◆ Human Brain Mapping 36:996–1009 (2015) ◆

Distributed Neural Representations of Logical Arguments in School-Age Children

Romain Mathieu,[1] James R. Booth,[2,3]* and Jérôme Prado[1,3]*

[1] *Laboratoire Langage, Cerveau et Cognition (L2C2), Centre National de la Recherche Scientifique (CNRS) & Université de Lyon, 69675 Bron, France*
[2] *Department of Communication Sciences and Disorders, University of Texas, Texas*
[3] *Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois*

Abstract: Children's understanding of linear-order (e.g., *Dan is taller than Lisa, Lisa is taller than Jess*) and set-inclusion (i.e., *All tulips are flowers, All flowers are plants*) relationships is critical for the acquisition of deductive reasoning, that is, the ability to reach logically valid conclusions from given premises. Behavioral and neuroimaging studies in adults suggest processing differences between these relations: While arguments that involve linear-orders may be preferentially associated with spatial processing, arguments that involve set-inclusions may be preferentially associated with verbal processing. In this study, we used functional magnetic resonance imaging to investigate whether these processing differences appear during the period of elementary school in development. Consistent with previous studies in adults, we found that arguments that involve linear-order and set-inclusion relationships preferentially involve spatial and verbal brain mechanisms (respectively) in school-age children (9–14 year olds). Because this neural sensitivity was not related to age, it likely emerges before the period of elementary education. However, the period of elementary education might play an important role in shaping the neural processing of logical reasoning, as indicated by developmental changes in frontal and parietal regions that were dependent on the type of relation. *Hum Brain Mapp 36:996–1009, 2015.*

Key words: deductive reasoning; fMRI; development; set-inclusion relations; linear relations

INTRODUCTION

Deductive reasoning is the ability to reach conclusions that necessarily follow from given premises. For instance, given the linear-order relations *Dan is taller than Lisa* and *Lisa is taller than Jess*, one can deduce that *Dan is taller than Jess*. Similarly, the relations of set-inclusion *All Tulips are Flowers* and

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: National Institute of Child Health and Human Development; Contract grant number: HD069781 (to J.P. and J.R.B.); Contract grant sponsor: Fondation de France; Contract grant number: 2012–00033701 (to J.P.).

Conflict of Interest: The authors declare no competing financial interests.

*Correspondence to: James Booth, Department of Communication Sciences and Disorders, University of Texas, 2504A Whitis Ave, Austin, TX 78712. E-mail: j-booth@austin.utexas.edu or Jérôme Prado, Laboratoire Langage, Cerveau et Cognition, Centre National de la Recherche Scientifique, 67 Boulevard Pinel, 69675 Bron cedex, France. E-mail: jprado@isc.cnrs.fr

Received for publication 23 May 2014; Revised 18 September 2014; Accepted 21 October 2014.

DOI: 10.1002/hbm.22681

Published online 30 October 2014 in Wiley Online Library (wileyonlinelibrary.com).

All Flowers are Plants enable the conclusion *All Tulips are Plants*. Deductive reasoning is a central component of human thinking [Stanovich and West, 2000]. It is also a scaffold for the acquisition of many important scientific concepts in children, such as ordinality, categorization, measurements, and mathematical proof [Bryant and Kopytynska, 1976; Michal and Ruhama, 2008; Nunes et al., 2007]. Impairments in deductive reasoning may thus be detrimental to academic growth and have been recently associated with developmental disabilities affecting both language [Katsos et al., 2011] and mathematical learning [Morsanyi et al., 2013]. Therefore, increasing our understanding of the psychological and neurological mechanisms contributing to deductive reasoning acquisition in children is important from both theoretical and clinical perspectives.

To our knowledge, all that is known about the neural bases of deductive reasoning comes from studies on adult participants. These studies indicate that mature deductive reasoning relies on an extensive brain system encompassing frontal, temporal, and parietal regions [Goel, 2007]. Activity in this system, however, appears to be heavily dependent on task-specific factors [Goel, 2007; Prado et al., 2011a]. For example, studies have shown that the neural correlates of deductive reasoning in adults are modulated by the semantic content of the premises [e.g., abstract vs. concrete; Goel et al., 2000, 2004], the absence/presence of conflicting information [e.g., consistent or inconsistent with prior beliefs; Goel and Dolan, 2003], the amount of information available to evaluate a conclusion [e.g., complete vs. incomplete information; Goel et al., 2009], and the argument complexity [modus ponens vs. modus tollens; Monti et al., 2007; Noveck et al., 2004]. Recently, it has been shown that the mature brain system for reasoning is also sensitive to the logical form of the premises [Prado et al., 2010, 2013; Reverberi et al., 2010]. A quantitative meta-analysis of the neuroimaging literature demonstrated a differential reliance on frontal and parietal regions for reasoning with linear-order versus set-inclusion relations [Prado et al., 2011a]: while spatial regions of the parietal cortex are more consistently associated with arguments containing linear-order than set-inclusion relations, verbal regions of the left frontal cortex are more consistently associated with arguments containing set-inclusion than linear-order relations [see also, Prado et al., 2013]. Critically, this neural dissociation is consistent with a large body of behavioral studies indicating that adults might often adopt different strategies when encoding and manipulating different types of relations: while integrating linear-order relations are likely to be associated with the construction of a unified mental representation in spatial working memory [Goodwin and Johnson-Laird, 2005, 2008; Potts, 1972, 1974; Prado et al., 2008], integrating set-inclusion relations might more frequently be associated with the step-by-step coordination of sentences in verbal working memory [Barrouillet, 1996; Carlson et al., 1992; Favrel and Barrouillet, 2000; Griggs, 1976; Griggs and Osterman, 1980; Newstead et al., 1985].

To date, it remains unclear how and when brain regions become sensitive to the type of logical relation during deductive reasoning. On the one hand, several studies have demonstrated that logical abilities can be observed in children well before they enter elementary school (around the ages of 4 or 5) [Bryant and Trabasso, 1971; Dias and Harris, 1988, 1990; Hawkins et al., 1984; Noveck, 2001]. This is in line with the idea that the human mind is equipped with a "mental logic" [Braine and O'Brien, 1998; Rips, 1994], and that the ability to make inferences is either innate [Macnamara, 1986] or at least develops very early [Moshman, 2004]. Thus, it is possible that the brain system for reasoning might become sensitive to the type of logical relation relatively early during children's development (i.e., before they enter elementary school). On the other hand, behavioral studies have also shown that the period of elementary school is associated with (i) an increase in the ability to use external cues when solving logical tasks (suggesting an increase in underlying deductive competence) [Mueller et al., 2001; O'Brien and Overton, 1982; Overton et al., 1985] and (ii) an increase in the number of mental models that children can construct from given premises [Gauffroy and Barrouillet, 2009; Markovits and Barrouillet, 2002]. This is consistent with the claim that formal deductive reasoning emerges around the period of elementary school [Piaget, 1981; Ricco and Overton, 2011] and increases with working-memory ability [Markovits and Barrouillet, 2002]. Therefore, it is also possible that the sensitivity in the brain to the type of logical relation may emerge around the period of elementary school.

The goal of this study was to test between these possibilities. Using functional magnetic resonance imaging (fMRI), we measured brain activity of school-age children (from ages 9 to 14) while they evaluated the validity of arguments such as (1) and (2) below:

1. Bud is slower than Joe
Joe is slower than Liz
Liz is slower than Rex
Therefore, Bud is slower than Rex

2. All larns are white
All white things are tall
Bud is a larn
Therefore, Bud is tall

Both arguments are valid, have a similar structure (three premises and a conclusion), and are about the same imaginary character (Bud). However, they differ with respect to the type of relation that they contain: whereas the premises of argument (1) involve linear-order relations, the premises of argument (2) involve relations of set-inclusion. In this study, we tested (i) whether brain regions involved in the processing of spatial and verbal information are sensitive to the type of relation that an argument contains (linear-order vs. set-inclusion) and (ii) whether this sensitivity increased with age during the elementary school period. To precisely identify regions involved in spatial and verbal processing, all subjects were presented with independent localizer tasks [Prado et al., 2013]. If the brain system for reasoning is already mature and sensitive to the type of logical relation before the time of elementary school, the differential reliance on spatial versus verbal mechanisms by linear-order versus set-inclusion arguments should be observed in children as young as 9. However, if this sensitivity emerges during the period of elementary school, this differential recruitment should increase with age.

