC and the cuneus in the ASD group. These ECs were in a stronger direction. Reduced EC was also found from the TP to the vmPFC. In addition, increased EC was found from the MTG to the hippocampus. These ECs were also in a stronger direction. Furthermore, increased EC was found from the hippocampus and the AMY to the MTG in the group with autism. These ECs were mainly in a weaker direction. Finally, increased EC was found from frontal cortical regions (FrontalSup and Supmedial) to the MFG, while decreased EC was found from the same frontal cortical regions to the SFG.

In another study, Hillebrandt, Friston, and Blakemore (2014) analyzed T-fMRI data from 132 TD individuals (43 males, 89 females; M_{age} = 30.5 years) obtained from the Human Connectome Project (HCP). Furthermore, they used DCM to investigate the EC between the middle temporal visual area (MT/V5) and the pSTS when participants observed triangles moving in an animate and intentionally way, compared to watching triangles moving inanimate and mechanically (the Frith-Happé Animations Test). Hillebrandt, Friston, and Blakemore (2014) found that forward connectivity from the MT/V5 to the pSTS increased, and that the inhibitory self-connection in the pSTS decreased when participants observed the triangles moving in an animate and intentional way.

The Human Connectome Project and the Autism Brain Imaging Data Exchange

Data repositories containing large samples freely shared with the public, such as the HCP and the ABIDE, are much needed to increase our understanding of the mechanisms underlying typical and atypical brain development.

The HCP consortium directed by Washington University, University of Minnesota, and Oxford University (WU-Minn HCP) was a five-year project aiming to outline the patterns of structural and functional brain connectivity, its variation, and its relationship to behavior in 1200 healthy monozygotic and dizygotic twin pairs and their non-twin siblings, as well as data from individuals not related to each other (Barch et al., 2013; Elam, 2014; Van Essen et al., 2012; Van Essen et al., 2013).

The HCP used imaging modalities such as R-fMRI, T-fMRI, diffusion magnetic resonance imaging (dMRI), T1- and T2-weighted magnetic resonance imaging (MRI), as well as electroencephalography (EEG) and magnetoencephalography (MEG) (Van Essen et al., 2013). To get a deeper understanding of the connection between brain connectivity and human functioning, the HCP gathered behavioral measures of a wide range of domains in cognition, emotion, motor function, and sensation (Barch et al., 2013; Van Essen et al.,

2012). Behavioral measures were mainly drawn from the NIH Toolbox and supplemented by non-Toolbox behavioral measures developed by the Gur laboratory group at the University of Pennsylvania (Van Essen et al., 2012; Barch et al., 2013).

T-fMRI was used by the HCP to outline the relationship between individual differences in the neurobiological substrates of mental processing, and functional and structural connectivity. T-fMRI was also used to characterize and validate the connectivity analyses administered on the structural and functional connectivity data. The Frith-Happé Animations Test was used as a measure of social cognition (ToM) (Barch et al., 2013).

The ABIDE includes resting-state functional and structural brain imaging data, as well as matching phenotypic information, of individuals with ASD and TD controls collected from multiple research sites (ABIDE, 2007; Di Martino et al., 2014). The ABIDE I consists of 1112 datasets in total from 539 individuals with ASD and 573 TD controls (ABIDE, 2016). The second iteration of the ABIDE (ABIDE II) contains even larger and better-characterized samples than ABIDE I and includes 1114 datasets from 521 individuals with ASD and 593 TD controls (age range: 5-64 years) (Di Martino et al., 2017; ABIDE, 2017).

Summary

In summary, several brain regions seem to be involved in ToM processing during story-based and nonstory-based ToM tasks, including the mPFC, the MFG, the SFG, the IFG, the ITG, the TP, the STS, the pSTS, the pSTG, the MTG, the TPJ, the PreC, the pCC, the AnG, and the AMY, among others (Mar, 2011; Schurz et al., 2014). Further, results from studies examining FC in individuals with ASD point towards mixed patterns of both reduced and increased FC between brain regions involved in ToM during task performance and during rest (Cheng et al., 2015; Cherkassky, Kana, Keller, & Just, 2006; Gotts, Ramot, Jasmin, & Martin, 2019; Kana et al., 2009; Mizuno et al., 2006; Shukla, Keehn, & Müller, 2010; Vissers, Cohen, & Geurts, 2011). Furthermore, far less knowledge exists about the

directionality (EC) of these connections in individuals diagnosed with ASD (Deshpande et al., 2013; Kana et al., 2014).

Thus, results from neuroimaging studies are inconsistent and sometimes contradictory. This can be due to the heterogeneous nature of ASD, due to small sample sizes, due to age-related and gender-related factors, or due to the great variability in the neuroscientific methods, analyses, paradigms, stimulus, and designs applied (Hull et al., 2017; Dufour et al., 2013; Kana et al., 2014; Mar, 2011; Moessnang et al., 2020; Schaafsma, Pfaff, Spunt, & Adolphs, 2015). Furthermore, the etiology of ASD is still not well known (Libero & Kana, 2013). This is unfortunate, given that it hampers the chance to detect signs of ASD even earlier in the development, to select the best treatment, and to predict the effect of treatment (ABIDE, 2017), which in turn is important to target the plasticity of the brain (Dawson et al., 2012).

Current Study and Hypotheses

This study aims to examine differences in EC between brain regions involved in ToM processing during T-fMRI (the Frith-Happé Animations Test) in TD adults and R-fMRI in TD adults and adults diagnosed with ASD. Data are obtained from the HCP and the ABIDE II.

At first, we need to identify which brain regions are involved in ToM processing. Given that the ABIDE database only contains R-fMRI data, T-fMRI data from healthy adult participants obtained from the HCP will be analyzed. The regions identified will further be used to define a DCM model that represents the ToM network. Next, the DCM will be applied to examine the EC between these regions. To examine differences in EC during T-fMRI and R-fMRI, the same DCM model will be applied to R-fMRI data from TD adult participants obtained from the ABIDE II database. In the last, and perhaps the most relevant

part of the study, the same DCM model will be applied to R-fMRI data from ASD adult participants obtained from the ABIDE II database.

Based on findings from previous research on brain regions involved in ToM processing during animated shapes tasks/social animations, we *(Hypothesis 1)* expect our analyzes to show activation in at least some of those brain regions most frequently seen to be activated, including the mPFC, the STS, the TP, the IFG, the ITG, the MTG, the pSTS, the TPJ, the SMG, the PreC, the pCC, the IOG, the right AMY, the insula, the SFS, the MFG, the fusiform gyrus, the lingual gyrus, the putamen, and the thalamus.

Further, when comparing R-fMRI in TD adult participants from the ABIDE II to T-fMRI in TD adult participants from the HCP, we *(Hypothesis 2)* expect the EC in the ToM network not to be significantly different for most connections, given that both groups are TD control participants.

When comparing R-fMRI in adult participants diagnosed with ASD from the ABIDE II to T-fMRI in TD adult participants from the HCP, we *(Hypothesis 3a)* expect to see greater deviations in EC than when compared against *Hypothesis 2*.

Finally, when comparing R-fMRI in TD adult participants from the ABIDE II and R-fMRI in adult participants diagnosed with ASD from the ABIDE II, we *(Hypothesis 3b)* expect to see similar deviations in EC as we expect to see in *Hypothesis 3a*.

Method

Ethics

This study is a contribution to the Re:State Research Group at the Department of Biological and Medical Psychology at the University of Bergen. The Regional Committees for Medical and Health Research Ethics (REC) have approved the use of data obtained from open-access neuroimaging databases.

The Human Connectome Project ethics. Participants' privacy was protected by the HCP by implementing a two-tiered strategy for data sharing. Information regarding family structure, handedness, exact age by year, body weight, height, and more, are considered information that increases the risk for identifying the subjects and is therefore restricted. In this study, we only used Open Access Data. This includes imaging data and most of the behavioral data not considered de-identifying. The Open Access Data is made freely available for those who register an account at "ConnectomeDB". However, all users must get approval from their local Institutional Review Board (IRB) or Ethics Committee to use the data, and all users must agree to the Open Access Data Use Terms written by the HCP (HCP, n.d.).

The Autism Brain Imaging Data Exchange II ethics. Prior to their research, contributing research sites to the ABIDE II had to verify that their local IRB or Ethics Committee had approved the data collection and the further distribution of the datasets. They also had to anonymize data by changing local subject identification numbers with the 1000 Functional Connectomes Project (FCP) or International Neuroimaging Data-Sharing Initiative (INDI) identifiers, and by removing any information regarding facial features from the images by using the FCP/INDI anonymization script. Information regarding phenotypic information was also reviewed and de-identified (Di Martino et al., 2017).

Participants

The Human Connectome Project participants. The HCP dataset included in our analyses were obtained from the WU-Minn HCP Young Adult study. The dataset, named "U100", included data from 100 unrelated young adults (Elam, 2014). Data from 50 unrelated young adults (27 females, 23 males, age range: 22-35) were further pseudo randomly drawn from the "U100" dataset to ensure an equal gender balance. All subjects gave informed consent prior to participation (Van Essen et al., 2013).

Participants had no documented history of psychiatric, neurologic, neuropsychiatric, or neurodevelopmental disorders based on the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA). Individuals with diabetes and high blood pressure were excluded (Van Essen et al., 2013). For more information regarding the inclusion and exclusion criteria, see Supplemental Table S1 in Van Essen et al. (2013).

The Autism Brain Imaging Data Exchange II participants. The ABIDE II dataset included in our analyses were provided by the Olin Neuropsychiatry Research Center at the Institute of Living at Hartford Hospital in Hartford. The dataset, abbreviated as "ABIDEII-ONRC_2", included data from 59 young adults, in which 35 individuals were TD controls (15 females, 20 males, M_{age} = 24 years, age range: 18-30 years) and 24 individuals were diagnosed with ASD (four females, 20 males, M_{age} = 22 years, age range: 18-31 years) (ABIDE, 2017; Supplementary Table 3 in Di Martino et al., 2017). This dataset was chosen because the characterization of the included participants, in terms of age and gender distribution, was the closest to those from the HCP database. Informed consent was given by all participants in accordance with the Institute of Living at Hartford Hospital Institutional Review Board oversight (ABIDE, 2017).

Full scale IQ was assessed for both groups using the Wechsler Adult Intelligence Scale - Third Edition (WAIS-III): Vocabulary and Block Design subscales. Handedness was assessed with the Edinburgh inventory, version 1 (ABIDE, 2017). IQ range for the TD participants were 85-146 (M = 111, SD = 13), while IQ range for the ASD participants were 80-146 (M = 114, SD = 16). Sixteen individuals with ASD were right-handed, four individuals with ASD were left-handed, while four individuals with ASD were mixed handed (hand preference changes with different tasks), while 32 TD individuals were right-handed, zero TD individuals were left-handed, while three TD individuals were mixed handed (hand preference changes with different tasks) (Supplementary Table 3 in Di Martino et al., 2017).

The TD participants were healthy and had no documented history of psychiatric disorders based on the Structured Clinical Interview for DSM-IV Axis I Disorders-Research Version (SCID-I RV) and the Autism Diagnostic Observation Schedule-Generic (ADOS-G). The Social Communication Questionnaire (SCQ)-lifetime form was completed by the participants' caregivers whenever feasible (ABIDE, 2017). Scores from the SCID-I RV, the ADOS-G, and the SCQ-lifetime form are not openly shared by the ABIDEII-ONRC_2 and are therefore not presented.

The ASD participants had to meet the cutoff score for ASD at the ADOS-G. In addition, the Autism Diagnostic Interview-Revised (ADI-R) and the SCQ-lifetime form was completed by the participants' caregivers whenever feasible. Comorbidity was also assessed using the Schedule of Affective Disorders and Schizophrenia for Children-Present and Lifetime Version (KSADS-PL). Individuals were excluded if they three months prior to the evaluation had met the criteria for Posttraumatic Stress Disorder, Manic or Depressive episode, Bipolar Disorder, and Schizophrenia (ABIDE, 2017). Scores from the ADOS-G, the SCQ-lifetime, and the KSADS-PL are not openly shared by the ABIDEII-ONRC_2 and is therefore not presented. The ASD participants had relatively low comorbidity (Di Martino et al., 2017).

Scan Procedures and Experimental Paradigm

The Human Connectome Project scan procedures and experimental paradigm.

