ABSTRACT

Title of Document:	THE EFFECT OF SOCIAL INTERACTION ON THE NEURAL CORRELATES OF LANGUAGE PROCESSING AND MENTALIZING
	Katherine Ann Rice, Master of Science, 2014
Directed By:	Assistant Professor Elizabeth Redcay Department of Psychology

Recent behavioral and neuroscience evidence suggests that studying the social brain in detached and offline contexts (e.g., listening to prerecorded stories about characters) may not capture real-world social processes. Few studies, however, have directly compared neural activation during live interaction to conventional recorded paradigms. The current study used a novel fMRI paradigm to investigate whether real-time social interaction modulates the neural correlates of language processing and mentalizing. Regions associated with social engagement (i.e., dorsal medial prefrontal cortex) were more active during live interaction. Processing live versus recorded language increased activation in regions associated with narrative processing and mentalizing (i.e., temporal parietal junction). Regions associated with intentionality understanding (i.e., posterior superior temporal sulcus) were more active when mentalizing about a live partner. These results have implications for quantifying and understanding the neural correlates of real-world

social behavior in typical adults, in developmental populations, and in individuals with social disabilities such as autism.

THE EFFECT OF SOCIAL INTERACTION ON THE NEURAL CORRELATES OF LANGUAGE PROCESSING AND MENTALIZING

By

Katherine Ann Rice

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2014

Advisory Committee:
Professor Elizabeth Redcay, Chair
Professor Jonathan Beier
Professor Jude Cassidy

Acknowledgements

Thank you to all of the members of the Developmental Social Cognitive Neuroscience lab for their assistance, specifically Laura Anderson, Viviana Bauman, Seleste Braddock, Tara Feld, Toby Hamovitz, Daniel O'Young, Kayla Velnoskey, & Brieana Viscomi. Thank you also to the Maryland Neuroimaging Center and staff for their assistance. And many, many thanks to Dr. Elizabeth Redcay for her invaluable guidance throughout the project.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Humans live in a world filled with reciprocal social interaction, but social neuroscientists have almost exclusively studied the brain in offline, non-interactive contexts. Paradigms in which participants view static photographs of strangers or listen to recorded stories about a character's false beliefs have yielded insight into how the brain processes social stimuli, but such paradigms may misrepresent the real-time workings of the human mind (Gallagher, 2012; Gallotti & Frith, 2013; Schilbach et al., 2013). Recently, Schilbach and colleagues (2013) urged social neuroscientists to take a second-person perspective. As compared to a third-person perspective, in which a participant observes social stimuli, second-person experiments examine the brain when a participant is actively involved in real-time, contingent interaction with another person. This approach is predicated on the theory that the brain operates differently during an interaction than during traditional offline paradigms. Direct tests of this second-person hypothesis, however, are limited. Careful and well-controlled comparisons of brain activity during interaction versus observation will vet the second-person hypothesis and will improve understanding of how the brain processes real-time social interaction, which has implications for understanding typical development and social disabilities.

There is preliminary evidence that live interaction engages social brain regions differentially than recordings of social stimuli (Redcay et al., 2010). This study compared participant neural activity when playing a live game over video feed with an experimenter versus watching a recorded video of the experimenter from a previous game. Several different components of live interaction, however, were conflated during the live segment, including eye contact, spoken language, turn taking, and contingent responding. These components often overlapped and the study was unable to disentangle which specific factors were responsible for differential brain activation in the live condition. Thus, although this study suggests that live interaction affects neural responses, it was unable to pinpoint the specific influence of live interaction on particular cognitive processes, such as processing live speech outside of the context of also seeing the live speaker.

Two such social cognitive processes of special interest to second-person neuroscience are language processing and mentalizing, which is also known as theory of mind, and refers to thinking about the thoughts, goals, and desires of others. Rich behavioral evidence from both language processing and mentalizing research indicates that these processes operate differentially in the presence of live interaction; in contrast, the neuroscience investigation of these processes almost exclusively employs third-person approaches. No study has yet determined how, or even if, the brain processes language or mental states differently in a real-time, contingent social interaction as compared to listening to similar, but recorded stimuli.

This Master's thesis uses a novel functional magnetic resonance imaging (fMRI) paradigm to directly examine the hypothesis that live interaction modulates neural activation to social stimuli. I compare brain activation for two events: first, listening to a social partner in real-time versus listening to recorded speech, and, second, mentalizing about this social partner versus mentalizing about characters in recorded stories. This new paradigm provides three main contributions to the social neuroscience literature: first, it resolves issues of internal validity present in past studies of live interaction (e.g., stimuli characteristics, attention, motivation; see Redcay, Rice, & Saxe, 2013 for a review);

second, it extends the second-person perspective to two novel social cognitive domains (i.e., language processing and mentalizing); and, third, this paradigm not only examines live interaction, but directly compares this live interaction to a traditional third-person condition, addressing the question of whether live interaction in and of itself changes the neural processes identified by conventional social neuroscience.

In addition to these contributions, the ultimate goal of this study is to extend a second-person neuroscience perspective to developmental populations. Developmental research examining the relationship between social cognition and the brain—developmental social cognitive neuroscience—has yielded insights into the mechanisms behind behavior (Diamond & Amso, 2008). Findings from developmental neuroimaging have influenced education (e.g., Dubinsky, 2010), social policy (e.g., Johnson, Blum & Giedd, 2009), and pediatrics (e.g., Diamond, 2001). In spite of these advances, limitations remain in translational neuroscience (e.g., Giedd & Rapoport, 2010), especially in identifying the brain bases of social developmental disorders such as autism. One potential reason for this limitation is that the offline, third-person approaches used to investigate social cognition do not reflect the contingent and reciprocal interactions that compose a child's real-world social life. Given the paucity of current research on second person neuroscience, the current study with adults is a necessary preliminary step to developmental work for two reasons: first, to resolve practical concerns with this novel paradigm, and second, to establish developmental endpoints in order to contextualize eventual studies with children.

In this thesis, I first introduce the relevant behavioral and neuroscience literature on the effect of live interaction on social cognition. I focus on language processing and mentalizing, and discuss the limitations with the current literature related to these social cognitive processes. I also outline the research aims of the current study. I next present the methods for the present study and explain the novel fMRI task. I end the Method section by reviewing my data analysis plan, explaining how I will test my specific hypotheses. In the subsequent Results section, I present the results from the current study, which indicate that a live context does alter the brain's response to processing language and mental states. Finally, I conclude with a discussion that contextualizes these findings, focusing on how this study provides the cornerstone of a larger developmental research program.

The Role of Live Interaction in Modulating Behavior

Live social interaction—defined as real-time, contingent, and reciprocal interaction with a real human partner—is privileged from early infancy. Newborns are drawn to faces (Farroni et al., 2005) and biological motion (Simion, Regolin, & Bulf, 2008), and will imitate the facial expressions of conspecifics (Meltzoff & Moore, 1983). By one month of age, infants behave differentially toward people versus objects (Trevarthen, 1998), and two-month-old infants are sensitive to contingent social interaction presented over live video-feed (Murray & Trevarthen, 1985; Stormark & Braarud, 2004; although see Rochat, Neisser, & Marian, 1998). In studies by Murray and Trevarthen (1985; 1986), infants and mothers viewed videos of each other via connected monitors, and, depending on the condition, the video was either live or recorded. In the live condition, both parties were able to interact over the live video feed, and in the recorded condition, they saw video from their previous interaction. While viewing recordings, both infants and mothers acted differently: infants were more distressed, and mothers used less infant-directed speech. These findings indicate that the contingent, live response of both the mother and the young infant are crucial to the dynamics of interaction. The importance of live interaction grows over the first year of life, as infants engage in games, emotional interactions, and protoconversations (non-linguistic back-and-forth vocal exchanges) with caregivers (Reddy, 2003; Trevarthen, 1998).

These early perceptual and behavioral biases for live social interaction build to the capacity for joint action, in which shared goals represent more than the sum of individual minds or actions (Gallotti & Frith, 2013; Tomasello, Carpenter, Call, Behne, & Moll, 2005). In the second year of life, infants' capacity for joint action can be observed in their ability to understand and coordinate pointing gestures (Liebal, Bane, Carpenter, & Tomasello, 2009), one of several early roots of joint action (reviewed in Carpenter, 2009). By preschool, children collaborate with adults and appear to find collaboration intrinsically rewarding, even when it is not necessary to complete a task (Gräfenhain, Behne, Carpenter, & Tomasello, 2009; Warneken, Chen, & Tomasello, 2006). Of special interest to the proposed study is the finding that preschoolers will coordinate their drum beats with a live social partner, but not a recorded sound (Kirschner & Tomasello, 2009).

Joint action remains a crucial part of social cognition into adulthood, and it has been examined in maturity with well-controlled laboratory tasks. For example, individuals achieve the same motoric goal more quickly when they believe that they are working as part of a group as opposed to working alone (Tsai, Sebanz, & Knoblich, 2011). Even when the social interaction is reduced to two participants moving cursors on a screen, participants are more likely to find each other's cursors than distractor targets (including a target that shadows their own movements; Auvray, Lenay, & Stewart, 2009).

The latter study suggests that live interaction is more than contingency alone; the shadowed target operated contingently, but did not fool the participants. In sum, mounting evidence from behavioral psychology indicates that social interaction produces emergent patterns of behavior that are not reducible to the individual level (De Jaeger, Di Paolo, & Gallagher, 2010). This is especially for true for two specific social cognitive domains: language processing and mentalizing.

Language processing.

Live interaction is necessary for first language acquisition. Very early perceptual biases (e.g., sensitivity to contingency and faces) help foster attention to speech (Locke, 1995), and, by the second half of the first year of life, live interaction is crucial for phonetic development. In a classic study, Kuhl and colleagues (2003) found that American infants from monolingual English-speaking homes were able to retain their ability to discriminate Mandarin phonemes only when they interacted live with a Mandarin speaker—watching videos of that speaker did not preserve the contrast. Later studies showed that live interaction also increased infant language production (Goldstein & Schwade, 2008) and that live interaction could even improve second language learning (Verga & Kotz, 2013). Kuhl (2007) has theorized that these effects, at least for infants, may be due to increased motivation (via increased arousal and attention) and joint attention, perhaps due to the ability of both participants to follow and respond to each other's gaze. This view is supported by more recent evidence from online video-based interaction, in which toddlers are able to learn from a social partner over video, as long as that partner is responding in a contingent way (Roseberry, Hirsh-Pasek, & Golinkoff, in press). This finding is consistent with the evidence that children as young as two months are sensitive to contingencies over video feed (Murray and Trevarthan, 1985).

Beyond language learning, live interaction influences mature language processing. For example, the previously reviewed theory of joint action has been extended to language processing in its prototypical live context—back-and-forth conversation (Brennan, Galati, & Kuhlen, 2010; Clark, 1996; Garrod & Pickering, 2004; Garrod & Pickering, 2009). When engaged in live conversation, speakers use similar grammatical structures (Branigan, Pickering, & Cleland, 2000), agree on conversational common ground (Brennan & Clark, 1996), and engage in near-seamless turn-taking (Sacks, Schegloff, & Jefferson, 1974). This synchronous activity extends beyond language; speakers in a conversation coordinate their eye movements (Richardson, Dale, & Kirkham, 2007) and bodily motions (Chartrand & Bargh, 1999; Shockley, Santana, & Fowler, 2003). These phenomena seem to emerge out of conversation and cannot be decomposed into individual behaviors outside of an interactive context.

Mentalizing.

Mentalizing, or theory of mind, is the understanding that others have mental states and that these states can be different from one's own and from reality. This ability has been traditionally indexed by the false belief task, which children pass around age four (Wellman, Cross, & Watson, 2001). During this task, children see a character, Sally, place a toy in a box. Then, while Sally is out of the room, another character moves the toy. To pass, children must explicitly indicate, either verbally or via pointing, that Sally will still look for the toy in the original box. That is, they must understand that she has a mental state different from reality.

Intriguingly, recent work from young infants and toddlers suggests that children pass an implicit version of this test at a much earlier age (Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007). In these implicit tasks, children show anticipatory looks to the box where the character believes the object should be. Related mentalizing research has been conducted in samples of adults and children with autism, which is a developmental disorder characterized by social impairment. These patient groups fail to display anticipatory real-time looking behavior, but, at least in adulthood, often pass third-party explicit versions of the false belief task (Senju, Southgate, White, & Frith, 2009; Senju et al., 2010). Such findings indicate that there may be dissociable mentalizing systems: one may be more cognitive and reflective, and the other may be more automatic and more frequently used in spontaneous, naturalistic, second-person interaction (Apperly & Butterfill, 2009). One theory suggests that, rather than computing a social partner's internal states in a detached way (which can be a route to success in a standard explicit theory of mind paradigm), humans use online information from the ongoing interaction to guide their real-time mentalizing (Klin, Jones, Schultz, & Volkmar, 2003).

Despite potential alternative mechanisms of real-world mentalizing, behavioral paradigms probing the understanding of mental states almost exclusively employ recorded, third-person stimuli. Mentalizing, however, is a complex process, and there is direct evidence that its subcomponents, such as language and person perception, are affected by live interaction. Several studies have found that participants look more at videotaped stimuli of people than real people (Foulsham, Walker, & Kingstone, 2011; Laidlaw, Foulsham, Kuhn, & Kingstone, 2011), perhaps out of a consideration of the other person's mental state and a desire to avoid being thought of as impolite or socially inappropriate. Further, there is evidence that conversation hinges on theory of mind, because individuals need to take into account another's mental state to chose appropriate topics, take turns, and understand referents (Brennan et al., 2010). Thus, there is suggestive evidence that mentalizing is different in live contexts, but there has yet to be systematic study of this phenomenon.

Live interaction and social disability.

Evidence from various social disabilities underscores the importance of investigating response to live stimuli. Individuals with autism show great impairment in real-world social interaction, but often show preserved skills when interpreting offline, third-person social stimuli (Klin et al., 2003: Senju et al., 2009). This deficit in live interaction appears early in development, as evidenced by the fact that infants at risk for autism show atypical patterns of visual fixation to their mothers' eyes during a live video paradigm (Merin, Young, Ozonoff, & Rogers, 2007). Such interaction-specific deficits are also present in adults; individuals with autism show an atypical arousal response to live versus videotaped social interaction (Riby, Whittle, & Doherty-Sneddon, 2012). Evidence from the eye-tracking literature suggests that the differences between typical individuals and those with autism are accentuated as the stimuli become more naturalistic (e.g., adding motion, sound, realistic scenes; Speer, Cook, McMahon, & Clark, 2007). Other social disorders, such as social anxiety, may also be best understood in the context of live interaction. For example, a review found that measures of social competence that employed complex interactions better predicted symptomatology and outcomes than isolated laboratory measures (Bierman & Welsh, 2000). Studies suggest that live interaction alters behavior differently in typical individuals as compared to individuals with social disorders, such as autism. Therefore, systematic examination of what components of live interaction alter social cognition could help illuminate the deficits underlying these disorders.

Neuroscience of Live Interaction

Social interactions can be studied in a variety of ways, including sophisticated dynamic computational models, more naturalistic studies, or the behavioral paradigms reviewed in the previous section. Perhaps unsurprisingly, given the ability of neuroimaging to illuminate mechanisms underlying behavior (Diamond & Asmo, 2008), several recent papers have argued for the application of a live interactive perspective to neuroimaging studies (e.g., Guionnet et al., 2010; Tylén, Allen, Hunter, & Roepstorff, 2012), and some limited literature has pursued this approach. Although not all of the extant second-person neuroscience studies directly target the question of processing of live versus recorded stimuli, they do elucidate some of the processes involved in social interaction. In this section, I review these interactive neuroscience studies as they relate to language processing and mentalizing.

Language processing.

Despite the strong behavioral evidence that live interaction can alter language processing, socially interactive neuroscience studies of language processing are quite limited and mostly involve conversational paradigms with limited experimental control. Suda and colleagues (2010) used functional near-infared spectroscopy (fNIRS) with two participants engaged in face-to-face conservation. Compared to the control condition of repeating nonsense syllables, conversation engaged frontal and temporal regions often implicated in mentalizing and language. Researchers found that the effect was larger for participants who were rated as more cooperative. Later replications with the same design found that increased activation during conversation was not present in individuals with schizophrenia (Takei et al., 2013), and that typical individuals with more autistic-like traits showed smaller increases in activation (Suda et al., 2011). These studies are intriguing, but fNIRS has limited spatial resolution compared to fMRI, and these specific paradigms have limited experimental control.

Researchers have also used fMRI hyperscanning techniques—where two participants are scanned simultaneously—to study linguistic processing. Stephens and colleagues (2010) had one participant tell an unrehearsed 15-minute story in the scanner, while a participant in another scanner listened to the live speech. The control stimulus was a 15-minute unrehearsed story in a language that the listener did not understand. The researchers modeled the similarity of the neural patterns of the speaker and listener in both basic auditory and language processing regions and found that the neural patterns of the participants became coupled only when the participant understood the live speech, with greater coupling associated with improved comprehension.

Another recent fMRI hyperscanning study investigated back-and-forth conversation (AbdulSabur, 2013). In the experimental condition, participants engaged in a conversation, and in the control condition, the participants talked about unrelated topics, interrupting each other whenever they chose. The author found increased activity in left temporal regions (including regions often associated with mentalizing) when participants were having a conversation and found that this increase was accentuated immediately prior to conversational turn-taking. Together, these studies converge with the behavioral evidence that the brain is sensitive to the interactive properties of conversation; however, results from hyperscanning paradigms are difficult to directly compare to third-person, traditional approaches to studying language.

Mentalizing.

The neural correlates of classic mentalizing paradigms (e.g., false belief reasoning) have not been investigated in a second-person context. Joint attention, however, has been examined in a contingent, interactive framework (e.g., Redcay et al., 2010; Schilbach et al., 2010). Although joint attention—or the coordination of two individuals' attention on an object—does not require explicit belief computation, many researchers have argued that joint attention is more than reflexive gaze following (reviewed in Carpenter & Liebal, 2011). In this framework, true joint attention involves shared knowledge that the other party is also directing attention to the same object, and thus entails processing and responding to the internal state of another individual. Consistent with this perspective, second-person neuroscience paradigms have found that joint attention increases activity in traditional mentalizing regions.

Schilbach and colleagues (2006; 2010) designed a paradigm in which participants in an fMRI scanner interacted with a computer avatar—which they believed was controlled by a real person—in a joint attention task. Participants either initiated joint attention (by looking in a particular direction) or responded to joint attention bids by following the avatar's gaze in a particular direction. As an experimental control, on certain trials, the avatar or the participant would respond noncontingently (i.e., look in the opposite direction). Researchers found that following the avatar's gaze increased activation in dorsal medial prefrontal cortex (dMPFC), a region associated with social processing, and that directing the gaze of the avatar via initiation of joint attention activated reward regions.

Redcay and colleagues (2010; 2012) examined true face-to-face interaction via video feed, in which a participant in the scanner played a live joint attention game with an experimenter outside of the scanner. Each individual could see the other's eye movements, and, depending on the trial, either initiated or responded to joint attention cues. As with the previous study, the control involved participants responding to non-social attention cues. Live joint attention, as compared to the control, differentially activated regions involved in mentalizing—namely the posterior superior temporal sulcus (pSTS) and temporal parietal junction (TPJ). This finding is especially relevant to the current study's paradigm given that these live joint attention games do not involve the explicit belief computation required by traditional theory of mind paradigms. Such findings indicate that perhaps spontaneous low level mentalizing in real-life contexts engages similar regions to more complex mentalizing processes.

Another relevant line of research comes from game theory and neuroeconomics, in studies where researchers examine participants' brain activity when they are playing, or believe they are playing, against a live social partner, as compared to playing against a computer. Such comparisons consistently show increased activation in social engagement regions when participants are playing against a human. For example, increased activation in dMPFC has been observed both in "rock-paper-scissors" and in prisoner dilemma games (Gallagher, Jack, Roepstorff, & Frith, 2002; McCabe, Houser, Ryan, Smith, & Trouard, 2001). More recent studies employing a variety of game theory paradigms have found similar patterns of increased activation in multiple regions linked to mentalizing (e.g., Fukui et al., 2006; reviewed in Camerer, 2008). For example, when participants played a human in a multi-step strategy game, they show increased STS activity compared to playing a computerized opponent (Coricelli & Nagel, 2007). Further, participants who exhibited more high-level strategies also exhibited higher activation in social brain regions linked to mentalizing.

Limitations in the current literature.

In response to this neuroimaging and behavioral evidence, Schilbach and colleagues (2013) argued for second-person neuroscience, postulating that the social brain operates differently in interaction as opposed to traditional studies of observation (e.g., watching a cartoon of Anne moving a toy while Sally is out of the room). Most of the extant studies, however, fail to directly address the question of whether the neural bases of mentalizing or language are different when considering a third party character versus a live social partner. For example, the language studies do not directly contrast listening to live versus recorded speech; rather, they contrast simple syllables or noncontingent, unrelated speech with back-and-forth conversation. The studies of joint attention examine trials in which the avatar (ostensibly representing a real person; Schilbach et al., 2010) or experimenter (via videofeed; Redcay et al., 2010) responds to joint attention cues and compare these trials to those in which this social partner was nonresponsive. Such studies focus on the role of contingency, rather than examining whether participating in contingent joint attention with a perceived live social partner is different than following identical joint attention cues when such cues are perceived to come from a prerecorded computer program. Similarly, the studies involving playing against a computer versus against a human did not compare how individuals respond to recorded social stimuli versus live social stimuli, and instead contrasted real-time social contingencies with real-time non-social contingencies.

