CHAPTER 15

Cultural effects on the neural basis of theory of mind

Chiyoko Kobayashi Frank¹ and Elise Temple²,*

¹School of Psychology, Fielding Graduate University, Santa Barbara, CA, USA
²Department of Education Graduate Faculty in Psychological & Brain Sciences, Dartmouth College, Hanover, NH, USA

Abstract: “Theory of mind” has been described as the ability to attribute and understand other people’s desires and intentions as distinct from one’s own. It has been found to develop as early as between 3 and 4 years old, with precursor abilities possibly developing much earlier. There has been debate about the extent to which the developmental trajectory of theory of mind may differ across cultures or language systems. Although very few neuroimaging studies have directly compared different groups from different culture and language systems, across studies of a number of cultural/language groups have been used to explore the neural correlates of theory of mind. A summary of these findings suggests that there may be both universal and culture or language-specific neural correlates related to theory of mind. These studies, while still preliminary in many ways, illustrate the importance of taking into account the cultural background of participants. Furthermore these results suggest that there may be important cultural influence on theory of mind and the neural correlates associated with this ability.

Keywords: theory of mind; culture; social cognition; fMRI; medial prefrontal; temporoparietal junction

Introduction

Theory of mind has been defined as the ability to impute mental states to both oneself and others (Premack and Woodruff, 1978; Wimmer and Perner, 1983). Since the first test of theory of mind in a chimpanzee (Premack and Woodruff, 1978), a number of paradigms have been devised to test theory of mind in humans (Baron-Cohen, 2000). Among those tasks, false-belief tasks have been among the most widely used both for testing normally developing (Wimmer and Perner, 1983) as well as atypical pediatric populations (Baron-Cohen et al., 1985; Baron-Cohen, 2000).

False-belief tasks come in a variety of versions but all involve a protagonist who has some false belief and a subject who has actually knows the true facts of the situation. One of the most common has been termed the “unexpected location” task (Baron-Cohen et al., 1985, 1986). In this task, a protagonist sees an object being placed in a certain location. The protagonist then leaves and the object is moved. When the protagonist returns, he or she mistakenly believes the object is still in its initial location. The key concept to be demonstrated is that the protagonist understands that the object is still in its initial location. The question to the subject is where the protagonist thinks the object is located. The key concept to be demonstrated is that the subject understands that the protagonist has a false belief about the location. False-belief tasks like this have been used as a litmus test to determine if a person has developed a theory of mind (Happe and Loth, 2002).

*Corresponding author.
Tel.: (603)646-9051; Fax: (603)646-3968;
E-mail: Elise.Temple@dartmouth.edu

DOI: 10.1016/S0079-6123(09)17815-9 213
After many variations of these tasks, a general pattern of results has emerged. Four-year-old children are generally able to succeed at false-belief tasks, whereas 3-year-old children tend to fail. In addition, older children or adolescents with autistic spectrum disorders often perform poorly (Baron-Cohen, 2000). This pattern has been hypothesized to be universal regardless of culture (Wellman, 1998). However, a recent meta-analysis comparing non-Western (China and Hong Kong) and North American (United States and Canada) children found similar development trajectories but wide variations in developmental timing across cultures (Liu et al., 2008). For example, the onset of false-belief understanding in Canadian children appeared around 3 years of age, whereas Hong Kong children showed similar performance as much as 2 years later (Liu et al., 2008). These results suggest both universal influences on theory of mind development as well as effects of experience. In addition, Liu et al.’s study highlights the importance of including non-Western cultures when exploring possible cultural effects on the development of any concept or process.

Over the last, almost, 15 years, a number of studies have explored the possible neural correlates of theory of mind. Many of these studies have worked to understand possible universal neural correlates of theory of mind (e.g., Gallagher et al., 2000; Frith and Frith, 2001, 2003). However, as indicated by this volume, recently researchers have begun to shift their attention to the possible cultural variations in the neural underpinnings of various social cognitive functions including theory of mind. In this article, we explore three primary topics in this ongoing effort. These include (1) the findings of cultural and cross-cultural behavioral studies of theory of mind; (2) results of neuroimaging studies of theory of mind and related socio-cognitive functions; (3) implications, limitations, and future directions of these results.