Data were collected at the Washington University in St. Louis over two days. All participants were trained in a mock scanner prior to the actual scanning. A structural MRI session, a R-fMRI session, and a T-fMRI session were completed during the first day. A dMRI scan and a second combined R-fMRI and T-fMRI session were completed during the second day. The total duration of the sessions was about four hours, not counting set-up time. All scans were done in a consistent order unless they were considered unusable due to factors such as head

movements. The behavioral assessments were completed by the participants outside of the scanner (Van Essen et al., 2013).

The Frith-Happé Animations Test (Figure 1) was used as a measure of social cognition (ToM). Inside the scanner, participants watched short video clips (20 s) of geometric shapes interacting with each other ("Mental" condition) or moving randomly ("Random" condition). These video clips were developed by Castelli, Happé, Frith, and Frith (2000) and Wheatley, Milleville, and Martin (2007). The participants were asked to evaluate whether the shapes were interacting with each other as if they were taking into account each other's feelings and thoughts, whether they had no interaction (i.e., there is no apparent interaction and their movement seems random), or whether they were "not sure". Each of the two task runs had five video blocks (two "Mental" and three "Random" in one run, three "Mental" and two "Random" in the other run) and five fixation blocks (15 s each) (Barch et al., 2013).

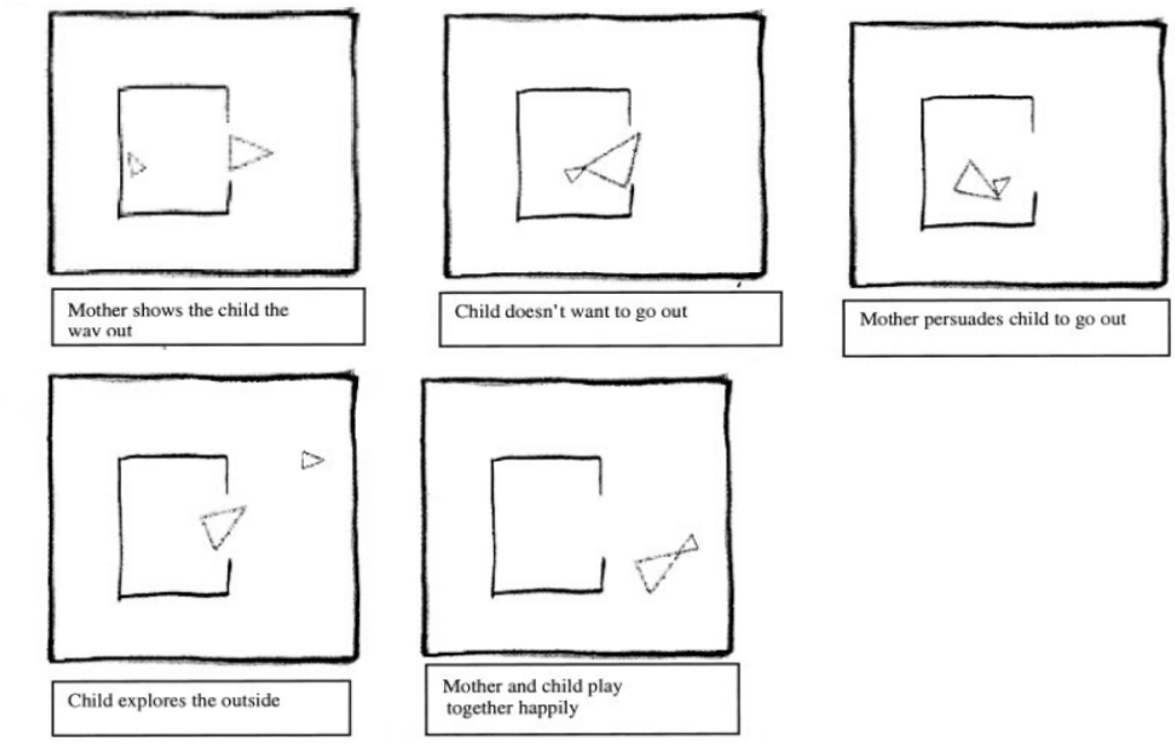


Figure 1. The stills illustrate the Frith-Happé Animation Test which was used as a measure of social cognition (ToM) in the HCP. The Big Triangle is coaxing the reluctant Little Triangle to come out of an enclosure. Participants do not see the captions. Illustration and descriptions are adapted from "Movement and Mind: A Functional Imaging Study of Perception and Interpretation of Complex Intentional Movement Patterns" by F. Castelli, F. Happé, U. Frith, and C. Frith, 2000, *NeuroImage, 12*, 3, p. 323. Copyright 2000, with permission from Elsevier.

The Autism Brain Imaging Data Exchange II scan procedures and experimental paradigm. Data from participants included in the ABIDEII-ONRC_2 dataset were based on R-fMRI. Neither of the participants were trained in a mock scanner. Prior to the scan, participants were given the instructions verbally. Participants were instructed to lie still with their eyes open while fixating on a central cross presented on a mirror attached to the head coil inside the scanner (ABIDE, 2017).

Data Acquisition

The Human Connectome Project data acquisition. The HCP data were acquired with a modified Siemens 3 Tesla (3T) "Connectome Skyra" scanner, using a standard 32 channels head coil. The modified hardware included gradient power amplifiers and a gradient coil that, compared to a standard Skyra scanner, increased the maximum gradient strength from 40 mT/m to 100 mT/m (Uğurbil et al., 2013; Van Essen et al., 2013). Echo-planar imaging (EPI) of the whole brain had the following acquisition parameters: Repetition time (TR) = 720 ms, echo time (TE) = 33 ms, flip angle (FA) = 52° (reduced from 90° to correspond to the Ernst angle and to increase the signal to noise ratio (SNR)), bandwidth (BW) = 2290 Hz/pixels, field-of-view (FOV) = 208 x 180 mm, 72 slices, a multiband acceleration factor of eight, a spatial resolution of 2 mm isotropic voxels, and a matrix size of 104 x 90. Each task was run twice; one with a right-to-left and one with a left-to-right phase encoding direction (Barch et al., 2013; Glasser et al., 2013; Uğurbil et al., 2013; Van Essen et

al., 2013). For a comprehensive overview of the HCP data acquisition, see Uğurbil et al. (2013).

The Autism Brain Imaging Exchange II data acquisition. The ABIDEII-ONRC_2 data were acquired with a Siemens 3T "MAGNETOM Skyra syngo MR D13", using a 32-channels head coil (Di Martino et al., 2017; ABIDE, 2016). The EPI sequences had the following acquisition parameters: TR = 475 ms, TE = 30 ms, FA = 60°, BW = 2604 Hz/pixels, 48 slices, a multiband acceleration factor of eight, and a spatial resolution of 3 mm isotropic voxels (Di Martino et al., 2017).

To increase the SNR for the images, a minimum of three anatomical scans were gathered whenever feasible. Further, a minimum of one R-fMRI scan was gathered. Images were inspected before the subject left the scanner, and scans of low quality were repeated (ABIDE, 2017). Only the scans with the best quality were used for this study.

Preprocessing

The Human Connectome Project preprocessing. The data obtained from the HCP were preprocessed by the HCP using their "minimal preprocessing pipelines", based on the fMRIVolume pipeline. This pipeline adjusts the gradient-nonlinearity-induced distortion and is followed by a realignment of the timeseries to the single-band reference (SBRef) image to adjust for participants' movement in the scanner. Also, the EPI distortion was adjusted by registering the SBRef image to the T1-weighted image. The transformation into the Montreal Neurological Institute (MNI) space was applied to the original fMRI timeseries by incorporating all transformations described above in a single spline interpolation resampling step. The MNI space timeseries were masked and normalized to a 4D whole-brain mean of 10.000. The outputs were resampled with a cubic voxel size of 2 mm (Glasser et al., 2013). For a detailed description of the HCP preprocessing pipelines, see Glasser et al. (2013).

After downloading these minimally preprocessed data, the data were smoothed with an 8 mm full-width-at-half-maximum (FWHM) Gaussian kernel, using SPM12.

The Autism Brain Imaging Exchange II preprocessing. The data from ABIDEII-ONRC_2 was not preprocessed in advance. Thus, to make data from the HCP and the ABIDE II as similar as possible, preprocessing methods comparable to those used on the HCP data were applied for the ABIDEII-ONRC_2 dataset as well. The preprocessing and denoising tools implemented in the software package SPM12 Statistical Parametric Mapping (https://www.fil.ion.ucl.ac.uk/spm/) running on MATLAB R2015 for Windows were used. The functional images were realigned, movement-related image distortions were removed ("unwarping"), co-registered with the structural images, normalized to MNI space, and smoothed with an 8 mm FWHM Gaussian kernel.

Volume-based Statistical Analysis

The Human Connectome Project volume-based statistical analysis. The first-level analysis was conducted for each participant separately, using a general linear model applied to the T-fMRI timeseries. The model specified the two main conditions: the fixation-blocks and the moving triangles. The latter condition was further divided into a "Random" and a "Mental" condition. Additionally, the initial countdown at the beginning of the fMRI run, the response trials, and the six rigid body motion parameters were included as nuisance covariates. A difference contrast was defined that estimated the differences between the "Mental" and "Random" condition, highlighting the regions involved when an interaction between the triangles was observed.

The second-level analysis was performed as a one-sample t-test that rested on the difference contrast. This analysis revealed for the entire group which brain regions were more active during the "Mental" than the "Random" condition. The results were explored at a family-wise error (FWE) corrected threshold of p (FWE) < .05 with at least ten voxels per

displayed cluster. Based on the group result derived from this second-level analysis of the HCP data, the following brain regions in the right temporal and right occipital lobe were chosen: the pSTS [48 -24 -6], the right TP [50 6 -22], and the right AnG [54 -52 16] with additional sub maxima within the right IOG [34 -88 -4], and the right ITG [52 -62 -8]. The blood-oxygenation-level-dependent (BOLD) timeseries were extracted from these five ROIs using an 8 mm radius sphere.

The Autism Brain Imaging Exchange II volume-based statistical analysis. The first-level analysis was conducted using a general linear model applied to the R-fMRI timeseries. Nuisance covariates included the six rigid body motion parameters, average white matter and cerebrospinal fluid signal timeseries. This procedure was necessary as a pre-analysis to the sub-sequent DCM analysis, and for regressing out the global signals from the white matter and cerebrospinal fluid signal, which are considered as nuisance signals in R-fMRI (Liu, Nalci, & Falahpour, 2017). Accordingly, each individual from the TD and ASD groups were analyzed separately. Further, a second-level analysis was not necessary for R-fMRI. BOLD timeseries were extracted from the same five ROIs as described for the HCP data.

Effective Connectivity Analysis

Dynamic Causal Modelling. In DCM for fMRI, bilinear differential equations explain the fluctuations in neuronal activity x(t) in terms of linearly distinguishable components that reflect the influence of other regional state variables (Friston, 2009). Further, it describes how the BOLD signal x at a given timepoint t changes over time dx/dt, as seen in the following bilinear differential equation (Stephan et al., 2007): $dx / dt = (A + B * u(t)) * x(t) + C * u(t)$.

Thereby, a known deterministic external stimulus u(t), like observing the moving triangles, produces a change in neuronal states directly through a region-specific function (C-

matrix), or increase the coupling parameters of the A-matrix in proportion to the bilinear coupling parameters of the B-matrix. While the time course of the BOLD signal x is measured and the input u (the task) is known, the A-, B-, and C-matrix have to be estimated, using a region-specific hemodynamic model. Thereby, the A-matrix describes the general EC between the regions of interest, the B-matrix describes how the given task influence this connectivity, and the C-matrix describes how and where the external stimulus enters the described network (Friston, 2009; Stephan et al., 2007; Zeidman et al., 2019).

Dynamic Causal Modelling in Human Connectome Project data. The DCMs were created and estimated using DCM12 (version 5370) as implemented in SPM12. As outlined above, the ROIs included were the pSTS, the TP, the AnG, the IOG, and the ITG. The DCM model contained the overall ECs of these regions while watching the Frith-Happé Animations (A-matrix), and a specific modulation of these connections when an interaction between the triangles was detected (B-matrix). Because DCM for T-fMRI is specified as a set of different alternative models (Stephan et al., 2010), the individual DCM models were jointly estimated using the Parametric Empirical Bayes (PEB) framework implemented in SPM12, which includes Bayesian model reduction. Further, we used Bayesian Model Selection (BMS), a post-hoc model selection routine, which allowed us to investigate all possible DCM models in order to select the model that represents the data best (Friston & Penny, 2011; Hillebrandt, Dumontheil, Blakemore, & Roiser, 2013). In the following comparisons between the HCP and ABIDE datasets, only the A-matrix was comparable.

