Neurocognitive development of relational reasoning

Eveline A. Crone[1,2,*], **Carter Wendelken**[2,*], **Linda van Leijenhorst**[1,2], **Ryan D. Honomichl**[2], **Kalina Christoff**[4], and **Silvia A. Bunge**[2,5]

[1]*Department of Developmental Psychology, Leiden University, the Netherlands* [2]*Center for Mind and Brain, University of California in Davis, USA* [3]*Department of Psychology, Hiram College, USA* [4]*Department of Psychology, University of British Columbia, Canada* [5]*Department of Psychology & the Helen Wills Neuroscience Institute, University of California at Berkeley*

Abstract

Relational reasoning is an essential component of fluid intelligence, and is known to have a protracted developmental trajectory. To date, little is known about the neural changes that underlie improvements in reasoning ability over development. In this event-related functional magnetic resonance imaging (fMRI) study, children aged 8-12 and adults aged 18-25 performed a relational reasoning task adapted from Raven's Progressive Matrices. The task included three levels of relational reasoning demands: REL-0, REL-1, and REL-2. Children exhibited disproportionately lower accuracy than adults on trials that required integration of two relations (REL-2). Like adults, children engaged lateral prefrontal cortex (PFC) and parietal cortex during task performance; however, they exhibited different timecourses and activation profiles, providing insight into their approach to the problems. As in prior studies, adults exhibited increased rostrolateral PFC (RLPFC) activation when relational integration was required (REL-2 > REL-1, REL-0). Children also engaged RLPFC most strongly for REL-2 problems at early stages of processing, but this differential activation relative to REL-1 trials was not sustained throughout the trial. These results suggest that the children recruited RLPFC while processing relations, but failed to use it to integrate across two relations. Relational integration is critical for solving a variety of problems, and for appreciating analogies; the current findings suggest that developmental improvements in this function rely on changes in the profile of engagement of RLPFC, as well as dorsolateral PFC and parietal cortex.

Introduction

Relational reasoning, or the ability to consider relationships between multiple mental representations, is directly linked to the capacity to think logically and solve problems in novel situations (Cattell, 1971; Halford, Wilson, & Phillips, 1998). Relational reasoning is an important component of fluid intelligence (Duncan, 2003). This capacity is typically considered to be useful for solving almost any problem, largely insensitive to cultural influences, rising and falling at its own rate across the lifespan, and affected in no specific behavioral domain by brain injury (Cattell, 1987). Relational reasoning is thought to be instrumental in the learning of tasks requiring complex spatial, numerical, or conceptual relations (Cattell, 1987). This capacity emerges in the first two or three years of life (Cattell, 1971, 1987), after the development of general perceptual, attentional, and motoric capabilities (Horn & Noll, 1997). Children begin to integrate multiple relations at about 5 years of age

Address correspondence to: Carter Wendelken, Ph.D., Helen Wills Neuroscience Institute, University of California at Berkeley, 210 Barker Hall, Berkeley, CA, 94720, cwendelken@berkeley.edu.
*Shared first-authorship

(Halford et al., 1998), but improvements in relational reasoning are observed throughout childhood (Richland, Morrison, & Holyoak, 2006; Sternberg & Rifkin, 1979; Vodegel Matzen, 1994).

Relational reasoning can be measured with the Raven Progressive Matrices (RPM) task (see Figure 1 for an adaptation of the RPM task). This is a visuospatial task that requires participants to identify relevant stimulus features based on the spatial organization of an array of stimuli, and then select the choice stimulus that matches one or more of these identified features (Raven, 1941). The relational complexity of a problem can be defined by the number of related dimensions that need to be considered jointly to arrive at the correct solution (Christoff et al., 2001). On 0-relational (REL-0) problems, no dimensional variation need be considered, and participants are simply asked to complete the array with the matching figure. On 1-relational (REL-1) problems, a single dimension of variation must be considered, and participants must identify a vertical or horizontal relationship between items in the array in order to complete the task. REL-1 problems thus require 1st-order relational processing: the consideration of relationships between items. On 2-relational (REL-2) problems, participants must consider two dimensions of variation: horizontal and vertical. The solution to a REL-2 problem involves the integration of information from both dimensions. Thus, performance of a REL-2 problem requires 2nd-order relational processing, or *relational integration*: the joint consideration of multiple relations. The goal of the current study was to use functional MRI to test whether developmental differences in relational reasoning are specifically associated with an increased capability to perform relational integration.

Neuropsychological and neuroimaging studies have shown that prefrontal cortex (PFC) is strongly implicated in relational reasoning. Studies of patients with early stages of the frontal variant of frontotemporal dementia have revealed deficits in relational reasoning on RPM and other tasks (Morrison et al., 2004; Waltz et al., 1999). Further, fMRI studies have implicated both dorsolateral PFC (DLPFC; lateral BA 9, 46) and rostrolateral PFC (RLPFC; lateral BA 10) in performance of the RPM task (Christoff et al., 2001).

RLPFC has been consistently implicated in fMRI studies of relational integration, including the RPM task (Christoff et al., 2001; Kroger et al., 2002; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997) and verbal propositional analogy tasks (Bunge, Wendelken, Badre, & Wagner, 2005). In the RPM task, RLPFC has been found to be engaged more strongly on REL-2 problems - in which participants must jointly process two dimensions of change in the visual arrays - than on either REL-0 or REL-1 problems (Christoff et al., 2001; Kroger et al., 2002). In the verbal propositional analogy task (Bunge et al., 2005; Wendelken et al., in press) RLPFC is engaged when participants must consider whether two semantic relations are analogous (e.g. "shoe is to sock as glove is to hand?"). Thus, these studies of reasoning in the visuospatial and verbal domains show that RLPFC is modulated by the need to jointly consider, or integrate, multiple relations.

Based on these and other data, several research groups have argued that relational integration is the basic task requirement that drives RLPFC (Bunge et al., 2005; Christoff & Gabrieli, 2002; Ramnani & Owen, 2004). Alternative theoretical accounts have been proposed to account for the involvement of RLPFC across a variety of cognitive tasks (Braver & Bongiolatti, 2002; Burgess, Gilbert, & Dumontheil, 2007; Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999; Ramnani & Owen, 2004). It is beyond the scope of this manuscript to discuss these distinct, but interrelated, accounts (for our view, see (Wendelken et al., in press) see also (Ramnani & Owen, 2004).

In contrast to RLPFC, DLPFC has shown sensitivity to other factors that affect task difficulty on reasoning tasks, rather than relational complexity *per se*. For example, DLPFC showed

sensitivity to the level of interference from competing response alternatives in an RPM-type task (Kroger et al., 2002), and to the need to override a response bias on a verbal propositional analogy task (Bunge et al., 2005). Thus, it has been argued that RLPFC plays a unique role in the representation of relational structures, whereas DLPFC is sensitive to task difficulty in a manner that is not specifically related to complexity level (Christoff et al., 2001; Wagner, Koch, Reichenbach, Sauer, & Schlosser, 2006). DLPFC may support reasoning by organizing representations in working memory, selecting between competing response alternatives, and monitoring performance (Christoff et al., 2001). Developmental changes in relational reasoning could be associated with maturation of either of these processes. In this study we tested whether children show immature activation in RLPFC, DLPFC or both, when performing relational reasoning problems based on the RPM (Christoff et al., 2001).

