Abstract: Researchers have hypothesized that autistics are missing core modules of the brain, critical neural tissue necessary for accomplishing various processes. In this article, we critically review the evidence supporting two such hypothesized deficits. We ask whether autistic brains lack a module for understanding the behavior of others (i.e., theory of mind) and whether they lack a module for processing faces. We illustrate that successful performance on theory of mind tasks depends on linguistic ability; therefore, it is not surprising that autistics are more likely to fail theory of mind tasks because a qualitative impairment in communication is one of the primary diagnostic criteria for autism. Similarly, we illustrate that autistics are less likely to fixate the eye region of facial photographs and that the amount of time spent fixating the eye region correlates with activation in the face processing “module”; therefore, it is not surprising that autistics are less likely to activate the putative face processing area. These illustrations cast doubt on the arguments that the autistic brain is missing the core modules responsible for understanding theory of mind and for processing faces.

During the 1990s, which were deemed the “Decade of the Brain,” less invasive techniques for imaging the human brain were developed. For centuries before, we humans had wanted to look inside our heads, to see what makes us tick, and to identify commonalities that unite us and differences that distinguish us. Therefore, starting in the late 1990’s, these less invasive brain imaging techniques began to be applied to numerous special populations, including individuals with autism. The application of less-invasive brain imaging techniques to special populations has allowed testing of previously conceived hypotheses. In this article we review critically two such hypotheses: Whether the brains of autistics lack two fundamental brain modules, one that controls the ability to mentalize about other people, which is known as Theory of Mind, and one that allows the ability to recognize faces.
Uta Frith, John Morton, and Alan Leslie (1991) have boldly proposed that “what all people with autism have in common is a particular cognitive deficit that gives rise to the core symptoms in the course of development” (p. 434). Frith et al.’s (1991) proposed cognitive deficit was “the development of the theory of mind, or mentalizing,” which is the “ability to predict and explain the behaviour of other humans in terms of their mental states” (p. 434). According to Frith et al. (1991) “the ability to mentalize is dependent on a specific mechanism that does not manifest itself from birth; neither can it be explained by learning” (p. 434).

Thus, this core deficit—the ability to mentalize about others, called a Theory of Mind—was believed to be (a) universal among all people with autism; (b) innate, neither manifested at birth nor learnable; and (c) biological, dependent on a specific neural mechanism (Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998). Let us examine each of these three claims by first reviewing the evidence for what it means—in research terms—to lack a theory of mind.

In 1985, Baron-Cohen, Leslie and Frith were the first to ask the question of whether autistic children have a Theory of Mind (Baron-Cohen, Leslie, & Frith, 1985). These researchers answered their research article’s title question negatively; they concluded that children with autism have neither a theory of their own mind nor a theory of other people’s minds. Since the publication of Baron-Cohen et al.’s (1985) landmark paper, no fewer than one hundred research articles have asked the same question. Although the answer has not always been as resoundingly negative as that claimed by Baron-Cohen et al. (1985), the idea that persons with autism lack a theory of mind has nonetheless become integrated throughout the academic and professional literature and has pervaded our society’s collective knowledge.

For example, a few years ago, when Newsweek magazine focused its cover article on autism, it ran the following byline, “Why more kids and families are facing the challenge of mindblindness” (Cowley, Foote, & Tesoriero, 2000). The assumed importance of the ability to attribute mental states to one’s self and to others, and the perceived pervasiveness of the lack of this ability among persons with autism was also captured in a quote by a well-known autism researcher in a 2002 USA Today article: “It’s as if they [persons with autism] do not understand or are missing a core aspect of what it is to be human” (Falcon & Shoop, 2002). But what is the empirical evidence that persons with autism lack a theory of mind?