Dynamic Causal Modelling in Autism Brain Imaging Exchange II data. DCM for R-fMRI was based on the same framework as DCM for T-fMRI. However, in the absence of an external input u, the B-matrix and C-matrix was not estimated. Furthermore, DCM for R-fMRI was performed as spectral DCM (spDCM), in which the EC is estimated from the cross spectra of the random intrinsic (spontaneous) fluctuations in the neuronal states than from

their time courses directly (Razi, Kahan, Rees, & Friston, 2014; Razi et al., 2017). The spDCM model was a fully connected model in which each region was connected to the other region. Again, the individual spDCM models were jointly estimated, using the PEB framework implemented in SPM12. This was followed by a Bayesian model reduction to restrict the number of parameters. Finally, connectivity strengths in the A-matrix were extracted from the estimated spDCM models.

Statistical Analysis

The International Business Machines (IBM) SPSS Statistics 25 software was used to investigate the EC strengths in the DCM A-matrixes. Preliminary analyses revealed that the data were not normally distributed when evaluated with the Shapiro-Wilk´s test ($p >$. 05). Furthermore, a visual inspection of the boxplots revealed outliers in the data. Both the violation of normality and the outliers were especially evident in the HCP TD data, in which all parameters had a significant *p*-value less than .001. Further, 20 of 25 connections had a minimum of one extreme outlier, defined as data points that extend more than three interquartile range from the edge of the box (IBM, n.d.). Due to the results of the preliminary analyses, a non-parametric One-Sample Wilcoxon Signed Rank Test was conducted to identify the parameters significantly different from a median of zero in each of the three DCM models (HCP TD, ABIDE II TD, ABIDE II ASD). To compare the DCM models with each other, a non-parametric Independent-Samples Mann-Whitney U Test was conducted. The level of significance was $p < .05$ (two-tailed) for all tests.

A multiple comparison Bonferroni correction was calculated by dividing the *p*-value .05 with the number of tests, which for the Wilcoxon Signed Rank Test was 25. The parameters were considered significantly different from a median of zero if the *p*-value was < .002. Furthermore, for the Mann-Whitney U Test, a multiple comparison Bonferroni

correction was calculated by dividing the *p*-value .05 by 50. As a result, the parameters of the DCM models were considered significantly different if $p < .001$.

The effect size estimate for the Mann-Whitney U Test was calculated as eta squared ($\eta^2 = z^2 / N$) (Fritz, Morris, & Richler, 2012). Here, .01 is interpreted as a small effect size, .06 as a medium effect size, and .14 as a large effect size (Cohen, 1988, pp. 286-287).

Results

Volume-based Statistical Analysis

The second-level analysis of the HCP data revealed for the entire group which brain regions were more active during the "Mental" than the "Random" condition when TD participants watched the Frith-Happé Animation. The coordinates are presented in Table 1, while the distribution of activation is displayed in Figure 2.

The clusters with the highest activation significant at a corrected threshold of *p* (FWE) < .05, and with at least 10 voxels per cluster, included the left MTG [-54 -58 12], the right IFG [52 30 4], the left IOG [-28 -96 -4], the right supplementary motor area (SMA) [8 12 64], the left cerebellum [-24 -76 -36], and the left IFG [-46 14 22]. More importantly, activation was also found within the right pSTS [48 -24 -6], the right TP [50 6 -22], and the right AnG [54 -52 16], with additional sub maxima within the right IOG [34 -88 -4], and the right ITG [52 -62 -8]. Given that the DCM analysis is limited to only a small number of co-activated brain regions (Hillebrandt, Friston, & Blakemore, 2014), we chose the five latter regions located in the right temporal and occipital lobe that was hypothesized to belong to the core areas of the ToM network of the right hemisphere.

Table 1

General Linear Model Results from the Second-Level Analysis

Localization		MNI Coordinates					
Region	Left/Right	*x*	*y*	*z*	Cluster Size[a]	*T*(49)[b]	*p*(FWE)
MTG	L	-54	-58	12	743	9.82	<.001
IFG	R	52	30	4	974	9.77	<.001
IOG	L	-28	-96	-4	169	9.30	<.001
SMA	R	8	12	64	88	8.94	<.001
Cerebellum	L	-24	-76	-36	377	8.73	<.001
IFG	L	-46	14	22	10	5.84	.012
pSTS	R	48	-24	-6	157	8.01	<.001
TP	R	50	6	-22	180	8.06	<.001
AnG[c]	R	54	-52	16	1636	10.97	<.001
IOG[c]	R	34	-88	-4		8.75	<.001
ITG[c]	R	52	-62	-8		8.64	<.001

Note. MTG = middle temporal gyrus; IFG = inferior frontal gyrus; IOG = inferior occipital gyrus; SMA = supplementary motor area; pSTS = posterior superior temporal sulcus; TP = temporal pole; AnG = angular gyrus; ITG = inferior temporal gyrus. Significance level of corrected *p* (FWE) < .05.

[a] Number of voxels (voxel size 2.0 mm x 2.0 mm x 2.0 mm). [b] Peak *T*-test value in clusters (degrees of freedom = 49). [c] Belong to the same activation cluster.

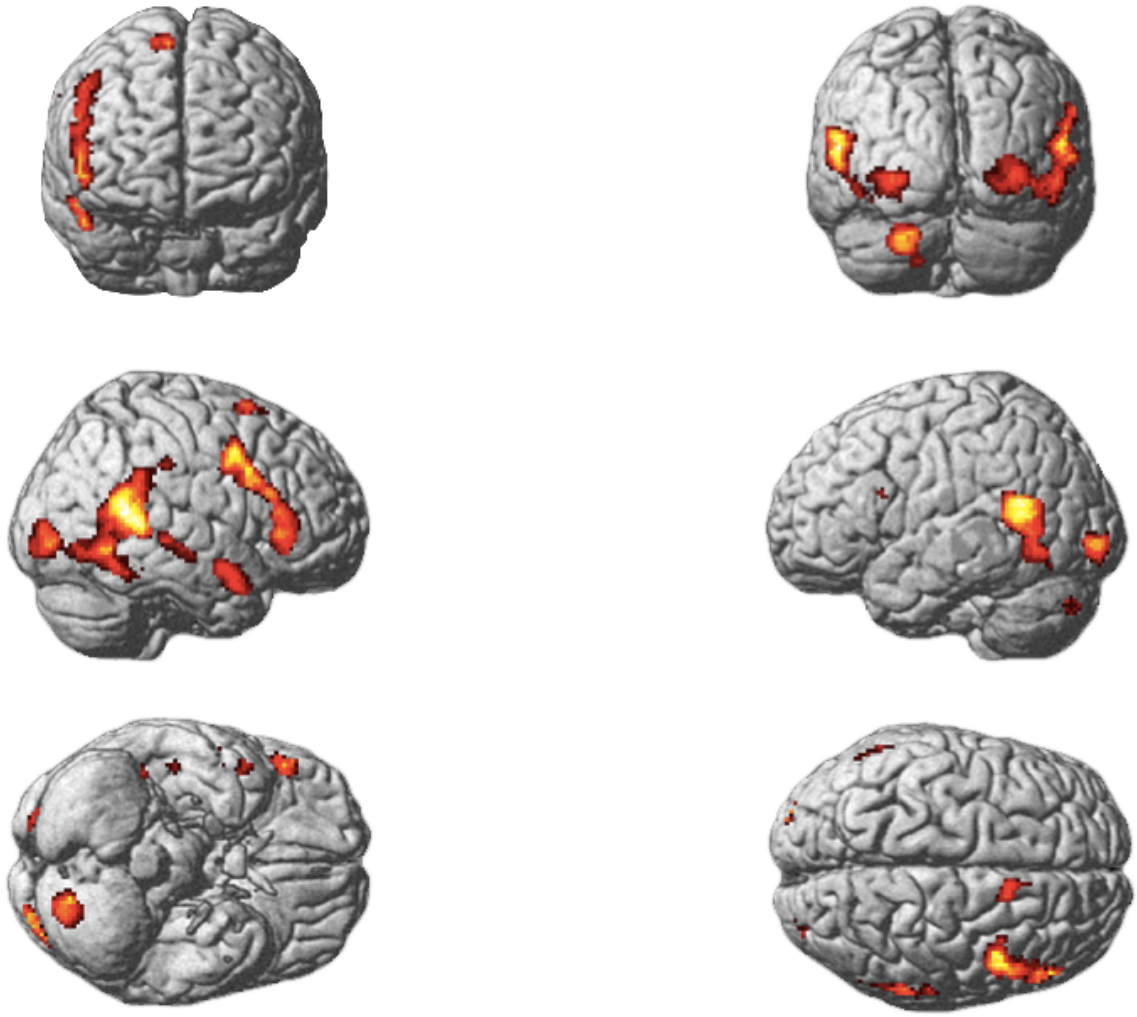


Figure 2. Group activation map. The figure displays those areas that were more active when observing an interaction between the triangles ("Mental") compared to random movements of them ("Random"). The activations are displayed at a corrected threshold of p (FWE) < .05, and at least ten voxels per cluster. The color of the activated areas indicates the significance of the activation, with increasing significance from red to yellow.

Effective Connectivity

Effective Connectivity in the Dynamic Causal Modelling A-matrix. The results from the Wilcoxon Signed Rank Test investigating the EC strengths in the DCM A-matrixes are presented in Table 2. In the HCP TD group, all five regions (the pSTS, the TP, the AnG, the IOG, and the ITG) demonstrated significant self-connections at a significance level of p <

.05 (two-tailed). Three of the five self-connections were significant even at the corrected significance level of $p < .002$. In the ABIDE II TD and the ABIDE II ASD group, all five regions showed significant self-connections at the corrected significance level of $p < .002$. However, the HCP TD groups' A-matrix had the most significant connections in total at both the uncorrected ($p < .05$) and the corrected significance level ($p < .002$). The ABIDE II ASD groups' A-matrix had the least significant connections in total. The five self-connections in the HCP TD group had all negative values, while the five self-connections for both the ABIDE II groups had all positive values. Self-connection parameters are unitless and represent an exponential scaling factor that varies up and down the default value of -0.5 Hz.

If the self-connection has positive values, it is interpreted as if the self-inhibition is increased. The more positive a self-connection parameter is, the more inhibited the region will be, and thus, its sensitivity to inputs from other regions of the network will be reduced. If the self-connection has negative values, it is interpreted as if the self-inhibition is less inhibited. The more negative a self-connection parameter is, the less inhibited the region will be (Zeidman et al., 2019).

All of the between-region connections in the HCP TD group were in the excitatory (stronger) direction, while most of the between-regions connections in both ABIDE II groups were in the inhibitory (weaker) direction. The parameters of the between-region connections are in Hz. Further, the connections with positive values are excitatory, while connections with negative values are inhibitory (Zeidman et al., 2019).

The significant connections in the DCM A-matrix at the corrected significance level of $p < .002$ are displayed in Figure 3, in which each group (HCP TD, ABIDE II TD, ABIDE II ASD) is presented separately.