PFC interacts closely with parietal cortex (Fuster, 2003; Petrides & Pandya, 1984), and this region is consistently engaged in the RPM task (Gray, Chabris, & Braver, 2003). The involvement of the superior parietal lobule (SPL) is likely to be related, at least in part, to the visuospatial demands of the task (Gross & Graziano, 1995). However, the role of parietal cortex in reasoning may extend beyond the representation of visuospatial features. Gray and colleagues showed that the level of activation in lateral PFC and inferior parietal lobule (IPL) - specifically, BA 40 - together accounted for the majority of the variance in the relationship between fluid reasoning ability and performance on a challenging working memory task (Gray et al., 2003; Lee et al., 2006). We have previously argued that, through its representation of spatial information, parietal cortex supports the organization of information and the maintenance of information in an organized state (Wendelken et al., 2007). Such a function could be involved in the ability to integrate relationships between items on the RPM task.

Additionally, Ramachandran and colleagues have argued that the left IPL - or, more specifically, the angular gyrus (BA 39) - is necessary for abstract thought. Their unpublished preliminary data from four stroke patients suggest, intriguingly, that the angular gyrus may be necessary for correctly interpreting metaphors (e.g. "All that glitters is not gold"), and for making connections between shapes and sounds that are linked by an abstract concept (e.g. readily associating an image of speckled dots and the sound 'shhhhh', because they are linked by the concept of fuzziness) (Ramachandran, presentation at American Psychological Society conference in April, 2005). These findings motivated us to examine the roles of IPL subregions in the RPM task, in addition to RLPFC and DLPFC. A prior study examining developmental changes in working memory indicated that besides PFC, children also under-recruited superior and inferior parietal cortex (Crone, Wendelken, Donohue, van Leijenhorst, & Bunge, 2006). Based on these findings, we sought to test whether changes in the function of BA 39 and/or 40 are associated with age-related improvements in reasoning ability from childhood to adulthood.

The differential brain development hypothesis builds on research examining changes in brain structure across development within individuals. These studies have shown that cortical white matter increases approximately linearly with age throughout childhood and adolescence, and differs little across regions (Giedd, 2004; Gogtay et al., 2004; Sowell et al., 2004). In contrast, much cortical gray matter follows an inverted-U shape over development, peaking at different ages depending on the region. DLPFC, RLPFC and parietal cortex show a relatively protracted developmental time course, in that cortical gray matter loss continues until the early 20s (Giedd, 2004). A cross-sectional structural MRI study focusing specifically on RLPFC and DLPFC indicated that cortical thickness decreases between the ages of 8 and 20 at a similar rate in both of these regions (O'Donnell, Noseworthy, Levine, & Dennis, 2005). These structural data bolster our hypothesis that developmental changes in relational reasoning are associated with immature functioning of the PFC and parietal cortex.

In the present study, we acquired event-related fMRI data while healthy right-handed participants performed a relational reasoning test based on the RPM task. Prior research indicates that large changes in relational reasoning take place before the age of 12 years, and that mature levels are generally reached in adolescence (Vodegel Matzen, 1994). Thus, we included children aged 8-12 (n=15) and adults aged 18-25 (n=17) in the present study. REL-0, REL-1, and REL-2 trials were pseudorandomly intermixed during the scans (Figure 1).

Methods

Participants

Nineteen adults aged 18-25 and twenty-three children aged 8-12 were recruited through local advertisements and from the University of California in Davis. Participants' consent was obtained according to the Declaration of Helsinki, and the study was approved by the Internal Review Board at the University of California in Davis. Four children were excluded because of equipment failure, and two adults and four children were excluded because of excessive head movement (> 3mm of translation in any direction). A total of thirty-two, right handed participants were included in the study: seventeen 18-25-year-olds (mean age 21.8; 8 men) and fifteen 8-12-year-olds (mean age 9.7; 11 boys). A Chi square analysis indicated that the proportion of women was higher in the adult group than in the children's group, $X^2(1) = 4.16$, $p < .05$. However, follow-up analyses indicated that condition effects did not differ between genders.

Behavioral Assessment

Children and adults participated in a separate behavioral testing session before scanning. Cognitive functioning was assessed using the Kaufmann Brief Intelligence Test (KBIT-2; (Kaufman & Kaufman, 1990). Estimated IQs were 120.5 for 8-12-year-olds and 111.2 for 18-25-year-olds. The differences in IQ between groups were marginally significant ($F(1, 32) = 3.37$, $p = .08$). Children and adults practiced the behavioral tasks in a quiet laboratory. Children were also trained to lie still in a mock scanner, which simulated the environments and sounds of an actual MRI scanner. Parents filled out behavioral questionnaires in this session. Participants were screened for psychiatric conditions with the Child Behavior Check List (CBCL) for children aged 8-17 (Achenbach, 1991), or the symptoms checklist-revised (SCL-R) for 18-25-years. All participants had scores within 1 SD of the mean of a normative standardized sample.

Tasks

Problems had the general form of the RPM test (Raven, 1941). Each problem consisted of a 3 × 3 matrix of stimuli, with the bottom right stimulus missing (Figure 1). Some of the problems were derived from actual RPM questions; others were devised in the laboratories of Drs. Christoff and Bunge. Problems were selected for inclusion based on the results of pilot behavioral data; an effort was made to minimize, to the extent possible, the differences in RTs between the three conditions.

After considering the relationship(s) among the given matrix stimuli, participants had to infer the missing stimulus and select it from among the three choices alternatives presented on the right side of the matrix. There were three levels of relation integration demands (Figure 1). REL-0 trials required no relational processing, because participants were simply asked to perform a visual match to identify the missing stimulus. REL-1 trials involved a change in either the horizontal or the vertical dimension, and required processing of a single relation. REL-2 trials involved two dimensions of change, in both the horizontal and the vertical direction; therefore, inferring the correct answer required consideration of two relationships

between items in the array. During pilot behavioral testing of the problems, problems that could not be solved by children within a 12-second window were eliminated.

On each trial, a relational problem was presented with three possible answers for up to 12 seconds. The three possible answers were mapped to the three buttons that were mapped to the index, middle and ring finger of the left hand. If a response was not made after 10 seconds, a green border appeared around the problem, prompting the participant to respond. Following the response - or after 12 seconds if there was no response - the participant responded to arrows pointed left or right for the remainder of the 22-second trial. The purpose of the arrows task was to keep participants occupied with a low-level cognitive task while the hemodynamic response function associated with RPM task performance returned to baseline. This procedure has been used successfully in prior studies (Christoff, Ream, Geddes, & Gabrieli, 2003) to avoid activating regions associated with introspective mental activity during a low-level baseline condition (Raichle et al., 2001). Prior to scanning, participants were trained extensively on the experimental task to make sure that they understood the task instructions. Before each fMRI scan, the experimenter reminded the participants of the task requirements.

Data acquisition

Trials were presented in five scans of 8 minutes each. The position of the correct stimulus - first, second, or third in the array of possible responses - was counterbalanced within and across conditions. During scanning, 35 REL-0 trials, 35 REL-1 trials and 35 REL-2 trials were presented in pseudorandomized order. The order of trials within each scan was determined by using an optimal sequencing program designed to maximize efficiency of recovery of the blood oxygenation level-dependent (BOLD) response (Dale, 1999).