The original Baron-Cohen et al. (1985) paper and the vast majority of its successors used as an experimental task that is considered the classic assay of Theory of Mind: the “false belief task.” For example, in a false contents belief task, a research participant is shown a common container, such as a box that typically holds a particular brand of candy, and the research participant is asked to predict what is inside. Then, the research participant is shown that the contents do not fit the expectations; for example, the experimenter pulls pencils rather than candy out of the box. After these “false contents” are exposed to the research participant, he is asked to predict what he thought the contents would have been prior to the false contents being exposed (e.g., “What did you think was inside the box before I opened it?”). If the research participant identifies the actual content of the container (e.g., pencils) rather than the expected content (e.g., candy), then he fails the first phase of the false belief
task. Failing the first phase of a false contents task reputedly demonstrates that the individual lacks a theory of his own mind.

In the second phase of a false belief task, a fictional or real person is introduced who is presumably not privy to the exposure of the false contents. The research participant is then asked to predict what this other person would think the contents would be prior to the false contents being exposed (e.g., “What do you think that Jamie will think is inside the box before I open it?”). If the research participant again identifies the actual content of the container (e.g., pencils) rather than the expected content (e.g., candy), then he fails the second phase of the false contents belief tasks. Failing the second phase of a false contents task reputedly demonstrates that the individual lacks a theory of another person’s mind. Thus, performance on the false contents belief task hinges on the research participant’s ability to answer two critical questions: “What did you think was inside the box before I opened it?” and “What do you think [another person] will think is inside the box before I open it?”

Prior to its use in the autism literature, the false belief task was used predominantly with preschool children who demonstrated that before age four, they typically do not answer both questions correctly, but sometime after age four, they typically do (Astington & Gopnik, 1991; Perner, 1991; Wellman, 1990; Wellman, Cross, & Watson, 2001; Wellman & Lagattuta, 2000; Wimmer & Perner, 1983). Because it is around age four that typically developing children begin saying mentalizing expressions such as “think that,” “know that,” and “believe that,” it is presumed that around age four typically developing children become aware of their own minds and the minds of others (Bartsch, 1995; Bartsch & Wellman, 1997).

Is Theory of Mind a Universal Deficit in Autism?

Some theorists, most notably Baron-Cohen, believe that a lack of Theory of Mind is the core deficit in autism (Baron-Cohen, 1995). However, even in the original Baron-Cohen et al. (1985) investigation, only 80% (16 out of 20) of the autistic children failed the false belief task; 20% of the autistic children passed the false belief task, and therefore 20% presumably demonstrated that their theory of mind was intact. Other autism researchers have argued that such data demonstrate that theory of mind deficits are not universal in autism (Happe, 1995; Ozonoff, Rogers, & Pennington, 1991).

In a further study, Baron-Cohen (1989) presented a more complex theory of mind task, what is called a second-order false belief task, in which a second individual’s beliefs are queried, for example, “What will Jamie think that Mary thinks is inside the box before I open it?” In this case, only 10% of the autistic children passed the false-belief task. However, other researchers have found success rates in this task ranging up to 50%, particularly when adults are tested (Tager-Flusberg and Sullivan, 1994), leading one group of researchers to draw the rather circular conclusion that “[p]eople with autism have a selective theory of mind (ToM) deficit. . . . Traditional ToM tests . . . are not subtle enough to detect deficits in adults of normal intelligence. . . . More subtle tests . . . are needed” (Rutherford, Baron-Cohen, & Wheelwright, 2002, p. 189). Rather than continue around that circle, one can ask whether individuals with other clinical diagnoses fail theory of mind tests. They do.
Numerous populations have been observed to fail tests of theory of mind, such as false belief tasks, including deaf children (Peterson & Siegal, 1995), blind children (Tager-Flusberg, 2001), non-autistic children and adolescents with mental retardation (Benson, Abbeduto, Short, Nuccio, & Maas, 1993), minimally verbal children with Cerebral Palsy (Dahlgreen, 2002), children with Down’s Syndrome (Zelazo, Burack, Benedetto & Frye, 1996), Parkinson’s patients (Saltzman, Strauss, Hunter, & Archibald, 2000), frontal lobe patients (Rowe, Bullock, Polkey, & Morris, 2001), and, rather curiously, children with specific language impairment (Miller, 2001). Children with specific language impairment have—by diagnostic definition—no disabilities in social or emotional processes and must score in the average range on every other measure of cognitive function save language skill. It is only their language that is impaired, hence the name, specific language impairment. So, why should children with specific language impairment appear to lack a theory of mind?