Table 2

Non-Significant and Significant Connections Different from Zero Based on the Wilcoxon Signed-Rank Test

	HCP TD[a]		ABIDE II TD[b]		ABIDE II ASD[c]	
Regions	*Mdn*	*z*	*Mdn*	*z*	*Mdn*	*z*
pSTS ↺	-0.0004*	-2.74	2.20**	5.16	2.25**	4.29
TP ↺	-0.002**	-3.53	2.34**	5.16	2.14**	4.29
AnG ↺	-0.01**	-4.28	1.16**	5.00	1.06**	4.29
IOG ↺	-0.0003*	-2.93	1.31**	4.95	1.22**	4.20
ITG ↺	-0.01**	-5.52	1.32**	5.11	1.41**	4.29
pSTS → TP	0.004**	3.91	-0.09*	-2.42	-0.15*	-2.40
pSTS → AnG	0.01**	4.01	0.09*	2.25	0.12*	1.97
pSTS → IOG	0.002*	3.16	-0.13*	-2.31	-0.27*	-2.40
pSTS → ITG	0.01**	4.83	-0.13*	-2.38	-0.13	-1.80
TP → pSTS	0.001**	3.48	-0.02	0.67	0.01	-0.06
TP → AnG	0.01**	3.99	0.02	1.44	0.04	0.34
TP → IOG	0.002**	4.15	-0.27**	-3.56	-0.37**	-3.71
TP → ITG	0.01**	4.53	-0.30*	-3.16	-0.25*	-2.71
AnG → pSTS	0.001**	3.43	0.44**	4.34	0.85**	3.60
AnG → TP	0.004**	3.90	0.15	1.75	-0.03	0.31
AnG → IOG	0.002**	3.84	-0.13	-0.49	-0.17*	-2.09
AnG → ITG	0.01**	4.54	0.12	1.17	0.32	1.09
IOG → pSTS	0.001*	2.00	-0.21**	-3.78	-0.17*	-2.09
IOG → TP	0.001*	2.69	-0.54**	-4.77	-0.49**	-4.26
IOG → AnG	0.01*	2.71	-0.45*	-2.57	-0.54*	-2.26

	HCP TD[a]		ABIDE II TD[b]		ABIDE II ASD[c]	
Regions	*Mdn*	*z*	*Mdn*	*z*	*Mdn*	*z*
IOG → ITG	0.02**	5.12	0.46**	3.36	0.26	1.34
ITG → pSTS	0.004**	3.82	-0.28*	-3.01	-0.21	-1.40
ITG → TP	0.004**	3.98	-0.49**	-4.55	-0.46**	-3.86
ITG → AnG	0.01**	3.95	-0.50*	-2.21	-0.33	-0.66
ITG → IOG	0.004**	4.08	0.63**	4.13	0.61	1.91

Note. Mdn = median; *z* = standardized test statistic; ↺ = self-connection; → = between-region connection. pSTS = posterior superior temporal sulcus; TP = temporal pole; AnG = angular gyrus; IOG = inferior occipital gyrus; ITG = inferior temporal gyrus. Self-connection parameters are unitless, and they vary up and down the default value of -0.5 Hz. The parameters of the between-region connections are in Hz (Zeidman et al., 2019). The HCP TD column includes decimals to the first number other than zero.

[a] $n = 50$. [b] $n = 35$. [c] $n = 24$.

* $p < .05$, two-tailed. ** $p < .002$, two-tailed.

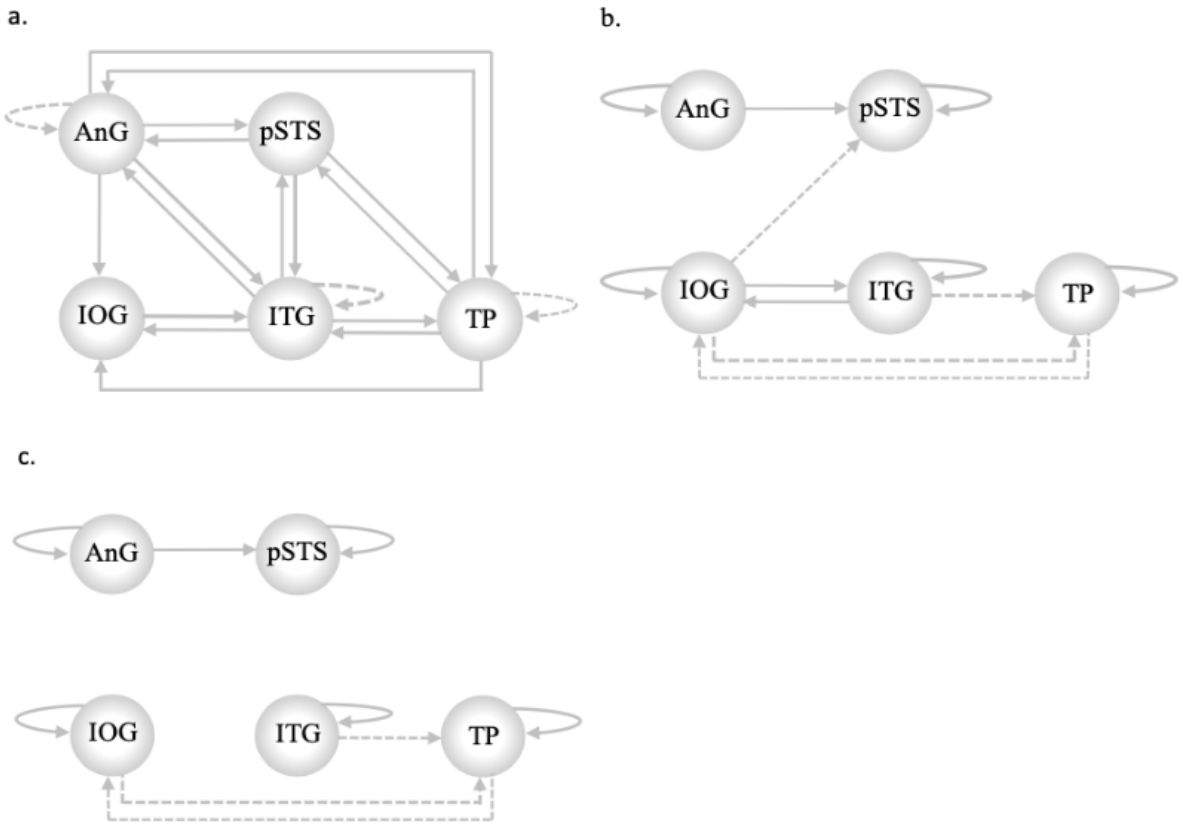


Figure 3. (a) the HCP TD group, (b) the ABIDE II TD group, and (c) the ABIDE II ASD group. Connections with a significant difference from a median of zero based on the one-sample Wilcoxon Signed Rank Test. Only connections significantly different at a corrected significance level of $p < .002$ are displayed. pSTS = posterior superior temporal sulcus; TP = temporal pole; AnG = angular gyrus; IOG = inferior occipital gyrus; ITG = inferior temporal gyrus. Dashed lines indicate inhibition. A higher negative self-connection value indicates less inhibition, while a higher negative between-region connection value indicates more inhibition. Solid lines indicate positive values. A higher positive self-connection value indicates more inhibition, while a higher positive between-region connection value indicates more excitation. The lines' thickness reflects the standardized test statistic (z-score), in which thicker lines indicate a stronger *z*-value (positive and negative) (see Table 2).

Comparison of the Effective Connectivity strength in the Dynamic Causal Modelling Models A-matrixes. The results from the Mann-Whitney U Test comparing the ABIDE II TD group and the HCP TD group are presented in Table 3. Given that the

distributions in the ABIDE II TD group and the HCP TD group were not similar when visually inspected, only the mean ranks were compared. The results revealed a significant difference in all five self-connections (the pSTS, the TP, the AnG, the IOG, and the ITG), even at the corrected significance level of $p < .001$. The eta squared effect size (η^2) indicated large effect sizes for all five self-connections, but especially for the pSTS and the TP ($\eta^2 = .72$). Thirteen of the 20 between-region connections were significantly different at a significance level of $p < .05$, but this was reduced to 11 with the corrected significance level of $p < .001$. The differences when compared to the corrected significance level ($p < .001$), were the connection between the pSTS and the IOG, as well as the connection between the ITG and the AnG. These two connections were significant only at the uncorrected significance level of $p < .05$. The effect size for the between-region connections varied from medium ($\eta^2 = .09$) to large ($\eta^2 = .41$) effect.

The significant differences in the DCM A-matrix between the ABIDE II TD group and the HCP TD group at the corrected significance level of $p < .001$ are displayed in Figure 4.

Table 3

Significant Connections When Comparing the ABIDE II TD Group and the HCP TD Group Based on the Mann-Whitney U Test

	ABIDE II TD[a]	HCP TD[b]				
Regions	*Mean Rank*	*Mean Rank*	*U*	*z*	*p*	η^2
pSTS ↺	68.00	25.50	1750.00	7.81	<.001	.72
TP ↺	68.00	25.50	1750.00	7.81	<.001	.72
AnG ↺	66.54	26.52	1699.00	7.36	<.001	.64
IOG ↺	65.00	27.60	1645.00	6.88	<.001	.56
ITG ↺	65.74	27.08	1671.00	7.11	<.001	.59
pSTS → TP	29.74	52.28	411.00	-4.14	<.001	.20
pSTS → IOG	32.91	50.06	522.00	-3.15	.002	.12
pSTS → ITG	30.97	51.42	454.00	-3.76	<.001	.17
TP → IOG	26.77	54.36	307.00	-5.07	<.001	.30
TP → ITG	26.40	54.62	294.00	-5.19	<.001	.32
AnG → pSTS	60.31	30.88	1481.00	5.41	<.001	.34
IOG → pSTS	28.49	53.16	367.00	-4.54	<.001	.24
IOG → TP	24.29	56.10	220.00	-5.85	<.001	.40
IOG → AnG	31.17	51.28	461.00	-3.70	<.001	.16
ITG → pSTS	27.66	53.74	338.00	-4.80	<.001	.27
ITG → TP	24.06	56.26	212.00	-5.92	<.001	.41
ITG → AnG	33.94	49.34	558.00	-2.83	.005	.09
ITG → IOG	55.80	34.04	1323.00	4.00	<.001	.19

Note. U = Mann-Whitney U; *z* = standardized test statistic; η^2 = eta squared; ↺ = self-connection; → = between-region connection. pSTS = posterior superior temporal sulcus; TP = temporal pole; AnG = angular gyrus; IOG = inferior occipital gyrus; ITG = inferior temporal gyrus. Only connections at a significance level of $p < .05$ are presented in the table. A higher mean rank value in the self-connections indicates that, compared to the other group,

the self-connections are more inhibited. A higher mean rank value for the between-region connections indicates that the connections are more excited or less inhibited (Zeidman et al., 2019).

[a] $n = 35$. [b] $n = 50$.

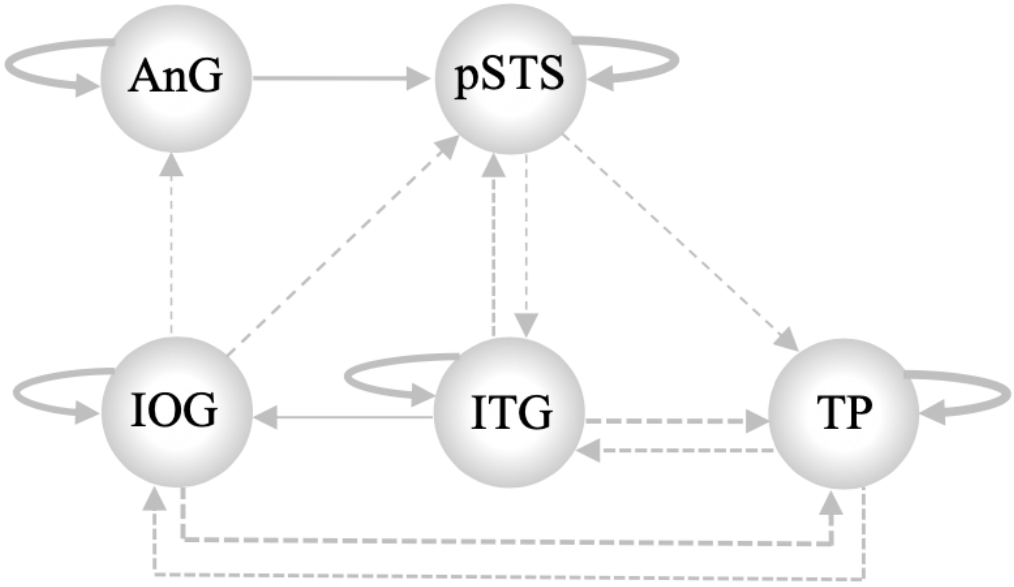


Figure 4. Connections with a significant difference in EC when comparing the ABIDE II TD group to the HCP TD group. Only connections significantly different at a corrected significance level of $p < .001$ are displayed. pSTS = posterior superior temporal sulcus; TP = temporal pole; AnG = angular gyrus; IOG = inferior occipital gyrus; ITG = inferior temporal gyrus. Solid lines indicate a higher mean rank value in the ABIDE II TD group compared to the HCP TD group, while dashed lines indicate a lower mean rank value in the ABIDE II TD group compared to the HCP TD group. A higher mean rank value in the self-connections indicate stronger inhibition, while a higher mean rank value in the between-region connections indicate stronger excitation or less significant inhibition (see Table 1 for the direction of the parameter). The lines' thickness reflects the effect size estimates, in which thicker lines indicate a stronger effect size (see Table 3).