MATERIALS AND METHODS

Participants

Thirty-six right-handed children with a full scale Intelligence Quotient (IQ) between 80 and 120 participated in the study. Children were recruited from schools in the Chicago metropolitan area. All were native English speakers. According to parental reports, participants did not have prior history of neurological disease, psychiatric disorders, learning disabilities, or attention deficits. All children and parents provided written informed consent to participate in the study, which was approved by the Northwestern University Institutional Review Board. Data from eight subjects were excluded because of excessive head movement in the scanner (see criteria below, $n = 6$), poor whole-brain coverage (i.e., insufficient coverage of the frontal, temporal, or parietal lobe, $n = 1$), and unacceptably low behavioral performance during the task (i.e., lower than 40% accuracy in the reasoning tasks, $n = 1$). Additionally, three subjects were not included in the analyses because they did not complete the entire experiment. Therefore, the final sample consisted of 25 children aged from 9.16 to 14.32 years (mean age = 11.48 years, standard deviation [SD] = 1.49, 14 males; mean IQ = 107.32, SD = 9.49). One child was in second grade, two children were in third grade, six children were in fourth grade, five children were in fifth grade, five children were in sixth grade, three children were in seventh grade, and three children were in eighth grade.

Cognitive Assessment

Children were administered with standardized tests of intellectual and reading abilities to ensure that there were no age differences with respect to those measures. IQ was measured with two verbal (similarities and vocabulary) and two performance (block design and matrix reasoning) tests of the Wechsler Abbreviated Scale of Intelligence [Wechsler, 1999]. Reading skills were evaluated by the Word Identification and Passage Comprehension subtests of the Woodcock Johnson III Tests of Achievement (WJ-III) [Woodcock et al., 2001].

Reasoning Trials

In each trial, participants were presented with a deductive argument that contained three premises and one conclusion [see (1) and (2) above]. Participants had to decide whether the conclusion of each argument necessarily followed from the premises. Arguments contained either linear-order or set-inclusion relations. Overall, linear-order arguments described a linear-ordering of four imaginary characters, while set-inclusion arguments described the inclusion of an imaginary character within two different classes. In linear-order arguments, the same comparative adjective was used throughout and consisted of one of the following: slower, faster, shorter, taller, younger, older, smaller, bigger. In set-inclusion arguments, the first class was characterized by a one-syllable name that was different in each problem (e.g., gofs, trabs, larns, progs). The second class was described by the following adjectives: tall, short, big, small, old, young, fast, slow, brown, red, black, blue, green, white, pink. Conclusions of arguments required either the integration of all three premises [e.g., arguments (1) and (2) above], or the integration of only two of the premises [e.g., consider the conclusions *Bud is slower than Liz* and *Bud is white* in arguments (1) and (2) above]. Furthermore, conclusions could be (i) valid and affirmative (18 linear-order and 18 set-inclusion arguments) [e.g., arguments (1) and (2) above], (ii) invalid and affirmative (six linear-order and six set-inclusion arguments) (e.g., *Bud is slower than Joe, Joe is slower than Liz, Liz is slower than Rex, therefore Rex is slower than Bud; All larns are white, All white things are tall, Bud is tall, therefore Bud is a larn*), (iii) valid with negation (six linear-order and six set-inclusion arguments) (e.g., *Bud is slower than Joe, Joe is slower than Liz, Liz is slower than Rex, therefore Rex is not slower than Bud; All larns are white, All white things are tall, Bud is not tall, therefore Bud is not a larn*), or (iv) invalid with negation (six linear-order and six set-inclusion arguments) (e.g., *Bud is slower than Joe, Joe is slower than Liz, Liz is slower than Rex, therefore Bud is not slower than Rex; All larns are white, All white things are tall, Bud is a larn, therefore Bud is not tall*) (see Supporting Information Table I for a complete list of experimental stimuli). We included such a variety of arguments to make the task as unpredictable as possible and encourage children to genuinely engage in reasoning during the experiment. Only arguments for which a correct response was provided were included in the analyses [Prado et al., 2010; Reverberi et al., 2007, 2010]. However, given that behavioral studies show age-related increases of accuracy for the most complex deductive arguments [Gauffroy and Barrouillet, 2009], age is likely to be confounded with number of correct responses if arguments are too difficult for younger children. To address this issue and ensure that the same number of arguments were analyzed in younger and older children, we focused on the most straightforward arguments for which we expected no (or minimal) age-related difference in accuracy: arguments for which the conclusion was valid and affirmative. Other arguments were considered fillers.

TABLE I. Brain regions activated in the localizer trials

Anatomical location	~BA	MNI coordinates X	Y	Z	Z-score
Spatial maintenance					
L. Inferior Parietal Lobule	40	−50	−44	50	3.60
R. Superior Temporal Gyrus	42	42	−36	2	3.47
L. Superior Parietal Lobule	7	−22	−64	42	3.41
R. Inferior Parietal Lobule	40	28	−50	46	3.35
L. Precuneus	7	−16	−60	38	2.92
R. Superior Parietal Lobule	7	28	−72	46	2.92
R. Precuneus	7	6	−64	58	2.66
Verbal maintenance					
L. ventral Inferior Frontal Gyrus	45/47	−48	38	6	4.45
L. Middle Temporal Gyrus	21/22	−52	−34	2	3.74
L. dorsal Inferior Frontal Gyrus	44	−52	16	22	3.44
L. Superior Temporal Gyrus	22	−60	−36	6	3.40
L. Precentral Gyrus	6	−48	8	10	3.20
L. Middle Frontal Gyrus	11	−36	34	−10	3.00
L. Angular Gyrus	39	−34	−60	26	2.76

L. = left; R. = right; ~BA = approximate Brodmann's area; MNI = Montreal Neurological Institute.

Localizer Trials

Each subject was run on functional localizer scans that included verbal and spatial maintenance trials. In spatial maintenance trials, two dot arrays were presented sequentially. The first dot array had to be maintained and compared with the second array to decide which array had the largest number of dots. One of the arrays always had 36 dots (reference array). The number of dots in the other array (test array) was varied parametrically to increase the difficulty of comparing the arrays with each other. The test array could have 12 dots (easy trials, $n = 24$), 18 dots (medium trials, $n = 24$), or 24 dots (hard trials, $n = 24$). The dot array with 36 dots was presented in the first position in half of the trials and in the second position in the other half. Six dot sizes were used and stimuli were controlled for differences in cumulative surface areas and distribution of dot sizes to ensure that judgments were solely based on numerical information [Prado et al., 2011b].

In verbal maintenance trials, two monosyllabic English words were presented sequentially. The first word had to be maintained and compared with the second word to decide whether the words rhymed or not. Orthography and phonology were manipulated independently to ensure that judgments were not based solely on orthographic similarities between words. The two words could have similar orthography and similar phonology (e.g., dime–lime; 12 trials), similar orthography but different phonology (e.g., pint–mint; 12 trials), different orthography but similar phonology (e.g., jazz–has; 12 trials), or different orthography and different phonology (e.g., press–list; 12 trials). These trials were contrasted with a control condition in which two symbol strings (i.e, rearranged parts of lower case Courier letters) were presented sequentially on the screen instead of words (12 trials). Participants had to determine whether the symbol strings matched (the symbols matched in half of the trials).

Experimental Protocol

After participants gave informed consent and were administered standardized tests, they participated in a practice session in which they practiced all trials and learned to minimize head movement in a mock fMRI scanner. This practice session included five arguments with linear-order relations, five arguments with set-inclusion relations, and 12 trials of each of the localizer trials (different sets of stimuli were used in the practice and fMRI session). In the fMRI scanner, participants performed two runs of each type of arguments. Participants were also presented with one run of verbal maintenance trials and two runs of spatial maintenance trials. The order of the tasks was counterbalanced across participants. The timing and order of trial presentation within each run was optimized for estimation efficiency using optseq2 (http://surfer.nmr.mgh.harvard.edu/optseq/). Behavioral responses were recorded using a keypad placed below the right hand. In reasoning trials, participants responded with their index finger if the conclusion was valid and with their middle finger if it was invalid. In verbal maintenance trials, participants responded with their index finger if the words rhymed and with their middle finger if they did not rhyme. In spatial maintenance trials, participants responded with their index finger if the first array was composed of more dots than the second array, and with their middle finger if the second array was composed of more dots than the first array. Stimuli were generated using E-prime software (Psychology Software Tools, Pittsburgh, PA) and projected onto a screen that was viewed by the participants through a mirror attached to the head coil.

