The Neural Bases of Social Cognition and Story Comprehension

Raymond A. Mar

Department of Psychology, York University, Toronto M3J 1P3 Canada; email: mar@yorku.ca

Key Words
neuroimaging, theory-of-mind, reading, narrative, fMRI, PET

Abstract
A great deal of research exists on the neural basis of theory-of-mind (ToM) or mentalizing. Qualitative reviews on this topic have identified a mentalizing network composed of the medial prefrontal cortex, posterior cingulate/precuneus, and bilateral temporal parietal junction. These conclusions, however, are not based on a quantitative and systematic approach. The current review presents a quantitative meta-analysis of neuroimaging studies pertaining to ToM, using the activation-likelihood estimation (ALE) approach. Separate ALE meta-analyses were conducted for story-based and nonstory-based studies of ToM. The conjunction of these two meta-analyses reveals a core mentalizing network that includes areas not typically noted by previous reviews. A third ALE meta-analysis was conducted with respect to story comprehension in order to examine the relation between ToM and stories. Story processing overlapped with many regions of the core mentalizing network, and these shared regions bear some resemblance to a network implicated by a number of other processes.
INTRODUCTION

One of the most fundamental tools we have for social cognition is the ability to infer the mental states of others, known as theory-of-mind (ToM) or mentalizing (Carruthers & Smith 1996, Premack & Woodruff 1978). Humans demonstrate great proficiency at this, identifying in others beliefs, emotions, and motivations similar to their own from about the age of 4 years onward (Aston et al. 1988). Our capacity to do so is essential for successful navigation of the social world, enabling collaboration with our peers (Watson et al. 1999) and buttressing the social relations that form the basis of both local community and broader society (Tomasello et al. 2005). Neuroscientific research has begun to uncover the key brain areas that support our ability to mentalize, with numerous reviews identifying what has become known as the mentalizing network. Brain areas that contribute to this network include the medial prefrontal cortex (mPFC), posterior cingulate cortex (pCC) and precuneus, bilateral temporoparietal junction (bTPJ), and, somewhat less commonly identified, the superior temporal sulcus (STS), temporal poles, and amygdala (e.g., Adolphs 2003, 2009; Allison et al. 2000; Beer & Ochsner 2006; Blakemore et al. 2004; Carrington & Bailey 2009; Decety & Chaminade 2003; Decety & Somerville 2003; C.D. Frith & Frith 1999; U. Frith & Frith 2001, 2003; Gallagher & Frith 2003; Lieberman 2007; Mitchell 2008a; Ochsner 2004; Saxe 2006a; Siegal & Varley 2003) (see Table 1 for acronym definitions). Together, however, these reviews do not present a consensus on the components of this network. Individual articles commonly mention some brain areas but neglect to mention others, identify areas not mentioned by other authors, and place a particular emphasis on different areas. The root of this heterogeneity likely lies in the method employed by most previous reviewers. These authors produce reviews based on a qualitative, narrative-based approach, which is vulnerable to various subjective biases that may influence any attempt to summarize the extant literature. Although some reviews provide a rerepresentation of data from previous studies, it is rare that any statistical analysis is undertaken to investigate whether the apparent patterns revealed reflect more than chance clustering.1 A more

1For an exception, see the analysis of Van Overwalle (2009), which employs statistical tests for the presence or absence of activations in a region. This article identifies the mPFC and the TPJ as key areas.
Table 1  Acronyms for brain regions

<table>
<thead>
<tr>
<th>X/Y</th>
<th>Z</th>
<th>Lobe/region</th>
<th>Gyrus/sulcus/cortex</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>S</td>
<td>= anterior</td>
<td>F = frontal</td>
</tr>
<tr>
<td>p</td>
<td>M</td>
<td>= posterior</td>
<td>T = temporal</td>
</tr>
<tr>
<td>l</td>
<td>I</td>
<td>= left</td>
<td>P = parietal</td>
</tr>
<tr>
<td>m</td>
<td></td>
<td>= medial</td>
<td>C = cingulate</td>
</tr>
<tr>
<td>dm</td>
<td></td>
<td>= dorsomedial</td>
<td>PF = prefrontal</td>
</tr>
<tr>
<td>r</td>
<td></td>
<td>= right</td>
<td>TPJ = temporoparietal junction</td>
</tr>
</tbody>
</table>

reliable and objective identification of the mentalizing network can be achieved only through the application of a quantitative, systematic, and statistical meta-analytic approach (Costafreda 2009, Wager et al. 2007). This review presents such an analysis, employing the activation likelihood estimation (ALE) method (Eickhoff et al. 2009, Turkeltaub et al. 2002) to determine which brain areas are consistently implicated by neuroimaging studies of ToM. Not only does this allow for a statistically based identification of ToM regions, this approach also allows us to address some interesting questions regarding studies of ToM: how the results of story-based and nonstory-based tasks compare, for example. This issue is important, given concerns that story-based examinations of ToM may be confounded by mental-state language processing and not adequately represent real-world applications of ToM (Apperly et al. 2004).

An additional question that may be investigated through an ALE meta-analysis of ToM is how ToM relates to other processes (Spreng et al. 2009). To this end, an ALE meta-analysis is conducted for story comprehension studies to investigate the relation between ToM (examined using both story and nonstory methods) and narrative comprehension. A number of theorists and researchers have postulated that readers employ ToM in the understanding of stories (Bruner 1986, Keen 2007, Mar & Oatley 2008, Zunshine 2006), and this is certainly the assumption of story-based studies of ToM. Teasing out the exact relation between story comprehension and ToM, however, has proven to be a difficult and under-recognized necessity (Ferstl & von Cramon 2002).

NEUROIMAGING STUDIES OF THEORY-OF-MIND

Story-Based Studies

The use of stories to explore the neural bases of ToM began with one of the earliest neuroimaging studies conducted on this topic, by Fletcher and colleagues (1995). These researchers adopted stories written by Happé (1994), which were designed to tap a reader’s capacity for mental inference. For example, one such “ToM story” describes how a captured prisoner of war, under interrogation, intentionally told his captors the truth about the location of some tanks with the knowledge that his enemies would assume he was lying and thus look in the only other likely location. Comprehending this story requires a clear understanding of how second-order belief reasoning (understanding how another person can hold a belief about someone else’s mental state) can be employed for the purposes of deception. As a control task, Fletcher and colleagues (1995) developed a series of stories that required no mental inference for comprehension, what they called “physical stories.” These were closely matched to the ToM stories in terms of content and difficulty but dealt with purely physical or mechanical causes rather than psychological ones. So, in a parallel example to the ToM story mentioned above, one physical story describes how foggy conditions led to the success of one army over another, due to the latter’s superiority in air power being neutralized by the weather. In a positron emission tomography (PET) study, participants read stories of both kinds and answered a question regarding...
each story that tested comprehension of either the mental states of characters or the physical causes of events.

This design quickly became popular among ToM researchers, with many employing these same stories or creating new variants (Table 2). Increasingly, neuroscientists have come to utilize these stories as a localizer task that identifies the areas thought to be employed in ToM for each participant; a mask is then used to isolate the relevant areas in the analysis of a subsequent task (Saxe et al. 2006a; cf. Friston & Henson 2006).

Some concerns about the use of stories in ToM studies have begun to arise, however. In the minds of some researchers, the use of verbal story stimuli to study ToM may problematically confound mental-state language processing or executive functioning with ToM (e.g., Apperly et al. 2004; cf. Saxe 2006b). The role that language and language areas of the brain play in ToM has been the subject of much debate. This is true not just in neuroimaging circles, but also in neuropsychological investigations of brain-damaged patients (Varley et al. 2001), as well as in

---

Table 2 Story-based theory-of-mind studies

<table>
<thead>
<tr>
<th>Article</th>
<th>Task</th>
<th>Control</th>
<th>Method</th>
<th>N</th>
<th>Foci</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aichorn et al. (2009)</td>
<td>False belief stories</td>
<td>False photograph stories</td>
<td>fMRI</td>
<td>21</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Berthoz et al. (2002a)</td>
<td>Intentional social violation story</td>
<td>Unintentional social violation story</td>
<td>fMRI</td>
<td>12</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Fletcher et al. (1995)</td>
<td>ToM story judgment</td>
<td>Physical causation story judgment</td>
<td>PET</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Gallagher et al. (2000)</td>
<td>ToM story judgment</td>
<td>Physical causation story judgment</td>
<td>fMRI</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Gobbini et al. (2007)</td>
<td>False belief stories</td>
<td>Physical stories</td>
<td>fMRI</td>
<td>12</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Jenkins &amp; Mitchell (2010)</td>
<td>ToM stories</td>
<td>Non-ToM stories</td>
<td>fMRI</td>
<td>15</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Kobayashi et al. (2006)</td>
<td>Second-order false-belief story judgment</td>
<td>Physical causation story judgment</td>
<td>fMRI</td>
<td>16</td>
<td>5</td>
<td>American participants</td>
</tr>
<tr>
<td></td>
<td>Second-order false-belief story judgment</td>
<td>Physical causation story judgment</td>
<td>fMRI</td>
<td>16</td>
<td>4</td>
<td>Japanese participants</td>
</tr>
<tr>
<td>Mitchell (2008b)</td>
<td>False belief stories</td>
<td>False photograph stories</td>
<td>fMRI</td>
<td>20</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nieminen-von Wendt et al.</td>
<td>ToM story judgment</td>
<td>Physical causation story judgment</td>
<td>PET</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Peron et al. (2006)</td>
<td>False belief stories</td>
<td>False photograph stories</td>
<td>fMRI</td>
<td>19</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Saxe &amp; Kanwisher (2003)</td>
<td>ToM story judgment</td>
<td>Physical causation story judgment</td>
<td>fMRI</td>
<td>25</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Saxe et al. (2006b)</td>
<td>False belief stories</td>
<td>False photograph stories</td>
<td>fMRI</td>
<td>12</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Saxe &amp; Wexler (2005)</td>
<td>False belief stories</td>
<td>False photograph stories</td>
<td>fMRI</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Spengler et al. (2009)</td>
<td>ToM stories</td>
<td>Physical stories</td>
<td>fMRI</td>
<td>18</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Vogeley et al. (2001)</td>
<td>ToM story judgment</td>
<td>Physical causation story judgment</td>
<td>fMRI</td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Young et al. (2007)</td>
<td>False belief stories</td>
<td>False photograph stories</td>
<td>fMRI</td>
<td>10</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>False belief stories</td>
<td>False photograph stories</td>
<td>fMRI</td>
<td>17</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: fMRI, functional magnetic resonance imaging; PET, positron emission tomography; ToM, theory-of-mind.
work with normally developing children (Astoning & Jenkins 1999) and abnormally developing children (Garfield et al. 2001). An additional concern with this story-based method is the assumption that inferring the mental states of fictional characters is identical to the process of understanding our peers. Although this assumption has not fully been examined, one functional magnetic resonance imaging (fMRI) study found that observing a cartoon person resulted in less blood-oxygenation-level-dependent (BOLD) response in key social processing areas compared to the exact same movements enacted by a real person [i.e., in the right temporoparietal junction (rTPJ), right superior temporal sulcus (rSTS), and pCC/precuneus (Mar et al. 2007)]. This finding brings into question whether our processing of fictional social agents is truly identical (in magnitude and form) to the processing of actual social agents. In light of the concerns about story-based approaches to examining ToM, a number of researchers have embarked on the development of nonstory-based methods for investigating this process.