In addition to not directly comparing observation versus interaction, the current literature is frequently limited by methodological issues, such as experimental control. Live interactions can be challenging to script, and it is difficult to determine which specific events are of interest in a fast-paced real-time interaction (De Jaeger et al., 2010). Hyperscanning data are particularly problematic to interpret, since it is difficult to have a priori hypotheses about what coordinated brain activity would look like on a neural level beyond simple simultaneous changes in activation. For example, as Konvalinka and Roperstorff (2012) discuss, social interaction involves multiple networks operating on different time scales, which may require more sophisticated modeling processes than those employed by most previous hyperscanning studies.

Beyond these analytic issues, fMRI analyses require strict control conditions in order to ensure that researchers isolate the effect of interest. This consideration, combined with the fact that the fMRI scanner is a noisy and isolated environment, make fMRI particularly ill-suited to study difficult-to-control live interaction. For example, without controlling for attention or motivation, it is possible that any effects of live interaction are instead due to differences in attention or stimuli (Redcay et al., 2013). Studies of conversation are especially sensitive to confounding variables, as that approach makes it difficult to control for low-level characteristics (e.g., length of pauses, speech rate, etc.). For example, the control condition in some conversation studies is nonsense syllables, which differ from conversation for many other reasons besides being less socially interactive. Thus, much of the current neuroscience literature on social interaction in the context of language processing and mentalizing fails to address a main claim of the second-person neuroscience movement—that observation is different from interaction—either because it is not the direct question of interest or due to methodological limitations.

Direct comparisons of live and recorded social stimuli.

A very small number of studies have directly compared neural responses to live versus matched, recorded social stimuli. For example, Pönkänen and colleages (2011) used event-related potentials to compare response to live faces versus static photographs of the same faces, for both direct and averted gaze. The live faces were presented from behind a shutter that became transparent for the same length of time that the photographs were presented. Researchers found that the brain was more responsive to direct gaze for only the live condition, and suggested that this finding could reflect increased self-relevance or spontaneous mentalizing about an imminent social partner. These findings are consistent with similar electroencephalography paradigms, which have also found differential brain activity for direct versus averted gaze only in the live condition (Hietanen, Leppänen, Peltola, Linna-aho, & Ruuhiala, 2008; Pönkänen, Peltola, & Hietanen, 2011).

In the previously mentioned study by Redcay and colleagues (2010), researchers contrasted the participant's brain activity during segments of live interaction over an audio and video feed with the experimenter to the activity that occurred while the participant viewed recordings of the experimenter from previous sessions (from a prior interaction with the same or a different participant). This contrast controlled for the basic characteristics of the social stimuli, but varied the interactive context. Compared to the recorded condition, social regions of the participants' brains (TPJ and pSTS) were more active during the live condition.

Although such studies are informative, they do not directly address how a live real-time social partner changes language processing or mentalizing. For example, in the Redcay et al. (2010) paradigm, researchers examined continuous social interaction over live video feed, and this interaction was composed of several, often overlapping events, including a social partner's gaze shifts, eye contact, turn taking, and contingent responding. Given this design, it was not possible to determine the effect of these separate live cues on the resulting neural response. Disentangling these separate processes—for example, examining live language processes separately from eye contact or contingency—could help pinpoint the effect of live interaction on separate cognitive processes. Further, this prior study did not explicitly attempt to match attention or motivation across conditions beyond instructions. Thus, the question of whether live interaction alone modulates the neural correlates of mentalizing or language processing is unresolved. Answering this question not only has methodological and theoretical relevance, but also may have broader clinical applications. For example, during third-person mentalizing about recorded stories, individuals with autism activate the same regions as typical individuals, despite real-life deficits in mental state attribution (Dufour et al., 2013). The characterization of the neural networks involved in real-life social processes may improve understanding of social disabilities by elucidating deficits obscured by third-person paradigms.

Overview of the Socio-Communicative Network

In order to contextualize the current study, it is necessary to briefly review the known neural correlates of language processing and mentalizing. The second-person approaches reviewed in the previous section constitute the minority of neuroimaging investigations in these areas; both processes have been almost entirely understood in a third-person context. For example, a common paradigm to examine language processing is to contrast responses to prerecorded sentences versus control sentences (e.g., backwards speech), and a common paradigm to study mentalizing is to compare participant responses to stories about false beliefs to responses to control stories (e.g., a photograph depicting an untrue representation of the world, such as a house that later burned down). Although these approaches have limitations, they represent an important starting point for hypotheses about regions that might be differentially activated by live interaction.

Neural correlates of language processing.

Several meta-analyses (Friederici, 2011; Price, 2010) have revealed a consistent set of regions involved in language processing, a 'language network.' Regions frequently implicated in this network are inferior frontal gyrus (IFG, or Broca's area), superior temporal gyrus (STG, or Wernicke's area), medial temporal gyrus (MTG), inferior parietal regions, and angular gyrus. For most participants, especially those who are right-handed, language is left-lateralized.

As the present study is focused on language comprehension, specifically with regard to sentence comprehension and prosody, I will focus on these facets of the speech processing literature. Although many overlapping regions are implicated in speech perception, there is evidence that right temporal regions are involved in more social aspects of speech, including prosodic processing (Friederici, 2002), although there is some individual variability (Ethofer, Van De Ville, Scherer, & Vuilleumier, 2009). Broadly, the language processing fMRI literature distinguishes between those regions involved more in pitch and prosody (i.e., STG) and those regions involved in sentence comprehension (i.e., left IFG; reviewed in Price, 2010, although see Frederici, Rüschemeyer, Hahne, & Fiebach, 2003 for evidence of superior temporal involvement in syntax).

Neural correlates of mentalizing.

The 'mentalizing network' (Frith & Frith, 2003), or the brain regions activated by mentalizing, is incredibly robust, with the same core brain regions activated for almost all individuals regardless of the specific paradigm. This core network involves bilateral temporal parietal junction (TPJ), dorsal medial prefrontal cortex (dMPFC), anterior superior temporal sulcus (aSTS), and precuneus/posterior cingulate cortex (PC/PCC; Saxe, 2009). As reviewed by Koster-Hale and Saxe (2013), this network is consistently activated by stimuli ranging from sentences to cartoons to videos, and by tasks ranging from personality judgments to explicit false belief processing to spontaneous mentalizing, although the relative strength and pattern of activation may depend on the task.

Although the regions of this network frequently co-activate, there is evidence to suggest that TPJ, specifically right TPJ, is most selective for mental state processing (Perner, Aichhorn, Kronbichler, Staffen, & Ladurner, 2006; Saxe & Kanwisher, 2003). Evidence for this comes not just from fMRI studies, but also lesion (Apperly, Samson, Chiavarino, & Humphreys, 2004) and transcranial magnetic simulation studies (Young, Camprodon, Hauser, Pascual-Leone, & Saxe, 2010). Although there has been controversy over whether the TPJ's selectivity is due to its role in processing unexpected stimuli, more recent research indicates that the region's role in mentalizing is dissociable from its role in other cognitive processes (Young, Dodell-Feder & Saxe, 2010).

In contrast to the highly selective TPJ, evidence indicates that dMPFC may be involved in more general social processing (Saxe & Powell, 2006), including personality trait judgments (Mitchell, Cloutier, Banaji, & Macrae, 2006), gaze detection (Pierno, Becchio, Turella, Tubaldi, & Castiello, 2008), communicative intent processing (Kampe, Frith, & Frith, 2003; Schilbach et al., 2006), and animacy detection, even when the animate agents are moving triangles (Tavares, Lawrence, & Barnard, 2008). As with the TPJ, there have been some claims that the dMPFC is more involved in general cognitive processes, but a meta-analysis suggests that it does play a specific role in social processing (Van Overwalle, 2011).

Finally, although not consistently activated in traditional false belief paradigms, the posterior region of the STS is activated by a variety social stimuli, including speech sounds (Möttönen et al., 2006), faces (Ethofer et al., 2013; Haxby, Hoffman, & Gobbini, 2000), biological motion (Kaiser et al., 2010; Puce & Perrett, 2003), gaze shifts (Kingstone, Tipper, Ristic, & Ngan, 2004; Pelphrey, Singerman, Allison, & McCarthy, 2003), and animate agents (Shultz & McCarthy, 2012). Especially relevant to the current paradigm is the role of the pSTS, particularly the right pSTS, in processing others' goals and intentions, beyond simply detecting motion or animacy (e.g., Gao, Scholl, & McCarthy, 2012; Pelphrey, Morris, & McCarthy, 2004; Shultz, Lee, Pelphrey, & McCarthy, 2010).

Socio-communicative network.

Language processing and mental state processing are often investigated separately, but evidence suggests that many regions in the language network and mentalizing network overlap as part of a larger socio-communicative network. For example, the TPJ is involved in both narrative comprehension and mentalizing (see Mar, 2011 for a review). Although right IFG is strongly linked to syntactic processing, it is also implicated in social behaviors such as irony comprehension (Wang, Lee, Sigman, & Dapretto, 2006), and dMPFC, implicated in many studies of social engagement, has also been linked to linguistic processes such as metaphor comprehension (Benedek et al., 2014). Similarly, pSTS is involved in processing syntactic movement (Ben-Shachar, Palti, & Grodzinsky, 2004), semantics (Xu, Kemeny, Park, Frattali, & Braun, 2005), intention (Pelphrey et al., 2004), and non-linguistic vocal communication (e.g., laughter; Shultz et al., 2012). Given the fact that brain regions in the socio-communicative network serve multiple functions, assuming that activation in a particular area indexes a particular domain-specific cognitive process—language processing versus mentalizing—can be problematic (cf. Poldrack, 2011).

In a review of the varied tasks linked to STS activation, Redcay (2008) posited that a common denominator is a task's social or communicative saliency. That is, the broader socio-communicative network may be evidence for a common factor in both language processing and mentalizing tasks that modulates activity across these regions. One possibility is that this network may ontogenetically result from the fact that language and mentalizing are frequently intertwined with social interaction. Given the hypothesis that live interaction could serve as a salient factor for the socio-communicative network, my aim in the current study was to determine if live context could change how the brain processes social stimuli. That is, I was interested in determining if the brain's pattern of activation in response to language changed as a result of this specific extralinguistic factor (i.e., listening to a live speaker). Similarly, my research aim regarding mentalizing was to determine if factors beyond the cognitive content of a belief representation were able to influence whole brain activation. In this study, the pattern of activated regions was not used to infer something about mentalizing versus linguistic processes. Rather, the pattern of results was used to determine if live interaction alone was enough to modulate activity in regions traditionally assumed to be predominately responsive to the cognitive computational aspects of language and mentalizing.

Current Study

Despite behavioral evidence that live interaction modulates language processing and mentalizing, the neural correlates of these processes have been almost exclusively characterized in offline contexts. Of the few studies that have taken a more interactive approach, most have not directly addressed the question of how live interaction, as compared to the observation of recorded stimuli, differently activates the brain. The current study aimed to close this gap by examining, in a novel and well-controlled paradigm, the neural correlates of language processing and mentalizing in a live, interactive context.

Participants completed a novel fMRI task with three conditions. In the Live condition, participants believed they were making choices for a live social partner who was conveying her preferences with them over a real-time audio feed. In the two other Not-Live conditions, participants listened to prerecorded audio and made choices based on the expressed preferences of third-party characters. In one of these prerecorded conditions, Social, the voice was high-pitched and friendly, and was thus matched to the voice in the Live condition. In the second prerecorded condition, Standard, the voice was lower-pitched with less prosodic variation.

In order to provide for tightly controlled comparisons between the Live and Not-Live conditions, each individual trial proceeded in a specific order. First, the participant heard the Story portion, which presented a short spoken vignette (e.g., "There are two things on the breakfast menu. One is pancakes and one is a bowl of fruit."). Then the participant heard the audio from the Social Prediction portion, which consisted of a spoken preference and question (e.g., "I/Megan am/is trying to eat healthy. Which food should I/she eat?"). The participant then chose one of two possible on-screen responses and received feedback. The feedback was designed to match attention and motivation across all three conditions. To further ensure matched comparison between conditions, the stimuli corresponding to the Live condition were prerecorded, although participants were led to believe that the presentation was live. This level of control was necessary to isolate the effect of a real-time interactive context from other factors such as attention. This paradigm aimed to address important gaps in the social neuroscience literature by allowing for a direct examination of how a live interaction can change behavioral and neural processing.

Research aims.

My broad research aim for the current study was to compare social cognition in offline versus interactive contexts using both behavioral and neuroimaging data. In this section, I set forth the specific research aims, and later in the Method section, I delineate the specific hypotheses and measures (also presented in Table 1).

Research aims related to behavioral data. The behavioral literature indicates that a live interactive context can improve task performance (e.g., Tsai et al., 2011). Further, this effect appears to be attenuated in individuals with autism or autistic-like personality traits (Freeth, Foulsham, & Kingstone, 2013). My first research aim was to quantify and account for individual differences in participants' subjective perceptions of how live each condition seemed. My second research aim was to determine how a live context changed the processing of language and mental states, as measured by behavior (i.e., reaction time).

Research aim related to the neural correlates of live language processing. Past research has found increased activation in socio-communicative regions during conversation, but such research in limited both theoretically and methodologically. Specifically, it seems unable to answer the question of whether a live context in and of itself changes neural processing when compared to recorded contexts. For this set of analyses, I examined the Story portion of the experiment. The only difference between conditions was the speakers' voice and whether the participant believed the speaker is live. My research aim was to identify how a live context changed the neural processing of language.

Research aim related to the neural correlates of live mentalizing. Past research with joint attention and computer-versus-human strategy games has found evidence for increased activity in mentalizing regions in response to human or contingent partners. Such research, however, has not directly compared mentalizing about a live social partner

Table 1. *Research Aims & Hypotheses*

Research Aim 1: Identify how live context behaviorally changes language processing & mentalizing		
Hypothesis 1	Participants' subjective experiences of "liveness" (based on the post-test questionnaire) will be higher in the live condition than the two not-live conditions	Supported
Hypothesis 2	Participants' "liveness" ratings will be negatively related to autistic-like personality traits for the live condition only.	Supported
Hypothesis 3	Performance in the live condition will be different than in the not-live condition, as measured by reaction time.	Not supported
Hypothesis 4	The magnitude by which the live condition affects performance will be significantly related to perceived "liveness"	Not supported
Research Aim 2: Identify how a live context changes neural processing of language		
General hypothesis	Processing speech from a live social partner will modulate neural activity compared to matched recorded speech	Supported
	Hypotheses about Candidate Regions for Specific Contrasts	
Hypothesis 5	Live vs. Standard Condition: The Live condition will more strongly engage regions of the brain associated with social engagement and differentially engage brain regions related to pitch and prosody, but will not differentially engage sentence comprehension regions.	Partially supported—predicted social engagement regions not engaged
Hypothesis 6	Live vs. Social Condition: The Live condition will more strongly engage regions of the brain associated with social engagement, but will not differentially engage language-processing regions.	Supported
Hypothesis 7	Social vs. Standard Condition: There will be no difference between the Social condition and the Standard condition in brain regions associated with social engagement or sentence comprehension, but there will be a difference in language regions related to pitch and prosody.	Supported
Research Aim 3: Identify how live interactive context changes neural processing of mental states		
General Hypothesis	Mentalizing about a live social partner will modulate neural activity as compared to an identical mental state computation about a third-party character	Supported
	Hypotheses about Candidate Regions for Specific Contrasts	
Hypothesis 8	Live vs. Standard: The live condition will more strongly engage the brain's mentalizing network and differentially engage language processing regions related to pitch and prosody, but will not differentially engage sentence comprehension or attention regions.	Partially supported—whole mentalizing network not engaged
Hypothesis 9	Live vs. Social: The Live condition will more strongly engage the brain's mentalizing network, but not differentially engage any language processing or attention regions.	Partially supported—whole mentalizing network not engaged
Hypothesis 10	Social vs. Standard: The Social condition will show differential activity in regions associated with pitch and prosody, but will not show a difference in mentalizing, sentence comprehension, or attention regions.	Supported

versus a mentalizing about a character in a story, controlling for the content of the mental state attribution. To address this question in the present study, I examined the Social Prediction period. This included the expression of the social partner (i.e., “my”) or character’s preference, the question as to what the social partner (i.e., “me”) or the character should do, and the period where the participant saw the answer choices and made a response. My research aim was to identify how a live interactive context changed the neural processing of mental states.

Chapter 2: Methods

Participants

Thirty-three adults (15 males), aged 18-27 years (*M*=21.58, *SD*=2.31), participated in the study in exchange for course credit or payment. Participants were members of the University of Maryland community, and extensive prescreening determined that all participants were native English speakers, had normal hearing, normal or corrected-to-normal vision, no first-degree relatives with autism or schizophrenia, and no personal history of any neurological impairments or psychological disorders. Two participants were not scanned: one was unable to enter the scanner, and one was unable to complete the behavioral questionnaires and thus was not scanned. Two further participants were excluded due to the fact that they failed to believe that the live interaction was live (i.e., in a post-test questionnaire, they explicitly indicated that they believed the live stimuli were prerecorded). Thus, the final sample was 29 adults (12 males) aged 18-26 years (*M*=21.46, *SD*=2.09). This sample size is consistent with literature on fMRI power, and should reliably detect effects with a fairly conservative correction for multiple comparisons, and still preclude spurious results (Desmond & Glover, 2002; Thirion et al., 2007; Yarkoni, 2009). The University of Maryland Institutional Review Board approved all consent forms and fMRI screening forms, as well as all study protocols (Appendix A).

Procedures

Overview of fMRI stimuli and design.

For the current experiment, all of the stimuli, even for the Live condition, were prerecorded in order to control for potential stimuli differences between or within subjects. My main research question, however, was predicated on the participant believing that some of these prerecorded interactions were, in fact, live. Therefore, I created the illusion that some of the stimuli were live. In the following sections, I first explain how this was achieved, and then describe the experimental design and stimuli. I end by describing the fMRI image acquisition parameters.

Creating the illusion.

When a participant arrived at the Maryland Neuroimaging Center (MNC), he or she completed the study consent form and re-completed the fMRI safety screening form. Next, I took the participant to the "experimenter room" (Figure 1A). There, I explained that, for this study, sometimes I would be talking with him or her directly over a live audio feed, and sometimes he or she would hear pre-recorded audio from other speakers (see Appendix B for the full script). I showed the participant how the laptop in the experimenter room was set up with a live video chat connection. I explained that participants would be asked questions via audio, and that, for the trials where I spoke to the participant live, I would see his or her responses using the web camera in the scan room (this camera was pointed at a display that shows which button the participant pressed in response to a question; Figure 1B). I explained that I would then open up a video link and give feedback (e.g., a thumbs up). For the other trials, I explained that only the computer would see the participant's answer, and would the computer automatically show feedback (e.g., a gold star or smiling face). I checked participants' comprehension of these instructions, to ensure that each participant understood that there would be a live condition and a prerecorded condition.

Once the participant was in the scanner, I talked to him or her from the experimenter room over a truly live video chat (Figure 1C), and gave clues to indicate the chat was live (e.g., “Hi [participant name], glad you were able to wear your contacts today”). During this portion of truly live interaction, we practiced some of the button presses (e.g., I said “Press 1 if it’s Tuesday and 2 if it’s Wednesday”). The participant saw the button that he or she pressed light up on our mutual chat screen, and then I gave live contingent feedback (e.g., thumbs up).