Scanning was performed with a standard whole-head coil on a 1.5 Tesla General Electric scanner at the University of California at Davis Imaging Research Center. Functional data were acquired by using a gradient-echo echo-planar pulse sequence (TR=2s, TE=40 ms, 24 axial slices, $3.44 \times 3.44 \times 5$ mm, 0-mm interslice gap, 235 volumes per run). Before each scan, four volumes were discarded to allow for T1-equilibration effects. High-resolution T1-weighted anatomical images were collected. Head motion was restricted using a pillow and foam inserts that surrounded the head. Visual stimuli were projected onto a screen that was viewed through a mirror.

fMRI data analysis

Data were preprocessed using SPM2 (Welcome Department of Cognitive Neurology, London). Images were corrected for differences in timing of slice acquisition, followed by a rigid motion correction. For all participants, head movement was 3 mm or less across the entire scan session. Structural and functional volumes were spatially normalized to T1 and echo planar imaging (EPI) templates, respectively. The normalization algorithm used a 12-parameter affine transformation together with a nonlinear transformation involving cosine basis functions. During normalization, the volumes were resampled to 3-mm cubic voxels. Templates were based on the MNI305 stereotaxic space (Cocosco, Kollokian, Kwan, & Evans, 1997), an approximation of Talairach space (Talairach & Tourneaux, 1988). These procedures have been validated for use in children in this age range (Burgund et al., 2002; Kang et al., 2003). Functional volumes were spatially smoothed with an 8-mm full width at half maximum isotropic Gaussian kernel.

Several analytic approaches were undertaken, including 1) timecourse analyses for functionally derived regions-of-interest (ROIs), 2) finite impulse response (FIR) analyses, 3) whole-brain two-sample t-tests comparing children and adults, and 4) whole-brain multiple regression analyses including age and estimated IQ as factors. In a standard GLM analysis, used to derive

ROIs and also to support follow-up tests, the fMRI time series data were modeled as a series of events convolved with a canonical hemodynamic response function (HRF). ROIs were derived from a preliminary analysis in which each trial was modeled as an event-related function starting at the beginning of the trial. Additional tests were performed using an analysis in which each trial was modeled as an event occurring at the midpoint of the trial, between stimulus onset and response. The FIR analysis differs from the standard GLM analysis, in that multiple events are associated with each trial, and these events are not convolved with the HRF. Specifically, each trial was modeled as a series of four impulse functions, at 4-second intervals following trial onset.

In all cases, error trials were modeled separately and were excluded from the analyses. The correct trials functions were used as covariates in a general linear model, along with a basis set of cosine functions that high-pass filtered the data, and a covariate for each session. Motion parameters (translation and rotation) were included as covariates of no interest. The least-squares parameter estimates in height of the best fitting canonical HRF for each condition were used in pair-wise contrasts. The resulting contrast images, computed on a subject-by-subject basis, were submitted to group analyses. At the group level, contrasts between conditions were computed by performing one-tailed t tests on these images, treating subjects as a random effect. Task-related responses were reported if they consisted of at least 5 contiguous voxels that exceeded an uncorrected threshold of $p < .001$ (for the adult group) or $p < .005$ (for the child group). We expected that failure in children to activate regions that were active in adults would be a prime indicator of developmental differences; for this reason, in order to reduce the risk of false negatives in the child data, we selected the less conservative threshold for the child group.

ROI timecourse analyses were performed with the Marsbar toolbox created for use with SPM2 (Brett, Anton, Valabregue, & Poline, 2002). DLPFC, RLPFC, and inferior parietal ROIs were created from activation clusters for the contrast REL-2 > REL-0 at $p<.005$ across all subjects. For the parietal ROI, an anatomical mask of inferior parietal cortex (from the Marsbar distribution) was used to limit the extent of the analyzed region. BOLD time series, averaged across all voxels in an ROI, were extracted for each experimental session. Mean time courses for each condition were then constructed by averaging together appropriate trial time courses, which were defined as 20-second windows of activity after each trial onset. These condition-averaged time courses were then averaged across sessions and subjects.

Results

Performance

Age Group (2) x Relational complexity (3) ANOVAs confirmed that responses to trials with high relational reasoning demands were associated with more errors (F (2, 56) = 40.20, $p < .001$) and slower response times (RT) ($F(2, 58) = 190.48$, $p < .001$) across participants (Figure 2). Post hoc comparisons showed that these differences were significant for all condition comparisons for both accuracy and RT (all p's < .001). Children made more errors than adults on REL-0 trials ($F(1, 30) = 12.56$, $p < .005$), REL-1 trials ($F(1, 30) = 7.89$, $p < .005$), and REL-2 trials ($F(1, 30) = 20.02$, $p < .001$). Critically, as expected, the differences in accuracy across age groups were more pronounced for REL-2 than for REL-1 trials (F (1, 29) = 10.59, $p < .001$), but not for REL-1 than for REL-0 trials ($F < 1$). No age differences were observed in RTs, $F(2, 58)= 2.08$, $p = .15$. These results confirm that developmental changes on this RPM-like task are greatest on the REL-2 problems, for which there is a need to integrate across two relations (see Figure 2).

Region-of-interest timecourse analyses

Due to uncertainty about the timing of prefrontal activation associated with the RPM task in children, we examined the timecourse of BOLD activation in selected regions of interest (ROIs). Functionally defined ROIs for left DLPFC, RLPFC, and Inferior Parietal Cortex (BA 39 and 40) were derived from a whole-brain REL-2 > REL-0 contrast across all participants, in which neural activation was modeled as an event at cue onset (see Methods). The average BOLD activation timecourse associated with each condition was extracted for each ROI and age group.

Because the activation profiles were not constant over time, we included time as a factor in the statistical analysis, examining the interaction between age group, condition, and timepoint separately for each ROI. For ROIs in which this interaction was significant ($p<.016$, adjusted for multiple comparisons across the three ROIs; alpha = .05), we sought to determine the source of the interaction. As such, follow-up ANOVAs were conducted at each timepoint to better understand the evolving pattern of activation (Figure 3).

Within RLPFC, a significant group x condition x timepoint interaction was observed, F (16, 480) = 1.96, $p < .05$. Post hoc comparisons revealed that this interaction was driven by differences during the latter part of the timeseries (between 10 and 18 seconds post-stimulus; see Figure 3). Both groups exhibited early differentiation between REL-2 and REL-1, between 4 and 8 seconds post-stimulus. However, whereas adults demonstrated a consistently higher level of BOLD signal for REL-2 throughout the period of measurement, children demonstrated increased signal for REL-1 during the later phase.

Within DLPFC, a strong group x condition x timepoint interaction was evident, F (16, 480) = 2.62, $p < .005$. Here again, children but not adults demonstrated increased BOLD signal linked to the solution of REL-1 problems during the latter portion of the timeseries. However, adults did show a larger increase for REL-1 relative to REL-0 in DLPFC during the earlier part of the timeseries, between 6 and 10 seconds post-stimulus, as was evident from the group x condition interaction for the peak and surrounding timepoints (between 6 and 10 seconds post-stimulus; $F(2,60) = 5.60$, $p < .01$). The group x condition x timepoint interaction in DLPFC was driven by an increased peak activation in adults relative to children for the REL-2 condition ($F(1,30) = 5.20$, $p < .05$).

The patterns observed in RLPFC and DLPFC were dissociable, as indicated by a significant region x relation x timepoint interaction, $F(16,480) = 2.87$, $p < .001$. Follow-up analyses indicated that the timecourse of activation differed between these regions for the comparison REL-1 vs. -REL-0 (region x relation x timepoint, F (8, 240) = 2.63, $p < .05$), and for the comparison REL-2 vs. -REL-1 (region x relation x timepoint, F (8, 240) = 2.29, $p < .05$). Follow-up analyses confirmed that REL-1 activation was observed later for RLPFC than DLPFC. In contrast, REL-2 activation was present earlier for RLPFC than for DLPFC (see Figure 6). Additionally, there was a significant group x region x condition x timepoint interaction, F (16, 480) = 1.99, $p < .05$. This quadruple interaction was due to the fact that children exhibited lower DLPFC activation relative to adults for REL-1 and REL-2 trials during the early part of the trials. In contrast, children showed a non-significant difference from adults in RLPFC for REL-2 trials early in the trial, but increased sustained activation for REL-1 trials in the latter part of the trial. These findings confirm prior research in adults by showing that RLPFC and DLPFC have different activation profiles during performance of an RPM-type task, and extend these findings by showing distinct patterns of age-related changes in the two regions.