The results from the Mann-Whitney U Test comparing the ABIDE II ASD group and the HCP TD group are presented in Table 4. Given that the distributions in the ABIDE II

ASD group and the HCP TD group were not similar when visually inspected, only the mean ranks were compared. The results revealed a significant difference in all five self-connections (the pSTS, the TP, the AnG, the IOG, and the ITG), even at the corrected significance level of $p < .001$. Similar to the comparison of the ABIDE II TD group to the HCP TD group, the eta squared effect size indicated large effect sizes for all five self-connections. Twelve of the 20 between-region connections were significantly different at a significance level of $p < .05$, but these were reduced to seven at the corrected significance level of $p < .001$. The differences when compared to the corrected significance level ($p < .001$), were the connection between the pSTS and the ITG, the AnG and the IOG, the IOG and the pSTS, the IOG and the AnG, as well as the ITG and the pSTS. These connections were significant only at the uncorrected significance level of $p < .05$. The effect size for the between-region connections varied from medium ($\eta^2 = .10$) to large ($\eta^2 = .54$) effect.

The significant differences in the DCM A-matrix between the ABIDE II ASD group and the HCP TD group at the corrected significance level of $p < .001$ are displayed in Figure 5.

Table 4

Significant Connections When Comparing the ABIDE II ASD Group and the HCP TD Group Based on the Mann-Whitney U Test

	ABIDE II ASD[a]	HCP TD[b]				
Regions	*Mean Rank*	*Mean Rank*	*U*	*z*	*p*	η^2
pSTS ↺	62.50	25.50	1200.00	6.93	<.001	.65
TP ↺	62.50	25.50	1200.00	6.93	<.001	.65
AnG ↺	62.50	25.50	1200.00	6.93	<.001	.65
IOG ↺	59.88	26.76	1137.00	6.20	<.001	.52
ITG ↺	62.46	25.52	1199.00	6.92	<.001	.65
pSTS → TP	23.50	44.22	264.00	-3.88	<.001	.20
pSTS → IOG	24.04	43.96	277.00	-3.73	<.001	.19
pSTS → ITG	25.21	43.40	305.00	-3.41	.001	.16
TP → IOG	20.79	45.52	199.00	-4.63	<.001	.29
TP → ITG	21.33	45.26	212.00	-4.48	<.001	.27
AnG → pSTS	50.08	31.46	902.00	3.49	<.001	.16
AnG → IOG	25.12	43.44	303.00	-3.43	.001	.16
IOG → pSTS	27.92	42.10	370.00	-2.66	.008	.10
IOG → TP	14.58	48.50	50.00	-6.35	<.001	.54
IOG → AnG	26.21	42.92	329.00	-3.13	.002	.13
ITG → pSTS	26.79	42.64	343.00	-2.97	.003	.12
ITG → TP	15.08	48.26	62.00	-6.21	<.001	.52

Note. U = Mann-Whitney U; *z* = standardized test statistic; η^2 = eta squared; ↺ = self-connection; → = between-region connection. pSTS = posterior superior temporal sulcus; TP = temporal pole; AnG = angular gyrus; IOG = inferior occipital gyrus; ITG = inferior temporal gyrus. Only connections at a significance level of $p < .05$ are presented in the table. A higher mean rank value in the self-connections indicates that, compared to the other group, the self-connections are more inhibited. A higher mean rank value for the between-region

connections indicates that the connections are more excited or less inhibited (Zeidman et al., 2019).

[a] $n = 24$. [b] $n = 50$.

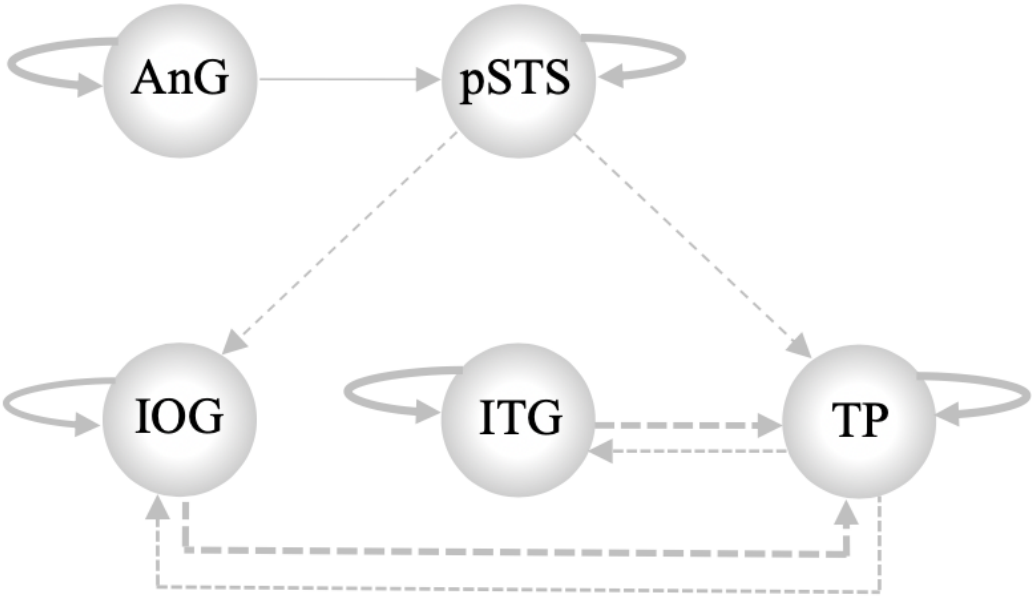


Figure 5. Connections with a significant difference in EC when comparing the ABIDE II ASD group to the HCP TD group. Only connections significantly different at a corrected significance level ($p < .001$) are displayed. pSTS = posterior superior temporal sulcus; TP = temporal pole; AnG = angular gyrus; IOG = inferior occipital gyrus; ITG = inferior temporal gyrus. Solid lines indicate a higher mean rank value in the ABIDE II ASD group compared to the HCP TD group, while dashed lines indicate a lower mean rank value in the ABIDE II ASD group compared to the HCP TD group. A higher mean rank value in the self-connections indicate stronger inhibition, while a higher mean rank value in the between-region connections indicate stronger excitation or less significant inhibition (see Table 1 for the direction of the parameter). The lines' thickness reflects the effect size estimates, in which thicker lines indicate a stronger effect size (see Table 4).

Finally, the Mann-Whitney U Test comparison between the ABIDE II TD group to the ABIDE II ASD group did not reveal any significant results at the uncorrected level of $p < .05$.

Discussion

This study aimed to examine differences in EC between brain regions involved in ToM processing during T-fMRI in TD adults and R-fMRI in TD adults and adults diagnosed with ASD.

We followed a hierarchical strategy by, first, analyzing T-fMRI data from the HCP in which participants performed the Frith-Happé Animations Test. This step was done to identify regions critically involved in ToM processing. Next, a DCM analysis of these data was conducted to identify the EC of the regions involved in ToM processing. Finally, we compared data from a R-fMRI study conducted in TD adults and adults diagnosed with ASD to these results.

Based on findings from previous research on brain regions involved in ToM processing during animated shapes tasks/social animations tasks, we *(Hypothesis 1)* expected our analyzes to show activation in at least some of those brain regions most frequently seen to be activated, including the mPFC, the STS, the TP, the IFG, the ITG, the MTG, the pSTS, the TPJ, the SMG, the PreC, the pCC, the IOG, the right AMY, the insula, the SFS, the MFG, the fusiform gyrus, the lingual gyrus, the putamen, and the thalamus. Results from the second-level analysis revealed activation in several different areas (see Table 1). However, in our DCM model, we chose to include only the pSTS, the TP, the AnG, the IOG, and the ITG, located in the right temporal and occipital lobe, which form a core ToM network. The other activated areas of the second-level analysis are mainly related to more general processes, like executive function and motor response (IFG, SMA, cerebellum) (Hugdahl, Raichle, Mitra, & Specht, 2015), or they are the left-hemisphere homologues to the areas selected for the DCM analysis.

As expected, activation was found in the pSTS. The pSTS is located in the posterior part of the STS. It should be noted that some researchers define or consider the pSTS as a

part of the TPJ (e.g., Patel, Sestieri, & Corbetta, 2019; Castelli, Happé, Frith, & Frith, 2000; Frith & Frith, 2003) or even as the AnG (Bzdok et al., 2016; Seghier, 2013). However, when it comes to mentalizing, the pSTS and the TPJ have been found to take dissociable roles (Gobbini, Koralek, Bryan, Montgomery, & Haxby, 2007).

The pSTS activates during the perception of biological motion and during the processing of socially relevant information (Allison, Puce, & McCarthy, 2000; Dasgupta, Tyler, & Grossman, 2011). The right pSTS, in particular, activates by goal-directed movement by animate, as well as inanimate, agents (Shultz, Lee, Pelphrey, & McCarthy, 2011; Shultz & McCarthy, 2012). Furthermore, the right pSTS is involved in detecting, encoding, predicting, and reasoning about social actions and the goals underlying them, which in turn is essential when it comes to making social interaction and communication effective (Gallagher & Frith, 2003; Frith & Frith, 2003; Kana et al., 2015; Rubio-Fernández & Geurts, 2013; Shultz, Lee, Pelphrey, & McCarthy, 2011).

Next, activation was found in the TP. The TP covers the most rostral part of the temporal lobe and is considered an extended part of the limbic system (Chabardes, Kahane, Minotti, Dominique, & Benabid, 2002; Olson, Plotzker, & Ezzyat, 2007). The TP is engaged in social and emotional processing (Olson, Plotzker, & Ezzyat, 2007). Further, in a study by Michel et al. (2013), they found that damage to the left TP did not affect a patient's ability to attribute mental states to others, measured by various nonverbal ToM tasks, as long as the right TP was intact. These results might suggest that the left TP is not crucial for ToM reasoning, at least not when presented with a nonverbal ToM task.

Further, activation was found in the AnG. The AnG is located in the posterior part of the inferior parietal lobule, at the junction between the occipital, temporal, and parietal lobes (Seghier, 2013). The AnG is involved in visuospatial attention (Studer, Cen, & Walsh, 2014), autobiographical and episodic memory (Berryhill, Phuong, Picasso, Cabeza, & Olson, 2007),

language (Nagaratnam, Phan, Barnett, & Ibrahim, 2002), and the recognition of visual patterns (Herath, Kinomura, & Roland, 2001), among other functions.

As previously mentioned, the AnG is often referred to as TPJ (Bzdok et al., 2016; Seghier, 2013). Further, it is also often referred to as the MTG or as the pSTS in the social cognition literature. In the language literature, the AnG is often referred to as the "pMTG" or the "temporal-parietal-occipital cortex" (Binder, Desai, Graves, & Conant, 2009; Bzdok et al., 2016; Seghier, 2013).

Next, activation was found in the IOG. The IOG is located in the occipital lobe, which forms the caudal part of the brain (Rehman & Khalili, 2020). The occipital lobe is predominantly involved in the processing of visual information (Rehman & Khalili, 2020; Rossion, Schiltz, & Crommelinck, 2003). Visual information will further be directed from the occipital lobe to more specialized regions for ToM reasoning (Castelli, Frith, Happé, & Frith, 2002).

Finally, activation was found in the ITG. The ITG is located on the lateral and inferior surfaces of the temporal lobe (Onitsuka et al., 2004). The ITG is involved in visual perception of objects and the recognition of visual patterns, among other functions (Herath, Kinomura, & Roland, 2001; Onitsuka et al., 2004).

Although several imaging studies have reported robust involvement of the mPFC across different ToM tasks (e.g., Frith & Frith, 2003; Mar, 2011; Schurz et al., 2014), our analyses revealed no activity in this region. Given that the mPFC is also involved in working memory (Smith et al., 2018), cognitive control (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004), and in processing socially and emotionally relevant information about others (Grossmann, 2013), one may question whether the mPFC has a more general role in social cognition. Thus, the mPFC might not be critically involved in ToM reasoning (Saxe & Powell, 2006; Otti, Wohlschlaeger, & Noll-Hussong, 2015).

