The expression of autism spectrum disorder (ASD) is highly heterogeneous, owing to the complex interactions between genes, the brain, and behavior throughout development. Here we present a model of ASD that implicates an early and initial failure to develop the specialized functions of one or more of the set of neuroanatomical structures involved in social information processing (i.e., the ‘social brain’). From this early and primary disruption, abnormal brain development is canalized because the individual with an ASD must develop in a highly social world without the specialized neural systems that would ordinarily allow him or her to partake in the fabric of social life, which is woven from the thread of opportunities for social reciprocity and the tools of social engagement. This brain canalization gives rise to other characteristic behavioral deficits in ASD including deficits in communication, restricted interests, and repetitive behaviors. We propose that focused efforts to explore the brain mechanisms underlying the core, pathognomonic deficits in the development of mechanisms for social engagement in ASD will greatly elucidate our understanding and treatment of this complex, devastating family of neurodevelopmental disorders. In particular, developmental studies (i.e., longitudinal studies of young children with and without ASD, as well as infants at increased risk for being identified with ASD) of the neural circuitry supporting key aspects of social information processing are likely to provide important insights into the underlying components of the full-syndrome of ASD. These studies could also contribute to the identification of developmental brain endophenotypes to facilitate genetic studies. The potential for this kind of approach is illustrated via examples of functional neuroimaging research from our own laboratory implicating the posterior superior temporal sulcus (STS) as a key player in the set of neural structures giving rise to ASD. **Keywords:** Social perception, social cognition, autism, functional neuroimaging, social brain.

The considerable heterogeneity in the expression and severity of the core and associated symptoms is a challenge that has hindered progress towards understanding autism spectrum disorder (ASD). To illustrate, within autistic disorder, variability in the social domain ranges from a near absence of interest in interacting with others to more subtle difficulties managing complex social interactions that require an understanding of other people’s goals and intentions and other cues of social context. Similarly, stereotyped and repetitive behaviors range from simple motor stereotypies and/or a preference for sameness to much more complex and elaborate rituals, accompanied by emotional dysregulation or ‘meltdowns’ when these rituals are interrupted. Some individuals with ASD lack basic speech abilities, while others can have language deficits that are mild and limited to language pragmatics. While a majority of individuals with ASD exhibit some level of intellectual impairment, intelligence quotients vary from the severe and profoundly impaired range to well above average.

Mirroring the phenotypic heterogeneity, there is also great genetic variability in ASD (Abrahams & Geschwind, 2008; Geschwind & Levitt, 2007; Gupta & State, 2007). For instance, many susceptibility loci have been identified, yet each is thought to account for only a small number (1–2%) of overall cases (e.g., Weiss et al., 2008). Various defined mutations, genetic syndromes, and de novo copy number variation account for as many as 10–20% of cases (Abrahams & Geschwind, 2008). Given the large number of potential genetic mechanisms, it is likely that no single molecular pathophysiological explanation will be sufficient to explain the majority of cases (Abrahams & Geschwind, 2008). The pace of discovery is accelerating toward the identification of a myriad of genetic mechanisms, all leading through a smaller but still substantial number of molecular pathophysiological mechanisms, to a set of neurodevelopmental disorders that comprise ASD.

**Constraining heterogeneity**

Reflecting upon the genetic and phenotypic heterogeneity, Geschwind and Levitt (2006) articulated an influential three-part hypothesis stating that:
(1) there are many ‘autisms’; (2) a common brain-systems-level feature across the autisms is the concept of ‘developmental disconnection’; (3) developmental disconnection manifests as a failure to develop normal connections between higher-order association areas of the temporal and parietal cortices and regions of the frontal cortices. This developmental disconnection or ‘underconnectivity’ component of the hypothesis builds from and complements other similar viewpoints (e.g., Brock, Brown, Boucher, & Rippon, 2002; Courchesne & Pierce, 2005; Hughes, 2007; Just, Cherkassky, Keller, & Minshew, 2004).

Functional neuroimaging studies of adolescents and adults with ASD have provided considerable evidence in support of underconnectivity in ASD (e.g., Castelli, Happé, Frith, & Frith, 2000; Just et al., 2004; Kana, Keller, Cherkassky, Minshew, & Just, 2006; Kana, Keller, Minshew, & Just, 2007; Kleinmans et al., 2008; Koshino et al., 2008; Koshino, Carpenter, Minshew, Cherkassky, Keller, & Just, 2005; Mason, Williams, Kana, Minshew, & Just, 2008; Monk et al., 2009; Noonan, Haist, & Müller, 2009; Villalobos, Mizuno, Dahl, Kemmotsu, & Müller, 2005; Wicker et al., 2008). Similarly, structural neuroimaging studies of older children, adolescents, and adults with ASD support the underconnectivity hypothesis (e.g., Barnea-Goraly et al., 2004; Sivarswamy et al., 2010; Fletcher et al., 2010; Sahyoun, Belliveau, Soulières, Schwartz, & Mody, 2010; Kumar et al., 2010; Cheung et al., 2009; Keller, Kana, & Just, 2007). Underconnectivity appears to be present in the brains of older children, adolescents, and adults with ASD. The appeal of the underconnectivity notion lies partly in the fact that it appears to offer a systems-level model of brain dysfunction that purports to account for the specific symptoms of ASD as well as the heterogeneity of etiology, behaviors and cognition (Geschwind & Levitt, 2006). However, it is not yet clear how the underconnectivity perspective accounts for the specific patterns of dysfunction in individuals with ASD. That is, how might this perspective explain what is common among individuals with ASD and what separates ASD from other neurodevelopmental disorders that also feature underconnectivity?