### Nonstory-Based Studies

Nonstory-based studies of ToM have employed a greater variety of stimuli and designs than their story-based counterparts (Table 3). One popular approach is to present cartoon versions of a false-belief test, or comics drawn to invoke ToM processing. Another approach employs a novel stimulus set initially designed by Heider & Simmel (1944). These stimuli consist of short animations depicting simple geometric shapes moving in a way that implies social behavior and intentional actions (e.g., hiding, chasing, teasing). Activations associated with these animations are contrasted with those observed during the presentation of shapes moving randomly and in a nonintentional manner. Variations of these animations have been constructed by researchers to pinpoint more specific hypotheses, such as the relation between emotional empathy and cognitive perspective-taking (Hynes et al. 2006). Although these types of animations reduce the possibility that language (mental state or otherwise) is confounded with the task, these designs still involve adopting the intentional stance (Dennet 1987) toward targets that clearly lack intentions (Mar & Macrae 2006), something possibly different from (though clearly similar to) our everyday application of ToM to our peers.

In another popular design, participants play a simple game (e.g., rock, paper, scissors) against a human opponent, contrasted with playing either by themselves or against a computer opponent (Table 3). In the first condition, participants are presumably employing mental inference in order to predict the next move of their opponent, whereas such a strategy is unlikely to be usefully employed against a computer (typically described as choosing randomly). Although actual conspecifics are not visually presented in these designs, their presence is assumed, akin to a computer-mediated interaction with another person (e.g., instant messaging or a game of online chess). Other common research designs include (a) asking participants to make mental-state judgments or form an impression based on pictures of human faces, contrasted with non-ToM judgments such as gender or sequence of presentation; (b) having participants take another’s perspective when making judgments or describing preferences contrasted with their own perspective; (c) viewing actions that are communicative or involve pretense contrasted with actual instrumental actions; and (d) viewing actions and judging the mental state behind them (Table 3). Although some of these studies involve the presentation of words or short phrases, none involve stories or mental-state language in the same way as the story-based studies mentioned above.

### Identifying the Mentalizing Network

With the diversity of opinion on how to properly study ToM using neuroimaging comes a parallel uncertainty of what is the proper meta-analysis required to identify the mentalizing network. Grouping all ToM

---

**Table 3**

<table>
<thead>
<tr>
<th>Design Features</th>
<th>Example Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>False-belief test</td>
<td>The primary measurement made during fMRI, corresponding to changes in blood oxygenation found within the brain (functional magnetic resonance imaging (fMRI))</td>
</tr>
<tr>
<td>Blood-oxygenation-level-dependent (BOLD) response</td>
<td>a test of ToM that relies on an understanding that others may hold beliefs not supported by reality</td>
</tr>
</tbody>
</table>

---

PS62CH05-Mar ARI 22 November 2010 9:29
Table 3  Nonstory-based theory-of-mind studies

<table>
<thead>
<tr>
<th>Article</th>
<th>Task</th>
<th>Control</th>
<th>Method</th>
<th>N</th>
<th>Foci</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bahneman et al. (2010)</td>
<td>ToM judgment</td>
<td>Appearance judgment</td>
<td>fMRI</td>
<td>25</td>
<td>9</td>
</tr>
<tr>
<td>Baron-Cohen et al. (1999)</td>
<td>Judge mental state from face</td>
<td>Judge gender from face</td>
<td>fMRI</td>
<td>12</td>
<td>53</td>
</tr>
<tr>
<td>Bhatt &amp; Camerer (2005)</td>
<td>Second-order belief judgment in game</td>
<td>Belief judgment</td>
<td>fMRI</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Blakemore et al. (2003)</td>
<td>Directed attention to animated shapes</td>
<td>Undirected attention to animated shapes</td>
<td>fMRI</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Brune et al. (2008)</td>
<td>ToM cartoon sequencing</td>
<td>Non-ToM cartoon sequencing</td>
<td>fMRI</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Brunet et al. (2000)</td>
<td>ToM cartoon judgment</td>
<td>Physical causality cartoon judgment</td>
<td>PET</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>Brunet et al. (2003)</td>
<td>ToM cartoon judgment</td>
<td>Physical causality cartoon judgment</td>
<td>PET</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Castelli et al. (2000)</td>
<td>ToM animated shapes</td>
<td>Randomly moving shapes</td>
<td>PET</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Ciaramidaro et al. (2007)</td>
<td>ToM cartoon judgment</td>
<td>Physical cartoon judgment</td>
<td>fMRI</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Decety et al. (2004)</td>
<td>Play game against a human</td>
<td>Play game alone</td>
<td>fMRI</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Elliott et al. (2006)</td>
<td>Play game with a human</td>
<td>Play game alone</td>
<td>fMRI</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Fukui et al. (2006)</td>
<td>Play game against human</td>
<td>Play game against computer</td>
<td>fMRI</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Gallagher et al. (2000)</td>
<td>ToM cartoon judgment</td>
<td>Physical causality cartoon judgment</td>
<td>fMRI</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Gallagher et al. (2002)</td>
<td>Play game against a human</td>
<td>Play game against computer</td>
<td>PET</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>German et al. (2004)</td>
<td>View pretense actions</td>
<td>View real actions</td>
<td>fMRI</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Gilbert et al. (2007)</td>
<td>Task with human</td>
<td>Task with computer</td>
<td>fMRI</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Grèzes et al. (2004a)</td>
<td>Judge target to have false expectation</td>
<td>Judge target to have true expectation</td>
<td>fMRI</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Grèzes et al. (2004b)</td>
<td>Judge action to be deceptive</td>
<td>Judge action to be honest</td>
<td>fMRI</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Hooker et al. (2008)</td>
<td>Infer emotion during false belief</td>
<td>Recognize emotion during false belief</td>
<td>fMRI</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Kana et al. (2009)</td>
<td>ToM animated shapes</td>
<td>Randomly moving shapes</td>
<td>fMRI</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Kircher et al. (2009)</td>
<td>Play game against a human</td>
<td>Play against a computer</td>
<td>fMRI</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Krach et al. (2009)</td>
<td>Play game against a human</td>
<td>Play against a computer</td>
<td>fMRI</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Lombardo et al. (2010)</td>
<td>Make mental judgment about other</td>
<td>Make mental judgment about self</td>
<td>fMRI</td>
<td>33</td>
<td>3</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Article</th>
<th>Task</th>
<th>Control</th>
<th>Method</th>
<th>N</th>
<th>Foci</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malhi et al. (2008)</td>
<td>ToM animated shape</td>
<td>Randomly moving shapes</td>
<td>fMRI</td>
<td>20</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Marjoram et al. (2006)</td>
<td>ToM cartoon judgment</td>
<td>Physical causality cartoon judgment</td>
<td>fMRI</td>
<td>13</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Mitchell et al. (2005a)</td>
<td>Judge mental state from face</td>
<td>Judge symmetry of face</td>
<td>fMRI</td>
<td>18</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Ochsner et al. (2005)</td>
<td>Judge other’s evaluation of self</td>
<td>Self-judgment</td>
<td>fMRI</td>
<td>16</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Platek et al. (2004)</td>
<td>Judged mental state based on pictures of eyes</td>
<td>Fixation on crosshairs</td>
<td>fMRI</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Rabin et al. (2010)</td>
<td>Infer mental state from photo</td>
<td>Recollect autobiographical memory from photo</td>
<td>fMRI</td>
<td>20</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Rilling et al. (2004)</td>
<td>Play game against a human</td>
<td>Press button for monetary reward</td>
<td>fMRI</td>
<td>19</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Rilling et al. (2008)</td>
<td>Play game against a human</td>
<td>Play against a computer</td>
<td>fMRI</td>
<td>20</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Russell et al. (2000)</td>
<td>Judge mental state from face</td>
<td>Judge gender from face</td>
<td>fMRI</td>
<td>7</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Samson et al. (2008)</td>
<td>ToM cartoon judgment</td>
<td>Semantic cartoon judgment</td>
<td>fMRI</td>
<td>17</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Sommer et al. (2007)</td>
<td>False belief cartoon</td>
<td>True belief cartoon</td>
<td>fMRI</td>
<td>16</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Spiers &amp; Maguire (2006)</td>
<td>Spontaneous ToM events during virtual reality</td>
<td>Non-ToM events</td>
<td>fMRI</td>
<td>20</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Spreng &amp; Grady (2010)</td>
<td>Infer mental state from photo</td>
<td>Recollect past or imagine future from photo</td>
<td>fMRI</td>
<td>16</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Sripada et al. (2009)</td>
<td>Play game against a human</td>
<td>Play against a computer</td>
<td>fMRI</td>
<td>26</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Vanderwal et al. (2008)</td>
<td>ToM animated shapes</td>
<td>Physical animation judgment</td>
<td>fMRI</td>
<td>17</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Vollm et al. (2006)</td>
<td>ToM cartoon judgment</td>
<td>Physical causality cartoon judgment</td>
<td>fMRI</td>
<td>13</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Walter et al. (2004)</td>
<td>Comic strips with communicative intentions</td>
<td>Comic strips with physical causality</td>
<td>fMRI</td>
<td>13</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>—</td>
<td>Comic strips with communicative intentions</td>
<td>Comic strips with physical causality</td>
<td>fMRI</td>
<td>12</td>
<td>15</td>
<td>Different sample</td>
</tr>
<tr>
<td>Wolf et al. (2010)</td>
<td>View ToM movie scenes</td>
<td>View physical inference movie scenes</td>
<td>fMRI</td>
<td>18</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: fMRI, functional magnetic resonance imaging; PET, positron emission tomography; ToM, theory-of-mind.
Expository nonfiction: discourse-level language that presents a series of arguments in order to communicate or persuade (e.g., essays)

studies to date is one option, collapsing across both story-based and nonstory-based investigations to provide an unbiased representation of the peak locations identified by previous studies of ToM. However, because nonstory-based studies currently predominate (Tables 2 and 3), these types of investigations will have greater influence on the overall result. It might also be argued that only a meta-analysis of the nonstory studies is appropriate, as it is these studies that are not confounded by mental-state language. This approach is not likely to satisfy other researchers who believe that story-based investigations of ToM are a necessary part of understanding the process of mental-inference, however (Saxe 2006b). A third possibility is to examine where story- and nonstory-based studies of ToM overlap, with the presumption that whatever is pivotal to the mentalizing process will be reflected by the convergence of these two separate methodologies, with idiosyncratic design elements (e.g., language or facial processing) not contributing to the overlap. Rather than adjudicate this debate, this review presents the results for story-based and nonstory-based studies along with their overlap; readers may decide for themselves which approach is most compelling.