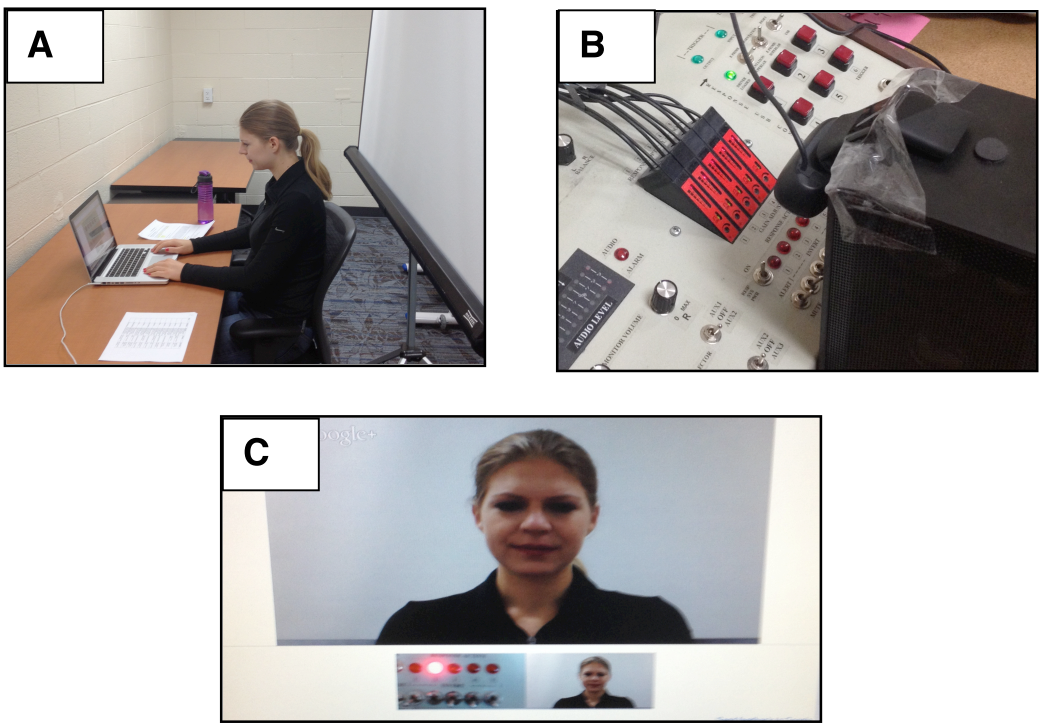

Figure 1. Experimental setup for creating the live illusion. Before the study begins, the participant views the experimenter room (A), where I have set up a live video chat. The webcam on the laptop in the experimenter room is trained on my face. The second webcam in the video chat (B) is trained on a button display, which will light up when the participant inside the scanner presses a button. When inside the scanner, before the main experiment begins, the participant views a screen depicted in (C). The participant sees my face and hears me talking to them. I ask them a question (e.g., press 1 if it is summer and 2 if it is winter). Immediately after pressing the button (in the case of (C), the button is 2), the participant sees their choice illuminated and thus knows that I have also seen their choice. I then give live contingent feedback (e.g., thumbs up).

After this portion of truly live interaction was over, we did a brief set of practice trials in the style of the actual experiment (i.e., prerecorded "Live" trials). After these trials, we returned to the truly live video chat, and I commented on the question I just "asked", reinforcing that I was talking to the participant live and that I was able to see his or her responses. After this final portion to solidify the live illusion, the main experiment began.

Stimuli presentation in the scanner.

Participants viewed 36 individual trials in the scanner (Appendix C), evenly distributed across 4 runs. For each trial, participants listened to a short story that presented two options, then heard a question about what I, or a third-party character, should do (Figure 2). The participant picked one of the options via button press. After making the choice, participants saw contingent feedback. The task was programmed and presented using the Psychophysics Toolbox Extension for MATLAB 7.6 (PTB-3; Brainard, 1997; Pelli, 1997). For the purposes of the current study, two events were of interest: first, the Story (e.g., "There are two things on the breakfast menu. One is pancakes and one is a bowl of fruit."), and second, the Social Prediction period. The latter consisted of the prompt (e.g., "I/Megan am/is trying to eat healthy"), the question (e.g., "Which food should I/Megan eat?"), and the participant's choice. The feedback portion was designed to ensure that motivation and attention were the same in both Live and Not-Live trials; the issue of mismatched attention has been a problem in past studies of live interaction (reviewed in Redcay et al., 2013). The accuracy and reaction time (RT) of the participant's choice was recorded. Participants had three seconds to respond before their answer was marked as incorrect.

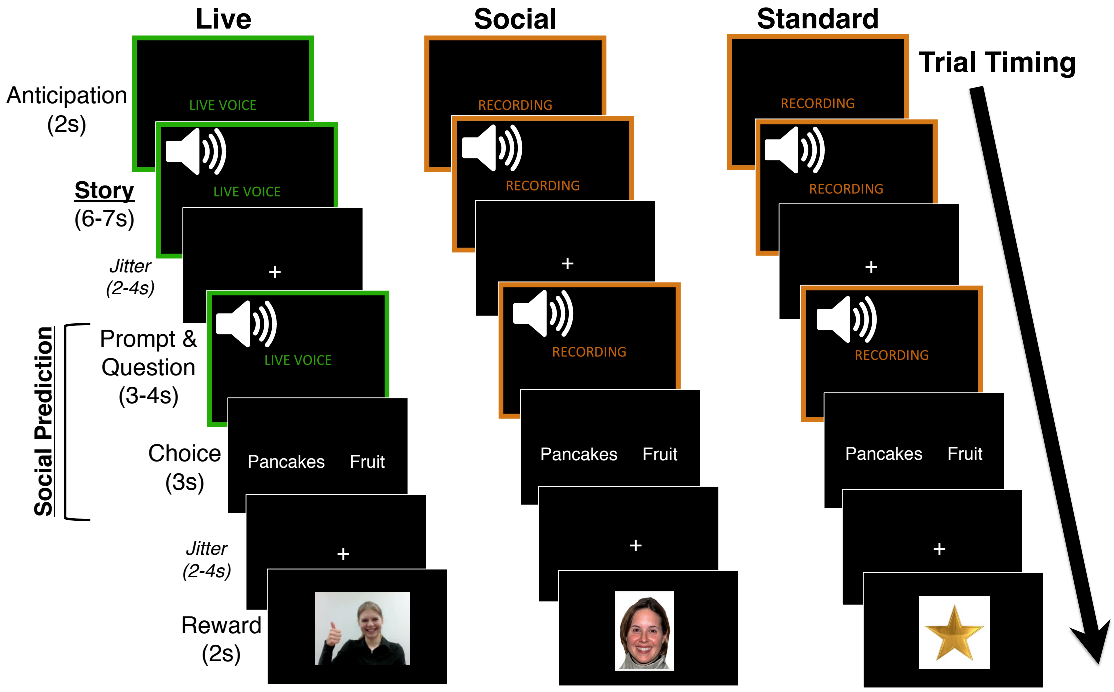

Figure 2. Experimental trial structure. The three conditions (Live, Social, and Standard) are depicted in the three columns. The specific events of interest for the current study (story and social prediction) are bolded and underlined. All participants were presented twelve unique trials in each condition. s=seconds.

Each participant viewed trials from the three conditions discussed previously (Live, Social, Standard). For the Live condition, participants heard my prerecorded voice and, for feedback, saw a video of me (e.g., giving a thumbs up or down). For the Social condition, participants heard a friendly speaker and feedback consisted of a standardized picture of a happy or sad Caucasian female (NimStim; Tottenham et al., 2009). For the Standard condition, participants heard a lower-pitched female speaker and feedback consisted of a gold star or red “x”.

During the audio portion of the Live trials, a black screen reading “LIVE VOICE” in green text was present. For the Social and Standard trials, a black screen reading “RECORDED” in orange text was displayed. The same screen was also displayed for two seconds before the start of each trial’s audio.

A 2-4 second jitter, consisting of a fixation cross, was present between the Story and the Social Prediction, and between the Social Prediction and seeing the feedback. This jitter allowed us to estimate the effects of these events. There were also 20 seconds of baseline (i.e., a fixation cross) at the beginning and end of each run, and an additional 20 seconds of baseline in the middle of each run. The distribution of the trial types (i.e., Live, Social, Standard) and the timing of the jitters and inter-trial intervals were determined by the program OptSeq (http://surfer.nmr.mgh.harvard.edu/optseq/), to ensure the optimal timing allowing for independent analysis of the events. This model was further tested for colinearity using AFNI's 3dDeconvolve (Cox, 1996; http://afni.nimh.nih.gov/), and analysis revealed that all beta values of interest were estimable.

Stimuli characteristics.

Item piloting. The final items that made up the 36 experimental trials were selected from a set of 126 candidate items. These 126 candidate items were piloted on a sample of eight typical adults, recruited from the University of Maryland graduate student population. Items were excluded if any of the following conditions were met: accuracy of 50% or less, the mean RT or standard deviation (SD) of mean RT was more than two SD away from the mean RT for all 126 items, or the number of syllables was more than two SD away from the mean. Consistent with the ultimate aim of extending this paradigm to a developmental population, the selected subset of 103 items was then tested with a group of seven typical children, aged 8-11. After this testing, 30 easy items were selected (on which child and adult accuracy was 100%) and six hard items were selected (on which child accuracy ranged from 43% to 72%). The high number of easy items was designed to ensure that participants would see mostly positive feedback.

The 30 easy items were selected out of a set of 64 items on which both age groups got 100%. These 30 items were selected because they were closest to the mean child RT. The six hard items were chosen out of the 19 items on which children got between 43% and 72% accuracy and were selected such that average RT was matched between the easy and hard items. For the trials with a third-party character name, each name was picked at random from the US Social Security Administration's list of the top 100 American names for male and female infants born in 2000-2009 (roughly the years in which the planned eventual child sample for the current study would have been born; http://www.ssa.gov/oact/babynames/).

Audio stimuli. Each of the three trial types (i.e., Live, Social, Standard) had a single speaker, and this speaker was held constant across participants. All 36 items were recorded by each speaker (although any individual participant heard each item read by only one speaker, according to the randomization process). In order to ensure that differences in speaker audio characteristics would not drive any effects, each speech file was normalized to 60 dB. After normalization, the speakers were compared on pitch, volume, and length of the audio segment (Table 2). This is an approach that is consistent with other literature on fMRI language processing (e.g., Shultz, Vouloumanos, & Pelphrey, 2012). All features were extracted using Praat 5.3 (Boersma, 2002; http://www.fon.hum.uva.nl/praat/). As expected, the Standard speaker was significantly lower-pitched. The Live and Social speaker were matched on volume and pitch for the Social Prediction audio (prompt and question). For the Story, the Social speaker had higher pitch, which is a feature of infant-directed speech (e.g., Trainor & Desjardins, 2002). Therefore, the Social speaker was either matched to, or even friendlier, than the Live speaker for both events of interest. With respect to the length of the Story portion, the difference between conditions for any particular item was always less than 500 milliseconds. The Live social prediction audio was the shortest in duration, due to the fact that the Live speaker is using the monosyllabic "I" instead of a potentially multi-syllabic name.

Table 2. *Audio Characteristics of the Three Experimental Conditions.*

	Live	Social	Standard	$F_{(2,105)}$	Pairwise Comparisons
STORY AUDIO FEATURES					
Length (s)	6.26 (.38)	6.25 (.39)	6.26 (.39)	.089	Live=Soc=Std
Volume (dB)	60.00 (.01)	60.00 (.01)	60.00 (.01)	.062	Live=Soc=Std
Pitch (Hz)	235.3 (8.5)	259.0 (14.4)	180.9 (5.7)	555.29***	Soc>Live>Std
SOCIAL PREDICTION AUDIO FEATURES					
Length (s)	3.53 (.39)	3.85 (.40)	3.66 (.37)	5.79**	Std=Soc>Live=Std
Volume (dB)	60.00 (.01)	60.00 (.01)	60.00 (.01)	.027	Live=Soc=Std
Pitch (Hz)	288.6 (46.3)	295.7 (35.6)	173.9 (7.7)	290.82***	Std>Live=Soc

Note. Values are mean (standard deviation). Post-hoc pairwise comparisons were made using a Tukey's test of multiple comparisons with an alpha of .05. Abbreviations: s=seconds; dB=decibels; Hz=Hertz. **, p <.01; ***, p<.001.

Visual stimuli. The screens reading “LIVE VOICE” and “RECORDED” were matched on number of letters and on luminosity. The feedback stimuli were not matched on visual characteristics, but all of the stimuli were silent and presented for two seconds.

Randomization.

Each participant was assigned one of three stimuli sets. The three sets differed based on which of the 36 items were presented in the Live, Social, and Standard conditions. There were always 12 items in each condition and each item was represented in each condition once. The allocation of these items within each of the three potential stimuli sets ensured that the total amount of time for each condition was matched (e.g., participants heard each speaker for the same amount of time). Additionally, each participant was be assigned to one of four run orders. Within each run, the timing and order of the trial types was predetermined, using the OptSeq procedure described previously (e.g., a run might start with two Social items, followed by a Live item). Between one and two of the hard items were selected for each run, and the rest of the items for that run were classified as easy (based on the piloting data with children). Based on these constraints, items were randomly assigned to different positions within the runs (e.g., the Live item selected could be any of the 12 possible Live items for that participant). The participant’s stimuli set (i.e., which items were assigned to which condition) and run order were predetermined to ensure that all possibilities were represented over the course of the study.

fMRI image acquisition procedure.

fMRI imaging data were collected using a 12-channel head coil on a single Siemens 3.0-T scanner at the Maryland Neuroimaging Center (MAGNETOM Trio Tim

System, Siemens Medical Solutions). The scanning protocol for each participant consisted of four runs of the experiment (T2-weighted echo-planer gradient-echo; 36 interleaved axial slices; voxel size=3.0 x 3.0 x 3.3 mm; repetition time=2200ms; echo time=24ms; flip angle=90°; pixel matrix=64 x 64), and a single structural scan (three-dimensional T1 magnetization-prepared rapid gradient-echo sequence; 176 contiguous sagittal slices, voxel size=1.0 x 1.0 x 1.0 mm; repetition time=1900ms; echo time=2.52ms; flip angle=9°; pixel matrix= 256 x 256).

The parameters for the functional scans were selected after piloting with four typical adults. The four combinations of two potential echo times (TE=24ms and TE=28ms) and two potential bandwidth lengths (2234 and 2442) were tested across each participant's four runs. This allowed intra- and inter-individual comparisons of the four protocols. I compared within-subject signal levels across all four combinations, specifically in regions prone to signal dropout (e.g., reward regions in ventral striatum). The parameters that best allowed signal preservation while maximizing specificity were TE=24ms and a bandwidth of 2234, and these parameter values were used for the experimental scanning.

Behavioral assessments.

Participants completed three behavioral assessments after completing the fMRI portion of the experiment.

Autism Quotient. (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubbley, 2001; Appendix D). The 50-item autism quotient (AQ) is used to measure autistic-like traits in the typical population. Scores on this measure for typical adults have been shown to correlate with visual perspective-taking ability (Brunyé et al., 2012), responsivity to emotional faces (Gayle, Gal, & Kieffaber, 2012), and perceptual learning (Reed, Lowe, & Everett, 2011). For each AQ item (e.g., "I find social situations easy"), participants answered on a 1 (definitely agree) to 4 (definitely disagree) Likert scale. Half the items were reversed scored. Answers of strongly agree and slightly agree were collapsed together, as were answers of strongly disagree and slightly disagree, and participants received either a 1 or a 0 for each item. Higher total scores indicate more autistic-like traits. For the current study, participants completed a pen-and-paper version of this assessment and were not told what the questionnaire was assessing.

Demographic questionnaire. (Appendix E) Participants completed a standard demographics questionnaire, including measures of handedness, race, and parental education.

Post-test questionnaire (Appendix F) Past studies of live interaction with an element of deception have employed post-test questionnaires to ensure that the illusion was successful (e.g., Schilbach et al., 2010; Teufel, Fletcher & Davis, 2010). For this current study, participants completed a 40-item post-test questionnaire to assess how live, likeable, and direct each participant found each of the three speakers (Live, Social, Standard) on a 1 to 7 point Likert scale. Participants were also asked to describe their impressions of the experiment, including whether or not they felt there was anything more to the study than they were told about. This procedure is standard in studies involving deception, because it does not prime a particular type of answer (e.g. Fazio & Zanna, 1978). Participants completed this assessment using a computer-administered Qualtrics survey, and no member of the experimental team was present while the assessment was administered. The first 17 participants who completed the study were given a slightly different version of the survey than its final form, and Appendix F denotes which items were included for each participant.

Debriefing.

At the end of the study and the post-test questionnaire, participants were debriefed (see Smith and Richardson, 1983 for more on the importance of debriefing). They were told that what they thought was live was actually prerecorded, and that this deception was necessary in order to control the exact characteristics of the audio and video stimuli. Participants were also asked specifically at this time if they had any suspicions about the live stimuli. Finally, they were given the IRB-approved debriefing sheet (included in the IRB materials in Appendix A).

Data Analysis & Hypotheses

Preliminary behavioral data analysis.

RT data were cleaned in a two-step process. First, RTs with values of less than 50ms were excluded because such short responses likely do not reflect true performance (Baayen & Milin, 2010). RTs were then inversely transformed, which is preferable to trimming longer RTs, given the limited variability in the current dataset due to the three-second response window (Ratcliff, 1993). Finally, since participants completed slightly different versions of the post-test questionnaire (e.g., some of the participants answered fewer questions attempting to measure subjective perceptions of how live each speaker seemed), all Likert scale items relating to how live a particular condition seemed were averaged together, yielding a single composite score, or liveness score, for each participant for each condition's speaker. A similar composite scale was constructed for the how much the participant liked each condition's speaker (likeability), and much the participant reported being attentive to and motivated by each condition (engagement). After these preliminary steps, the specific hypotheses, detailed below and in Table 1 could be tested.

Testing behavioral hypotheses.

For the behavioral data, I hypothesized that, based on the post-test questionnaire, participants would rate their subjective experiences of live interaction as higher in the live condition than the two not-live conditions (Hypothesis 1; Table 1). Further, given difficulties with live social interaction for people with autism, I hypothesized that the subjective ratings of liveness for the Live condition would be negatively related to AQ scores, which measures autistic-like personality traits (Hypothesis 2). To test Hypothesis 1, I compared the liveness composite score from the post-test questionnaire (detailed in the previous section) across all three conditions using a repeated measures ANOVA, and used a linear contrast to test the explicit prediction that that perceived liveness would be highest for Live, followed by Social and then Standard. To test Hypothesis 2, I examined the correlation between subjective liveness ratings for the three conditions (Live, Social, and Standard) and AQ scores. I also examined this relation controlling for scores on post-test questionnaire items measuring condition engagement and likeability, to ensure that any effects were driven by perceived liveness alone.

I also hypothesized that individuals would be faster to respond to questions in the Live conditions as compared to the two Not-Live conditions (Hypothesis 3), and that the magnitude of improved performance would be related to perceived liveness (Hypothesis 4). To test the effect of condition on performance, I used a repeated measures ANOVA to test a linear contrast for RT differences between all three conditions. Similar to

Hypothesis 1, this was based on the a priori hypothesis that Social performance will fall in-between Live and Standard performance. To test Hypothesis 4, I first calculated the degree to which the live condition affected performance by subtracting the average (cleaned) Live RT from the averaged Not-Live RT. More positive numbers indicated a larger time advantage for the Live condition. I next examined the relation between this variable and each participant's liveness rating for the Live condition using a Pearson's correlation.

Finally, for both Hypothesis 3 and 4, I did not have a hypothesis about the specific direction of the effects. Although there is evidence that belief computation can be faster than similar types of physical computations (Saxe, Shultz, & Jiang, 2006), such studies did not directly compare belief computations under different degrees of social saliency. It is possible that the live context could induce anxiety, superfluous thoughts, or more in-depth processing, which could slow performance. Thus, all behavioral hypotheses were tested using two-tailed tests with an alpha of .05 with SPSS 20.0.

fMRI preprocessing.

Image preprocessing was performed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm), and preprocessing steps were similar to another study that examined the neural correlates of live joint attention (Redcay et al., 2012). For the present study, the data were collected using an interleaved scheme, and were thus slice-time corrected, then realigned to the original volume from the first functional run, and then normalized to the Montreal Neurological Institute (MNI) template brain using both linear and non-linear transformations. Data were low-pass filtered (1/128 Hz) and spatially smoothed using a 5mm full-width half-maximum (fwhm) Gaussian kernel.

Outlying volumes (i.e., due to motion or global signal) for each participant were detected using the artifact detection toolbox (http://nitrc.org/projects /artifact_detect). A volume was identified as a motion outlier if the difference between two consecutive volumes exceeded 1mm (averaging across translational and rotational movements). The threshold for a global signal outlier was three SD from the mean global signal. These outlying time points were considered as individual regressors in subsequent analysis (see following section). Participants were to be excluded if, on two or more runs, their number of motion or global signal outliers exceeded 15% of collected volumes or if total motion exceeded 4mm. No runs, however, met these criteria, so all runs were included for all participants (*n*=29).

fMRI data analysis.

I conducted both first-level and second-level analyses. First-level analyses were conducted for individual participants; the resulting estimates of condition effects for each participant were analyzed at the group level in second-level analyses. For the first-level analyses, for each run, I compiled a data file with the onsets and durations for the events of interest (i.e., for all three conditions, the language processing Story period and the mentalizing Social Prediction period) and the events of no interest (i.e., the two seconds before the trial started and the feedback periods). Using SPM8, I convolved the stimulus duration and onsets with the canonical hemodynamic response function. Then I created a general linear model for each voxel across all four runs and all time points. This model estimated the fit of each subject's convolved blood-oxygen-level dependent (BOLD) signal with our model of when stimuli were presented. That is, in this model, all six events of interest were included as regressors (i.e., Live Story, Social Story, Standard

Story, Live Social Prediction, Social Social Prediction, Standard Social Prediction). This model also included, as regressors of no interest, the anticipatory periods and the reward periods. All six motion parameters (roll, pitch, yaw, x, y, and z) were included as regressors in this model, as was each individual outlier (such that the value for that outlier was 1, and all other values in the vector were zero), although these were not events of interest.