We believe that it is premature to conclude that underconnectivity reflects a primary, unifying neural-systems-level mechanism in ASD. There are several reasons for our caution. First, only a handful of fMRI studies have tested children with ASD (e.g., Lee et al., 2009; Brito et al., 2009). The results of these studies have been somewhat mixed, if not tending to run counter to key aspects of the disconnection hypothesis. For instance, using functional magnetic resonance imaging (fMRI), Lee and colleagues (2009) examined functional connectivity in data collected during a Go/No-go task in samples of 8- to 12-year-old children with and without an ASD. They focused on the connectivity values between the left and right inferior frontal cortices (IFC; BA 47) and regions of the respective frontal, striatal, and parietal cortices. The two groups of children did not differ in their functional connectivity. Intriguingly, in the ASD group, there was a significant negative correlation between age and two long-range IFC correlation pairs: the right IFC ↔ bilateral presupplementary motor area (BA 6) and right IFC ↔ right caudate. These findings indicate normal prefrontal cortical functional connectivity in school-age children with ASD, but also suggest that some functional connections may abnormally decrease with age in the children with ASD. In other words, the brains of these children become more ‘disconnected’ over time instead of starting that way.

Second, the specificity of reduced functional connectivity for ASD relative to other neurodevelopmental and neuropsychiatric disorders has not been demonstrated. Reduced functional connectivity, including reductions in the long-range frontal ↔ temporal cortical and frontal ↔ parietal cortical connections, have been reported in a variety of disparate neurological, neuropsychiatric, and neurodevelopmental disorders and conditions including: Alzheimer’s disease (e.g., Supekar, Menon, Rubin, Musen, & Greicius, 2008), schizophrenia (e.g., Esslinger et al., 2009; Friston & Frith, 1995), adolescent depression (Cullen et al., 2009), chronic heroin use (Liu et al., 2009), post-traumatic stress disorder (Shaw et al., 2001), and dyslexia (e.g., Richards & Berninger, 2008).

Germane to the point of specificity, one study of Romanian orphans reported prominent effects on brain connectivity from the experience of profound, early, and severe socioemotional deprivation. Eluvathingal and colleagues (2006) examined, using diffusion tensor imaging (DTI) tractography, the integrity of white matter tracts that connect limbic and paralimbic structures, including the orbital frontal gyrus, infralimbic prefrontal cortex, hippocampus/amygda, lateral temporal cortex, and the brainstem. These regions were selected a priori on the basis of a prior positron emission tomography (PET) study that identified glucose hypometabolism in each of these neuroanatomical structures in children with ASD (Chugani et al., 2001). The children exhibited a relatively mild specific cognitive impairment and impulsivity, but they did not have an ASD. Fractional anisotropy values in the left uncinate fasciculus (which connects the gyri of the frontal lobe with the anterior end of the temporal lobe) were decreased significantly in the early deprivation group compared with typically developing (TD) comparison children. This finding highlights the possibility that the observed findings of reduced long-range functional connectivity in adults with ASD are actually the result of the ASD. If individuals with an ASD lack the necessary, early developing mechanisms for social engagement that ensure normative social development, then opportunities for
social interaction are inherently reduced, particularly experiences sought out by the individual.

Our field has also overemphasized the heterogeneity in ASD, at the expense of recognizing the homogeneity of core disruptions in social information processing. While diverse with regard to phenotypic expression and genetic etiology, the ASDs share the common, pathognomonic feature of dysfunctional reciprocal social interaction. In this review we present recent evidence from our program of research of inherent and lifelong impairments in the structure and function of neuroanatomical systems that are related to the universal and pathognomonic ASD deficits in reciprocal social interaction. We recognize that these disruptions in brain structure and function may arise from a number of different genetic and molecular etiologies and are also further transformed across development by the experiences and activity of the individual in the world (Kandel, 1998). We argue that the reciprocal relationship between brain disruption and atypical social development drives homogeneity in the syndrome’s presentation even in the presence of enormous phenotypic and genotypic heterogeneity.

Our contention is that despite the etiological heterogeneity and despite such high phenotypic variability, it might be the case that the various factors contributing to the now well-characterized triad of impairments in ASD exert their effect through a circumscribed set of neural structures and their interconnections that both give rise to and are shaped by social development. That is, it is possible that the simplest and potentially most powerful marker of ASD will be found at the level of brain systems, particularly those regions devoted to social information processing, and their development (i.e., the dynamic interactions among genes, brain, and behavior over time).