An advantage of the ALE method to meta-analysis is the potential to compare the results of different meta-analyses and thus examine how different processes relate. In order to better elucidate the relation between ToM and narratives, an additional meta-analysis was undertaken in this review to uncover the brain areas commonly associated with story comprehension.

THE RELATION BETWEEN THEORY-OF-MIND AND NARRATIVE COMPREHENSION

A number of theorists have hypothesized that ToM is employed in the understanding of fiction (for a review, see Mar & Oatley 2008). The basic premise of this idea is that, in understanding fictional others (e.g., characters in a novel or a film), we employ the same or similar processes used to understand the mental states of real others (Gerrig 1993, Oatley 1999). Although this idea certainly seems intuitive, to date there has not been a great deal of empirical work that investigates this topic. If story comprehension involves social cognitive processes, then we would expect individuals who frequently engage with stories to benefit socially in some way from these repeated experiences. Consistent with this idea, Mar and colleagues (2006) have demonstrated that lifetime exposure to narrative fiction, controlling for exposure to expository nonfiction, is positively associated with social abilities. Unfortunately, the correlational nature of this study excludes the possibility of any causal inference. But additional support comes in the form of a subsequent study that ruled out the role of individual differences in explaining this phenomenon (Mar et al. 2009). Also consistent with this idea are studies on preschool-aged children that demonstrate a relation between exposure to storybooks and social development (Adrian et al. 2005, Aram & Aviram 2009, Mar et al. 2010). Although there has been some related work on this question from other perspectives (e.g., whether stories can foster empathy toward out-group members; for a review, see Paluck & Green 2009), gaps still exist in the empirical literature regarding whether ToM and narrative comprehension are related, in precisely what way, and what the ramifications of such a relation might be (Mar & Oatley 2008).

Some neuroimaging investigations of this topic do exist. Mar (2004), for example, conducted a qualitative review of discourse-processing studies (both neuroimaging and patient-based) and noted a similarity in the areas commonly associated with narrative and ToM (i.e., mPFC, bTPJ, temporal poles, and pCC). Ferstl and colleagues (2008) conducted an ALE meta-analysis of text comprehension

2Exposure to expository nonfiction, controlling for narrative fiction, demonstrates the opposite relation (Mar et al. 2006). Expository nonfiction constitutes a useful control because it embodies discourse-level text but does not include the characters and settings that are likely to draw upon ToM processes.
and also observed a similar pattern (i.e., mPFC, bTPJ, and anterior temporal lobes). This meta-analysis included a number of improvements over the earlier review (Mar 2004), most notably its employment of a quantitative approach. One other important distinction is that Ferstl and colleagues (2008) chose not to include any contrasts that employed stories known to have ToM content, favoring the physical stories used as controls in these studies. This removes an important confound, forming a stronger argument for a shared neural network subserving ToM and story comprehension. However, these physical stories bear little resemblance to the everyday narratives that we typically read, which are known to contain a great deal of ToM content. Even simple children’s stories regularly deal in situations that involve the inference of belief, and these stories are full of mental-state language (Cassidy et al. 1998, Dyer et al. 2000). Although it is interesting to investigate whether any piece of connected discourse will engage ToM brain areas, it is also interesting to examine whether the stories we actually read draw upon ToM processes in a similar way as real-world social comprehension (and if so, what the social and cognitive ramifications of this might be).

Ferstl and colleagues (2008) have suggested that in order to confidently determine the existence and nature of an overlap between ToM and story comprehension, two ALE analyses for these separate processes should be conducted and compared, and this approach is taken here. The meta-analysis for narrative comprehension reported in this review differs from that performed by Ferstl and colleagues (2008) in important ways, however. Unlike the previous analysis, the current one includes no single-sentence studies, only one contrast is taken from each paper (Turkeltaub et al. 2002), studies published since that paper are included, and studies that explicitly employed a nonfiction or nonstory text are excluded. Most importantly, a new algorithm for the ALE analysis is employed (Eickhoff et al. 2009), improving the formerly fixed-effects analysis to a random-effects analysis. Adopting this algorithm means that the results of this analysis can be generalized to the larger population of similar studies from which this sample is drawn, whereas the results of the earlier paper can only be generalized to the particular set of studies examined. In this meta-analysis of story comprehension, no stories devoted to examining ToM are included (these are found in the story-based ToM meta-analysis), nor are control stories taken from ToM studies (recall that only one contrast is taken per experiment). Instead, this meta-analysis focuses on story comprehension studies with no particular stance toward ToM in an effort to gather a sample more akin to the stories we actually read in our daily lives. By comparing the results of this analysis to those for story-based and nonstory-based ToM studies, we can more clearly examine how stories and ToM are related.

**METHOD**

**Selection of Studies**

One possible criticism of meta-analytic studies is that they are prone to bias as qualitative reviews; bias simply enters the equation when studies are selected, shaping the results of any analysis. In order to reduce this possibility, a systematic approach was adopted for the selection of studies. Briefly, a set of face-valid search criteria was employed to gather relevant studies through a search of three online databases: PsycInfo, Medline, and the Science Citation Index. Searches were limited to articles published (or

www.annualreviews.org • Stories and Social Comprehension 111
available through online advance access) in English, between January 1985 and January 2010, and identified by the following keywords: “neuroimaging” OR “fMRI” OR “PET” AND “theory-of-mind” OR “mentalizing” OR “mindreading.” The results of this search were then culled by removing review papers and theoretical papers as well as studies that did not report data for nonclinical samples (e.g., Castelli et al. 2002), studies with child participants (e.g., Ohnishi et al. 2004), and studies that employed a region-of-interest approach and did not report individual coordinates (e.g., Hynes et al. 2006). Additionally, studies were excluded if they did not examine a basic process of mental inference but instead pursued a more specific goal (e.g., identifying the role of self-other similarity; Mitchell et al. 2006) or if it was unclear if mental inference was taking place (e.g., response to hearing one’s own name; Kampe et al. 2003). Studies with emotional stimuli (e.g., facial displays of emotions or emotionally disturbing pictures; e.g., Dolan & Frith 2004) were excluded in order to compose a more uniform sample. Although many of these emotion studies examined the process of inferring an individual’s mental state, the nature of the stimuli involved in these studies is likely to evoke a very different neural response compared to studies using neutral targets. In cases where papers included more than one independent sample, all appropriate data were included in the analysis (e.g., Walter et al. 2004). A total of 63 samples were considered in this meta-analysis (N = 897), which were categorized as either story-based (20 samples, N = 274; Table 2) or nonstory-based (43 samples, N = 623; Table 3).

A parallel procedure was followed to gather neuroimaging studies on narrative comprehension, using the keywords “neuroimaging” OR “fMRI” OR “PET” AND “narrative” OR “story” AND “comprehension.” Reviews, theoretical papers, and clinical studies (e.g., McNeil et al. 2006) were excluded. In addition, studies were excluded if they involved child participants (e.g., Schmithorst et al. 2006); explicitly mentioned employing expository nonfiction texts (an important exclusion in light of our interest in narrative processing; e.g., articles from a journal of popular science; Alho et al. 2003); employed a region-of-interest approach (e.g., Cherney & Small 2006); examined single sentences and not connected discourse (e.g., Friese et al. 2009); employed multiple sentences but the text did not constitute a story (e.g., syllogisms, questions and answers, action scripts; e.g., Caplan & Dapretto 2001); or was a more narrow investigation that did not provide appropriate contrasts for identifying story-specific activations due to lack of an appropriate baseline (e.g., contrasts between types of stories; Ferstl et al. 2005). This analysis included 23 samples (N = 355; Table 4). There was no overlap in the studies used for each of the three meta-analyses.

**Activation Likelihood Estimation Analysis**

The ALE approach to meta-analysis for neuroimaging data involves collecting 3-D coordinate locations of peak activations from multiple studies (Eickhoff et al. 2009, Laird et al. 2005, Turkeltaub et al. 2002; see also Chein et al. 2002). Only one analysis was selected from each study in order to limit the degree to which individual studies may unequally influence the overall pool of foci (Turkeltaub et al. 2002). In all cases, statistical contrasts were selected to maximize the degree to which they isolated the process of interest. For example, if a study reported a contrast between stories and single words, and a contrast between stories and sentences, the latter would be chosen due to its superior control condition. A total of 550 foci were gathered from the ToM studies (146 story-based, 404 nonstory-based; Tables 2 and 3), and 216 foci were collected from the narrative comprehension studies (Table 4). In cases in which coordinates were reported using the coordinate system of Talairach & Tournoux (1988), these were converted to that of the Montréal Neurological Institute (MNI) using...
### Table 4  Narrative comprehension studies

