NeuroImage 163 (2017) 264–275

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

The neural development of conditional reasoning in children: Different mechanisms for assessing the logical validity and likelihood of conclusions

Flora Schwartz *, Justine Epinat-Duclos, Jessica Léone, Jérôme Prado **

Institut des Sciences Cognitives Marc Jeannerod, UMR 5304, Centre National de la Recherche Scientifique (CNRS) & Université de Lyon, France

ARTICLE INFO

Keywords:
Reasoning
Higher-level cognition
Functional MRI
Development
Deduction

ABSTRACT

Scientific and mathematical thinking relies on the ability to evaluate whether conclusions drawn from conditional (if-then) arguments are logically valid. Yet, the neural development of this ability – termed deductive reasoning – is largely unknown. Here we aimed to identify the neural mechanisms that underlie the emergence of deductive reasoning with conditional rules in children. We further tested whether these mechanisms have their roots in the neural mechanisms involved in judging the likelihood of conclusions. In a functional Magnetic Resonance Imaging (fMRI) scanner, 8- to 13-year-olds were presented with causal conditional problems such as "If a baby is hungry then he will start crying; The baby is crying; Is the baby hungry?". In *Validity* trials, children were asked to indicate whether the conclusion followed out of necessity from the premises. In *Likelihood* trials, they indicated the degree of likelihood of the conclusion. We found that children who made accurate judgments of logical validity (as compared to those who did not) exhibited enhanced activity in left and medial frontal regions. In contrast, differences in likelihood ratings between children were related to differences of activity in right frontal and bilateral parietal regions. There was no overlap between the brain regions underlying validity and likelihood judgments. Therefore, our results suggest that the ability to evaluate the logical validity of conditional arguments emerges from brain mechanisms that qualitatively differ from those involved in evaluating the likelihood of these arguments in children.

1. Introduction

Deductive reasoning describes the ability to infer logically valid conclusions from prior information. For example, in the context of conditional rules, deductive reasoning allows one to draw the conclusion "The baby will start crying" from the premises "If a baby is hungry then she will start crying" and "The baby is hungry" (an inference termed *Modus Ponens* or *MP*). Not only is the ability to make such conditional deductions at the heart of scientific and mathematical thinking (Michal and Ruhama, 2008), but impairments in deductive reasoning are observed in children with math learning disability (Morsanyi et al., 2013). Therefore, understanding the neuro-cognitive mechanisms enabling the emergence of deductive reasoning in children is important from theoretical, clinical, and educational perspectives.

Although there is evidence that simple inferences such as Modus Ponens are made relatively early in development (i.e., as early as in kindergarten; Hawkins et al., 1984; Byrnes and Overton, 1986), studies also indicate that young children's deductive behavior is limited. For instance, those children usually fail to detect that the conclusion "The baby is hungry" cannot be logically drawn from the premises "If a baby is hungry then she will start crying" and "The baby starts crying" (an inference termed *Affirmation of the Consequent* or *AC*). This is because there may be other reasons leading to a baby crying, and the conclusion "The baby is hungry" does not follow out of necessity (e.g., the baby may be too cold). Increased ability to reject that AC form is usually observed in young adolescents, at least when premises have a concrete content (Markovits and Vachon, 1990; Barrouillet et al., 2002). Therefore, and although rejecting the AC form remains difficult even for educated adults (Cummins et al., 1991; Markovits and Doyon, 2004), increased ability to reject the AC form is often considered a hallmark of the emergence of deductive reasoning in children.

Over the past decades, studies have found that deductive reasoning

* Corresponding author. Institut des Sciences Cognitives Marc Jeannerod, UMR 5304, Centre National de la Recherche Scientifique (CNRS) & Université de Lyon, 67 Boulevard Pinel, 69675, Bron cedex, France.

** Corresponding author. Institut des Sciences Cognitives Marc Jeannerod, UMR 5304, Centre National de la Recherche Scientifique (CNRS) & Université de Lyon, 67 Boulevard Pinel, 69675, Bron cedex, France.

E-mail addresses: schwartz@isc.cnrs.fr (F. Schwartz), jprado@isc.cnrs.fr (J. Prado).

https://doi.org/10.1016/j.neuroimage.2017.09.029
Received 9 May 2017; Accepted 14 September 2017
Available online 19 September 2017

primarily engages frontal and parietal brain regions in adults (for a review, see Prado et al., 2011). Yet, to our knowledge, only one prior neuroimaging study has investigated deductive reasoning in children (Mathieu et al., 2015), and that study did not differentiate between children who exhibited accurate deductive performance and those who did not. Thus, the brain mechanisms that underlie the emergence of deductive reasoning in children remain unknown. The present functional magnetic resonance imaging (fMRI) study had two main goals. The first one was to identify the brain regions in which activity differs between children who reject the AC form and those who do not, thereby shedding light on the brain regions that underlie the emergence of deductive reasoning with conditional rules in children. We presented children between 8 and 13 with conditional problems of the AC form in a fMRI scanner (note that this relatively wide age range was chosen so that a relatively large variability in deductive responses could be observed). We asked them to indicate whether conclusions followed out of necessity from the premises (i.e., *Validity* trials, see Table 1). Activity during the evaluation of the AC form was systematically compared to a baseline in which children evaluated conclusions of more simple problems of the MP form, which children older than 8 should uniformly endorse (Markovits et al., 1996; Janveau-Brennan and Markovits, 1999). This was done to ensure that reasoning activity during evaluation of the AC form was isolated from activity related to reading a conditional rule and activity associated with selecting between two responses (both of these components were similar in AC and MP forms, see Monti et al., 2007 for a discussion of baseline issues in neuroimaging studies of reasoning). Activity during the evaluation of the AC form (compared to the MP form) was then related to rate of acceptance of the AC form across subjects, thereby identifying the brain regions underlying the emergence of deductive reasoning with conditional rules (i.e., the brain regions in which activity increased as rate of acceptance decreased).

The second goal of our study was to shed light on a debate between *one-process* and *two-process* theories about the nature of the mechanisms allowing for the emergence of deductive reasoning in children. On the one hand, one-process theories assume that deductive reasoning is an extension of the type of inductive reasoning used in everyday life, i.e., when one infers conclusions that are more or less likely given prior information and knowledge (Heit and Rotello, 2010; Rotello and Heit, 2009). For example, proponents of the influential Mental Model theory of deductive reasoning – which assumes that reasoners construct spatial mental representations of the premises when drawing deductive inferences – have argued that the same type of mechanisms underlie deductive and inductive reasoning (Johnson-Laird, 1994). More recently, researchers have proposed Bayesian accounts of deductive reasoning that assume that judgments of logical validity are determined by judgments of likelihood (Oaksford et al., 2000; Oberauer, 2006). For instance, consider the problem "If a baby is hungry then she will start crying; The baby starts crying; Is the baby hungry?". Bayesian theories posit that reasoners may intuitively calculate the likelihood of a baby being hungry given that she starts crying. This likelihood is relatively high, so the conclusion that the baby is hungry might be drawn by some individuals when faced with the premise "The baby starts crying". However, likelihood estimates vary between individuals and will change with experience (Evans and Over, 2013; Evans et al., 2015; Oaksford and Chater, 2001). Thus, Bayesian theories can explain why rejection rates of the AC form varies between individuals and typically increase over development: Because older children have a broader knowledge base than younger children, a given premise (e.g., "The baby starts crying") is more likely to evoke multiple associated causes in older than younger children (e.g., "The baby is hungry, "The baby is cold" etc.). This will lower the certainty of the conclusion. Overall, then, Bayesian theories assume that judgments of likelihood may translate into likelihood of rejection of a deductive argument, either directly (e.g., a conclusion associated with a probability of 60% might be accepted about 60% of the time; Liu and Song, 2003), or indirectly using an internal threshold (i.e., a conclusion may only be accepted if the associated probability is above a certain threshold; Oberauer, 2006). In sum, one-process theories do not see any major qualitative difference between evaluating the logical validity and the likelihood of a conclusion.

On the other hand, two-process theories assume that the mechanisms supporting judgments of logical validity are different from those underlying judgments of likelihood. For instance, the Mental Logic theory posits that deductive reasoning relies on formal rules of inference that are specific to logic and therefore cannot account for inductive reasoning (Braine and O'Brien, 1998). A developmental variant of the Mental Model theory proposed by Markovits and Barrouillet (2001) also makes a distinction between evaluating the logical validity and the likelihood of conclusions in children. Specifically, this model emphasizes that deductive reasoning relies on the retrieval of relevant knowledge stored in long-term memory. For instance, in the problem mentioned earlier (i.e., "If a baby is hungry then she will start crying; The baby starts crying; Is the baby hungry?"), reasoners may search for an alternate cause leading to a baby crying (i.e., other than the baby being hungry). If at least one alternative is found (e.g., a baby may cry because she is too cold), the conclusion that the baby is hungry will not be made. Thus, unlike Bayesian theories, this theory does not posit that reasoners intuitively compute likelihoods when assessing conclusions. Rather, they may search for counterexamples and a conclusion will be rejected if at least one counterexample is found. Clearly, the representation and maintenance of such counterexamples relies on working-memory resources (Markovits and Doyon, 2004). Therefore, a developmental increase in working-memory capacity may be at the heart of the increased ability to reject the AC form with age (Markovits and Barrouillet, 2001; Barrouillet and Lecas, 1999; De Neys and Everaerts, 2008). Finally, the idea that judgments of logical validity and likelihood are different is broadly consistent with dual-system theories of reasoning, which posit that two types of cognitive processing underlie human reasoning (Evans and Stanovich, 2013). The first type (often referred to as 'heuristic' or 'intuitive') is fast, unconscious, and autonomous. The second type of processing (often referred to as 'analytical' or 'deliberate') is slow, conscious, and controlled. It has been proposed that judgments of likelihood, which rely on associative information and similarity, are more likely to involve heuristic than analytic processing. In contrast, judgments of logical validity, which require deliberative and accurate reasoning, are more likely to rely on the analytic than the heuristic processing (Heit and Rotello, 2010; Rotello and Heit, 2009; Heit, 2014). In sum, two-process theories posit that there is a difference between evaluating the logical validity and the likelihood of a conclusion, either because these processes rely on entirely separate mechanisms or on a different mixture of heuristic and analytic processing.