We propose that ASD begins with a failure in the emergence of the specialized functions of one or more of the set of neuroanatomical structures involved in social information processing. This failure happens early in ontogeny, within the first nine months to one year of life, if not earlier. In turn, because the affected regions do not generate the normal stream of both intrinsic and stimulus-driven signals, the normal developmental pattern of connections among these brain regions is greatly altered. Abnormal brain development is canalized because the individual with an ASD must develop in a highly social world without the specialized neural systems that would ordinarily allow him or her to partake in the fabric of social life, which is woven from opportunities for social reciprocity. These opportunities provide the basis for normal cognitive development in typically developing children, but are greatly decreased in ASD.

Our proposal borrows from and builds upon contributions beginning even before neuroscientists recognized the existence of a ‘social brain’. In the late 1980s, researchers began to discuss an early form of what we now call the social motivation hypothesis (e.g., Sahley & Panksepp, 1987). In 1991, Fotheringham put forward an early form of the amygdala hypothesis of ASD, followed by Simon Baron-Cohen’s groundbreaking exposition on this hypothesis Baron-Cohen et al., 1999). More broadly, Baron-Cohen (1995) formulated a theory of social brain dysfunction in Mindblindness. At approximately the same time, Uta and Chris Frith integrated the available neuroscientific evidence to put forward a social brain hypothesis to explain theory of mind deficits in ASD (Frith & Frith, 1999). Indeed, one might argue that Leo Kanner captured the essence of our argument, without empirical data on the central nervous system, in his original descriptions of ASD, wherein he states: ‘We must, then, assume that these children have come into the world with innate inability to form the usual, biologically-provided affective contact with other people, just as other children come into the world with innate physical or intellectual handicaps’ (1943, p. 250). These are just a few of the ideas predating our own conceptualization, which have served to shape our ideas of abnormal social brain development in ASD.

The social brain and social perception

Social perception refers to the initial stages of evaluating the intentions and psychological dispositions of others by using their gaze direction, body movements, hand gestures, facial expressions, and other biological-motion cues (Allison, Puce, & McCarthy, 2000). Neuroscientists became deeply interested in social perception when it was discovered that neurons within the temporal cortex and amygdaloid complex of monkeys were sensitive to and selective for social objects (e.g., faces and hands) and complex social stimuli (actions in a social context and direction of gaze) (Brothers & Ring, 1993; Brothers, Ring, & Kling, 1990; Desimone, 1991; Desimone, Albright, Gross, & Bruce, 1984; Perrett, Rolls, & Caan, 1979, 1982; Perrett et al., 1984, 1985; Rolls, 1981, 1995). On the basis of these pioneering findings, the field began to think seriously about the possibility of a network of brain regions dedicated to processing social information. The label ‘the social brain’ was coined by Leslie Brothers (1990), and it served to elegantly capture the core, emerging idea. The social brain is defined as the complex network of areas that enable us to recognize other individuals and to evaluate their mental states (e.g., intentions, dispositions, desires, and beliefs). The key idea is that human beings, in response to the unique computational demands of their highly social environments, have evolved cognitive mechanisms and an associated, dedicated neural system that support such abilities as recognizing other agents and their actions, individuating others, perceiving the emotional states
of others, analyzing the intentions and dispositions of others, sharing attention with one another, and representing another person’s perceptions and beliefs.

In an early and influential model of the social brain, Brothers (1990) emphasized the contributions of the superior temporal sulcus (STS), amygdala, orbital frontal cortex (OFC), and fusiform gyrus (FFG) to social perception. These four neuroanatomical structures are illustrated in Figure 1. In humans, the STS region, particularly the posterior STS in the right hemisphere, analyzes biological motion cues, including eye, hand, and other body movements, to interpret and predict the actions and intentions of others (e.g., Bonda, Petrides, Ostry, & Evans, 1996; Pelphrey, Morris, Michielich, Allison, & McCarthy, 2005). The FFG, located in the ventral occipitotemporal cortex, contains a region termed the fusiform face area (FFA), which has been implicated in face detection (identifying a face as a face) and face recognition (identifying one’s friend versus a stranger) (e.g., Kanwisher, McDermott, & Chun, 1997; Puce, Allison, Asgari, Gore, & McCarthy, 1996). The OFC has been strongly implicated in social reinforcement and reward processes more broadly (e.g., Rolls, 2000, 2009). Finally, the amygdala, a complex structure that is highly interconnected with cortical and other subcortical brain structures, has been implicated in helping to recognize the emotional states of others through analysis of facial expressions (e.g., Morris et al., 1996), as well as in multiple aspects of the experience and regulation of emotion (e.g., Davis & Whalen, 2001; Klüver & Bucy, 1997; LeDoux, 2000).

To understand social brain function, we must be as attentive to the interconnections of neuroanatomical structures as we are to their individual contributions. Currently, in humans, much is known about the roles played by the individual brain regions, but very little is known about the ways in which they are interconnected, and thus even less is known about how they interact functionally. However, we can take some initial guidance from the monkey brain; where it is known that the STS region has reciprocal connections to the amygdala (Amaral, Price, Pitkanen, & Carmichael, 1992). The amygdala, in turn, is connected to the OFC (Amaral et al., 1992), and the STS is also connected with the OFC (Barbas, 1988). The OFC is connected to prefrontal cortex (Pandya & Yeterian, 1996), which is connected to motor cortex and the basal ganglia, thus completing what Allison and colleagues (2000) described as a pathway from perception to action.