For the second-level analysis, individual participant results were analyzed at the group level for the following contrasts: Story (across all conditions) vs. baseline (fixation crosshair), Social Prediction (across all conditions) vs. baseline, and all six individual events of interest vs. baseline. I also analyzed specific contrasts to specifically examine the effect of live interaction on language processing (i.e., Live Story vs. Social Story, Live Story vs. Standard Story) and the effect of live interaction on mentalizing (i.e., Live Social Prediction vs. Social Social Prediction, Live Social Prediction vs. Standard Social Prediction). I also examined two contrasts to determine the effect of a recorded speaker's prosody and pitch on language processing (Social Story vs. Standard Story) and mentalizing (Social Social Prediction vs. Standard Social Prediction). For those contrasts, and the contrasts with baseline, I used a whole-brain random effects analysis, using SPM8 and in-house MATLAB scripts. To determine whether contrasts were significant, I used a two-tailed t-test. I first used a cutoff of $p < .001$ to threshold the contrasts maps. Then, to correct for multiple comparisons, I used SPM's false discovery rate (FDR) algorithm to determine which cluster size threshold that maintained an overall alpha of .05 (k=93). This procedure is consistent with other social neuroimaging paradigms.

Testing neural hypotheses.

Given the preliminary nature of this study, I employed whole-brain analyses, instead of limiting the analysis to specific regions of interest. Social brain regions are difficult to precisely anatomically define, and it is challenging to extrapolate specific activation coordinates from previous studies, as these previous studies have almost exclusively taken a third-person perspective. Thus, using anatomical or functional maps to restrict data analysis to particular regions could obfuscate important trends in the data. For this study, I created tables with all significant regions for all contrasts, in order to fully examine the effects of the experimental conditions, and I considered the full activation pattern and regions in tandem. With that caveat, I did have a set of candidate regions that I was interested in for each individual hypothesis, detailed below.

Language processing. For the Story portion of the experiment (i.e., the portion matched on linguistic content), I conducted three pairwise comparisons, such that each condition (Live, Standard, Social) was compared to each other condition. These comparisons contrasted listening to speech from a live social partner with listening to prerecorded speech, and also differed in terms of prosody between the Standard versus the Live and Social conditions. I hypothesized that the comparison between live and prerecorded speech would differentially activate regions associated with social engagement (i.e., dMPFC), and that contrasts between speakers with different prosody levels would differentially engage pitch and prosody regions (i.e, STG).

Specifically, for the comparison of the Live and Standard conditions, I predicted that the Live condition would more strongly engage regions of the brain associated with social engagement and differentially engage brain regions related to pitch and prosody

(Hypothesis 5; Table 2). I predicted that the Live and Social conditions would not differentially activate pitch and prosody regions, but that the Live condition would more strongly activate regions associated with social engagement (Hypothesis 6). Finally, for the comparison of the two Not-Live conditions, I hypothesized that there would be no differential activation in social engagement regions, but differential activity in auditory speech processing regions (Hypothesis 7). Additionally, given the matched syntactic content between conditions, for each comparison, I also predicted that there would be no difference between conditions in activation for regions related to sentence comprehension (i.e., left IFG; Hypotheses 5-7).

Mentalizing. For the Social Prediction part of the experiment, which consisted of the prompt, question, and choice segments, I again compared each condition to each other condition. These comparisons contrasted mentalizing about a character in a story versus mentalizing about a real-life social partner, and I predicted the mentalizing network (i.e., TPJ, dMPFC, PCC) would be more activated for the real-life partner. This network subsumes the social engagement region (i.e., dMPFC) from the language processing comparisons in Hypotheses 5-7. The predictions regarding mentalizing are consistent with past literature suggesting that dMPFC is engaged in broader person perception and social engagement, and that regions such as the TPJ are selectively activated by belief processing (e.g., Saxe & Powell, 2006). Given the strong role of executive function in mentalizing (Carlson, Moses, & Breton, 2002), combined with the potential differential attention to second-person versus third-person descriptors (e.g., “I” vs. “Megan”), I included specific hypotheses about the attention network for these contrasts.

Specifically, comparing the Live and Standard conditions, I predicted differential involvement of the pitch and prosody regions outlined previously, and increased activation during the Live condition in the mentalizing network (Hypothesis 8). For the comparison between the pitch-matched Live and Social conditions, I hypothesized equivalent activation in pitch and prosody regions, but increased activation for the Live condition in the socio-communicative network, specifically in regions implicated in mentalizing studies (Hypothesis 9). Finally, for the comparison between the two Not-Live conditions, I predicted equivalent activation in mentalizing regions, but increased activation for the higher pitched Social condition in pitch and prosody regions (Hypothesis 10). Given efforts to match linguistic and attentional demands, I did not predict differential activity in the syntactic processing regions detailed for the previous hypotheses, nor did I predict differential activation in the attentional network (Hypotheses 8-10).

Chapter 3: Results

Behavioral Results

Subjective impressions of perceived liveness.

A repeated measures ANOVA revealed a significant effect of condition (Live, Social, Standard) on subjective ratings of liveness ($F(2,56)=39.99$, $p<.001$), and the linear contrast was also significant ($F(1, 28)=65.23$, $p<.001$; Table 3).

Table 3. *Descriptive Statistics for the Three Experimental Conditions.*

	Live	Social	Standard	$F_{(2,56)}$	Pairwise Comparisons
BEHAVIORAL PERFORMANCE					
Accuracy (%)	95.66 (.06)	94.25 (.08)	94.20 (.08)	.395	Live=Soc=Std
RT (ms)	1088 (199)	1121 (188)	1133 (242)	.520	Live=Soc=Std
POST-TEST QUESTIONNAIRE RATINGS					
Liveness	5.76 (1.23)	3.99 (1.50)	2.61 (1.51)	39.99***	Live>Soc>Std
Likeability	6.45 (.69)	4.76 (1.52)	3.52 (1.27)	46.07***	Live>Soc>Std
Engagement	6.50 (.81)	5.60 (1.21)	4.76 (1.50)	27.98***	Live>Soc>Std

Note. Values are mean (standard deviation). Post-hoc pairwise comparisons were made using a Sidak correction for multiple comparisons at an alpha of .05. Untransformed RTs are reported in the table, but statistical tests were performed on inverse transformed RTs. All post-test questionnaire ratings are on a 1 to 7 scale. When assumptions of spherecity were violated, a Greenhouse-Geisser correction was employed. Abbreviations: RT=reaction time; ms=milliseconds. ***, $p<.001$.

Thus, as predicted in Hypothesis 1, liveness ratings were highest for the Live condition. Consistent with Hypothesis 2, AQ scores were significantly correlated with perceived liveness for the Live condition ($r(27)=-.40$, $p=.03$; Figure 3), but not the Social ($r(27)=-.27$, $p=.17$) nor Standard ($r(27)=.13$, $p=.49$) conditions, such that more autistic-like personality traits were related to less of a subjective sense of live interaction during the Live condition. The correlation between AQ scores and perceived liveness remained significant after controlling both for self-reported engagement during the live condition and for participant's individual ratings of the likeability of the live speaker ($r(25)=-.54$, $p=.003$).

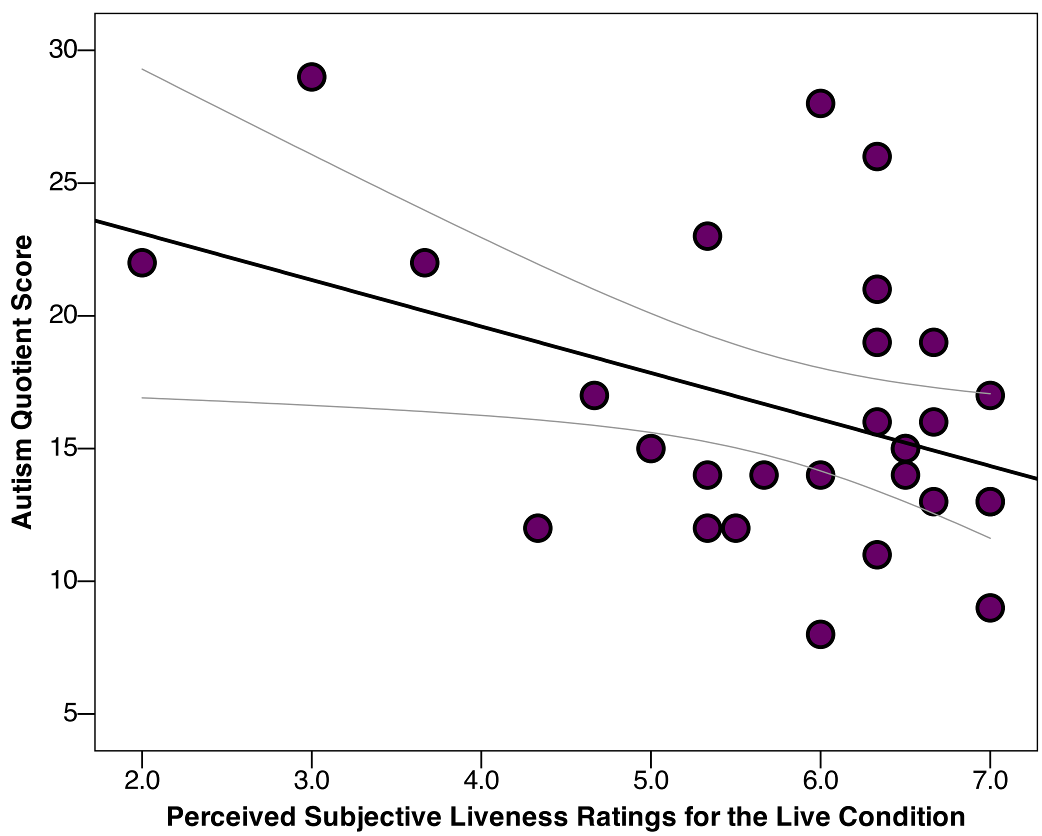

Figure 3. Correlation between participants' liveness ratings for the Live condition and autism quotient (AQ) scores. Higher AQ scores indicate more autistic-like personality traits. Participant liveness ratings were averaged across several questions that assessed how direct the interaction seemed and how much it seemed like someone was in the scanner talking to the participant versus sounding like a prerecording. Possible scores ranged from 1-7. The correlation was significant ($r=-.40$; $p<.05$).

Behavioral effect of live context on processing language and mental states.

In the test of the effect of condition on RT, a repeated measures ANOVA revealed no significant effect ($F(2,56)=.52$, $p=.60$), and the linear contrast was also not significant ($F(1,28)=.14$, $p=.71$). Thus, Hypothesis 3 was not supported; there was no effect of live interaction on behavioral performance when mentalizing about a live social partner versus a third-person character. Consistent with this lack of an effect of condition on behavioral performance, Hypothesis 4 was also not supported; differences in RT between the Live and Not-Live conditions were not correlated with the perceived liveness of the Live condition ($r(27)=.26$, $p=.18$). That is, the relative change in performance during the live interaction versus recorded segments was not related to a participant's subjective judgment of how live the interaction seemed.

Neuroimaging Results

The effect of a live context on the neural processing of language.

As hypothesized, processing speech from a live social partner modulated neural activity compared to processing content-matched recorded speech (Table 4). The specific pattern of results for the candidate regions was broadly consistent with Hypotheses 5-7: whole-brain comparisons of the Live condition to the two Not-Live conditions revealed activation in regions often associated with social processing and engagement, but the comparison of the two Not-Live conditions only implicated regions involved in pitch and prosody processing. Corrected whole-brain comparisons for the Live versus Social conditions revealed significant activation in dMPFC and left TPJ (Figure 4). The comparison between Live and Standard conditions implicated a cluster of significant

Table 4. *Regions Showing Differences Between Conditions for the Syntactically-Matched Story Portion.*

Region	Side	Peak *t*	Cluster *k*	MNI Coordinates *x*	*y*	*z*
EFFECT OF LIVE INTERACTION (MATCHING PITCH/PROSODY)						
Live>Social						
Angular gyrus	L	4.54	93	-34	-78	42
TPJ	L	5.09	382	-38	-56	32
dMPFC	L	5.08	108	-8	50	28
dMPFC	R	5.39	152	12	56	28
Social>Live						
Lingual gyrus	L	-5.56	482	-18	-94	-4
Lingual gyrus	R	-4.82	164	22	-94	4
EFFECT OF LIVE INTERACTION (MISMATCHED PITCH/PROSODY)						
Live>Standard						
mid-STS/pSTS	R	7.18	903	48	-30	2
Standard>Live						
Putamen	L	-5.45	146	-28	-4	6
Putamen	R	-5.25	180	32	-8	4
EFFECT OF PITCH/PROSODY (MATCHED ON NOT-LIVE CONTEXT)						
Live>Standard						
STG	R	5.72	184	64	-22	6
Standard>Live						
None						

Note. TPJ=temporal parietal junction; dMPFC=dorsal medial prefronal cortex; pSTS=posterior superior temporal sulcus; STG=superior temporal gyrus. All coordinates are in Montreal Neurologic Institute (MNI) space.

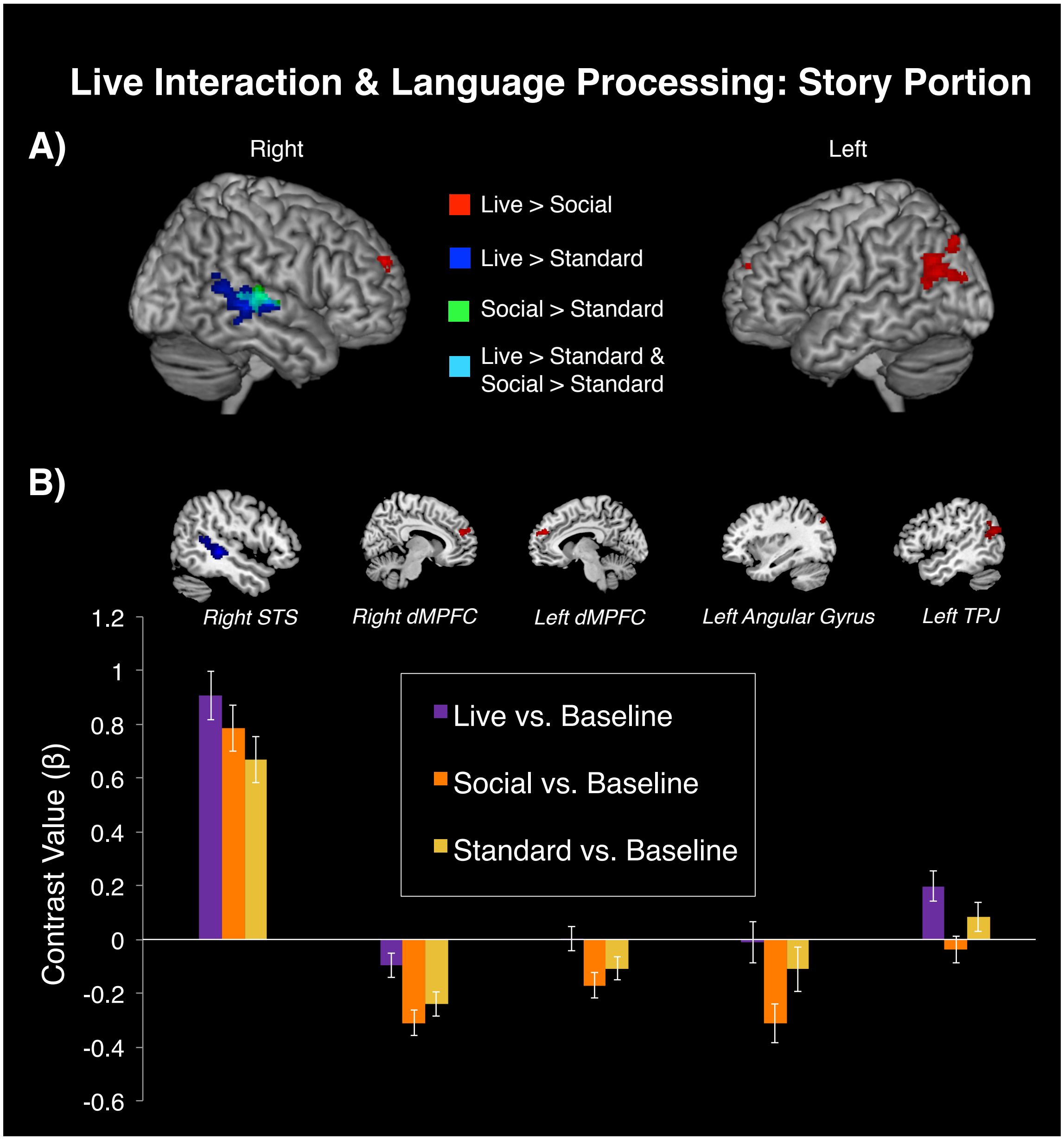

Figure 4. Effect of live context on the neural correlates of language processing. (A) The contrasts for the Story portion between all three conditions. This lateral image depicts false detection rate corrected ($p<.05$) activation up to 16 voxels deep. (B) Activation relative to baseline. Right STS is defined based on the Live>Standard contrast, and all other regions are defined based on the Live>Social contrast. Baseline was a fixation crosshair on a black screen. Bar graphs are intended to provide an illustration of activation versus baseline and are not presented for statistical purposes. Abbreviations: STS= superior temporal sulcus; dMPFC=dorsal medial prefronal cortex; TPJ= temporal parietal junction.

activation in temporal cortex, encompassing mid-STS and extending toward pSTS. Surprisingly, dMPFC was not more active in the Live than Standard condition, although this activation did emerge at a lower threshold (p<.001, k>25; [12 58 24], t=4.99). The comparison of the two Not-Live conditions, Social versus Standard, revealed activation only in STG.

The effect of a live context on the neural correlates of mentalizing.

Again, as hypothesized, mentalizing about a live social partner modulated neural activity as compared to an identical mental state computation about a third-party character (Table 5). Broadly consistent with Hypotheses 8-10, whole-brain comparisons of the Live versus Not-Live conditions revealed increased activity in brain regions frequently implicated in social engagement and goal and intention processing. Again, these regions were not active when comparing the two Not-Live conditions, which varied on speaker pitch, prosody, and perceived likability, but not liveness. Specifically, the comparison between Live and Social Conditions revealed activity in bilateral pSTS and right superior posterior frontal gyrus (pSFG; Figure 5), and there was some overlap with the left pSTS activation from the Story portion of the experiment (Figure 6a). The whole-brain comparisons for the Live versus Standard conditions revealed increased activation in dMPFC, bilateral pSTS, right mid-STS, and right pSFG, but not TPJ. The contrast between the Social and Standard conditions only revealed increased bilateral STG activation, and even a more liberal threshold (p>.001, uncorrected), did not indicate activation differences outside of regions associated with auditory processing (Figure 6b).

Table 5. *Regions Showing Differences between Conditions for the Social Prediction Portion.*

Region	Side	Peak *t*	Cluster *k*	MNI Coordinates *x*	*y*	*z*
EFFECT OF LIVE INTERACTION (MATCHING PITCH/PROSODY)						
Live>Social						
pSTS	L	7.91	346	-52	-58	16
pSTS	R	7.23	536	48	-66	12
pSFG	R	5.18	106	12	10	70
Social>Live						
STG	L	-6.04	566	-60	-8	14
STG	R	-5.34	174	56	-10	6
EFFECT OF LIVE INTERACTION (MISMATCHED PITCH/PROSODY)						
Live>Standard						
pSTS	L	8.20	424	-46	-66	10
pSTS	R	8.19	872	44	-64	10
dMPFC	R	5.18	233	6	50	26
pSFG	R	5.51	148	8	14	64
STG	R	5.77	467	58	-26	-10
Standard>Live						
None						
EFFECT OF PITCH/PROSODY (MATCHED ON NOT-LIVE CONTEXT)						
Live>Standard						
STG	L	5.67	332	-60	-20	-4
STG	R	5.86	114	60	-12	-10
Standard>Live						
None						

Note. pSTS=posterior superior temporal sulcus; pSFG=posterior superior frontal gyrus; STG=superior temporal gyrus; dMPFC=dorsal medial prefronal cortex. All coordinates are in Montreal Neurologic Institute (MNI) space.