Stimulus Timing

In reasoning trials, each premise and conclusion appeared on the screen one at a time and remained on the screen until the end of the trial. Each sentence was also simultaneously spoken through headphones to facilitate comprehension. The first premise was presented at 0 s, the second at 2 s, the third at 4 s, and the conclusion was displayed at 6 s (see Fig. 1). Response time (RT) was calculated from the presentation of the conclusion to the button press. The end of the trial occurred either when a button was pressed or 8 s after the onset of the conclusion if the participant did not respond. Variable periods of passive visual fixation (ranging from 2,600 to 3,400 ms) were added between each trial. Furthermore, each run ended

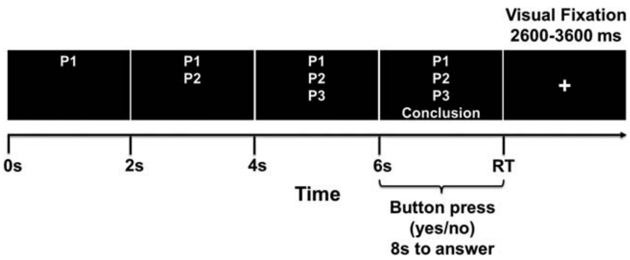

Figure 1.
Experimental procedure. In each trial, three premises (P1, P2, and P3) and one conclusion appeared one at a time every 2 s and remained on the screen until the end of the trial. Subjects were asked to evaluate the conclusion by pressing one of the two response keys. Each trial was followed by a period of visual fixation during which a cross remained at the center of the screen.

with 22 s of passive visual fixation. Those fixation periods (during which participants fixated a cross at the center of the screen) constituted the baseline.

Behavioral Analyses

Mean accuracy and mean RT associated with the evaluation of arguments of interest were analyzed in a general linear model (GLM) with the within-subject factor logical argument (linear-order vs. set-inclusion) and the continuous predictor age. Behavioral data associated with spatial and verbal maintenance trials were analyzed in additional GLM analyses. Accuracy and RT data were normalized using an arcsine and a logarithmic transformation (respectively) prior all analyses to improve the conformity of the data to the standard assumptions of parametric testing [Howell, 2011]. Mean RT was calculated based on correct trials only.

fMRI Data Acquisition

Images were collected with a Siemens Trio 3T MRI scanner (Siemens Healthcare, Erlangen, Germany). The fMRI blood oxygenation level dependent signal was measured with a susceptibility weighted single-shot echo planar imaging sequence. Imaging parameters were: time repetition (TR) = 2,000 ms, time echo (TE) = 20 ms, flip angle = 80°, matrix size = 128 × 120, field of view = 220 × 206.25 mm, slice thickness = 3 mm (0.48 mm gap), number of slices = 32, voxel size = 2 × 2 × 4 mm^3. In addition to the functional scans, a high-resolution T1-weighted whole-brain anatomical volume was collected for each participant (TR= 1,570 ms, TE= 3.36 ms, matrix size = 256 × 256, field of view = 240 mm, slice thickness = 1 mm, number of slices = 160).

fMRI Data Preprocessing

Data analysis was performed using the Statistical Parametric Mapping software (SPM8; Functional Imaging Laboratory, UCL, London, UK, http://www.fil.ion.ucl.ac.uk/spm). The first six scans of each run were removed to allow for magnetization equilibration effects. The remaining functional images were corrected for slice acquisition delays and realigned to the first image of the first run to correct for head movements. Images were then spatially smoothed with a Gaussian filter equal to twice the voxel size (4 × 4 × 8 mm^3 full width at half maximum). ArtRepair, an artifact repair software [Mazaika et al., 2009] (http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html), was used to help remove motion from the functional images prior to normalization. ArtRepair improves the quality of fMRI data containing high motion by removing residual motion fluctuation and detecting scans with significant artifact. Volumes with rapid scan-to-scan movements of greater than 3 mm were repaired by interpolation of the two nearest nonrepaired scans. We verified whether the repaired volumes corresponded to arguments of interest or filler arguments. A subject was excluded from further analysis if more than four arguments of interest were associated with repaired volumes. Finally, functional images were normalized into the standard Montreal Neurological Institute (MNI) space. This was done in two steps. First, after coregistration with the functional data, the structural image was segmented into gray matter, white matter, and cerebrospinal fluid by using a unified segmentation algorithm [Ashburner and Friston, 2005]. Second, the functional data were normalized to the MNI space by using the normalization parameters estimated during unified segmentation (normalized voxel size, 2 × 2 × 4 mm^3).

fMRI Data Analyses

Statistical analysis was performed according to the GLM [Josephs et al., 1997]. Reasoning arguments were modeled as epochs with onsets time locked to the presentation of the first premise and offsets time locked to the button press. Localizer trials were modeled as epochs with onsets time locked to the presentation of the first stimulus and a duration equal to trial length (2 s). Arguments of interest in which subjects provided a correct response were sorted by type of relation (linear-order, set-inclusion). Regressors of no interest coded all the other trials (i.e., fillers and incorrect responses on arguments of interest). Additionally, a parametric regressor coding for RTs across trials was included to rule out the possibility that any difference between conditions could be explained by trial-by-trial variation in performance. All correct verbal and spatial maintenance trials were also sorted by trial type and incorrect responses were coded as regressors of no interest. Epochs were convolved with a canonical hemodynamic response function. The time series data were high-pass filtered (1/128 Hz), and serial correlations were corrected using an autoregressive AR (1) model.

The goal of this study was to assess the sensitivity of the children's brain to the type of logical relation that an argument contains (linear-order vs. set-inclusion). Thus, whole-brain activity associated with arguments of interest was analyzed in a random-effect GLM with the within-subject factor logical argument (linear-order vs. set-inclusion) and the continuous predictor age. T-contrasts were created to determine the significance and directionality of (i) the main effect of logical argument and (ii) the interaction of logical argument X age. Note that these contrasts do not identify any activity that would be common to linear-order and set-inclusion arguments. Independent localizer contrasts were also calculated to identify regions involved in spatial and verbal maintenance. Regions involved in spatial maintenance were those in which activity increased with the difficulty of the comparison between dot arrays. In other words, these were regions activated in at least one of the following random-effect t-contrasts: Hard > Medium trials, Medium > Easy trials, and Hard > Easy trials. Regions involved in verbal maintenance were identified using the contrast Words > Symbol strings.

Statistical maps were controlled for a family-wise error (FWE) threshold of $P < 0.05$ across the whole brain, via a combination of uncorrected voxelwise height threshold of $P < 0.05$ and cluster extent threshold. The cluster extent threshold was calculated separately for each map by (i) estimating the spatial correlation across voxels using the program 3dFWHM and (ii) using this information as input to whole-brain Monte Carlo simulations (5,000 iterations calculated with the AlphaSim program, http://afni.nimh.nih.gov/afni/). One exception was made for the RostroLateral PreFrontal Cortex (RLPFC). This region (typically defined as the lateral part of Brodmann areas 10 and 47) was found to be involved in processing linear-order and set-inclusion arguments in a previous study on adult participants [Prado et al., 2013]. Thus, we used a small volume correction (SVC) procedure to examine activity in this region. Activity in the bilateral RLPFC was considered significant if it was below a FWE threshold of $P < 0.05$ across an anatomical mask including the rostral Middle Frontal Gyrus (MFG; approximately BA 10) and the Inferior Frontal Pars Orbitalis (approximately the lateral portion of BA 47), bilaterally (defined with the WFU PickAtlas Tool). This corrected threshold was achieved via a combination of uncorrected voxelwise height threshold of $P < 0.05$ and cluster extent threshold calculated within the anatomical mask (by using the Alphasim procedure described above). All coordinates are reported in MNI space and approximate Brodmann areas are identified by the Talairach Daemon software (http://www.talairach.org/daemon.html).