**Behavioral explorations of cultural variation in theory of mind development**

There has been debate about the extent to which the developmental timing and trajectory of theory of mind ability depends on culture and language. Some of the cross-cultural studies have supported the universal developmental pattern of theory of mind (Avis and Harris, 1991; Naito et al., 1994; Lee et al., 1999; Tardiff and Wellman, 2000; Callaghan et al., 2005; Yazdi et al., 2006), whereas others have found differences in non-Anglo children compared to European/American children (delays in non-Anglo children: Chen and Lin, 1994; Wahi and Johri, 1994; Vinden, 1999; Koyasu, 1997; Louis, 1998; Goushiki, 1999; Naito, 2003; Naito and Koyama, 2006; Liu et al., 2008 and advancements in non-Anglo children: Shatz et al., 2003). For example, Callaghan et al. (2005) found that local children in Canadian, Indian, Peru, Samoa, and Thailand all pass the false-belief task at about the same time. They concluded that onset of false-belief understanding is tightly synchronous across different cultures. However, other studies show dramatically different onsets, with fewer than 60% of Japanese 5–6-year olds passing the standard false-belief task (e.g., Koyasu, 1997; Goushiki, 1999; Naito, 2003). Moreover, the results that seem to show no difference across culture/language do not necessarily rule out that there may be cultural/linguistic influence on “how” theory of mind is understood. The forced-choice style false-belief tasks used in the majority of the cross-cultural studies make it difficult to detect strategic differences. For instance, in a study with Mandarin-speaking children, Chinese children’s performance for the false-belief task was comparable to Western children’s performance. However, their performance was influenced by the choice of verb (i.e., three verbs that all mean “think”) used in the false-belief task (Lee et al., 1999). These results indicate possible strategic differences influenced by linguistic variations.

Compared to possible differences due to linguistic variation, those due to cultural variation can be difficult to detect because it is harder to specify exactly what cultural variables may affect theory of mind performance. Vinden (1999) examined emotional false-belief understanding in children from Papua New Guinea, Africa, and Western-influenced countries. While all children came to understand emotion based on desire, only
Western culture children came to understand emotion based on belief by 6 years of age. These results may indicate that at least these non-Western children progress differently in the development of their emotion understanding and the relationship between emotion and internal states of minds.

Similarly, Naito and Koyama (2006) found a striking delay in false-belief understanding in Japanese children (as much as 2 years delayed from Western children). Moreover, even when answering correctly, the justifications the Japanese children gave differed from those typically given by Western children. Most Japanese children referred to behavior (e.g., “she was there first”) or social rules (e.g., “he said to wait there”), rather than the internal and personal justifications (e.g., “he wanted the toy”) commonly given by children from Western cultures (e.g., Bartsch and Wellman, 1989; Wimmer and Mayringer, 1998). These types of results suggest that children across cultures may understand theory of mind in different ways. While Western children may conceptualize theory of mind as being personal and intentional, at least some non-Western cultures may conceptualize theory of mind as being mostly situational and interpersonal.

Brain imaging explorations of neural correlates of theory of mind

To date, a number of brain imaging studies have examined the neural correlates of theory of mind using a variety of paradigms (Fletcher et al., 1995; Goel et al., 1995; Happé et al., 1996; Brunet et al., 2000, 2003; Gallagher et al., 2000; Vogeley et al., 2001; Ferstl and von Cramon, 2002; Saxe and Kanwisher, 2003; Walter et al., 2004; Kobayashi et al., 2006, 2007a, b, 2008; Saxe and Powell, 2006; Sommer et al., 2007; Abraham et al., 2008; Lissek et al., 2008). Many of these studies have found significant activity in the medial prefrontal cortex (mPFC) or anterior cingulate cortex (ACC) during false-belief conditions (e.g., Fletcher et al., 1995; Goel et al., 1995; Happé et al., 1996; Brunet et al., 2000, 2003; Ferstl and von Cramon, 2002; Kobayashi et al., 2006, 2007b, 2008). Additionally, the temporoparietal junction (TPJ) has also been found to be important for theory of mind processing (Saxe and Kanwisher, 2003; Saxe and Powell, 2006). Other brain areas that have sometimes been shown to be involved in theory of mind tasks include dorsolateral prefrontal cortex (DLPFC), precuneus/posterior cingulate cortex (PCC), superior temporal sulcus (STS), and temporal pole (TP).