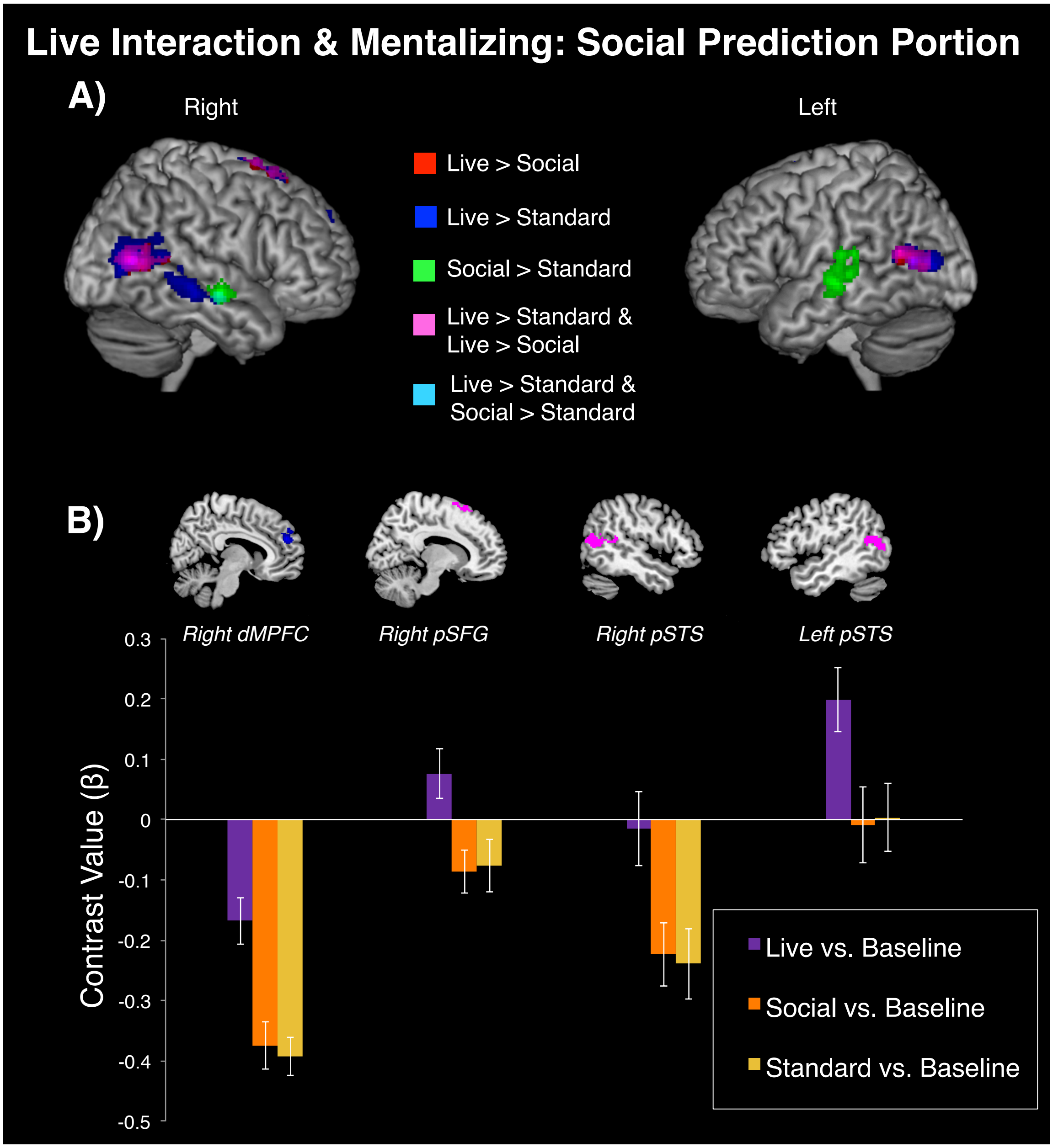

Figure 5. Effect of live context on the neural correlates of mentalizing (A) The contrasts for the Social Prediction portion between all three conditions. This lateral image depicts false detection rate corrected (*p*<.05) activation up to 16 voxels deep. (B) Activation relative to the baseline fixation cross. Right dMPFC is defined from the Live>Standard contrast, and right pSFG and bilateral pSTS are defined based on overlap between Live>Social and Live>Standard. Bar graphs are intended to provide an illustration of activation versus baseline and are not presented for statistical purposes. Abbreviations: dMPFC=dorsal medial prefrontal cortex; pSFG=posterior superior frontal gyrus; pSTS= posterior superior temporal sulcus.

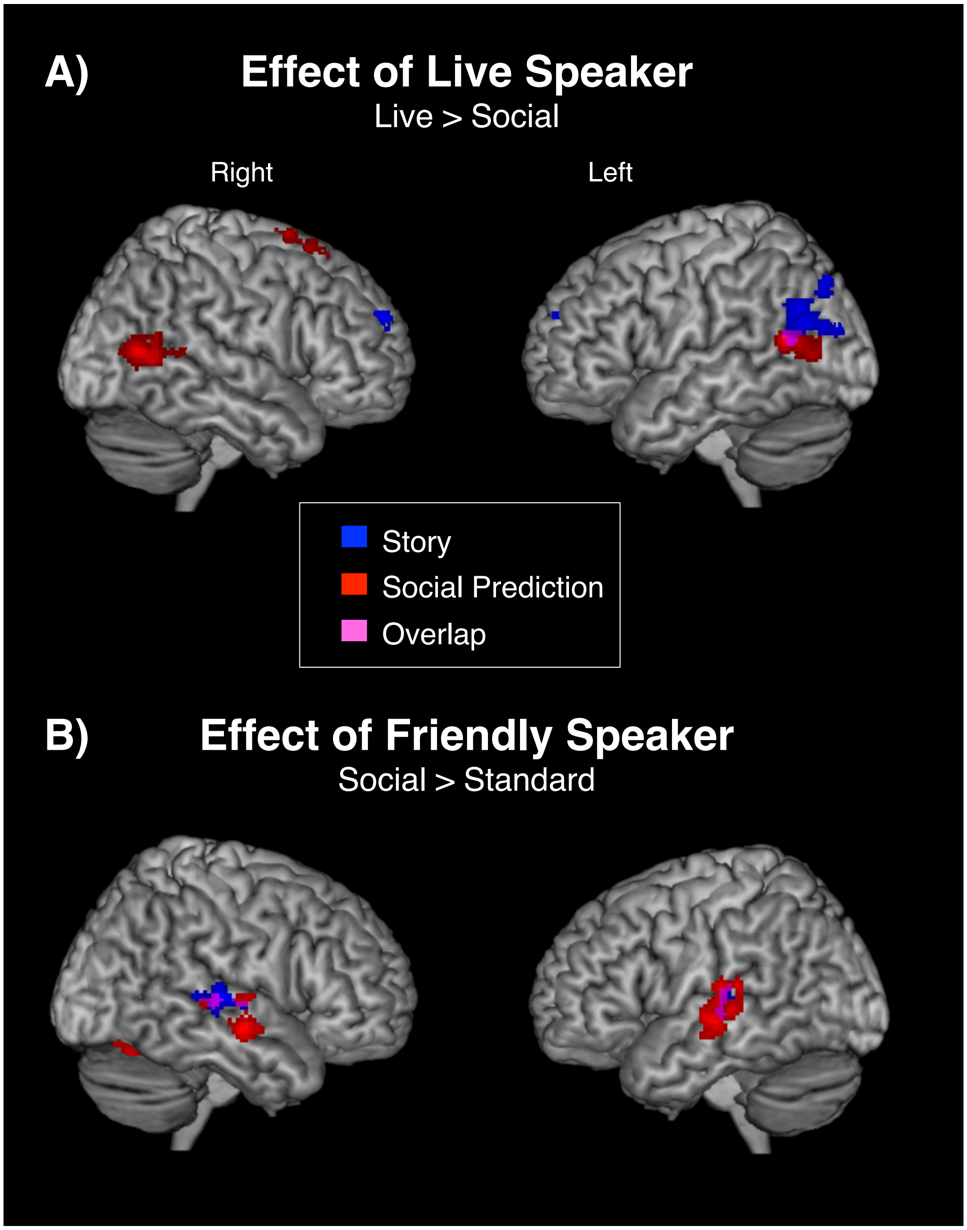

Figure 6. The effect of speaker across both the Story and Social Prediction portions. (A) The false discovery rate corrected (p<.05) effect of the Live versus Social speaker (matched on pitch and prosody). Overlap between language processing (Story) and mentalizing (Social Prediction) experimental portions was found in left posterior superior temporal sulcus. (B) The effect of a high-pitched recorded speaker with high prosody (Social) versus a lower-pitched, more monotone speaker (Standard). The higher-pitched speaker was subjectively judged as more friendly by participants. The figure depicts uncorrected (p<.001, k>25) activation for the Social>Standard speaker contrast. Activation was limited to secondary auditory regions in bilateral superior temporal gyrus.

Chapter 4: Discussion

The current study used a novel fMRI paradigm to determine if the neural correlates of listening to and thinking about another person were altered when typical adults believed they were part of a live interaction, as compared to listening to a recording. This novel paradigm was successful in convincing participants they were part of a live interaction, and the lower a participant's level of autistic-like personality traits, the more he or she believed they were part of a live interaction. As hypothesized, the fMRI data indicated that, although the social cognitive demands were computationally identical between recorded and live conditions, the neural correlates of language processing and mentalizing were significantly altered by the live context, specifically in several regions associated with socio-communicative processing.

Interpreting Behavioral Results

Consistent with my hypotheses, on a post-test questionnaire, participants' subjective ratings of perceived liveness were higher in the Live than in the Social and Standard conditions. For the Live condition, but not the other two conditions, perceived liveness was significantly correlated with AQ scores, even after controlling for other subjective ratings including the Live speaker's likeability. Past research with typical populations has linked AQ scores to personality traits, such as extraversion (Austin, 2005), and social cognitive competence (e.g., detecting mental states based on photographs of the eyes; Voracek & Dressler, 2006). In the current study, perhaps individuals with higher levels of autistic-like personality traits were less sensitive to the cues signaling a live interaction, or the experience of interacting with a person was less salient. The link between sensitivity to live interaction and autistic-like personality traits is also consistent with research that suggests that live or real-time interactions pose a greater difficulty for individuals with autism than offline social cognitive computations (e.g., Klin, 2000; Klin et al., 2003).

Participant ratings of liveness, likeability, and engagement were highest for the Live speaker, followed by the Social, and then Standard speaker. The difference between the Live and Social speaker on measures of likeability and engagement is somewhat surprising, given efforts to match speaker friendliness and attention during each condition. However, given the centrality of live interaction to human social experiences, it may not be possible to match live and recorded speakers on these dimensions, at least without making that live speaker aversive. That is, by virtue of being live, a speaker appears more likeable. Future studies could have naïve participants directly compare the recordings of the Live and Not-Live speakers, to determine if the Live speaker's greater likeability remains outside of a live context.

Although participants claimed to be more attentive and motivated in the Live condition, there were no behavioral differences in performance on the Social Prediction portion of the task. This null effect was not predicted, but may have been due to the ease of the task, as participants were near ceiling for the questions in the current task. Future studies employing a live paradigm could include more difficult tasks to rigorously examine the effect of a live interaction on behavior (e.g., Auvray et al., 2009), or could employ implicit tasks that produce continuous performance measures (e.g., Senju et al., 2009). Such tasks may also uncover individual differences in such an effect, and future work could investigate whether any individual differences may also relate to scores on measures like the AQ.

Interpreting Neuroimaging Results

Language processing.

As hypothesized, during the matched Story portion of the experiment, bilateral dMPFC, a region linked to social engagement (Van Overwalle, 2011), was more active for the Live than Social speaker. This increased activation may be a marker of the social salience of the situation, or the perceived increased demand of social processing during a live interaction. Interestingly, during the Story portion, no explicit social demand was placed on the participant. Although contingency is an important part of live interaction, the Story segment ostensibly represents entirely passive listening. Thus, it is noteworthy that, even during this portion, the impression of listening to a live speaker was sufficient to alter neural activity.

The left TPJ, as part of the mentalizing network, was not hypothesized to be activated in the contrasts for the Story portion of the experiment, although it was hypothesized to be involved in the Social Prediction portion. However, left TPJ was significantly more active in the Live than Social conditions during the Story. Although TPJ is commonly implicated in mentalizing tasks, particularly those that are story-based (e.g., false belief stories; Saxe, 2009), a recent meta-analysis found that left TPJ was an area of overlapping activation among studies of story-based mentalizing, non-story based mentalizing (e.g., cartoon pictures), and general narrative comprehension (Mar, 2011). Thus, for this current study, several possible explanations for TPJ activation emerge, highlighting the difficulty of inferring whether changes in activation specifically implicate language processing or mentalizing. One possibility is that listening to a live speaker alters the narrative processing of two identical stories. Another possibility is that, although the Story portion did not contain any explicit mental states, participants were spontaneously representing the live social partner's potential beliefs and goals. Both processes may also be operating in tandem: narrative processing may be altered due to participant's enrichment of the narrative by consideration of the live partner's mental states. Although the current paradigm is unable to dissociate these possibilities, these findings provide the first evidence that a live context alone alters the neural correlates of language processing.

In the Live versus Standard comparison, a wide swath of right STS was more active, extending from more mid to posterior regions. The peak coordinates of this STS activation have been implicated in pitch and prosody perception, which does vary between the Live and Standard conditions, but the more posterior activation is spatially closer to regions involved in intention and animacy perception, especially given its rightward lateralization (Gao, Scholl, & McCarthy, 2012; Pelphrey, Morris, & McCarthy, 2004; Shultz, Lee, Pelphrey, & McCarthy, 2011). Thus, as suggested by the role of pSTS in processing a variety of communicative behaviors, this region's patterns of activation appear to be modulated by a live context.

As hypothesized, the contrast between the Social and Standard conditions only revealed increased activation in STG, which has been consistently implicated in studies of pitch and prosody (reviewed in Price, 2010). Given that the Social speaker was rated as significantly friendlier and more engaging than the Standard speaker, the lack of an effect outside of secondary auditory regions seems to imply that brain is not broadly sensitive to the effects of friendliness or sociability, at least in a recorded context. This comparison between the Social and Standard conditions is especially important, given that the goal of the present study was to isolate the effect of liveness, and the Live condition had significantly higher ratings of engagement and likeability than either Not-Live condition. The Social condition, however, was also rated as significantly more engaging than likable than the Standard condition. Thus, if neural activation outside of auditory regions had appeared similar in both the Live versus Standard and Standard versus Social contrasts, it would have indicated to a general effect of attention or likeability. Such an effect, however, was not present, suggesting the paradigm was successful in isolating the effect of liveness.

Mentalizing.

As hypothesized, the neural correlates of making predictions about a social partner were significantly different in the Live versus Not-Live conditions. Contrary to hypotheses, however, the region most frequently implicated in other studies of mental state representation (i.e., right TPJ; Saxe, 2009) was not significantly more active during the Live condition than either recorded condition. Rather, bilateral pSTS activation was significantly greater in the Live condition than either Not-Live condition. The specific comparison between the Live and Standard conditions also revealed, for the Live condition, increased activation in regions linked to social engagement (i.e., dMPFC) and auditory (i.e., STG) processing. The lack of differential TPJ activity during the Social Prediction portion, however, may not be surprising considering the content of the mental representations in the current paradigm. The TPJ is most frequently implicated in studies of belief representations (e.g., false belief; Koster-Hale & Saxe, 2013), but the current paradigm involves communicated preferences and goals (e.g., “I like watching team sports”; “I am trying to eat healthy”), which may be more related to the pSTS’s role in goal and intention processing (e.g., Vander Wyk et al., 2009). Past studies have dissociated the role of the pSTS and TPJ in mentalizing (Gobbini, Koralek, Bryan, Montgomery, & Haxby, 2007), and, intriguingly, the specific pattern of pSTS activation in this study resembles that found in studies of biological motion processing (Ethofer, Gschwind, & Veilleumier, 2011; Grossmann et al., 2000; Thompson, Hardee, Panayiotou, Crewther, & Puce, 2007). One possible explanation is that participants are mentally visualizing the live, but not recorded, social partners. Another possibility is that although this region is also sensitive to biological motion, that sensitivity is indicative of a broader sensitivity to processing complex social stimuli (e.g., Redcay, 2008). Although future studies could further isolate the role of this specific region, importantly, this is the first well-controlled study illustrating that pSTS is sensitive to a live, interactive context, even without visual stimuli depicting actions.

Contrary to specific hypotheses, both contrasts of Live versus Not-Live conditions implicated the same superior frontal region. This region has been linked to a diverse set of motoric and attentional functions, including both action observation and imitation (Caspers, Zilles, Laird, & Eickhoff, 2010). There was no reaction time difference between conditions, however, indicating that differences the motor behavior of pressing the response button are unlikely to explain this activation difference. Further, in an uncorrected comparison ($p<.001$, $k=24$), a similar region ([12 22 62]; $t=4.05$) was active in the matched Live versus Social *Story* portion, which did not contain a motor response. Superior frontal regions have also been linked to autobiographical processing (Fink et al. 1996; Lee et al., 2002; Ryan et al., 2001), which may provide another explanation for increased activity. Future analyses may be able to dissociate some of these possibilities by considering the time course of activation during the Social Prediction period. If activation in this region peaks toward the end of the period, when participants are making a motor response, it may indicate a motoric explanation, whereas earlier activation, when hearing about a person's goals and preferences, may indicate that this region is modulated by socio-communicative context.

Interpreting patterns of activation and deactivation.

Interpretation of the current results is complicated by the existence of the default mode network (DMN), an interconnected set of regions that become less active when a participant is completing a task (Raichle et al., 2001). The regions involved in this network overlap with many of the same regions consistently identified in social tasks (Buckner & Carroll, 2007; Schilbach et al., 2012; Spreng, Mar, & Kim, 2009), and some have theorized that the DMN is implicated in self-reflective processes (e.g., awareness, recollection; Andrews-Hanna, Reidler, Hunag, & Buckner, 2010; Buckner et al., 2008; Spreng & Grady, 2010). The high levels of DMN activity at rest often create difficulties when interpreting condition-specific changes relative to baseline (e.g., a fixation cross). For example, a study of the effect of person versus object knowledge found that much of the increased activation in social brain regions when considering people versus objects was driven by relatively less deactivation when thinking about people compared to baseline (Mitchell, Heatherton, & Macrae, 2002).

Findings relating to the DMN may explain some results in the current study. For example, although the Live condition significantly activated dMPFC compared to the Social condition, baseline activation in that region was higher than for either experimental condition. Such findings may account for a relative lack of a difference between the Live and Standard condition during the Story portion. Participants reported being significantly less engaged by the Standard than the Social condition. Increased mind wandering or decreased attention during the Standard condition could result in relatively more dMPFC or TPJ activity in the Standard than Social condition, attenuating the differences between the Standard and Live conditions. With the current paradigm, this possibility is difficult to disentangle, but is lent some support by the significant differences in pSTS activity between the Live and Standard conditions, as this region is not consistently implicated in the DMN. Interestingly, for certain regions the Live condition resulted in deactivation compared to baseline. This decrease may be due to the highly constrained nature of the interaction in the present experiment, which perhaps limited the full engagement of self-reflective monitoring systems. It is possible that real-world, unscripted social interactions more closely resemble spontaneous cognition during baseline conditions and would thus not show this pattern of deactivation.

Comparison of Live vs. Social and Story vs. Social Prediction.

In terms of comparing the two Not-Live conditions, given the above evidence, and given that the Live and Social conditions are matched on pitch and prosody, the Live versus Social comparison likely provides a most precise test of the effect of a live interaction than Live versus Standard contrasts. In terms of comparing the mentalizing and language processing results, although a direct comparison of the Story and Social Prediction portions is difficult, general findings seem to indicate that the Story portion more strongly engages regions involved in mentalizing and narrative comprehension (i.e., TPJ) and representing a social partner (i.e., dMPFC), and the Social Prediction portion more strongly engages regions associated with representing goals and intentions (i.e., pSTS). The Story portion also appears to be more left lateralized, compared to bilateral activation in the Social Prediction portion. This may be consistent with research suggesting that left TPJ is more commonly engaged than right TPJ in non-story based mentalizing and narrative comprehension (Mar, 2011) and that left TPJ is selectively activated by processing communicative intentions (Ciaramidaro et al., 2007). There was also some overlap in left pSTS activation between both Story and Social Prediction segments, which may indicate a common representation of a social partner, regardless of the specific task.

One potential criticism of the current study's significance is that the regions identified as sensitive to live interaction have already been implicated in offline approaches to social neuroscience, limiting the insights of a second-person approach. This framing, however, overlooks the goal of second-person neuroscience: to reconceptualize the brain regions typically studied and discussed in cognitive computational framework. That is, conventional framing of the role of TPJ or pSTS often involves a focus on the computational load shouldered by these region—for example, parsing the representations that make up a belief state or considering how a pattern of actions depicts an intention. The current study illustrates that these regions are sensitive to a live context across invariant social cognitive computational demands. For example, when holding narrative content constant, activity in mentalizing and narrative processing regions like pSTS/TPJ is modulated by listening to words spoken by a live social partner.

Consistent with the present findings, second-person neuroscience conceptualizes social brain regions based on their role in social interaction. As Schilbach and colleagues argue (2013), it is possible that these regions' role in social interaction is ontogenetically prior to their role in explicit cognitive processing (e.g., the role of TPJ in explicit theory of mind). That is, rather than specific regions serving specific cognitive processes, and these processes coming together to support social interaction, these brain regions first support real-world social engagement. Such a perspective aligns with evidence from atypical development that real-world mental state understanding implicitly emerges from dynamic social interaction (Klin et al., 2003) and evidence from infancy that regions later involved in more explicit mentalizing initially subserve social interaction (e.g., joint attention; Grossmann & Johnson, 2010). Thus, although individual paradigms devoid of social interaction may find activation in similar social brain regions to the present study, such activation may simplify or even misrepresent the more fundamental role of the neural networks in supporting real-world social engagement.

Future data analysis.