Recall the two key questions asked during the false belief task. The syntactic form of these two questions is one of the most complex in the English language. These sentences exhibit sentential complement constructions, in which a complement clause is embedded in the matrix clause. Indeed, all mentalizing statements require sentence complement constructions, which are some of the most complex syntactic structures in the English language.

Does performance on false belief tasks within the general population depend on linguistic sophistication? Correlational studies document significant correlations between language comprehension measures and performance on false belief tasks (Cutting & Dunn, 1999; Hughes & Dunn, 1997; Jenkins & Astington, 1996). Cross-linguistic studies, that is, studies comparing across different languages, document that children acquiring languages in which the analog of the English sentential complement structure is acquired earlier versus later demonstrate earlier versus later success on false belief tasks (de Villiers & de Villiers, 2000; Perez-Leroux, 1998). Longitudinal studies investigating which comes first—successful comprehension of complement structures or passing false belief tasks—document that successful comprehension of complement structures must occur first (de Villiers, 2000; de Villiers & Pyers, 1997).

Recall that a primary diagnostic criterion for autism is a qualitative impairment in communication that can be manifested by a “delay in or total lack of spoken language” (American Psychiatric Association, 1994). One of Tager-Flusberg’s longitudinal studies (Steele, Joseph, & Tager-Flusberg, 2003) investigated theory of mind among 57 children with autism between the ages of 4 to 14 years, at the start of the study. Over a one-year period, there were significant developmental improvements in theory of mind ability, and those improvements were primarily related to the children’s developing language abilities. Other cross-sectional studies have demonstrated the same relation: Theory of mind ability in autism is tightly coupled developmentally with language ability (Tager-Flusberg, 1997).

Furthermore, Tager-Flusberg and Sullivan (1994) have demonstrated that when autistic children are compared with non-autistic children who are matched to the autistic children’s language skills, the difference between autistic and non-autistic children in their success rate of passing first and second-order false belief tasks disappears. In other words, if one controls for language abilities, theory of mind deficits are not unique to autism. Moreover, if one creates a false drawing task that tests theory of
mind without reliance on language, one finds that children with autism and children with deafness actually outperform children with normal hearing (Peterson, 2002).

**Is Theory of Mind Innate?**

Although training in mind reading has become a popular intervention for autism (Howlin, Baron-Cohen, & Hadwin, 1998), Tager-Flusberg and her colleagues (Hale & Tager-Flusberg, 2003) have demonstrated that grammatical training on sentential complement structures—low-frequency grammatical constructions such as “what will Jesse think is inside the box before I open it”—improves performance of false belief tasks as successfully as training on only false belief tasks. And training on sentential complement constructions has an added benefit: Improving the understanding of other sentential complement sentences.

Thus, to return to the first two claims made about Theory of Mind as a core deficit in autism: Theory of Mind does not pass the universality criterion. Unsuccessful Theory of Mind performance is neither characteristic of all persons with autism nor characteristic of only persons with autism. Theory of Mind performance is not impervious to training either. Recall that Frith et al.’s original proposal was that the ability to mentalize, while not manifest at birth, could not be explained by learning. Let us now consider the third criterion, the putative existence of a specialized neural mechanism.