Although rarely mentioned in the synthesis of the above work, the studies completed thus far have used participants from a few different cultural and/or linguistic backgrounds (though often from Western cultures). In their seminal study using positron emission tomography (PET), Happé et al. (1996) examined the neural basis of theory of mind in Swedish adults with and without Asperger syndrome. They found greater activity in the mPFC in the typically developed control group compared to the Asperger group (Fig. 1). Brunet et al. (2000, 2003) tested French adults with comic strips depicting intentional and non-intentional (control) stories and found more brain activity in several regions including the mPFC during intentional processing. A number of studies have tested German adults with various theory of mind paradigms and implicated several areas including the mPFC and STS (Vogeley et al., 2001; Ferstl and von Cramon, 2002; Walter et al., 2004; Sommer et al., 2007; Abraham et al., 2008; Lissek et al., 2008). Our examinations of adults and children with non-Western background using a false-belief paradigm have shown theory of mind sensitive activity in the mPFC (among other regions) (Kobayashi et al., 2006, 2007a, b, 2008).

Putative neural correlates of theory of mind: universal or culture dependent?

The studies discussed above suggest neural correlates of theory of mind that are possibly both culture dependent and culture independent. In what follows, we discuss some of the brain regions implicated in the literature, in terms of their putative roles in processing either universal or culture/language-specific aspects of theory of mind.
Medial prefrontal cortex and anterior cingulate cortex

The ACC/mPFC has been one of the most consistently implicated region using a variety of theory of mind tasks across a number of cultures (Fig. 1). The mPFC has been suggested to be involved in a self-referential component of theory of mind (Ochsner et al., 2004; Mitchell et al., 2005). According to one review, the mPFC may be further divided into functional subregions: the posterior rostral mPFC, the anterior rostral mPFC, and the orbital mPFC involving cognitive, emotional, and rewarding aspect of theory of mind, respectively (Amodio and Frith, 2006). The ACC has also been conceptualized as potentially divisible into cognitive (dorsal ACC) and emotional (ventral ACC) subregions (Bush et al., 2000). Among these regions, the anterior rostral mPFC (together with a part of ventral ACC) seems to be the one that is most consistently recruited for different theory of mind tasks across cultures (Fig. 1). Thus, we hypothesize that the anterior rostral mPFC and ventral ACC may be involved in theory of mind processing, specifically the emotional and self-related aspects, regardless of culture. There is debate, however, about the role of mPFC in theory of mind. Two studies with patients with frontal damage have found intact theory of mind performance (Fine et al., 2001;...
Bird et al., 2004). Saxe (2006) had argued that although the mPFC is involved in reasoning about mental states, it is related more to the understanding of relationships — the emotional aspects of relationships such as empathy in ventral prefrontal areas and the collaborative aspects in dorsal regions. Clearly, more research across different tasks, cultures, and age groups need to be done to clarify these theories.

**Temporoparietal junction**

While the majority of the earlier neuroimaging studies of theory of mind implicated the mPFC as the core region for theory of mind (e.g., Fletcher et al., 1995; Goel et al., 1995; Happé et al., 1996; Gallagher et al., 2000; Vogeley et al., 2001), several later fMRI studies found more robust brain activity in the TPJ than in the mPFC (Saxe and Kanwisher, 2003; Saxe and Powell, 2006). In addition, disrupted theory of mind performance has been shown in TPJ-damaged patients (Apperly et al., 2004; Samson et al., 2007). Saxe (2006) has argued that the TPJ is crucial for representing mental states and theory of mind. However, the specialized involvement of the TPJ in theory of mind has been questioned in studies that have found activity in the TPJ while subjects reoriented their attention to any novel stimuli (Corbetta et al., 2008; Mitchell, 2008).

The generalization of TPJ’s importance in theory of mind across cultures is still uncertain. TPJ activity specific to theory of mind has not consistently been found across cultures other than American and British adults (see Fig. 1). We did not find theory of mind specific TPJ activity (at $p < 0.05$, uncorrected) in Japanese children and found significantly less theory of mind TPJ activity in Japanese adults (compared to American adults) (Kobayashi et al., 2006; Kobayashi et al., 2007b; see also Perner and Aichhorn, 2008).

One possibility is that potential cultural differences in TPJ activity in theory of mind tasks are due to different cultural approaches to theory of mind. Some research has suggested that the TPJ may be involved in distinguishing self-agency from other agency (Blakemore and Frith, 2003; Jackson and Decety, 2004; Decety and Grézes, 2006). If one culture had a more self–other distinction of theory of mind, perhaps the TPJ would be more involved in their processing.

Some studies have suggested that Japanese culture may encourage intersubjective or situational mentalizing over a subjective approach (Naito, 2007; Naito and Koyama, 2006). While Indo-European language speakers may conceive of an event based on the action-agent model (Werner and Kaplan, 1963), Japanese speakers may tend to frame the event as a situation that is beyond the agent’s control (Maynard, 1997).