The role of the STS region in social perception and its dysfunction in ASD

Representing actions

In an initial fMRI study of typically developing young adults, we compared the response from the posterior STS to four different types of motion conveyed via animated virtual-reality characters (Pelphrey, Mitchell, et al., 2003). Participants viewed walking, a biological motion conveyed by a robot or a human. They also viewed a nonmeaningful but complex nonbiological motion in the form of a disjointed mechanical figure as well as a complex, meaningful, and nameable nonbiological motion involving the movements of a grandfather clock. We reasoned that a region selectively responsive to biological motion should respond strongly to both the man and the robot walking, but not respond to the mechanical figure or the grandfather clock.

We found strong and equivalent activity in the right hemisphere posterior STS to the human and robot walking and very little activity to the moving clock and the mechanical figure. This pattern of results was very different from that observed in the nearby motion-responsive visual area called MT or V5 (Zeki et al., 1991), which responded robustly to all four of our stimulus conditions. We concluded that the posterior STS selectively processes biological motion. This finding led us to begin to view the posterior STS as one component of the neural system supporting social perception, via its identification and representation of observed human actions as compared to the movements of other objects.

Perceiving intentions

Having established a ‘baseline’ role for the posterior STS region in social perception, we went on to examine whether this region simply serves as a biological motion detector or if it is involved more broadly in aspects of social perception through the evaluation of the intentions and dispositions that are conveyed by biological motions. For example, we asked if this region might be involved in representing...
another person’s intentions with respect to the objects in his or her visual field (Vander Wyk, Hudac, Carter, Sobel, & Pelphrey, 2009). This kind of intention understanding involves integrating actions with the social and physical context and it is possible that the STS may show regional sensitivity to environmental or cognitive sources of information that inform this context. An especially salient component of social context is the perceived emotions of other individuals, particularly when a perceived emotion indicates another person’s like or dislike of an object or event. We examined whether the posterior STS exhibits differences in activity depending on the previous emotional context related to understanding another’s preferences, which would thus inform the viewer about that person’s underlying intentions. This experiment used a paradigm adapted from a study of young children by Phillips et al. (2002), in which the participants observed an actress on video express positive or negative regard towards one of two cups. She then reached and picked up that same cup or the other one. Viewing the actress’s emotional expression allowed the participant to attribute an intention to her: positive expressions toward an object warranted the attribution of the intent to pick up the object, while negative expressions warranted the opposite attribution. The subsequent reaching gesture was then interpreted as being either congruent or incongruent with the intention. As illustrated in Figure 2, by crossing emotional expression and congruency of the action, we created four experimental conditions: Positive-Incongruent, Negative-Incongruent, Positive-Congruent, and Negative-Congruent.

Our results indicated that the right posterior STS exhibited significantly more activity for the incongruent trials than for the congruent trials. This showed that the activity of the right posterior STS to a given biological motion is sensitive to prior emotional context. Specifically, the posterior STS showed a greater response when participants viewed a reach that was incongruent with a prior emotional expression (i.e., when the actress reached for the object not targeted by a prior positive expression or when she reached for the object targeted by a prior negative expression) than when they viewed a reach congruent with the actresses’ expression. In both cases in which positive and negative emotions set up these expectations about the actress’s future behavior towards an object, the response in the posterior STS was greater upon viewing unanticipated actions.

We concluded that the STS does not merely represent the surface features of biological motion, but also participates in integrating biological motion, or actions, with the social and/or physical context, and thereby plays a role in analyzing and interpreting the intentions underlying observed biological motions. This view of the functional role of the posterior STS is further bolstered by additional findings regarding a role for the posterior STS in analyzing the intentions and dispositions of others as conveyed by biological motion (e.g., Brass, Schmitt, Spengler, & Gergely, 2007; Castelli et al., 2000; Pelphrey, Singerman, et al., 2003; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004).

Prior to our study by Vander Wyk and colleagues (2009), potential limitations in the experimental designs employed by our group and others left the findings open to an alternative interpretation. Corbetta, Patel, and Shulman (2008) argued that the prior results could be more parsimoniously explained by uncontrolled stimulus differences in demands for shifting attention. After all, the posterior STS region is known to be involved in the allocation of spatial attention in response to visual cues (e.g., Hopfinger, Buonocore, & Mangun, 2000). When we scrutinized our own study designs (e.g., Pelphrey, Singerman, et al., 2003) we recognized that our conditions could plausibly differ in the number of attention shifts evoked. However, in the Vander Wyk et al. (2009) study described above, by using both positive and negative emotions directed toward an

![Figure 2](https://example.com/figure2.png)