**Does Theory of Mind Depend on a Specific Neural Mechanism?**

While at least ten studies have concluded that the putative Theory of Mind module resides in the medial prefrontal cortex (MPC; Brunet, Sarfati, Hardy-Bayle, & Decety, 2000; Calarge, Andreassen, & O’Leary, 2003; Castelli, Frith, Happe, & Frith, 2002; Castelli, Happe, Frith, & Frith, 2000; Fletcher et al., 1995; Gallagher, Happe, Brunswick, Fletcher, Frith, & Frith, 2000; Goel, Grafman, Sadato, & Hallett, 1995; Happe et al., 1996; Nieminen-von Wendt et al., 2003; Vogeley et al., 2001), other studies have suggested the temporal parietal junction (Castelli et al., 2000; Gallagher et al., 2000; Saxe & Kanwisher, 2003), the orbito-frontal cortex (Baron-Cohen, Ring, Moriarty, Schmitz, Costa, & Ell, 1994), the temporal pole (Calarge et al., 2003; Castelli et al., 2000, 2002; Gallagher et al., 2000; Happe et al., 1996), the extrastriate cortex (Brunet et al., 2000; Castelli et al., 2000, 2002; Nieminen-von Wendt et al., 2003), the precuneus (Gallagher et al., 2000; Saxe & Kanwisher, 2003), the thalamus (Nieminen-von Wendt et al., 2003), the anterior cingulate gyrus (Brunet et al., 2000; Calarge et al., 2003; Fletcher et al., 1995; Nieminen-von Wendt et al., 2003; Vogeley et al., 2001), and the cerebellum (Brunet et al., 2000; Calarge et al., 2003). Figure 1 illustrates the range of neuroanatomical locations claimed to be the seat of mentalizing (and Theory of Mind) abilities. The inability of dozens of brain imaging studies to localize a consistent neural area casts strong doubt on there being a single neural module that is missing in autism.
We turn now to a deficit for which there is consistent agreement, at least with regard to the neural tissue involved. Several brain imaging studies have demonstrated that when viewing facial photographs, autistics show less brain activation in the right fusiform gyrus (FG), an area which has been dubbed the “face processing area” because it is prominently activated when non-autistics view facial photographs (Critchley et al., 2000; Hall, Szechtman, & Nahmias, 2003; Pierce, Muller, Ambrose, Allen, & Courchesne, 2001; Schultz et al., 2000). For example, Pierce et al., (2001) reported a robust region of activation in the right fusiform gyrus of non-autistics that was not observed in the right fusiform gyrus of autistics. Such data have led some researchers to speculate that “autism and [Asperger’s Syndrome] involve a congenital abnormality in face ensembles within the [fusiform gyrus] region” (Schultz et al., 2000, p. 338).

However, it is not too surprising that autistics are less likely to activate the putative face processing area; autistics are less likely to look at faces. Indeed, one of the DSM IV criteria for autism is infrequency of eye contact; so by diagnostic definition autistics are less likely to look at faces. But why? Why do persons with autism avoid eye contact? Is it because of social indifference, as some theorists have suggested? Consider instead the words of autistics about this important issue.

In her book, Through the Eyes of Aliens: A Book about Autistic People, Jasmine Lee O’Neill, a mute autistic, gives the following admonition to non-autistics, “Autistics often avoid eye contact, so don’t assume you’re being ignored or treated rudely, if you’re not looked at directly” (O’Neill, 1999, p. 26). Matthew Ward, an honors math
major at the University of Wisconsin, and also an autistic who participated in an experiment described below, stated the following when he was addressing a seminar, “It’s painful for me to look at other people’s faces. Other people’s eyes and mouths are especially hard for me to look at. My lack of eye contact sometimes makes people, especially my teachers and professors, think that I’m not paying attention to them” (Ward, personal communication).

Thus, rather than being socially indifferent, these autistics are fully aware that eye contact is not only expected but also that its lack can be interpreted as rudeness or apathy. But as Matthew Ward points out, eye contact can be painful. Jasmine O’Neill writes that “eyes are very intense and show emotions. It can feel creepy to be searched with the eyes. Some autistic people don’t even look at the eyes of actors or news reporters on television” (O’Neill, 1999, p. 26).

In her book, *The World of the Autistic Child*, Bryna Siegel, a non-autistic clinician, gave the following analogy: “For autistic children, making eye contact with most people seems to be as difficult as [non-autistic people] staring down someone very threatening. One way I sometimes explain this to parents is to say that for an autistic child, giving eye contact is like it might be for you, if you suddenly found yourself at a crowded party, in a strange country where everyone felt it was quite normal to talk to you from within four inches of your face, and ignored signals you might make to indicate you wished to move farther away. In that case, you would probably try to avoid eye contact and turn away, too” (Siegel, 1998, p. 47).

