

# Increased Functional Selectivity over Development in Rostrolateral Prefrontal Cortex

Carter Wendelken,<sup>1\*</sup> Elizabeth D. O'Hare,<sup>1\*</sup> Kirstie J. Whitaker,<sup>1</sup> Emilio Ferrer,<sup>3</sup> and Silvia A. Bunge<sup>1,2</sup>

<sup>1</sup>Helen Wills Neuroscience Institute and <sup>2</sup>Department of Psychology, University of California, Berkeley, Berkeley, California 94720, and <sup>3</sup>Department of Psychology, University of California, Davis, Davis, California 95616

Relational reasoning, or the ability to identify and consider relationships between multiple mental representations, is a fundamental component of high-level cognition (Robin and Holyoak, 1995). The capacity to reason with relations enables abstract thought and may be at the core of what makes human cognition unique (Penn et al., 2008). This capacity improves throughout childhood and adolescence (Ferrer et al., 2009). Here, we sought to better understand the neural mechanisms that support its emergence. We have hypothesized previously, based on fMRI research in adults, that (1) inferior parietal lobe (IPL) plays a central role in representing relationships between mental representations (first-order relations) and (2) rostrolateral prefrontal cortex (RLPFC) integrates inputs from IPL to build second-order relational structures (i.e., relations between relations). In the present study, we examined fMRI and cortical thickness data from 85 children and adolescents (ages 6–18 years). Participants performed a relational matching task in which they viewed arrays of four visual stimuli and determined whether two stimuli shared a particular feature (a first-order relational judgment) or whether two pairs of stimuli matched according to the same feature (a second-order relational judgment). fMRI results provide evidence for increased functional selectivity across ages 6–18 years in RLPFC and IPL. Specifically, young children engaged RLPFC and IPL indiscriminately for first-order and second-order relational judgments, and activation for first-order relations diminished with age whereas activation for second-order relations stayed elevated. Examination of cortical thickness revealed that increased functional selectivity in RLPFC could be partly accounted for by cortical thinning in IPL.

## Introduction

Relational reasoning, or the ability to consider relationships between multiple mental representations, is a fundamental component of high-level cognition (Robin and Holyoak, 1995). Neuropsychological studies have demonstrated that improvements in relational reasoning occur throughout childhood and adolescence (Sternberg and Rifkin, 1979; McArdle et al., 2002; Richland et al., 2006; Ferrer et al., 2009), and several recent studies have begun to probe the neural basis of these age-related changes (Wright et al., 2007; Crone et al., 2009; Eslinger et al., 2009; Dumontheil et al., 2010). The current study builds on this work, examining the changes in brain structure and function that underlie developmental improvements in relational reasoning from age 6 to 18 years.

Relational reasoning problems can vary in the complexity of the relational task to be performed (Halford, 1992). First-order relational processing involves consideration of individual relations, in isolation. Second-order relational processing involves

consideration of relations in combination, or relations of relations. It is the capacity for higher-order relational processing that is thought to underlie abstract thought and to support many of the complex cognitive operations that are thought to be uniquely human (Penn et al., 2008).

Three brain areas stand out as being particularly involved in relational reasoning: rostrolateral prefrontal cortex (RLPFC), dorsolateral prefrontal cortex (DLPFC), and the inferior parietal lobe (IPL) in the vicinity of the intraparietal sulcus (for review, see Krawczyk et al., 2011). Based on our previous research (Wendelken and Bunge, 2010; Wendelken et al., 2011), we have hypothesized that (1) IPL and DLPFC represent and process first-order relations, and (2) RLPFC integrates inputs from these regions to process second-order relations.

Several recent fMRI studies have reported developmental changes in these regions associated with relational reasoning. In one study, children demonstrated increasing activation with age in parietal cortex concomitant with decreasing activation with age in lateral PFC, for first-order problems (Eslinger et al., 2009). In another study, children and adults demonstrated similar engagement of RLPFC for second-order problems, but children, unlike adults, also engaged RLPFC for first-order problems (Crone et al., 2009). In another study, older adolescents demonstrated greater activation in left RLPFC for second-order relative to first-order problems than mid-adolescents or adults (Dumontheil et al., 2010). We have hypothesized that RLPFC undergoes a process of specialization during middle childhood, changing from a region that is involved in general relational

Received Sept. 21, 2011; revised Sept. 13, 2011; accepted Sept. 16, 2011.

Author contributions: E.F. and S.A.B. designed research; E.D.O. and K.J.W. performed research; C.W., E.D.O., E.F., and S.A.B. analyzed data; C.W., E.D.O., E.F., and S.A.B. wrote the paper.

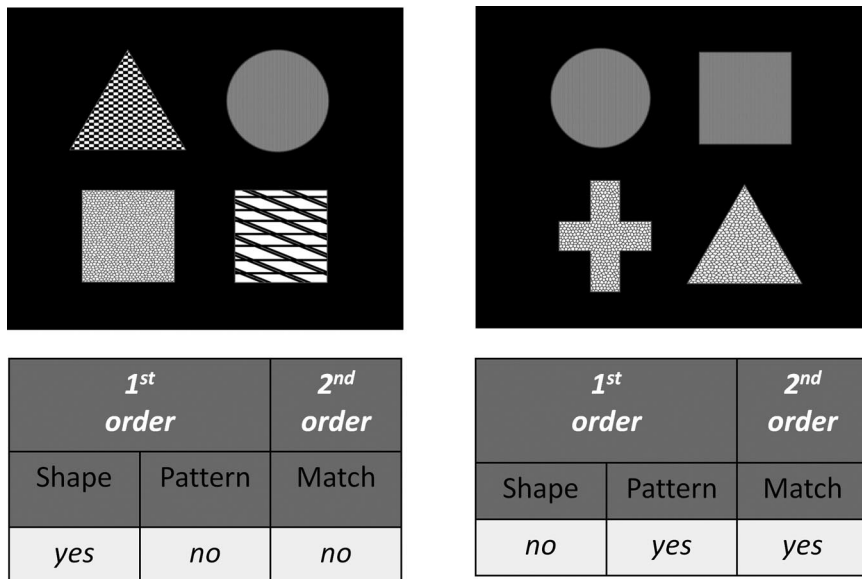
This work was supported by NIH/NINDS Grant R01 NS057146 (S.A.B., E.F.). We thank Chloe Green, Mehdi Boudhaddou, Ori Elis, and Alexis Ellis for assistance with data collection.

\*C.W. and E.D.O. contributed equally to this work.

Correspondence should be addressed to Carter Wendelken, 34A Barker Hall, Berkeley, CA 94720-1650. E-mail: cwendelken@berkeley.edu.

DOI:10.1523/JNEUROSCI.1193-10.2011

Copyright © 2011 the authors 0270-6474/11/3117260-09\$15.00/0



**Figure 1.** Task illustration. Each trial consisted of a yes/no judgment based on a stimulus array that contained four patterned shapes. On Shape trials, participants determined whether there was a shape match in either pair. On Pattern trials, participants determined whether there was a pattern match in either pair. On Match trials, participants decided whether the bottom pair of stimuli matched along the same dimension (i.e., Shape or Pattern) as the top pair of stimuli. Shape and Pattern trials require first-order relational processing, whereas Match trials require second-order relational integration.

processing to one that is specialized for second-order relational processing (Crone et al., 2009). The first goal of the current study is to test this hypothesis in a large sample of typically developing children.