Brain activity in activated clusters was extracted for visualization using the SPM toolbox Marsbar (http://marsbar.sourceforge.net/). Regions of Interest (ROIs) included all significant voxels within a 6-mm radius of each coordinate of interest. For each participant, we calculated the average parameter estimate for each condition within an ROI by averaging the fMRI signal across all significant voxels within that ROI.

RESULTS

Behavioral Data

Cognitive assessment

Across all participants, mean performance IQ and mean verbal IQ were 107.6 (SD = 12.5) and 105.8 (SD = 8.8), respectively. Mean scores for the Word-Identification and Passage Comprehension subtests of the WJ-III were 106.6 (SD = 8.9) and 101.6 (SD = 7.2), respectively. Correlation analyses did not indicate any significant relationship between age and standardized score for measures of performance IQ ($r = -0.09$, $P = 0.68$), verbal IQ ($r = -0.16$, $P = 0.44$), Word-Identification ($r = -0.20$, $P = 0.34$) and Passage Comprehension ($r = -0.23$, $P = 0.27$). Therefore, children were comparable in terms of their age-normalized cognitive abilities.

Performance on reasoning trials

Accuracy in the scanner was 90% (range = 70%–100%) for all arguments of interest, with an average of 84% (range = 53%–100%) and 97% (range = 65%–100%) for linear-order and set-inclusion arguments, respectively. Accuracy did not change with age ($F_{1,23} = 1.79$, $P = 0.19$) and age did not interact with logical relation ($F_{1,23} = 0.17$, $P = 0.68$). Thus, the number of linear-order and set-inclusion relations included in the fMRI analyses was similar in younger and in older children. However, there was a main effect of logical argument ($F_{1,23} = 35.41$, $P = 4.57 \times 10^{-6}$), indicating that arguments that contained set-inclusion relations were evaluated more accurately than those with linear-order relations. To ensure that none of the fMRI results could be explained by a difference in the number of linear-order versus set-inclusion arguments analyzed, we included in all analyses the mean accuracy difference between linear-order and set-inclusion arguments as nuisance covariate.

RT across all arguments of interest was 2,372 ms (range = 1,253–3,482 ms), with an average of 3,031 ms (range = 1,561–4,910 ms) and 1,714 ms (range = 871–2,465 ms) for linear-order and set-inclusion arguments, respectively. There was no main effect of age ($F_{1,23} = 1.09$, $P = 0.31$) and no interaction of age by logical argument ($F_{1,23} = 0.46$, $P = 0.50$). However, there was a main effect of logical argument ($F_{1,23} = 113.35$, $P = 2.31 \times 10^{-10}$), indicating that conclusions of arguments that contained set-inclusion relations were evaluated faster than those that contained linear-order relations. To ensure that none of the fMRI results could be affected by this difference of response times, we included in all analyses the mean

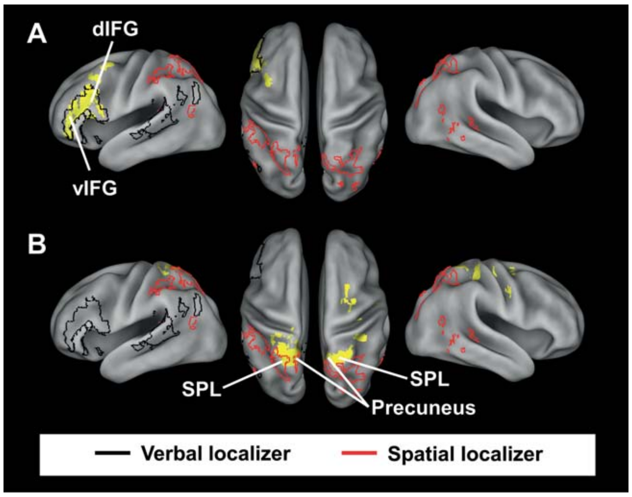

Figure 2.

Main effect of logical argument. **A**: Brain regions showing greater activation for set-inclusion than linear-order arguments. **B**: Brain regions showing greater activation for linear-order than set-inclusion arguments. Black outlines represent regions involved in verbal maintenance and red outlines represent regions involved in spatial maintenance. Activations are overlaid on a three-dimensional (3D) rendering of the MNI-normalized anatomical brain. vIFG, ventral Inferior Frontal Gyrus; dIFG, dorsal Inferior Frontal Gyrus; SPL, Superior Parietal Lobule. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

response-time difference between linear-order and set-inclusion arguments as nuisance covariate.

Performance on localizer trials

Performance on localizer trials was high. For spatial maintenance trials, accuracy and RT were 86% and 1,022 ms. Reaction times increased with comparison difficulty (easy trials, 987 ms; medium trials, 1,058 ms; hard trials, 1,021 ms; $F_{2,48} = 8.16$, $P = 8.93 \times 10^{-4}$). No difference was observed in terms of accuracy (easy trials, 87%; medium trials, 87%; hard trials, 85%; $F_{2,48} = 1.71$, $P = 0.19$).

For verbal maintenance trials, accuracy and RT were 79% and 1,227 ms for word trials. They were 92% and 659 ms for symbol string trials. Behavioral performance was lower when participants evaluated word pairs than when they evaluated symbol strings (accuracy: t_{24}=3.54, $P = 0.002$; reaction times: $t_{24} = 9.71$, $P = 3.21 \times 10^{-9}$).

fMRI Data

Localizer trials

Localizer trials were used to identify brain regions involved in the maintenance of spatial and verbal information (see Methods). Regions involved in spatial maintenance were found in the bilateral Superior Parietal Lobule (SPL), the bilateral Inferior Parietal Lobule bordering the Supramarginal Gyrus (SMG), the bilateral Precuneus, and the right Superior Temporal Gyrus (STG) (see Table I and red outlines in Fig. 2). Regions involved in verbal maintenance included the left ventral and dorsal Inferior Frontal Gyrus (vIFG/dIFG), the left Precentral Gyrus (PG), the left MFG, the left STG, the left Middle Temporal Gyrus, and the left Angular Gyrus (see Table I and black outlines in Fig. 2).

Main effect of logical argument

First, we calculated the contrast set-inclusion arguments > linear-order arguments to identify the regions more activated for set-inclusion than linear-order arguments across all children. We found activity in a cluster encompassing the left ventral and dorsal IFG (vIFG and dIFG), as well as in the left MFG and the PG (see Fig. 2A and Table II). A Boolean intersection between the contrast set-inclusion arguments > linear-order arguments and the verbal localizer map (both thresholded at $P < 0.05$, corrected for multiple comparisons) revealed 395 voxels (6,320 mm^3) in common between the maps (see Fig. 3A). We did not find any common voxels between the contrast set-inclusion arguments > linear-order arguments and the spatial localizer map.

