



## PAPER

# Neural correlates of belief- and desire-reasoning in 7- and 8-year-old children: an event-related potential study

Lindsay C. Bowman,<sup>1</sup> David Liu,<sup>2</sup> Andrew N. Meltzoff<sup>3</sup> and Henry M. Wellman<sup>1</sup>

1. Department of Psychology, University of Michigan, USA

2. Department of Psychology, University of California, San Diego, USA

3. Institute for Learning and Brain Sciences, University of Washington, USA

## Abstract

*Theory of mind requires belief- and desire-understanding. Event-related brain potential (ERP) research on belief- and desire-reasoning in adults found mid-frontal activations for both desires and beliefs, and selective right-posterior activations only for beliefs. Developmentally, children understand desires before beliefs; thus, a critical question concerns whether neural specialization for belief-reasoning exists in childhood or develops later. Neural activity was recorded as 7- and 8-year-olds (N = 18) performed the same diverse-desires, diverse-beliefs, and physical control tasks used in a previous adult ERP study. Like adults, mid-frontal scalp activations were found for belief- and desire-reasoning. Moreover, analyses using correct trials alone yielded selective right-posterior activations for belief-reasoning. Results suggest developmental links between increasingly accurate understanding of complex mental states and neural specialization supporting this understanding.*

## Introduction

Having a ‘Theory of Mind’ – the understanding that people’s actions are guided by internal mental states such as beliefs, desires, and intentions – is a cornerstone of social-cognitive development (Harris, 2006; Wellman, 2002). Theory of mind goes beyond social perception; it requires conceptualization of and reasoning about people’s mental states in order to accurately predict and explain behaviour.

Important achievements in development of theory of mind are manifest during the preschool years (e.g. Wellman & Liu, 2004). An often-studied, clear example of these achievements is the development of false-belief understanding; for example, understanding that a girl, Sarah, could believe her dog was in the shed when in reality it was at the playground. Numerous studies show that children transition from consistently failing to consistently passing standard false-belief tasks in the preschool and early school years (e.g. Wellman, Cross & Watson, 2001), demonstrating development of an explicit understanding of beliefs and of the mind more generally. However, focus on beliefs and false beliefs alone is limited. Theory of mind – often termed a belief-desire or belief-desire-emotion naïve psychology – involves understanding multiple causally interconnected mental concepts, and developmentally, children’s theory of mind proceeds in a progression of mental-state understandings.

A crucial, well-documented progression is that children consistently develop an explicit understanding of desires before developing an explicit understanding of beliefs – they can understand that Sarah *wants* to find her dog before understanding that she *thinks* it is in the shed (e.g. Bartsch & Wellman, 1995; Gopnik & Slaughter, 1991; Wellman & Liu, 2004). The purpose of the present study is to examine, using event-related brain potential (ERP) measures, the neural mechanisms underlying understanding of both desires and beliefs in children – thereby providing insights into children’s developing understanding of the mind.

The extant findings of understanding desires before beliefs hold across tasks matched on procedural methodology, linguistic structure, and materials. For example, one can directly compare diverse-desires versus diverse-beliefs tasks (used by Wellman & Liu, 2004; and also Wellman & Woolley, 1990). For the diverse-desires task, children are told about a character who *likes* a particular snack opposite to what the child prefers (e.g. likes carrots not cookies). Children are asked to predict which food the character will choose for a snack (the carrot or the cookie). For the matching diverse-beliefs task, children are told about a character who *thinks* his cat is hiding in a location opposite to what the child thinks (e.g. he thinks the cat is in the garage but the child thinks the cat is in the tree). Children are asked to predict where the character will look for his cat (in the garage or in the

Address for correspondence: Lindsay C. Bowman, Department of Psychology, 1012 East Hall, 530 Church St., Ann Arbor, MI, 48109-1043, USA; e-mail: lcbowman@umich.edu

tree). The demands and format for the two tasks are virtually identical except that children must predict behaviour based on different mental concepts (desires versus beliefs). Children consistently pass diverse-desires tasks at an earlier age than diverse-beliefs tasks. Recent looking-time studies on infants' false-belief understanding (e.g. Onishi & Baillargeon, 2005) have added fuel to the debate about the age at which children possess the capacity to process others' beliefs, at least implicitly. Nevertheless, studies of explicit performance strongly support a consistent developmental progression of explicit understanding of desires before explicit understanding of beliefs.

What mechanisms underlie this progression of understanding desires before understanding beliefs? To date, the neurocognitive mechanisms underlying this change remain essentially unexplored. We address this issue in the current research.

Functional neuroimaging and electrophysiological studies with adults converge on findings that theory-of-mind reasoning recruits neural substrates in medial pre-frontal cortex (MPFC) and temporoparietal junction (TPJ) (see Gallagher & Frith, 2003; Saxe, Carey & Kanwisher, 2004, for reviews). Some of this research suggests that some substrates in particular (e.g. TPJ) are especially recruited to process information about complex mental-states (e.g. beliefs). For example, fMRI data show that for adults, TPJ is recruited for belief-processing over processing of physical or general social characteristics (i.e. a person's appearance) (Saxe & Powell, 2006), and that this recruitment is independent of recruitment for executive functioning (Saxe, Shulz & Jiang, 2006).

Most relevant to the current study, Liu, Meltzoff and Wellman (2009a) recorded ERPs as adult participants performed tasks requiring reasoning about diverse-desires and diverse-beliefs (and as a control, about parallel physical situations). A late slow wave (LSW) with mid-frontal scalp distribution was associated with desire- and belief-judgments. However, an LSW with right-posterior scalp distribution was associated *only* with belief-judgments. These findings demonstrate neural overlap as well as critical differences in reasoning about desires and beliefs, and point to a possible explanation for the progression of understanding desires to understanding beliefs seen in children. Children may need to recruit additional neural processes (within posterior parietal regions) for reasoning about beliefs beyond a common neural system (within medial frontal regions) for reasoning about mental states more generally (Liu *et al.*, 2009a).

It is not possible to determine the developmental process of neural specializations solely from adult research (Karmiloff-Smith, 1997). Recent data from adults and children, however, begin to suggest that neural circuitry associated with TPJ may become increasingly specialized for mental-state reasoning with development. Sabbagh, Bowman, Evraire and Ito (2009)

found that, in 4-year-old children, increased functional maturation of the dorsal MPFC and the right TPJ was associated with better performance on a battery of representational theory-of-mind tasks. In their fMRI study, Saxe, Whitfield-Gabrieli, Scholz and Pelphrey (2009) found that children 6 to 11 years old recruited both MPFC and TPJ for processing mental states and social interactions (compared to processing physical events); however, as children aged, specifically the TPJ was found to increase in selectivity for mental-state reasoning (in comparison to both reasoning about physical events *and* social interactions), with the oldest children (10- to 11-year-olds) showing the greatest selectivity (see also Kobayashi, Glover & Temple, 2007; Sommer, Meinhardt, Eichenmüller, Sodian, Döhnell & Hajak, 2010).

### *The present study*

These prior neuroscientific data leave much unknown. Only two studies, both with adults, have focused on specifically comparing belief- and desire-reasoning (Liu *et al.*, 2009a; Abraham, Rakoczy, Werning, von Cramon & Schubotz, 2010). We investigated the neural correlates of children's desire- and belief-reasoning. Given that the developmental trajectories for desire- and belief-reasoning differ, the crucial questions we tackled are the extent to which the underlying neural circuitries supporting belief- and desire-reasoning differ in childhood, and the nature of neural changes, if any, from childhood to adulthood. The Liu *et al.* (2009a) ERP study with adults provides an approach to address these questions and forms the backdrop for the present study; we used the same tasks and stimuli used by Liu *et al.* (diverse-desires and diverse-beliefs tasks with physical reasoning tasks as control) to collect ERP data from 7- and 8-year-old children. These children are old enough to endure the lengthy experimental procedure required by ERP methods (e.g. Liu, Sabbagh, Gehring & Wellman, 2009b), but are young enough that their belief- and desire-reasoning should not yet be as expert as adults' (see Friedman & Leslie, 2004a, 2004b).

Several informative patterns of results might appear. We expect activity from mid-frontal regions to be associated with both children's belief- and desire-reasoning, adding to the body of evidence demonstrating that medial frontal cortex is important for children's theory-of-mind reasoning in general (e.g. Liu *et al.*, 2009b; Sabbagh *et al.*, 2009). A critical focus, however, concerns patterns of neural processing that might differentiate belief- from desire-reasoning. The neural data, to the extent that they exist, suggest that increasing posterior-parietal specialization for belief-reasoning (i.e. a more focused and narrowed recruitment of neural substrates for that particular type of reasoning specifically) occurs only at later developmental time points (e.g. Saxe *et al.*, 2009). Thus, we may or may not see distinct neural activation associated with belief- versus desire-reasoning in posterior parietal regions for 7- and 8-year-old

children. In either event, we can compare the results of the present study to those of the parallel study conducted with adults (Liu *et al.*, 2009a). One possibility is that children's neural activation would show the same dissociable patterns found by Liu *et al.* (2009a). This pattern would suggest that distinct regions for belief- versus desire-reasoning exist robustly in children at least as young as 7 years old.