Many brain regions undergo changes in cortical thickness during childhood (Shaw et al., 2008; Giedd and Rapoport, 2010). We hypothesized that synaptic pruning of local circuitry—reflected in cortical thinning (O'Donnell et al., 2005; Tamnes et al., 2010)—might be a mechanism of increasing functional selectivity (Ferrer et al., 2009). The second goal of this study is thus to test this hypothesis by examining the connection between cortical thickness and functional activation in the regions that are most involved in relational reasoning.

## Materials and Methods

**Study design.** As part of a larger study, participants completed a battery of cognitive assessments, two fMRI tasks (a relational matching task and a visual analogy task), one or more anatomical scans, and a diffusion tensor imaging scan.

The present investigation is limited to analysis of the fMRI and behavioral data from one of the two tasks (relational matching) and cortical thickness data extracted from the anatomical scans. To make the point that the relational matching task used here is a good measure of reasoning capacity, we report the correlation across individuals between performance on this task and a fluid reasoning factor score (FR) calculated on the basis of four standardized cognitive assessments: “Block Design” and “Matrix Reasoning” subtests of the Wechsler Intelligence Scale for Children (Wechsler, 1974) and “Analysis-Synthesis” and “Concept Formation” subtests of the Woodcock–Johnson Test of Achievement (Woodcock and Johnson, 1990). The factor scores were extracted from a model in which the four observed scales loaded freely on a single factor representing fluid reasoning.

**Participants.** The present study includes fMRI data from 85 typically developing individuals (aged 6–18 years; mean  $\pm$  SD age,  $12.8 \pm 3.1$  years), including 35 females (mean  $\pm$  SD age,  $12.2 \pm 3.0$  years) and 50 males (mean  $\pm$  SD age,  $13.3 \pm 3.1$  years). Male participants were marginally older than female participants ( $p = 0.09$ ). There was no relationship between age and IQ ( $r = -0.01$ ). Cortical thickness data are also included for 79 participants (mean  $\pm$  SD age,  $12.9 \pm 2.9$  years; 45 males),

because six subjects who were included in the fMRI analysis were excluded from the structural analysis as a result of poor-quality anatomical reconstruction (see below).

Data from an additional 24 individuals were excluded because of excessive scanner movement ( $n = 17$ ; aged 6–12 years, mean  $\pm$  SD age,  $8.6 \pm 1.7$  years) or chance-level performance on any condition in the fMRI task ( $n = 7$ ; aged 6–18 years; mean  $\pm$  SD age,  $10.6 \pm 4.4$  years). Exclusion as a result of movement was determined on the basis of analysis with the ArtRepair tool, described below.

All participants were screened for neurological impairments, psychiatric illness, history of learning disability, and developmental delay. All participants and their parents gave their informed assent or consent to participate in the study, which was approved by the Committee for Protection of Human Subjects at the University of California, Berkeley.

**Experimental task.** Participants performed a relational matching task similar to one used previously in our laboratory (Bunge et al., 2009; cf. Christoff et al., 2003; Smith et al., 2007). The task was designed to require either first-order or second-order relational judgments. Arrays of four patterned shapes were presented on every trial, and participants were required to make a yes/no judgment regarding

the presence or absence of a specific kind of match using a button box (Fig. 1). For first-order trials, participants were instructed to match either shape or pattern. After the shape instruction, participants decided whether there was a shape match for either the top or bottom pair of items. Similarly, after the pattern instruction, participants decided whether there was a pattern match for either pair. For second-order trials, participants determined whether the top and bottom item pairs matched along the same dimension (shape or pattern). Thus, on second-order trials but not on first-order trials, participants had to integrate information across two item pairs.

The experiment was run as a blocked design with three scans of 5 min, 25 s each. Each scan consisted of three 90-s blocks, one for each of the three instructions (“Shape” or “Pattern” for first-order blocks, “Match” for second-order blocks). Each 15-trial block began with an instruction screen, and the instruction for each block remained onscreen for the duration of the block. Each trial lasted 6 s, and the stimulus array remained on screen for the duration of the trial. Two 20-s rest blocks, during which participants fixated on a crosshair, alternated with the task blocks, and the scan started with 6 s of rest and ended with 10 s of rest. A total of 45 Shape trials, 45 Pattern trials, and 45 Match trials were acquired per session. Across all trials, 50% of stimuli within a condition warranted a “yes” response. There were 45 stimulus arrays total, and each stimulus array was paired once with each of the three instructions; thus, only the instructions, and not the stimulus arrays, differed between the different trial types.

Participants were trained on the task before the start of the scanning session. Training included a verbal explanation of the task with example stimuli. Once in the scanner, participants completed a short practice run. All instructions were reviewed again before the start of functional data collection.

**Data acquisition.** Brain imaging data were collected on a Siemens 3 T Trio system at the University of California, Berkeley Brain Imaging Center. Participants viewed stimuli backprojected onto a projection screen with a mirror mounted on the head coil and responded using a button box held in their right hand. Stimulus presentation and response acquisition were controlled with Presentation psychological experimentation software (Neurobehavioral Systems).

Two high-resolution T1-weighted MPRAGE anatomical scans (TR, 2300 ms; TE, 2.98 ms;  $1 \times 1 \times 1$  mm voxels) were acquired for cortical morphometric analyses. These two MPRAGE scans were averaged dur-

ing post-processing to increase the signal-to-noise ratio. Multi-slice echo-planar imaging (EPI) was used to collect functional imaging data (gradient-echo EPI sequence; TR, 2000 ms; TE, 25 ms; 33 axial slices;  $2.0 \times 1.8 \times 3.0$  mm voxels; no interslice gap; flip angle,  $90^\circ$ ; field of view, 230 mm; 165 volumes per run).

**Image and statistical analysis.** Functional imaging data were preprocessed and analyzed using SPM5 (Wellcome Trust Center for Neuroimaging). The first three volumes from each functional scanning run were discarded to allow for T1 equilibration. Functional images were corrected for differences in slice acquisition timing and were realigned to the first volume by means of rigid-body motion transformation. Motion parameters were extracted from this process and were used to inform a volume repair procedure (ArtRepair; Stanford Psychiatric Neuroimaging Laboratory).

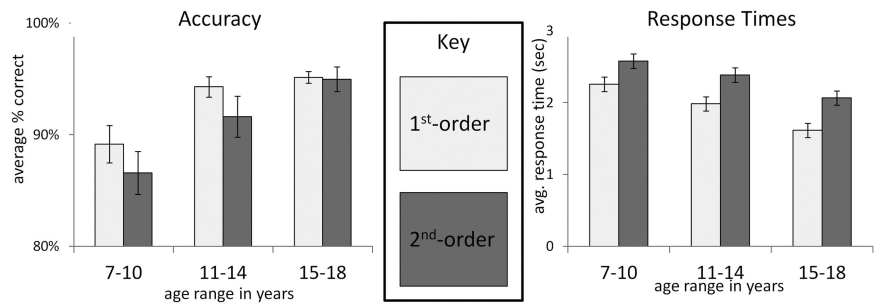
ArtRepair identified bad volumes on the basis of within-scan movement and signal fluctuations and then corrected bad signal values via interpolation and also via deweighting in the subsequent analyses. Any run with  $>20\%$  repaired volumes was discarded, and any participant with more than one run discarded in this manner was excluded from additional preprocessing or analysis.