Although, for the purposes of the current study, analyses were restricted to the proposed hypotheses, this live interaction fMRI data set is rich with possibilities for future analysis. For example, participant ratings of liveness and likeability could be added to the fMRI analyses as covariates, to attempt to distinguish response patterns that are most specific to perceived liveness. Changes in neural responses over the course of the experiment could also be examined, as perhaps some of the effect of liveness begins to dissipate due to the lack of full reciprocity of interaction in the Live condition. Examining activation during the feedback portion of the current paradigm would also be informative, given that the literature of social reward has almost exclusively focused on static photographs of the faces of strangers (e.g., Delmonte et al., 2012; Izuma, Saito, &

Sadato, 2008; Izuma, Saito, & Sadato, 2010; Lin, Adolphs, & Rangel, 2012; Zeeland et al., 2010).

The analytic approach for the current study was to examine activation in the whole brain, but other approaches could also be informative. A subset (*n*=23) of participants completed a theory of mind localizer task in the scanner. This task was unrelated to the live paradigm, and involved reading stories about false beliefs and stories about false physical representations (e.g., a photograph; Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011). Future work could consider, for individual participants, the regions activated by this localizer, and define these as regions of interest (ROIs) for the live paradigm. Examining changes for Live versus Not-Live speech in these specific ROIs would increase power and increase the ability to conclude effects are specific to regions that individual participants use while mentalizing. Other analyses could examine measures of functional connectivity, as the brain activates as a unit, and changing connectivity could be a more fine-grained measure of sensitivity to live contexts (e.g., Rissman, Gazzaley, & D'Esposito, 2004). Item-level analysis could also compare the response to mentalizing items that involve an explicit goal (e.g., "I am trying"), those that involve a desire (e.g., "I like"), and those that do not involve an explicit mental state verb (e.g., "I own").

Limitations and Future Directions

In addition to future investigations with the collected data set, some methodological limitations of current study's design could guide future work. First, the Social Prediction portion concatenates together several different events (e.g., hearing a preference, making a choice, a motor response), and the current design cannot dissociate the effects of each portion. Further, this portion is not linguistically matched between the Live and Not-Live conditions (e.g., “I want” vs. “Megan wants”). A future study could put jitter time between these events, and could also include a control where the recorded portion includes first-person sentences (e.g., “I want”). It is possible that simply hearing first-person language has similar neural correlates to a more complete live interaction.

Another limitation with the current design is that participants briefly met the live social partner before the experiment when she explained the video set up. Although this interaction is minimal, it makes it impossible to fully dissociate any effect of speaker familiarity. Future experiments could limit interaction with the experimenter before the fMRI scan begins. Alternatively, future studies could systematically vary the quality of the interaction with the experimenter before the study, such that sometimes she follows a warm and engaging script, and sometimes she follows a script in which she is cold and brusque. This experimental manipulation could also help isolate the effects of liveness and likeability.

Finally, although participants rated their experiences with the manipulation, and completed the AQ, they did not complete more implicit, ecologically-valid measures of real-world social functioning. Future research could better connect the results of this live paradigm with real-world interactions by extending this paradigm to individuals with autism or other social disabilities. There are two potential explanations for the unique difficulty of individuals with autism in live social interaction, as opposed to laboratory-based assessments (e.g., Klin et al., 2003): first, a lack of motivation for live interaction, resulting in neural hypoactivation (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012), and second, aversion and overstimulation to live interaction, resulting in hyperactivation to live stimuli (Dalton et al., 2005; Tanaka & Sung, in press). Extending this paradigm to individuals with autism could help disambiguate these two possibilities.

Relevance to Development

Implications of the current project.

As mentioned in the introduction, the ultimate goal of this project is to chart the development of the social brain from a second-person neuroscience perspective. The present study laid the groundwork for this goal in two important ways: first, it tested and validated this novel live paradigm; and second, the current results characterized the developmental endpoints of second-person neural processing. This characterization was a necessary step to contextualize and interpret data from children.

Practical concerns. The paradigm was previously untested, making it important to establish the workability of the paradigm with an adult sample. If children had been unable to complete the paradigm, failed to believe the live illusion, or provided data that was difficult to interpret, it would have been unclear if this was due to something intrinsic to the design, or rather due to the shorter attention spans of children, or their increased movement while in the scanner (Van Dijk et al., 2012). Further, our behavioral piloting indicated that children often reported that they did not have any suspicions about the live set-up, but when debriefed, claim they "knew it all along." By beginning with adults, we were able to use extensive post-testing to determine that 94% of participants believed they were in a live interaction. Further, evidence from the current study suggests that the comparison of the Live versus Social conditions best captures the effect of a live interaction on neural processing. Thus, the future studies with children can eliminate the Standard condition, which will boost power by increasing the number of trials in the other two conditions. Of course, there may still be difficulties with extending the paradigm to children, but the current study provides the foundation for developmental studies.

Characterizing the developmental trajectory. Knowledge of developmental endpoints is crucial to a developmental approach (e.g., Apperly et al., 2009). In fields as varied as approximate number understanding and implicit memory, developmental scientists have sampled adults in order to map developmental trajectories (e.g., Cycowicz, Feldman, Snodgrass & Rothstein, 2000; Farrell & Barrett, 2006; Halberda & Feigenson, 2008). For example, there are several developmental neural phenomena that show U-shaped developmental patterns. Adolescents, as compared to children and adults, show heightened nucleus accumbens activation compared to orbitofrontal activation in response to rewards (Galvan et al., 2006). In a similar vein, there is evidence that frontal activation to fearful faces increases between ages 8 and 15 (Yurgelun-Todd & Killgore, 2006), but diminishes again in adulthood (Monk et al., 2003). In both examples, studies of developmental populations without context from adults would lead to spurious conclusions; researchers could infer that reward sensitivity or activation to fearful faces increases throughout development. Given such examples, and given that second-person neuroscience is almost completely unexplored, the current study helps to anchor and contextualize future developmental findings.

Specifically, the current study provides evidence that there is specialized processing of live stimuli, and that this specialization is present in brain regions (e.g., pSTS, TPJ) that show protracted structural (e.g., Mills et al., 2014) and functional (Carter & Pelphrey, 2006) development. For example, work by Carter and Pelphrey (2006) suggests that between the ages of 7 and 11, the pSTS specializes for biological motion. Given that the pSTS activation in the current experiment is spatially similar to activation during biological motion perception, the results from the current study yield a clear hypothesis that, in late childhood, the processing of live interaction will specialize. Similarly, TPJ, another region activated by live interaction in the current study, shows specialization between ages 8 and 12, and this specialization correlates with improved theory of mind behavioral performance (Gweon, Dodell-Feder, Bedny, & Saxe, 2012; Saxe, Whitfield-Gabrieli, Scholz, & Pelphrey, 2009). In contrast, dMPFC, another region implicated in the current study, does not seem to show this same pattern of functional specialization for belief states in this age range. Thus, an a priori hypothesis for dMPFC may be that it is already selectively actives to live conditions by middle childhood. The current study makes an important contribution by characterizing the neural correlates of live interaction in adulthood, allowing for clearer interpretation of results from future studies with developmental populations.

Future developmental studies.

The next step in this research program is to extend this paradigm to late childhood (roughly ages 8-12), which is a time of increased functional specialization in social processing regions and great changes in social behavior. During this age range, children begin to spend less time with their parents and more time with peers (Lam, McHale, & Crouter, 2012), and their social interactions with peers become increasingly complex (Farmer et al., in press; Feiring & Lewis, 1991). Socio-emotional understanding increases (Carr, 2011), and children also begin to take more risks (e.g., Chilcoat & Anthony, 1996; Jackson, Henriksen, Dickinson, Messer, & Robertson, 1998). In addition to these real-world changes, children improve in laboratory social cognitive tasks, including measures of mentalizing (Dumontheil, Apperly, & Blakemore, 2010) and sarcasm and irony comprehension (Capelli, Nakagawa, & Madden, 1990; Pexman & Glenwright, 2007). Overall, late childhood represents a time when individual variability in social competence widens (Monahan & Steinberg, 2011).

This individual variability in social competence and experience in late childhood lays a foundation for adolescent social functioning, which can predict disorders such as anxiety and depression (Bress, Smith, Foti, Klein, & Hajack, 2012; Carr, 2011; Eccles, Roeser, Wigfield, & Freedman-Doan, 1999; Meyer, Weinberg, Klein, & Hajack, 2012). Late childhood also represents a time when the social skills gap widens between children with autism spectrum disorders and their typical peers, as the adaptive skills of children with autism, but not necessarily their offline cognitive skills, begin to plateau (Anderson, Maye & Lord, 2011; Bal et al., 2013; Kanne et al., 2011). Despite the central role of interactive social experience in late childhood, social brain development in this age range has been almost exclusively studied in a third-person context. Given this evidence, extending this paradigm to developmental populations, and eventually to atypical developmental populations, is an important future step.

Conclusion

Although social cognitive research has suggested that a real-time, interactive context can alter behavior, most social neuroscience studies have used "human" as a proxy for social, studying the processing of speech by using recordings and studying the processing of mental states with stories about characters. The current study's novel and well-controlled paradigm provides evidence that a live social context alone is sufficient to change the brain’s response when processing another’s speech or mental states. This finding has implications for studying and quantifying the neural underpinnings of social cognitive processes in both typical and atypical populations.

Appendix A

Institutional Review Board Approval Letter

Date: May 17, 2013
To: Elizabeth Redcay
From: University of Maryland College Park (UMCP) IRB

Project Title: [344916-11] Functional brain organization in typical development
Reference #: 11-0415
Submission Type: Continuing Review/Progress Report

Action: APPROVED
Approval Date: May 17, 2013
Expiration Date: June 25, 2014
Review Type: Expedited Review
Review Category: Expedited review category # 4, 6, and 7

Thank you for your submission of Continuing Review/Progress Report materials for this project. The University of Maryland College Park (UMCP) IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the project and insurance of the participant understanding followed by a signed consent form. Informed consent must continue throughout the project via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this committee prior to initiation. Please use the appropriate revision forms for this procedure which are found on the IRBNet Forms and Templates Page.

All UNANTICIPATED PROBLEMS involving risks to subjects or others (UPIRSOs) and SERIOUS and UNEXPECTED events must be reported promptly to this office. Please use the appropriate reporting forms for this procedure. All FDA and sponsor reporting requirements should also be followed.

All NON-COMPLIANCE issues or COMPLAINTS regarding this project must be reported promptly to this office.

This project has been determined to be a Minimal Risk project. Based on the risk, this project requires continuing review by this committee on an annual basis. Please use the appropriate forms for this procedure. Your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date of June 25, 2014.

Please note that all research records must be retained for a minimum of three years after the completion of this project.

If you have any questions, please contact the IRB Office at 301-405-4212 or irb@umd.edu. Please include your project title and reference number in all correspondence with this committee.

Debriefing Sheet

The purpose of the present study was to determine how individuals respond to live versus recorded speech and how individuals respond to live social, non-live social and non-social rewards. However, due to the need to have each participant experience exactly the same live stimuli, the stimuli that you were told were live were actually prerecorded. The results from this study are going to be used to help us better understand human social development.

There were varying degrees of difficulty, so if you thought that some of the trials were easier or harder, you're not alone.

Thanks again for your participation in this study! If you have any questions, concerns, comments, or would like to receive a copy of the report summarizes our findings, please leave your contact information with the experimenter or contact either of the following researchers:

Kate Rice
krice@umd.edu
0112C BPS
(301)405-6712

Elizabeth Redcay
redcay@umd.edu
2147D BPS
(301)405-2884

For more information regarding research in this area, you can go to these references:

De Jaegher, H., et al. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Science,* 14, 441-447.

Izuma, K., Saito, D., & Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience,* 22, 621-631.

Risko, E., et al. (2011). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, 6, May, 1-11.

If you have questions regarding your participation in research, please contact:
University of Maryland Institutional Review Board.
1204 Marie Mount
College Park, Maryland 20742
(301) 405-4212
FAX: (301) 314-1475
irb@umd.edu

Appendix B

Script for fMRI Session

Kate: Welcome to the Live Language study. We are interested in how the brain processes information differently when I'm talking to you directly versus when the computer is playing prerecorded audio. To measure your brain responses, you'll be in the MRI scanner. In this setup, you're laying down in a long tube, with a computer screen directly over your face. This makes it difficult to have regular, face-to-face conversation! So we've constructed a set-up to come close to real-world interaction.

This study has two conditions. In the recorded condition, the computer will play audio clips of short stories and questions, and then two answer choices will appear on the screen. You'll press a button to pick the right answer. The computer will then give you visual feedback about whether your answer was right or wrong, like a picture of a smiling face or a gold star, or a frowning face and a red x. No one else sees your answers, just the computer program.

For the live condition, we've set up a link between the computer in this room, and the computer screen that you'll be looking at in the scanner. Our goal is to mimic the basic characteristics of the recorded stuff, so that the only thing that's different is the fact that it's live—I'll be talking directly to you in real time even though you won't see my face. We can't just have a normal back and forth conversation, since that would be too different from the prerecorded condition.

So I'm going to ask you questions in the same audio-only format as the computer's prerecorded questions. Basically, the computer screen will just tell you whether it's live or recorded. And on my computer in here, I'll see a green light when it's time to ask you a question.

Because the rewards from the computer don't have any sound, I'm just going to give you silent feedback. After I get a cue from my computer to ask you a question, I'll see your button feedback in a small part of my display. So I'll see if you got it right or wrong. And then I'll nod or give a thumbs up or whatever else I can think of when you get something right and I'll shake my head or give a thumbs down or make a sad face when you get something wrong.

We don't have a two-way audio or video hookup between this computer and you in the scanner, so I won't be able to see your face or hear what you say. There will be a scanner operator who will talk to you over the scanner microphone, to remind you to stay still, etc., but since I'm not in the official scanner room, I'm not allowed to give safety tips or monitor how you're doing in the scanner. That's what the operator will be doing.

The questions today should be pretty easy, since eventually we'd like to extend this project to kids. We still want you to try your best on all the questions, whether I'm the one asking or not. You will only have a few seconds to respond to each question, or your answer will be marked wrong.

At the end of this main part, you'll read some short stories and answer questions about them. That should take about 10 minutes. And then you'll be all done.
Do you have any questions for me?

So I'm going to get set-up in my control room and [Operator name] is going to get you set up in the scanner.

Kate walks the participant into the control room and goes back to the side room. Operator turns on the volume on the computer in the control room and Kate turns on the volume on her computer. The gchat window should already be open.

She then takes the participant into the scan room and gets them all set up. Kate is watching the response box via webcam.

Kate: Okay [subject name], I hope you're okay in there. Let's test out if I'm getting your key presses over at this computer. Press 1 if it's a weekday and 2 if it's a weekend.

Participant presses the key and sees either the 1 or the 2 light up on their portion of the chat screen.

Assuming a correct answer, Kate gives a positive thumbs up or okay gesture.

Press 1 if it's May and 2 if it's June.

Participant presses the key and sees either the 1 or the 2 light up on their portion of the chat screen.

Assuming a positive answer, Kate gives a positive thumbs up or okay gesture.

Kate: Looks good over here. [Operator name] will get the experiment started for you. For the real experiment, you're not going to be able to see the light corresponding to your answer, since we want your first feedback to be the goldstar or the smiling face or me giving a thumbs up. And I only see your answers for the Live items-otherwise it's just the computer. Let's practice what the real experiment would look like. This time, I'll just give you a thumbs up if you get it right or a thumbs down if you get a wrong. It looks like the first question here on my sheet is about what kind of phone I should get, so when I see the light on my computer, I'll know we've reached the part of the experiment when it's time to ask you a question.

fMRI operator is listening for this cue in the control room and runs the practice trials.

Kate: All right, awesome job on that one. I wish I didn't have to check my email all the time, data plans are so expensive. Hopefully you did just as well on the computer questions! Okay, now it's time to begin the real experiment. Before I get to talk to you again, you'll watch some pictures of animals and other things. So I'll talk to you soon!

fMRI operator begins the rest of the experiment.

Appendix C

Individual Items for fMRI Task

The Story text is the same for all conditions (Live, Social, and Standard), but that the prompt and question are different between the Live condition and the Social and Standard condition. For the Social and Standard condition, in the prompt, "I" is replaced by a name (e.g., "Daniel wants to do some exercise", and in the question, "I" is replaced by the appropriate pronoun (e.g., "he").

Story Text	LIVE prompt	LIVE question
There are two activities after work. One is a pick-up soccer game and one is a movie premiere.	I want to do some exercise.	Which activity should I do?
The zoo has two animal shows. There is an indoor tiger show and a dolphin show in water.	I do not want to get wet.	Which show should I watch?
There are two apartments for rent. One has a big bedroom closet and one has a big kitchen.	I have a lot of clothes.	Which apartment should I rent?
The are two phone apps that cost a dollar. One has cooking recipes and one has movie reviews.	I am hosting a big dinner party next week.	Which app should I buy?
There are two things on the breakfast menu. One is pancakes and one is a bowl of fruit.	I am trying to eat healthy.	Which food should I eat?
There are two restaurants for brunch. One is a big restaurant and one is a small coffee shop.	I want to have lots of guests at brunch.	Where should I go for brunch?
There are two cafeterias in the office building. One has just tea and the other has juice drinks.	I don't like tea that much.	Which cafe should I visit?
There are two ways to carry things. Purses look nice and backpacks can hold a lot of stuff.	I have to carry a lot of books.	Which should I bring with me to school?
There are two kinds of desserts in the cookbook. Cake takes an hour and fruit takes 15 minutes.	I am in a rush to school.	Which kind of dessert should I make?
There are two restaurants for dinner. One serves fancy French food and one is a fast food restaurant.	I want to go somewhere nice on a date.	Where should I go for dinner?
There are two types of exercise. Running is good for speed and lifting weights is good for strength.	I want to move heavy things.	Which type of exercise should I do?
There are two hotels in the town. One has a waterpark and one is a bed and breakfast.	I want a quiet room to sleep in.	Which hotel should I stay at?
There are two kinds of books at the library book sale. There are short stories and long novels.	I like spending a lot of time reading.	What kind of book should I read?

There is a math test soon. There is a month-long review class and a one week review class.	I got an A on the last math test.	Which class should I take?
There are two microwaves to buy. One is large red microwave and one is a small white microwave.	I have an all-white kitchen.	Which microwave should I buy?
There are two things for a snack. One is strawberry yogurt and one is a bagel with butter.	I am trying to eat more dairy.	Which food should I eat?
There are two ways to move. Things can be tightly packed in cardboard boxes or carried by hand.	I need to move an expensive vase.	How should I move my vase?
There are two showings of the new comedy movie. One is at 7pm and one is at 10pm.	I like going to bed early.	What time should I see the movie?
There are two necklaces for sale. One is a long silver chain and one is a gold choker.	I need more gold jewelry.	Which necklace should I pick?
There are two new TV shows. One is on the History Channel and one is on Cartoon Network.	I want to watch true stories.	Which TV channel should I watch?
There are two new night classes. Pottery class is on Tuesdays and the music class in on Mondays.	I have to work late on Mondays.	Which class should I take?
Two places host parties. During the day, the pool is open and at night, the restaurant is open.	I want to have my party at night.	Where should I have my party?
The new phone has two cases. One has bright colors and patterns and one is black and gray.	I like phones with dark colors.	What phone should I pick?
There are two popular TV shows on at 9pm. One is a comedy and one is a drama.	I like shows that are funny.	Which type of show should I watch?
There is a science test coming up. There is an expensive review class and a cheap review class.	I am trying to save money.	Which class should I take?
Two kinds of shirts are on sale at the department store. T-shirts are cheap and turtlenecks are expensive.	I need new winter clothes.	Which shirt should I buy?
There are two recently released computer games. One is a solitaire game and one is a racing game.	I like to play competitive games.	What game should I buy?
There are two sofas for sale. There is a soft brown sofa and a small brown wooden sofa.	I need comfortable furniture for my living room.	Which sofa should I buy?
There are two teams at the gym. There is a beginning basketball team and an expert golf team.	I like being very active.	Which sport should I play?
There are two kinds of pets for sale at the store. There are colorful goldfish and beagle puppies.	I like playing with animals.	What kind of pet should I pick?
There are two suburbs outside the city. One is close to the metro and one is far away.	I don't own a car.	Which suburb should I pick?

There are two things to wear for summer exercise. There are athletic shorts and there are bathing suits.	I am going to a water park.	Which clothes should I wear?
There are two new movies in theaters. One is a romantic comedy and one is a scary movie.	I really like to laugh.	Which movie should I see?
There are two sports to watch. There is a college football rivalry game and there is women's gymnastics.	I like watching team sports.	Which sport should I watch?
There are two types of wallpaper. One has lots of flowers and leaves and one has blue stripes.	I really like being outdoors.	Which wallpaper should I pick?
There are two places to buy the book. The bookstore is more expensive and the website is cheaper.	I would like to save money.	Where should I buy the book?

Appendix D

Autism Quotient

Below is a list of statements. Please read each statement very carefully and rate how strongly you agree or disagree with it by circling the number with your answer.