Like the PFC subregions discussed above, regions in IPL (BA 39 and 40) exhibited effects of relational complexity that differed between children and adults (data for left BA 40 displayed

in Figure 3). An ANOVA was conducted with group (children, adults) as a between-subjects factor, and region (left BA 39, left BA 40), condition (REL-0, REL-1, REL-2) as within-subject factors. Overall, parietal activation was greater in adults than in children: there was a main effect of group across regions ($F(1,30) = 8.0$, $p < .01$) but no region x group or region x condition x group interaction (F's < 1). There was a main effect of region ($F(1,30) = 30$, $p < .001$), driven by greater overall activation in BA 39, but no region x condition interaction ($p = .28$). There was a significant group x condition x timepoint interaction ($F(16,480)=3.2$, $p<.001$), indicating group differences in the pattern of activation across conditions. This interaction was driven by the relatively increased activation for REL-1 in children, in both parietal ROIs. In summary, both regions in IPL exhibited similar patterns, with greater overall activation and a more pronounced effect of relational integration in adults than in children.

Finite Impulse Response (FIR) analysis

To visualize the dynamics of activation across the whole brain in children and adults, we performed finite impulse response (FIR) analyses. This analysis, described more fully in the Methods, provides a whole-brain view of the pattern of BOLD activation changes over time. We tested for differing levels of BOLD activation for REL-1 relative to REL-0 and for REL-2 relative to REL-1 trials during each four-second time-window following stimulus onset (Figure 5).

In adults, BOLD signal increases related to 1st-order relational processing were most evident between 4 and 8 seconds, bilaterally in DLPFC, VLPFC, and superior parietal cortex. The timing of this signal indicates that 1st-order relational processing was engaged shortly after stimulus onset, and for a relatively short duration. Relational integration was associated with similar BOLD increases in DLPFC, VLPFC, and parietal cortex, and uniquely with increases in RLPFC. In addition, DLPFC, RLPFC, and superior parietal BOLD signal increases associated with relational integration were sustained over a longer duration, through the 8-12 second and 12-16 second time-bins. Thus, relational integration appears to involve activity that is more sustained than, or delayed relative to, the processing of 1st-order relational information. This activity includes RLPFC as well as most if not all of the regions involved in 1st-order relational processing.

In children, the pattern of activation over time, across the two principal contrasts, was markedly different than in adults. For children, BOLD increases uniquely associated with relational integration were only seen in the initial two time-bins, and were most apparent in the first one (0-4 seconds post-stimulus). However, we observed late increases (between 8 and 16 seconds post-stimulus) in bilateral DLPFC, associated with 1st-order relational processing. Thus, while adults appeared to engage DLPFC and RLPFC in a sustained manner in order to support relational integration, sustained engagement of DLPFC in children was observed even for the simpler 1st-order relational processing task.

Whole-brain analyses: group comparisons and multiple regression analyses

Because the timecourse and FIR analyses revealed that effects of interest were associated mainly with the middle and later parts of a trial, rather than with the stimulus onset, we modeled neural activity at the midpoint of each trial in subsequent analyses. Overall activation patterns for the basic midpoint analysis were consistent with the FIR analysis (see Supplementary Table). In a two-sample comparison for the REL-2>REL-0 contrast, we found that adults showed significantly greater activation in right RLPFC (Figure 5a). Children exhibited significantly greater activation than adults in precentral gyrus. No significant group differences were observed in whole-brain two-sample t-tests for REL-2>REL-1 or REL-1>REL-0.

To better understand the neural correlates of relational integration in children, we included estimated IQ (KBIT-2 score) and age as covariates in a second-level analysis of the REL-2>REL-1 contrast for children. This analysis revealed a strong positive relationship (at $p<.001$) between age and right RLPFC activation (Figure 5b). There was no significant effect of KBIT-2 score on brain activation.

Discussion

This study examined the neural correlates of developmental differences in relational reasoning. As expected, children performed well on the REL-0 and REL-1 problems, which required, respectively, perceptual matching and consideration of one dimension of change across the stimulus array. In contrast, children made disproportionately more errors than adults on the REL-2 problems as compared to REL-0 or REL-1 (see also (Richland et al., 2006) for similar findings in a picture analogies task). This difference in performance was observed despite the slow presentation of the trials, so differences could not be associated with a speed instruction. In effect, children exhibited similar RTs to adults on REL-2 problems, despite performing more slowly than adults on REL-0 and REL-1 problems. This finding suggests that children failed to allocate sufficient time to the REL-2 problems, tending to select an answer without having considered both dimensions of relational change.

Indeed, prior behavioral studies have suggested that children select answers without considering all the possible relations of a problem. Vodegel Matzen, and van der Molen (Vodegel Matzen, 1994) analyzed the types of errors that 9-14-year olds tended to make on a set of complex RPM-type problems with 9-stimulus arrays, three possible dimensions of change (horizontal, vertical, and diagonal), and eight response options. Based on the participants' choices, the researchers found that lower-scoring children tended to use some but not all of the relevant rules to solve a problem. This finding is consistent with prior results of Carpenter et al. (1990) in college students, which suggested that lower-scoring adults experience more difficulty systematically organizing information in working memory.

The FIR analyses demonstrated that, in adults, relational integration was associated with sustained activation in RLPFC and DLPFC relative to the conditions that did not require relational integration. Analyses of ROI timecourses confirmed that adults recruited RLPFC and DLPFC more strongly on REL-2 than REL-1 and REL-0 problems. RLPFC did not differentiate between REL-0 and REL-1 trials (REL-2 > REL-0, REL-1), whereas DLPFC exhibited an increase from REL-0 to REL-1 and from REL-1 to REL-2. These results for adults replicate those of Christoff et al. (Christoff et al., 2001), and support the hypothesis that RLPFC is important for integrating the outcomes of two or more separate cognitive operations (Bunge et al., 2005; Bunge & Zelazo, 2006; Ramnani & Owen, 2004; Wendelken et al., in press). For DLPFC, by comparison, the evidence does not suggest specific involvement in relational integration. Rather, increased activation in DLPFC during relational reasoning has been associated with increased working memory and response selection requirements (Christoff et al., 2001; Kroger et al., 2002).

Children aged 8-12 exhibited a different activation pattern from adults in both RLPFC and DLPFC. Both groups recruited RLPFC specifically for REL-2 problems. However, children recruited RLPFC more for REL-2 relative to REL-1 problems early in the trial, but did not show the sustained increase for REL-2 over REL-1 seen in adults. This pattern of strong early differentiation followed by a lack of sustained differentiation suggests that children solve the problems differently than adults. They recruited RLPFC, but did not engage it preferentially for REL-2 problems, suggesting that they failed to persevere in considering multiple relations, and therefore tended to select a response without having considered all aspects of a problem. This apparent difference in the way children and adults tackle the RPM problems is supported

by the pattern of behavior. As noted previously, children made disproportionately more errors on REL-2 trials relative to the other conditions than adults, but - if anything - exhibited a trend towards a disproportionately *lower* RT cost on these trials. Finally, the evidence that activation in RLPFC associated with relational integration increases as a function of age suggests that development of the RLPFC integration mechanism occurs over the age range (8-12 years) that was studied.