There may be variability in this type of interdependent versus independent thought even in Western cultures. A large-scale comparative study involving more than 5000 participants from 29 nations (including Asian, European, and North American nationals) indicated that interdependency is correlated with low English-language fluency (Fernández et al., 2005). In addition, it has been demonstrated that, unlike Anglo-Americans, 3–5-year-old French children rarely use the subjective “belief” concept to justify the behaviors in false-belief task (Bradmetz, 1998). It is not clear why several neuroimaging studies of theory of mind conducted in continental European countries (French, Swedish, and several German) did not find the theory of mind specific activity in the TPJ. One hypothesis, that clearly needs more explicit testing, is that at least individuals from continental European cultures may also conceptualize theory of mind in less self-referential way. We believe that more work in different cultures needs to be done to determine whether TPJ theory of mind specific activity is truly culture independent or whether it is especially engaged in Anglo-American cultures.

**Other putative theory of mind regions (PCC/Precuneus, DLPFC, and TP)**

The PCC or precuneus area has been among the most frequently implicated regions in the theory
of mind neuroimaging studies. Given that not only American/British (Fletcher et al., 1995; Gallagher et al., 2000; Saxe and Powell, 2006) but also German (Ferstl and von Cramon, 2002; Vogeley et al., 2001; Walter et al., 2004; Sommer et al., 2007) and Japanese (Kobayashi et al., 2006) adults showed activity in this area for theory of mind (vs. a variety of non-ToM conditions), the PCC’s involvement in theory of mind may also be culture independent. Vogeley and Fink (2003) suggest that the PCC or medial parietal cortex is important for the formation of first person perspective. Moreover, along with the mPFC, this region has been hypothesized to be active during a default mode or baseline condition when subjects may be self-ruminating (Gusnard and Raichle, 2001; Gusnard et al., 2001; den Ouden et al., 2005). Perhaps, the precuneus/PCC is involved in processing some first person-related intentional aspects of theory of mind or the self-ruminating aspects of the default network in a culture-independent manner.

The DLPFC has been implicated as being important for executive function (e.g., Frith and Dolan, 1996; MacDonald et al., 2000). DLPFC activity may be related to inhibitory control involved in theory of mind (Saxe et al., 2004; Kain and Perner, 2005). For instance, in the false-belief scenario, one has to inhibit the first (or second in the case of the second-order false-belief task) character’s belief or thought. At least five brain imaging studies that tested French (Brunet et al., 2000), German (Vogeley et al., 2001; Sommer et al., 2007), British (Gallagher et al., 2000), and Japanese and American adults (Kobayashi et al., 2006) have found activity in this area for cartoon or story-based theory of mind. It is possible that involvement of DLPFC is related to inhibitory control in certain theory of mind tasks but in a culture-independent manner.

Several theory of mind imaging studies across cultures found activity in the anterior tip of STS or TP (Gallagher et al., 2000; Vogeley et al., 2001; den Ouden et al., 2005; Abraham et al., 2008). However, we found significant difference between American and Japanese groups in this area (Kobayashi, 2007). In our study, American adults and children had more activity in this area than the Japanese adults and children during the same cartoon theory of mind task. Therefore, theory of mind activity in the TP (especially the right TP) may be culture dependent.

The TP have been implicated in a number of related operations. TP regions are activated when one retrieves autobiographical or episodic memory (Fink et al., 1996; Maguire and Mummery, 1999; Maguire et al., 2000). Frith and Frith (2003) have suggested that this area may be responsible for accessing social knowledge in the form of scripts, which aid interpretation of social situations. In Naito’s 2003 study on Japanese children’s theory of mind, she not only found a delay in theory of mind, but also a correlation between the children’s theory of mind and performance in self-related episodic memory. One possibility, again that needs confirmation, is that people from Japanese culture recruit the TP to a lesser magnitude than Americans during theory of mind processing because their access to episodic memory is more automatic than Americans.