		Definitely Agree	Slightly Agree	Slightly Disagree	Definitely Disagree
1.	I prefer to do things with others rather than on my own.	1	2	3	4
2.	I prefer to do things the same way over and over again.	1	2	3	4
3.	If I try to imagine something, I find it very easy to create a picture in my mind.	1	2	3	4
4.	I frequently get so strongly absorbed in one thing that I lose sight of other things.	1	2	3	4
5.	I often notice small sounds when others do not.	1	2	3	4
6.	I usually notice car number plates or similar strings of information.	1	2	3	4
7.	Other people frequently tell me that what I've said is impolite, even though I think it is polite.	1	2	3	4
8.	When I'm reading a story, I can easily imagine what the characters might look like.	1	2	3	4
9.	I am fascinated by dates.	1	2	3	4
10.	In a social group, I can easily keep track of several different people's conversations.	1	2	3	4
11.	I find social situations easy.	1	2	3	4
12.	I tend to notice details that others do not.	1	2	3	4
13.	I would rather go to a library than a party.	1	2	3	4
14.	I find making up stories easy.	1	2	3	4

		Definitely Agree	Slightly Agree	Slightly Disagree	Definitely Disagree
15.	I find myself more strongly drawn to people than to things.	1	2	3	4
16.	I tend to have very strong interests which I get upset about if I can't pursue.	1	2	3	4
17.	I enjoy social chit-chat.	1	2	3	4
18.	When I talk, it isn't always easy for others to get a word in edgewise.	1	2	3	4
19.	I am fascinated by numbers.	1	2	3	4
20.	When I'm reading a story, I find it difficult to work out the characters' intensions.	1	2	3	4
21.	I don't particularly enjoy reading fiction.	1	2	3	4
22.	I find it hard to make new friends.	1	2	3	4
23.	I notice patterns in things all the time.	1	2	3	4
24.	I would rather go to the theater than a museum.	1	2	3	4
25.	It does not upset me if my daily routine is disturbed.	1	2	3	4
26.	I frequently find I don't know how to keep a conversation going.	1	2	3	4
27.	I find it easy to "read between the lines" when someone is talking to me.	1	2	3	4
28.	I usually concentrate more on the whole picture, rather than the small details.	1	2	3	4

		Definitely Agree	Slightly Agree	Slightly Disagree	Definitely Disagree
29.	I am not very good at remembering phone numbers.	1	2	3	4
30.	I don't usually notice small changes in a situation, or a persons' appearance.	1	2	3	4
31.	I know how to tell if someone listening to me is getting bored.	1	2	3	4
32.	I find it easy to do more than one thing at once.	1	2	3	4
33.	When I talk on the phone, I'm not sure when it's my turn to speak.	1	2	3	4
34.	I enjoy doing things spontaneously.	1	2	3	4
35.	I am often the last to understand the point of a joke.	1	2	3	4
36	I find it easy to work out what someone is thinking or feeling just by looking at their face.	1	2	3	4
37.	If there is an interruption, I can switch back to what I was doing very quickly.	1	2	3	4
38.	I am good at social chit-chat.	1	2	3	4
39.	People often tell me that I keep going on about the same thing.	1	2	3	4
40.	When I was young, I used to enjoy playing games involving pretending with other children.	1	2	3	4
41.	I like to collect information about categories of things (e.g., types of car, types of bird, types of train, types of plant, etc.)	1	2	3	4
42.	I find it difficult to imagine what it would be like to be someone else.	1	2	3	4

		Definitely Agree	Slightly Agree	Slightly Disagree	Definitely Disagree
43.	I like to plan any activities I participate in carefully.	1	2	3	4
44.	I enjoy social occasions.	1	2	3	4
45.	I find it difficult to work out people's intentions.	1	2	3	4
46.	New situations make me anxious.	1	2	3	4
47.	I enjoy meeting new people.	1	2	3	4
48.	I am a good diplomat.	1	2	3	4
49.	I am not very good at remembering people's date of birth.	1	2	3	4
50.	I find it very easy to play games with children that involve pretending.	1	2	3	4

Appendix E

Demographics Questionnaire

Project Title: Functional brain organization in typical development

Date: ___________________ Participant ID # __________

This information will only be used to determine the demographics of our sample population.
All responses will be kept confidential and will only be identified with a number that is not
connected to your name.

Demographics

1. Participant's Age (Years and months): ______________________

2. Participant's Gender: ___Male ___Female

3. Participant's Ethnicity (please check one):
   ____ Hispanic or Latino
   ____ Not Hispanic or Latino
   ____ Does not wish to disclose

4. Participant's Race: (More than one option may be checked)
   _____ American Indian or Alaskan Native
   _____ Asian
   _____ Native Hawaiian or Pacific Islander
   ____ White or Caucasian
   _____ Black or African American
   _____ Does not wish to disclose

5. Number and ages of siblings:
____

6. Have you been exposed to a language other than English? ____ yes ____ no
If yes, in what way (please note extent of exposure, duration, and fluency level)?

5. What is the highest level of education you have completed?
Some High School ______
High School ______
Some College ______
Technical or AA Degree ______
College Degree ______
Some Graduate School ______

Post Graduate Degree ______

6. What is the highest level of education completed by your parents?

	Parent 1	Parent 2
Some High School	______	______
High School	______	______
Some College	______	______
Technical or AA Degree	______	______
College Degree	______	______
Some Graduate School	______	______
Post Graduate Degree	______ _	_____

Hand Usage

1. Which is your dominant hand? ___Left ___Right ___Use both hands
2. Please indicate which hand you would ordinarily use for each activity.

a. Write	Left	Right	Either hand
b. Draw	Left	Right	Either hand
c. Use a bottle opener	Left	Right	Either hand
d. Throw a snowball	Left	Right	Either hand
e. Use a hammer	Left	Right	Either hand
f. Use a toothbrush	Left	Right	Either hand
g. Use a screwdriver	Left	Right	Either hand
h. Use an eraser	Left	Right	Either hand
i. Use a pair of scissors	Left	Right	Either hand
j. Use a tennis racket	Left	Right	Either hand
k. Hold a match when striking it	Left	Right	Either hand
l. Stir a can of paint	Left	Right	Either hand

3. At any point in your life, were you forced to switch from your dominant hand to your less
dominant hand? ___Yes ___No

4. Is anyone in your family (e.g., parents, brothers, sisters) or any other blood relative left handed? ___Yes ___No

a. If yes, which family member(s) are left handed? ________________________________

5. Are there times when you find yourself using your less dominant hand for the very same
things that you would normally do with your dominant hand?
___Yes ___No
a. If yes, please indicate which activities: ____
____

b. Which of the following is the reason for use of the less dominant hand?
- Consciously trying to use it more ___Yes ___No
- Due to injury or impairment ___Yes ___No
- Necessary to play a sport or a musical instrument ___Yes ___No

Appendix F

Post-test Questionnaire

Note: Different versions of this questionnaire were given to the participants from whom we have already collected data. Items given to all N=17 participants are denoted with an asterisk and additional items given to N=8 of those participants are denoted with a double asterisk. The blocks about the different conditions (i.e., Live, Social, Standard) were randomized. Items which will compose the "liveness" index for each condition are bolded.

1*) What were your overall impressions?

2*) How many different people's voices did you hear?

1 2 3 4 More than 4

We actually designed the study so that in addition to hearing the live speaker, you heard two other people: an older woman with a more monotone voice and a younger speaker with a more engaging voice. We're now going to ask you some questions about those speakers you heard.

Think about that first speaker, who had a lower pitched voice. [play clip

]

3*) How much did you like this speaker?

1	2	3	4	5	6	7
Disliked A lot			Neutral			Liked A lot

4*) How much did you pay attention when she was talking?

1	2	3	4	5	6	7
Very Little			Neutral			Very Much

5*) How motivated were you to get the questions asked in her voice right?

1	2	3	4	5	6	7
Very Unmotivated			Neutral			Very Motivated

6*) How much did your experiences with her feel live?

1	2	3	4	5	6	7
Not at all live			Neutral			Very Live

When she asked the questions, you saw either gold stars or red Xs.

7*) How did you feel when you received feedback that you got the right answer?

1	2	3	4	5	6	7
Very Negative			Neutral			Very Positive

8*) Do you remember if you ever got one of those questions wrong and saw a red X?

Yes No

[If yes]
9*) How did you feel when you received feedback that you got the wrong answer?

1	2	3	4	5	6	7
Very Negative			Neutral			Very Positive

10) How much did it feel like this speaker was talking directly to you?

1	2	3	4	5	6	7
Not at all direct		Neutral			Felt like she was talking to me directly	

Think about that second recorded speaker, who had a more high pitched voice [play clip]

11*) How much did you like this speaker?

1	2	3	4	5	6	7
Disliked A lot		Neutral				Liked A lot

12*) How much did you pay attention when she was talking?

1	2	3	4	5	6	7
Very Little		Neutral				Very Much

13*) How motivated were you to get the questions asked in her voice right?

1	2	3	4	5	6	7
Very Unmotivated			Neutral			Very Motivated

14*) How much did your experiences with her feel live?

1	2	3	4	5	6	7
Not at all live			Neutral			Very Live

When she asked the questions, you saw either smiling faces or sad faces.

15*) How did you feel when you received feedback that you got the right answer?

1	2	3	4	5	6	7
Very Negative			Neutral			Very Positive

16*) Do you remember if you ever got one of those questions wrong and saw a sad face?

Yes No

[If yes]
17*) How did you feel when you received feedback that you got the wrong answer?

1	2	3	4	5	6	7
Very Negative			Neutral			Very Positive

18) How much did it feel like this speaker was talking directly to you?

1	2	3	4	5	6	7
Not at all direct			Neutral		Felt like she was talking to me directly	

Think about the person who was talking to you live over the computer.

19*) How much did you like this speaker?

1	2	3	4	5	6	7
Disliked A lot			Neutral			Liked A lot

20*) How much did you pay attention when she was talking?

1	2	3	4	5	6	7
Very Little			Neutral			Very Much

21*) How motivated were you to get the questions asked in her voice right?

1	2	3	4	5	6	7
Very Unmotivated			Neutral			Very Motivated

22*) How much did your experiences with her feel live?

1	2	3	4	5	6	7
Not at all live			Neutral			Very Live

When she asked the questions, she would give you positive or negative feedback.

23*) How did you feel when you received feedback that you got the right answer?

1	2	3	4	5	6	7
Very Negative			Neutral			Very Positive

24*) Do you remember if you ever got one of those questions wrong?

Yes No

[If yes]

25*) How did you feel when you received feedback that you got the wrong answer?

1	2	3	4	5	6	7
Very Negative			Neutral			Very Positive

26*) How much did it feel like this speaker was talking directly to you?

1	2	3	4	5	6	7
Not at all direct			Neutral		Felt like she was talking to me directly	

27*) Did any particular items or questions seem strange?

28*) How hard was the task? How hard was it to pay attention?

29*) Did the volume and audio characteristics seem fairly consistent?

30*) Were there times when the live speaker felt like she was interacting with you more than other times?

31*) Do you think there was anything more to the study than what we told you about?

32) If yes, what was more to the study than we told you about?

33*) Do you have any other thoughts about the task to share with us?

34) How much did it feel like the Live speaker was interacting with you directly in real time versus sounding like a recording?**

1	2	3	4	5	6	7
Felt like I was listening to a recording			Neutral		Felt like she she was talking to me directly	

35) How was your interaction with the live speaker before the start of the main part of the experimenter?

1	2	3	4	5	6	7
Very Negative			Neutral			Very Positive

36) How much do you think you would like interacting with her again in real life?

1	2	3	4	5	6	7
Not at all			Neutral		Very much	

37) How much did it feel like the higher pitched speaker was interacting with you directly in real time versus sounding like a recording?**

1	2	3	4	5	6	7
Felt like I was listening to a recording		Neutral			Felt like she she was talking to me directly	

38) How much do you think you would like interacting with her again in real life?

1	2	3	4	5	6	7
Not at all			Neutral		Very much	

39) How much did it feel like the lower pitched speaker was interacting with you directly in real time versus sounding like a recording?**

1	2	3	4	5	6	7
Felt like I was listening to a recording		Neutral			Felt like she she was talking to me directly	

40) How much do you think you would like interacting with her again in real life?

1	2	3	4	5	6	7
Not at all			Neutral		Very much	

References

AbdulSabur, N. (2013). *The neural mechanisms supporting structure and inter-brain connectivity in natural conversation.* (Unpublished doctoral dissertation). University of Maryland, College Park.

Anderson, D. K., Maye, M. P., & Lord, C. (2011). Changes in maladaptive behaviors from mid-childhood to young adulthood in autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, *116*, 381-397. doi: 10.1352/1944-7558-116.5.381

Andrews-Hanna, J. R., Reidler, J. S., Huang, C., & Buckner, R. L. (2010). Evidence for the default network's role in spontaneous cognition. *Journal of Neurophysiology*, *104*(1), 322. doi: 10.1152/jn.00830.2009

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953. doi:10.1037/a0016923

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, *16*(10), 1773-1784. doi:10.1162/0898929042947928

Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology*, *45*, 190-201. doi: 10.1037/a0014098

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, *38*(2), 451-460. doi: 10.1016/j.paid.2004.04.022

Auvray, M., Lenay, C., & Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, *27*(1), 32-47. doi:10.1016/j.newideapsych.2007.12.002

Baayen, R. H. and Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3(2), 12-28.

Bal, E., Yerys, B. E., Sokoloff, J. L., Celano, M. J., Kenworthy, L., Giedd, J. N., & Wallace, G. L. (2013). Do social attribution skills improve with age in children with high functioning autism spectrum disorders? *Research in Autism Spectrum Disorders*, *7*, 9-16. doi: 10.1016/j.rasd.2012.07.004

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5-17. doi:10.1023/A:1005653411471

Ben-Shachar, M., Palti, D., & Grodzinsky, Y. (2004). Neural correlates of syntactic movement: converging evidence from two fMRI experiments. *Neuroimage*, *21*(4), 1320-1336. doi: 10.1016/j.neuroimage.2003.11.027

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., ... & Neubauer, A. C. (2014). Creating metaphors: the neural basis of figurative language production. *Neuroimage*, *90*, 99-106. doi: 10.1016/j.neuroimage.2013.12.046

Bierman, K. L., & Welsh, J. A. (2000). Assessing social dysfunction: The contributions of laboratory and performance-based measures. *Journal of Clinical Child Psychology, 29*(4), 526-539. doi:10.1207/S15374424JCCP2904_6

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341-345.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433-436. doi:10.1163/156856897X00357

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*(2), B13-B25. doi:10.1016/S0010-0277(99)00081-5

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482. doi:10.1037/0278-7393.22.6.1482

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, *53*, 301-344. doi:10.1016/S0079-7421(10)53008-1

Bress, J. N., Smith, E., Foti, D., Klein, D. N., & Hajack, G. (2012). Neural response to reward and depressive symptoms in late childhood to early adolescence. *Biological Psychology*, *89*, 156-162. doi: 10.1016/j.biopsycho.2011.10.004

Brunyé, T. T., Ditman, T., Giles, G. E., Mahoney, C. R., Kessler, K., & Taylor, H. A. (2012). Gender and autistic personality traits predict perspective-taking ability in typical adults. *Personality and Individual Differences*, *52*(1), 84-88. doi:10.1016/j.paid.2011.09.004

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, *1124*(1), 1-38. doi: 10.1196/annals.1440.011

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49-57. doi: 10.1016/j.tics.2006.11.004

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73-92. doi:10.1002/icd.298

Camerer, C. F. (2008) Behavioral game theory and the neural basis of strategic choice. In P. Glimcher, E. Fehr, C. Camerer, & Poldrack, R. (Eds.), *Neuroeconomics: Decision Making and the Brain* (pp. 193-206). San Diego: Academic Press.

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, *1*(2), 380-392. doi:10.1111/j.1756-8765.2009.01026.x

Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 159-181). Cambridge, MA: MIT Press.

Carr, A. (2011). Social and emotional development in middle childhood. In D. Skuse, H. Bruce, L. Dowdney, & D. Mrazek (Eds.), *Child psychology and psychiatry: Frameworks for practice* (2nd ed., pp. 56-61). Hoboken, NJ: Wiley-Blackwell.

Carter, E. J., & Pelphrey, K. A. (2006). School-aged children exhibit domain-specific responses to biological motion. *Social Neuroscience*, 1, 396-411. doi: 10.1080/17470910601041382

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage*, *50*(3), 1148-1167. doi: 10.1016/j.neuroimage.2009.12.112

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893. doi:10.1037/0022-3514.76.6.893

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231-239. doi: 10.1016/j.tics.2012.02.007

Chilcoat, H. D., & Anthony, J. C. (1996). Impact of parent monitoring on initiation of drug use through late childhood. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(1), 91-100. doi: 10.1097/00004583-199601000-00017

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., & Walter, H. (2007). The intentional network: How the brain reads varieties of intentions. *Neuropsychologia*, *45*(13), 3105-3113. doi: 10.1016/j.neuorpsychologia.2007.05.011

Clark, H. H. (1996). *Using Language* (Vol. 4). Cambridge: Cambridge University Press.

Coricelli, G., & Nagel, R. (2009). Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proceedings of the National Academy of Sciences*, *106*(23), 9163-9168. doi:10.1073/pnas.0807721106

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162-173. doi:10.1006/cbmr.1996.0014

Cycowicz, Y. M., Friedman, D., Snodgrass, J. G., & Rothstein, M. (2000). A developmental trajectory in implicit memory is revealed by picture fragment completion. *Memory*, *8*, 19-35. doi: 10.1080/096582100387687

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., ... & Davidson, R. J. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, *8*(4), 519-526. doi:10.1038/nn1421

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Does social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10): 441-447. doi:10.1016/j.tics.2010.06.009

Delmonte, S., Balsters, J. H., McGrath, J., Fitzgerald, J., Brennan, S., Fagan, A. J., & Gallagher, L. (2012). Social and monetary reward processing in autism spectrum disorders. *Mol. Autism*, *3*(7). doi: 10.1186/2040-2392-3-7

Desmond, J. E., & Glover, G. H. (2002). Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *Journal of Neuroscience Methods, 118*(2), 115-128. doi:10.1016/S0165-0270(02)00121-8

Diamond, A. (2001). A model system for studying the role of dopamine in prefrontal cortex during early development in humans. In C. Nelson & M. Luciana (eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 433-472). Cambridge, MA: MIT Press.

Diamond, A., & Amso, D. (2008). Contributions of neuroscience to our understanding of cognitive development. *Current Directions in Psychological Science*, *17*, 136-141. doi: 10.1111/j.1467-8721.2008.00563.x

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage*, *55*(2), 705-712. doi: 10.1016/j.neuroimage.2010.12.040

Dubinsky, J. M. (2010). Neuroscience education for prekindergarten–12 teachers. *The Journal of Neuroscience*, *30*, 8057-8060. doi: 10.1523/jneurosci.2322-10.2010

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), e75468. doi:10.1371/journal.pone.0075468

Eccles, J. S., Roeser, R., Wigfield, A., & Freedman-Doan, C. (1999). Academic and motivational pathways through middle childhood. In L. Balter & C. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 287-317). Philadelphia: Psychology Press.

Ethofer, T., Bretscher, J., Wiethoff, S., Bisch, J., Schlipf, S., Wildgruber, D., & Kreifelts, B. (2013). Functional responses and structural connections of cortical areas for processing faces and voices in the superior temporal sulcus. *Neuroimage, 76, 45-56.* doi:10.1016/j.neuroimage.2013.02.064

Ethofer, T., Gschwind, M., & Vuilleumier, P. (2011). Processing social aspects of human gaze: a combined fMRI-DTI study. *Neuroimage*, *55*(1), 411-419. doi: 10.1016/j.neuroimage.2010.11.033

Ethofer, T., Van De Ville, D., Scherer, K., & Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Current Biology*, *19*(12), 1028-1033. doi:10.1016/j.cub.2009.04.054

Farmer, T. W., Irvin, M. J., Motoca, L. M., Leung, M. C., Hutchins, B. C., Brooks, D. S., & Hall, C. M. (in press). Externalizing and Internalizing Behavior Problems, Peer Affiliations, and Bullying Involvement Across the Transition to Middle School. *Journal of Emotional and Behavioral Disorders*.

Farrell, L., & Barrett, P. (2006). Obsessive-compulsive disorder across developmental trajectory: Cognitive processing of threat in children, adolescents and adults. *British Journal of Psychology*, *97*, 95-114. doi: 10.1348/000712605X58592

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 17245-17250. doi:10.1073/pnas.0502205102

Fazio, R. H., & Zanna, M. P. (1978). Attitudinal qualities relating to the strength of the attitude-behavior relationship. *Journal of Experimental Social Psychology*, *14*(4), 398-408. doi:10.1016/0022-1031(78)90035-5

Feiring, C., & Lewis, M. (1991). The transition from middle childhood to early adolescence: Sex differences in the social network and perceived self-competence. *Sex Roles*, *24*, 489-509. doi: 10.1007/BF00289335

Fink, G. R., Markowitsch, H. J., Reinkemeier, M., Bruckbauer, T., Kessler, J., & Heiss, W. D. (1996). Cerebral representation of one's own past: neural networks involved in autobiographical memory. *The Journal of Neuroscience*, *16*(13), 4275-4282.