The late engagement of RLPFC in children, for both REL-1 and REL-2 problems, may not be entirely explained by pre-response neural activity. Post-response processes such as response evaluation may also contribute to the observed BOLD activation. In this task, response evaluation should involve similar relational processing demands as response generation. If children are engaging post-response evaluative processes to a greater extent than adults, it would suggest greater uncertainty about the accuracy of their responses. However, it should be noted that prefrontal and parietal activation peaks occurred before motor cortex peaks, as was observed in a comparative analysis of BOLD timecourses (not shown), so post-response processing is unlikely to be a dominant factor in the activation of these regions.

In contrast to the pattern in RLPFC, children showed initially reduced activation relative to adults in DLPFC. This reduced activation is consistent with studies demonstrating more DLPFC activation in adults relative to children for other cognitive processes, such as working memory manipulation (Crone et al., 2006) and spatial working memory maintenance (Klingberg, Forssberg, & Westerberg, 2002), and with the slow structural development of DLPFC. However, children did show sustained activation of DLPFC for REL-1 problems. This finding, combined with the fact that their error rate for these problems was similar to that of adults, suggests that children do successfully, though perhaps less efficiently, employ DLPFC for the solution of 1st-order relational problems. Unlike adults, however, they also appear to employ RLPFC to solve REL-1 problems.

Summarizing the prefrontal results, these data are consistent with prior work in adults in showing that RLPFC is more selectively engaged by 2nd-order relational processing than DLPFC (Christoff et al., 2001), and further indicate that the RLPFC specialization for relational integration that is seen in adults is still undergoing changes in children between the ages of 8 and 12. In children, unlike adults, this region is still engaged for 1st-order as well as for 2nd-order relational processing. The involvement of DLPFC in relational processing has largely emerged in children 8-12; however, further development of this functionality is assumed to take place during adolescence, given the differences observed between age groups here.

Like PFC, parietal cortex has been implicated in prior studies of fluid reasoning (Christoff et al., 2001; Kroger et al., 2002; Lee et al., 2006). As expected, SPL was modulated by relational complexity; this finding may be explained at least in part by increased visuospatial processing demands accompanying the increases in relational complexity. We sought to test the hypothesis that the IPL, not associated in the literature with visuospatial processing, would be sensitive to the number of relations being considered. Indeed, IPL in adults was, like lateral PFC, sensitive to relational complexity - and exhibited an immature pattern of activation in children. Thus, developmental improvements in relational reasoning are likely to be due to changes in both prefrontal and parietal function. These results are consistent with individual differences data revealing that variance in RPM performance among adults could be accounted for in large part by differences in the level of engagement of lateral PFC and IPL during performance of a working memory task (Gray et al., 2003).

In the present study, children displayed different temporal dynamics in RLPFC relative to adults. Prior studies that have compared BOLD response time courses between children and adults have not revealed systematic age-related differences in hemodynamic response

functions (Kang et al., 2003), and, indeed, our studies have typically observed similar rise times and time-to-peak in children and adults. In a developmental fMRI study that involved visual analogies, we have found that children aged 7-12 can show delayed activation of RLPFC (and VLPFC) relative to adults (Wright, Matlen, et al, under review). But in this study as in the current one, children showed reduced activation in left RLPFC for relational integration relative to a simple relational task, as well as increasing levels of RLPFC activation with increasing age.

Prior studies tapping other cognitive processes than relational reasoning have reported that age differences in performance can be associated with both increased and reduced brain activation (for review, see (Bunge & Wright, 2007). Within the context of a single study, Olesen et al. (2006) reported that working memory performance in the face of distraction was associated with reduced DLPFC activation during a delay period in children relative to adults, but increased DLPFC activation during distraction (Olesen, Macoveanu, Tegner, & Klingberg, 2006). This differential pattern of activation indicates that across development, some cognitive mechanisms - which rely on distinct yet partially overlapping neural circuits - mature relatively late in development, as revealed by an age-related increase in activation, and others become more efficient with age, as revealed by an age-related increase in sustained activation.

To our knowledge, this is the first study examining the neural correlates of the development of relational reasoning, and provides evidence that the development of relational reasoning is associated with changes in the contribution of RLPFC to relational integration, and in the contribution of DLPFC to relational processing more generally. Our laboratory is embarking on a longitudinal study to measure within-person improvements in reasoning ability during development, and how these changes relate to changes in brain structure and function.
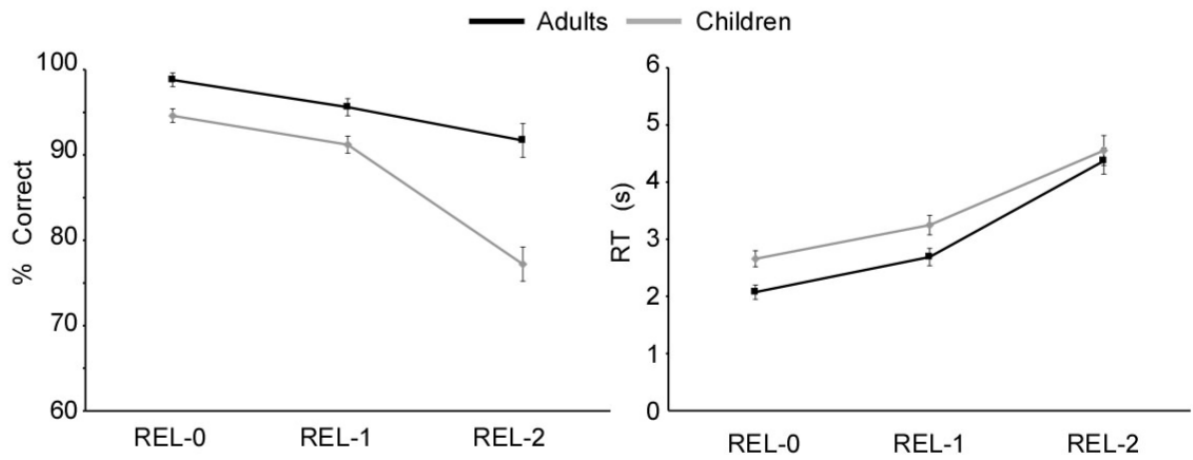
References

Achenbach, TM. Manual for the child behavior checklist/4-18 and 1991 profile. 1991.

Braver TS, Bongiolatti SR. The role of frontopolar cortex in subgoal processing during working memory. Neuroimage 2002;15(3):523–536. [PubMed: 11848695]

Brett, M.; Anton, JL.; Valabregue, R.; Poline, JB. Region of interest analysis using an spm toolbox; Paper presented at the 8th International Conference on Functional Mapping of the Human Brain; Sendai, Japan. 2002;

Bunge SA, Wendelken C, Badre D, Wagner AD. Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. Cereb Cortex 2005;15(3):239–249. [PubMed: 15238433]

Bunge SA, Wright SB. Neurodevelopmental changes in working memory and cognitive control. Curr Opin Neurobiol 2007;17(2):243–250. [PubMed: 17321127]

Bunge SA, Zelazo PD. A brain-based account of the development of rule use in childhood. Current Directions in Psychological Science 2006;15:118–121.