Dalton et al. (2005) recently investigated the biological basis of these intuitions. Their experiment involved 14 right-handed males, all of whom had a DSM IV diagnosis of autism, whose ages ranged from 10–25 years. Included also was a comparison group of 12 right-handed males, none of whom had any DSM IV diagnosis, whose ages ranged from 13–23 years. Participants were acclimated to the brain scanner via participation in a simulation session using a mock scanner, complete with mock scanner noise and many sample trials of the task. Some of the autistic participants chose to participate in more than one simulation session. The experimental task involved viewing 40 facial photographs. Sixteen of the photographs were neutral in their expression of emotion in contrast to 24 that were demonstratively expressive of happiness, anger, or fear. Thus, the emotional expressiveness of the facial photographs was one of two variables that were manipulated.

The other variable manipulated by Dalton et al. (2005) was motivated by a lecture given by Temple Grandin that the senior author of this article attended a few years ago. Grandin is the autistic about whom neurologist Oliver Sacks wrote the essay “An Anthropologist on Mars” in his best seller of the same title. Although Grandin has written several books about autism, such as *Thinking in Pictures and Other Reports from My Life with Autism*, and although Grandin tours the world speaking about autism, the lecture that inspired the experimental manipulation was based on Grandin’s animal science research, which has received international attention as well (Weise, 2003). Grandin’s analysis of animal behavior provides the basis of specific recommendations for more humane interaction. For instance, Grandin advises that when working with a new or a familiar-but-slightly-skittish animal, handlers should avoid making direct eye contact, because direct eye contact can be too threatening. Instead, Grandin advises handlers to avert their gaze, or wear a wide-brimmed baseball cap to occlude their eyes.
Therefore, in the Dalton et al. (2005) experiment, half of both the neutral and emotional faces had been photographed with their eyes straight ahead, and half were photographed with their eyes averted by a quarter turn of the head. The participants' task was to judge whether each facial photograph expressed emotion or whether it was neutral. An event-related functional MRI design was employed, in which each facial photograph was shown for 3 seconds followed by a 5, 6, or 7-second interval. Each participant's electrodermal activity—that is, his or her skin conductance—was measured throughout the experiment, and each participant's eye movements were measured, using an iView eye tracker inside the magnet.

The motivation for including eye tracking was drawn from a couple of recent non-imaging studies that demonstrated that autistic participants were less likely to fixate the eye region of facial photographs (Klin, Jones, Schultz, Volkmar, & Cohen, 2002a, 2002b). Again, this is not too surprising if you pay attention to the words of autistics. For example, Lars Perner is an assistant professor of marketing at San Diego State and an autistic who gives presentations at autism conferences. During one such presentation, Dr. Perner fielded the following comment from a 60 year-old autistic audience member, “For all my life, my brothers and everyone up 'til very recently, have been trying to make me look at them straight in the face. And that is about the hardest thing that I, as an autistic person, can do, because it’s like hypnosis. And you’re looking at each other square in the eye, and it’s very draining.” Perner replied, “Eye contact is very draining indeed. I have developed a strategy for that. I look at people’s noses instead. That works. And people don’t notice” (Perner, 2002).

Similar advice is offered in the Self-Help Guide for Special Kids and Their Parents, co-authored by a mother and her autistic son (Matthews & Williams, 2000). These authors suggest that to practice making eye contact, “first look at a person’s nose. Noses are not as scary as eyes because they do not change their expression or convey a person’s feelings” (Matthews & Williams, 2000, p. 22). These authors’ insight is correct: Noses do not convey very much emotional expression. So, in real life it could be quite adaptive to look at a person’s nose, but in a laboratory experiment, in which the task is to judge whether a facial photograph expresses emotion, you are apt to perform more poorly.