Second, we calculated the contrast linear-order arguments > set-inclusion arguments to identify the regions more activated for linear-order than set-inclusion arguments across all children. We found activity in the bilateral SPL and bilateral Precuneus, as well as in the right MFG

TABLE II. Brain regions differentially activated during linear-order versus set-inclusion arguments across all subjects

		MNI coordinates			
Anatomical location	~BA	*X*	*Y*	*Z*	*Z*-score
Set-inclusion > Linear-order					
L. dorsal Inferior Frontal Gyrus	44	−48	10	14	3.15
L. Middle Frontal Gyrus	8/9	−44	22	46	3.02
L. ventral Inferior Frontal Gyrus	47	−46	40	−10	2.73
L. Precentral Gyrus	6	−42	4	58	2.48
Linear-order > Set-inclusion					
R. Superior Parietal Lobule/ Precuneus	7	18	−54	58	3.85
L. Precuneus	7	−12	−52	54	3.70
L. Superior Parietal Lobule	7	−24	−46	66	2.98
R. Lingual Gyrus	18	12	−66	6	2.94
R. Middle Frontal Gyrus	6	26	−10	46	2.93
R. Precentral Gyrus	6	40	−10	42	2.84

L. = left; R. = right; ~BA = approximate Brodmann's area; MNI = Montreal Neurological Institute.

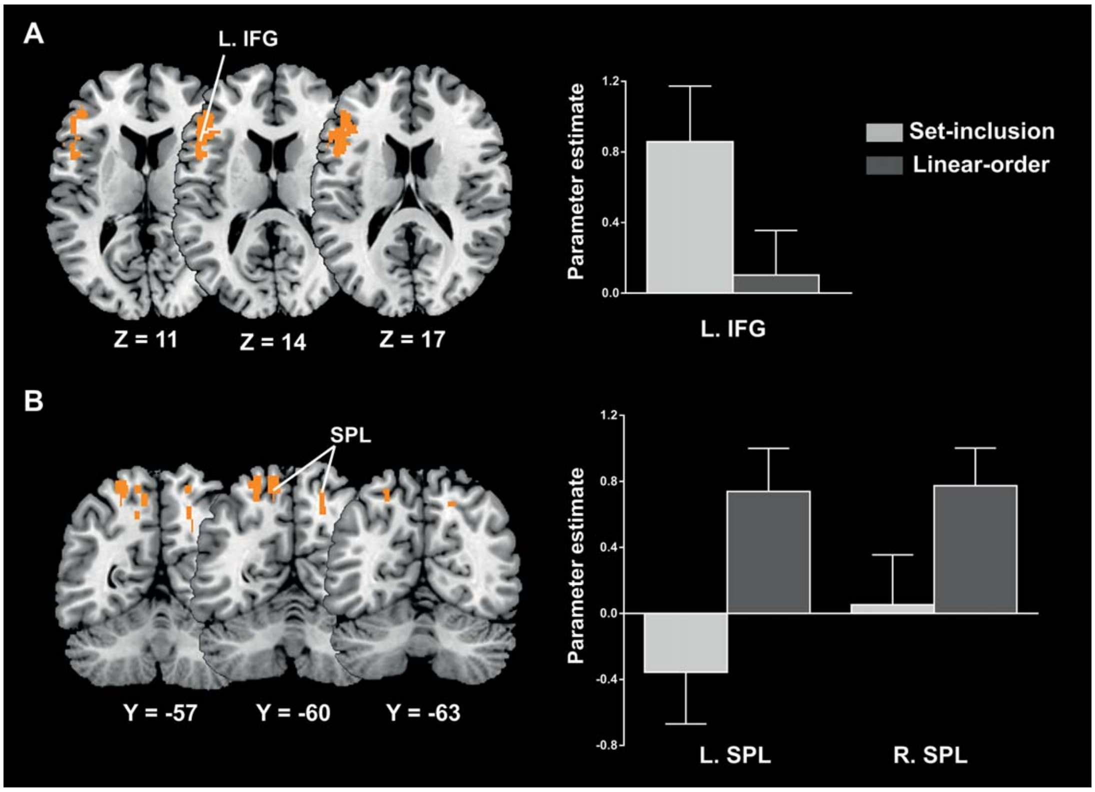

Figure 3.

Overlap between regions involved in verbal and spatial maintenance and regions sensitive to the type of logical argument. **A**: Left: Boolean intersection between the regions showing greater activity for set-inclusion than linear-order arguments and the regions involved in verbal maintenance. Right: Plot of the brain activity observed for set-inclusion (light gray) and linear-order (dark gray) arguments in the overlapping left IFG cluster (for visualization only). 0 corresponds to the baseline (**B**) Left: Boolean intersection between the regions showing greater activity for linear-order than set-inclusion arguments and the regions involved in spatial maintenance. Right: Plot of the brain activity observed for set-inclusion (light gray) and linear-order (dark gray) arguments in the overlapping left and right SPL clusters (for visualization only). 0 corresponds to the baseline. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

and the right PG (see Fig. 2B and Table II). A Boolean intersection between the contrast linear-order arguments > set-inclusion arguments and the spatial localizer map (both thresholded at $P < 0.05$, corrected for multiple comparisons) revealed 123 voxels (1,968 mm^3) in common between the maps (see Fig. 3B). We did not find any common voxels between the contrast linear-order arguments > set-inclusion arguments and the verbal localizer map.

Interaction of logical argument by age

To explore the interaction between logical argument and age, we regressed the activity associated with the contrasts set-inclusion arguments > linear-order arguments and linear-order arguments > set-inclusion arguments with age. We did not find any regions in which age-related increases of activity were greater for linear-order than set-inclusion arguments. However, age-related increases of activity were greater for set-inclusion than linear-order arguments in several frontal and parietal regions. These included clusters in the bilateral PG, bilateral medial Superior Frontal Gyrus (mSFG), bilateral Precuneus, and right SPL (see Fig. 4 and Table III). Additionally, we found greater age-related increases of activity for set-inclusion than linear-order arguments in the left RLPFC after SVC (see Methods). None of these clusters overlapped with regions involved in verbal or spatial maintenance.

DISCUSSION

There is accumulating evidence that deductive reasoning does not rely on a homogeneous brain system in adults. Instead, it appears to involve neural systems that are sensitive to task-related factors, such as semantic content of premises [Goel et al., 2000, 2004] or argument complexity [Monti et al., 2007; Noveck et al., 2004]. Recent evidence indicates

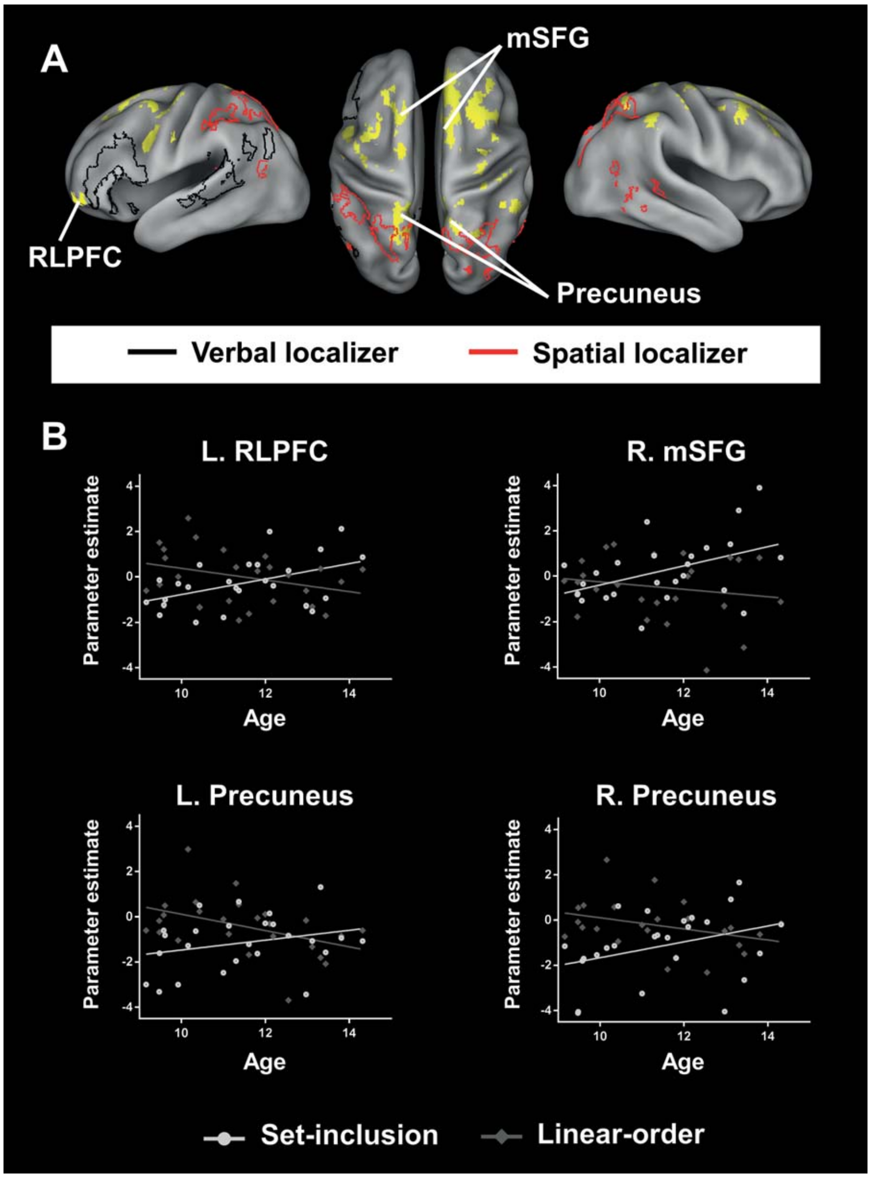

Figure 4.