A second possibility is that children's neural activation would show use of the same neural circuitry to reason about both beliefs and desires. This pattern would suggest that neural specialization for belief-reasoning (over reasoning about other mental states) appears only later in development, as a function of further development and use. A third possibility is that children's neural activation would only *partly* resemble neural activation in adults, which would suggest both similarities and differences in children's neural circuitry for mental-state reasoning compared to adults. Regardless of which pattern occurs, the present study will help illuminate the cognitive and neural processes supporting a developing theory of mind.

## Methods

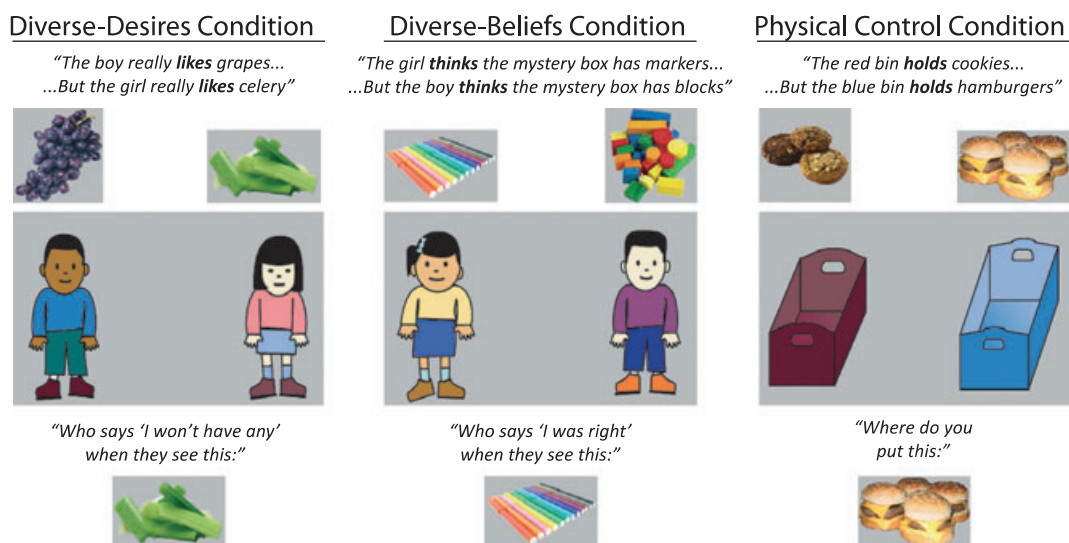
### Participants

Forty-one 7- to 8-year-old typically developing children (26 males) participated in the study. All participants were right handed, with normal or corrected-to-normal vision. Twenty-three participants did not provide at least 15 usable, artifact-free trials of electrophysiological data for each of the three conditions, yielding a final sample of 18 participants (range = 91–107 months;  $M = 97.28$  months,  $SD = 4.76$ ; 13 males) used in our full-trial

analyses (as described below). Exclusions of this magnitude are common for child EEG/ERP data collection (see Sabbagh *et al.*, 2009) due to limits in the total number of trials children can undertake and children's greater susceptibility for artifact (e.g. eye blinks, eye movements, head/body movements) compared to adults.

### Stimuli and procedure

To parallel Liu *et al.* (2009a) as closely as possible, we used the same tasks, stimuli, EEG recording system, and conditions: multi-trial diverse-desires, diverse-beliefs, and diverse-physical judgment tasks created to collect ERPs from participants. The structure of all 144 trials (48 in each of three conditions) was the same. In each trial, participants first received information about two characters with different desires for food/toys, two characters with different beliefs about food/toys, or two locations to put food/toys away. The experimenter sat beside participants and read information simultaneously presented on screen (e.g. 'the boy likes grapes, but the girl likes celery') as participants viewed accompanying pictures. On a random third of trials, the initial information was followed by a memory check to ensure that participants paid attention to each trial (e.g. 'who likes grapes?'). If participants answered the memory question incorrectly, the information phase was repeated. After the information phase of each trial (7300 ms in duration), participants were read the target question (details provided below for each condition) and were presented with a picture of one of the two food/toys (e.g. celery) for 2000 ms. This pictorial presentation of a single food/toy was the target visual event to which ERP data were time-locked. Participants then answered the target question (via button press). See Figure 1 for schematic summary of the tasks.



**Figure 1** A schematic of the theory-of-mind Diverse-Desires (left), Diverse-Beliefs (middle), and Diverse-Physical (right) tasks showing examples of information phase (top) and target questions (bottom) as well as sample graphics for both food and toy condition types.

### Desire reasoning

For the Desires condition, in each trial, the experimenter read about a boy who likes a particular food/toy (e.g. grapes/markers), a girl who likes a different food/toy (e.g. celery/blocks), and a closed box said to contain either a snack or a toy (with boy or girl first being counter-balanced). The experimenter then read one of four target questions about what would happen when the box was opened: 'Who says "I'll have some" when they see this?' / 'Who says "I won't have any" when they see this?' (when the story was about food), or 'Who says "I'll play with it" when they see this?' / 'Who says "I won't play with it" when they see this?' (when the story was about toys). The question presented (either positive or negative wording) was randomized in each trial. After the target question, participants were immediately presented with a picture of one of the two foods/toys (e.g. celery/blocks). Note that in this and other conditions, participants were not able to answer the target question until presented with a picture of the food/toy that was in the closed box. After 2000 ms of seeing the revealed food/toy, participants answered by choosing one of the two characters.

### Belief reasoning

The Beliefs condition followed the same presentation format (including counter-balancing and randomization) as just described, except that the content of the information and target questions was about beliefs. The experimenter read about a mystery box containing food/toys for a guessing game, and participants were told about a girl who thinks the box contains a particular food/toy and a boy who thinks the box contains a different food/toy. The target questions for this condition were: 'Who says "I was right" when they see this?' or 'Who says "I was wrong" when they see this?' After the target question, participants were presented with a picture of one of the two foods/toys, and then answered by choosing one of the two characters.

### Physical reasoning

The Physical condition provided a non-mental control condition (again following the same presentation format as the other two mental-state conditions). The experimenter read about a closed box containing food/toys to put away, and participants were told that the red bin should receive a particular food/toy while the blue bin should receive a different food/toy. Target questions for this condition were: 'Where do you put this?' or 'Where do you not put this?' After the target question, participants were presented with a picture of one of the two foods/toys that was in the closed box, and answered by choosing one of the two bins.

### Commonalities across conditions

Because trials for all three conditions were constructed to have the same perceptual and linguistic structure, any differences between conditions would point to the *mental-state processing* beyond these perceptual and task similarities.

Trials for each condition and type (e.g. desires condition about toys) were presented in blocks of six. Blocks semi-randomly alternated between the conditions, with the stipulation that no condition and type repeated successively. After every 2–3 blocks, there was a break to play with stickers that provided physical and mental reprieves for the children protecting against fatigue (10 sticker breaks overall, average duration 1–2 minutes). The EEG net was not removed during breaks although no EEG data were recorded. The total length of experiment, including breaks, was approximately 40–45 minutes.

### *Electrophysiological recording and analysis*

Participants sat while the experimenter applied and adjusted the EEG sensor net. Participants were given task instructions (described above) and EEG recording began. Parents were allowed to remain in the room with their child during recording, and both parent and child were asked to be as quiet as possible throughout the procedure. Net placement, adjustment, and instructions took approximately 15 minutes.

EEG was recorded continuously from scalp electrodes using the Geodesic Sensor Net (Tucker, 1993), a network of 128 Ag/AgCl electrodes embedded in an elastic geodesic tension structure. Impedance for all electrodes was kept below 50 K $\Omega$  (this EEG system used high-impedance amplifiers, thus the relatively high electrode impedances), and all recordings were referenced to the vertex (Cz). Signals were amplified with a 0.1 Hz to 100 Hz elliptical bandpass filter and digitized at a 250 Hz sampling rate. Continuous EEG data were segmented to epochs of 1500 ms after stimulus onset with a 100 ms pre-stimulus baseline.

Artifacts were identified in the EEG via the following steps. For each trial, channels were marked for artifact if a running average of activity exceeded 40  $\mu$ V (detecting sharp transitions in the signal). Subsequent to this automated process, each trial was manually inspected. Trials with more than 15 channels with artifact were excluded. For trials with less than 15 channels with artifact, an algorithm that derives values from neighbouring channels via spherical spline interpolation replaced bad channels. EEG data were corrected for eye-blink and eye-movement artifacts using the Gratton, Coles and Donchin (1983) algorithm. EEG data were re-referenced off-line against the average reference. Epochs of EEG data in the same condition were averaged to derive the ERP data. An average of 94 usable trials total (no less than 15 usable trials per condition)

per participant were used to derive the ERP data; the number of usable trials is consistent with previous developmental ERP research on social cognition (Liu *et al.*, 2009a; Liu *et al.*, 2009b). Prior to analysis, the ERP data were corrected to the 100 ms pre-stimulus baseline and digitally filtered with a 30 Hz low-pass filter.