After volume repair, the mean structural image was coregistered to the mean realigned functional image and then spatially normalized to the T1 anatomical template of SPM5. Normalization parameters obtained from this process were then applied to the un-normalized functional images to produce a set of functional images in SPM standard space (MNI152), with  $3 \times 3 \times 3$  mm voxels. As a final preprocessing step, functional images were smoothed with an 8 mm FWHM isotropic Gaussian kernel.

Data analysis was performed using a general linear model (GLM) that incorporated task effects, session effects, and a general linear trend. Task effects were modeled as 90-s block regressors, convolved with the canonical hemodynamic response function in SPM. This GLM was used to compute the least-squares parameter estimate of the height of the best-fitting synthetic response function for each condition at each voxel. Separate parameter estimates were initially obtained for the Shape and Pattern blocks, but these were averaged to produce a single first-order parameter estimate, alongside the single second-order (Match) parameter estimate, for subsequent analysis.

ROI analyses were performed using MarsBaR (Brett et al., 2002) to characterize the activation profiles of RLPFC, DLPFC, and IPL, a priori regions of interest based on previous studies of relational reasoning (Bunge et al., 2009; Crone et al., 2009). Each region was defined anatomically using templates from the automated anatomical labeling set (Tzourio-Mazoyer et al., 2002) and was further restricted to voxels within the anatomical region that demonstrated task-related activation across all participants (first-order  $>$  null or second-order  $>$  null, at  $p < 0.01$  uncorrected). The left and right RLPFC ROIs consisted of task-active voxels in the middle frontal gyrus, anterior to  $y = 48$  mm. The DLPFC ROIs consisted of task-active voxels in the middle frontal gyrus, anterior to  $y = 15$  and posterior to  $y = 42$ . The IPL ROIs consisted of all task-active voxels in the IPL (BA 40). For each ROI, the mean signal across all voxels in the ROI was submitted to GLM analysis, as described above, to produce an ROI parameter estimate for each condition for each participant. These ROI parameter estimates, along with other measures, were then submitted to statistical analysis in MATLAB (MathWorks).

Anatomical imaging data were preprocessed and analyzed with the FreeSurfer software package (version 4.4.0; <http://surfer.nmr.mgh.harvard.edu>). Details of these procedures have been described previously (Fischl and Dale, 2000). Briefly, the MRI data from each individual were processed with a series of automated and semi-automated steps, including (1) removal of non-brain tissue using a hybrid watershed/surface deformation procedure (Ségonne et al., 2004), (2) automated Talairach transformation, (3) segmentation of the subcortical white matter and deep gray matter volumetric structures (Fischl et al., 2002), (4) intensity normalization (Sled et al., 1998), (5) tessellation of the gray matter/white



**Figure 2.** Behavioral performance on the relational matching task. Scatter plots show accuracy (top) and response time (bottom) versus age for the first-order and second-order conditions.

matter boundary, (6) automated topology correction (Ségonne et al., 2007), and (7) surface deformation after intensity gradients to optimally place the gray/white and gray/CSF borders at the location where the greatest shift in intensity defines the transition to the other tissue class (Dale et al., 1999).

Cortical thickness was calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface (Fischl and Dale, 2000). Thickness measures were mapped onto the inflated surface of each participant's reconstructed brain (Fischl et al., 1999), enabling the visualization of data across the entire cortical surface, independent of cortical folding. Maps were smoothed with a 10 mm Gaussian kernel, and non-rigid high-dimensional spherical averaging was used to align cortical folding patterns across participants (Fischl et al., 1999), allowing for the creation of average surface models while accounting for cortical sulcal variability across participants. Statistical comparisons of surface maps and thickness measures were generated by computing a GLM of the effects of different variables (e.g., age) on thickness at each cortical surface location. Of the 85 participants included in the study, six had to be excluded from the structural analyses because of excessive motion during the high-resolution anatomical MRI scan, resulting in poor cortical reconstructions.

FreeSurfer ROIs (surface labels) were created by transforming SPM volume ROIs into the average subject space of FreeSurfer, which involved the following: (1) flipping the orientation of an ROI mask and the T1 template of SPM, (2) coregistering the SPM T1 template to the average FreeSurfer T1 image, (3) applying these coregistration parameters to the ROI mask, and then (4) manually tracing the outline of these volumes to create surface labels. FreeSurfer average-space surface ROIs were transformed into individual subject surface ROIs, average cortical thickness values were extracted from each region, for each subject, and these values were submitted to additional statistical analysis.

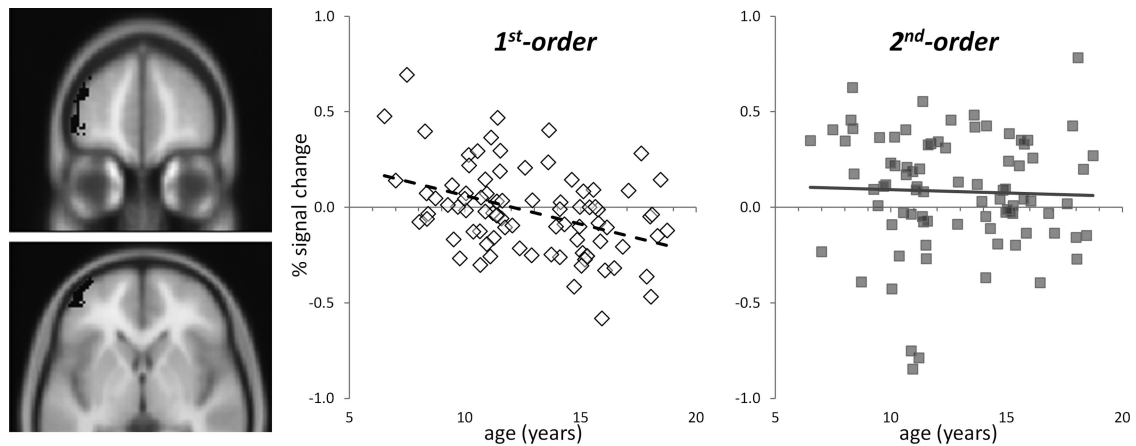
Our statistical analysis includes tests of specific theoretically driven hypotheses as well as some more exploratory tests. For the former, significance was assessed at an  $\alpha$  of 0.05. For the latter, we performed correction for multiple comparisons across each collection of tests, using the Bonferroni's correction, and adjusted our standard for significance accordingly. In all cases, the original (uncorrected)  $p$  values are reported.