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*(17), 1920-1931. doi:10.1016/j.visres.2011.07.002

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PloS One*, *8*(1), e53286. doi:10.1371/journal.pone.0053286

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78-84. doi:10.1016/S1364-6613(00)01839-8

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, *91*(4), 1357-1392. doi:10.1152/physrev.00006.2011

Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex*, *13*, 170-177. doi: 10.1093/cercor/13.2.170

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

Fukui, H., Murai, T., Shinozaki, J., Aso, T., Fukuyama, H., Hayashi, T., & Hanakawa, T. (2006). The neural basis of social tactics: An fMRI study. *Neuroimage*, *32*(2), 913-920. doi:10.1016/j.neuroimage.2006.03.039

Gallagher, S. (2012). In defense of phenomenological approaches to social cognition: Interacting with the critics. *Review of Philosophy and Psychology*, *3*, 187-212. doi:10.1007/s13164-011-0080-1

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage*, *16*(3), 814-821. doi:10.1006/nimg.2002.1117

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences, 17,* 160-165. doi:10.1016/j.tics.2013.02.002

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *The Journal of Neuroscience*, *26*, 6885-6892. doi: 10.1523/JNEUROSCI.1062-06.2006

Gao, T., Scholl, B. J., & McCarthy, G. (2012). Dissociating the detection of intentionality from animacy in the right posterior superior temporal sulcus. *The Journal of Neuroscience, 32*(41), 14276-14280. doi:10.1523/JNEUROSCI.0562-12.2012

Garrod, S., & Pickering, M.J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*, 8–11. doi:10.1016/j.tics.2003.10.016

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*(2), 292-304. doi:10.1111/j.1756-8765.2009.01020.x

Gayle, L. C., Gal, D. E., & Kieffaber, P. D. (2012). Measuring affective reactivity in individuals with autism spectrum personality traits using the visual mismatch negativity event-related brain potential. *Frontiers in Human Neuroscience*, *6*(334), 1-7. doi:10.3389/fnhum.2012.00334/abstract

Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: what have we learned and where are we going? *Neuron*, *67*, 728-734. doi: 10.1016/j.neuron.2010.08.040

Gobbini, M. I., Koralek, A. C., 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

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*, 515-523. doi:10.1111/j.1467-9280.2008.02117.x

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, *45*(5), 1430. doi:10.1037/a0016122

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*(5), 711-720. doi: 10.1162/089892900562417

Grossmann, T., & Johnson, M. H. (2010). Selective prefrontal cortex responses to joint attention in early infancy. *Biology Letters*, *6*(4), 540-543. doi:10.1098/rsbl.2009.1069

Guionnet, S., Nadel, J., Bertasi, E., Sperduti, M., Delaveau, P., & Fossati, P. (2012). Reciprocal imitation: toward a neural basis of social interaction. *Cerebral Cortex*, *22*(4), 971-978. doi:10.1093/cercor/bhr177

Gweon, H., Dodell-Feder, D., Bedny, M., & Saxe, R. (2012). Theory of Mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. *Child Development*, *83*, 1853-1868. doi: 10.1111/j.1467-8624.2012.01829.x

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the" Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44, 1457. doi: 10.1037/a0012682

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223-233. doi:10.1016/S1364-6613(00)01482-0

Hietanen, J.K., Leppänen, J.M., Peltola, M.J., Linna-aho, K., Ruuhiala, H.J. (2008). Seeing direct and averted gaze activates the approach-avoidance motivational brain systems. *Neuropsychologia*, 46(9), 2423–30. doi:10.1016/j.neuropsychologia.2008.02.029

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, *58*(2), 284-294. doi: 10.1016/j.neuron.2008.03.020.

Izuma, K., Saito, D. N., & Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience*, *22*(4), 621-631. doi:10.1162/jocn.2009.21228

Jackson, C., Henriksen, L., Dickinson, D., Messer, L., & Robertson, S. B. (1998). A longitudinal study predicting patterns of cigarette smoking in late childhood. *Health Education & Behavior*, *25*(4), 436-447. doi: 10.1177/109019819802500403

Johnson, S. B., Blum, R. W., & Giedd, J. N. (2009). Adolescent maturity and the brain: the promise and pitfalls of neuroscience research in adolescent health policy. *Journal of Adolescent Health*, *45*(3), 216-221. doi: 10.1016/j.jadohealth.2005.05.016

Kaiser, M. D., Hudac, C. M., Shultz, S., Lee, S. M., Cheung, C., Berken, A. M., ... & Pelphrey, K. A. (2010). Neural signatures of autism. *Proceedings of the National Academy of Sciences*, *107*(49), 21223-21228. doi:10.1073/pnas.1010412107

Kampe, K. K., Frith, C. D., & Frith, U. (2003). "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *The Journal of Neuroscience*, *23*, 5258-5263.

Kanne, S. M., Gerber, A. J., Quirmbach, L. M., Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2011). The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *Journal of Autism and Developmental Disorders*, *41*, 1007-1018. doi: 10.1007/s10803-010-1126-4

Kingstone, A., Tipper, C., Ristic, J., & Ngan, E. (2004). The eyes have it!: An fMRI investigation. *Brain and Cognition*, *55*(2), 269-271. doi:10.1016/j.bandc.2004.02.037

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*(3), 299-314. doi:10.1016/j.jecp.2008.07.005

Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: the Social Attribution Task. *Journal of Child psychology and Psychiatry*, *41*(7), 831-846. doi: 10.1111/1469-7610.00671

Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*, 345-360. doi:10.1098/rstb.2002.1202

Konvalinka, I., & Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Frontiers in Human Neuroscience*, *6*, 60-69. doi:10.3389/fnhum.2012.00215

Koster-Hale, J., & Saxe, R. (2013) Functional neuroimaging of theory of mind. In S. Baron-Cohen, M. Lombardo, & H. Tager-Flusberg (Eds.), *Understanding Other Minds* (pp. 132-163). Oxford: Oxford University Press.

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, 10, 110-120. doi: 10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096-9101. doi:10.1073/pnas.1532872100

Laidlaw, K. E., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proceedings of the National Academy of Sciences*, *108*(14), 5548-5553. doi:10.1073/pnas.1017022108

Lam, C. B., McHale, S. M., & Crouter, A. C. (2012). Parent–child shared time from middle childhood to late adolescence: Developmental course and adjustment correlates. *Child Development*, *83*, 2089-2103. doi: 10.1111/j.1467-8624.2012.01826.x

Lee, T., Liu, H. L., Tan, L. H., Chan, C. C., Mahankali, S., Feng, C. M., ... & Gao, J. H. (2002). Lie detection by functional magnetic resonance imaging. *Human Brain Mapping*, *15*(3), 157-164. doi: 10.1002/hbm.10020

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, *12*(2), 264-271. doi:10.1111/j.1467-7687.2008.00758.x

Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social cognitive and affective neuroscience*, *7*(3), 274-281. doi: 10.1093/scan/nsr006

Locke, J. (1995). *The child's path to spoken language*. Harvard University Press.

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual review of psychology*, *62*, 103-134. doi: 10.1146/annurev-psych-120709-145406

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*(20), 11832-11835. doi:10.1073/pnas.211415698

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development, 54*, 702-709. doi:10.2307/1130058

Merin, N., Young, G. S., Ozonoff, S., & Rogers, S. J. (2007). Visual fixation patterns during reciprocal social interaction distinguish a subgroup of 6-month-old infants at-risk for autism from comparison infants. *Journal of Autism and Developmental Disorders*, *37*(1), 108-121. doi:10.1007/s10803-006-0342-4

Meyer, A., Weinberg, A., Klein, D. N., & Hajcak, G. (2012). The development of the error-related negativity (ERN) and its relationship with anxiety: Evidence from 8 to 13 year-olds. *Developmental Cognitive Neuroscience*, *2*, 152. doi: 10.1016/j.dcn.2011.09.005

Mills, K. L., Lalonde, F., Clasen, L. S., Giedd, J. N., & Blakemore, S. J. (2014). Developmental changes in the structure of the social brain in late childhood and adolescence. *Social Cognitive and Affective Neuroscience*, *9*, 123-131. doi: 10.1093/scan/nss113

Mitchell, J. P., Cloutier, J., Banaji, M. R., & Macrae, C. N. (2006). Medial prefrontal dissociations during processing of trait diagnostic and nondiagnostic person information. *Social Cognitive and Affective Neuroscience*, *1*(1), 49-55. doi:10.1093/scan/nsl007

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences*, *99*(23), 15238-15243. doi: 10.1073/pnas.232395699

Monk, C. S., McClure, E. B., Nelson, E. E., Zarahn, E., Bilder, R. M., Leibenluft, E., Charney, D. S., Ernst, M., & Pine, D. S. (2003). Adolescent immaturity in attention-related brain engagement to emotional facial expressions. *Neuroimage*, *20*, 420-428. doi: 10.1016/S1053-8119(03)00355-0

Möttönen, R., Calvert, G. A., Jääskeläinen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J., & Sams, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage*, *30*(2), 563-569. doi:10.1016/j.neuroimage.2005.10.002

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177-197). Norwood, NJ: Ablex.

Murray, C. T., & Trevarthen, C. C. (1986). The infant's role in mother–infant communication. *Journal of Child Language*, *13*, 15-29. doi:10.1017/S0305000900000271

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437-442. doi:10.1163/156856897X00366

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, *16*(10), 1706-1716. doi:10.1162/0898929042947900

Pelphrey, K. A., Singerman, J. D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia*, *41*(2), 156-170. doi:10.1016/S0028-3932(02)00146-X

Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Thinking of mental and other representations: The roles of left and right temporo-parietal junction. *Social Neuroscience*, *1*(3-4), 245-258. doi:10.1080/17470910600989896

Pierno, A. C., Becchio, C., Turella, L., Tubaldi, F., & Castiello, U. (2008). Observing social interactions: the effect of gaze. *Social Neuroscience*, *3*(1), 51-59. doi:10.1080/17470910701563269

Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*, *72*(5), 692-697. doi: 10.1016/j.neuron.2011.11.001

Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., & Hietanen, J. K. (2011). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Social Cognitive and Affective Neuroscience*, *6*(4), 486-494. doi:10.1093/scan/nsq068

Pönkänen, L. M., Peltola, M. J., & Hietanen, J. K. (2011). The observer observed: Frontal EEG asymmetry and autonomic responses differentiate between another person's direct and averted gaze when the face is seen live. *International Journal of Psychophysiology, 82*(2), 180-187. doi:10.1016/j.ijpsycho.2011.08.006

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, 1191(1), 62-88. doi:10.1111/j.1749-6632.2010.05444.x

Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 435-445. doi:10.1098/rstb.2002.1221

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*, *98*(2), 676-682. doi: 10.1073/pnas.98.2.676

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*(3), 510-532. doi:10.1037/0033-2909.114.3.510

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neuroscience & Biobehavioral Reviews*, *32*(1), 123-142. doi:10.1016/j.neubiorev.2007.06.004

Redcay, E., Dodell-Feder, D., Mavros, P. L., Kleiner, M., Pearrow, M. J., Triantafyllou, C., ... & Saxe, R. (2012). Atypical brain activation patterns during a face-to-face joint attention game in adults with autism spectrum disorder. *Human Brain Mapping*, *34*(10), 2511-2532. doi:10.1002/hbm.22086

Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D., & Saxe, R. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage*, *50*(4), 1639-1647. doi:10.1016/j.neuroimage.2010.01.052

Redcay, E., Rice, K., & Saxe, R. (2013). Interaction versus observation: A finer look at this distinction and its importance to autism. *Behavioral and Brain Sciences*, *36*, 435-435. doi:10.1017/S0140525X12002026

Reddy, V. (2003). On being the object of attention: implications for self–other consciousness. *Trends in Cognitive Sciences*, *7*(9), 397-402. doi:10.1016/S1364-6613(03)00191-8

Reed, P., Lowe, C., & Everett, R. (2011). Perceptual learning and perceptual search are altered in male university students with higher Autism Quotient scores. *Personality and Individual Differences*, *51*(6), 732-736. doi:10.1016/j.paid.2011.06.016

Riby, D. M., Whittle, L., & Doherty-Sneddon, G. (2012). Physiological reactivity to faces via live and video-mediated communication in typical and atypical development. *Journal of Clinical and Experimental Neuropsychology*, *34*(4), 385-395. doi:10.1080/13803395.2011.645019

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*(5), 407-413. doi:10.1111/j.1467-9280.2007.01914.x

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage*, *23*(2), 752-763. doi: 10.1016/j.neuroimage.2004.06.035

Rochat, P., Neisser, U., & Marian, V. (1998). Are young infants sensitive to interpersonal contingency? *Infant Behavior and Development*, *21*(2), 355-366. doi:10.1016/S0163-6383(98)90012-1

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (In press). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*. doi:10.1111/cdev.12166

Ryan, L., Nadel, L., Keil, K., Putnam, K., Schnyer, D., Trouard, T., & Moscovitch, M. (2001). Hippocampal complex and retrieval of recent and very remote autobiographical memories: evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus*, *11*(6), 707-714. doi: 10.1002/hipo.1086

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 696-735. doi:10.2307/412243

Saxe, R. (2009). Theory of Mind (Neural Basis). In W. Banks (Ed.), *Encyclopedia of consciousness* (pp. 401-409). Oxford: Elsevier.

Saxe, R. & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in theory of mind. *Neuroimage*, 19, 1835-1842. doi:10.1016/S1053-8119(03)00230-1

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, 692-699. doi:10.1111/j.1467-9280.2006.01768.x

Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Social Neuroscience*, *1*(3-4), 284-298. doi:10.1080/17470910601000446

Saxe, R. R., Whitfield-Gabrieli, S., Scholz, J., & Pelphrey, K. A. (2009). Brain Regions for Perceiving and Reasoning About Other People in School-Aged Children. *Child Development*, *80*, 1197-1209. doi: 10.1111/j.1467-8624.2009.01325.x

Schilbach, L., Bzdok, D., Timmermans, B., Fox, P. T., Laird, A. R., Vogeley, K., & Eickhoff, S. B. (2012). Introspective minds: using ALE meta-analyses to study commonalities in the neural correlates of emotional processing, social & unconstrained cognition. *PloS one*, *7*(2), e30920. doi: 10.1371/journal.pone.0030920

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36, 393-414. doi:10.1017/S0140525X12000660

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., Shah, N. J., Fink, G. R., & Vogeley, K. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, *22*, 2702-2715. doi:10.1162/jocn.2009.21401

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718-730. doi:10.1016/j.neuropsychologia.2005.07.017

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., Osani, H., Csibra, G. (2010). Absences of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology*, 22, 353-260. doi:10.1017/S0954579410000106

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science*, *325*, 883-885. doi:10.1126/science.1176170

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 326-332. doi:10.1037/0096-1523.29.2.326

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

Shultz, S., Vouloumanos, A., & Pelphrey, K. (2012). The superior temporal sulcus differentiates communicative and noncommunicative auditory signals. *Journal of Cognitive Neuroscience*, *24*(5), 1224-1232. doi:10.1162/jocn_a_00208

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences*, *105*(2), 809-813. doi:10.1073/pnas.0707021105

Smith, S. S., & Richardson, D. (1983). Amelioration of deception and harm in psychological research: The important role of debriefing. *Journal of Personality and Social Psychology*, *44*(5), 1075. doi:10.1037/0022-3514.44.5.1075

Southgate, V., Senju, A.,& Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science,* 18, 587-592. doi:10.1111/j.1467-9280.2007.01944.x

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism*, *11*(3), 265-277. doi:10.1177/1362361307076925

Spreng, R. N., & Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *Journal of Cognitive Neuroscience*,*22*(6), 1112-1123. doi: 10.1162/jocn.2009.21282

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of cognitive neuroscience*,*21*(3), 489-510. doi: 10.1162/jocn.2008.21029

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, *107*(32), 14425-14430. doi:10.1073/pnas.1008662107

Stormark, K. M., & Braarud, H. C. (2004). Infants' sensitivity to social contingency: a "double video" study of face-to-face communication between 2-and 4-month-olds and their mothers. *Infant Behavior and Development*, *27*, 195-203. doi:10.1016/j.infbeh.2003.09.004

Suda, M., Takei, Y., Aoyama, Y., Narita, K., Sakurai, N., Fukuda, M., & Mikuni, M. (2011). Autistic traits and brain activation during face-to-face conversations in typically developed adults. *PloS One*, *6*(5), e20021. doi:10.1371/journal.pone.0020021

Suda, M., Takei, Y., Aoyama, Y., Narita, K., Sato, T., Fukuda, M., & Mikuni, M. (2010). Frontopolar activation during face-to-face conversation: An in-situ study using near-infrared spectroscopy. *Neuropsychologia*, *48*(2), 441-447. doi:10.1016/j.neuropsychologia.2009.09.036

Takei, Y., Suda, M., Aoyama, Y., Yamaguchi, M., Sakurai, N., Narita, K., ... & Mikuni, M. (2013). Temporal lobe and inferior frontal gyrus dysfunction in patients with schizophrenia during face-to-face conversation: A near-infrared spectroscopy study. *Journal of Psychiatric Research*, *47*(11), 1581-1589. doi:10.1016/j.jpsychires.2013.07.029

Tanaka, J. W., & Sung, A. (in press). The "Eye Avoidance" Hypothesis of Autism Face Processing. *Journal of Autism and Developmental Disorders*. doi: 10.1007/s10803-013-1976-7

Tavares, P., Lawrence, A. D., & Barnard, P. J. (2008). Paying attention to social meaning: an fMRI study. *Cerebral Cortex*, *18*(8), 1876-1885. doi:10.1093/cercor/bhm212

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: attributed mental states influence perception. *Trends in Cognitive Sciences*, *14*(8), 376-382. doi:10.1016/j.tics.2010.05.005

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., & Poline, J. B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *Neuroimage, 35*(1), 105-120. doi:10.1016/j.neuroimage.2006.11.054

Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., & Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *Neuroimage*, *37*(3), 966-973. doi: 10.1016/j.neuroimage.2007.06.058

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675-690. doi:10.1017/S0140525X05000129

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242-249. doi:10.1016/j.psychres.2008.05.006

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, *9*(2), 335-340. doi:10.3758/BF03196290

Trevarthen, C. (1998). The concept and foundations of infant intersubjectivity. In S. Braten (Ed.), *Intersubjective communication and emotion in early ontogeny,* (pp. 15-46). New York: Cambridge University Press.

Tsai, J. C. C., Sebanz, N., & Knoblich, G. (2011). The GROOP effect: Groups mimic group actions. *Cognition*, *118*(1), 135-140. doi:10.1016/j.cognition.2010.10.007

Tylén, K., Allen, M., Hunter, B. K., & Roepstorff, A. (2012). Interaction vs. observation: distinctive modes of social cognition in human brain and behavior? A combined fMRI and eye-tracking study. *Frontiers in Human Neuroscience*, *6*(331)*, 1-11*. doi:10.3389/fnhum.2012.00331

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

Van Overwalle, F. (2011). A dissociation between social mentalizing and general reasoning. *Neuroimage*, *54*(2), 1589-1599. doi:10.1016/j.neuroimage.2010.09.043

Vander Wyk, B. C., Hudac, C. M., Carter, E. J., Sobel, D. M., & Pelphrey, K. A. (2009). Action understanding in the superior temporal sulcus region. *Psychological Science*, *20*(6), 771-777. doi: 10.1111/j.1467-9280.2009.02359.x

Verga, L., & Kotz, S. A. (2013). How relevant is social interaction in second language learning? *Frontiers in Human Neuroscience*, *7*(550), *1-7*. doi:10.3389/fnhum.2013.00550

Voracek, M., & Dressler, S. G. (2006). Lack of correlation between digit ratio (2D: 4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences*, *41*(8), 1481-1491. doi: 10.1016/j.paid.2006.06.009

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain*, *129*(4), 932-943. doi: 10.1093/brain/awl032

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*(3), 640-663. doi:10.1111/j.1467-8624.2006.00895.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*(3), 655-684. doi:10.1111/1467-8624.00304

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage*, *25*(3), 1002-1015. doi: 10.1016/j.neuroimage.2004.12.013

Yarkoni, T. (2009). Big correlations in little studies: Inflated fMRI correlations reflect low statistical power—Commentary on Vul et al. (2009). *Perspectives on Psychological Science*, *4*(3), 294-298. doi:10.1111/j.1745-6924.2009.01127.x

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences, 107*(15), 6753-6758. doi:10.1073/pnas.0914826107

Young, L., Dodell-Feder, D., & Saxe, R. (2010). What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia, 48*(9), 2658-2664. doi:10.1016/j.neuropsychologia.2010.05.012

Yurgelun-Todd, D. A., & Killgore, W. D. (2006). Fear-related activity in the prefrontal cortex increases with age during adolescence: a preliminary fMRI study. Neuroscience Letters, 406, 194-199. doi: 10.1016/j.neulet.2006.07.046

Zeeland, S. V., Ashley, A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A., & Bookheimer, S. Y. (2010). Reward processing in autism. *Autism Research*,*3*(2), 53-67. doi: 10.1002/aur.122