## Results

### *Preliminary analyses*

Given the exclusions in our sample, we confirmed that included and excluded children did not differ on demographic or performance variables – pairwise *t*-tests indicated that these groups were equivalent (all *ps* > .40) on age, gender, and task performance variables (mean accuracy and mean reaction time for both memory control and target questions). Thus, our target group is representative of the entire sample.

As expected based on previous behavioural research, children were better at solving diverse-desires (84.6% correct) and physical control tasks (88.3%) compared to diverse-beliefs (65.0%); Beliefs condition versus Desires and Physical conditions,  $t(16) = -5.57, p < .001$ ,  $t(16) = -7.96, p < .001$ , respectively. Desires and Physical conditions did not differ from each other,  $t(16) = -1.38, ns$ . This pattern is consistent with numerous findings demonstrating that, compared to desire-understanding, belief-understanding emerges later in early childhood (see meta-analysis in Wellman & Liu, 2004), and can be less accurate/fluent in older children and adults (e.g. Malle, 2004). We dealt with this lesser accuracy for Beliefs by examining ERP patterns on all trials (full-trial analysis), as well as just correct trials (correct-trial analysis). Due to the reduced number of correct trials overall, two participants included in the full-trial analyses did not have at least 15 artifact-free correct trials per condition, and were excluded from correct-trial analyses. The final sample for *correct-trial* analyses was 16 children (12 males; age range = 91–106 months;  $M = 97.08$  months,  $SD = 4.12$ ;) – 16 of the same 18 participants as those used for *full-trial* analyses.

### *Full-trial analyses*

The Physical condition was designed as a control for the two mental-state conditions. Thus, differences in waveforms between Beliefs and Physical or between Desires and Physical conditions reveal components associated with reasoning about *mental states* over and above common components for processing these parallel task formats and making these comparative judgments; whereas differences between the Beliefs and Desires conditions reveal differences in processes for reasoning about *belief* states versus *desire* states.

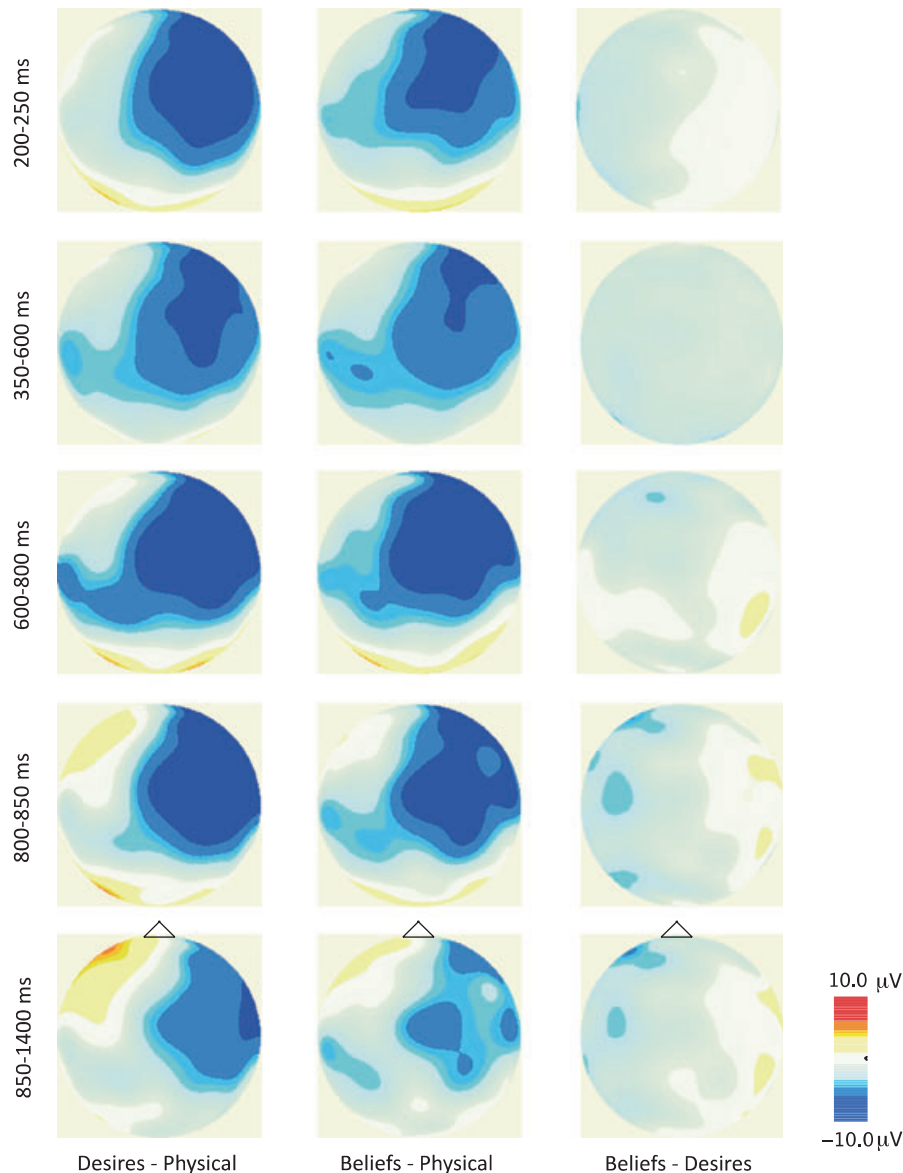
As a first step, difference waves were calculated by subtracting the mean amplitude in one condition from

the mean amplitude in another condition (i.e. Beliefs minus Physical, Desires minus Physical, and Beliefs minus Desires), at each of the 128 channels. Topographical maps of the difference waves show a clear difference between both Belief and Physical conditions and Desire and Physical conditions concentrated in right mid-frontal scalp regions (see Figure 2). These differences appear early around 200 ms post-stimulus onset and remain late beyond 800 ms post-stimulus. In contrast, no clear differences are apparent between Beliefs and Desires conditions at any point along the time-course. To confirm the condition differences revealed in the topographic maps, we conducted two analyses: analysis of a subset of channels systematically encompassing locations from left to right and top to bottom over the scalp, and a more focused region of interest (ROI) analysis.

### Subset analysis

Following a common analytic approach to avoid 128 multiple comparisons (e.g. Liu *et al.*, 2009a), we selected a grid of channels from the larger group of 128 channels to systematically sample locations from left to right (laterality) and from anterior to posterior (caudality) over the scalp. Figure 3 displays the grand average waveforms for all three conditions from all channels in this 5 × 5 grid. Visual inspection of the waveforms in Figure 3 matches the pattern displayed in the topographic maps: (a) no difference between the Beliefs and Desires conditions, but (b) a clear difference between the Physical condition and both mental-state conditions (more negative), concentrated in the right mid-frontal channels (e.g. channel FC2). Again, this difference appears early around 200 ms, and remains late beyond 800 ms post-stimulus. We examined these differences through analyses of the mean amplitude of the ERP data for all 25 channels in the grid within five different epochs: 200–250 ms, 350–600 ms, 600–800 ms, 800–850 ms, and 850–1400 ms post-stimulus onset. Each epoch represents a distinct portion of the waveform (i.e. peak, gradual increase in slope, attenuation of slope, horizontal slope, and tail-end of waveform, respectively). When necessary, *p*-values were adjusted using the Greenhouse-Geisser correction – a common adjustment in most ERP research (e.g. Liu *et al.*, 2009a; Van der Cruyssen, Van Duynslaeger, Cortoos & Van Overwalle, 2009). We adopted an alpha of .01 or less to be conservative, given the number of comparisons and relatively small sample size.