Figure 2 Schematic illustration of the four conditions in the experiment. In the positive conditions, the actress expressed positive regard toward one of the cups, and in the negative conditions, she expressed negative regard toward one of the cups. In the congruent conditions, the actress’s reach was expected, given her previous expression. In the incongruent conditions, her reach was unexpected, given her previous expression. The figure is borrowed with permission from Vander Wyk et al., (2009) *Psychological Science*
object, we ensured that the number of attention shifts was fully balanced between congruent and incongruent trials. Our results supported an attention understanding as opposed to an attention-shifting hypothesis. That is, the pattern of findings we observed in this study fully contradicted the argument that increased activity in the posterior STS region while viewing unexpected actions in earlier findings was related to additional shifts in attention required when observed movements differ from expectations (Corbetta et al., 2008). By our estimation, the attention-reorienting account would predict greater activity for the Positive-Incongruent than for the Negative-Incongruent condition, because the former requires the participant to reorient attention twice and the latter only once. Likewise, the Positive-Congruent condition would be expected to involve only one shift in attention, while the Negative-Congruent condition should demand two shifts in attention (i.e., shifting attention to the cup at which the actress directed negative regard and then shifting attention to the cup picked up by the actress). This attention account would therefore predict equivalent activation in the Positive-Incongruent and Negative-Congruent conditions and equivalent activity in the Positive-Congruent and Negative-Incongruent conditions. We observed a very different pattern of effects: within the right posterior STS, the incongruent condition was greater than the congruent condition, regardless of the affect, a pattern of effects that an attention-reorienting account cannot explain.

**Failing to read intentions: a role for STS dysfunction in ASD**

A large body of research demonstrates that individuals with ASD exhibit characteristic and early appearing deficits in the use of gaze information to understand the intentions and mental states of others, as well as to coordinate joint attention (Baron-Cohen et al., 1999; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998; Frith & Frith, 1999; Leekam, Hunnissett, & Moore, 1998; Leekam, López, & Moore, 2000; Loveland & Landry, 1986; Mundy, Sigman, Ungerer, & Sherman, 1986). This failure to understand the mentalistic significance of eye gaze and the focus of another person’s attention motivated a recent study of posterior STS function in individuals with ASD.

Adolescent and young adult participants with (N = 12) and without (N = 14) high-functioning ASD viewed actions that were congruent or incongruent with expectations given positive or negative emotional content. The two participant groups were matched on IQ, age, and gender. The individuals with ASD (not pervasive developmental disorder – not otherwise specified (PDD-NOS) or Asperger’s syndrome) were evaluated by an expert clinician and met gold-standard research criteria using the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) and the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994). Neurotypical participants scored below the threshold for ASD using the Autism Screening Questionnaire (ASQ; Berument, Rutter, Lord, Pickles, & Bailey, 1999) or the ADOS. This experiment made use of the identical reach-to-cups paradigm as our prior study (Vander Wyk, et al., 2009). To recap, as illustrated in Figure 2, participants were instructed to watch videos in which an actress was seated in front of a red cup and a green cup. At the start of a trial, the actress shifted her head and gaze to show preference to a cup by either smiling (Positive regard) or frowning (Negative regard), followed by a return to a neutral expression and eye contact with the camera. She maintained this pose while reaching towards, lifting, and setting down either cup to end the trial. Thus, given the initial preference and whether the actress chose to reach toward and pick up the ignored or attended cup, four conditions were created: (1) a Positive-Congruent condition where the actress expressed a positive emotion towards an object and then reached for that object; (2) a Positive-Incongruent condition, in which the actress reached for the ignored object; (3) a Negative-Congruent conditions where the actress expressed a negative emotion and reached to the ignored object; and (4) a Negative-Incongruent condition, in which the actress expressed a negative emotion to an object, but then reached to that object.

Based on our prior work, we hypothesized that the right posterior STS would not differentiate incongruent and congruent actions in ASD (Pelphrey, Morris, et al., 2005). As illustrated in the right panel of Figure 3, the new sample of typically developing participants demonstrated a strong effect of congruency (Incongruent > Congruent) within the right posterior STS. This area of activation nicely overlaps with the previously published reference region, clearly replicating our prior findings (Vander Wyk et al., 2009). In sharp contrast, individuals with ASD failed to exhibit Incongruent > Congruent activation in the posterior STS region. As shown in the left panel of Figure 3, participants with ASD responded with significant activation to the two stimulus categories; however, the activation levels did not differ as a function of congruency.

Consistent with our findings, other neuroimaging studies have revealed dysfunction of the STS in ASD during tasks involving eye movement perception (Pelphrey et al., 2005), the attribution of intentions to moving geometric figures (Castelli et al., 2002), and human speech perception (Boddaert et al., 2003; Gervais et al., 2004). Bilateral hypoperfusion of temporal lobe areas at rest has been observed in children with ASD (Ohnishi et al., 2000; Zibovicvics et al., 2000). A positron emission tomography (PET) study of speech perception reported abnormal
laterality of responses and hypoactivation of the left superior temporal gyrus (Boddaert et al., 2003), and an fMRI study observed abnormal responses in the STS to human voices (Gervais et al., 2004). Finally, a study comparing cortical sulcal maps in individuals with and without ASD found anterior and superior displacements of the STS (Levitt et al., 2003), and Boddaert and colleagues (2004) reported abnormal STS volumes in ASD.