Burgess PW, Gilbert SJ, Dumontheil I. Function and localization within rostral prefrontal cortex (area 10). Philos Trans R Soc Lond B Biol Sci 2007;362(1481):887–899. [PubMed: 17403644]

Cattell, RB. Abilities: Their structure, growth and action. Houghton-Mifflin; Boston: 1971.

Cattell, RB. Intelligence: Its structure, growth and action. North-Holland; Amsterdam: 1987.

Christoff K, Gabrieli JDE. The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. Psychobiology 2002;28(2):168–186.

Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, et al. Rostrolateral prefrontal cortex involvement in relational integration during reasoning. Neuroimage 2001;14(5):1136–1149. [PubMed: 11697945]

Christoff K, Ream JM, Geddes LP, Gabrieli JD. Evaluating self-generated information: Anterior prefrontal contributions to human cognition. Behav Neurosci 2003;117:1161–1168. [PubMed: 14674837]

Cocosco CA, Kollokian V, Kwan RKS, Evans AC. Brain web: Online interface to a 3d mri simulated brain database. Neuroimage 1997;5:S452.

Crone EA, Wendelken C, Donohue S, van Leijenhorst L, Bunge SA. Neurocognitive development of the ability to manipulate information in working memory. Proc Natl Acad Sci U S A 2006;103(24):9315–9320. [PubMed: 16738055]

Dale AM. Optimal experimental design for event-related fmri. Hum Brain Mapp 1999;8(23):109–114. [PubMed: 10524601]

Duncan J. Intelligence tests predict brain response to demanding task events. Nature Neuroscience 2003;6:207–208.

Fuster, JM. Cortex and mind: Unifying cognition. Oxford University Press; 2003.

Giedd JN. Structural magnetic resonance imaging of the adolescent brain. Ann N Y Acad Sci 2004;1021:77–85. [PubMed: 15251877]

Gray JR, Chabris CF, Braver TS. Neural mechanisms of general fluid intelligence. Nat Neurosci 2003;6(3):316–322. [PubMed: 12592404]

Gross CA, Graziano MS. Multiple representations of space in the brain. The Neuroscientist 1995;1:43–50.

Halford GS, Wilson WH, Phillips S. Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. Behav Brain Sci 1998;21(6):803–831. [PubMed: 10191879]

Horn, JL.; Noll, J. Human cognitive capabilities: Gf-gc theory. Guilford Press; New York: 1997.

Kaufman, AS.; Kaufman, NL. Manual for the kaufman brief intelligence test. 1990.

Klingberg T, Forssberg H, Westerberg H. Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. J Cogn Neurosci 2002;14(1):1–10. [PubMed: 11798382]

Koechlin E, Basso G, Pietrini P, Panzer S, Grafman J. The role of the anterior prefrontal cortex in human cognition. Nature 1999;399(6732):148–151. [PubMed: 10335843]

Kroger JK, Sabb FW, Fales CL, Bookheimer SY, Cohen MS, Holyoak KJ. Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. Cereb Cortex 2002;12(5):477–485. [PubMed: 11950765]

Lee KH, Choi YY, Gray JR, Cho SH, Chae JH, Lee S, et al. Neural correlates of superior intelligence: Stronger recruitment of posterior parietal cortex. Neuroimage 2006;29(2):578–586. [PubMed: 16122946]

Morrison RG, Krawczyk DC, Holyoak KJ, Hummel JE, Chow TW, Miller BL. A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. J Cogn Neurosci 2004;16(2):260–271. [PubMed: 15068596]

O'Donnell S, Noseworthy MD, Levine B, Dennis M. Cortical thickness of the frontopolar area in typically developing children and adolescents. Neuroimage 2005;24(4):948–954. [PubMed: 15670671]

Olesen PJ, Macoveanu J, Tegner J, Klingberg T. Brain activity related to working memory and distraction in children and adults. Cereb Cortex. 2006

Petrides M, Pandya DN. Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. J Comp Neurol 1984;228(1):105–116. [PubMed: 6480903]

Prabhakaran V, Smith JAL, Desmond JE, Glover GH, Gabrieli JDE. Neural substrates of fluid reasoning: An fmri study of neocortical activation during performance of the raven's progressive matrices test. Cognitive Psychology 1997;33:43–63. [PubMed: 9212721]

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. Proc Natl Acad Sci U S A 2001;98:676–682. [PubMed: 11209064]

Ramnani N, Owen AM. Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. Nat Rev Neurosci 2004;5(3):184–194. [PubMed: 14976518]

Raven JC. Standardization of progressive matrices. Brit J Med Psychol 1941;19:137–150.

Richland LE, Morrison RG, Holyoak KJ. Children's development of analogical reasoning: Insights from scene analogy problems. J Exp Child Psychol 2006;94(3):249–273. [PubMed: 16620867]

Sternberg RJ, Rifkin B. The development of analogical reasoning processes. J Exp Child Psychol 1979;27(2):195–232. [PubMed: 438723]

Talairach, J.; Tourneaux, P. Co-planar stereotaxic atlas of the human brain. 1988.

Vodegel Matzen, L. Unpublished masters thesis. Faculty of Psychology, University of Amsterdam; 1994. Performance on raven's progressive matrices: What makes a difference?.

Wagner G, Koch K, Reichenbach JR, Sauer H, Schlosser RGM. The special involvement of rostrolateral prefrontal cortex in planning abilities: An event-related fmri study with the tower of london paradigm. Neuropsychologia 2006;44:2337–2347. [PubMed: 16806311]

Waltz JA, Knowlton BJ, Holyoak KJ, Boone KB, Mishkin FS, de Menenzes Santos M. A system for relational reasoning in human prefrontal cortex. Psychol Sci 1999;10(2):119–125.

Wendelken C, Nakhabenko D, Donohue SE, Carter CS, Bunge SA. Brain is to thought as stomach is to...?: Specifying the role of rostrolateral prefrontal cortex in analogical reasoning. Journal of Cognitive Neuroscience. in press

Wendelken C, Bunge SA, Carter CS. Maintaining Structured Information: an Investigation into the Roles of Parietal and Lateral Prefrontal Cortices. Neuropsychologia. 2007doi:10.1016/j.neuropsychologia. 2007.09.015