<table>
<thead>
<tr>
<th>Article</th>
<th>Task</th>
<th>Control</th>
<th>Method</th>
<th>N</th>
<th>Foci</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crinion &amp; Price (2005)</td>
<td>Auditory fables</td>
<td>Fables reversed</td>
<td>fMRI</td>
<td>18</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Crinion et al. (2003)</td>
<td>Auditory stories</td>
<td>Stories reversed</td>
<td>PET</td>
<td>11</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>Auditory stories</td>
<td>Stories reversed</td>
<td>PET</td>
<td>6</td>
<td>10</td>
<td>Different sample</td>
</tr>
<tr>
<td>Dick et al. (2009)</td>
<td>Auditory fables</td>
<td>Fixation</td>
<td>fMRI</td>
<td>24</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Ferstl &amp; von Cramon (2001)</td>
<td>Read related sentences</td>
<td>Read unrelated sentences</td>
<td>fMRI</td>
<td>12</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Giraud et al. (2000)</td>
<td>Auditory story</td>
<td>Auditory sentences</td>
<td>PET</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Kansaku et al. (2000)</td>
<td>Auditory story</td>
<td>Story reversed</td>
<td>fMRI</td>
<td>22</td>
<td>2</td>
<td>Males</td>
</tr>
<tr>
<td>–</td>
<td>Auditory story</td>
<td>Story reversed</td>
<td>fMRI</td>
<td>25</td>
<td>2</td>
<td>Females</td>
</tr>
<tr>
<td>Kuperberg et al. (2006)</td>
<td>Read related sentences</td>
<td>Read unrelated sentences</td>
<td>fMRI</td>
<td>15</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Lindenberg &amp; Scheef (2007)</td>
<td>Written and auditory stories</td>
<td>Language-like stimuli</td>
<td>fMRI</td>
<td>19</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Maguire et al. (1999)</td>
<td>Written stories</td>
<td>Incomprehensible stories</td>
<td>PET</td>
<td>13</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Miura et al. (2005)</td>
<td>Written stories</td>
<td>Fixation</td>
<td>fMRI</td>
<td>30</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Pappathanassiou et al. (2000)</td>
<td>Auditory story</td>
<td>Rest</td>
<td>PET</td>
<td>8</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Perani et al. (1996)</td>
<td>Auditory story (native language)</td>
<td>Auditory story (foreign language)</td>
<td>PET</td>
<td>9</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Perani et al. (1998)</td>
<td>Auditory story (native language)</td>
<td>Backward story (foreign language)</td>
<td>PET</td>
<td>9</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>Auditory story (native language)</td>
<td>Backward story (native language)</td>
<td>PET</td>
<td>12</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Robertson et al. (2000)</td>
<td>Written sentences with definite articles</td>
<td>Sentences with indefinite articles</td>
<td>fMRI</td>
<td>8</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Siebörger et al. (2007)</td>
<td>Read related sentences</td>
<td>Read distantly related sentences</td>
<td>fMRI</td>
<td>14</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Tzourio et al. (1998)</td>
<td>Auditory stories</td>
<td>Rest</td>
<td>PET</td>
<td>10</td>
<td>17</td>
<td>Right-handers</td>
</tr>
<tr>
<td>–</td>
<td>Auditory stories</td>
<td>Rest</td>
<td>PET</td>
<td>5</td>
<td>16</td>
<td>Left-handers</td>
</tr>
<tr>
<td>Xu et al. (2005)</td>
<td>Written fables</td>
<td>Unrelated sentences</td>
<td>fMRI</td>
<td>22</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Yarkoni et al. (2008a)</td>
<td>Written story</td>
<td>Scrambled sentences</td>
<td>fMRI</td>
<td>29</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Yarkoni et al. (2008b)</td>
<td>Written story</td>
<td>Fixation</td>
<td>fMRI</td>
<td>28</td>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: fMRI, functional magnetic resonance imaging; PET, positron emission tomography.

Tools provided by the GingerALE software (BrainMap GingerALE 2.0; Research Imaging Center, University of Texas Health Science Center at San Antonio). Individual coordinates were then modeled by a 3-D Gaussian probability distribution using this same program, with the characteristics of each distribution based on an empirical derivation (Eickhoff et al. 2009). A probability was calculated for each voxel based on the uncertainty distributions for all foci in each study. This value represents the probability that an activation is located at that exact position, taking into account the spatial uncertainty associated with different...
normalization procedures and between-subject variability. A probability map is then created for each study, and ALE scores are calculated by taking the union of all the individual study maps. These ALE scores represent the probability that a voxel is activated based on interstudy convergence. In order to test for statistical significance these probabilities are compared against a null distribution, calculated using an analytic method that approximates $10^{11}$ permutations of randomly selected voxels (one from each study map). ALE values for each voxel were determined to be statistically significant using an alpha threshold of $p < 0.05$ (false-discovery rate controlled; Laird et al. 2005), with clusters possessing a minimum volume of 100 mm$^3$ reported. (Clusters smaller than this volume are not listed but may appear in figures.) This version of GingerALE improves on past versions by using an algorithm to test above-chance clustering between experiments (a random-effects analysis), allowing for generalization of the results beyond the sample examined to the entire population of similar studies. Additional improvements found in this version of GingerALE include the use of a gray matter mask to exclude coordinates located in deep white matter and the consideration of sample size for each study.

Clusters were localized with respect to anatomical landmarks using neuroanatomy atlases (e.g., Scarabino et al. 2003) and Brodmann areas for each extrema noted using the Talairach daemon (Research Imaging Center, University of Texas Health Science Center at San Antonio). As always, Brodmann area designations determined in this way are very rough estimates that are unlikely to correspond to microanatomical distinctions; they should be interpreted with a great deal of caution (Devlin & Poldrack 2007).

**Conjunction Analysis**

Overlap between two or more ALE meta-analyses was determined using a conjunction analysis in which ALE maps for each analysis were overlaid upon one another. Colors were then assigned to visualize where clusters from each analysis overlapped. This approach does not involve a statistical test (see Spreng et al. 2009).

**RESULTS**

**Story-Based ToM Studies**

Table 5 and Figures 1 and 2 present the ALE results for story-based ToM studies (clusters in blue, turquoise, purple, and black). Complete results images for all ALE analyses are available via the Supplemental Material link from the Annual Reviews home page, http://www.annualreviews.org/ (nii format).

The cluster most likely to be activated by these studies was located at the right angular gyrus, including the pSTS, TPJ, and pSTG (Figure 2B,D), with a similar cluster appearing on the left (but of slightly lower likelihood; Figure 2A,C,E). A cluster in the mPFC was next most likely to be activated (Figure 1A,B), similar in probability to the left TPJ. This prefrontal cluster was the largest observed, spanning from $x = -19$ to 13, progressing dorsally as it extended from the left hemisphere into the right, including portions of the frontal pole and the most anterior aspect of the cingulate cortex. Next in likelihood was the precuneus, extending to the pCC (Figure 1A,B). Additional areas were found in other social-processing areas such as the STS (bilaterally), less frequently mentioned areas such as the left temporal pole and left amygdala (Figure 1C), and also in a rarely mentioned region, the left superior frontal gyrus (SFG).

**Nonstory-Based Theory-of-Mind Studies**

Table 6 and Figures 1 and 2 present the ALE results for the nonstory-based ToM studies (clusters in green, turquoise, yellow, and

---

1Statistical subtraction was not used to compare ALE results because this analysis is currently only available for an earlier version of GingerALE, one that does not support a random-effects analysis.
Table 5 Clusters resulting from the ALE analysis of story-based theory-of-mind studies

<table>
<thead>
<tr>
<th>Peak: cluster extent</th>
<th>BA</th>
<th>Vol. (mm$^3$)</th>
<th>ALE ($\times 10^{-3}$)</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>r: 1</td>
<td>mPFC: frontal pole, aCC, mSFG</td>
<td>9</td>
<td>7624</td>
<td>3.96</td>
<td>4</td>
<td>58</td>
</tr>
<tr>
<td>r</td>
<td>aCC</td>
<td>32</td>
<td>312</td>
<td>1.39</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>l</td>
<td>Angular gyrus: pSTS/TPJ, pSTG, pMTG</td>
<td>39</td>
<td>4072</td>
<td>3.79</td>
<td>−52</td>
<td>−58</td>
</tr>
<tr>
<td>l</td>
<td>Angular gyrus</td>
<td>39</td>
<td>152</td>
<td>1.08</td>
<td>−42</td>
<td>−72</td>
</tr>
<tr>
<td>r</td>
<td>Angular gyrus: pSTS/TPJ, pSTG</td>
<td>39: 22</td>
<td>3696</td>
<td>5.02</td>
<td>54</td>
<td>−54</td>
</tr>
<tr>
<td>l</td>
<td>MTG: STS</td>
<td>21</td>
<td>328</td>
<td>1.39</td>
<td>−56</td>
<td>−28</td>
</tr>
<tr>
<td>r</td>
<td>MTG: STS</td>
<td>21</td>
<td>1552</td>
<td>1.69</td>
<td>56</td>
<td>−16</td>
</tr>
<tr>
<td>l</td>
<td>Precuneus: pCC</td>
<td>31: 23</td>
<td>3392</td>
<td>2.61</td>
<td>−10</td>
<td>−50</td>
</tr>
<tr>
<td>l</td>
<td>Temporal pole</td>
<td>38</td>
<td>120</td>
<td>1.18</td>
<td>−52</td>
<td>12</td>
</tr>
<tr>
<td>l</td>
<td>aMTS: MTG, ITG</td>
<td>20</td>
<td>288</td>
<td>1.40</td>
<td>−52</td>
<td>−4</td>
</tr>
<tr>
<td>l</td>
<td>SFG</td>
<td>9</td>
<td>496</td>
<td>1.38</td>
<td>−20</td>
<td>54</td>
</tr>
<tr>
<td>l</td>
<td>SFG</td>
<td>8</td>
<td>336</td>
<td>1.47</td>
<td>−20</td>
<td>34</td>
</tr>
<tr>
<td>l</td>
<td>Amygdala: parahippocampal gyrus</td>
<td>168</td>
<td>168</td>
<td>1.19</td>
<td>−26</td>
<td>−2</td>
</tr>
</tbody>
</table>

Notes: Peak is location of peak voxel in cluster; cluster extent refers to other regions in contact with cluster; BA is Brodmann area. All BAs estimated using the Talairach daemon and are rough approximations that should be interpreted with caution. Additional BAs refer to peaks for subclusters. Coordinates are reported in MNI space. Abbreviations: a, anterior; ALE, activation likelihood estimation probability; BA, Brodmann area; CC, cingulate cortex; ITG, inferior temporal gyrus; m, medial; MTG, middle temporal gyrus; MTS, middle temporal sulcus; p, posterior; PFC, prefrontal cortex; SFG, superior frontal gyrus; STG, superior temporal gyrus; STS, superior temporal sulcus; TPJ, temporoparietal junction; Vol., cluster volume. See Supplemental Material for all results images (.nii format).

The cluster with the greatest likelihood and largest volume was located in the mPFC, ranging from $x = −19$ to 15 (Figure 1A,B). The peak voxel for this cluster was located in the left hemisphere, with the cluster progressing ventrally as it moved closer to the midline, extending into the right hemisphere and encompassing the most anterior part of the cingulate cortex, and also extending superiorly to the dorsal aspect of the SFG. Next in likelihood was a cluster in the right anterior STS (Figure 2B,D), approaching the temporal pole, with a similar activation occurring on the left (but of lower likelihood; Figure 2C). Additional clusters were found in frontal areas (SFS, MFG), ventral regions (fusiform and lingual gyrus), and the subcortex (putamen, thalamus).