Amygdala and fusiform gyrus contributions to disrupted face perception in ASD

In comparison to typically developing participants, hypoactivation of the amygdala (e.g., Baron-Cohen et al., 1999; Ogai et al., 2003; Pelphrey et al., 2007) and the FFG (e.g., Critchley et al., 2000; Pelphrey et al., 2007; Schultz et al., 2000) have often been observed in people with ASD relative to typically developing participants. However, the question of whether or not hypoactivation in these regions is an aspect of the brain phenotype in ASD remains open and widely debated. Indeed, critical questions have been raised regarding the mechanisms underlying the amygdala. In particular, to what degree is the hypoactivation accounted for by known differences in the visual scanpaths exhibited by individuals with and without autism in response to faces?

In a recent study (Perlman, Hudac, Pegors, Minshew, & Pelphrey, 2010), we experimentally manipulated activity in the face processing system of individuals with ASD by compelling them to look at the eyes of a face to varying degrees. This allowed us to determine whether hypoactivation in the amygdala and FFG can be accounted for by known differences in the visual scanpaths exhibited by individuals with and without autism in response to faces. We directly manipulated visual scanpaths across four experimental conditions including free viewing, central fixation, low, medium, or high amounts of fixating on the eyes in a group of high-functioning adults with autistic disorder and a matched (on IQ, age, and sex) group of typically developing participants during fMRI.

Comparison of activation during free viewing revealed hypoactivation of the right FFG and bilateral amygdala in the ASD group relative to typically developing participants. As illustrated in the top panel of Figure 4, activity in the right FFG of individuals with ASD increased from a slightly negative response during central fixation to a robustly
positive response for each of the eye fixation conditions. In contrast, central fixation reduced activity in typically developing participants to below-zero levels, suggesting an inhibition of activity in the FFG when typically developing individuals have their scanpaths artificially constrained. Within the amygdala, there was no effect of the scanpath manipulation for individuals with autism and the scanpath manipulation reduced amygdala activation in typically developing participants (see Figure 4, bottom panel) relative to the free viewing condition. Our findings concerning increases in FFG activity when participants with ASD are compelled to look at the eyes were generally consistent with the two prior studies that, using varying methodology, constrained fixation to a central crosshair on a face (Hadjikhani et al., 2004; Pierce, Haist, Sedaghat, & Courchesne, 2004). However, we demonstrated that activity in the right FFG of individuals with ASD increases from a below-zero level during the central fixation condition to robustly positive levels during execution of a scanpath involving eye contact. By simply manipulating visual attention to the eyes, we were able to ‘normalize’ activity in one component of the face processing system in individuals with ASD. Furthermore, our findings from the free viewing condition (Figure 4, top panel) reveal that the effect is not attributable to decreases in activity in the FFG of our typically developing participants as a function of artificially constraining their visual scanpaths during the low, medium, and high eye fixation conditions. It is noteworthy that when the fixation patterns were constrained to the center of the face during our Central Fixation condition, as they were constrained to the center of the eyes in the study by Hadjikhani et al. (2004) and to the center of the screen by Pierce et al. (2004), we found reduced activity to the point of significant deactivation in TD participants. In contrast to the FFG, our manipulation of eye contact had no effect on amygdala activation in our group of participants with ASD. In typically developing individuals, constraining eye movements to the center of the screen or directing eye movements to the eyes of the stimulus face served to significantly decrease amygdala activity (bottom panel of Figure 4). This observation is consistent with previous research conceptualizing the role of the amygdala as directing attention towards salient environmental stimuli such as the eyes of faces (e.g., Adolphs et al., 2005). It is possible that we do not observe amygdala activity when we constrain the eye movements of typically developing participants because we are circumventing the need for the amygdala to direct eye movements in response to the presentation of an emotionally expressive face. In the case of ASD, the lack of amygdala activity during free...
viewing, combined with the observation that this hypoactivation cannot be reversed via the manipulation of eye movements, suggests that disruption in this brain mechanism for directing eye movements to the key, most socially relevant features of a face is a core feature of the brain phenotype in autism.

Comparing ASD and schizophrenia: Overlapping and distinct social brain dysfunction

We suggested earlier the necessity of demonstrating specificity for ASD relative to other neurodevelopmental disorders in construing theories of brain system dysfunction. To this end, Pinkham and her colleagues (Pinkham, Hopfinger, Pelphrey, Piven, & Penn, 2008) conducted a study that directly compared age- and IQ-matched adults with schizophrenia (paranoid and non-paranoid), adults with ASD, and typically developing adults. ASD and schizophrenia were compared because individuals with either of these very distinct neurodevelopmental disorders present with some overlapping deficits in aspects of social cognition and social behavior. A trustworthiness task, developed by Adolphs and his colleagues (1998), was used. The task required participants to look at a series of gray-scale frontal images of faces and to perform a forced-choice rating as trustworthy or untrustworthy. As a control task, participants viewed the same kind of faces, but judged age, classifying faces as ‘30 years of age or younger’ or ‘over 30 years of age’. The design was partially modeled after an fMRI study by Winston and colleagues (Winston, Strange, O’Doherty, & Dolan, 2002) that found explicit trustworthiness judgments to robustly activate key components of the social brain including the right posterior STS, bilateral VLPFC/AI, amygdala, and lateral FFG.