This statement is supported by Bird, Castelli, Malik, Frith, and Husain (2004), who discovered that a patient who had suffered a stroke in the medial frontal region did not have any significant impairments across a wide range of ToM tasks. However, since the results from our second-level analysis only revealed which brain regions were more active during the "Mental" condition, we cannot exclude the possibility that the mPFC might have been active during both the "Mental" condition and the "Random" condition.

When comparing R-fMRI in TD adult participants from the ABIDE II to T-fMRI in TD adult participants from the HCP, we *(Hypothesis 2)* expected the EC in the ToM network not to be significantly different for most connections, given that both groups were TD control participants. Our analysis revealed differences in 16 of 25 connections at the corrected significance level of $p < .001$. As shown in Table 3, the main differences (i.e., the strongest effect size) between the groups were in the self-connections, in which the self-connections in the ABIDE II TD group had a higher mean rank value, and thus, were more inhibited, than the self-connections in the HCP TD group.

Although both groups were TD control participants, different paradigms were used. The HCP TD participants watched the Frith-Happé Animations inside the scanner (T-fMRI), while the ABIDE II TD participants lay inside the scanner in the absence of a concrete task or stimulus (R-fMRI). Thus, the differences observed are likely to be explained by the fact that the HCP TD participants performed a task. Further, one may question whether the regions of the ToM network are less inhibited and more sensitive to input from other regions in the network because the participants were watching the Frith-Happé Animations. However, given that we did not compare R-fMRI data from the HCP with R-fMRI data from the ABIDE II, this question remains to be answered.

When comparing R-fMRI in adult participants diagnosed with ASD from the ABIDE II to T-fMRI in TD adult participants from the HCP, we *(Hypothesis 3a)* expected to see

greater deviations in EC than when compared against *Hypothesis 2*. However, at the corrected significance level of $p < .001$, our analysis revealed significant differences in 12 of 25 connections in the ABIDE II ASD and the HCP TD comparison, against 16 of 25 connections in the ABIDE II TD and the HCP TD comparison. Thus, greater deviations in EC were not found.

Fewer significant different connections where found when comparing the ABIDE II ASD group to the HCP TD group than when comparing the ABIDE II TD group to the HCP TD group (see Figure 4 and Figure 5). These results might be explained by the fact that the group sizes were different, in which data from 24 individuals diagnosed with ASD (ABIDE II) and data from 50 TD individuals (HCP) were included. On this note, a small sample size reduces the study's power, and low statistical power will reduce the chance to detect a statistically significant effect (Button et al., 2013; Elsayir, 2018).

As in the ABIDE II TD and the HCP TD group comparison, the main differences (i.e., the strongest effect size) between the groups were in the self-connections, in which the self-connections in the ABIDE II ASD group had a higher mean rank value, and thus, were more inhibited than the self-connections in the HCP TD group (Table 4). These findings are likely to be explained by the fact that different paradigms were used, which put distinctive processing demands on the investigated network.

Previous R-fMRI studies have found significant differences in EC, as well as FC, between individuals diagnosed with ASD and TD controls during R-fMRI in the DMN (e.g., Cheng et al., 2015; Rolls et al., 2020). Thus, when comparing R-fMRI in TD adult participants from the ABIDE II to R-fMRI in adult participants diagnosed with ASD from the ABIDE II, we *(Hypothesis 3b)* expected to see similar deviations in EC as we expected to see in *Hypothesis 3a.* However, no significant differences were found.

Again, a possible explanation for not finding any significant differences is that data from only 24 individuals with ASD were included in the ABIDE dataset. When it comes to ASD, a bigger sample size is needed to generalize findings to larger populations of individuals with the disorder. In addition, a small sample size of individuals with ASD can be problematic due to the heterogeneous nature of the disorder (Dufour et al., 2013; Hayden, 2018).

Further, another possible explanation for not finding any significant differences is that the regions included in the present study were based on T-fMRI. In T-fMRI, brain activity at higher frequencies is measured, evoked by an explicit task or external stimulus (Glover, 2011). Consequently, greater fluctuations in the BOLD signal are detected and can be measured more reliably (Kristo et al., 2014). In R-fMRI, however, endogenous (spontaneous) brain activity at low frequencies (0.01-0.1 Hz) is measured while the individual lies inside the scanner in the absence of any sensory or cognitive stimulus (Kristo et al., 2014; Razi et al., 2017; Smitha et al., 2017). This can result in larger between-subject variability and less statistical power (Specht, 2020). Further, although some resting-state networks found through R-fMRI resemble those found through T-fMRI (Biswal, Yetkin, Haughton, & Hyde, 1995; Smith et al., 2009), differences in activity in the ToM network might be less visible during R-fMRI when compared to T-fMRI. Thus, one may question whether R-fMRI provides the same information as T-fMRI, and whether greater deviations would have been detected if both datasets were based on T-fMRI.

A third possible explanation for not finding any significant differences between the ABIDE II groups might be that, in DCM analysis of T-fMRI data, DCM is based on estimating the model evidence. Numerous models have to be estimated, and their likelihood have to be compared to find the best model (Specht, Baumgartner, Stadler, Hugdahl, & Pollmann, 2014; Stephan et al., 2010). Thus, there is a possibility that we did not select the

best DCM model to represent the ToM network in the HCP data. Although this could explain some of the general differences between the HCP group and the ABIDE II groups, it does not explain why there are no significant differences between the two ABIDE II groups. Note, DCM for R-fMRI data is not based on comparing several models, but uses Bayesian methods, like PEB and Bayesian model reduction, to infer on the most likely model space starting from a fully connected model.

Further, research suggests that human intellectual performance is related to how efficiently the brain integrates information between regions (van den Heuvel, Stam, Kahn, & Hulshoff Pol, 2009). Also, lower IQ is associated with decreased connectivity within the DMN (Gabrielsen et al., 2018). Thus, one may question whether a significant difference in EC would have been detected between the ABIDE II groups if the ABIDE II ASD participants were less functioning. In the present study, IQ range for the ABIDE II TD participants were 85-146 (M = 111, SD = 13), while IQ range for the ABIDE II ASD participants were 80-146 (M = 114, SD = 16) (Supplementary Table 3 in Di Martino et al., 2017).

Finally, individuals were excluded from the ABIDEII-ONRC_2 if they three months prior to the evaluation had met the criteria for Posttraumatic Stress Disorder, Manic or Depressive episode, Bipolar Disorder, and/or Schizophrenia (ABIDE, 2017). Also, the ASD participants had relatively low comorbidity (Di Martino et al., 2017). Because many individuals with ASD have additional problems and because symptom severity correlates with lower brain activity (Happé, Booth, Charlton, & Hughes, 2006; Libero & Kana, 2013; Mannion, Leader, & Healy, 2013), the individuals in this dataset might not be representable for the ASD population in general.

Limitations and Future Research

One limitation of the present study is related to the small group size. Although the ABIDE II includes 1114 datasets from 521 individuals with ASD and 593 TD controls, the individual datasets are relatively small. Although we chose the dataset with one of the largest samples of adult individuals with ASD, the sample size could have been too small to detect significant differences and to address the challenges imposed by the heterogeneity of ASD.

Further, the criteria for participating in the HCP study were rather strict, in which none of the participants had any documented history of psychiatric, neurologic, neuropsychiatric, or neurodevelopmental disorders. Even individuals with diabetes and high blood pressure were excluded (Van Essen et al., 2013). Thus, the variation within the HCP group was relatively small, and the group might not have been fully representative of the general population.

The criteria for participating in the ABIDEII-ONRC_2 was also relatively strict, making us question the representativeness of the ASD individuals. Further, data from only four female individuals diagnosed with ASD were included in the ABIDEII-ONRC_2 dataset. Although more males than females are diagnosed with the disorder (Green, Travers, Howe, & McDougle, 2019), an equal gender distribution is needed to generalize the findings and to conclude that the findings are also representative of females with ASD (Milner, McIntosh, Colvert, & Happé, 2019).

Limitations can also be discussed regarding the data acquisition, the preprocessing methods, and the datasets used. First, the data were acquired with different scanners. Second, different acquisition parameters were used by the HCP and the ABIDE, in which shorter TR time was used in the ABIDE II (475 ms) compared to HCP (720 ms). Also, different parameters for the FA, the TE, and matrix size could result in different types of artifacts and different sensitivities to the BOLD signal (Sunaert, 2016). Third, data from the ABIDEII-

ONRC_2 were not preprocessed in advance. Although we tried to make the data as similar as possible to the HCP, the preprocessing methods might have been slightly different.

On a final note, Di Martino et al. (2017) argue that researchers should ensure a balanced design by selecting participants from the same collection. Although this argument is mainly directed to those using datasets from the ABIDE I and the ABIDE II, the same rule could apply for those using datasets collected from different data repositories, such as the ABIDE II and the HCP.

Conclusion

Many fMRI studies have examined FC in individuals diagnosed with ASD compared to TD controls during ToM task performance and during rest. However, relatively few studies have examined EC in individuals diagnosed with ASD compared to TD controls (Deshpande et al., 2013; Kana et al., 2014). Thus, this study is a contribution to understanding the differences in EC between brain regions involved in ToM processing during T-fMRI and R-fMRI in TD adults and adults diagnosed with ASD.

Results from the analysis of T-fMRI data obtained from the HCP revealed activation in some of the same brain regions as we expected to see *(Hypothesis 1)*. In our DCM model, we chose to include the pSTS, the TP, the AnG, the IOG, and the ITG, located in the right temporal and occipital lobe.

Further, contrary to what was expected *(Hypothesis 2)*, our analysis revealed several significant differences in the connections of the ToM network when comparing R-fMRI in TD adult participants from the ABIDE II to T-fMRI in TD adult participants from the HCP. We argue that these results could be explained by the fact that different paradigms were used, and that the regions of the ToM network are less inhibited and more sensitive to input from other regions in the network when performing a task.

Further, when comparing R-fMRI in adult participants diagnosed with ASD from the ABIDE II to T-fMRI in TD adult participants from the HCP, we *(Hypothesis 3a)* expected to see greater deviations in EC than when compared against *Hypothesis 2*. However, greater deviations were not found. We argue that these findings are likely to be explained by the fact that different paradigms were used. Further, the fewer significant different connections that were found when comparing the ABIDE II ASD group to the HCP TD group than when comparing the ABIDE II TD group to the HCP TD group, might be explained by the fact that the group sizes were different.

Finally, when comparing R-fMRI in TD participants from the ABIDE II and R-fMRI in ASD participants from the ABIDE II, we *(Hypothesis 3b)* expected to see similar deviations as we expected to see in *Hypothesis 3a*. However, no significant differences were found. Thus, we question whether greater deviations would have been detected if the sample size were bigger in the ASD group, if both datasets were based on T-fMRI, whether we actually selected the best DCM model to represent the ToM network in the HCP data, and whether significant difference would have been detected if the ABIDE II ASD participants were less functioning and perhaps more representative of the ASD population.

Overall, our results highlight that findings from neuroimaging studies can be inconsistent and contradictory. Additionally, our results highlight the heterogeneity in the ASD population. Further, while results from neuroimaging studies often emphasize the differences in both EC and FC between brain regions involved in ToM processing in individuals with ASD and TD individuals, the results from this study highlight the similarities. Although we did not find what we expected, understanding the similarities between individuals diagnosed with ASD and TD individuals can be of equal importance to understand the organization and the functioning in the brain of individuals diagnosed with ASD.