**Story- Versus Nonstory-Based Theory-of-Mind Studies**

In order to clarify the similarities between story- and nonstory-based studies of ToM, a conjunction analysis was performed. The outcome of this analysis is portrayed in Figures 1 and 2, with overlapping voxels appearing in turquoise and black. A large overlap between the two types of ToM occurred in the mPFC, primarily in the right hemisphere (Figure 1B). At $x = −6$ in the left hemisphere, the two clusters are separate with the nonstory-based region superior to the story-based one (Figure 1A). By $x = −3$, the two clusters are almost completely overlapping and remain so as the cluster progresses through the right hemisphere until...
Table 6  Clusters resulting from the ALE analysis of nonstory-based theory-of-mind studies

<table>
<thead>
<tr>
<th>Peak: cluster extent</th>
<th>BA</th>
<th>Vol. (mm$^3$)</th>
<th>ALE ($\times 10^{-3}$)</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>l: r mPFC: aSFG, aCC</td>
<td>8: 9, 32</td>
<td>6920</td>
<td>4.16</td>
<td>−10</td>
<td>50</td>
<td>34</td>
</tr>
<tr>
<td>r</td>
<td>aCC: dmPFC</td>
<td>32</td>
<td>304</td>
<td>1.79</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>l pMTG: pSTS/TPJ, STG, supramarginal gyrus</td>
<td>22: 39</td>
<td>4912</td>
<td>3.05</td>
<td>−56</td>
<td>−48</td>
<td>4</td>
</tr>
<tr>
<td>l Angular gyrus</td>
<td>39</td>
<td>320</td>
<td>1.88</td>
<td>−46</td>
<td>−62</td>
<td>32</td>
</tr>
<tr>
<td>r pSTG: pSTS/TPJ, MTG, IPG</td>
<td>13*: 39, 22, 37</td>
<td>5608</td>
<td>2.44</td>
<td>50</td>
<td>−46</td>
<td>18</td>
</tr>
<tr>
<td>r Angular gyrus</td>
<td>39</td>
<td>600</td>
<td>2.13</td>
<td>52</td>
<td>−54</td>
<td>34</td>
</tr>
<tr>
<td>l STS: MTG, STG</td>
<td>21</td>
<td>400</td>
<td>1.80</td>
<td>−54</td>
<td>−32</td>
<td>−6</td>
</tr>
<tr>
<td>r: l Precuneus: pCC, paracentral lobule</td>
<td>7</td>
<td>4904</td>
<td>2.64</td>
<td>2</td>
<td>−56</td>
<td>38</td>
</tr>
<tr>
<td>l IFG (orbitalis): IFG (triangularis)</td>
<td>47</td>
<td>2040</td>
<td>3.64</td>
<td>−48</td>
<td>30</td>
<td>−12</td>
</tr>
<tr>
<td>l IFG (orbitalis)</td>
<td>47</td>
<td>144</td>
<td>1.53</td>
<td>−32</td>
<td>20</td>
<td>−12</td>
</tr>
<tr>
<td>l IFG</td>
<td>44</td>
<td>528</td>
<td>2.29</td>
<td>−48</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>r IFG (opercularis): IFG (triangularis)</td>
<td>44: 45</td>
<td>2432</td>
<td>2.23</td>
<td>54</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>l aMTG: aSTS, aMTS</td>
<td>21</td>
<td>1856</td>
<td>2.50</td>
<td>−54</td>
<td>0</td>
<td>−26</td>
</tr>
<tr>
<td>r aSTS: MTG, STG</td>
<td>21</td>
<td>1952</td>
<td>3.86</td>
<td>60</td>
<td>−8</td>
<td>−18</td>
</tr>
<tr>
<td>l SFS: SFG</td>
<td>6</td>
<td>152</td>
<td>1.66</td>
<td>−28</td>
<td>2</td>
<td>60</td>
</tr>
<tr>
<td>r MFG</td>
<td>6</td>
<td>576</td>
<td>2.15</td>
<td>44</td>
<td>20</td>
<td>44</td>
</tr>
<tr>
<td>r Fusiform gyrus</td>
<td>20</td>
<td>496</td>
<td>1.97</td>
<td>−38</td>
<td>−46</td>
<td>−18</td>
</tr>
<tr>
<td>l Lingual gyrus</td>
<td>18</td>
<td>448</td>
<td>2.11</td>
<td>2</td>
<td>−80</td>
<td>0</td>
</tr>
<tr>
<td>r Insula</td>
<td>47</td>
<td>1288</td>
<td>3.09</td>
<td>34</td>
<td>22</td>
<td>−12</td>
</tr>
<tr>
<td>r Amygdala: parahippocampal gyrus, uncus</td>
<td>560</td>
<td>2.34</td>
<td>26</td>
<td>0</td>
<td>−22</td>
<td></td>
</tr>
<tr>
<td>l Putamen: caudate nucleus</td>
<td>176</td>
<td>1.67</td>
<td>−14</td>
<td>10</td>
<td>−8</td>
<td></td>
</tr>
<tr>
<td>l Thalamus</td>
<td>224</td>
<td>1.98</td>
<td>−8</td>
<td>−16</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>r Thalamus</td>
<td>424</td>
<td>2.02</td>
<td>4</td>
<td>−28</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Peak is location of peak voxel in cluster; cluster extent refers to other regions in contact with cluster. All BAs estimated using the Talairach daemon and are rough approximations that should be interpreted with caution. Additional BAs refer to peaks for subclusters. Coordinates are reported in MNI space. Asterisk indicates BA only present in nonhuman primates, Talairach daemon designation. Abbreviations: a, anterior; ALE, activation likelihood estimation probability; BA, Brodmann area; CC, cingulate cortex; ITG, inferior temporal gyrus; m, medial; MTG, middle temporal gyrus; MTS, middle temporal gyrus; p, posterior; PFC, prefrontal cortex; SFG, superior frontal gyrus; STG, superior temporal gyrus; STS, superior temporal sulcus; TPJ, temporoparietal junction; Vol., cluster volume. See Supplemental Material for all results images (.nii format).

$x = 13$. A separate region of overlap also occurs in the left mPFC ($x = −19$ to $−13$), between the nonstory cluster and a different story-based cluster more dorsal than the cluster that bridges the hemispheres. Clusters also overlapped in the right pSTS/TPJ and angular gyrus (Figure 2B,D), the left pSTS/TPJ and angular gyrus (including the supramarginal gyrus) (Figure 2A,C,E), the right MTG and STS (Figure 2D; MTG not shown), left MTG and aMTG (Figure 2C/E), and the precuneus and pCC (Figure 1A,B). One additional point of overlap was observed between a large nonstory cluster in the IFG and a very small story-based cluster in pars orbitalis (32 mm$^3$, ALE $= 1.00 \times 10^{-3}$, MNI: $−46$ $−30$ $−14$; Figure 2A). This cluster was not reported earlier because of its size (below the 100 mm$^3$ cluster threshold), but its probability is above the threshold for statistical significance ($p < 0.05$, false-discovery rate controlled).

Although these areas were shared across the two approaches to studying ToM, it is important to note that the likelihood associated with a region often differed depending on the approach. For story-based studies, the rTPJ was the most likely region, followed by the mPFC. In the case of nonstory-based studies,
the mPFC was the most likely region, with the rTPJ having a lower likelihood than a number of other areas, including the right aSTS, left IFG, right insula, and even the ITPJ. Differences in approach to studying ToM may explain some of the current disagreement regarding which areas are key for mentalizing.

One other similarity is worth noting. Both types of studies were likely to activate the amygdala, although it was the left amygdala in the case of story-based studies and the right amygdala for nonstory-based studies (Figure 1C). Keeping in mind that the amygdala is a heterogenous structure that includes several distinct nuclei (Davis et al. 2010, Markowitsch 1998), the peak locations for story- and nonstory-based studies were in very similar locations aside from their hemispheric lateralization, differing by only 2 mm on the Y-axis with respect to their peak voxels (Tables 5 and 6). This implies that the difference in amygdala involvement for story- and nonstory-based studies is exclusively one of lateralization and not the engagement of different subregions.

Table 7  Clusters resulting from the activation likelihood estimation analysis of narrative comprehension studies

<table>
<thead>
<tr>
<th>Peak: cluster extent</th>
<th>BA</th>
<th>Vol. (mm&lt;sup&gt;3&lt;/sup&gt;)</th>
<th>ALE (× 10&lt;sup&gt;−3&lt;/sup&gt;)</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>r: 1</td>
<td>mPFC: aSFG</td>
<td>9</td>
<td>952</td>
<td>2.20</td>
<td>6</td>
<td>60</td>
</tr>
<tr>
<td>l</td>
<td>mPFC: aSFG</td>
<td>8</td>
<td>328</td>
<td>1.86</td>
<td>−10</td>
<td>50</td>
</tr>
<tr>
<td>l</td>
<td>aSTS: IFG (pars orbitalis), temporal pole, MTG, STG, STS, pSTS/TPJ</td>
<td>21: 47, 38, 22, 41, 39</td>
<td>18,288</td>
<td>3.44</td>
<td>−58</td>
<td>−6</td>
</tr>
<tr>
<td>r</td>
<td>pSTS/TPJ: pSTG, pMTG</td>
<td>39</td>
<td>696</td>
<td>2.22</td>
<td>60</td>
<td>−58</td>
</tr>
<tr>
<td>r</td>
<td>STG: STS</td>
<td>22: 41</td>
<td>2728</td>
<td>2.27</td>
<td>58</td>
<td>−2</td>
</tr>
<tr>
<td>r</td>
<td>STS</td>
<td>22</td>
<td>288</td>
<td>1.61</td>
<td>54</td>
<td>−42</td>
</tr>
<tr>
<td>l</td>
<td>IFG: pars triangularis</td>
<td>46: 45</td>
<td>2792</td>
<td>3.29</td>
<td>−48</td>
<td>26</td>
</tr>
<tr>
<td>r</td>
<td>IFG</td>
<td>45</td>
<td>320</td>
<td>1.80</td>
<td>58</td>
<td>32</td>
</tr>
<tr>
<td>l</td>
<td>aITG</td>
<td>20</td>
<td>312</td>
<td>1.53</td>
<td>−42</td>
<td>−16</td>
</tr>
<tr>
<td>r</td>
<td>Temporal pole: aSTS, aSTG, aMTG</td>
<td>38</td>
<td>3256</td>
<td>3.37</td>
<td>52</td>
<td>10</td>
</tr>
<tr>
<td>r</td>
<td>aMTG: STS, MTS</td>
<td>20</td>
<td>840</td>
<td>1.85</td>
<td>56</td>
<td>−10</td>
</tr>
<tr>
<td>l</td>
<td>Dorsal precentral gyrus</td>
<td>6</td>
<td>1568</td>
<td>3.30</td>
<td>−46</td>
<td>2</td>
</tr>
<tr>
<td>r</td>
<td>MTG</td>
<td>20</td>
<td>112</td>
<td>1.31</td>
<td>54</td>
<td>−34</td>
</tr>
<tr>
<td>r</td>
<td>Cerebellum</td>
<td>136</td>
<td>1.45</td>
<td>26</td>
<td>−82</td>
<td>−34</td>
</tr>
<tr>
<td>l</td>
<td>Medial geniculate nucleus: parahippocampal gyrus</td>
<td>35</td>
<td>112</td>
<td>1.30</td>
<td>−18</td>
<td>−22</td>
</tr>
</tbody>
</table>