Wright, SB.; Matlen, BJ.; Baym, CL.; Ferrer, E.; Bunge, SA. Neural Correlates of Fluid Reasoning in Children and Adults. under review
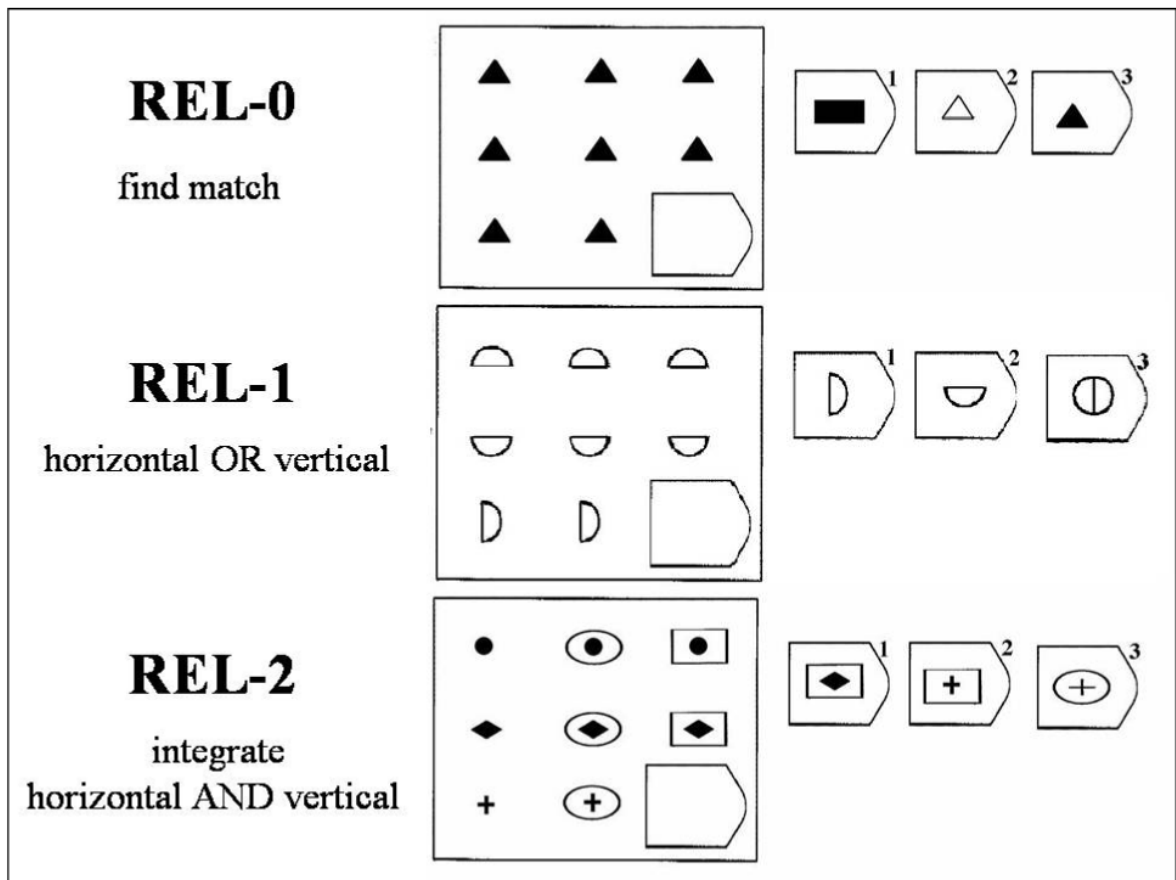
Figure 1.
The Raven's Progessive Matrices task, including sample REL-0, REL-1, and REL-2 problems.
Participants were instructed to indicate the correct solution from the three possible answers on
the right with a button press.