References

Akhtar, N., & Gernsbacher, M. A. (2007). Joint Attention and Vocabulary Development: A Critical Look. *Language and Linguistics Compass*, *1*(3), 195–207. doi:10.1111/j.1749-818X.2007.00014.x

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4(*7), 267-278. doi:10.1016/S1364-6613(00)01501-1

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: Author

Arslan, B., Taatgen, N. A., & Verbrugge, R. (2017). Five-year-olds' systematic errors in second-order false belief tasks are due to first-order theory of mind strategy selection: A computational modeling study. *Frontiers in psychology, 8*(275). doi:10.3389/fpsyg.2017.00275

Autism Brain Imaging Exchange. (2007, March 27). Welcome to the Autism Brain Imaging Data Exchange. Retrieved from http://fcon_1000.projects.nitrc.org/indi/abide/

Autism Brain Imaging Exchange. (2016, June 24). Autism Brain Imaging Data Exchange I. Retrieved from http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html

Autism Brain Imaging Exchange. (2017, March 27). Autism Brain Imaging Data Exchange II. Retrieved from http://fcon_1000.projects.nitrc.org/indi/abide/abide_II.html

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110-118. doi:10.1016/j.tics.2009.12.006

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M.,

… WU-Minn HCP Consortium. (2013). Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage, 80,* 169-189. doi:10.1016/j.neuroimage.2013.05.033

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46. doi:10.1016/0010-0277(85)90022-8

Bauminger-Zviely, N. (2013). False-Belief Task. In F. R, Volkmar (Eds.), Encyclopedia of Autism Spectrum Disorders (2013). New York: Springer. doi:10.1007/978-1-4419-1698-3_91

Berryhill, M. E., Phuong, L., Picasso, L., Cabeza, R., & Olson, I. R. (2007). Parietal lobe and episodic memory: Bilateral damage causes impaired free recall of autobiographical memory. *Journal of Neuroscience, 27*(52), 14415-14423. doi:10.1523/JNEUROSCI.4163-07.2007

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-Analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*(12), 2767-2796. doi:10.1093/cercor/bhp055

Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on "theory of mind" and cognition. *Brain, 127*, 914-928. doi:10.1093/brain/awh108

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine, 34*(4), 537-541. doi:10.1002/mrm.1910340409

Brewer, N., Young, R. L., & Barnett, E. (2017). Measuring theory of mind in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*(7), 1927-1941. doi:10.1007/s10803-017-3080-x

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network:

anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124*(1), 1-38. doi:10.1196/annals.1440.011

Button, K. S., Ioannidis, J., Mokrysz, C., Nosek, B., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365-376. doi:10.1038/nrn3475

Byom, L. J., & Mutlu, B. (2013). Theory of mind: mechanisms, methods, and new directions. *Frontiers in Human Neuroscience*, *7*, 413. doi:10.3389/fnhum.2013.00413

Bzdok, D., Hartwigsen, G., Reid, A., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2016). Left inferior parietal lobe engagement in social cognition and language. *Neuroscience and Biobehavioral Reviews, 68*, 319-334. doi:10.1016/j.neubiorev.2016.02.024

Campbell, S. B., Leezenbaum, N. B., Mahoney, A. S., Moore, E. L., & Brownell, C. A. (2016). Pretend play and social engagement in toddlers at high and low genetic risk for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(7), 2305-2316. doi:10.1007/s10803-016-2764-y

Carrington, S. J., & Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping, 30*(8), 2313-2335. doi:10.1002/hbm.20671

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes, *Brain, 125*(8), 1839-1849. doi:10.1093/brain/awf189

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and Mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage, 12*(3), 314-325. doi:10.1006/nimg.2000.0612.

Chabardes, S., Kahane, P., Minotti, L., Dominique, H., & Benabid, A. L. (2002). Anatomy of

the temporal pole region. *Epileptic disorders: international epilepsy journal with videotape, 4*(1), 9-15. Retrieved from https://pubmed.ncbi.nlm.nih.gov/12424085/

Cheng, W., Rolls, E. T., Gu, H., Zhang, J., & Feng, J. (2015). Autism: reduced connectivity between cortical areas involved in face expression, theory of mind, and the sense of self. *Brain, 138*, 1382-1393. doi:10.1093/brain/awv051

Cherkassky, V. L., Kana, R. K., Keller, T. A., & Just, M. A. (2006). Functional connectivity in a baseline resting-state network in autism. *Neuroreport, 17*(16), 1687-1690. doi:10.1097/01.wnr.0000239956.45448.4c

Cohen, J. (1988). The Analysis of Variance. *In Statistical Power Analysis for the Behavioral Sciences (2nd ed.)* [Internet]. Hillsdale (NY): Lawrence Erlbaum Associates. Retrieved from http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf

Dasgupta, S., Tyler, S. C., & Grossman, E. D. (2011). Co-localization of human posterior STS during biological motion, face and social perception. *Journal of Vision, 11*(11), 629. doi:10.1167/11.11.629

Dawson, G., Jones, E. J., Merkle, K., Venema, K., Lowy, R., Faja, S., … Webb, S. J. (2012). Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(11), 1150-1159. doi:10.1016/j.jaac.2012.08.018

Deshpande, G., Libero, L. E., Sreenivasan, K. R., Deshpande, H. D., & Kana, R. K. (2013). Identification of neural connectivity signatures of autism using machine learning. *Frontiers in Human Neuroscience, 7*(670). doi:10.3389/fnhum.2013.00670

Dufour, N., Redcay, E., Young, L., Mavros, P. L., Moran, J. M., Triantafyllou, C, … Saxe, R. (2013). Similar brain activation during false belief tasks in a large sample of adults with and without autism. *PLoS ONE, 8*(9). doi:10.1371/journal.pone.0075468

Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., … Milham, M. P. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry, 19*(6), 659-667. doi:10.1038/mp.2013.78

Di Martino, A., O´Connor, D., Chen, B., Alaerts, K., Anderson, J. S., Assaf, M., … Milham, M. P. (2017). Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Scientific Data, 4*(170010). doi:10.1038/sdata.2017.10

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage, 59*(3), 2349-2361. doi:10.1016/j.neuroimage.2011.09.017

Elam, J. (June 09, 2014). HCP announces release of data from 500 subjects. Retrieved from https://www.humanconnectome.org/study/hcp-young-adult/article/hcp-announces-release-data-500-subjects

Elsayir, H. (2018). Factors Determining the Power of a Statistical Test for the Difference between Means and Proportions. *American Journal of Mathematics and Statistics, 8*(6), 171-178. doi:10.5923/j.ajms.20180806.02

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. *Cognition, 57*(2), 109-128. doi:10.1016/0010-0277(95)00692-r

Fox, M., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in Systems Neuroscience, 4*(19). doi:10.3389/fnsys.2010.00019

Friston K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology, 7*(2), doi:10.1371/journal.pbio.1000033

Friston, K., & Penny, W. (2011). Post hoc Bayesian model selection. *NeuroImage*, *56*(4),

2089-2099. doi:10.1016/j.neuroimage.2011.03.062

Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind". *Cognition, 50*(1-3), 115-132. doi:10.1016/0010-0277(94)90024-8

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 358*(1431), 459-473. doi:10.1098/rstb.2002.1218

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology, 141*(1), 2-18. doi:10.1037/a0024338

Gabrielsen, T. P., Anderson, J. S., Stephenson, K. G., Beck, J., King, J. B., … South, M. (2018). Functional MRI connectivity of children with autism and low verbal and cognitive performance. *Molecular Autism, 9*(67). doi:10.1186/s13229-018-0248-y

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences, 7*(2), 77-82. doi:10.1016/S1364-6613(02)00025-6

Girli, A., & Tekin, D. (2010). Investigating false belief levels of typically developed children and children with autism. *Procedia - Social and Behavioral Sciences, 2*(2), 1944-1950. doi:10.1016/j.sbspro.2010.03.261

Glasser, M. F, Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., … WU-Minn HCP Consortium. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage, 80,* 105-124. doi:10.1016/j.neuroimage.2013.04.127

Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery Clinics of North America*, *22*(2), 133-139. doi:10.1016/j.nec.2010.11.001

Gobbini, M. I., Koralek, A., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two

Takes on the Social Brain: A Comparison of Theory of Mind Tasks. *Journal of Cognitive Neuroscience, 19*(11), 1803-1814. doi:10.1162/jocn.2007.19.11.1803

Gotts, S. J., Ramot, M., Jasmin, K., & Martin, A. (2019). Altered resting-state dynamics in autism spectrum disorder: Causal to the social impairment? *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 2*(90), 28-36. doi:10.1016/j.pnpbp.2018.11.002

Green, R. M., Travers, A. M., Howe, Y., & McDougle, C. J. (2019). Women and autism spectrum disorder: Diagnosis and implications for treatment of adolescents and adults. *Current Psychiatry Reports, 21*(22). doi:10.1007/s11920-019-1006-3

Grossmann, T. (2013). The role of medial prefrontal cortex in early social cognition. *Frontiers in Human Neuroscience, 7*, 340. doi:10.3389/fnhum.2013.00340

Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., … Szatmari, P. (2015). Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority. *Molecular Autism*, *6*(36). doi:10.1186/s13229-015-0019-y

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition, 61*(1), 25-39. doi:10.1016/j.bandc.2006.03.004

Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., … Frith, C. (1996). “Theory of mind” in the brain. Evidence from a PET scan study of Asperger syndrome. *NeuroReport, 8*(1), 197-201. doi:10.1097/00001756-199612200-00040

Hassan, M. M., & Mokhtar, H. M. O. (2019). Investigating autism etiology and heterogeneity by decision tree algorithm. *Informatics in Medicine Unlocked, 16*, 100215. doi:10.1016/j.imu.2019.100215

Hayden, E., C. (2018, October 17). For studies, size matters: Let us count the ways. Retrieved from https://www.spectrumnews.org/news/studies-size-matters-let-us-count-ways/

Herath, P., Kinomura, S., & Roland, P. E. (2001). Visual recognition: Evidence for two distinctive mechanisms from a PET study. *Human Brain Mapping, 12*(2), 110-119. doi:10.1002/1097-0193(200102)12:2<110::AID-HBM1008>3.0.CO;2-0

Hillebrandt, H., Dumontheil, I., Blakemore, S. J., & Roiser, J. P. (2013). Dynamic causal modelling of effective connectivity during perspective taking in a communicative task. *NeuroImage, 76*, 116-124. doi:10.1016/j.neuroimage.2013.02.072

Hillebrandt, H., Friston, K. J., & Blakemore, S. J. (2014). Effective connectivity during animacy perception - dynamic causal modelling of Human Connectome Project data. *Scientific Reports, 4*(6240), 1-9. doi:10.1038/srep06240

Hugdahl, K., Raichle, M. E., Mitra, A., & Specht, K. (2015). On the existence of a generalized non-specific task-dependent network. *Frontiers in Human Neuroscience, 9*, 430. doi:10.3389/fnhum.2015.00430

Hull, J. V., Dokovna, L. B., Jacokes, Z. J., Torgerson, C. M., Irimia, A., & Van Horn, J. D. (2017). Resting-State functional connectivity in autism spectrum disorders: A review. *Frontiers in Psychiatry, 7*(205). doi:10.3389/fpsyt.2016.00205

Human Connectome Project. (n.d.). HCP Young Adult. Retrieved from https://www.humanconnectome.org/study/hcp-young-adult

Human Connectome Project (2017, November 30). WU-Minn HCP Consortium Restricted Data Use Terms. Retrieved from https://www.humanconnectome.org/study/hcp-young-adult/document/wu-minn-hcp-consortium-restricted-data-use-terms

International Business Machines. (n.d.). BOX Subcommand (IGRAPH command).