## Results

### Behavioral data

Figure 2 shows average accuracy and response times on the relational matching task for three age ranges, including 7–10 year olds, 11–14 year olds, and 15–18 year olds. To assess performance, accuracy, and response time, values were submitted to a mixed ANOVA with condition (first-order, second-order) as a within-subject factor, gender as a between-subjects factor, and age as a covariate. We note that, although the graphs in Figure 2 show data for discrete age groups, we examine age here as a continuous covariate. There was a main effect of condition on accuracy, such that participants responded with greater accuracy on first-order trials (mean  $\pm$  SD,  $92.8 \pm 6.6\%$ ) than on second-order trials (mean  $\pm$  SD,  $90.0 \pm 9.6\%$ ;  $F = 4.3$ ,  $p < 0.05$ ). Similarly, there was a main effect of condition on response times,



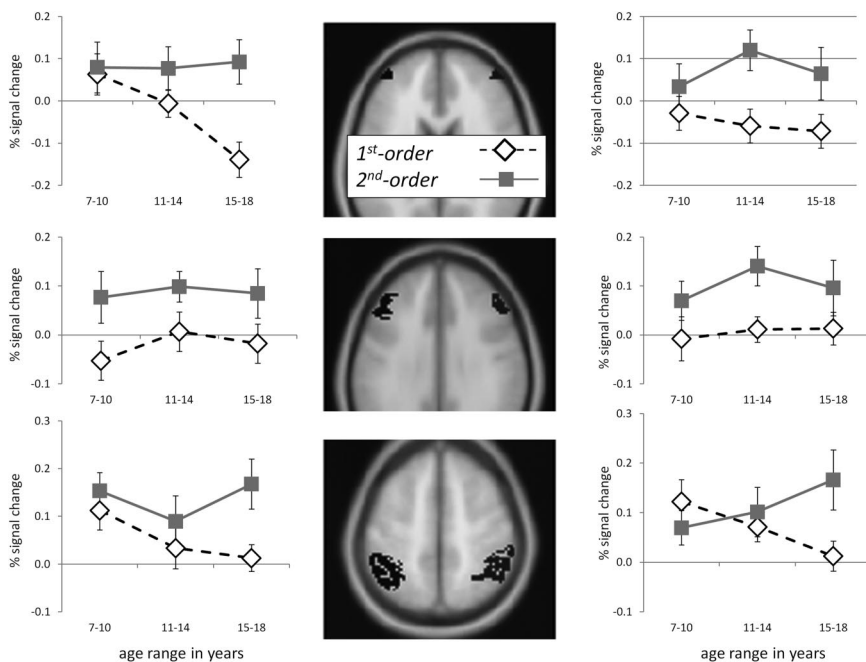


**Figure 3.** Functional activation versus age for left RLPFC. A significant decrease with age is observed for first-order (left plot) but not for second-order (right plot).

**Table 1.** Partial correlations involving fMRI activation and age for six functional ROIs

Region (size in cm <sup>3</sup> )	First-order versus age		Selectivity (second – first) versus age	
	Controlling for gender	Controlling for gender, RT, and accuracy	Controlling for gender	Controlling for gender, RT, and accuracy
L. RLPFC (2.05)	<b>−0.42 (<i>p</i> &lt; 0.001)</b>	<b>−0.32 (<i>p</i> = 0.002)</b>	<b>0.24 (<i>p</i> = 0.02)</b>	<b>0.23 (<i>p</i> = 0.02)</b>
R. RLPFC (1.99)	—	—	—	—
L. DLPFC (2.18)	—	—	—	—
R. DLPFC (2.20)	—	—	—	—
L. IPL (8.26)	−0.23 ( <i>p</i> = 0.02)	−0.15 ( <i>p</i> = 0.09)	—	—
R. IPL (4.75)	−0.23 ( <i>p</i> = 0.02)	−0.25 ( <i>p</i> = 0.02)	0.27 ( <i>p</i> = 0.006)	0.25 ( <i>p</i> = 0.01)

All significant and marginally significant correlations are listed. Bold type indicates results that survived correction for multiple comparisons (see Results). — indicates not significant (*p* > 0.1). L, Left; R, right; RT, reaction time.



**Figure 4.** Functional activation by age group for six ROIs, including left and right RLPFC, DLPFC, and IPL. Data are plotted for younger children (ages 7–10 years), middle-aged children (ages 11–14 years), and older children (ages 15–18 years).

such that participants responded more rapidly to first-order trials (mean ± SD, 1.96 ± 0.53 s) than to second-order trials (mean ± SD, 2.35 ± 0.56 s; *F* = 164.2, *p* < 0.001). There were no significant interactions between condition and either age or gender, for accuracy or response time (all *p* > 0.1). Overall performance did improve with age, as indicated by positive correlations between accuracy and age for both conditions (first-order, *r* = 0.40, *p* < 0.001; second-

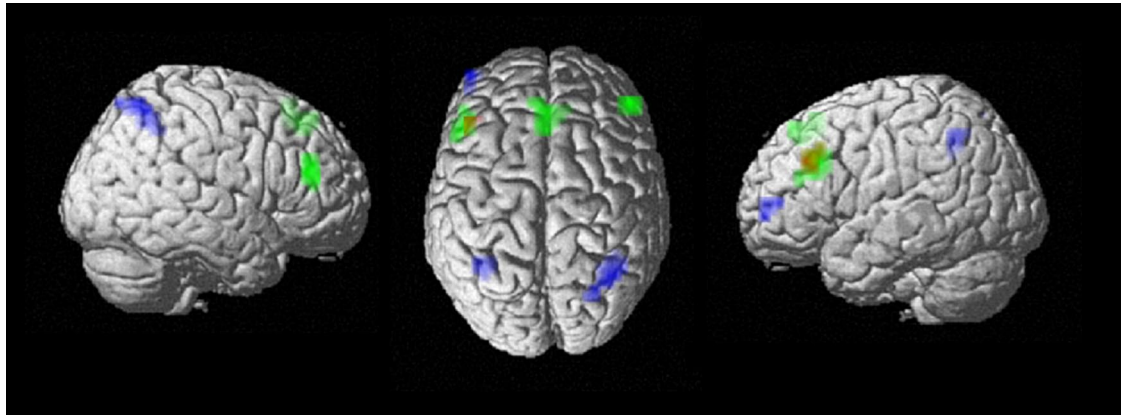
order, *r* = 0.42, *p* < 0.001; Fig. 2A) and by negative correlations between response times and age (first-order, *r* = −0.54, *p* < 0.001; second-order, *r* = −0.42, *p* < 0.001).

To relate our laboratory relational matching task to standardized measures of fluid reasoning, we computed an FR for each participant, based on four standardized reasoning tests (see Materials and Methods).

As expected, the FR was correlated with age (*r* = 0.35, *p* < 0.001). Even when controlling for age, FR was positively correlated with accuracy on the relational matching task. This effect was marginal for first-order trials (*r* = 0.20, *p* = 0.09) but highly significant for second-order trials (*r* = 0.34, *p* = 0.002).

**Does RLPFC demonstrate increasing specialization for relational integration with development?**

To test our primary hypothesis concerning the development of functional selectivity in RLPFC, we conducted a detailed examination of functional activation patterns in left and right RLPFC ROIs. The RLPFC ROIs were constructed by taking all voxels within a left or right RLPFC anatomical template (middle frontal gyrus, anterior to *y* = 48) that demonstrated task-related activation (first-order > null or second-order > null) across all participants (at *p* < 0.01 uncorrected). For each ROI, we obtained the partial correlations between functional activation (first-order > null, sec-