The behavioral results indicated that the groups did not differ on the trustworthiness or age judgments in terms of the percentages rated as trustworthy/over 30 years of age, nor did they differ in reaction times. Each group exhibited activation of key, a priori selected regions of the social brain, including the bilateral amygdala, lateral FFG, and VLPFC/AI. The ASD group and the paranoid schizophrenia group exhibited significantly reduced activity in the right amygdala, FFG, and left VLPFC/AI as compared to the typically developing adults, and in the left VLPFC/AI as compared to non-paranoid schizophrenics. Strikingly, activity in the right posterior STS distinguished the participants with ASD from the three other participant groups. That is, during trustworthiness judgments, activation of the right posterior STS was observed for the typically developing participants, the non-paranoid schizophrenic group, and the paranoid schizophrenic group, but not in the group of participants with ASD. This suggests that dysfunction in this region might represent a ‘neural signature’ of ASD.

On the one hand, the finding of overlapping areas of dysfunction in three components of the social brain (the amygdala, the VLPFC/AI, and the lateral FFG) across two very different neurodevelopmental disorders suggests that social cognition deficits are the result of specific neurofunctional impairments in the social brain. These deficits appear to be deficit-specific as opposed to disorder-specific (Pinkham et al., 2008). On the other hand, the finding that activity in the right posterior STS differentiates individuals with ASD from individuals with schizophrenia suggests that this aspect of social brain dysfunction is disorder specific. Note, too, that there was no reduced activity in any brain area in any of the clinical groups relative to the typically developing participants during age judgments, demonstrating that the reduced activity observed for social judgments cannot be attributed to overall reductions in activation, a failure to view stimuli, or a failure to engage in the task.

An emerging developmental perspective on the social brain

The work reviewed above indicates that the posterior STS in typically developing young adults is highly specialized for detecting biological motion and interpreting the actions and intentions of others. This region is not specialized for these functions in young adults with ASD. A fundamental question remains to be addressed: How does the posterior STS become specialized across development for processing biological motion and intentional actions, and when does this process go astray in ASD? Answering this question would constrain distinct, but not mutually exclusive, interpretations of findings of posterior STS dysfunction in young adults with ASD. One interpretation is that very early disruptions in the posterior STS may drive the deficits in social perception, characteristic of ASD. Alternatively, posterior STS dysfunction in ASD may emerge as a consequence of the failure to follow the course of normative social development. From early on, children with ASD demonstrate deficits in basic mechanisms that predispose the allocation of attention to socially relevant signals that facilitate engagement with others (Jones & Klin, 2009). For example, while infants as young as two days old preferentially attend to point-light displays of biological motion (Simion et al., 2008), 2-year-olds with ASD fail to do so (Klin, Lin, Gorrindo, Ramsay, & Jones, 2009). Derailment of these basic social mechanisms early in life likely has cascading effects on subsequent

---

1 Participants were young adults, and thus it was unlikely that they would be familiar with Jack Weinberger’s suggestion: ‘Don’t trust anyone over thirty.’
development, causing increasing divergence in the processes impacting brain development (Jones & Klin, 2009). Given what is known from the behavioral literature regarding the early development of key aspects of social perception, determining precisely how the specialization of the posterior STS emerges will require longitudinal studies of infants in the first two years of life and methodological approaches that allow for the consideration and measurement of the impact of genes and experience (and their interaction) on the developing social brain.

At present, our knowledge of early typical and atypical development of social brain structures such as the posterior STS remains limited, in part due to technological and practical limitations. Imaging data is difficult to acquire for infants and young children who lack the ability to remain still while awake. Despite these limitations, studies have begun to elucidate the emergence of the specialization of the posterior STS for detecting agents and evaluating intentions.

**Detecting agents**

Carter and Pelphrey (2006) investigated the development of social brain regions involved in the perception of biological motion in 7- to 10-year-old typically developing children using fMRI. A key aim of the study was to test the specificity of the response to biological motion in the posterior STS by presenting participants with four sets of stimuli described earlier: a walking human, a walking robot, nonmeaningful but complex nonbiological motion (a disjointed mechanical figure), and meaningful, complex, and nameable nonbiological motion (a grandfather clock). As in adults (Pelphrey, Mitchell, et al., 2003), blood-oxygen-level dependence (BOLD) activity in the posterior STS of this sample of school-age children clearly differentiated biological from nonbiological motion. However, individual differences in the magnitude of this differentiation were apparent. A significant positive correlation was found between the magnitude of differentiation between biological and nonbiological motion in the right posterior STS and age. A recent functional near-infrared spectroscopy study examined whether 5-month-old infants show a differential hemodynamic response over the posterior temporal lobe (a region encompassing the posterior STS) in response to dynamic biological motion stimuli (an actress moving her hands, eyes and mouth) compared with dynamic nonbiological motion (moving machinery) (Lloyd-Fox et al., 2009). Significant hemodynamic changes were observed in response to the dynamic biological motion stimuli in both left and right posterior sensors. Additionally, more activation was observed in response to biological compared with nonbiological motion, suggesting that in 5-month-old infants there are already areas of the temporal lobe, located roughly in an area corresponding to the adult posterior STS, dedicated to the perception of biological motion. Together these studies suggest early specialization of a cortical network for the perception of biological motion, which becomes increasingly fine-tuned throughout development.