In the present study, in addition to identifying the brain regions underlying the emergence of deductive reasoning with conditional rules, we aimed to shed some light on the debate between one-process and two-process theories. That is, we tested whether the brain regions that underlie the emergence of deductive reasoning overlap with the brain circuits associated with judging the likelihood of conclusions. This was done by presenting children – in a second part of the experiment – with conditional problems of the AC form and asking them to indicate on a scale the likelihood of the conclusion (i.e., *Likelihood* trials) (Markovits and Thompson, 2008). Activity during the evaluation of the AC form

Table 1
Examples of AC and MP forms in *Validity* and *Likelihood* trials for the rule "If a baby is hungry then she will start crying".

Logical form	Premise	Question
Validity trial		
MP	A baby is hungry	Is it certain that she will start crying?
AC	A baby starts crying	Is it certain that she is hungry?
Likelihood trial		
MP	A baby is hungry	How sure it is that she will start crying?
AC	A baby starts crying	How sure it is that she is hungry?

Notes: MP: Modus Ponens, AC: Affirmation of the Consequent.

(compared to the MP form) was then related to likelihood rating of the AC form across subjects. Brain regions in which activity increased as likelihood rating of the AC form decreased in *Likelihood* trials were then compared to brain regions in which activity increased as judgments of logical validity of the AC form decreased in *Validity* trials.

2. Material and methods

2.1. Participants

Twenty-seven typically developing children from 8 to 13 participated in the experiment. All children were native French speakers, had no MRI counter-indications, had no history of neurological and psychiatric disorder, and were right handed. All participants had a full-scale IQ above 80 (mean: 113, standard deviation (sd): 11) as measured by the NEMI 2 (Nouvelle Echelle Metrique de l'Intelligence; Cognet, 2006). Parents gave their written informed consent and children gave their assent to participate in the experiment. Families were paid 80 € for their participation. The experiment was approved by the local ethics committee (CPP Lyon Sud-Est II). Ten participants were excluded from the analyses because of (1) withdrawal from the experiment before scanning (n = 1), (2) technical issues during the scanning session (n = 2), (3) poor brain coverage (n = 2), or (4) excessive head movements (n = 5). Thus, 17 participants (7 females) from 8.2 to 13.7 were included in the analyses (mean age: 10.8, sd: 1.7).

2.2. Task

In the scanner, participants were presented with a conditional reasoning task in which they evaluated the conclusion of an argument either based on "validity" or "likelihood" instructions. The task was adapted from Markovits and Handley (2005) and Markovits and Thompson (2008). It involved the presentation of causal conditional rules (e.g., "If a baby is hungry then he will start crying"), followed by a premise (e.g., "A baby starts crying"). This premise was systematically followed by a question whose format differed as a function of type of instructions. In *Validity* trials, the question always started with "Is it certain that … ?" (e.g., "Is it certain that she is hungry?") and participants had to answer yes or no. In *Likelihood* trials, the question always started with "How sure is it that … ?" (e.g., "How sure is it that she is hungry?") and participants had to move a cursor along a horizontal scale to indicate the degree of likelihood of the conclusion (see Fig. 1). There were 5 possible degrees of likelihood, represented from left to right on the scale (not sure at all, not very sure, medium sure, very sure, very very sure). The initial position of the cursor was always at the center of the scale.

Overall, 24 different causal rules were presented to participants. The exact same rules were used in *Validity* and *Likelihood* trials. Sixteen rules were taken from Grosset et al. (2005). Following Grosset et al. (2005), rules varied with regard to the strength of association between cause and consequence. In half of the rules, the cause was strongly associated with the consequence, such that individuals frequently think of that specific cause in relation to the consequence (e.g., "If a baby is hungry then she will start crying") (Grosset et al., 2005). In the other half, the cause was more weakly associated with the consequence, such that participants less frequently evoke that specific cause in relation to the consequence (e.g., "If a baby is too cold then she will start crying"). Frequencies of report of the cause given the consequence for these 16 rules are given in Grosset et al. (2005). To increase the number of trials and maximize signal-to-noise ratio in the fMRI scanner, we created 8 additional rules similar to those in Grosset et al. (2005). Four of these rules had a strong association between the cause and the consequence and four had a weaker association. These 8 new rules were pretested on another group of children before the experiment. Across all 24 rules, the same consequence was presented with two different causes (one strongly and one weakly associated) over the course of the experiment.

Each rule was presented with an AC form (If P then Q; Q) and with a MP form (If P then Q; P). Examples of AC and MP forms in *Validity* and *Likelihood* trials are given in Table 1. Overall, there were 18 *Validity* trials with the AC form, 18 *Validity* trials with the MP form, 18 *Likelihood* trials with the AC form, and 18 *Likelihood* trials with the MP form. In 24 additional trials (12 *Validity* and 12 *Likelihood*), premises negated the cause or the consequence of the rule (e.g., If a baby is hungry then she will start crying; A baby is not crying). This was done to introduce some variability in the task and discourage participants from developing expectations. These 24 trials were considered fillers and were not further analyzed.

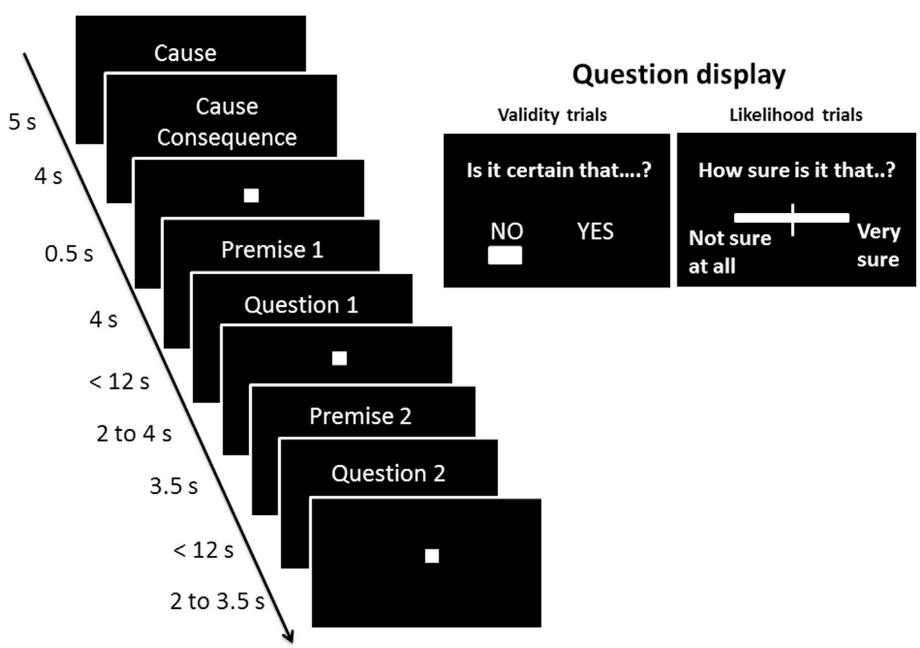

Fig. 1. Experimental procedure. A conditional rule was displayed (e.g., "If a baby is hungry then she will start crying"), followed by a premise (e.g., "The baby starts crying"). This premise was systematically followed by a question whose format depended upon the type of instruction (Validity or Likelihood). Each rule was always followed by two premises (and their associated questions).

The task was presented in 6 runs. A single run was entirely composed of either *Validity* or *Likelihood* trials to avoid interferences between instructions. Half of the subjects started with 3 runs of *Validity* trials before being presented with 3 runs of *Likelihood* trials, whereas the other half started with 3 runs of *Likelihood* trials before being presented with 3 runs of *Validity* trials. To maximize design efficiency and minimize the time spent in the scanner, a single rule was always followed by two different premises and their associated questions, leading to two trials per rule (see Fig. 1). There were 8 rules per run, which corresponded to 16 trials. Trials were presented in a pseudo-randomized order, such that the number of AC and MP forms (and the number of weak and strong causes) was counterbalanced within and between runs. The logical forms AC and MP (and the strength of associations between cause and consequence in the rules) were also counterbalanced between the first and the second trials of each rule for each run. Four different scenarios were generated that met the requirements listed above.