Indeed, as shown in Figure 2 the autistics in the Dalton et al. (2005) experiment were less accurate in their judgments about whether the photographed faces portrayed emotion or whether they were neutral. The non-autistics were close to ceiling in their accuracy, with a mean of 39.4 out of 40 correct, whereas the autistics’ mean was around 34 out of 40, or 85%. But there was great variability among the autistics’ accuracy. There were no significant differences between the two groups’ judgment times when the photographed faces were neutral, regardless of whether the photographed persons’ eyes were looking straight ahead or were averted to the side. In contrast, there were significant group differences in judgment time when the photographed faces displayed emotion: The autistic participants made their judgments less rapidly, particularly when the photographed eyes were looking straight ahead, but also when the photographed eyes were averted.

Dalton et al. (2005) observed both qualitative and quantitative differences between the two groups of participants’ patterns of eye tracking, as expected from autistic persons’ insights. Typically non-autistic participants began in the eye region
and remained relatively close to the eyes. In contrast, autistic participants did not always begin in the eye region and might instead scan the mouth or the nose. An examination of the average time that the participants spent fixating on the photographed face’s right eye revealed that the autistics spent significantly less time fixating the photographs eyes.

Although the autistics and non-autistics did not differ significantly in the number of fixations they made on either the eyes or the mouths, the non-autistics were significantly more likely to fixate the eyes rather than the mouths. This was not a significant difference for the autistics; they were just as likely to fixate the mouths as the eyes. The number of eyes fixated was correlated with the number of correctly judged faces, although this correlation was significant only for the autistics, most likely because the non-autistics were near ceiling in the number of faces they judged correctly, and, therefore, the non-autistics’ range was not as variable as that of the autistics. The autistics were also more variable in the number of fixations they made on the right eyes, the faces in general, and both eyes combined.

Turning to the brain imaging data, Dalton et al. (2005) replicated the finding that when viewing facial photographs non-autistics activate the right fusiform gyrus more than autistics do, as shown in Figure 3. Significant between-group differences in percent MRI signal change were observed for both the right fusiform and the left fusiform: The non-autistic participants showed greater activation in the right than the left-fusiform, whereas the autistic participants showed equivalent activation in the right- and left-fusiform. Very importantly, for both groups, the magnitude of the right fusiform activation was correlated with the time that the participants spent fixating the photographs’ right eye; indeed, the correlation coefficient for the autistics \((r=.75)\) was numerically higher than that for the non-autistics \((r=+.42)\).

Thus, the more time the participants spent fixating the photographs’ eyes—in other words, the more time the participants spent making eye contact—the greater
the activation in the fusiform gyrus of their brains. This finding alone can explain why previous studies have suggested that autistic brains are missing the critical face processing area of the brain. It is not that autistics’ face processing “module” is broken or missing—it is simply that they use it for briefer periods of time.

CONCLUSION

To substantiate the claim that theory of mind is a core deficit in autism, three requirements must be met: individuals with autism must universally fail tests of theory of mind; theory of mind must be innate; and theory of mind must depend on a specific neural mechanism. However, in some studies up to 50% of autistics succeed on theory of mind tasks, and other populations including individuals with specific language impairment fail. Therefore, a “lack of theory of mind” is neither universal in autism nor specific to autism. Training in mind reading or grammatical constructions improves performance on theory of mind tasks, suggesting that theory of mind is not innate. Finally, numerous brain imaging studies have failed to pinpoint one singular neural mechanism.

Previous research has been interpreted as demonstrating that autistic individuals do not activate the putative face area while viewing faces to the same extent as do non-autistics has led to speculation that autistics are missing or have broken “fusiform face processing areas.” However, autistics are less likely to look at faces, especially at the eye region. In fact, Dalton et al. (2005) reported that the time autistics spent fixating the eye region of facial photographs correlates with the neural activity in the putative face region.
Taken together, the evidence presented in this paper casts doubt on the arguments that the autistic brain is missing the core modules responsible for understanding theory of mind and for processing faces.

References


Weise, E. (2003, August 12). Prodding slaughterhouses; Animal-behavior expert is helping packing plants lead cattle to slaughter more humanely. *USA Today*, D.05.


Contact Information:

*Morton Ann Gernsbacher, Ph.D.*

*Department of Psychology*

*University of Wisconsin-Madison*

*Madison, WI 53706*

*MAGernsb@wisc.edu*