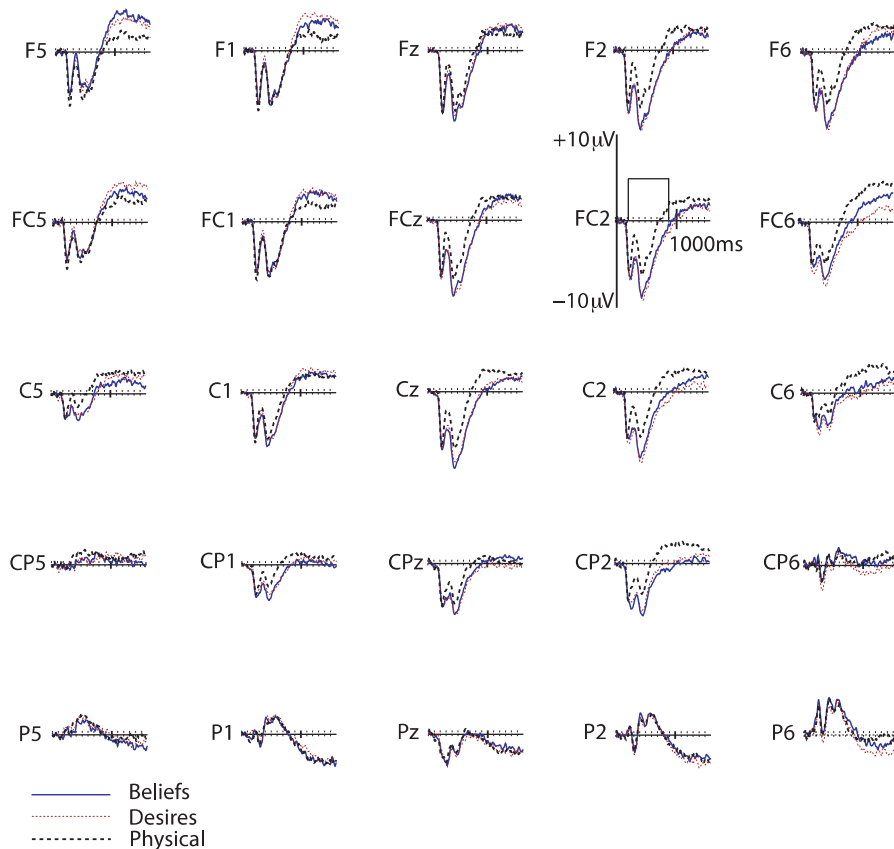
As expected, five separate omnibus 3 (condition) × 5 (laterality) × 5 (caudality) repeated measures analyses of variance (ANOVAs) (one for each of the five epochs) comparing the mean amplitude across all three conditions yielded significant condition effects in the 200–250, 350–600, 600–800, and 800–850 ms epochs (all *ps* < .002), but not in the 850–1400 epoch (*p* = .082). To further examine the condition differences, three



**Figure 2** Full-trial analyses topographical maps of the scalp electrical activity at all 128 channels. Maps depict mean amplitude difference for Physical subtracted from Desire (left), Physical subtracted from Belief (middle), and Desire subtracted from Belief (right) in the 200–250 ms (top), the 350–600 ms, 600–800 ms, 800–850 ms, and 850–1400 ms (bottom) post-stimulus epochs. Maps are oriented with frontal position up from an overhead scalp view. Darker regions indicate greater negative differences between two conditions. The maps show a clear difference in the 200–250, 350–600, 600–800, and 800–850 ms epochs for the Desire–Physical and Belief–Physical comparisons (Desires and Beliefs more negative compared to Physical) but do not show any clear difference for the Belief–Desire comparison.

2 (condition)  $\times$  5 (laterality)  $\times$  5 (caudality) repeated measures ANOVAs on the mean amplitude in each of the five epochs directly compared Desires versus Physical, Beliefs versus Physical, and Beliefs versus Desires conditions. We focus on main effects of condition as well as three-way condition  $\times$  laterality  $\times$  caudality interactions because our focal questions concern differences in ERPs across *conditions* and where these condition differences concentrate on the scalp. Main effects of caudality or laterality and two-way interactions (subsumed by the focal three-way patterns) are not considered further.

Both Beliefs–Physical and Desires–Physical comparisons yielded main effects of condition in the 200–250, 350–600, 600–800, and 800–850 ms epochs (top panel of Table 1). Both comparisons also yielded significant condition  $\times$  laterality  $\times$  caudality interactions in the 200–250, 350–600, 600–800, and 800–850 ms epochs (bottom panel of Table 1), evidencing a similar right-frontal activation in all of these epochs. In contrast, comparisons of Beliefs and Desires conditions showed no significant condition effects in any epoch and no significant condition  $\times$  laterality  $\times$  caudality interactions (see Table 1). Amplitude differences across conditions



**Figure 3** Full-trial analysis grand average event-related brain potential waveforms for the Desires (red dotted lines), Beliefs (blue solid lines), and Physical (black dashed lines) conditions from 25 channels selected to encompass locations from left to right and front to back (top to bottom in the grid) over the scalp. The 128 EGI channels can be grouped and labeled in terms of the 10–10 system convention. This figure adopts that more standard convention, with the selected channels reflecting 10–10 labels and positions. The boxed section indicates epochs in which analyses were performed. Waveforms are displayed with positive amplitudes above the axis and negative amplitudes below.

**Table 1** Full-trial subset analysis results for the 2(condition) × 5(laterality) × 5(caudality) repeated measures ANOVAs for all 5 epochs

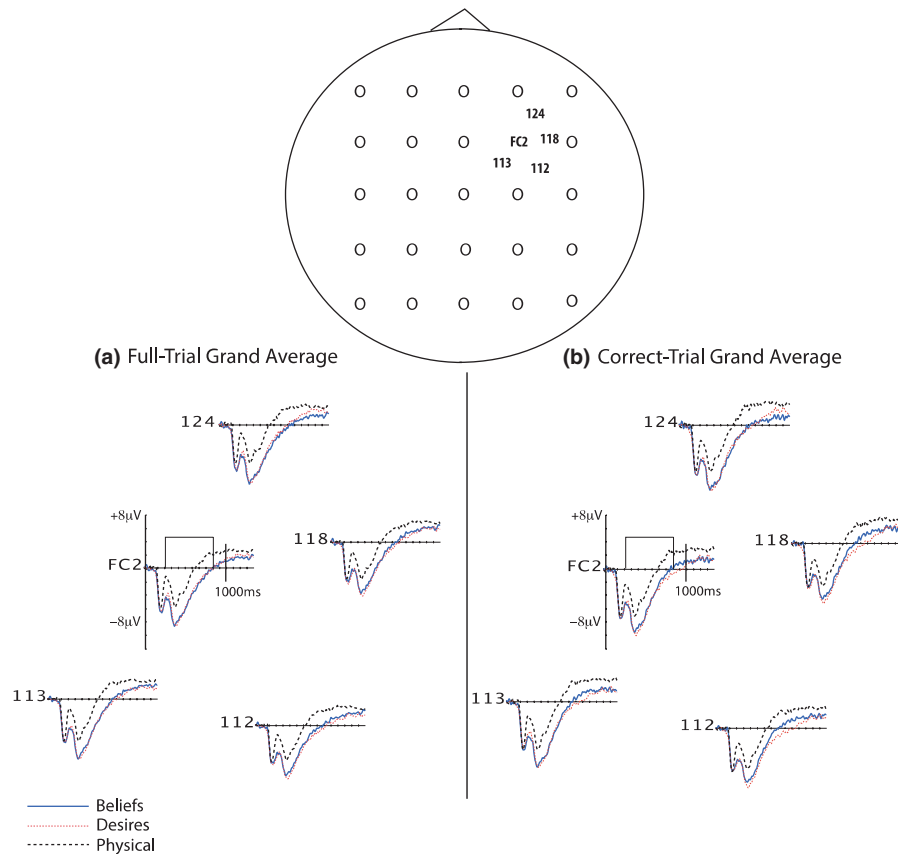
Comparison	Epoch (ms post stimulus onset)				
	200–250	350–600	600–800	800–850	850–1400
Condition effects					
Belief vs. Desire	$F(1,17) = 1.58$	$F(1,17) = .04$	$F(1,17) = .50$	$F(1,17) = .58$	$F(1,17) = .23$
Desire vs. Physical	$F(1,17) = 23.07^{***}$	$F(1,17) = 40.55^{***}$	$F(1,17) = 38.83^{***}$	$F(1,17) = 11.48^{**}$	$F(1,17) = 2.93$
Belief vs. Physical	$F(1,17) = 24.84^{***}$	$F(1,17) = 34.25^{***}$	$F(1,17) = 26.15^{***}$	$F(1,17) = 15.20^{**}$	$F(1,17) = 3.79$
Condition × laterality × caudality interaction effects					
Belief vs. Desire	$F(16,272) = 2.05$	$F(16,272) = 1.29$	$F(16,272) = 1.50$	$F(16,272) = 2.67$	$F(16,272) = 1.61$
Desire vs. Physical	$F(16,272) = 7.80^{**}$	$F(16,272) = 7.01^{***}$	$F(16,272) = 6.90^{***}$	$F(16,272) = 11.10^{***}$	$F(16,272) = 7.57^{***}$
Belief vs. Physical	$F(16,272) = 5.44^{**}$	$F(16,272) = 5.89^{***}$	$F(16,272) = 4.40^{**}$	$F(16,272) = 4.57^{**}$	$F(16,272) = 3.51^*$

Notes. \* indicates  $p < .05$ , \*\* indicates  $p < .01$ , \*\*\* indicates  $p < .001$

diminish after 850 ms; there were no significant condition effects in the 850–1400 ms epoch, and only one significant condition × laterality × caudality interaction (for Desires versus Physical). Thus as is clear in Figure 3, from as early as 200 ms post-stimulus and sustained up to 850 ms post-stimulus, the ERP waveforms for both mental-state conditions (Beliefs and Desires) differ significantly from the Physical control condition, but they do not differ significantly from each other.