2.3. Stimulus timing

Stimuli were generated with Presentation software (Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). In each problem, the cause and the consequence of the rule appeared on the screen one at a time (cause at 0 s and consequence at 5 s). After 4 s, the rule was replaced by a central fixation dot for 1 s. A premise was then presented for 4 s. This was immediately followed by the first question (whose form differed under “validity” and “likelihood” instruction). The question disappeared with the participant's response or after 12 s if no response was provided. A variable fixation period ranging from 2 to 4 s followed the disappearance of the question. Another premise was then presented for 3.5 s. This was immediately followed by the second question, which disappeared with the participant's response or after 12 s if no response was provided. A variable fixation period ranging from 2 to 3.5 s followed the disappearance of the question. All sentences were additionally spoken on headphones to facilitate comprehension. Responses were recorded using MRI compatible response buttons. The experimental procedure is illustrated in Fig. 1.

2.4. Mock scanner practice

Children were familiarized with the MRI environment in a mock scanner prior the actual fMRI session. The time between the training session and the fMRI session did not exceed 4 weeks. A motion tracker system (3D Guidance trak STAR, Ascension Technology Corporation) was used to measure head movements and provide feedback to participants. Children also heard a recording of the noises associated with all fMRI sequences. Children practiced the task in the mock scanner. During that practice session, children were presented with 4 trials (the rules differed from the rules used in the main experiment).

Participants received the following instructions to complete the task: “You will be presented with sentences. You will be able to read them and to listen to them. You have to consider that what is in the sentences is always true. For example, you will read the following sentence: “If the sky is dark then it will start raining”. It means that you have to consider that every time the sky is dark, it will start raining. Then, you will have to answer a question. Sometimes, the question will be like “The sky is dark. Is it certain that it will start raining? ”. If you are certain that it will start raining, you have to respond “yes”. Otherwise, you have to respond “no”. Sometimes, the question will be “The sky is dark. How sure is it that it will start raining?”. In this case, you will have to tell us how sure you are that it will start raining. A line will then appear on the screen. The more sure you are, the more you have to move the cursor on the right. If you are quite sure, you move once on the right. If you are very sure, you move twice on the right. The less sure you are, the more you have to move the cursor on the left. If you are not very sure, you move once on the left. If you are not sure at all, you move twice on the left. If you are medium sure, you can let the cursor at the center.” Each child performed two trials with the experimenter (one *Validity* and one *Likelihood* instruction) and two trials alone in the mock scanner. No performance feedback and no further instruction were given after training.

2.5. Imaging acquisition

Functional and anatomical images were acquired with a Philips 3T Achieva scanner (Philips Medical systems, Best, Netherlands) at Pierre Wertheimer Neurologic Hospital in Lyon. A high-resolution anatomical scan was collected for each participant (Field of view = 240 × 240 mm, 512 × 512 matrix, TR = 1278 ms, TE = 3.047 ms, flip angle = 8°, slice thickness = 0.9 mm, number of slices = 188). Functional sequences were collected with a gradient-echo, echo-planar sequence (TR = 2200 ms, TE = 30 ms, flip angle = 90°). Twenty-six axial slices were acquired per volume (slice thickness = 3.5 mm, Field of view = 220 mm, 128 × 128 matrix).

2.6. fMRI data preprocessing

Images were analyzed with SPM12 (Statistical Parametric Map 12, Welcome department of Cognitive Neurology, London, UK). The first 6 images of each run were not acquired to allow for T1 equilibration effects. Functional images were corrected for slice acquisition delays and realigned to the first image of the first run to correct for head movements. Volumes with excessive head motion were further identified using the ArtRepair software (Mazaika et al., 2009). Specifically, volumes showing rapid scan-to-scan movement greater than 1.5 mm were substituted by the interpolation of the 2 nearest non-repaired volumes. Runs with more than 15% of repaired volumes were excluded from the analysis (based on these criteria, one run was discarded for three participants). As recommended by the ArtRepair developers (http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html), realigned images were spatially smoothed with a Gaussian filter prior application of the ArtRepair algorithms. Because voxels were anisotropic, this filter was also anisotropic (4 × 4 × 8 mm full-width at half maximum). Finally, in order to compare brain activity across grades, all individual brains needed to be normalized into the same stereotactic space. This always constitutes a challenge because brain anatomy changes over development (Wilke et al., 2002). Critically, however, anatomical differences between children older than 7-8 year-olds and adults are small enough that they are beyond the resolution of fMRI experiments (Burgund et al., 2002; Kang et al., 2003). Therefore, considering the age of our participants and the resolution of our data, we decided to normalize all individual brains into the standard adult MNI space. This also allowed us to compare the results of the present study to the results of previous research on adults. This normalization was done in 2 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, functional data were normalized to the MNI space by using the normalization parameters estimated during unified segmentation. The quality of the normalization was verified in each participant by visually checking the registration and ensuring an adequate correspondence between each individual's brain and the MNI template.

2.7. fMRI data processing

Statistical analyses of fMRI data were performed according to the GLM. For both *Validity* and *Likelihood* trials, activation associated with each logical form was modeled as epochs with onsets time locked to the presentation of the premise and with duration matched to the offset of the question. In total, 6 different regressors were created to take into account the logical form and the strength of the association between the cause and the consequence (AC with strongly associated causes, AC with weakly associated causes, MP with strongly associated causes, MP with weakly associated causes, fillers with strongly associated causes, fillers with weakly associated causes). Two additional regressors of no interest modeled the rules themselves (one regressor for rules with a weak cause and one for rules with a strong cause) and one regressor modeled the average signal in the run. Therefore, there were 9 regressors in each run. Because a behavioral response was not recorded (or given in less than 0.5 s) in only 2% of the trials, all trials were included in the analyses. This model was applied to each participant. Finally, time series were high-pass filtered (1/128 Hz) and serial correlations were corrected for first-order (AR1) autocorrelations.

For each subject and trial type (*Validity* and *Likelihood*), brain activity associated with the AC form was contrasted to brain activity associated with the MP form. These subject-level contrasts were then entered into 2nd-level voxelwise regression analyses in which performance associated with the AC form (acceptance rate in *Validity* trials and likelihood rating in *Likelihood* trials) constituted the regressor of interest. In addition, we included age as a regressor of no interest in all 2nd-level models given that age interacted with some aspects of behavioral performance (see Supplemental Information). A FWE-corrected cluster-level threshold of $p = 0.05$ (defined using a voxel-level threshold of $p = 0.001$) was applied to all whole-brain statistical maps to assess brain activations.

2.8. Region of interest analyses

Region of interest (ROI) analyses were conducted using the SPM toolbox Marsbar (http://marsbar.sourceforge.net/). ROIs included all voxels within a 6-mm radius of each coordinate identified in (1) the contrast showing increases of brain activity associated with decreases in acceptance of the (strongly related) AC form in Validity trials and (2) the contrast showing increases in brain activity associated with decreases in likelihood rating of the (strongly related) AC form in Likelihood trials. For each participant, we calculated the average activity for each trial type within an ROI by averaging the fMRI signal across all voxels within that ROI. Importantly, statistical analyses of ROI activity are only performed when coordinates of interest are defined in an independent contrast. That is, activity associated with rules with a weak association between cause and consequence was analyzed in ROIs defined using rules with a strong association between cause and consequence. Similarly, activity associated with Likelihood trials was analyzed in ROIs defined using Validity trials (and vice versa).

3. Results

3.1. Validity and likelihood judgments are related at the behavioral level

An assumption of the study is that AC arguments are associated with a relatively large variability in performance between subjects. As shown on Fig. 2A (left), the AC form was indeed associated with large inter-individual differences in response rates in *Validity* trials: While some children rejected that form most of the time, others accepted it most of the time. Fig. 2B (left) indicates a similar variability in likelihood ratings of the AC form in *Likelihood* trials: While some children judged that form as unlikely most of the time, others judged it as likely most of the time. These patterns contrasted with performance associated with the MP form. That is, the large majority of children accepted the MP form most of the time in *Validity* trials (see Fig. 2A, right). Similarly, the majority of children judged the MP form as likely most of the time in *Likelihood* trials (see Fig. 2B, right).

We then tested whether there was a relationship between accepting the AC form in *Validity* trials and rating the AC form as likely in *Likelihood* trials across subjects. The correlation was significant ($r^2 = 0.49$, $p = 0.002$) (see Fig. 2C). Thus, those participants who were willing to reject the AC form in *Validity* trials were also those who were rating the AC form as unlikely in *Likelihood* trials. However, this correlation was not specific to the AC form. That is, the relationship between accepting the MP form in *Validity* trials and rating the MP form as likely in *Likelihood* trials across subjects was also highly significant ($r^2 = 0.46$, $p = 0.003$) (see Fig. 2D). Additional analysis of behavioral performance (not directly related to our hypotheses) can be found in Supplemental Information.

3.2. Left and medial frontal regions underlie the emergence of deductive reasoning in rules with strongly associated cause

In *Validity* trials, regions underlying the emergence of deductive reasoning were identified by comparing brain activity associated with the AC form in children who were rejecting that form most of the time versus in children who were accepting that form most of the time (with children providing mixed responses in the middle of this continuum). That is, we linearly related acceptance rate of the AC form to brain activity associated with evaluating that form across subjects. Activity associated with the MP form, for which relatively little variation in performance was observed (see Fig. 2A), was used as baseline. We failed to find any regions in which acceptance rate of the AC form was related to fMRI activity. To evaluate whether the strength of association between cause and consequence affected this (lack of) result, we then conducted separate analyses on rules with strong and weak associations between cause and consequence. For rules with a strong association between cause and consequence, we found that a decrease in acceptance rate of the AC form was related to an increase in activity in left frontal regions of the Middle Frontal Gyrus (MFG) and Precentral Gyrus (PreG). Increased activity was also observed in the dorso-medial Prefrontal Cortex (dmPFC) (see Table 2 and Fig. 3). There were no regions in which a decrease in acceptance rate of the AC form was related to a decrease in activity across subjects.