Figure 2.
Accuracy and response times associated with task performance during fMRI data acquisition. Children performed disproportionately worse on REL-2 problems than adults, but there were no group differences in RTs for this condition.
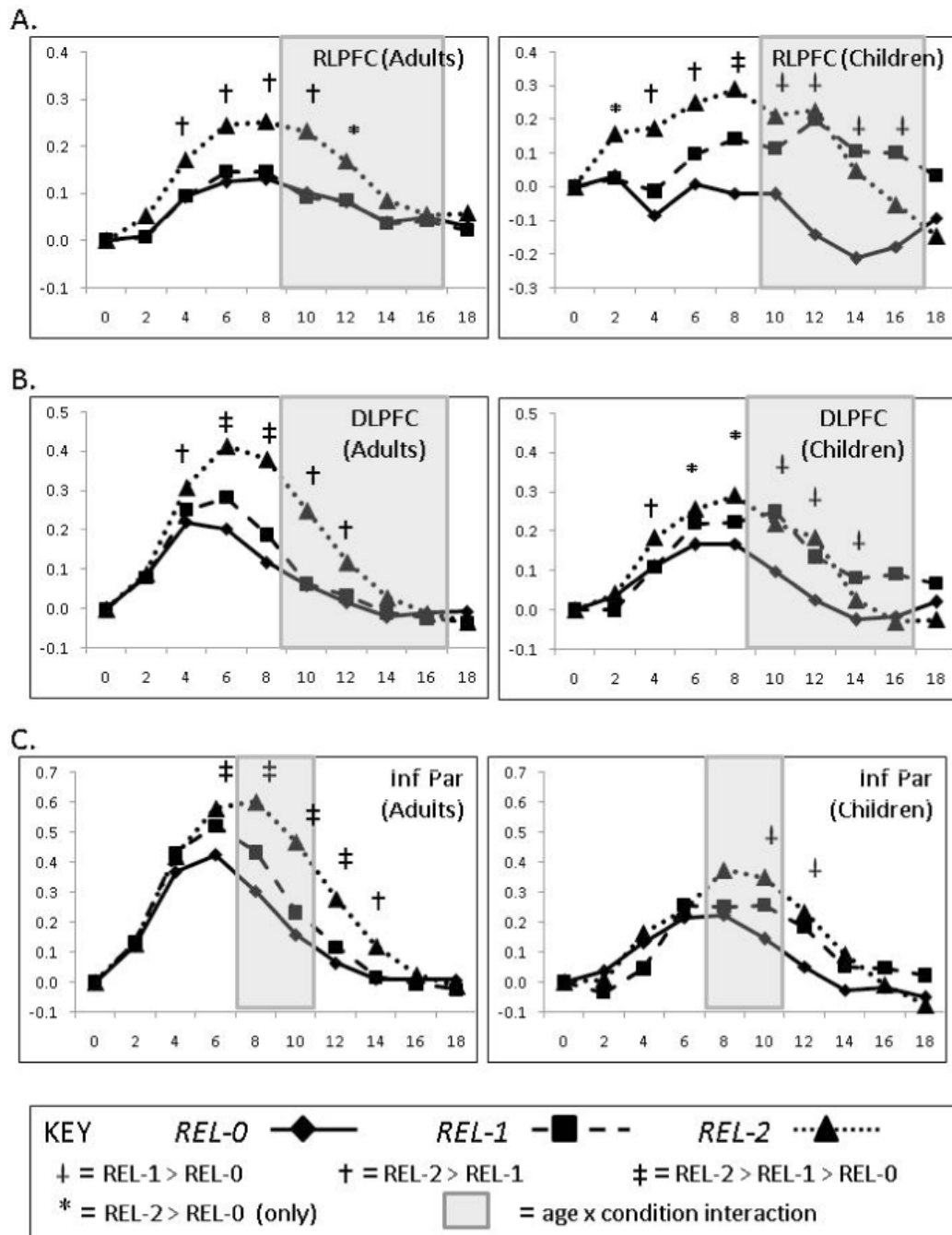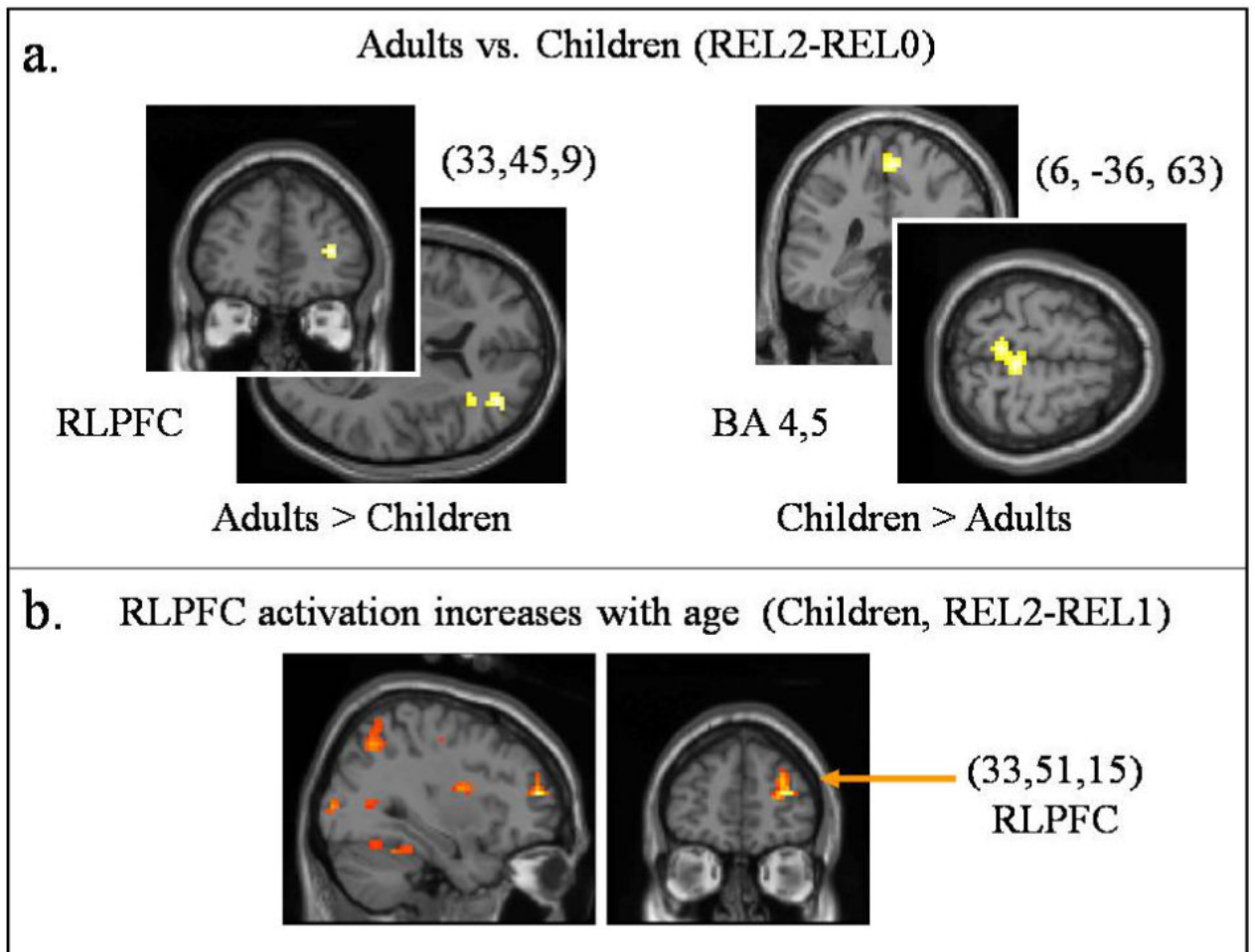
Figure 3.
ROI timecourses. (A) Timecourses from left RLPFC ROIs in adults and children. (B) Timecourses from left DLPFC ROIs in adults and children. (C) Timecourses from right inferior parietal cortex ROIs in adults and children. Significant effects of condition at each timepoint are indicated by means of a symbol above that timepoint (see key at bottom of figure). Timepoints exhibiting significant ($p<0.05$) age x condition interactions are indicated via shading. Marginal interactions ($p<0.1$) that are surrounded by significant interactions are also shaded.
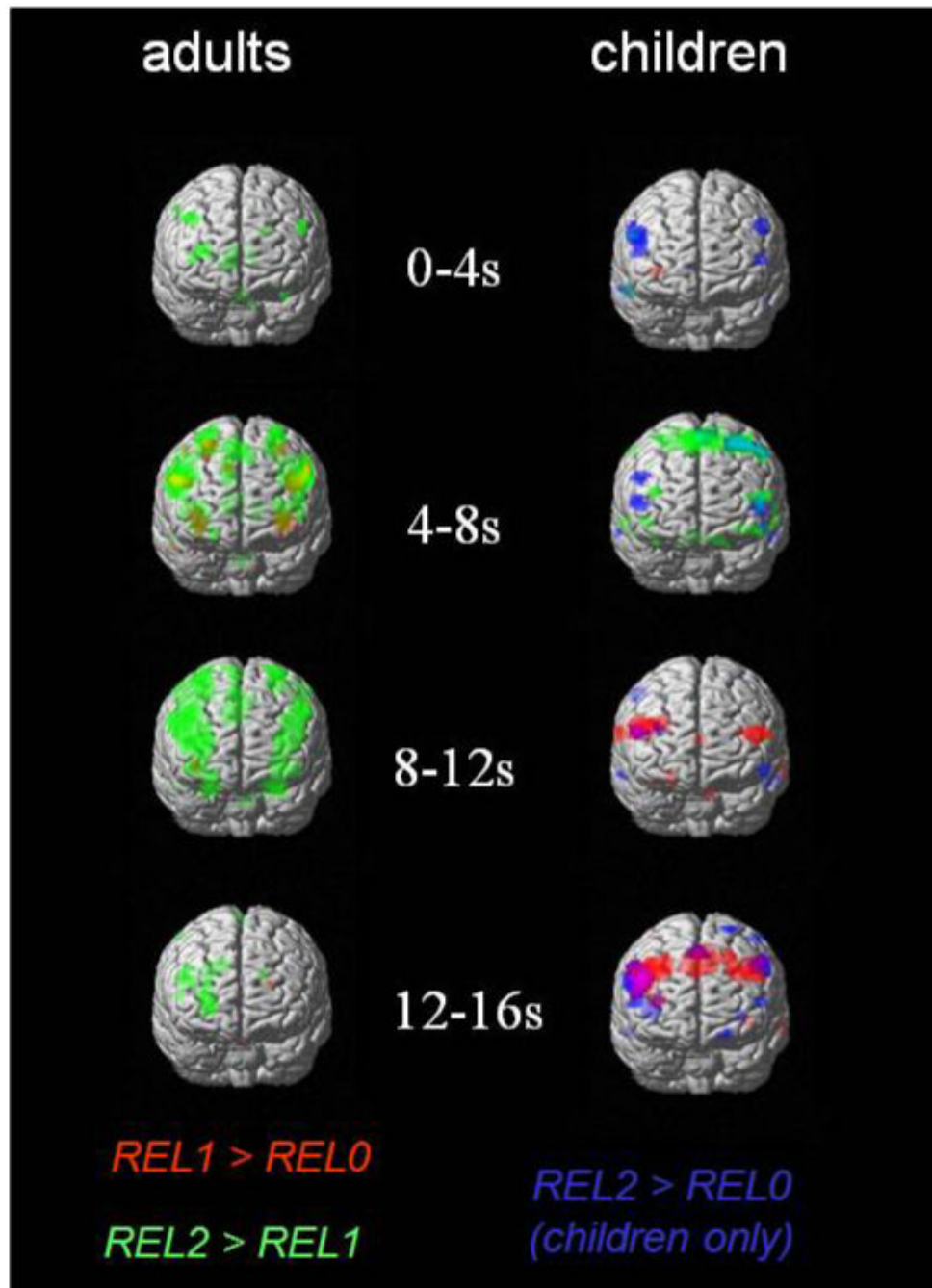
Figure 4.
FIR analysis results in adults and children, showing activation related to first-order relational processing (REL-1 > REL-0; shown in red) and relational integration (REL-2 > REL-1; shown in green). Images are thresholded at p < .001 for adults and at p<.005 for children.

Figure 5.
(A) Results of a two-sample t-test, comparing adults with children on the REL-2 > REL-0 contrast. Notably, this anlaysis revealed greater right RLPFC activation in adults than children. (B) Regions, including right RLPFC, that showed a pattern of increasing activation as a function of age in children.