Retrieved from https://www.ibm.com/support/knowledgecenter/en/SSLVMB_24.0.0/spss/base/syn_igraph_box.html

Kahan, J., & Foltynie, T. (2013). Understanding DCM: Ten simple rules for the clinical. *NeuroImage, 83*, 542-549. doi:10.1016/j.neuroimage.2013.07.008

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2009). Atypical frontal-posterior synchronization of theory of mind regions in autism during mental state attribution. *Social Neuroscience, 4*(2), 135-152. doi:10.1080/17470910802198510

Kana, R. K., Libero, L. E., & Moore, M. S. (2011). Disrupted cortical connectivity theory as an explanatory model for autism spectrum disorders. *Physics of Life Reviews, 8*(4), 410-437. doi:10.1016/j.plrev.2011.10.001

Kana, R. K., Maximo, J. O., Williams, D. L., Keller, T. A., Schipul, S. E., Cherkassy, V. L., … Just, M. A. (2015). Aberrant functioning of the theory-of-mind network in children and adolescents with autism. *Molecular Autism, 6*(59). doi:10.1186/s13229-015-0052-x

Kana, R. K., Uddin, L. Q., Kenet, T., Chugani, D., & Müller, R. A. (2014). Brain connectivity in autism. *Frontiers in Human Neuroscience, 8*(349), doi:10.3389/fnhum.2014.00349

Kristo, G., Rutten, G. J., Raemaekers, M., de Gelder, B., Rombouts, S. A. B., & Ramsey, N. F. (2014). Task and task-free FMRI reproducibility comparison for motor network identification. *Human Brain Mapping, 35*(1), 340-352. doi.org/10.1002/hbm.22180

Lai M. C., Lombardo, M. V., & Baron-Cohen, S. (2014). Autism. *Lancet, 383*(9920), 896-910. doi:10.1016/S0140-6736(13)61539-1

Landa, R., & Garrett-Mayer, E. (2006). Development in infants with autism spectrum

disorders: a prospective study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(6), 629-638. doi:10.1111/j.1469-7610.2006.01531.x

Leslie, A. (1987). Pretense and Representation: The origins of "theory of mind". *Psychological Review, 94*(4), 412-426. doi:10.1037//0033-295X.94.4.412

Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). Review of methods for functional brain connectivity detection using fMRI. *Computerized Medical Imaging and Graphics, 33*(2), 131-139. doi:10.1016/j.compmedimag.2008.10.011

Libero, L. E., & Kana R. K. (2013). Advancing our understanding of the brain in autism: contribution of functional MRI and diffusion tensor imaging. *Imaging in Medicine, 5*(5), 453-465. doi:10.2217/iim.13.46

Liu, T. T., Nalci, A., & Falahpour, M. (2017). The global signal in fMRI: Nuisance or Information? *NeuroImage*, *150*, 213-229. doi:10.1016/j.neuroimage.2017.02.036

Mahy, C. E. V., Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental Cognitive Neuroscience, 9*, 68-81. doi:10.1016/j.dcn.2014.01.002

Mannion, A., Leader, G., & Healy, O. (2013). An investigation of comorbid psychological disorders, sleep problems, gastrointestinal symptoms and epilepsy in children and adolescents with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders, 7*(1), 35-42. doi:10.1016/j.rasd.2012.05.002

Mar, R. A. (2011). The Neural Bases of Social Cognition and Story Comprehension. *Annual Review of Psychology, 62*, 103-34. doi:10.1146/annurev-psych-120709-145406

Michel, C., Dricot, L., Lhommel, R., Grandin, C., Ivanoiu, A., Pillon, A., & Samson, D. (2013). Extensive left temporal pole damage does not impact on theory of mind abilities. *Journal of Cognitive Neuroscience, 25*(12), 2025-2046. doi:10.1162/jocn_a_00488

Milner, V., McIntosh, H., Colvert, E., & Happé, F. (2019). A qualitative exploration of the female experience of autism spectrum disorder (ASD). Journal of autism and developmental disorders, 49(6), 2389-2402. doi.org/10.1007/s10803-019-03906-4

Mizuno, A., Villalobos, M. E., Davies, M. M., Dahl, B. C., & Müller, R. A. (2006). Partially enhanced thalamocortical functional connectivity in autism. *Brain Research, 1104*(1), 160-174. doi.10.1016/j.brainres.2006.05.064

Mohammad-Rezazadeh, I., Frohlich, J., Loo, S. K., & Jeste, S. S. (2016). Brain connectivity in autism spectrum disorder. *Current Opinion in Neurology, 29*(2), 137-147. doi:10.1097/WCO.0000000000000301

Moessnang, C., Baumeister, S., Tillmann, J., Goyard, D., Charman, T., Ambrosino, S., … Lythgoe, D. J. (2020). Social brain activation during mentalizing in a large autism cohort: The longitudinal European autism project. *Molecular Autism*, *11*(17). doi:10.1186/s13229-020-0317-x

Müller, R. A., & Fishman, I. (2018). Brain connectivity and neuroimaging of social networks in autism. *Trends in Cognitive Sciences, 22*(12), 1103-1116. doi:10.1016/j.tics.2018.09.008

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science, 16*(5), 269-274. doi:10.1111/j.1467-8721.2007.00518.x

Nagaratnam, N., Phan, T. A., Barnett, C., & Ibrahim, N. (2002). Angular gyrus syndrome mimicking depressive pseudodementia. *Journal of Psychiatry and Neuroscience, 27*(5), 364-368. Retrieved from https://pubmed.ncbi.nlm.nih.gov/12271792/

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing, *Brain*, *130*(7), 1718-1731. doi:10.1093/brain/awm052

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258. doi:10.1126/science.1107621

Onitsuka, T., Shenton, M. E., Salisbury, D. F., Dickey, C. C., Kasai, K., Toner, S. K., … McCarley, R. W. (2004). Middle and inferior temporal gyrus gray matter volume abnormalities in chronic schizophrenia: an MRI study. *The American journal of psychiatry*, *161*(9), 1603-1611. doi:10.1176/appi.ajp.161.9.1603

Otti, A., Wohlschlaeger, A. M., & Noll-Hussong, M. (2015). Is the medial prefrontal cortex necessary for theory of mind? *PLoS ONE, 10*(8). doi:10.1371/journal.pone.0135912

Padmanabhan, A., Lynch, C. J., Schaer, M., & Menon, V. (2017). The default mode network in autism. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2*(6), 476-486. doi:10.1016/j.bpsc.2017.04.004

Patel, G. H., Sestieri, C., & Corbetta, M. (2019). The evolution of the temporoparietal junction and posterior superior temporal sulcus. *Cortex, 118*, 38-50. doi: 10.1016/j.cortex.2019.01.026

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437-471. doi:10.1016/0022-0965(85)90051-7

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman. G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America, 98*(2) 676-682. doi:10.1073/pnas.98.2.676

Razi, A., Kahan, J., Rees, G., & Friston, K. J. (2014). Construct validation of a DCM for resting state fMRI. *NeuroImage, 106*, 1-14. doi:10.1016/j.neuroimage.2014.11.027

Razi, A., Seghier, M. L., Zhou, Y, McColgan, P., Zeidman, P., Park, H.-J., … Friston, K. J.

(2017). Large-scale DCMs for resting-state fMRI. *Network Neuroscience, 1*(3), 222-241. doi:10.1162/netn_a_00015

Rehman, A., & Khalili, Y. A. (2020). Neuroanathomy, Occipital Lobe. *In StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK544320/

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial prefrontal cortex in cognitive control. *Science, 306*(5695), 443-447. doi:10.1126/science.1100301

Rolls, E. T., Zhou, Y., Cheng, W., Gilson, M., Deco, G., & Feng, J. (2020). Effective connectivity in autism. *Autism Research, 13*(4), 32-44. doi:10.1002/aur.2235

Rossion, B., Schiltz, C., & Crommelinck, M. (2003). The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces. *NeuroImage, 19*(3), 877-883. doi:10.1016/S1053-8119(03)00105-8

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*(1), 27–33. doi:10.1177/0956797612447819

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science, 17*(8), 692-699. doi:10.1111/j.1467-9280.2006.01768.x

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences, 19*(2), 65-72. doi:10.1016/j.tics.2014.11.007

Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscientist, 19*(1), 43-61. doi:10.1177/1073858412440596

Shukla, D. K., Keehn, B., & Müller, R. A. (2010). Regional homogeneity of fMRI time series

in autism spectrum disorders. *Neuroscience Letters*, *476*(1), 46–51. doi:10.1016/j.neulet.2010.03.080

Shultz, S., Lee, S. M., Pelphrey, K., & McCarthy, G. (2011). The posterior superior temporal sulcus is sensitive to the outcome of human and non-human goal-directed actions. *Social Cognitive and Affective Neuroscience*, *6*(5), 602-611. doi:10.1093/scan/nsq087

Shultz, S., & McCarthy, G. (2012). Goal-directed actions activate the face-sensitive posterior superior temporal sulcus and fusiform gyrus in the absence of human-like perceptual cues. *Cerebral Cortex, 22*(5), 1098-1106. doi:10.1093/cercor/bhr180

Schurz, M., & Perner, J. (2015). An evaluation of neurocognitive models of theory of mind. *Frontiers in Psychology, 6,* 1610. doi:10.3389/fpsyg.2015.01610

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9-34. doi:10.1016/j.neubiorev.2014.01.009

Sidera, F., Perpiñà, G., Serrano, J., & Rostan, C. (2018). Why is theory of mind important for referential communication? *Current Psychology (New Brunswick, N.J.), 37*(1), 82-97. doi:10.1007/s12144-016-9492-5

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., … Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences, 106*(31), 13040-13045. doi:10.1073/pnas.0905267106

Smith, R., Lane, R. D., Alkozei, A., Bao, J., Smith, C., Sanova, A., … Killgore, W. D. S. (2018). The role of medial prefrontal cortex in the working memory maintenance of one’s own emotional responses. *Scientific Reports, 8*(1), 3460. doi:10.1038/s41598-018-21896-8

Smitha, K. A., Raja, K. A., Arun, K. M., Rajesh, P. G., Thomas, B., Kapilamoorthy, T. R., & Kesavadas, C. (2017). Resting state fMRI: A review on methods in resting state connectivity analysis and resting state networks. *The neuroradiology journal*, *30*(4), 305-317. doi:10.1177/1971400917697342

Specht, K. (2020). Current challenges in translational and clinical fMRI and future directions. *Frontiers in Psychiatry, 10*, 924. doi:10.3389/fpsyt.2019.00924

Specht, K., Baumgartner, F. J., Stadler, J., Hugdahl, K., & Pollmann, S. (2014). Functional asymmetry and effective connectivity of the auditory system during speech perception is modulated by the place of articulation of the consonant- A 7T fMRI study. *Frontiers in Psychologys, 5*(549). doi:10.3389/fpsyg.2014.00549.

Stephan, K. E., Harrison, L. M., Kiebel, S. J., David, O., Penny, W. D., & Friston, K. J. (2007). Dynamic causal models of neural system dynamics: Current state and future extensions. *Journal of Biosciences, 32*(1), 129-144. doi:10.1007/s12038-007-0012-5

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage, 49*(4), 3099-3109. doi:10.1016/j.neuroimage.2009.11.015

Studer, B., Cen, D., & Walsh, V. (2014). The angular gyrus and visuospatial attention in decision-making under risk. *NeuroImage, 103*, 75-80. doi:10.1016/j.neuroimage.2014.09.003

Sunaert, S. (October 12, 2016). BOLD fMRI and DTI: Artifacts, tips, and tricks. Retrieved from https://radiologykey.com/bold-fmri-and-dti-artifacts-tips-and-tricks/

Uğurbil, K., Xu, J., Auerbach, E. J., Moeller, S., Vu, A., Duarte-Carvajalino, J. M., … WU-Minn HCP Consortium. (2013). Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage, 80*, 80-104. doi:10.1016/j.neuroimage.2013.05.012

van den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(23), 7619-7624. doi:10.1523/JNEUROSCI.1443-09.2009

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., … WU-Minn HCP Consortium. (2012). The Human Connectome Project: a data acquisition perspective. *NeuroImage*, *62*(4), 2222–2231. doi:10.1016/j.neuroimage.2012.02.018

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Urgurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage, 80*, 62-79. doi:10.1016/j.neuroimage.2013.05.041

Vissers, M. E., Cohen, M. X., & Geurts, H. M. (2011). Brain connectivity and high functioning autism: A promising path of research that needs refined models, methodological convergence, and stronger behavioral links. *Neuroscience and Biobehavioral Reviews, 36*(1), 604-625. doi:10.1016/j.neubiorev.2011.09.003

Werling, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology, 26*(2), 146-53. doi:10.1097/WCO.0b013e32835ee548

White, S. J, Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé animations: a quick and objective test of theory of mind for adults with autism. *Autism Research, 4*(2), 149-154. doi:10.1002/aur.174

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: Distinct roles for the social network and mirror system. *Psychological Science, 18*(6), 469-474. doi:10.1111/j.1467-9280.2007.01923.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining

function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128. doi:10.1016/0010-0277(83)90004-5

Zalla, T., & Korman, J. (2018). Prior knowledge, episodic control and theory of mind in autism: Toward an integrative account of social cognition. *Frontiers in psychology, 9*, 752. doi:10.3389/fpsyg.2018.00752

Zeidman, P., Jafarian, A., Corbin, N., Seghier, M. L., Razi, A, Price, C. J., & Friston, K. J. (2019). A guide to group effective connectivity analysis, part 1: First level analysis with DCM for fMRI. *NeuroImage, 200*, 174-190. doi:10.1016/j.neuroimage.2019.06.031