Surprisingly, these results were specific to rules with a strong association between cause and consequence. That is, for rules with a weak association between cause and consequence, we failed to detect any regions in which a decrease in acceptance rate of the AC form was related to either an increase or a decrease in activity across subjects. Similarly, no negative relationship between brain activity and acceptance rate of the AC form could be observed for rules with a weak association between cause and consequence in the 3 regions that we identified using rules with a strong association (left MFG: $r^2 = 0.016$, $p = 0.31$; left PreG: $r^2 = 0.018$, $p = 0.30$; dmPFC: $r^2 = 0.019$, $p = 0.30$, one-tailed). The possible reasons for this null finding are discussed in the general discussion.

3.3. Right and medial frontal as well as bilateral parietal regions underlie differences in likelihood ratings for rules with strongly associated cause

In *Likelihood* trials, we compared brain activity associated with the AC form in children who were judging that form as unlikely most of the time versus in children who were judging that form as likely most of the time (with children providing mixed responses in the middle of this continuum). That is, we linearly related likelihood rating of the AC form to brain activity associated with that form across subjects. Activity associated with the MP form, for which relatively little variation in rating was observed (see Fig. 2B), was used as baseline. We failed to find any regions in which likelihood rating of the AC form was related to fMRI activity when all rules were considered. We then conducted separate analyses on rules with strong and weak associations between cause and consequence. For rules with a strong association between cause and consequence, we found that a decrease in likelihood rating of the AC form was related to an increase in activity in right lateralized regions of the Inferior Frontal Gyrus (IFG) and MFG. Increased activity was also observed in the left and right dmPFC, as well as in the parietal cortex at the level of the right Postcentral Gyrus, right Inferior Parietal Lobe (IPL), left Precuneus, left Angular Gyrus (AG), and left IPL (see Table 3 and Fig. 4). There were no regions in which a decrease in likelihood rating of the AC form was related to a decrease in activity across subjects.

Finally, for rules with a weak association between cause and consequence, whole-brain analyses revealed no region in which a decrease in likelihood rating of the AC form was related to either an increase or a

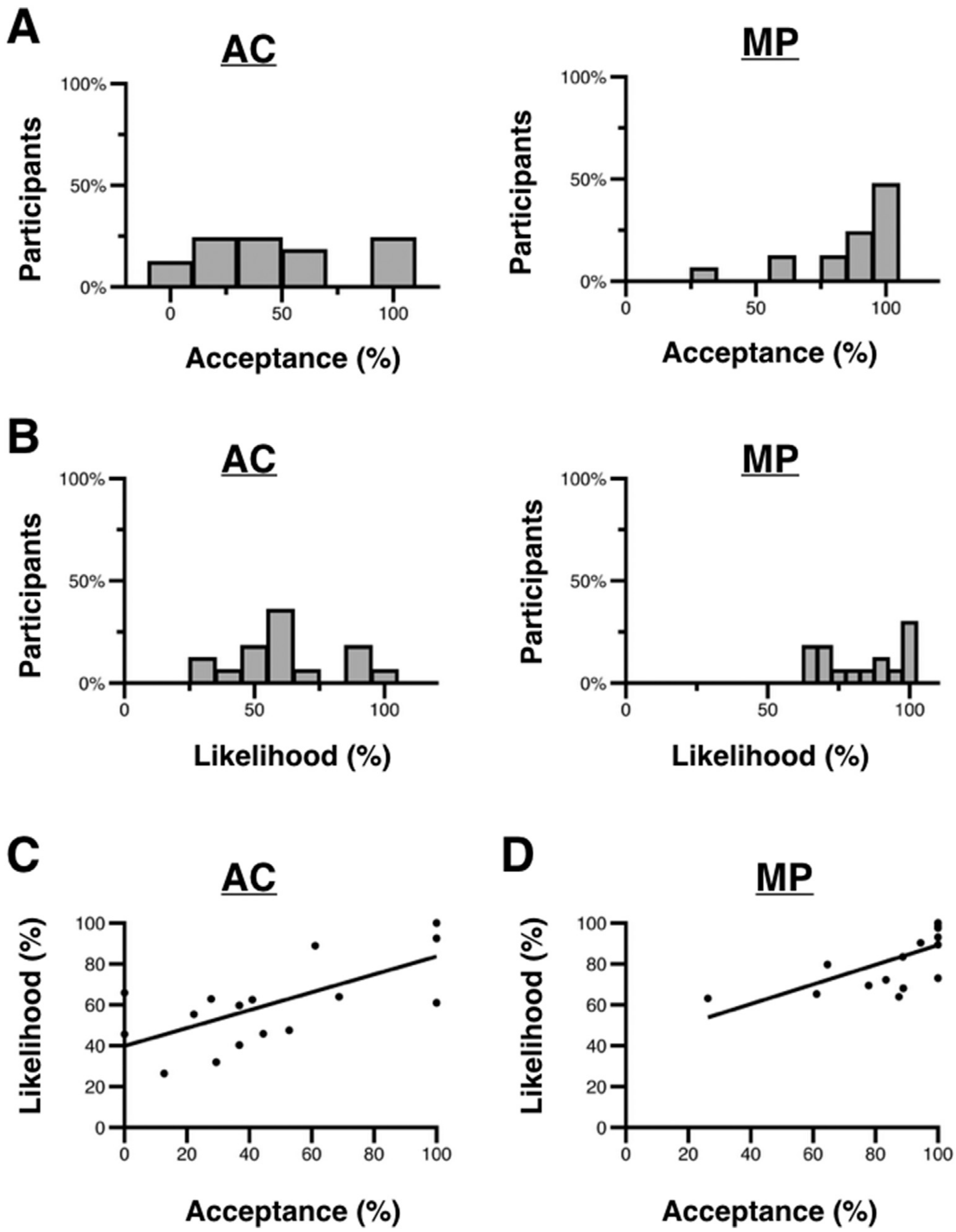

Fig. 2. Behavioral performance. (A) Frequency histograms of acceptance rates across participants for Affirmation of the Consequent (AC) and Modus Ponens (MP) in *Validity* trials. (B) Frequency histograms of likelihood ratings across participants for AC and MP in Likelihood trials. (C) Relationship between likelihood rating in *Likelihood* trials and acceptance rate in *Validity* trials for AC. (D) Relationship between likelihood rating in *Likelihood* trials and acceptance rate in *Validity* trials for MP.

Table 2
Brain regions in which a decrease in acceptance rate of the AC form was related to an increase of activity for that form (as compared to the MP form) in *Validity* trials.

Anatomical location	~BA	MNI coordinates			Z-score	Cluster size (mm3)
		X	Y	Z		
Rules with a strong association between cause and consequence						
L. MFG	9	−38	28	26	4.67	1136
L. PreG	4	−30	−10	42	3.96	1584
dmPFC	32	0	8	50	4.46	992
Rules with a weak association between cause and consequence						
No suprathreshold clusters						

BA: Brodmann area, MNI: Montreal Neurological Institute.

decrease in activity across subjects. However, for these rules, we did find a negative relationship between activity and likelihood rating of the AC form in a region that was identified using rules with a strong association, the left Precuneus ($r^2 = 0.19$, $p = 0.04$, one-tailed) (see Fig. 5B). Therefore, activity in this region appears to be associated with inter-individual differences in likelihood rating whether the rule has a strong or a weak association between cause and consequence.

3.4. The brain regions that are involved in likelihood and validity judgments differ

Finally, we tested whether there was any overlap among the brain regions that are involved in assessing the likelihood of a conclusion and those that underlie validity judgments. We tested that overlap in two ways. First, using only rules with a strong association between cause and consequence (as no whole-brain difference could be observed for rules with a weak association), we performed a Boolean intersection (corresponding to a conjunction null hypothesis in the case of two contrasts, Nichols et al., 2005) between (1) the brain regions in which activity