**Reading intentions**

Mosconi and colleagues (Mosconi, Mack, McCarthy, & Pelphrey, 2005) used fMRI to examine the neural circuitry underlying eye-gaze processing and intention understanding in 7- to 10-year-old typically developing children. Children viewed an animated actor who shifted her eyes toward a target object (a congruent gaze shift) or towards empty space (an incongruent gaze shift). Consistent with prior adult studies, the posterior STS was sensitive to the intentions underlying the actor’s eye movements, suggesting that by 7 years of age the neural circuitry underlying the detection of intentions through eye-gaze shifts may already be specialized for this particular social perception task. A similar study, using the same paradigm, was conducted in both adults and 9-month-old infants using event-related potential (ERP) measurements (Senju et al., 2006). In adults, incongruent gaze shifts elicited larger amplitudes at occipitotemporal sites (N330). A similar posterior component (N290) was observed in infants in response to incongruent gaze shifts. An additional frontal component (N200), showing higher amplitude in response to congruent gaze shifts, was observed in infants but not in adults. While these results suggest that infants and adults recruit a similar cortical network to encode intentions of others, the more widespread activation in infants may imply less specialization early in development.

These studies represent initial advances in developmental social neuroscience. However, much work needs to be done in order to advance our understanding of the emerging functional specialization of the posterior STS and other areas of the social brain network. First, although infant studies demonstrate neural activity in response to biological motion versus mechanical motion, the exact degree of specialization of this response remains unclear (Johnson et al., 2009). Second, while we have some understanding of brain development at the level of specific cortical regions (e.g., the posterior STS), few infant studies have examined how networks of brain regions become specialized for supporting social perception. Finally, studies have not yet addressed the disruption of neural systems for social perception in ASD. Because ASD are typically diagnosed after the second year of life, data on infants with ASD younger than two years of age are rare. However, research programs investigating infants at a high genetic risk for ASD may provide the opportunity for longitudinal studies of the development of cortical circuitry supporting social perception, an important step for developing optimal treatments for ASD early in life.

© 2011 The Authors

Conclusions

Given the nature of neurodevelopmental disorders, the expression of ASD is expected to be heterogeneous, particularly due to the pivotal interactions between gene, brain, and behavior throughout development (Cicchetti & Cohen, 2006). Focused efforts to explore the genetic and neural mechanisms underlying the core, homogeneous components of ASD and the deficits in tools for social engagement are necessary to elucidate this complex, devastating family of neurodevelopmental disorders. Future work in this area should focus on identifying the earliest departures from the typical development of the social brain and follow the complex interactions among genes, brain, and behavior that drive and constrain the atypical development of the social brain in ASD. Such work holds remarkable potential for informing the development of rational approaches to intervention and treatment of ASD.

Acknowledgements

The research reviewed herein was supported by grants from the Simons Foundation, the National Institute of Mental Health, the National Institute of Child Health and Development, the John Merck Scholars Fund, the US Veterans Affairs Administration, and Autism Speaks. Kevin Pelphrey was supported by a Career Development Award from the National Institutes of Health (NIMH Grant MH071284). We gratefully acknowledge our collaborators, especially Gregory McCarthy, James Morris, and Truett Allison. We are deeply grateful to the families, adults, and children who participate in our research. We confirm that all of the research from our laboratory presented in this review meets the ethical guidelines, including adherence to the legal requirements, of the USA. Informed consent was obtained from each of the study participants or their parents.

Correspondence to

Kevin Pelphrey, Harris Associate Professor, Child Psychiatry and Psychology, Yale Child Study Center, Yale School of Medicine, 230 South Frontage Rd., New Haven, CT 06520, USA; Email: kevin.pelphrey@yale.edu

Key points

• We present a model of autism spectrum disorders that implicates an early and initial failure to develop the specialized functions of one or more of the set of neuroanatomical structures involved in social information processing. From this early and primary disruption, abnormal brain development is canalized. This brain canalization gives rise to other characteristic behavioral deficits in autism spectrum disorders including deficits in communication, restricted interests, and repetitive behaviors.
• We illustrate how focused efforts to explore the brain mechanisms underlying the core, pathognomonic deficits in the development of mechanisms for social engagement in autism spectrum disorders will greatly elucidate our understanding and treatment of this complex, devastating family of neurodevelopmental disorders.

References


Jones, W., & Klin, A. (2009). Heterogeneity and homogeneity across the autism spectrum: The role of...
study of eye, mouth, and hand movements. Cerebral Cortex, 15, 1866–1876.


Manuscript accepted 21 September 2010

© 2011 The Authors