Interaction between age and logical argument. **A**: Brain regions showing greater age-related increases of activity for set-inclusion than linear-order arguments. **B**: Plots of the brain activity observed for set-inclusion (light gray) and linear-order (dark gray) arguments in the left RLPFC, right mSFG, left Precuneus and right Precuneus as a function of age (for visualization only). Black outlines represent regions involved in verbal maintenance and red outlines represent regions involved in spatial maintenance. Activations are overlaid on a 3D rendering of the MNI-normalized anatomical brain. RLPFC, RostroLateral PreFrontal Cortex; mSFG, medial Superior Frontal Gyrus. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

that the reasoning brain system may also be sensitive to the logical form of the premises [Prado et al., 2011a, 2013; Reverberi et al., 2010]. For example, we have shown arguments that contain linear-order relations are more likely to be associated with spatial regions of the parietal cortex than arguments that contain set-inclusion relations [Prado et al., 2011a, 2013]. In contrast, arguments that contain set-inclusion relations are more likely to be associated with verbal regions of the frontal cortex than arguments that contain linear-order relations [Prado et al., 2011a, 2013]. In line with these findings, we found here that reasoning with linear-order and set-inclusion relations is associated with differential engagement of spatial and verbal mechanisms in children as young as age 9, suggesting that the neural sensitivity to the type of logical

TABLE III. Brain regions showing greater age-related activity for set-inclusion than linear-order arguments

Anatomical location	~BA	MNI coordinates X	Y	Z	Z-score
L. Precentral Gyrus	6	−28	−14	70	4.13
L. Superior Frontal Gyrus	6/8	−16	10	58	3.51
R. Middle Frontal Gyrus	8	26	10	46	3.45
L. Middle Frontal Gyrus	6	−28	2	62	3.35
R. Supramarginal Gyrus	40	52	−44	54	3.15
L. RLPFC	10/11	−28	46	−10	3.12
R. Medial/Superior Frontal Gyrus	6/8	6	32	54	2.82
R. Postcentral Gyrus	3	48	−22	62	2.75
L. Precuneus/Superior parietal Lobule	7	−18	−50	66	2.62
R. Superior Parietal Lobule	7	28	−58	66	2.61
L. Medial/Superior Frontal Gyrus	6/8	−14	18	58	2.37
R. Precentral Gyrus	6	28	−20	66	2.37
R. Superior Frontal Gyrus	8	14	26	54	2.23
R. Precuneus	7	8	−52	62	2.17

L. = left; R. = right; ~BA = approximate Brodmann's area; MNI = Montreal Neurological Institute.

argument might appear early during development. However, the period of elementary school was associated with age-related changes of activity that were dependent on the type of argument in several frontal and parietal regions. Therefore, this period might play an important role in shaping the neural processing of deductive reasoning in children.

Spatial and Verbal Regions are Sensitive to the Type of Logical Argument Before Elementary School

A central finding of this study is that brain activity was modulated by the type of logical relation present in a deductive argument in school-age children. As in adults [Prado et al., 2011a, 2013], arguments that contained linear-order relations were associated with greater activity than set-inclusion relations in regions of the parietal cortices (SPL and Precuneus) that were also involved in the maintenance of spatial information. In contrast, arguments that contained set-inclusion relations were associated with greater activity than linear-order relations in a region of the left IFG that were also involved in the maintenance of verbal information. In other words, when both types of arguments are compared to each other, linear-order arguments elicit extra processing in spatial regions of the SPL/Precuneus, whereas set-inclusion arguments elicit extra processing in verbal regions of the left IFG. This neural dissociation was as strong in younger as it was in older school-age children (i.e., it was not modulated by age), and might have, thus, appeared well before the period of elementary education. Because a similar dissociation has been found in adults [Prado et al., 2011a, 2013], this finding suggests that the sensitivity of the brain system for reasoning to different types of logical relations emerges relatively early. This is broadly consistent with the idea that (at least some) logical competence might appear before the period of elementary school [Moshman, 2004; Noveck, 2001].

Why would spatial and verbal brain regions be differentially activated by arguments containing linear-order and set-inclusion relations? A first possibility is that the interaction might result from factors inherent to the lexical content of the arguments. For example, linear-order relations involve comparative adjectives. These are often thought to elicit spatial and/or visual mental images [Knauff and Johnson-Laird, 2002] and have been linked to the posterior parietal cortex in neuroimaging studies [Coventry et al., 2013; Noordzij et al., 2008; Wallentin et al., 2005]. In contrast, set-inclusion relations involve Aristotelian quantifiers (e.g., "all," "some"). Unlike quantifiers that refer to numerical concepts (e.g., "at least 3"), Aristotelian quantifiers have not been associated with spatial regions of the parietal cortex in neuroimaging and patient studies [Morgan et al., 2011; Troiani et al., 2009]. Rather, they appear to be associated with frontal regions, perhaps because they involve a form of simple perceptual logic [Morgan et al., 2011; Troiani et al., 2009]. A second possibility, suggested by Favrel and Barrouillet [2000], is that it might be more difficult to construct spatial representations of set-inclusion relations than spatial representations of linear-order relations. This could be because sets cannot be aligned along a single underlying dimension, such as age or size for linear-orderings [Favrel and Barrouillet, 2000]. This could also be because a relation such as *All Larns are white* is ambiguous and compatible with two spatial representations (it could depict either the identity or the inclusion of two sets), whereas a relation such as *Bud is slower than Joe* is only compatible with one model and much easier to represent as a mental image [Favrel and Barrouillet, 2000; Prado et al., 2011a, 2013]. In support of this idea, Mani and Johnson-Laird [1982] have demonstrated that it is easier to encode a statement compatible with two or more models in a verbal than spatial form. Our findings, as well as a large body of behavioral studies [Barrouillet, 1996; Carlson et al., 1992; Favrel and Barrouillet, 2000; Griggs, 1976; Griggs and Osterman, 1980; Newstead et al., 1985], are compatible with the idea that reasoners might not easily construct integrated spatial representations of set-inclusion relations and might, therefore, rely on verbal representations.

The period of Elementary School is Associated with Changes in the Sensitivity of Fronto-Parietal Regions to the Type of Logical Argument

As discussed above, the neural sensitivity to the type of logical argument changed with age in none of the regions involved in spatial or verbal maintenance. However, we found an interaction between age and type of argument in several regions of the frontal (i.e., RLPFC, mSFG/PG) and parietal (i.e., Precuneus/SPL, SMG) cortices. Specifically, these regions were associated with greater age-related increases of activity for set-inclusion arguments than for linear-order arguments. This interaction provides further evidence for processing differences between linear-order and set-inclusion relations. Because all of these clusters were located outside of the verbal and spatial brain regions identified by the localizer trials, we can only speculate on the possible functions of these regions.

First, it has been proposed that the role of the frontal regions we identified (i.e., RLPFC and mSFG/PG) during deductive reasoning is to coordinate inferential steps, while keeping track of the overall structure of the arguments [Monti et al., 2009; Reverberi et al., 2009, 2012a]. This claim is in line with the fact that activity in the RLPFC and/or the mSFG increases with the number of inferential steps in a deductive argument [Monti et al., 2007; Noveck et al., 2004; Prado et al., 2013]. It is also consistent with the more general proposal that these frontal regions are important when a task requires the coordination and integration of multiple subgoals to achieve a main goal [Koechlin et al., 2000; Ramnani and Owen, 2004]. Such coordination and monitoring operations are critical when premises of arguments are stored as isolated representations and subsequently integrated in a step-by-step manner, as might be the case for set-inclusion relations [Favrel and Barrouillet, 2000]. But they are unnecessary when statements elicit the construction of an integrated mental model, as is likely to be the case for linear-order relations [Goodwin and Johnson-Laird, 2005]. Thus, the interaction between age and type of logical argument in the RLPFC and mSFG/PG might be a consequence of the different formats the relations are stored in, and may reflect a maturation of the different cognitive strategies involved in processing different types of logical arguments.