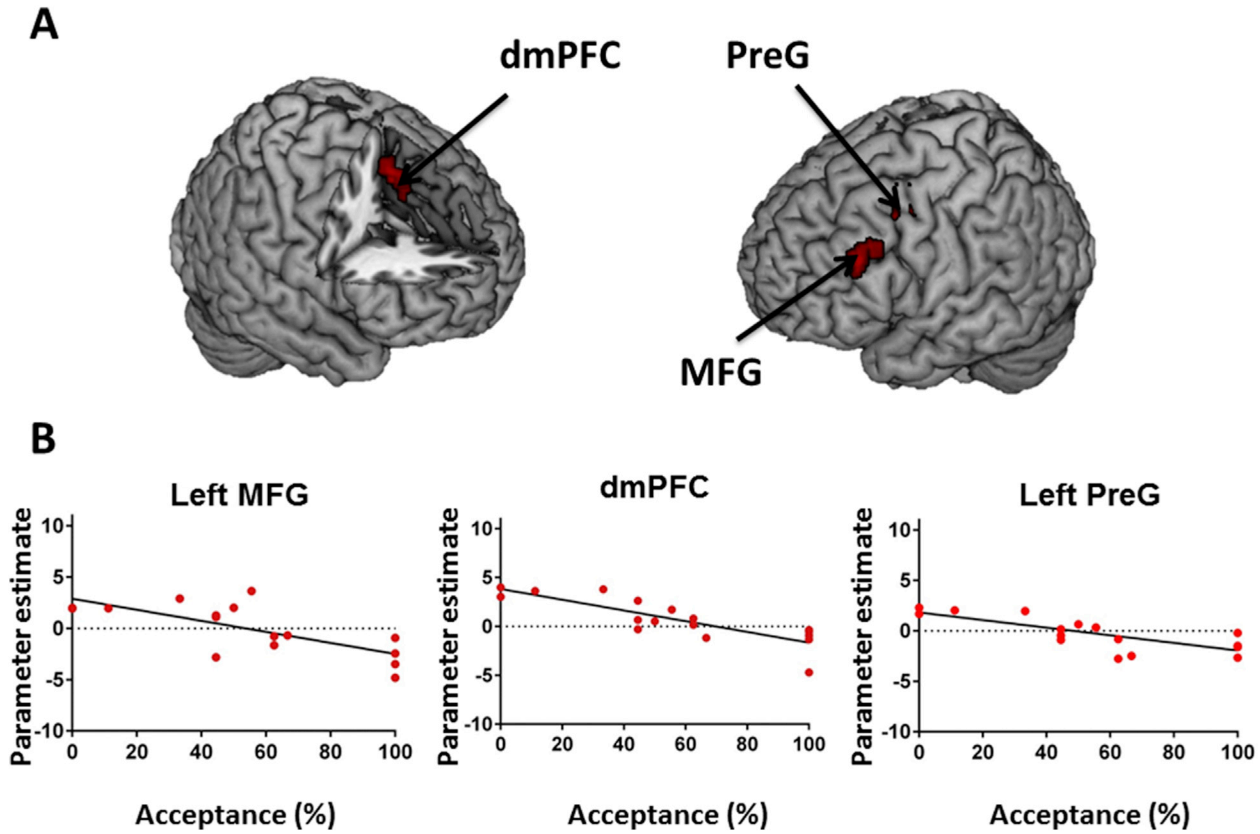

Fig. 3. For rules with a strong association between cause and consequence, brain regions in which a decrease in acceptance rate of the AC form was related to an increase of activity for that form (as compared to the MP form) in *Validity* trials. (A) Activations overlaid on a 3D rendering of the MNI-normalized anatomical brain. (B) For each activated cluster, scatter plot of brain activity (parameter estimate) associated with the AC form (compared to the MP form) as a function of acceptance rate of the AC form. Note that plots are based on average activity around coordinates of interest identified in a non-independent contrast (i.e., the whole-brain regression of acceptance rate of the AC form on brain activity associated with that form). Thus, effect sizes might appear inflated and plots are only displayed for illustration purpose.

Table 3
Brain regions in which a decrease in likelihood rating of the AC form was related to an increase of activity for that form (as compared to the MP form) in *Likelihood* trials.

Anatomical location	~BA	MNI coordinates			Z-score	Cluster size (mm3)
		X	Y	Z		
Rules with a strong association between cause and consequence						
L. Precuneus	7	−12	−74	46	4.61	2128
L. dmPFC	32	−12	22	38	4.61	1072
R. dmPFC	32	6	6	42	4.30	5024
L. Angular Gyrus	39	−30	−62	34	4.05	1952
L. Inferior Parietal Lobe	40	−34	−42	42	4.04	1360
R. Postcentral Gyrus	2	54	−30	46	4.80	2560
R. Inferior Parietal Lobe	40	42	−40	46	4.30	7216
R. Precentral Gyrus	6	50	4	34	4.06	2544
R. Inferior Frontal Gyrus	9	44	22	26	3.78	1056
Rules with a weak association between cause and consequence						
No suprathreshold clusters						

BA: Brodmann area, MNI: Montreal Neurological Institute.

increased with a decrease in acceptance rate of the AC form in *Validity* trials and (2) the brain regions in which activity increased with a decrease in likelihood rating of the AC form in *Likelihood* trials. Each contrast was thresholded using a FWE-corrected cluster-level threshold of $p = 0.05$ (defined using a voxel-level threshold of $p = 0.001$). We did not find any voxels in common between these two sets of regions (see Fig. 6).

Second, using an ROI approach, we tested whether a relationship between activity and acceptance rate of the AC form in *Validity* trials could be observed in any of the regions that showed a relationship between activity and likelihood rating of the AC form in *Likelihood* instruction. First, we focused on regions in which variations of activity could be observed for rules with a strong association between cause and consequence. We did not find any significant relationship between activity and acceptance rate of the AC form in *Validity* trials in any of the ROIs that showed a relationship between activity and likelihood rating in *Likelihood* trials (all ps > 0.12). Likewise, a relationship between activity and likelihood rating of the AC form in *Likelihood* trials was observed in none of the regions that showed a relationship between activity and acceptance rate of the AC form in *Validity* trials (all ps > 0.46). Second, we focused on the only region in which we found variations of activity for rules with a weak association between cause and consequence: the left Precuneus. In that region, although there was a relationship between activity and likelihood rating of the AC form for rules with a weak association in *Likelihood* trials (see Fig. 5B), there was no relationship between activity and acceptance rate of the AC form in deductive trials ($r^2 = 0.04$, $p = 0.21$, one-tailed) (see Fig. 5A). Therefore, the brain regions that are involved in assessing the likelihood of a conclusion do not appear to overlap with the brain regions underlying validity judgments, both for rules with a strong and weak association between cause and consequence.

4. Discussion

The goal of the present study was twofold. First, we aimed to identify the neural mechanisms that underlie the emergence of deductive reasoning with conditional rules in children. Second, we tested whether these mechanisms overlap with the neural mechanisms involved in judging the likelihood of conclusions.

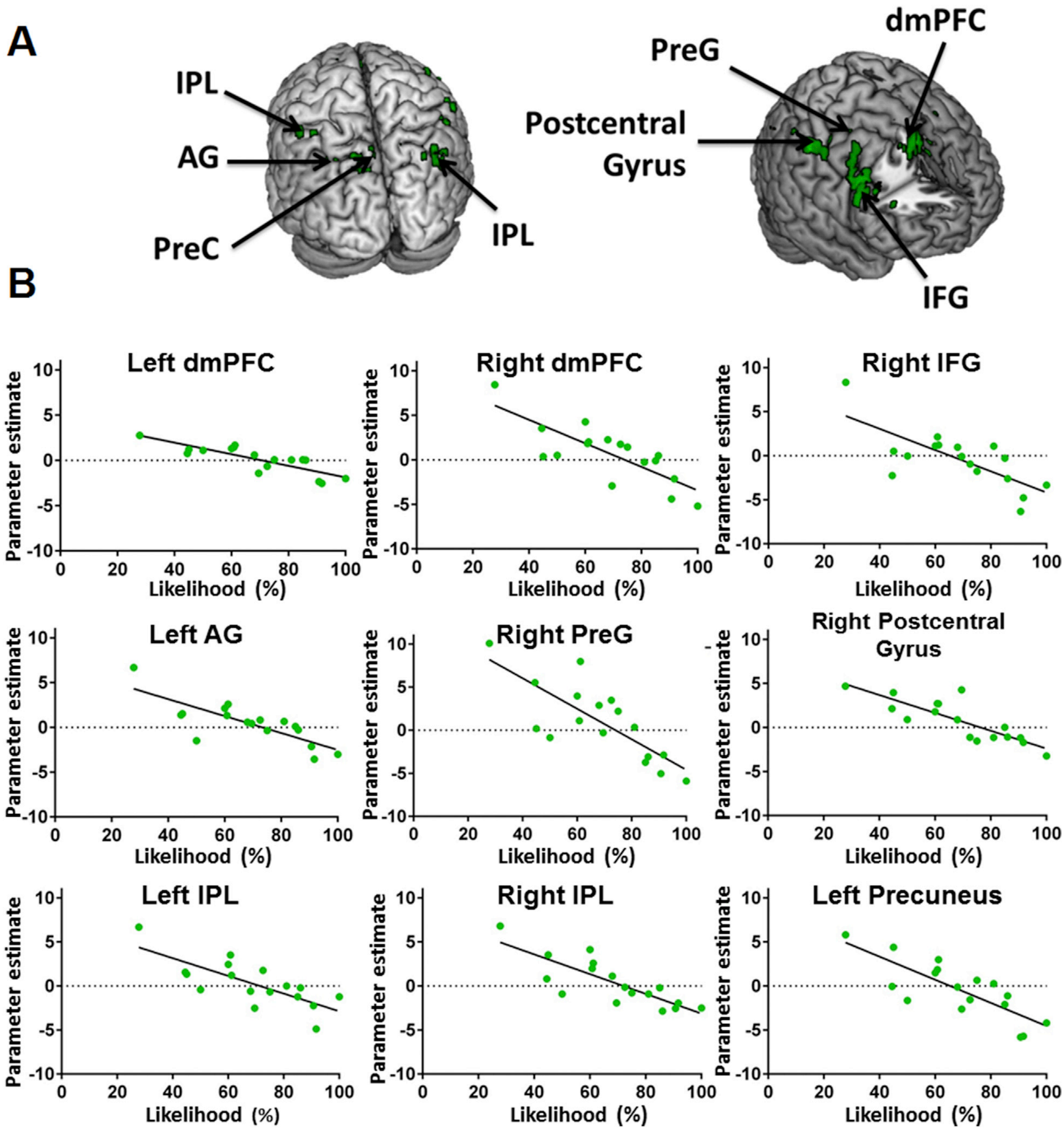

Fig. 4. For rules with a strong association between cause and consequence, brain regions in which a decrease in likelihood rating of the AC form was related to an increase of activity for that form (as compared to the MP form) in *Likelihood* trials. (A) Activations overlaid on a 3D rendering of the MNI-normalized anatomical brain. (B) For each activated cluster, scatter plot of brain activity (parameter estimate) associated with the AC form (compared to the MP form) as a function of likelihood rating of the AC form. Note that plots are based on average activity around coordinates of interest identified in a non-independent contrast (i.e., the whole-brain regression of likelihood rating of the AC form on brain activity associated with that form). Thus, effect sizes might appear inflated and plots are only displayed for illustration purpose.