Notes: Peak is location of peak voxel in cluster; cluster extent refers to other regions in contact with cluster. All BAs estimated using the Talairach daemon and are rough approximations that should be interpreted with caution. Additional BAs refer to peaks for subclusters. Coordinates are reported in MNI space. Abbreviations: a, anterior; ALE, activation likelihood estimation probability; BA, Brodmann area; CC, cingulate cortex; ITG, inferior temporal gyrus; m, medial; MTG, middle temporal gyrus; MTS, middle temporal sulcus; p, posterior; PFC, prefrontal cortex; SFG, superior frontal gyrus; STG, superior temporal gyrus; STS, superior temporal sulcus; TPJ, temporoparietal junction; Vol., cluster volume. See Supplemental Material for all results images (.nii format).
temporal pole (Figure 2B,D). Another cluster possessing high likelihood was found in the left dorsal precentral gyrus (Figure 2A,C,E), and additional clusters were found in the mPFC bilaterally (Figure 1A,B). Finally, clusters appeared in the cerebellum and in the medial geniculate nucleus extending to the parahippocampal gyrus.

Story- and Nonstory-Based Theory-of-Mind and Narrative Comprehension

In order to examine how social cognition and story comprehension relate, the overlap of story comprehension, story-based ToM, and nonstory-based ToM was examined. The three groups of studies converged in a number of areas, represented by the black regions in Figures 1 and 2. Overlap occurred in the mPFC (the paracingulate region) primarily in the right hemisphere (Figure 1B). At x = −6, the three analyses produce nonoverlapping clusters (Figure 1A), with story-based ToM studies producing the most inferior cluster directly anterior to the aCC, a smaller narrative cluster appearing superior to this region, and finally a nonstory-based ToM cluster superior to this, closest to the dorsal surface. At x = −3, however, all three clusters overlap, and this continues into the right hemisphere until x = 11. All three analyses also overlap in the right pSTS/TPJ (Figure 2D), but not as far caudally and dorsally as where the two ToM analyses overlap in the angular gyrus (Figure 2B,D). A similar relationship appears in the left hemisphere, with all three analyses overlapping in the pSTS/TPJ (Figure 2E) but not in the more caudal and dorsal region of ToM overlap (Figure 2A,C). An additional area of overlap occurred in the aMTG, bilaterally (Figure 2C,D,E). Finally, overlap was observed in the left IFG at pars opercularis (Figure 2A), taking into account the small cluster from the story-based analysis mentioned above.

In sum, narrative comprehension clusters overlapped with many of the regions identified by the overlap of story- and nonstory-based ToM studies. Perhaps the most notable exception for this pattern is the absence of any narrative comprehension cluster in the large ToM overlap found in medial parietal regions (pCC and precuneus; Figure 1A,B). Narrative comprehension clusters were also absent in some additional areas of ToM overlap, including the bilateral angular gyrus, left MTG and SFG, and right STS (Figures 1 and 2, turquoise clusters that do not include black).

DISCUSSION

Quantitative meta-analyses were conducted for both story- and nonstory-based neuroimaging studies of ToM, and an additional meta-analysis was performed for studies of narrative comprehension. The relation between these three groups of studies was examined with regard to overlapping areas of the brain. One primary goal of this undertaking was to identify the brain areas consistently activated by studies of ToM at probabilities greater than chance, in other words, to identify the mentalizing network.

Identifying the Mentalizing Network

As noted above, there are numerous ways that one could go about identifying the mentalizing network. There are arguments for examining only the story-based ToM studies, which tend to be more homogenous in design, or only the nonstory-based ToM studies, which do not rely on mental-state language, or perhaps only their overlap, which should control for idiosyncratic elements of each approach and provide convergent validity.

4Narrative comprehension regions that did not touch upon the overlap between ToM results include a large swath of the left STG, from the posterior STG to the temporal pole, an IFG cluster on the left (including pars opercularis) separate from the area of overlap, regions of the medial geniculate nucleus and parahippocampal gyrus, the right cerebellum, rSTG, right temporal pole, and rITG (Figures 1 and 2, red clusters that do not include black). Please see Supplemental Materials for full results files for all meta-analyses.
Story-based ToM studies. Employing stories written to tap mental inference is one of the most popular ways that researchers isolate the mentalizing network. A meta-analysis of these studies found a network of brain regions that was more extensive than that typically reported by qualitative reviews on this topic. Along with the typical areas associated with mentalizing (mPFC, bTPJ, pCC), this analysis revealed clusters in the bilateral STS, left temporal pole, left amygdala, and left SFG (Figures 1 and 2; Table 5). It is important to stress that all of these clusters appear at above-chance levels and are thus statistically reliable correlates of story-based ToM studies. Although it is tempting to disregard the areas that are least likely to be activated (i.e., have the lowest ALE values), all of the regions listed have probabilities that are above an appropriate alpha threshold and deserve some consideration. This does not mean that all of these areas are necessary for ToM, only that they are consistently found across studies of this type.

An advantage of this group of studies is its homogeneity; almost all of the studies employed nearly identical designs. The results thus provide a good picture of what story-based studies of ToM indicate. A potential concern, however, is that some brain areas identified in this analysis represent mental-state language processing, but not ToM specifically. It is therefore important to examine whether a similar network is observed with nonstory-based studies of ToM.

Nonstory-based ToM studies. The meta-analysis for nonstory-based studies of ToM revealed a similar pattern of brain areas compared to story-based studies, one that involved the social-processing areas typically mentioned by previous reviewers (i.e., mPFC, bTPJ, pCC). As before, however, clusters were also discovered in areas not traditionally identified as belonging to this network, most notably in the IFG bilaterally, anterior temporal regions, right amygdala, and bilateral STS. Additional clusters appeared in frontal and ventral temporal cortices along with some subcortical areas (Figures 1 and 2; Table 6). The presence of a large and probabilistically likely left IFG cluster associated with nonstory-based studies indicates that this area may not simply represent language-processing per se, but may contribute some related process that aids mental inference. Not all of the tasks included in this analysis were completely language free, however, and it is possible that this area is implicated solely because participants were asked to perform some simple reading during the task. That said, this possibility is not likely when one considers the fact that reading was involved in both the experimental and control tasks for these studies, and the difficulty of reading was equivalent. This point also pertains to the story-based studies of ToM: Very similar stories were read in the control condition as in the experimental condition, so the observed activations are unlikely to be due to general language processing.

Overlap between story- and nonstory-based ToM studies. A final perspective on identifying the true mentalizing network is to examine where the two approaches to examining ToM overlap, assuming that processes idiosyncratic to each will not be represented. This is the most conservative approach and may be the most appropriate one. Story-based and nonstory-based studies overlap in a number of areas, specifically the mPFC, bilateral pSTS (including separate areas of overlap in the angular gyrus bilaterally), bilaterally in more anterior temporal regions (MTG, STS), and the pCC and precuneus. An additional portion of overlap occurred in the left IFG when a smaller cluster (<100 mm³) from the story-based analysis was taken into account (Figure 2A). One additional area was also activated by both story-based and nonstory-based ToM studies, but with a key difference in lateralization. The amygdala was activated on the left for story-based studies but on the right for nonstory-based studies (Figure 1C). Language is typically lateralized to the left hemisphere for right-handed individuals (Bookheimer 2002, Frost et al. 1999), which may explain why these activations appear on the left for
story-based ToM studies. Complicating this interpretation is a meta-analysis of amygdalic activations that found no relation between language-related stimuli and hemispheric lateralization (Baas et al. 2004; cf. Markowitsch 1998, Olsson & Phelps 2007). A large proportion of the studies examined in this meta-analysis, however, employed nonverbal materials (42 of 54, only 8 employed strictly verbal stimuli; Bass et al. 2004). This imbalance may have hampered attempts to identify a systematic pattern of lateralization. Also, studies were only included in the analysis if they mentioned lateralization or asymmetry in the abstract or keywords, which means that studies reporting unilateral amygdala activation strictly within the body of a paper were not included. The current meta-analyses are an informative contribution to the debate on amygdalic lateralization, despite the fact they were not specifically designed to answer this question. With relatively large groups of story and nonstory-based studies and a quantitative approach with no a priori assumptions regarding the amygdala, the observed language-based lateralization of activation is a reliable and interesting result. This is particularly true in light of the closely mirrored locations for the left and right amygdala. The amygdala’s role during mentalizing may be predominately stimulus driven, reflecting evaluations of the motivational or arousal value of valenced stimuli (Olsson & Ochsner 2008).

Understanding the Components of the Mentalizing Network

A number of regions identified by the convergence of story- and nonstory-based studies of ToM have been the subject of extended debate within the ToM literature. Briefly, the mPFC has been hypothesized to support social cognitive processes such as mentalizing, person perception, and self processing, as well as nonsocial functions involving attention and coherence making (Amadio & Frith 2006, Ferstl & von Cramon 2002, Gilbert et al. 2007, Mitchell et al. 2005b, Saxe & Kanwisher 2003, Saxe & Powell 2006). A similar debate exists for the TPJ, with some arguing for a unique role in mental inference and others proposing a more general role related to the direction of attention (Decety & Lamm 2007, Mitchell 2008b, Saxe & Powell 2006, Scholz et al. 2009). In the case of the STS, this area has traditionally been associated with the perception of biological motion (Allison et al. 2000), although the results of a recent meta-analysis argue that this structure is also engaged during a variety of other tasks, with its function likely changing depending on the coactivating network (Hein & Knight 2008). (For more on the mPFC, TPJ, and STS, see the previously mentioned reviews on ToM, which typically discuss these areas at length.)