Second, it is surprising to note that set-inclusion arguments were associated with age-related increases of activity in parietal regions. Although none of these regions overlapped with the regions identified by our spatial localizer, they were relatively close. Therefore, it remains possible that such increased activity reflects an increase in the use of spatial mechanisms for set-inclusion arguments with age. Indeed, although neuroimaging findings indicate that set-inclusion relations are more likely to be encoded in a verbal rather than spatial format in most participants [Goel et al., 2000; Prado et al. 2010, 2013], this does not imply that spatial representations can never be used when reasoning with these relations. In fact, the use of spatial representations is likely to be constrained by spatial working-memory capacities. As such capacities are known to increase with age in children [Klingberg et al., 2002], older children might have access to both verbal and spatial mechanisms when reasoning with set-inclusion relations. Developmental increases of activity in parietal regions might thus index a greater use of spatial mechanisms by some individuals (or in some trials) in older than younger children.

Relationship of Present Work to Prior Studies on the Neural Development of Reasoning

To the best of our knowledge, this study is the first to explore some of the neural mechanisms associated with the comprehension of deductive arguments in children. Previous studies, however, have investigated the neural operations that allow children to make analogies (e.g., BRAIN is to THOUGHT as STOMACH is to?) [Crone et al., 2009; Dumontheil et al., 2010; Eslinger et al., 2009; Wendelken et al., 2008, 2011]. Analogical and deductive reasoning differ in several aspects and are accounted for by different theories in the cognitive literature [Gentner, 2003; Johnson-Laird and Byrne, 2002]. However, they also share important features, such as the need for representing relations between items and comparing these relations with each other. Using visual analogy tasks (e.g., Raven's progressive matrices), studies have found developmental changes of activity in both the parietal cortex and the RLPFC [Crone et al., 2009; Dumontheil et al., 2010; Eslinger et al., 2009; Wendelken et al., 2011]. Interestingly, it has been proposed that regions of the parietal cortex might be important for representing simple relationships between mental representations during analogical reasoning, whereas the RLPFC might build higher-order structures (where different relations are integrated with each other and need to be coordinated) [Wendelken et al., 2011]. This interpretation can be applied to the present findings as well: enhanced activity in the SPL for linear-order relations might reflect the construction of a unified, single, mental representation of the relations between each item, whereas enhanced activity in the RLPFC for set-inclusion relations might indicate the construction of a higher order structure in which separate mental representations are coordinated. Thus, deductive and analogical reasoning might share some common mechanisms at the neural level. Future neuroimaging studies are needed to explore the degree of similarity between both types of reasoning.

Limitations

This study aimed to examine whether the type of logical relation contained in an argument influences the brain activity associated with that argument. Therefore, our design and analyses focused on exploring differences between set-inclusion and linear-order arguments, rather than exploring brain activity common to both types of arguments. Future studies might investigate whether both types of arguments are associated with the development of common reasoning processes by comparing these arguments to a baseline controlling for other confounding factors (e.g., reading, response selection).

Another limitation is that it remains unclear whether the neural sensitivity we observe results from differences in premise representation or in reasoning per se (or both). Dissociating between these factors in our study is difficult for methodological and theoretical reasons. First, premises were immediately followed by conclusions in our design. The sluggishness of the hemodynamic response makes it difficult to clearly dissociate the premise processing from the conclusion stage. Second, activity associated with premises might not only reflect premise representation but also reasoning. Indeed, prior behavioral research suggests that participants automatically integrate premises on the fly and make spontaneous inferences based on these premises [Lea, 1995; Reverberi et al., 2012b]. Recent neuroimaging research has also shown that reasoning-related activity can be observed before the appearance of the conclusion (i.e., on the presentation of the premises) and might correspond to the generation of this conclusion [Prado et al., 2010; Reverberi et al., 2007, 2010]. For these reasons, future studies will be needed to further investigate the source of the effects we observed.

CONCLUSION

Our results demonstrate distributed representations of linear-order and set-inclusion arguments in spatial and verbal brain regions of school-age children. Specifically, we found that linear-order and set-inclusion arguments preferentially involve spatial and verbal brain mechanisms in these children (respectively). Because this neural sensitivity was not modulated by age, it likely emerges before the beginning of elementary school [Favrel and Barrouillet, 2000; Prado et al., 2013]. However, our results also indicate that the period of elementary education is associated with neurodevelopmental changes in several frontal and parietal regions that were dependent on the type of relation. Given the importance of deductive reasoning for academic learning, an important goal for future research is to evaluate whether language impairments or learning disabilities are associated with specific deficits of the brain systems supporting the representation of linear-order and set-inclusion reasoning [Katsos et al., 2011; Morsanyi et al., 2013].

ACKNOWLEDGMENT

We wish to thank two anonymous reviewers for their helpful comments on a previous version of this manuscript.

REFERENCES

Ashburner J, Friston KJ (2005): Unified segmentation. Neuroimage 26:839–851.

Barrouillet P (1996): Transitive inferences from set inclusion relations and working memory. J Exp Psychol Learn Mem Cogn 6: 1408–1422.

Braine MD, O'Brien D. 1998. Mental Logic. Mahwah, NJ: Lawrence Erlbaum Associates. 481 p.

Bryant PE, Kopytynska H (1976): Spontaneous measurement by young children. Nature 260:773–773.

Bryant PE, Trabasso T (1971): Transitive inferences and memory in young children. Nature 232:456–458.

Carlson RA, Lundy DH, Yaure RG (1992): Syllogistic inference chains in meaningful text. Am J Psychol 105:75–99.

Coventry KR, Christophel TB, Fehr T, Valdes-Conroy B, Herrmann M (2013): Multiple routes to mental animation: Language and functional relations drive motion processing for static images. Psychol Sci 24:1379–1388.

Crone E, Wendelken C, van Leijenhorst L, Honomichl R, Christoff K, Bunge S (2009): Neurocognitive development of relational reasoning. Dev Sci 12:55–66.

Dias MG, Harris PL (1988): The effect of make-believe play on deductive reasoning. Br J Dev Psychol 6:207–221.

Dias MG, Harris PL (1990): The influence of the imagination on reasoning by young children. Br J Dev Psychol 8:305–318.

Dumontheil I, Houlton R, Christoff K, Blakemore S-J. (2010): Development of relational reasoning during adolescence. Dev Sci 13:F15–F24.

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.

Favrel J, Barrouillet P (2000): On the relation between representations constructed from text comprehension and transitive inference production. J Exp Psychol Learn Mem Cogn 26:187–203.

Gauffroy C, Barrouillet P (2009): Heuristic and analytic processes in mental models for conditionals: An integrative developmental theory. Dev Rev 29:249–282.

Gentner D (2003): Psychology of analogical reasoning. In: Encyclopedia of Cognitive Science. London: Nature Publishing Group. pp 106–112.

Goel V (2007): Anatomy of deductive reasoning. Trends Cogn Sci 11:435–441.

Goel V, Dolan RJ (2003): Explaining modulation of reasoning by belief. Cognition 87:B11–B22.

Goel V, Buchel C, Frith C, Dolan RJ (2000): Dissociation of mechanisms underlying syllogistic reasoning. Neuroimage 12:504–514.

Goel V, Makale M, Grafman J (2004): The hippocampal system mediates logical reasoning about familiar spatial environments. J Cogn Neurosci 16:654–664.

Goel V, Stollstorff M, Nakic M, Knutson K, Grafman J (2009): A role for right ventrolateral prefrontal cortex in reasoning about indeterminate relations. Neuropsychologia 47:2790–2797.

Goodwin GP, Johnson-Laird PN (2005): Reasoning about relations. Psychol Rev 112:468–493.

Goodwin GP, Johnson-Laird PN (2008): Transitive and pseudo-transitive inferences. Cognition 108:320–352.