#### Right-frontal ROI analysis

A further ROI analysis confirmed that this frontal effect encompassed a cluster of channels. In line with other ERP investigations (e.g. Sabbagh, Moulson & Harkness, 2004), cluster effects were considered significant only when condition differences occurred in at least five adjacent channels, and each channel showed a similar pattern of activity that achieved significance at the .05



**Figure 4** **a and b** Grand average event-related brain potential waveforms for the Desires (red dotted lines), Beliefs (blue solid lines), and Physical (black dashed lines) for the right mid-frontal five-channel cluster for full-trial (A. left) and correct-trial (B. right) analyses. The boxed section indicates epochs in which the Beliefs and Desires conditions showed the strongest significant differentiation from the Physical Condition. Waveforms are displayed with positive amplitudes above the axis and negative amplitudes below. Scalp positions for the five-channel cluster (bold channels) relative to the  $5 \times 5$  grid are shown at the top.

level or better. As shown in Figures 3 and 4, right-frontal channel FC2 from within our grid is located roughly in the centre of the group of channels exhibiting the visible difference between mental-state and Physical conditions. Thus, channel FC2 along with four other neighbouring channels (EGI channels 112, 113, 118, and 124) constituted our five-channel cluster for this ROI analysis (see Figure 4a).

Three 2 (condition)  $\times$  5 (channel) repeated measures ANOVAs on the mean amplitude in each of the five epochs established above directly compared Desires–Physical, Beliefs–Physical, and Beliefs–Desires conditions. Main effects of condition (and not condition  $\times$  channel interactions) are focal because our question concerns the extent to which condition effects are observed within our channel *cluster*. These ANOVAs showed that both Beliefs–Physical and Desires–Physical comparisons yielded main effects of condition in the 200–250, 350–600, 600–800, and 800–850 ms epochs (see Table 2). These condition differences diminished after 850 ms (no condition effects were significant in the 850–1400 ms epoch). Beliefs–Desires comparisons showed no significant condition effects in any of the epochs. Paired-samples *t*-tests confirmed that Beliefs and Desires conditions did not differ from each other in any of the five

channels within the cluster, at any of the five epochs ( $ps > .23$ ). In contrast, at all five channels, both Beliefs and Desires differed significantly from the Physical condition in the 200–250, 350–600, 600–800, and 800–850 ms epochs ( $ps < .005$ ). This cluster of effects converges directly with the earlier analyses, and further confirms what is shown clearly in Figures 2, 3, and 4a: in right mid-frontal scalp regions, Beliefs and Desires conditions differ equally from the Physical control condition but do not differ from each other.<sup>1</sup>

Repeated measures ANOVAs for a contrasting cluster of five channels in equivalent positions, except for the

<sup>1</sup> Also clear in Figures 2 and 3, the mental-state versus control condition effect appears to include regions beyond a five-channel cluster. Indeed, considering a cluster of 10 channels – the five above and five additional neighbouring channels (EGI channels 3, 4, 106, 107, and 111) – additional 2 (condition)  $\times$  10 (channel) repeated measures ANOVAs along with *t*-tests at all 10 channels in each of the same five epochs yielded results identical to those of the five-channel cluster analyses: Beliefs–Physical and Desire–Physical condition differences occurred in the 200–250, 350–600, 600–800, and 800–850 epochs in the overall models and at all 10 individual channels (with effects diminishing in the 850–1400 ms epoch) whereas Beliefs and Desires conditions did not differ significantly from each other in any model or at any channel, in any epoch.



**Table 2** Full-trial ROI analysis condition effects for the 2(condition) × 5(channel) repeated measures ANOVAs: right mid-frontal 5-channel cluster

Comparison	Epoch (ms post stimulus onset)				
	200–250	350–600	600–800	800–850	850–1400
Belief vs. Desire	$F(1,17) = .23$	$F(1,17) = .06$	$F(1,17) = .75$	$F(1,17) = .03$	$F(1,17) = .05$
Desire vs. Physical	$F(1,17) = 32.09^{***}$	$F(1,17) = 33.71^{***}$	$F(1,17) = 23.92^{***}$	$F(1,17) = 19.92^{***}$	$F(1,17) = 6.68^*$
Belief vs. Physical	$F(1,17) = 19.10^{***}$	$F(1,17) = 23.22^{***}$	$F(1,17) = 21.63^{***}$	$F(1,17) = 17.34^{**}$	$F(1,17) = 6.27^*$

Notes. \* indicates  $p < .05$ , \*\* indicates  $p < .01$ , \*\*\* indicates  $p < .001$

important feature that they are on the left rather than right-frontal side of the scalp, yielded no significant condition effects for any of the condition pairs in any epoch. Thus, the mental-state versus control condition effect, though diffuse, is concentrated in the right mid-frontal scalp.

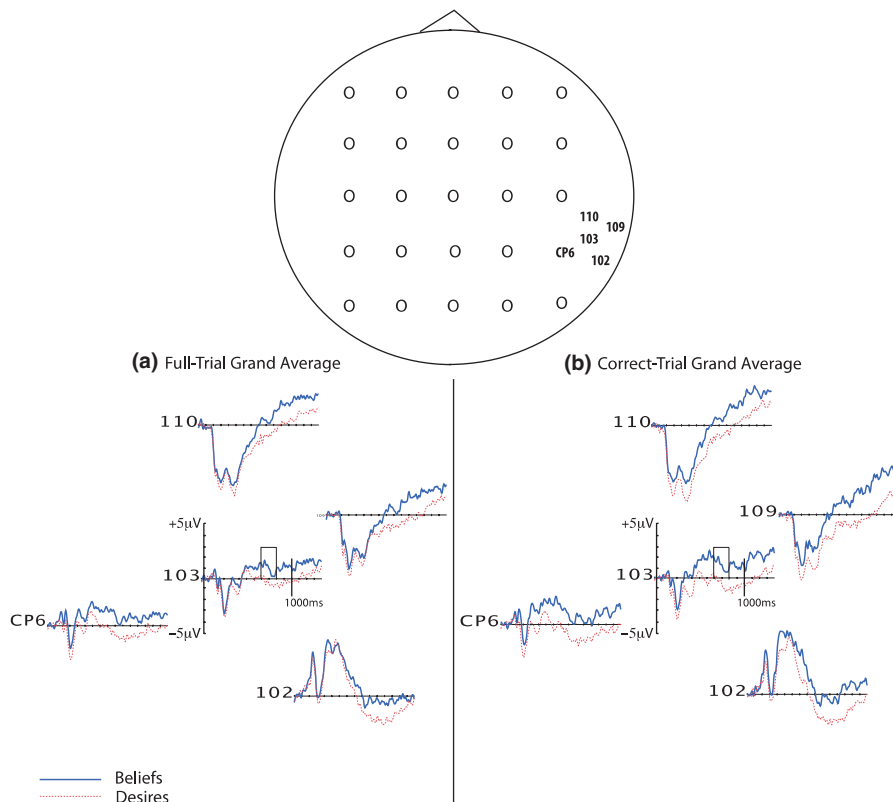
#### Right-posterior ROI analysis

As outlined earlier, there are *a priori* reasons for closely examining Beliefs versus Desires differences in right-posterior regions. Moreover, as can be seen in Figure 3 and still more clearly in Figure 5a, visual inspection of the grand average waveforms for the right-posterior channel CP6 in our data suggests greater mean amplitude for Beliefs compared to Desires. Thus, we selected channel CP6 along with four additional neighbouring channels

showing a similar effect (EGI channels 102, 103, 109, and 110) to serve as a right-posterior five-channel cluster (see Figure 5). Repeated measures ANOVAs comparing mean amplitude in Beliefs versus Desires conditions across all five epochs for this right-posterior cluster yielded no significant differences in any epoch. Thus, for the full-trial analysis, belief- and desire-reasoning did not show any distinction in any scalp regions.

#### Correct-trial analyses

The full-trial analyses have the advantage of including more ERP trials (94 trials per child across all three conditions, 65% of the 144 maximally presented) than correct-trial analyses (80 trials per child, 56% of the total). However, as noted earlier, children were accurate



**Figure 5** a and b Grand average event-related brain potential waveforms for the Desires (red dotted lines), and Beliefs (blue solid lines) conditions for the right-posterior five-channel cluster for full-trial (A. left) and correct-trial (B. right) analyses. The boxed section indicates the 600–800 ms epoch in which the correct-trial analyses first demonstrated significantly greater amplitude for Beliefs over Desires. Waveforms are displayed with positive amplitudes above the axis and negative amplitudes below. Scalp positions for the five-channel cluster relative to the 5 × 5 grid are shown at the top.

on only 65% of their belief trials. It is not clear that the incorrect trials access belief-reasoning. Moreover, a common concern in comparing neurophysiological and neuroimaging data across ages is differences in performance, where observed activation differences could be due to differences in either age or performance (Gaillard, Grandin & Xu, 2001). Given that in the parallel ERP study by Liu *et al.* (2009a) the adult participants performed with near-perfect accuracy across all conditions (i.e. 96.0%, 96.4%, and 96.5% correct for beliefs, desires and physical trials, respectively), a focus on correct trials allows a more direct comparison to that adult data by eliminating accuracy differences between the two age groups. Thus, following an often-recommended strategy (e.g. Casey, Giedd & Thomas, 2000), we analyzed ERP responses to the correct trials alone for Beliefs, Desires, and Physical conditions. We adopted the same approach as for the full-trial analyses, within the same five epochs.