4.1. Left and medial frontal regions support the emergence of deductive reasoning with conditional rules in children

In *Validity* trials, enhanced activity in children who reject the AC form (compared to those who accept that form) was observed in three brain regions: the left MFG (BA9), the left PreG (BA4) and the dmPFC (BA32). Therefore, these regions appear to support the emergence of deductive behavior in children. Activity in the left lateral and medial frontal cortex is frequently reported in studies that investigate the neural bases of deductive reasoning in adults (Fangmeier et al., 2006; Goel et al., 2000; Goel and Dolan, 2004; Monti et al., 2007; Monti et al., 2009; Reverberi et al., 2010, 2012, 2007). For example, in a previous quantitative meta-analysis of the literature on the neural bases of deductive reasoning, we found that the left MFG, left PreG and dmPFC were among the regions that were the most consistently activated in studies employing deductive tasks (Prado et al., 2011). Given the known role of the lateral prefrontal cortex (including the MFG and PreG) in working memory (Owen et al., 2005; Rottschy et al., 2012), it has been proposed that this area might have a general cognitive support role during deductive reasoning (Monti et al., 2007, 2009). For example, a lesion study found that patients with lesions to the left lateral frontal cortex exhibited deductive reasoning impairments, but only when they also had working-memory deficits (Reverberi et al., 2009). Interestingly, developmental studies have demonstrated age-related increases of activity in the lateral frontal cortex

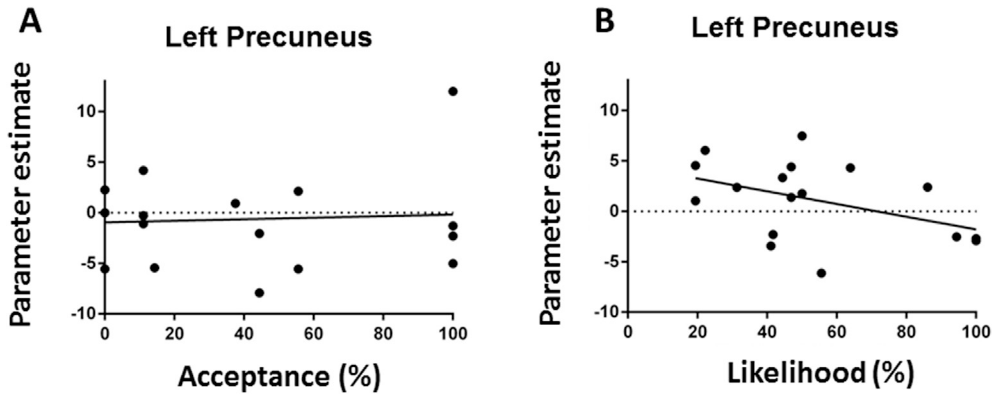

Fig. 5. For rules with a weak association between cause and consequence, relationship between left Precuneus activity and performance in *Validity* and *Likelihood* trials. (A) Scatter plot of brain activity (parameter estimate) associated with the AC form (compared to the MP form) as a function of acceptance rate of the AC form in *Validity* trials. (B) Scatter plot of brain activity (parameter estimate) associated with the AC form (compared to the MP form) as a function of likelihood rating of the AC form in *Likelihood* trials. Note that plots are based on average activity around coordinates of interest identified in an independent contrast (i.e., the whole-brain regression of likelihood rating of the AC form for strongly associated rules on brain activity associated with that form).

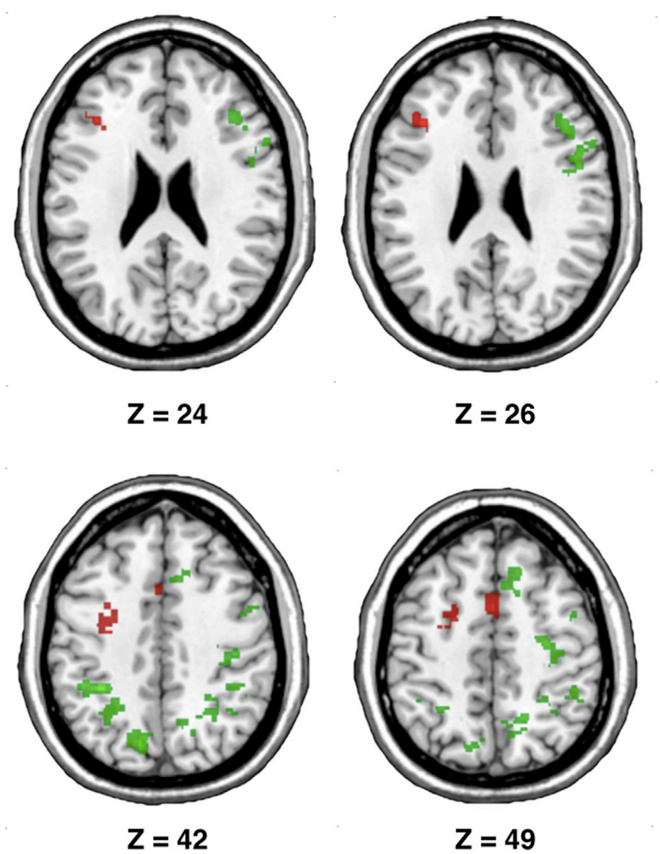

Fig. 6. Brain activity modulated by acceptance rate of the AC form in *Validity* trials (red) and likelihood rating of the AC form in *Likelihood* trials (green) for strongly associated rules, overlaid on the same axial slices of a MNI-normalized brain.

in working-memory tasks, and activity in these regions relates to working-memory abilities (Klingberg, 2006). Therefore, greater activity in the left MFG and PreG in children who exhibit deductive behavior (i.e., who reject the AC form) compared to those who do not (or less) exhibit that behavior (i.e., who accept the AC form) might reflect increased working-memory engagement. This would be consistent with counter-examples theories of reasoning (Markovits and Vachon, 1990; Markovits and Barrouillet, 2001; Barrouillet et al., 2002). However, given that (1) we did not independently localize the brain mechanisms involved in working-memory and that (2) the regions activated involve only a restricted portion of the brain system supporting working-memory (Owen et al., 2005), this claim remains speculative and needs to be confirmed by future studies.

Besides the left MFG and PreG, another region in which activity was greater for children who rejected the AC form than in children who accepted that form was the dmPFC (BA32). Lesions to this region have been shown to impair deductive reasoning performance in adults, even when working memory abilities are spared (Reverberi et al., 2009). In contrast to patients with lesions to the lateral prefrontal cortex, patients with a lesion to the dmPFC also have issues evaluating the difficulty of a deductive problem correctly (Reverberi et al., 2009). Therefore, it has been suggested that the dmPFC might be involved in the evaluation of task difficulty during reasoning (Reverberi et al., 2009) and more generally in performance monitoring (as reviewed by Ridderinkhof et al., 2004) and meta-cognitive operations (Reverberi et al., 2009).

4.2. *Different mechanisms for evaluating the logical validity and likelihood of conclusions in children*

In *Likelihood* trials, enhanced activity in children who judged the AC form as unlikely (compared to those who judged that form as likely) was observed in several brain regions encompassing the right frontal cortex and the bilateral parietal cortex. Although the specific role of these regions in our task remains unknown, a similar set of bilateral parietal and right frontal regions is consistently activated in studies that employ arguments involving relations of linear order (e.g., *Bill is taller than Sam*, *Sam is taller than Jim*, therefore *Bill is taller than Jim*) (Prado et al., 2011; Prado et al., 2013; Mathieu et al., 2015). We note that assessing the likelihood of a conclusion also requires participants to make a judgment of order, as reasoners must decide on where this likelihood falls on an internal continuum ordered from highly unlikely to highly likely. Most importantly, however, we did not find any overlap between the brain mechanisms that support rejecting the AC form in *Validity* trials (as compared to accepting that form) and the brain mechanisms that support judging the AC form as unlikely in *Likelihood* trials (as compared to judging that form as likely). Therefore, our fMRI results are inconsistent with the idea that deciding whether a conclusion logically follows from premises involves an assessment of likelihood and a direct or indirect conversion of such probability judgment onto a binary scale (Liu and Song, 2003; Oberauer, 2006). More generally, our results are difficult to reconcile with the claim from *one-process theories* of reasoning that deductive and inductive reasoning rely on the same resources (Oaksford and Chater, 2001; Evans et al., 2015). Rather, in line with *two-process theories*, these results suggest that inductive and deductive reasoning rely on qualitatively different neural mechanisms, at least in children. These results are broadly consistent with an increasing number of behavioral and imaging studies that have also observed dissociations between deductive and inductive reasoning in adults and children.