Griggs R (1976): Logical processing of set inclusion relations in meaningful text. Mem Cogn 4:730–740.

Griggs R, Osterman L (1980): Processing artificial set inclusion relations. J Exp Psychol Hum Learn Mem 6:39–52.

Hawkins J, Pea RD, Glick J, Scribner S (1984): Merds that laugh don't like mushrooms: Evidence for deductive reasoning by preschoolers. Dev Psychol 20:584–594.

Howell DC. 2011. Statistical Methods for Psychology. Belmont, CA: Wadsworth Cengage Learning.

Johnson-Laird PN, Byrne RM (2002): Conditionals: A theory of meaning, pragmatics, and inference. Psychol Rev 109:646.

Josephs O, Turner R, Friston K (1997): Event-related fMRI. Hum Brain Mapp 5:243–248.

Katsos N, Roqueta CA, Estevan RA, Cummins C (2011): Are children with specific language impairment competent with the pragmatics and logic of quantification? Cognition 119:43–57.

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

Knauff M, Johnson-Laird PN (2002): Visual imagery can impede reasoning. Mem Cogn 30:363–371.

Koechlin E, Corrado G, Pietrini P, Grafman J (2000): Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. Proc Nat Acad Sci USA 97:7651–7656.

Lea RB (1995) On-line evidence for elaborative logical inferences in text. J Exp Psychol Learn Mem Cogn 21:1469.

Macnamara J. 1986. A Border Dispute: The Place of Logic in Psychology. Cambridge, MA: MIT Press.

Mani K, Johnson-Laird PN (1982): The mental representation of spatial descriptions. Mem Cogn 10:181–187.

Markovits H, Barrouillet P (2002): The development of conditional reasoning: A mental model. Dev Rev 22:5–36.

Mazaika PK, Hoeft F, Glover GH, Reiss AL (2009): Methods and software for fMRI analysis of clinical subjects. Neuroimage 47: S58.

Michal A, Ruhama E (2008): Deductive reasoning: In the eye of the beholder. Educ Stud Math 69:235–247.

Monti MM, Osherson DN, Martinez MJ, Parsons LM (2007): Functional neuroanatomy of deductive inference: A language-independent distributed network. Neuroimage 37:1005–1016.

Monti MM, Parsons LM, Osherson DN (2009): The boundaries of language and thought in deductive inference. Proc Nat Acad Sci USA 106:12554–12559.

Morgan B, Gross RG, Clark R, Dreyfuss M, Boller A, Camp E, Liang TW, Avants B, McMillan CT, Grossman M (2011): Some is not enough: Quantifier comprehension in corticobasal syndrome and behavioral variant frontotemporal dementia. Neuropsychologia 49:3532–3541.

Morsanyi K, Devine A, Nobes A, Szűcs D (2013): The link between logic, mathematics and imagination: Evidence from children with developmental dyscalculia and mathematically gifted children. Dev Sci 16:542–553.

Moshman D (2004): From inference to reasoning: The construction of rationality. Think Reason 10:221–239.

Mueller U, Overton WF, Reene K (2001): Development of conditional reasoning: A longitudinal study. J Cogn Dev 2:27–49.

Newstead S, Keeble S, Manktelow K (1985): Children's performance on set inclusion and linear ordering relationships. Bull Psychonomic Soc 23:105–108.

Noordzij ML, Neggers SFW, Ramsey NF, Postma A (2008): Neural correlates of locative prepositions. Neuropsychologia 46:1576–1580.

Noveck IA (2001): When children are more logical than adults: Experimental investigations of scalar implicature. Cognition 78:165–188.

Noveck IA, Goel V, Smith KW (2004): The neural basis of conditional reasoning with arbitrary content. Cortex 40:613–622.

Nunes T, Bryant P, Evans D, Bell D, Gardner S, Gardner A, Carraher J (2007): The contribution of logical reasoning to the learning of mathematics in primary school. Br J Dev Psychol 25:147–166.

O'Brien DP, Overton WF (1982): Conditional reasoning and the competence-performance issue: A developmental analysis of a training task. J Exp Child Psychol 34:274–290.

Overton WF, Byrnes JP, O'Brien DP (1985): Developmental and individual differences in conditional reasoning: The role of contradiction training and cognitive style. Dev Psychol 21:692–701.

Piaget J. 1981. Intelligence and Affectivity: Their Relationship During Child Development. Palo Alto, CA: Annual Reviews.

Potts G (1972): Information processing stragies used in the encoding of linear ordering. J Verb Learn Verb Behav 11:727–740.

Potts G (1974): Storing and retrieving information about ordered relationship. J Exp Psychol 103:431–439.

Prado J, Van der Henst J-B, Noveck IA (2008): Spatial associations in relational reasoning: Evidence for a SNARC-like effect. Q J Exp Psychol 61:1143–1150.

Prado J, Van Der Henst J-B, Noveck IA (2010): Recomposing a fragmented literature: How conditional and relational arguments engage different neural systems for deductive reasoning. Neuroimage 51:1213–1221.

Prado J, Chadha A, Booth J (2011a): The brain network for deductive reasoning: A quantitative meta-analysis of 28 neuroimaging studies. J Cogn Neurosci 23:3483–3497.

Prado J, Mutreja R, Zhang H, Mehta R, Desroches AS, Minas JE, Booth JR (2011b): Distinct representations of subtraction and multiplication in the neural systems for numerosity and language. Hum Brain Mapp 32:1932–1947.

Prado J, Mutreja R, Booth JR (2013): Fractionating the neural substrates of transitive reasoning: Task-dependent contributions of spatial and verbal representations. Cereb Cortex 23:499–507.

Ramnani N, Owen A (2004): Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. Nat Rev Neurosci 5:184–194.

Reverberi C, Cherubini P, Rapisarda A, Rigamonti E, Caltagirone C, Frackowiak RSJ, Macaluso E, Paulesu E (2007): Neural basis of generation of conclusions in elementary deduction. Neuroimage 38:752–762.

Reverberi C, Shallice T, D'Agostini S, Skrap M, Bonatti LL (2009): Cortical bases of elementary deductive reasoning: Inference, memory, and metadeduction. Neuropsychologia 47:1107–1116.

Reverberi C, Cherubini P, Frackowiak RSJ, Caltagirone C, Paulesu E, Macaluso E (2010): Conditional and syllogistic deductive tasks dissociate functionally during premise integration. Hum Brain Mapp 31:1430–1445.

Reverberi C, Bonatti LL, Frackowiak RSJ, Paulesu E, Cherubini P, Macaluso E (2012a): Large scale brain activations predict reasoning profiles. Neuroimage 59:1752–1764.

Reverberi C, Pischedda D, Burigo M, Cherubini P (2012b): Deduction without awareness. Acta Psychol 139:244–253.

Ricco RB, Overton WF (2011): Dual systems competence - procedural processing: A relational developmental systems approach to reasoning. Dev Rev 31:119–150.

Rips L. 1994. The Psychology of Proof. Cambridge, MA: MIT Press.

Stanovich KE, West RF (2000): Individual differences in reasoning: Implications for the rationality debate? Behav Brain Sci 23:645–665.

Troiani V, Peelle JE, Clark R, Grossman M (2009): Is it logical to count on quantifiers? Dissociable neural networks underlying numerical and logical quantifiers. Neuropsychologia 47:104–111.

Wallentin M, Ostergaard S, Lund S, Ostergaard L, Roepstorff A (2005): Concrete spatial language: See what I mean? Brain Lang 92:221–233.

Wechsler D. 1999. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Psychological Corporation.

Wendelken C, Nakhabenko D, Donohue SE, Carter CS, Bunge SA (2008): Brain is to thought as stomach is to??: Investigating the role of rostrolateral prefrontal cortex in relational reasoning. J Cogn Neurosci 20:682–693.

Wendelken C, O'Hare E, Whitaker K, Ferrer E, Bunge S (2011): Increased functional selectivity over development in rostrolateral prefrontal cortex. J Neurosci 31:17260–17268.

Woodcock R, McGrew KS, Mather N. 2001. Woodcock-Johnson-III Tests of Achievement. Itasca, IL: Riverside Publishing.