#### Right-frontal effects

Visual inspection of the correct-trial grand average ERP waveforms from the 5 × 5 grid in all three conditions revealed a pattern nearly identical to the results of the full-trial analysis: From 200 ms through 850 ms post-stimulus, amplitude for the mental-state conditions (Beliefs and Desires) was clearly more negative compared to the Physical condition within right mid-frontal scalp regions (see Figure 4b). As with full-trial analyses, three 2 (condition) × 5 (laterality) × 5 (caudality) repeated measures ANOVAs on the mean amplitude of correct trials in each of the five epochs yielded significant condition and three-way interaction effects in the 200–250, 350–600, 600–800, and 800–850 ms epochs (as well as significant interaction effects in the 850–1400 ms epoch) for Belief–Physical, and Desire–Physical comparisons; whereas Belief–Desire comparisons showed no significant condition or interaction effects for any epoch (see Table 3).

The same ROI analyses for the same right mid-frontal channel cluster used in the full-trial analysis (see Fig-

ure 4) yielded results nearly identical to those of the full-trial analysis. Beliefs–Physical and Desire–Physical condition differences occurred in the 200–250, 350–600, 600–800, and 800–850 epochs at each of the five individual channels ( $ps < .008$ ), whereas Beliefs and Desires conditions did not differ significantly from each other in overall ANOVAs or at any channel, in any epoch ( $ps > .25$ ).

#### Right-posterior ROI analysis

As with the full-trial analyses, we conducted a further examination of differences between Beliefs and Desires conditions within right-posterior scalp regions. We used the same right-posterior cluster of five channels as in the full-trial analysis (see Figure 5), and conducted 2 (condition) × 5 (channel) repeated measures ANOVAs, as well as follow-up paired samples *t*-tests at each individual channel just as before. As seen in Figure 5b, the correct-trial grand average waveforms for the right-posterior cluster reveal greater mean amplitude for Beliefs compared to Desires conditions. This difference appears around 350 ms post-stimulus onset, and remains beyond 800 ms post-stimulus. Notably, the difference between Beliefs and Desires appears more pronounced in the correct-trial waveforms compared to the full-trial waveforms (Figure 5a and b). Indeed, ANOVAs comparing mean amplitude in Beliefs versus Desires conditions for correct trials alone yielded a significant difference within the 600–800 and 850–1400 ms epochs (see top panel of Table 4), and *t*-tests revealed that Beliefs and Desires conditions differed significantly at all five channels within the cluster in each of these epochs (see bottom panel of Table 4). Parallel ANOVAs indicated near significant differences in the 350–600, and 800–850 ms epochs (see Table 4) – and in those epochs many but not all channels reached significance at the .05 level. In contrast, repeated measures ANOVAs for an equivalent cluster of five left-posterior channels showed no significant difference between Beliefs and Desires conditions in any epoch. These results confirm the presence of a Beliefs versus Desires distinction (beginning most strongly in the

**Table 3** Correct-trial subset analysis results for the 2(condition) × 5(laterality) × 5(caudality) repeated measures ANOVAs for all 5 epochs

Comparison	Epoch (ms post stimulus onset)				
	200–250	350–600	600–800	800–850	850–1400
<b>Condition effects</b>					
Belief vs. Desire	$F(1,15) = 5.91$	$F(1,15) = .36$	$F(1,15) = .20$	$F(1,15) = .34$	$F(1,15) = .66$
Desire vs. Physical	$F(1,15) = 15.63^{**}$	$F(1,15) = 35.34^{***}$	$F(1,15) = 32.68^{***}$	$F(1,15) = 22.92^{***}$	$F(1,15) = 8.86^{**}$
Belief vs. Physical	$F(1,15) = 30.53^{***}$	$F(1,15) = 43.09^{***}$	$F(1,15) = 20.95^{***}$	$F(1,15) = 13.31^{**}$	$F(1,15) = 3.89$
<b>Condition × laterality × caudality interaction effects</b>					
Belief vs. Desire	$F(16,240) = 0.74$	$F(16,240) = 2.11$	$F(16,240) = 1.13$	$F(16,240) = 2.03$	$F(16,240) = 1.49$
Desire vs. Physical	$F(16,240) = 8.96^{***}$	$F(16,240) = 9.62^{***}$	$F(16,240) = 7.63^{***}$	$F(16,240) = 11.48^{***}$	$F(16,240) = 8.80^{***}$
Belief vs. Physical	$F(16,240) = 5.32^{**}$	$F(16,240) = 5.28^{**}$	$F(16,240) = 4.41^{**}$	$F(16,240) = 4.74^{**}$	$F(16,240) = 5.30^{**}$

Notes. \*\* indicates  $p < .01$ , \*\*\*indicates  $p < .001$

**Table 4** Correct-trial ROI analysis results of ANOVAs and *t*-tests for all 5 epochs: right posterior 5-channel cluster

Channel	Comparison	Epoch (ms post stimulus onset)				
		200–250	350–600	600–800	800–850	850–1400
2(condition) × 5(channel) repeated measures ANOVA condition effects						
All 5	Belief vs. Desire	$F(1,15) = .85$	$F(1,15) = 6.92^*$	$F(1,15) = 11.44^{**}$	$F(1,15) = 7.36^*$	$F(1,15) = 13.44^{**}$
Paired samples <i>t</i> -test individual channel condition effects						
CP6	Belief vs. Desire	$t(15) = .81$	$t(15) = 2.37^*$	$t(15) = 2.97^{**}$	$t(15) = 2.72^*$	$t(15) = 2.50^*$
102	Belief vs. Desire	$t(15) = .45$	$t(15) = 1.63$	$t(15) = 2.15^*$	$t(15) = 1.86$	$t(15) = 3.24^{**}$
103	Belief vs. Desire	$t(15) = .67$	$t(15) = 2.30^*$	$t(15) = 2.67^*$	$t(15) = 2.39^*$	$t(15) = 2.84^*$
109	Belief vs. Desire	$t(15) = 1.60$	$t(15) = 2.55^*$	$t(15) = 3.31^{**}$	$t(15) = 2.48^*$	$t(15) = 3.65^{**}$
110	Belief vs. Desire	$t(15) = 1.24$	$t(15) = 2.54^*$	$t(15) = 2.71^*$	$t(15) = 1.79$	$t(15) = 2.34^*$

Notes. \* indicates  $p < .05$ , \*\* indicates  $p \leq .01$

600–800 ms epoch) specific to right-posterior scalp regions.

## Discussion

The present study investigated the neural correlates of belief- and desire-reasoning in 7- and 8-year-old children. We found evidence for two distinct neural patterns associated with reasoning about mental states. First, neural activation for both mental-state conditions (Beliefs and Desires) differentiated equally from the physical control, but did not differentiate from each other. These effects appear as early as 200 ms and are sustained up to 850 ms post-stimulus onset, and are observed within right mid-frontal scalp regions. Second, the neural activation associated with belief-reasoning differentiated from that for desire-reasoning. Intriguingly, this differentiation appeared for analyses of only trials in which children performed correctly. These effects first reached peak significance within the 600–800 ms post-stimulus epoch, and were concentrated in right-posterior scalp regions.

The current study was designed to examine belief- and desire-reasoning in children, and employed the same stimuli and methodology used in a recent ERP investigation of adults (Liu *et al.*, 2009a) to provide a developmental comparison. Thus, the key results from the present study are discussed separately but also in conjunction with the adult findings from Liu *et al.*

As background for the adult–child comparisons, our ERP task was designed such that all three conditions (Beliefs, Desires, and Physical) had the same perceptual and linguistic structure including similar two-part comparisons. Following standard neuroimaging subtraction methodology then, any differences in neural activation across conditions can be attributed to differences in reasoning about the *content* of each condition (i.e. belief-reasoning, desire-reasoning, or reasoning about physical locations) beyond the memory and processing demands common in all conditions. Similarly, comparison of these condition differences across age are likely to reflect development in reasoning specific to mental states rather than information processing capacities alone.

## Mental-state versus control distinction in right mid-frontal scalp regions

As expected, we found that children exhibited a pattern of neural processing for reasoning about diverse-beliefs and diverse-desires that was separate from processing the physical control condition. This result parallels the adult ERP finding (Liu *et al.*, 2009a) in which adults also showed overlapping activation for belief- and desire-reasoning, with both mental-state conditions differing equally from activation for the physical control. For adults, this effect was concentrated in mid-frontal scalp regions, and occurred in the later epoch of 800–850 ms post-stimulus onset. Our child data showed this same pattern of neural activation within roughly the same scalp regions (slightly more right lateralized compared to adults, but within overlapping regions given the spatial resolution of EEG data; Ferree, Clay & Tucker, 2001), and the pattern also occurred in the 800–850 ms epoch (though the effect was not specific to this epoch in children). These findings suggest that by age 7 years, some adult-like neural processes for mental-state reasoning have already developed, and they are in line with existing research demonstrating that mental-state reasoning recruits specific neural substrates in medial-frontal regions of the brain in both adults and children (see Saxe *et al.*, 2004, for review). These substrates have been shown to be associated with mental-state reasoning even after controlling for domain-general skills (i.e. executive functioning; Sabbagh *et al.*, 2009, for 4-year-olds), and they have been shown to be distinct from the substrates supporting non-mental processing (i.e. physical reasoning; Saxe *et al.*, 2009, for 6-year-olds). Thus, our results follow the common pattern of specific neural substrates supporting mental-state reasoning, and add increasing evidence that such specificities start to develop even at young ages.