First, a growing number of behavioral studies suggest that deductive and inductive reasoning involve different mechanisms (Verschueren et al., 2005; Markovits and Handley, 2005; Markovits and Thompson, 2008; Markovits et al., 2013; Markovits et al., 2015; Rips, 2001; Heit and Rotello, 2010; Rotello and Heit, 2009). For instance, Markovits and Thompson (2008) presented children from 6 to 9 with a task similar to the task used here (with MP and AC arguments that children had to evaluate under either "validity" or "likelihood" instructions). Although the authors found an increase in the rate of rejection of the AC form with age, no age-related change in performance was observed under "likelihood" instruction. Thus, there was a developmental dissociation between the patterns of performance associated with evaluating the logical validity and likelihood of a conclusion. Studies in adults have also pointed to differences between these types of evaluation. For example, again using a task similar to that used in the present study, Markovits and Handley (2005) found different patterns of behavioral performance across different forms of conditional inferences, a result in line with that of Rips (2001). Finally, performance in deductive and inductive tasks is affected by different factors. For instance, whereas deductive judgments are more influenced by logical validity, inductive judgments are more influenced by argument length and premise-conclusion similarity (Heit and Rotello, 2010; Rotello and Heit, 2009). Interestingly, the weight of these factors depends upon task-related parameters. That is, increasing analytic processing by reducing fluency (using a font that is difficult to read) makes induction judgments more sensitive to logical validity, whereas increasing heuristic processing by speeding up judgments makes deductive judgments more sensitive to premise-conclusion similarity (Heit and Rotello, 2010; Rotello and Heit, 2009). This has led Heit and Rotello (2010) to propose that "induction and deduction judgments both tap into underlying heuristic and analytic processes but in different proportions" (p. 806). That is, whereas inductive reasoning might more heavily rely on heuristic processing, deductive reasoning might more heavily rely on analytic processing. It is possible that the different brain systems identified here (for judging the logical validity or the likelihood of a conclusion) reflect this different mixture of heuristic and analytic processing. Future studies in which analytic and heuristic processing are manipulated during the task (e.g., by reducing fluency or speeding up the task) are needed to shed light on this issue.

Second, although the few neuroimaging studies that have investigated the neural substrates of deductive and inductive reasoning have all been performed on adult participants and have generated inconsistent results, we note that they all point to some neural differences between both forms of reasoning (Goel et al., 1997; Osherson et al., 1998; Parsons and Osherson, 2001; Goel and Dolan, 2004). For example, Parsons and Osherson (2001) found little overlap between deductive and inductive reasoning, the two forms of reasoning mostly activating different hemispheres. In another study, Goel and Dolan (2004) found a neural dissociation between deductive and inductive reasoning in the prefrontal cortex. It is difficult to compare these results to ours because (1) the materials differed from those used in the present study (and this has been found to significantly affect the location of neural activity associated with deductive reasoning; Prado et al., 2011); (2) previous studies have not compared brain activity between participants who exhibit deductive behavior versus those who do not; and (3) only adults were investigated in those studies. However, the conclusions of these studies converge with ours in showing that there is some form of neural dissociation between deductive and inductive reasoning.

4.3. Limitations

It is important to acknowledge here several potential limitations of the present study. First, as we have shown in a previous meta-analysis (Prado et al., 2011), the brain mechanisms involved in deductive reasoning depend upon the type of argument involved. It is also important to point out that, given the focus of the present study on the emergence of deductive reasoning, the dissociation between validity and likelihood trials is observed with children participants. Therefore, future studies need to determine whether the results obtained here with conditional arguments may apply to (1) other types of arguments (e.g., categorical syllogisms) and (2) adult participants.

Second, our task (adapted from Markovits and Handley, 2005; Markovits and Thompson, 2008), involves a binary choice in *Validity* trials and a choice between 5 response options on a scale in *Likelihood* trial. Therefore, it might be argued that the neural differences observed between *Validity* and *Likelihood* trials may be due to differences in the procedures used for collecting a response rather than to differences in underlying computational processes. We think that this is unlikely because our baseline consisted in problems of the MP form that were presented with the exact same procedure as that employed for the AC form (notably in terms of number of response options). In other words, subtracting activity associated with the MP form from activity associated with the AC form in both *Validity* and *Likelihood* trials should isolate reasoning-related activity from activity associated with response selection. This notably makes it possible to compare activity related to the AC form (compared to the MP form) in *Validity* trials to activity related to the AC form (compared to the MP form) in *Likelihood* trials without that comparison being confounded by a difference in number of response options. Nonetheless, we acknowledge that this cognitive subtraction logic has its limitations (Friston et al., 1996) and future studies may investigate differences in the neural bases of deductive and inductive reasoning while equating the number of response options between conditions.

Third, an unexpected result is that we failed to detect any activity differences between participants who rejected the AC form and those who accepted that form when rules had a weak association between cause and consequence in *Validity* trials. That is, most of the results described above are limited to rules with a strong association between cause and consequence. It is always difficult to discuss a null result. One potential explanation is that counterexamples come to mind more easily in weakly than strongly associated rules (Quinn and Markovits, 1998; Grosset et al., 2005). For instance, consider how the counterexample "a baby who cries might also be hungry" may be relatively easily retrieved when faced with the argument "If a baby is too cold then he will start crying; The baby starts crying; Is the baby too cold?". Rejecting AC forms with weakly associated rules might thus require less working-memory capacity than rejecting the AC forms with strongly associated rules. It is possible that our experiment might have lacked power to detect the more subtle inter-individual differences in working-memory processing that may underlie weakly associated rules. Future studies might also explore this possibility.[1]

Fourth, although our results clearly point to dissociated neural mechanisms for likelihood and validity judgments, we did find that accepting the AC form in *Validity* trials was correlated with evaluating that form as likely in *Likelihood* trials. How can one reconcile these two findings? It is important to note that behavior often reflects the combined effects of multiple processing stages (Sternberg, 2001). As such, this behavioral correlation may not necessarily indicate that validity and likelihood judgments share the same core mechanisms. For instance, the correlation might be driven by inter-individual differences in the ability to inhibit an overall tendency to accept conclusions (and rate them as likely) in children. In fact, we also found that accepting the MP form in *Validity* trials was correlated with evaluating that form as likely in

[1] However, note that, even in rules with a weak association, we identified a region of the left Precuneus in which activity increased as likelihood ratings decreased in Likelihood trials, whereas no variation of activity was observed with acceptance rates in Validity trials. Thus, in weakly as in strongly associated rules, the brain mechanisms that are involved in evaluating the likelihood of a conclusion might not necessarily contribute to validity judgments.

Likelihood trials. Therefore, this correlation is not specific to the AC form and might be mediated by domain-general factors (e.g., differences in executive functioning or response bias) rather than by inter-individual differences in core reasoning mechanisms.

4.4. Conclusion

In sum, we investigated here the brain regions underlying the emergence of deductive reasoning with conditional rules in children and tested whether these brain mechanisms overlap with those involved in evaluating the likelihood of a conclusion. We found that children who made accurate judgments of logical validity (as compared to those who did not) exhibited enhanced activity in left and medial frontal regions. In contrast, differences in likelihood ratings between children were related to differences of activity in right frontal and bilateral parietal regions. Thus, there was a neural dissociation between the brain regions involved in judgments of likelihood and logical validity. These results appear in line with *two-process theories* of reasoning, which assume that deductive and inductive reasoning rely on distinct cognitive resources (Markovits and Barrouillet, 2001; Verschueren et al., 2005; Heit and Rotello, 2010; Rotello and Heit, 2009).

Conflicts of interest

The authors declare no competing financial interests.

Acknowledgments

This research was supported by grants from the European Union (Marie Curie Career Integration Grant n° PCIG12-GA-2012-333602) and the Agence Nationale de la Recherche (ANR-14-CE30-0002-01) to J.P. We thank the Hospices Civils de Lyon for sponsoring the research, as well as Romain Mathieu, Auriane Couderc and the MRI technicians' team at the Lyon Neurological Hospital for their assistance in collecting the fMRI data. Finally, we are grateful to Pr. Christian Scheiber for his help with the pre-MRI medical exams.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2017.09.029.