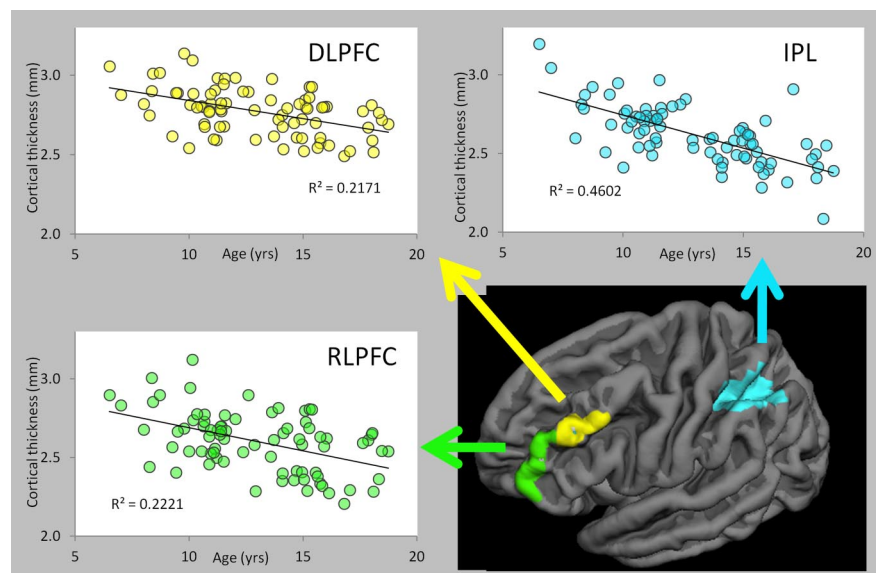


**Figure 5.** Whole-brain activation for the second-order > first-order contrast, for younger children (red), middle children (green), and older children (blue). Areas of overlap between the younger and middle children are shown in yellow.

ond-order > null) or functional selectivity (second-order > first-order) and age, controlling for gender. In addition, to ensure that results were not an artifact of differences in performance between the two conditions, we also obtained partial correlations after controlling for response time and accuracy. Specifically, for the correlation with first-order activation, we controlled for first-order response time and first-order accuracy. For the correlation with second-order activation, we controlled for second-order response time and second-order accuracy. For the correlation with selectivity, we controlled for the second-order – first-order response time and accuracy differences. For our primary hypothesis, concerning specialization of RLPFC, we assessed significance at an  $\alpha$  of 0.05. In accordance with our hypothesis, left RLPFC exhibited a significant age-related increase in the second-order – first-order contrast value. In contrast, right RLPFC demonstrated no significant age-related changes. Additional assessments of regional activation versus age were conducted with an  $\alpha$  of 0.004, taking into account 12 separate tests. In left RLPFC, we observed a highly significant decrease in activation for first-order trials with age but no change in activation for second-order trials (Fig. 3; for statistics, see Table 1). Thus, we observed a robust age-related change in functional selectivity in left but not right RLPFC.

#### Do other core components of the relational reasoning network—DLPFC and IPL—demonstrate developmental changes in function?

DLPFC and IPL (BA 40) were examined in a similar manner to RLPFC. Neither left nor right DLPFC demonstrated any change in activation as a function of age (Table 1). However, both left and right IPL demonstrated a developmental pattern that was similar to RLPFC: decreasing engagement for first-order, resulting in a pattern of age-related increase in the second-order – first-order contrast value. IPL did not exhibit as marked a decrease in first-order activation with age as did left RLPFC, and the IPL result did not survive correction for multiple comparisons.



**Figure 6.** Cortical thinning in RLPFC, DLPFC, and IPL. Results are shown for the left side only but were similar on the right.

To compare regions, we submitted data from all six ROIs (Fig. 4) to a mixed ANOVA with region (RLPFC, DLPFC, or IPL), side (left or right), and condition (first-order or second-order) as within-subject factors and age as a continuous covariate. We observed a significant four-way interaction ( $F_{(2,81)} = 5.7, p = 0.005$ ). This effect was primarily driven by a significant interaction between region, condition, and age in the left hemisphere ( $F_{(2,81)} = 6.6, p = 0.002$ ), in which the condition  $\times$  age interaction in DLPFC was significantly different from that observed in RLPFC ( $F_{(1,83)} = 9.0, p = 0.004$ ). Specifically, whereas first-order activation was flat or even increasing with age in DLPFC, it decreased with age in RLPFC.

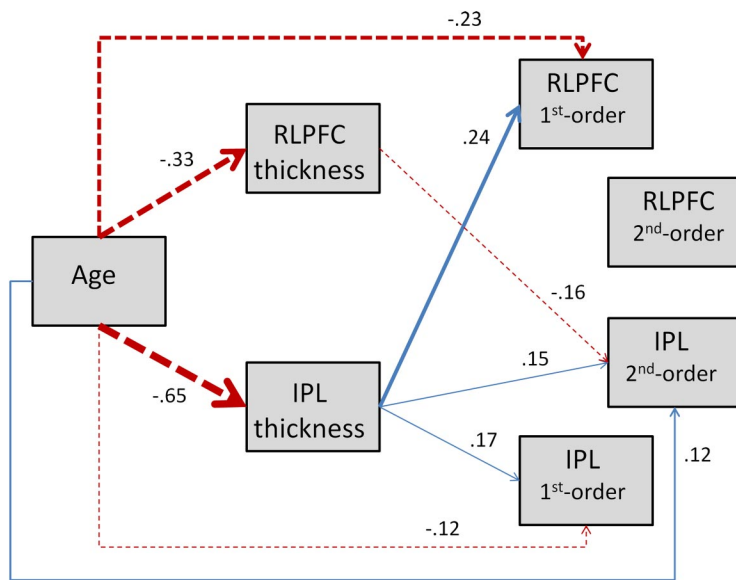
#### Whole-brain analyses

To probe more generally for regions involved in second-order relational processing and for regions undergoing developmental change related to this, we performed several different whole-brain analyses. First, we examined the second-order > first-order contrast across all participants. This whole-brain contrast yielded activation in bilateral DLPFC, right RLPFC, and left IPL, as well as in dorsomedial PFC. Next, we examined this contrast separately for each of three age groups (Fig. 5). For older children

**Table 2. Sample statistics for variables examined via structured equation modeling**

	Age	RLPFC thickness	IPL thickness	RLPFC first-order	RLPFC second-order	IPL first-order	IPL second-order
Mean (SE)	12.9 (0.3)	2.9 (0.03)	2.6 (0.02)	−0.02 (0.03)	0.08 (0.03)	0.05 (0.02)	0.14 (0.03)
Covariances (top right)							
Age		−0.232	−0.396	−0.280	−0.033	−0.155	−0.013
RLPFC thickness	−0.313		0.015	0.005	0.001	0.008	−0.011
IPL thickness	−0.664	0.314		0.018	0.000	0.010	0.005
RLPFC first-order	−0.410	0.087	0.406		0.008	0.022	0.005
RLPFC second-order	−0.037	0.016	−0.008	0.117		0.003	0.025
IPL first-order	−0.246	0.159	0.252	0.471	0.053		0.015
IPL second-order	−0.016	−0.167	0.100	0.084	0.310	0.273	
Correlations (bottom left)							

Note that entries in the bottom diagonal are correlations, and entries in the top diagonal are covariances.



**Figure 7.** Structured equation model relating age, cortical thickness, and fMRI activation in left RLPFC and IPL. Only connections that yielded standardized parameter estimates  $>0.1$  are depicted, with line thickness corresponding approximately to the magnitude of the parameter estimate and color indicating the direction of the relationship (solid/green for positive, dashed/red for negative).

(ages 15–18 years), this contrast revealed significant activation in left RLPFC, bilateral IPL, and dorsomedial PFC but not in DLPFC. For the middle group (ages 11–14 years), as well as for the younger group (ages 7–10 years), significant activation was observed in bilateral DLPFC and dorsomedial PFC but not in RLPFC or parietal cortex.