Inferior Frontal Gyrus. Although the left IFG has long been known as Broca’s area and associated with language production (Lorch 2008), the belief that this region only functions to produce language is now untenable (Bookheimer 2002). For one, Broca’s area also supports other aspects of language such as comprehension, responding to syntactic complexity (Caplan et al. 2000). The functions of this region also extend beyond the realm of language, to the processing of music, for example (Koelsch 2006, Patel 2003). Schubotz & Fiebach (2006) recently edited a special issue devoted to integrative models of Broca’s area, acknowledging that different research perspectives have adopted this area and ascribed separate domain-specific functions. In contrast to the language researchers mentioned above, those who study action perception have long identified Broca’s area as a key region for the perception of biological motion (e.g., Schubotz & von Cramon 2004). This belief is partly based on the observation that the IFG in humans is homologous to a motor region of the macaque brain (F5); this region has demonstrated “mirror” properties, becoming active both during observation of an action and during
performance of that same action (for human data, see Molnar-Szakacs et al. 2005; for a review, see Iacoboni 2009; for a critique, see Hickok 2009). A number of theorists have proposed that these “mirror neurons” support ToM and social understanding in humans through the simulation of a target person’s actions (e.g., Blakemore & Decety 2001, Blakemore & Frith 2005, Gallese et al. 2004, Iacoboni 2009, Keysers & Perrett 2004, Prinz 2006, Rizzolatti & Craighero 2004). Although the mirror properties of Broca’s area may appear to be an attractive explanation for why this area is so commonly observed in studies of mental inference (particularly nonstory-based paradigms), recent examinations have provided evidence that the mirror neuron network and the mentalizing network are quite separate (Van Overwalle & Baetens 2009, Wheatley et al. 2007; cf. Zaki & Ochsner 2009).

Previous reviewers of ToM may have been reluctant to list the IFG as part of the mentalizing network due to its traditional association with language or perhaps due to the current lack of consensus on its function. One other reason may be the type of studies examined, as the presence of this activation was far more reliable for nonstory-based studies of ToM than for story-based studies. Those who rely on the latter are thus less likely to observe activations in this region and may subsequently fail to consider it an element of this network.

An integrative account of IFG function may best explain how this structure can be involved in such seemingly disparate processes. Some have attempted to reconcile its role in language processing and action perception by proposing a link between the evolution of language and the emergence of gesture, imitation, and symbolic thinking (Rizzolatti & Arbib 1998). Others have hypothesized that a single process may underlie its role in all these tasks, such as high-level sequencing of complex information; this would be useful for decoding language, music, and causal actions (Fiebach & Schubotz 2006). The cluster observed in the present study tended to be more anterior (i.e., pars orbitalis and triangularis, BA 45/47) than is Broca’s area proper (pars triangularis and opercularis, BA 44/45), although similar theories have been put forth regarding these more anterior regions and high-level ordering, sequencing, or processing of temporal coherence (Levitin & Menon 2003). Attempts to integrate findings from various subdisciplines of neuroscience are likely to provide the most fruitful path to understanding how the IFG contributes to mental inference.

Temporal poles/anterior temporal lobes. Although there were no overlapping voxels in the temporal poles for story- and nonstory-based ToM studies, the two groups of studies overlapped in anterior temporal regions, and both had separate clusters in the left temporal pole. The temporal pole, like the IFG, has been linked to a variety of different processes. It is known as part of the limbic system (Heimer & Von Hoesen 2006) and has been linked to the processing of emotional and arousing stimuli (Beauregard et al. 2001, Berthoz et al. 2002b). Others have observed that it plays a major role in language comprehension, particularly at the discourse level (Maguire et al. 1999; for reviews, see Ferstl et al. 2008, Mar 2004). There are also some indications that it may be involved in the processing of music (Koelsch 2006) or auditory object features more generally (Zatorre et al. 2004). Some researchers have proposed functions more in keeping with social processing, such as face processing (Mesulam 1998, Seeck et al. 1993), the storage of social concepts (Zahn et al. 2007), or the processing of social semantic knowledge (Ross & Olson 2010, Simmons et al. 2010) (for a review, see Olson et al. 2007). Jung-Beeman and colleagues (2004, Kounios & Beeman 2009) have shown that the right anterior temporal lobe may be involved in insight problem solving, specifically the drawing of associations between remote pieces of information. The right aSTG region identified by these researchers is very close to the overlap observed for story-based and nonstory-based ToM studies (Figure 2D; Jung-Beeman et al. 2004, figure 2). Many of these processes (language, music, face processing, insight) appear to
involve a binding together of related information to produce a whole, which could form a core function of the anterior temporal lobes.

**Posterior cingulate/precuneus.** The pCC has often been omitted from reviews of ToM, and when it is noted it quite often remains undiscussed (Adolphs 2009, Beer & Ochsner 2006, Mitchell 2008a, Saxe 2006a). This omission is somewhat surprising, considering how likely it is to be implicated by studies of ToM according to the meta-analyses reported here. Perhaps this region has been neglected because it has been associated with a diverse range of possible functions, as has its closely interconnected neighbor the precuneus (Cavanna & Trimble 2006). Resting-state functional connectivity has revealed a number of distinct regions within this medial parietal region, with the cluster overlap for ToM appearing most centrally in the ventral aspect of the central precuneus extending into the pCC; this region has strong connections to the pCC, MTG, and angular gyrus (Margulies et al. 2009, seed 16). Cavanna & Trimble (2006) reviewed the functional correlates of the precuneus, naming as associate processes visuospatial imagery, episodic memory, ToM, and feelings of agency. These authors propose that the precuneus can be understood as having two separate functional centers: an anterior portion that supports mental imagery involving the self and a posterior region that supports the retrieval of episodic memories. The region of interest from this analysis is located within the central portion of the precuneus, a region that may be linked to cognition and the drawing of associations (Margulies et al. 2009). With respect to the pCC, a unique meta-analysis based on text analysis found that the posterior portion (closest to the precuneus) was linked to memory processes, with the more anterior portion associated with the perception of pain (Nielsen et al. 2005). Hagmann and colleagues (2008) have shown that together these medial parietal regions are highly connected to the rest of the brain, forming a “processing core” that is likely important for functional integration and the coordination of processes in both hemispheres. Thus, these medial parietal regions may support the imagery and imagination processes required to infer the mental states of another, integrating inputs from a wide variety of other brain regions that support memory, motor, and somatosensory processing.

**Understanding the Relation Between Stories and ToM**

Another goal of this review was to examine the relation between story processing and ToM. To this end, an ALE meta-analysis of narrative comprehension was conducted, identifying a number of areas that overlap with the mentalizing network. Perhaps somewhat unexpectedly, narrative comprehension illustrated greater overlap with nonstory-based studies of ToM relative to the story-based examinations. This may be attributed to the larger and more numerous clusters observed for the nonstory analysis, which in turn is a product of the greater number of foci and more diverse designs employed here compared to the story-based analysis.

The core mentalizing network, defined as the overlap between story- and nonstory-based studies, demonstrated an overlap with narrative comprehension in the mPFC, bilateral pSTS/TPJ, bilateral aMTG, and also the left IFG (pars opercularis) if a small story-based cluster is taken into account. There were some notable absences in this overlap. For one, the narrative comprehension analysis did not reveal any clusters in the pCC or precuneus. Although a qualitative review had previously identified this region as an important correlate of narrative comprehension studies (Mar 2004), a subsequent ALE meta-analysis on the topic found that the pCC was only present in contrasts that specifically isolated the processing of textual coherence (Ferstl et al. 2008). The earlier qualitative review included studies that were intended to measure ToM (Mar 2004), which may explain this divergent result. It may be that the pCC is only involved in narrative comprehension during parts of the story that
are especially demanding with respect to visual imagery or are likely to evoke personal memories (Larsen & Seilman 1988). Speer and colleagues (2009) found medial parietal activations during changes in character, time, objects, and goals, but not causal or spatial changes, demonstrating that engagement of this region may depend on story content. Similarly, Mano and colleagues (2009) found that the pCC was specifically engaged during perspective taking in stories.

Another core ToM region that did not overlap with narrative comprehension is the angular gyrus (bilaterally). The overlap of ToM studies revealed two distinct clusters in the posterior lateral temporal cortex, a pSTS/TPJ cluster and a separate cluster in the angular gyrus (for both hemispheres), with narrative comprehension only overlapping with the more anterior pSTS/TPJ cluster. The angular gyrus has clear associations with language comprehension (Carreiras et al. 2009, Dronkers et al. 2004), although it may support lower-level language processes such as semantic access (Graves et al. 2010) and therefore not appear in contrasts that isolate narrative-level processes. It is unclear how the angular gyri contribute to ToM and how this function differs from the pSTS/TPJ (which does overlap with narrative comprehension), but some possibilities include sequencing (Rosenthal et al. 2009) or an aspect of scene construction (Summerfield et al. 2010).

What could be the meaning of this overlap between ToM and narrative comprehension? One possibility is that ToM processes are employed during narrative comprehension, as readers infer the mental states of characters in a manner similar to how mental states are inferred in real-life conspecifics (Aastington 1990, Bruner 1986, Keen 2007, Mar 2004, Mar & Oatley 2008, Mason & Just 2009, Pe-skin & Aastington 2004, Zunshine 2006). This is in keeping with recent work on anthropomorphization (Kwan & Fiske 2008), which has shown that people can treat fictional persons as if they were real (Epley et al. 2007) and that these fictional others can serve a social function. The mere presence of fictional others can relieve feelings of loneliness and isolation (Derrick et al. 2009, Epley et al. 2008), for example, or produce social psychological phenomena such as social facilitation (Gardner & Knowles 2008). Similarity should not be confused with identity, however, and there is growing evidence that perceptions of intentionality may vary along a continuum for targets that are actually intentional as well as those seen as less intentional. The level of intentionality attributed to other minds appears to be determined by varying levels of two dimensions, labeled as either warmth and competence (Fiske et al. 2007) or experience and agency (Gray et al. 2007). Neural differences in the perception of real social agents and anthropomorphized nonsocial agents might be reflections of varying levels of attribution for these two dimensions (Chaminade et al. 2007, Mar et al. 2007; for dehumanized social agents, see Harris & Fiske 2008).

That ToM processes are employed during story comprehension is not the only possible explanation, however. Ferstl and colleagues (2008), for example, have argued against this explanation, proposing instead that a more general process (or set of processes) underlies the activation of the mPFC, perhaps the initiation and maintenance of volitional cognitive processes (Ferstl & von Cramon 2002). In light of the results reported here, however, it appears that the overlap in the ToM and narrative comprehension networks is greater than previously thought, necessitating an expansion of these ideas.