Two intriguing differences exist between our child ERP data and the comparison data collected from adults. In adults, the distinction between mental-state conditions and the physical condition occurred as an LSW within the 800–850 ms post-stimulus epoch, and the two mental-state conditions were more positive compared to the control condition. In children, there was a distinction

between mental-state and physical conditions within this epoch, but this distinction appears earlier, starting at the P200 peak (a standard landmark in ERP data characterized by a positive peak around 200 ms post-stimulus onset), and mental-state conditions were more negative compared to the control condition. The adult data also showed a P200 peak, but with no differentiation between any of the conditions evident at this point, only later.

An interesting question thus concerns why children show an effect for mental-state reasoning that begins earlier in the neural time-course, and is negative compared to the control condition. We discuss some possible explanations, but further research is needed to best understand these differences. One possibility is that these differences between adults and children reflect differences in the cognitive processes underlying children's and adults' desire- and belief-reasoning. Generally, the ERP literature suggests that the P200 is associated with automatic feature detection processes, encoding, and classification (see e.g. Van der Cruyssen *et al.*, 2009). Importantly, the P200 has been shown to be associated with detection and classification of mental versus non-mental features. For example, research suggests that the P200 is responsive to detecting others' goals and intentions (Van der Cruyssen *et al.*, 2009). Both our child ERP data and the comparison adult data show a P200 peak in all three conditions. Thus, the P200 occurring in both adults and children represents some sort of automatic classification of the presented stimuli. However, for children, beyond any such general stimulus classification effect, there is a distinction in this initial automatic classification that emphasizes mental versus non-mental content. One hypothesis is that the kind of mental-state understanding achieved by children in the preschool years gives mental states a special emphasis even very early in processing. Adults, however, may no longer emphasize the (now commonplace) broader distinction of mental versus non-mental content, but emphasize more importantly which *type* of mental content is present (i.e. beliefs versus desires) – an emphasis that recruits neural processes operating only later in the ERP time-course.

Relatedly, differences in neural activation between adults and children could also reflect differences in the underlying cognitive process of 'mental-state decoding'. Some researchers argue that theory-of-mind reasoning involves at least two component processes: mental-state decoding (on-line attribution of mental states to individuals), and mental-state reasoning (using representations of individuals' mental states in order to predict and make judgments about their actions) (e.g. Tager-Flusberg and Sullivan, 2000; Pineda & Hecht, 2009; Sabbagh, 2004). Event-related potential research on the neural correlates of the decoding component has shown that neural activation associated with mental-state decoding appears early in the ERP time-course (i.e. prior to 300 ms post-stimulus), and moreover, it is more negative compared to activation for non-mental decoding of

control conditions (Sabbagh *et al.*, 2004; Sabbagh, 2004). Indeed, the neural activation for both mental-state conditions was more negative compared to the control prior to 300 ms post-stimulus for children (our data), but *not* for adults (Liu *et al.*, 2009a, data). Thus, this differential could reflect a difference in how children and adults decoded the mental states presented in the Beliefs and Desires conditions.

Such an interpretation makes conceptual sense given the nature of our task and data analyses. That is, the information phases of our task (e.g. 'the boy really likes grapes, but the girl really likes celery') require mental-state decoding – attributing a specific desire and belief to each character. The target questions (e.g. 'who says: "I won't have any" when they see this?') require mental-state reasoning – using beliefs/desires to judge what a character would say. It was the neural responses to these target questions (and not the information phase) that were analyzed – and indeed, adults (who likely completed the decoding process in the information phase) did not show the early increased negative activation for mental-state conditions versus control that is characteristic of mental-state decoding. However, children are arguably less efficient/effective at mental-state processing, and may still be conducting the decoding process (or even just initiating it) at the time of the target question. Thus for children, the ERP waveforms may well be capturing both decoding (indexed by an early increased negative activation for mental-state conditions compared to the control) and reasoning (indexed by ERP differences similar to those for adults later in the waveform).

Findings of this sort highlight how neuroscientific data can reveal different underlying processes even when behavioural performances appear similar on the surface (in this case for children and adults). Future research investigating the neural circuitry of belief- and desire-reasoning in older children should look for an attenuation of the distinction between mental-state and control conditions early in the ERP time-course.

#### *Beliefs versus desires distinction in right-posterior scalp regions*

In addition to overlap in belief and desire neural activation in mid-frontal scalp regions, there was also evidence for a pattern of neural processing specific to belief-reasoning – distinct from desire-reasoning – concentrated in right-posterior scalp regions. These results are also in line with results from the parallel adult ERP study (Liu *et al.*, 2009a). The adult data showed a distinction between belief and desire neural activation, reflected in right-posterior scalp regions, within the 600–800 ms epoch. Given that the adults performed essentially perfectly, this result represents almost exclusively correct-trial data. The correct-trial data for our child participants also showed the same beliefs–desires distinction, in roughly the same scalp region, and first reached peak significance in the 600–800 ms epoch. However, for

children (who unlike adults had numerous incorrect belief trials) this pattern was not apparent for full-trial data. Nonetheless, the effect seems reliable. Given that overall performance accuracy for the Beliefs condition was 64% correct for children (compared to 96% correct for adults), in addition to capturing truly accurate reasoning, our child data may include some 'correct guesses' as well. Thus, if anything our results may slightly underestimate the presence of a beliefs–desires distinction in children. Adult and child data taken together thus demonstrate that by at least 7 years of age, children show adult-like neural processing for mental-state reasoning including a right-posterior beliefs–desires distinction, but this distinction is only evident in children when considering children's accurate performance trials.

Other research on the neurological correlates of mental-state reasoning also demonstrates that children show increasingly adult-like neural processing as their accuracy for mental-state reasoning improves. Liu *et al.* (2009b) found that children who correctly performed false-belief tasks showed a pattern of neural activity similar to adults (although with more diffuse regions of activation). However, children who did not show correct performance did not show adult-like neural activity for mental-state understanding. Sabbagh *et al.* (2009) found that 4-year-olds who exhibited more accurate performance on a battery of theory-of-mind tasks also exhibited increased functional maturation of their dorsal medial pre-frontal cortex (dMPFC) and right TPJ (two regions robustly recruited for mental-state reasoning in adults – see Saxe *et al.*, 2004, for review). Extending these initial findings, our ERP data also show that children's neural processing is *most* adult-like (including both a frontal mental-state versus control neural distinction, as well as a posterior beliefs versus desires neural distinction) *only* when considering solely accurate performance.

Two informative yet distinct interpretations could account for this signature result. One interpretation is that neural mechanisms for both belief- and desire-reasoning are functionally and anatomically mature in children, but that elements of the way in which children reason about mental states mask the presence of this already fully developed system. That is, children may have a mature neural system for mental-state reasoning, but they fail to reliably recruit this neural system to reason about mental states. This failure to activate mature neural mechanisms may be due to a still-developing understanding of mental states, or due to domain-general difficulties in selecting the appropriate mechanisms. Similarly, it is possible that a fully developed theory-of-mind system is present in children; however, children do not consistently employ an adult-like *strategy* for reasoning about mental states (e.g. they concentrate more heavily on mental-state decoding versus mental-state reasoning), and that this different strategy is supported by different neural mechanisms.

An alternative interpretation also consistent with these results is that our child data, in conjunction with adult

data from Liu *et al.* (2009a), suggest an age-related trend in specialization of right-posterior regions for belief-state reasoning – a specialization that is linked to increasing accuracy on the behavioural measures. That is, it is possible that at an early point in development, when accuracy for belief-reasoning is low compared to desire-reasoning, neural substrates are most robustly recruited for understanding mental states more generally. Neural specialization for reasoning specifically about more complex, representational mental states such as beliefs is still emerging, occurring infrequently, only when belief-reasoning is highly accurate. Thus, data from adults (Liu *et al.*, 2009a), in comparison to child data, show that later in development, once a better, more distinctive understanding of representational mental states (i.e. beliefs) is more firmly and consistently present, some neural substrates have more robustly specialized to support understanding of these complex mental states over and above the processing of other mental states (i.e. desires).