**Is cortical thinning the mechanism of increasing functional specialization in RLPFC and IPL?**

Having found evidence that RLPFC and IPL undergo a process of functional specialization for relational reasoning during development, we sought to determine whether or not cortical thinning in these regions contributes to the development of relational reasoning selectivity. To this end, we extracted cortical thickness values from the left and right RLPFC and IPL ROIs (and also from left and right DLPFC). As expected, we observed significant cortical thinning with age in each of these ROIs (Fig. 6).

Our primary question was whether or not cortical thinning could be the primary mechanism that drives functional specialization. To test this, we examined RLPFC and IPL cortical thickness and functional activation values in the context of structural equation modeling. We created a model that included the following variables: age, RLPFC cortical thickness, IPL cortical

thickness, RLPFC first-order activation, RLPFC second-order activation, IPL first-order activation, and IPL second-order activation. The full model included a link from age to every other variable. In effect, these links represent direct age-related mechanisms (i.e., maturation) that are not otherwise specified in the model. We also included a link from each cortical thickness variable to each activation variable. Thus, in this full model, we allowed for the possibility that RLPFC thickness might influence RLPFC activation and that IPL thickness might influence IPL activation. In addition, because there is considerable evidence of tight functional coupling between RLPFC and IPL (Vincent et al., 2008; Boorman et al., 2009; Wendelken et al., 2011), and because structural changes in one region that make it function more efficiently could translate to improved efficiency in the functioning of another region with which it interacts, we also allowed for the possibility that IPL thickness might affect RLPFC activation (and vice versa).

Sample statistics are provided in Table 2, and standardized parameter estimates are shown in Figure 7 and Table 3. Only links for which standardized parameter estimates exceeded 0.1 are included in the figure. Notably, RLPFC thickness had no impact on activation in RLPFC. However, IPL thickness did contribute to first-order activation in RLPFC; specifically, decreasing cortical thickness in IPL led to decreasing first-order activation in RLPFC. Cortical thinning in IPL had a similar impact on both first- and second-order IPL activation.

Starting with the full model described above, we created alternative models by removing select connections and compared these alternative models to the full model. Model comparison results are presented in Table 4. Because we sought to explain the developmental change associated with first-order activation in left RLPFC, we probed the results of eliminating each of the three inputs to this variable: age, RLPFC thickness, and IPL thickness. We observed that eliminating either age or IPL thickness produced a model that was marginally worse than the full model ( $p = 0.1$ ), whereas removing RLPFC thickness had no effect. Removing both age and IPL thickness produced a model that was much worse than the full model ( $p = 0.001$ ). These results support the idea that cortical thinning in IPL, but not in RLPFC, contributes to changes in functional activation patterns.



**Table 3. Modeling results for the full and age-excluded structured equation models**

Parameter	Est.	SE	<i>p</i> value	Std. Est.
<b>Full model</b>				
Age → RLPFC thickness	−0.027	0.009	0.005	−0.334
Age → IPL thickness	−0.041	0.005	<0.001	−0.651
Age → RLPFC first-order	−0.016	0.01	0.05	−0.227
RLPFC thickness → RLPFC first-order	−0.048	0.097	0.46	−0.055
IPL thickness → RLPFC first-order	0.266	0.151	0.05	0.239
Age → RLPFC second-order	−0.001	0.015	0.62	−0.011
RLPFC thickness → RLPFC second-order	−0.009	0.145	0.92	−0.008
IPL thickness → RLPFC second-order	−0.088	0.226	0.69	−0.058
Age → IPL first-order	−0.008	0.010	0.37	−0.117
RLPFC thickness → IPL first-order	0.035	0.098	0.49	0.042
IPL thickness → IPL first-order	0.185	0.153	0.35	0.174
Age → IPL second-order	0.010	0.013	0.69	0.120
RLPFC thickness → IPL second-order	−0.173	0.125	0.06	−0.163
IPL thickness → IPL second-order	0.197	0.195	0.16	0.148
RLPFC first-order ~ RLPFC second-order	0.004	0.006	0.25	0.064
IPL first-order ~ IPL second-order	0.018	0.006	0.01	0.339
RLPFC first-order ~ IPL first-order	0.017	0.005	0.01	0.382
RLPFC second-order ~ IPL second-order	0.027	0.009	0.01	0.358
<b>Age-excluded model</b>				
RLPFC thickness → RLPFC first-order	−0.053	0.105	0.61	−0.058
IPL thickness → RLPFC first-order	0.458	0.130	<0.001	0.394
RLPFC thickness → RLPFC second-order	0.020	0.150	0.89	0.016
IPL thickness → RLPFC second-order	−0.052	0.189	0.78	−0.034
RLPFC thickness → IPL first-order	0.071	0.097	0.46	0.085
IPL thickness → IPL first-order	0.241	0.123	0.05	0.225
RLPFC thickness → IPL second-order	−0.256	0.134	0.06	−0.229
IPL thickness → IPL second-order	0.193	0.170	0.26	0.136
RLPFC first-order ~ RLPFC second-order	0.009	0.007	0.21	0.139
IPL first-order ~ IPL second-order	0.015	0.006	0.01	0.292
RLPFC first-order ~ IPL first-order	0.017	0.005	<0.001	0.426
RLPFC second-order ~ IPL second-order	0.026	0.009	0.005	0.324

Est., Parameter estimate; Std. Est., standardized parameter estimate. Estimated parameters include regression coefficients (→) and covariances (−).

In addition to the full model, we also considered a model in which all connections from age to activation variables were removed (the age-excluded model). That this model was not significantly worse than the full model suggests that cortical thickness changes are a good proxy for age. As with the full model, we selectively eliminated effects on RLPFC first-order activation from the age-excluded model. Here again, we observed that the connection from RLPFC thickness could be removed without cost, whereas removing the connection from IPL thickness produced a model fit that was significantly worse ( $p = 0.001$ ).

Thus, regardless of whether age is explicitly included in the model, we observe a relationship between cortical thinning in IPL (but not RLPFC) and functional specialization in RLPFC.

## Discussion

The present study combined functional and anatomical imaging methods to characterize the neural bases of developmental improvements in relational reasoning. In our previous study that compared adults and children performing the Raven's progressive matrices task, we observed that, although children performed similarly to adults on first-order relational problems, they performed much worse than adults on second-order relational problems (Crone et al., 2009). Based on this previous finding, we expected in the current study to see large age-related improvements in performance for second-order reasoning and smaller improvements for first-order reasoning. In fact, we observed similar age-related performance improvements for both types of relational reasoning. Ceiling effects may limit the usefulness

of the accuracy scores in this case; however, the same pattern was observed for response times. A recent study that used a similar relational matching task reported a decrease in performance during middle childhood for second-order relational processing, such that children aged 9.8–11.4 years performed better than children aged 11.5–17.7 years (Dumontheil et al., 2010); we did not observe this pattern in our data.

### Emergence of functional selectivity in PFC

With the fMRI results, we were able to confirm our first hypothesis. Specifically, the present results demonstrate that RLPFC undergoes a developmental change in its pattern of engagement in relational reasoning. In younger children, RLPFC is equally engaged by first- and second-order relational processing. However, decreasing engagement of RLPFC for first-order relational processing with increasing age leads to specialization of the region for second-order processing in older teens and, as previous studies have shown, in adults. Previous studies have demonstrated similar patterns. In addition to our previous developmental study, which demonstrated that children engage RLPFC to a greater extent than adults on first-order matrix reasoning problems, two other studies are particularly relevant. First, in a sample of 16 children (ages 8–19 years) performing a first-order reasoning task, Eslinger et al. (2009) demonstrated decreasing engagement of lateral PFC, including RLPFC, with increasing age. Second, in a study that examined 37 adolescents and young adults, Dumontheil et al. (2010) observed a marginal increase in RLPFC specificity (second-order > first-order) from younger adolescents to older adolescents. Our results complement and extend the results of these previous studies.