**Limitations and future directions in cultural neuroimaging studies of theory of mind**

There are several limitations of the current cross-cultural research (both behavioral and neuroimaging) of theory of mind that need to be considered. One clear limitation is the paucity of cross-cultural neuroimaging research of theory of mind overall. Even though there are a number of within cultural brain imaging studies of theory of mind, so far, ours are the only cross-cultural studies that have tried to compare neural correlates of theory of mind between the two cultures: Japanese and American (Kobayashi et al., 2006, 2007b, 2008). However, ours is not a true cross-cultural study, since the Japanese subjects were bilingual and lived in the United States at the time of the experiment. Ideally, as in behavioral studies, cross-cultural variation should be tested against clearly defined within cultural variation. Cross-cultural theory of mind brain imaging studies that compare subjects from two (or more) distinct cultures need to be completed to gain a more thorough understanding.
Another clear limitation of cross-cultural theory of mind research overall (see Vinden, 1999 and Lillard, 1998a, b) is that the standard false-belief task may not be valid for all cultures. Many non-Western cultures do not construe behaviors as personal and intentional. Results of Naito and Koyama (2006) have shown that many Japanese children rarely give desire-based explanation to account for the false belief of the protagonist. These results (together with Vinden’s, 1999, results) call into question the applicability of the developmental order of theory of mind concepts — from desire-based understanding to belief-based understanding (Wellman and Liu, 2004) — to all cultures. As we mentioned earlier, the forced-choice false-belief task may not measure these qualitative differences in theory of mind understanding. Future cross-cultural brain imaging studies of theory of mind might target some clearly defined cultural variables (e.g., subjectivity vs. intersubjectivity) that might account for strategic differences in false-belief understanding.

Another major limitation is the scarcity of developmental neuroimaging research of theory of mind. Brain imaging studies of theory of mind related tasks in children are relatively few (Ohnish et al., 2004; Dapretto et al., 2006; Wang et al., 2006a, b; Kobayashi et al., 2007a, b). We found some cross-cultural differences in the neural basis of theory of mind in Japanese and American children (Kobayashi et al., 2007b) although they were much smaller than the differences found in adults (Kobayashi et al., 2006). Nativistic theories of theory of mind suggest that at least core set of mental concepts are resistant to cultural variation throughout development (Bruner, 1990; Wellman, 1990). They also presume that these core sets have clearly defined neural basis (Gallagher et al., 2000; Leslie, 2005; and see Barrett and Kurzban, 2001, for a different view that does allow cross-cultural variation in neural basis of theory of mind). In fact, our study found several brain regions that showed convergent activity across cultural and age groups (Kobayashi, 2007). Thus, more neuroimaging studies in children across different cultures are called for to examine whether or not these brain regions subserve the universal core sets of theory of mind.

Lastly, differences in theory of mind related brain activity across cultures could also be attributed to linguistic differences. Although the extent to which language affects neural basis of theory of mind is still debated (Siegal and Varley, 2002), increasing evidence suggests that there may be reciprocal influence between language and theory of mind throughout development (Malle, 2002; Miller, 2006). The processing of pragmatically coherent sentences also recruits the mPFC area (Ferstl and von Cramon, 2002). It may be that the pragmatics or communicative aspects of language profoundly affect theory of mind throughout development, consistent with the finding that people with autism are most impaired in the pragmatic aspects of language (Landau, 2000; Miller, 2006). Since pragmatics is the very aspect of language where cultural constructs often seep in (Fiedler, 2008), it may be difficult to tease apart linguistic effect on theory of mind from cultural one.

**Conclusion**

Both behavioral and brain imaging research have begun to suggest that some aspects of theory of mind may not be entirely universal. There seem to be differences among cultures in how people understand others’ behaviors. These differences may be reflected in at least some variability in the neural correlates of theory of mind across cultures. Several brain regions associated with theory of mind, such as the mPFC, show evidence for possibly being culture independent. Other putative theory of mind regions, such as the TPJ, may be employed by Anglo-Americans, but to a lesser degree by other cultural groups. An important task of future theory of mind research is to define which cultural factors may be affecting theory of mind and examine these factors in systematic ways. These factors may not be only dichotomous (e.g., independence vs. interdependence) but also multifaceted (see Kashima and Kashima, 1997). It may also be problematic to impose a Western assumption of intentionality on false belief to different cultures (see Vinden, 1999).
Many of the theories of theory of mind predict culturally invariant core biological bases of theory of mind at least during early years in life. Thus, an important goal of theory of mind brain imaging research is to explore a possible core neural basis that remains relatively free from cultural influence.

Currently, there are more questions than answers in terms of the neural correlates of theory of mind and how they may vary across development, across languages, and across cultures. However, given some differences have already been suggested in both the brain and behavior of people across different cultures, it seems imperative to at least consider that cultural differences may exist. As all the articles in this volume suggest, studying brain mechanism in one or even two cultures cannot necessarily give you a universal understanding of those processes. Taking culture into account is especially important in processes like theory of mind that may have both cultural and linguistic influences.

Acknowledgments

We thank Dr. Barbara C. Lust, Dr. Michael J. Spivey, Dr. Michael Siegal, and Dr. Simon Baron-Cohen for discussion, and Randall Frank for assistance.

References