Recent neuroscientific research with adults and children as well as developmental theory both provide support for the idea that a posterior system specialized in belief-reasoning is still developing in childhood – making the second interpretation particularly intriguing. Saxe *et al.* (2009) found activation in MPFC and TPJ as children (6- to 11-year-olds) listened to stories about mental states and social interactions (in comparison to stories about physical events). However, specifically the TPJ (and not the MPFC) increased in selectivity for mental-state reasoning (in comparison to both reasoning about physical events *and* social interactions) as children aged. Sommer *et al.* (2010) found that both adults and children 10–12 years recruited dorsal MPFC for reasoning about false beliefs, but only adults – not children – also selectively recruited the right TPJ for false-belief reasoning. In the present study, the direct comparison of belief and desire neural activation adds a crucial element to the characteristics of this possible neural specialization. That is, our data (in conjunction with adult data from Liu *et al.*, 2009a) suggest that with development, right-posterior regions show specialization beyond general mental-state reasoning, and are recruited for belief-reasoning over and above reasoning for other types of mental states such as desires. Interestingly, this neural specialization for belief-reasoning continues to develop after preschool age, after children demonstrate initial abilities to reason about beliefs. Thus our findings, and those of other emerging neuroscientific research, could be indicating that the neurocognitive systems supporting young children's initial abilities to reason about beliefs might not be entirely the same as the neurocognitive mechanisms that later become specialized in adults.

We favour this interpretation because it represents a straightforward developmental possibility that an understanding of beliefs may build on prior desire-understanding, recruiting additional substrates in right-posterior regions as belief-understanding becomes more distinct and accurate. Such a developmental scenario is

supported by a wealth of behavioural evidence that children reach an explicit understanding of desires before they come to an explicit understanding of beliefs (e.g. Wellman, 2002; Wellman & Liu, 2004). Indeed, our current data show that even by middle childhood, children still have greater accuracy on diverse-desires tasks compared to diverse-beliefs tasks (although children in our study are older than ages when children, on average, pass diverse-desires and diverse-beliefs tasks, our tasks presented information quickly and thus were more difficult overall). In short, it is possible that as children increase their accuracy for belief-reasoning to adult-like levels, the underlying neural substrates for reasoning about beliefs show an accompanying specialization to become adult-like as well. Future research, especially longitudinal research, is needed to directly address this hypothesis, and to determine whether neural specialization might precede increased accuracy or vice versa.

It is important to note that, due to the low-spatial resolution of ERP data, locations of scalp distribution effects do not pinpoint the location of underlying neural activity. Thus, with our current data, it is not possible to determine precisely how the location of neural circuitry for mental-state reasoning in children compares to circuitry in adults. Future research using neuroimaging techniques with higher spatial resolution such as fMRI, functional near-infrared spectroscopy (fNIRS), and magnetoencephalography (MEG), as well as EEG/ERP source localization analyses are needed to properly localize the neural substrates supporting belief- and desire-reasoning in children and adults.

## Conclusions

Results of the present study shed light on development of neural circuitry supporting belief- and desire-reasoning in 7- and 8-year-old children. Previous research has begun to uncover neural mechanisms for understanding of beliefs and false beliefs (e.g. Liu *et al.*, 2009b; Sabbagh *et al.*, 2009). However, given behavioural research demonstrating that an understanding of desires develops before understanding beliefs, investigations of mechanisms that underlie both belief- and desire-reasoning – and their interrelation – are necessary to uncover the process by which an explicit theory of mind develops. To this end, our results demonstrate that by ages 7 and 8 years, children have already developed neural specializations for reasoning about beliefs and desires that are distinct from the neural activations for physical reasoning – neural patterns that are similar to those found in adults. However, in contrast to adults, children at this age do not yet robustly recruit right-posterior regions for reasoning about beliefs over and above reasoning about mental-states more generally; in particular, children showed selective activation for belief versus desire reasoning *only* when analyses were restricted to correct trials. Two possible lines of reasoning could account for this pattern of results. A fully developed

neural system for mental-state reasoning (including specialization for beliefs) could exist in both adults and children, masked by children's less accurate reasoning and/or their different approaches to mental-state understanding. Alternatively, neuromechanisms for mental-state reasoning, and in particular belief-reasoning, may be still developing in childhood, pointing to a connection between the development of an increasingly accurate understanding of complex mental states, and the development of specialized neural substrates to support this understanding. Future neuroscientific research promises to provide further insight into how an understanding of different mental states builds to form the adult's explicit theory of mind.

## Acknowledgements

This research was supported by an NSERC PGS-D grant awarded to Bowman, an NICHD (HD066029) grant awarded to Liu, and an NICHD (HD022149) grant awarded to Wellman. We gratefully acknowledge the contribution of the parents and children who participated in this study.

## References

- Abraham, A., Rakoczy, H., Werning, M., von Cramon, D.Y., & Schubotz, R.I. (2010). Matching mind to world and vice versa: functional dissociations between belief and desire mental state processing. *Social Neuroscience*, *5*, 1–18.
- Bartsch, K., & Wellman, H.M. (1995). *Children talk about the mind*. New York: Oxford University Press.
- Casey, B.J., Giedd, J.N., & Thomas, K.M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology. Special Issue: Behavior, Event-Related Potentials, and Neuroimaging: A Lifespan Perspective*, *54*, 241–257.
- Ferree, T.C., Clay, M.T., & Tucker, D.M. (2001). The spatial resolution of scalp EEG. *Neurocomputing*, *38*, 1209–1216.
- Friedman, O., & Leslie, A.M. (2004a). A developmental shift in processes underlying successful belief-desire reasoning. *Cognitive Science*, *28*, 963–977.
- Friedman, O., & Leslie, A.M. (2004b). Mechanisms of belief-desire reasoning: inhibition and bias. *Psychological Science*, *15*, 547–552.
- Gaillard, W.D., Grandin, C.B., & Xu, B. (2001). Developmental aspects of pediatric fMRI: considerations for image acquisition, analysis, and interpretation. *NeuroImage*, *13*, 239–249.
- Gallagher, H.L., & Frith, C.D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, *7*, 77–83.
- Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development*, *62*, 98–110.
- Gratton, G., Coles, M.G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology*, *55*, 468–484.

- Harris, P.L. (2006) Social cognition. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology* (5th edn, pp. 811–858). New York: Wiley.
- Karmiloff-Smith, A. (1997). Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology*, **13**, 513–524.
- Kobayashi, C., Glover, G.H., & Temple, E. (2007). Children's and adults' neural bases of verbal and nonverbal 'theory of mind'. *Neuropsychologia*, **45**, 1522–1532.
- Liu, D., Meltzoff, A.N., & Wellman, H.M. (2009a). Neural correlates of belief- and desire-reasoning. *Child Development*, **80**, 1163–1171.
- Liu, D., Sabbagh, M.A., Gehring, W.J., & Wellman, H.M. (2009b). Neural correlates of children's theory of mind development. *Child Development*, **80**, 318–326.
- Malle, B.F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: MIT Press.
- Onishi, K.H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, **308**, 255–255.
- Pineda, J.A., & Hecht, E. (2009). Mirroring and mu rhythm involvement in social cognition: are there dissociable subcomponents of theory of mind? *Biological Psychology*, **80**, 306–314.
- Sabbagh, M.A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: implications for autism. *Brain and Cognition*, **55**, 209–219.
- Sabbagh, M.A., Bowman, L.C., Evraire, L.E., & Ito, J.M.B. (2009). Neurodevelopmental correlates of theory of mind in preschool children. *Child Development*, **80**, 1147–1162.
- Sabbagh, M.A., Moulson, M.C., & Harkness, K.L. (2004). Neural correlates of mental state decoding in human adults: an event-related potential study. *Journal of Cognitive Neuroscience*, **16**, 415–426.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, **55**, 87–124.
- Saxe, R., & Powell, L.J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science*, **17**, 692–699.
- Saxe, R., Schulz, L.E., & Jiang, Y.V. (2006). Reading minds versus following rules: dissociating theory of mind and executive control in the brain. *Social Neuroscience. Special Issue: Theory of Mind*, **1**, 284–298.
- Saxe, R., Whitfield-Gabrieli, S., Scholz, J., & Pelphrey, K.A. (2009). Brain regions for perceiving and reasoning about other people in school-aged children. *Child Development*, **80**, 1197–1209.
- Sommer, M., Meinhardt, J., Eichenmüller, K., Sodian, B., Döhl, K., & Hajak, G. (2010). Modulation of the cortical false belief network during development. *Brain Research*, **1354**, 123–131.
- Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: evidence from William's syndrome. *Cognition*, **76**, 59–90.
- Tucker, D.M. (1993). Spatial sampling of head electrical fields: the geodesic sensor net. *Electroencephalography and Clinical Neurophysiology*, **87**, 154–163.
- Van der Cruyssen, L., Van Duynslaeger, M., Cortoos, A., & Van Overwalle, F. (2009). ERP time course and brain areas of spontaneous and intentional goal inferences. *Social Neuroscience*, **4**, 165–184.
- Wellman, H.M. (2002). Understanding the psychological world: developing a theory of mind. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development*. (pp. 167–187). Malden, MA: Blackwell Publishing.
- Wellman, H.M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, **72**, 655–684.
- Wellman, H.M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, **75**, 523–541.
- Wellman, H.M., & Woolley, J.D. (1990). From simple desires to ordinary beliefs: the early development of everyday psychology. *Cognition*, **35**, 245–275.

Received: 6 April 2011

Accepted: 22 February 2012