The present results are consistent with our previous findings regarding differentiation of left and right RLPFC. Previously, in adults, we have observed that left RLPFC demonstrates strong specialization for second-order relational processing, in the context of the relational matching task, right RLPFC demonstrates weaker specialization attributable to relatively stronger engagement on the first-order task (Bunge et al., 2009). In the current study, left RLPFC underwent a dramatic change toward specialization with age; in right RLPFC, this pattern was relatively weak.

The functional profile of DLPFC, unlike that of RLPFC, did not significantly change with age. In fact, in younger children, DLPFC was more specifically engaged than RLPFC for second-order relational processing. We have suggested previously that, in adults, DLPFC is involved in the processing of first-order relations and that its activity is modulated by the number or individual relations processed or by the difficulty of the relational processing task (Wendelken et al., 2009). This possibility is consistent with the current results. Moreover, the lack of age-related change suggests that this functionality is well established even in younger children.

### Emergence of functional selectivity in parietal cortex

In adults, IPL in the vicinity of the intraparietal sulcus is the region that, after RLPFC, is most typically activated in contrasts between second- and first-order relational processing (Crone et al., 2009; Wendelken et al., 2011). Moreover, this is the region that in many studies demonstrates the strongest functional connectivity with RLPFC (Vincent et al., 2008; Boorman et al., 2009). However, unlike RLPFC, IPL also tends to be engaged by first-order relational tasks (Wendelken and Bunge, 2010). Thus, it was somewhat surprising that IPL, like RLPFC, demonstrated an age-related decrease in activation for the first-order task in the present study. In fact, this contrasts with our previous developmental

**Table 4. Model comparison results, with comparisons to the full model in the top part of the table and comparisons to the age-excluded model in the bottom part of the table (see Results)**

Model	$\chi^2$	df	$\chi^2$ (difference)	df (difference)	p value
Full	2.3	1			
Full, excluding age → RLPFC first-order	4.9	2	2.6	1	0.1
Full, excluding RLPFC thickness → RLPFC first-order	2.5	2	0.2	1	NS
Full, excluding IPL thickness → RLPFC first-order	5.2	2	2.9	1	0.1
Full, excluding [age and RLPFC thickness] → RLPFC first-order	5.0	3	2.7	2	NS
Full, excluding [age and IPL thickness] → RLPFC first-order	15.6	3	13.3	2	0.001
Age-excluded	5.7	5			
Age-excluded, excluding RLPFC thickness → RLPFC first-order	5.8	6	0.1	1	NS
Age-excluded, excluding IPL thickness → RLPFC first-order	16.3	6	10.6	1	0.001

study using Raven's progressive matrices task, which showed greater IPL activation in adults than in children for the first-order condition (Crone et al., 2009). This result also contrasts with a previous developmental study of first-order reasoning that reported increasing parietal activation as a function of age (Eslinger et al., 2009). Of course, it is possible that these differences could be explained by the substantial differences in task demands across the three studies. In particular, it is notable that, in both of the previous studies that had demonstrated increased parietal activation with age for first-order processing, the first-order relational condition is contrasted with a zero-order (i.e., non-relational) condition, whereas in the present study, there was no non-relational condition. We speculate that IPL would demonstrate an even more pronounced developmental decrease in activation associated with non-relational processing relative to that observed here for first-order processing.

### The role of cortical thinning

Numerous structural brain imaging studies have demonstrated that gray matter thinning occurs in dorsal parietal lobes during childhood and then spreads anteriorly to dorsal frontal regions during the adolescent and post-adolescent years (Gogtay et al., 2004; Sowell et al., 2004; Shaw et al., 2008). Early postmortem work has been leveraged to suggest that the changes in cortical composition observed with MRI are attributable to reductions in synaptic density (Huttenlocher, 1979; Huttenlocher et al., 1982) and increases in axonal myelination (Yakovlev and Lecours, 1967). Concomitant reductions in synaptic density and increases in axonal myelination are the hallmarks of experience-based neural plasticity. This process involves the initial overproduction of neurons and synaptic connections during infancy and early childhood followed by activity-dependent fine-tuning of neural activity via synaptic pruning that continues well into adolescence (Fuster, 2002; Tsujimoto, 2008).

It is tempting to speculate that changes in functional activation observed during cognitive development are related to maturational changes in brain structure (Casey et al., 2005; Ferrer et al., 2009), but little evidence linking developmental changes in brain structure and function exists (Lu et al., 2009). One attempt to address this question used fMRI data and a modeling approach rather than actual structural data and concluded that increases in frontoparietal synaptic connectivity are a major structural driver of changes in working memory-related brain activation (Edin et al., 2007). We hypothesized that, for a region that demonstrates a pattern of increasing functional specialization with age, this change might be driven by local synaptic pruning, as reflected in decreasing cortical thickness of the region in question.

Cortical thinning in RLPFC has been well characterized previously (O'Donnell et al., 2005; Tamnes et al., 2010), and, in the

present study, both cortical thickness and functional activation were highly correlated with age. However, contrary to our expectation, cortical thinning within RLPFC was not an important factor driving functional specialization in this region. Maturational changes in local brain structure may well explain the increasing functional specialization of RLPFC, but the nature of these changes have yet to be determined, and they are not reflected in local cortical thickness.

Changes in cortical thickness in IPL do appear to have some effect on changes in functional activation. With regard to functional activation within IPL, this effect was similar for first-order and second-order reasoning, suggesting that cortical thinning in IPL is associated not with greater selectivity in this region but rather with overall reduction in the level of functional activation. More intriguing, perhaps, is the fact that IPL thickness also affected functional activation in left RLPFC. In this case, the effect of cortical thinning was limited to first-order reasoning. Thus, cortical thinning in IPL appears to affect functional selectivity in left RLPFC.

We had hypothesized that local changes in cortical thickness would lead to local changes in functional selectivity. It is, at first look, somewhat surprising that the strongest effect of cortical thinning on functional selectivity is nonlocal, specifically, that cortical thinning in IPL affects functional selectivity in RLPFC.

However, there is considerable evidence for a close interaction between these two regions. In relational reasoning tasks, IPL, in the vicinity of the intraparietal sulcus, is the region that is most consistently activated alongside RLPFC (Crone et al., 2009; Wendelken et al., 2011). RLPFC and IPL show consistently strong functional connectivity in humans (Vincent et al., 2008). Based on anatomical tracing studies in primates, it has long been assumed that humans lack direct anatomical projections between anterior prefrontal and parietal cortices. However, a recent demonstration that macaques lack the pattern of resting-state connectivity between these regions that is observed in humans (Mars et al., 2011) suggests that this previous assumption might be incorrect. If RLPFC depends on outputs from IPL, then improvements in the overall efficiency of IPL (which might be a consequence of cortical thinning) can very plausibly lead to improved task specificity (i.e., functional selectivity) in RLPFC.