Along these lines, theorists have observed that this wider network of brain regions (mPFC, bTPJ, pCC) is similar across a range of cognitive processes (e.g., autobiographical memory, future thinking, spatial navigation, mind-wandering and stimulus-independent thought, creativity, sleep-related cognition, ToM, and narrative comprehension; Buckner & Carroll 2007; Christoff et al. 2009, 2010; Mason et al. 2007; Raichle et al. 2001). Empirical support for the existence of an overlapping network underlying many of these processes has begun to be established (Rabin et al.
2010, Spreng et al. 2009, Spreng & Grady 2010, Wilson et al. 2008). Various theoretical accounts for what underlies this network have also been put forth, including projection of the self (Buckner & Carroll 2007, Mitchell 2009), scene construction (Hassabis & Maguire 2007, Summerfield et al. 2010), associative processing (Bar et al. 2007), or the integration of motivational systems with language/categorization systems (Xu et al. 2005). The current meta-analysis provides evidence for a set of brain regions shared between ToM and story comprehension, linking the latter to this group of processes. This is consistent with a previous demonstration that the narrative comprehension network is similar to that observed during stimulus-independent thought (Wilson et al. 2008). Networks evoked by all these processes are similar but not necessarily identical, however. Future work should explore more carefully the divergences between various process networks.

Study Limitations
There are several notable limitations of this study. One obvious shortcoming is that only neuroimaging research is represented in this analysis, and the type of inferences that can be made regarding brain function using these methods is limited in some ways (Logothetis 2008). For example, areas identified using these methods can only be said to be associated with target processes and not necessarily essential components of a functional network. The findings presented here should be considered a complement to the important neuropsychological work that has been done with brain-damaged populations, work that can support causal inferences (although issues of generalizing to nondamaged persons cannot be ignored). Early research in the ToM domain implicated the frontal lobes (Bach et al. 1998, Happé et al. 2001), although recent work has found stronger evidence for the TPJ (Apperly et al. 2004, Bird et al. 2004). Some of this work also employed a story-based format, indicating that the right hemisphere plays a special role in understanding intentions in stories (Winner et al. 1998). Consistent with this idea, the patient literature on narrative comprehension appears to implicate the right hemisphere as well as the frontal lobes (Mar 2004). Some patient work has also begun to examine the relations between the many processes that draw on this shared network. Rosenbaum and colleagues (2007), for example, have shown that two patients incapable of retrieving episodic memories can successfully pass tests of ToM. Performance on tasks that share this core network is therefore dissociable.

A limitation also exists in employing the ALE methodology, with meta-analyses based on full statistic images demonstrating superior results to approaches that rely on aggregating peak coordinates (Salimi-Khorshidi et al. 2009). Unfortunately, gaining access to the full results images of a large group of studies is not currently feasible. A shared database of images from neuroimaging studies would be a great advantage for all in the neuroscience community.

CONCLUSIONS
A quantitative meta-analytic approach to examining ToM, story comprehension, and their interrelation has yielded some unique results. First, the core mentalizing network, as identified by the overlap of reliable activations from two approaches to ToM, is larger than most describe, including the mPFC, bilateral pSTS, bilateral angular gyri, bilateral anterior temporal areas, pCC and precuneus, and possibly the left IFG. Second, this mentalizing network overlaps with the narrative comprehension network in a number of areas, including the mPFC, bilateral pSTS/TPJ, bilateral anterior temporal areas, and possibly the left IFG. Together, these analyses have confirmed that a shared network exists for ToM and narrative comprehension, which resembles in some ways a network implicated in numerous other processes (Spreng et al. 2009). Adopting a quantitative approach to meta-analysis has allowed for the illumination of several issues pertaining to social cognition and story comprehension, but a number of future issues remain to be explored.
FUTURE ISSUES

1. The analyses presented here have identified the brain regions associated with measures of ToM, bringing to light a larger network of regions than is typically discussed. An important future project will be determining which of these areas are necessary for mentalizing. This may be especially pertinent for ToM areas identified here that are not commonly associated with this process, such as the IFG. Answering such a question will require a synthesis of various approaches, including research with brain-damaged patients, transcranial magnetic stimulation, repetition-suppression and parametric designs, along with animal work.

2. Individual studies have found medial parietal regions to be involved in narrative comprehension, yet the ALE meta-analysis found no reliable activations in this area, taking a number of studies into account. Identifying the moderators of this association will likely bring us closer to understanding the precise role of the pCC and precuneus in the context of narrative comprehension.

3. The core network involved in ToM overlaps with the regions associated with narrative comprehension, but the precise relation between these two processes remains to be elucidated. Examining the regions of overlap and nonoverlap is likely to provide helpful clues so that future studies can examine this question more directly.

4. Moreover, the shared regions of ToM and narrative comprehension resemble in many ways a network that has been associated with a great number of other processes (e.g., daydreaming, future-thinking, and autobiographical memory). Determining whether this core network supports a single function that underlies all of these processes, or whether it can support many different functions determined by the neural context and task at hand, is a key issue for future research.

5. Another key issue is how to better understand anthropomorphization. How is inferring the mental states of fictional others (e.g., characters in a book) similar to and different from inferring the mental states of real others? What determines our tendency to attribute minds to a target, whether that target is actually intentional (e.g., dehumanized conspecifics) or not (e.g., humanized machines, animals, and forces such as weather)? Much of the work on mentalizing has assumed that making inferences about intentional and nonintentional targets is an identical process, but this assumption remains to be tested.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Thanks to Colin DeYoung and R. Nathan Spreng for helpful comments on early versions of this manuscript and to Kathy Kotnowski for assistance preparing this manuscript. Finally, a special thank you to Susan Fiske for her support.
LITERATURE CITED


Davis FC, Johnstone T, Mazzulla EC, Oler JA, Whalen PJ. 2010. Regional response differences across the human amygdaloid complex during social conditioning. *Cerebral Cortex* 20:612–21


Logothetis NK. 2008. What we can and what we cannot do with fMRI. *Nature* 869–78


Saxe R. 2006b. Why and how to study theory of mind with fMRI. Brain Res. 1079:57–65


Schubotz RI, Fiebach CJ. 2006. Integrative models of Broca’s area and the ventral premotor cortex. *Cortex* 42:461–63


Yarkoni T, Speer NK, McAvoy MP, Zacks JM. 2008a. Pictures of a thousand words: investigating the neural mechanisms of reading with extremely rapid event-related fMRI. *NeuroImage* 42:973–87


Figure 1
ALE results for story-based and nonstory-based ToM studies along with narrative comprehension studies: medial surface and coronal slice. (A) Medial surface at $x = -8$. (B) Medial surface at $x = 7$. (C) Coronal slice at $y = -2$; the left hemisphere is on the left. Clusters of reliable activations across studies for the three meta-analyses and their overlap. Figures were created using Mango (http://ric.uthscsa.edu/mango). See Supplemental Materials for full results files for all meta-analyses (.nii format).
Figure 2

ALE results for story-based and nonstory-based ToM studies along with narrative comprehension studies: 
left and right lateral surfaces. (A) Left lateral surface at $x = -43$. (B) Right lateral surface at $x = 53$. (C) Left lateral surface at $x = -49$. (D) Right lateral surface at $x = 57$. (E) Left lateral surface at $x = -53$. Clusters of reliable activations across studies for the three meta-analyses and their overlap. Figures were created using Mango ([http://ric.uthscsa.edu/mango](http://ric.uthscsa.edu/mango)). See Supplemental Materials for full results files for all meta-analyses (.nii format).
Annual Review of Psychology
Volume 62, 2011

Contents

Prefatory
The Development of Problem Solving in Young Children: A Critical Cognitive Skill
Rachel Keen ................................................................. 1

Decision Making
The Neuroscience of Social Decision-Making
James K. Rilling and Alan G. Sanfey ........................................... 23

Speech Perception
Speech Perception
Arthur G. Samuel ............................................................... 49

Attention and Performance
A Taxonomy of External and Internal Attention
Marvin M. Chun, Julie D. Golomb, and Nicholas B. Turk-Browne ............ 73

Language Processing
The Neural Bases of Social Cognition and Story Comprehension
Raymond A. Mar .......................................................... 103

Reasoning and Problem Solving
Causal Learning and Inference as a Rational Process:
The New Synthesis
Keith J. Holyoak and Patricia W. Cheng .......................................... 135

Emotional, Social, and Personality Development
Development in the Early Years: Socialization, Motor Development, and Consciousness
Claire B. Kopp ............................................................... 165

Peer Contagion in Child and Adolescent Social and Emotional Development
Thomas J. Dishion and Jessica M. Tipsord ..................................... 189
Adulthood and Aging
Psychological Wisdom Research: Commonalities and Differences in a Growing Field
Ursula M. Staudinger and Judith Glück ......................................................... 215

Development in the Family
Socialization Processes in the Family: Social and Emotional Development
Joan E. Grusec ................................................................. 243

Psychopathology
Delusional Belief
Max Coltheart, Robyn Langdon, and Ryan McKay ........................................ 271

Therapy for Specific Problems
Long-Term Impact of Prevention Programs to Promote Effective Parenting: Lasting Effects but Uncertain Processes
Irwin N. Sandler, Erin N. Schoenfelder, Shartlene A. Wolchik, and David P. MacKinnon ................................................................. 299

Self and Identity
Do Conscious Thoughts Cause Behavior?
Roy F. Baumeister, E.J. Masicampo, and Kathleen D. Vohs ................................ 331

Neuroscience of Self and Self-Regulation
Todd F. Heatherton ................................................................. 363

Attitude Change and Persuasion
Attitudes and Attitude Change
Gerd Bohner and Nina Dickel ................................................................. 391

Cross-Country or Regional Comparisons
Culture, Mind, and the Brain: Current Evidence and Future Directions
Shinobu Kitayama and Ayse K. Uskul ................................................................. 419

Cognition in Organizations
Heuristic Decision Making
Gerd Gigerenzer and Wolfgang Gaissmaier ...................................................... 451

Structures and Goals of Educational Settings
Early Care, Education, and Child Development
Deborah A. Phillips and Amy E. Lowenstein ...................................................... 483
Psychophysiological Disorders and Psychological Dimensions on Medical Disorders

Psychological Perspectives on Pathways Linking Socioeconomic Status and Physical Health
Karen A. Matthews and Linda C. Gallo .................................................. 501

Psychological Science on Pregnancy: Stress Processes, Biopsychosocial Models, and Emerging Research Issues
Christine Dunkel Schetter .......................................................... 531

Research Methodology

The Development of Autobiographical Memory
Robyn Fivush .......................................................... 559

The Disaggregation of Within-Person and Between-Person Effects in Longitudinal Models of Change
Patrick J. Curran and Daniel J. Bauer .................................................. 583

Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP)
Marta Kutas and Kara D. Federmeier .................................................. 621

Indexes

Cumulative Index of Contributing Authors, Volumes 52–62 ......................... 000
Cumulative Index of Chapter Titles, Volumes 52–62 ................................ 000

Errata

An online log of corrections to Annual Review of Psychology articles may be found at http://psych.AnnualReviews.org/errata.shtml