Although changes in IPL cortical thickness do appear to affect functional activation, in RLPFC as well as in IPL, this is by no means the only developmental mechanism at work. A key goal for future research is to assess the role that strengthening of structural and/or functional connectivity between RLPFC and IPL play in development of relational reasoning.



## References

- Boorman ED, Behrens TE, Woolrich MW, Rushworth MF (2009) How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62:733–743.
- Brett M, Anton J, Valabregue R, Poline J (2002) Region of interest analysis using an SPM toolbox. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan, June 2–6.
- Bunge SA, Holskog EH, Wendelken C (2009) Left, but not right, rostralateral prefrontal cortex meets a stringent test of the relational integration hypothesis. *Neuroimage* 46:338–342.
- Casey BJ, Tottenham N, Liston C, Durston S (2005) Imaging the developing brain: what have we learned about cognitive development? *Trends Cogn Sci* 9:104–110.
- Christoff K, Ream JM, Geddes LP, Gabrieli JD (2003) Evaluating self-generated information: anterior prefrontal contributions to human cognition. *Behav Neurosci* 117:1161–1168.
- Crone EA, Wendelken C, van Leijenhorst L, Honomichl RD, Christoff K, Bunge SA (2009) Neurocognitive development of relational reasoning. *Dev Sci* 12:55–66.
- Dale AM, Fischl B, Sereno MI (1999) Cortical Surface-Based Analysis\* 1: I. Segmentation and Surface Reconstruction. *Neuroimage* 9:179–194.
- Dumontheil I, Houlton R, Christoff K, Blakemore SJ (2010) Development of relational reasoning during adolescence. *Dev Sci* 13:F15–F24.
- Edin F, Macoveanu J, Olesen P, Tegner J, Klingberg T (2007) Stronger synaptic connectivity as a mechanism behind development of working memory-related brain activity during childhood. *J Cogn Neurosci* 19:750–760.
- Eslinger PJ, Blair C, Wang J, Lipovsky B, Realmuto J, Baker D, Thorne S, Gamson D, Zimmerman E, Rohrer L, Yang QX (2009) Developmental shifts in fMRI activations during visuospatial relational reasoning. *Brain Cogn* 69:1–10.
- Ferrer E, O'Hare E, Bunge S (2009) Fluid reasoning and the developing brain. *Frontiers in Neuroscience* 3:46.
- Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A* 97:11050–11055.
- Fischl B, Sereno MI, Dale AM (1999) Cortical Surface-Based Analysis\* 1: II: Inflation, Flattening, and a Surface-Based Coordinate System. *Neuroimage* 9:195–207.
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM (2002) Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron* 33:341–355.
- Fuster JM (2002) Frontal lobe and cognitive development. *J Neurocytol* 31:373–385.
- Giedd JN, Rapoport JL (2010) Structural MRI of pediatric brain development: what have we learned and where are we going? *Neuron* 67:728–734.
- Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, Nugent TF 3rd, Herman DH, Clasen LS, Toga AW, Rapoport JL, Thompson PM (2004) Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A* 101:8174–8179.
- Halford G (1992) Analogical reasoning and conceptual complexity in cognitive development. *Hum Dev* 35:193–217.
- Huttenlocher PR (1979) Synaptic density in human frontal cortex - developmental changes and effects of aging. *Brain Res* 163:195–205.
- Huttenlocher PR, De Courten C, Garey LJ, and Van der Loos H (1982) Synaptic development in human cerebral cortex. *Int J Neurol* 16–17: 144–154.
- Krawczyk DC, Michelle McClelland M, Donovan CM (2011) A hierarchy for relational reasoning in the prefrontal cortex. *Cortex* 47:588–597.
- Lu LH, Dapretto M, O'Hare ED, Kan E, McCourt ST, Thompson PM, Toga AW, Bookheimer SY, Sowell ER (2009) Relationships between brain activation and brain structure in normally developing children. *Cereb Cortex* 19:2595–2604.
- Mars RB, Jbabdi S, Sallet J, O'Reilly JX, Crosson PL, Olivier E, Noonan MP, Bergmann C, Mitchell AS, Baxter MG, Behrens TE, Johansen-Berg H, Tomassini V, Miller KL, Rushworth MF (2011) Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *J Neurosci* 31:4087–4100.
- McArdle JJ, Ferrer-Caja E, Hamagami F, Woodcock RW (2002) Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Dev Psychol* 38:115–142.
- O'Donnell S, Noseworthy MD, Levine B, Dennis M (2005) Cortical thickness of the frontopolar area in typically developing children and adolescents. *Neuroimage* 24:948–954.
- Penn DC, Holyoak KJ, Povinelli DJ (2008) Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav Brain Sci* 31:109–130, discussion 130–178.
- Richland LE, Morrison RG, Holyoak KJ (2006) Children's development of analogical reasoning: Insights from scene analogy problems. *J Exp Child Psychol* 94:249–273.
- Robin N, and Holyoak KJ (1995) Relational Complexity and the Functions of Prefrontal Cortex. pp. 987–999. Cambridge, MA: MIT Press.
- Ségonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, Fischl B (2004) A hybrid approach to the skull stripping problem in MRI. *Neuroimage* 22:1060–1075.
- Ségonne F, Pacheco J, Fischl B (2007) Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Trans Med Imaging* 26:518–529.
- Shaw P, Kabani NJ, Lerch JP, Eckstrand K, Lenroot R, Gogtay N, Greenstein D, Clasen L, Evans A, Rapoport JL, Giedd JN, Wise SP (2008) Neurodevelopmental trajectories of the human cerebral cortex. *J Neurosci* 28:3586–3594.
- Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 17:87–97.
- Smith R, Keramatian K, Christoff K (2007) Localizing the rostralateral prefrontal cortex at the individual level. *Neuroimage* 36:1387–1396.
- Sowell ER, Thompson PM, Leonard CM, Welcome SE, Kan E, Toga AW (2004) Longitudinal mapping of cortical thickness and brain growth in normal children. *J Neurosci* 24:8223–8231.
- Sternberg RJ, Rifkin B (1979) The development of analogical reasoning processes. *J Exp Child Psychol* 27:195–232.
- Tamnes CK, Ostby Y, Fjell AM, Westlye LT, Due-Tønnessen P, Walhovd KB (2010) Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cerebral Cortex* 20:534–548.
- Tsujimoto S (2008) The prefrontal cortex: functional neural development during early childhood. *Neuroscientist* 14:345–358.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008) Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J Neurophysiol* 100:3328–3342.
- Wechsler D (1974) Manual for the Wechsler intelligence scale for children—Revised. New York: Psychological Corporation.
- Wendelken C, Bunge SA (2010) Transitive inference: distinct contributions of rostralateral prefrontal cortex and the hippocampus. *J Cogn Neurosci* 22:837–847.
- Wendelken C, Chung D, Bunge S (2011) Rostrolateral prefrontal cortex: domain-general or domain-sensitive? *Hum Brain Mapp*. Advance online publication. doi:10.1002/hbm.21336.
- Woodcock RW, Johnson MB (1990) Woodcock & Johnson psychoeducational battery-Revised, Examiner's manual. Chicago: Riverside.
- Wright SB, Matlen BJ, Baym CL, Ferrer E, Bunge SA (2007) Neural correlates of fluid reasoning in children and adults. *Front Hum Neurosci* 1:8.
- Yakovlev P, Lecours A (1967) The myelogenetic cycles of regional maturation of the brain. Regional development of the brain in early life. Oxford: Blackwell.