References

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26 (3), 839–851.
Barrouillet, P., Lecas, J.F., 1999. Mental models in conditional reasoning and working memory. Think. Reason. 5 (4), 289–302.
Barrouillet, P., Markovits, H., Quinn, S., 2002. Developmental and content effects in reasoning with causal conditionals. J. Exp. Child. Psychol. 81 (3), 235–248. https://doi.org/10.1006/jecp.2001.2652.
Braine, M.D.S., O'Brien, D.P. (Eds.), 1998. Mental Logic. Erlbaum, Mahwah, NJ.
Burgund, E.D., Kang, H.C., Kelly, J.E., Buckner, R.L., Snyder, A.Z., Petersen, S.E., Schlaggar, B.L., 2002. The feasibility of a common stereotactic space for children and adults in fMRI studies of development. Neuroimage 17 (1), 184–200.
Byrnes, James P., Overton, Willis F., 1986. Reasoning about certainty and uncertainty in concrete, causal and propositional contexts. Dev. Psychol. 22 (6), 793–799.
Cognet, G., 2006. Nouvelle Echelle Métrique de l'Intelligence. Editions du Centre de Psychologie Appliqué, Paris.
Cummins, Denise D., Lubbart, Todd, Alksnis, Olaf, Rist, Robert, 1991. Conditional reasoning and causation. Mem. Cognit. 19 (3), 274–282.
De Neys, W., Everaerts, D., 2008. Developmental trends in everyday conditional reasoning: the retrieval and inhibition interplay. J. Exp. Child. Psychol. 100 (4), 252–263. https://doi.org/10.1016/j.jecp.2008.03.003.
Evans, J.S.B.T., Over, D., 2013. Reasoning to and from belief: deduction and induction are still distinct. Think. Reason. 19 (3), 267–283. https://doi.org/10.1080/13546783.201.
Evans, J.S.B., Stanovich, K.E., 2013. Dual-process theories of higher cognition: advancing the debate. Perspect. Psychol. Sci. 8 (3), 223–241.
Evans, J.S.B.T., Thompson, V.A., Over, D., 2015. Uncertain deduction and conditional reasoning. Front. Psychol. 6 (398) https://doi.org/10.3389/fpsyg.2015.00398.
Fangmeier, T., Knauff, M., Ruff, C., Sloutsky, V., 2006. fMRI evidence for a three-stage model of deductive reasoning. J. Cogn. Neurosci. 18 (3), 320–334.
Friston, K.J., Price, C.J., Fletcher, P., Moore, C., Frackowiak, R.S.J., Dolan, R.J., 1996. The trouble with cognitive subtraction. Neuroimage 4 (2), 97–104.
Goel, V., Buchel, C., Frith, C., Dolan, R.J., 2000. Dissociation of mechanisms underlying syllogistic reasoning. Neuroimage 12, 504–514. https://doi.org/10.1006/nimg.2000.0636.
Goel, V., Dolan, R.J., 2004. Differential involvement of left prefrontal cortex in inductive and deductive reasoning. Cognition 93, B109–B121. https://doi.org/10.1016/j.cognition.2004.03.001.
Goel, V., Gold, B., Kapur, S., Houle, S., 1997. The seats of reason? An imaging study of deductive and inductive reasoning. NeuroReport 8, 1305–1310.
Grosset, N., Barrouillet, P., Markovits, H., 2005. Chronometric evidence for memory retrieval in causal conditional reasoning: the case of the association strength effect. Mem. Cogn. 33 (4), 734–741.
Hawkins, J., Pea, R.D., Glick, J., Scribner, S., 1984. "Merds that laugh Don't like mushrooms": evidence for deductive reasoning by preschoolers. Dev. Psychol. 20 (4), 584–594.
Heit, E., 2014. Brain imaging, forward inference, and theories of reasoning. Front. Hum. Neurosci. 8.
Heit, E., Rotello, C.M., 2010. Relations between inductive reasoning and deductive reasoning. J. Exp. Psychol. Learn. Mem. Cogn. 36 (3), 805.
Janveau-Brennan, J., Markovits, H., 1999. The development of reasoning with causal conditional. Dev. Psychol. 35 (4), 904–911.
Johnson-Laird, P.N., 1994. Mental models and probabilistic thinking. Cognition 50 (1), 189–209.
Kang, H.C., Burgund, E.D., Lugar, H.M., Petersen, S.E., Schlaggar, B.L., 2003. Comparison of functional activation foci in children and adults using a common stereotactic space. Neuroimage 19 (1), 16–28.
Klingberg, T., 2006. Development of a superior frontal-intraparietal network for visuo-spatial working memory. Neuropsychologia 44 (11), 2171–2177. https://doi.org/10.1016/j.neuropsychologia.2005.11.019.
Liu, W.Y., Song, N., 2003. Fuzzy functional dependencies and bayesian networks. J. Comput. Sci. Technol. 18 (55), 56–66.
Markovits, H., Barrouillet, P., 2001. The development of conditional reasoning: a mental model account. Dev. Rev. 22 (1), 5–36.
Markovits, H., Brisson, J., De Chantal, P.L., 2015. Deductive updating is not Bayesian. J. Exp. Psychol. Learn. Mem. Cognit. 41 (4), 949–956. https://doi.org/10.1037/xlm0000092.
Markovits, H., Brisson, J., & De Chantal, P.L. (2015). Deductive Updating Is Not Bayesian. J. Exp. Psychol. Learn. Mem. Cognit., 41(4), 949–956. doi: 10.1037/xlm0000092
Markovits, H., Brunet, M.L., Thompson, V.A., Brisson, J., 2013. Direct evidence for a dual process model of deductive inference. J. Exp. Psychol. Learn. Mem. Cogn. 39 (4), 1213–1222.
Markovits, H., Doyon, C., 2004. Information processing and reasoning with premises that are empirically false: interference, working memory, and processing speed. Mem. Cogn. 32 (4), 592–601.
Markovits, H., Handley, S., 2005. Is inferential reasoning just probabilistic reasoning in disguise? Mem. Cogn. 33 (7), 1315–1323.
Markovits, H., Thompson, V.A., 2008. Different developmental patterns of simple deductive and probabilistic inferential reasoning. Mem. Cogn. 36 (6), 1066–1078. https://doi.org/10.3758/MC.36.6.1066.
Markovits, H., Vachon, R., 1990. Conditional reasoning, representation, and level of abstraction. Dev. Psychol. 26 (6), 942–951.
Markovits, H., Venet, M., Janveau-Brennan, J., Malfait, N., Pion, N., Vadeboncoeur, I., 1996. Reasoning in young children: fantaisy and information retrieval. Child. Dev. 67, 2857–2872.
Mathieu, R., Booth, J.R., Prado, J., 2015. Distributed neural representation of logical arguments in school-age children. Hum. Brain Mapp. 36 (3), 996–1009.
Mazaika, P.K., Hoeft, F., Glover, G.H., Reiss, A.L., 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.
Monti, M.M., Osherson, D.N., Martinez, M.J., Parsons, L.M., 2007. Functional neuroanatomy of deductive inference: a language-independent distributed network. Neuroimage 37 (3), 1005–1016.
Monti, M., Parsons, L., Osherson, D., 2009. The boundaries of language and thought in deductive inference. PNAS 106 (30), 12554. https://doi.org/10.1073/pnas.0902422106.
Morsanyi, K., Devine, A., Nobes, A., Szücs, D., 2013. The link between logic, mathematics and imagination: evidence from children with developmental dyscalculia and mathematically gifted children. Dev. Sci. 16 (4), 542–553. https://doi.org/10.1111/desc.12048.
Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B., 2005. Valid conjunction inference with the minimum statistic. Neuroimage 25 (3), 653–660.
Oaksford, M., Chater, N., 2001. The probabilistic approach to human reasoning. Trends Cogn. Sci. 5 (8), 349–357.
Oaksford, Mike, Chater, Nick, Larkin, Joanne, 2000. Probabilities and polarity biases in conditional inference. J. Exp. Psychol. Learn. Mem. Cogn. 26 (4), 883–899.
Oberauer, K., 2006. Reasoning with conditionals: a test of formal models of four theories. Cogn. Psychol. 53 (3), 238–283. https://doi.org/10.1016/j.cogpsych.2006.04.001.
Osherson, D., Perani, D., Cappa, S., Schnur, T., Grassi, F., Fazio, F., 1998. Distinct brain loci in deductive versus probabilistic reasoning. Neuropsychologia 36 (4), 369–376.
Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E., 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum. Brain Mapp. 25 (1), 46–59. https://doi.org/10.1002/hbm.20131.
Parsons, L., Osherson, D., 2001. New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. Cereb. Cortex 11, 954–965.

Prado, J., Chadha, A., Booth, J.R., 2011. The brain network for deductive reasoning: a quantitative meta-analysis of 28 neuroimaging studies. J. Cogn. Neurosci. 23 (11), 3483–3497.
Prado, J., Mutreja, R., Booth, J.R., 2013. Fractionating the neural substrates of transitive reasoning: task-dependent contributions of spatial and verbal representations. Cereb. Cortex 23 (3), 499–507. https://doi.org/10.1093/cercor/bhr389.
Quinn, S., Markovits, H., 1998. Conditional reasoning, causality, and the structure of semantic memory: strength of association as a predictive factor for content effects. Cognition 68 (3), B93–B101. https://doi.org/10.1016/s0010-0277(98)00053-5.
Reverberi, C., Bonatti, L., Frackowiak, R.S., Paulesu, E., Cherubini, P., Macaluso, E., 2012. Large scale brain activations predict reasoning profile. Neuroimage 59, 1752–1764. https://doi.org/10.1016/j.neuroimage.2011.08.027.
Reverberi, C., Cherubini, P., Frackowiak, R.S., 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., Cherubini, P., Rapisarda, A., Rigamonti, E., Caltagirone, C., Frackowiak, R.S., Paulesu, E., 2007. Neural basis of generation of conclusions in elementary deduction. Neuroimage 28, 752–762. https://doi.org/10.1016/j.neuroimage.2007.07.060.
Reverberi, C., Shallice, T., D'Agostini, S., Skrap, M., Bonatti, L., 2009. Cortical bases of elementary deductive reasoning: inference, memory, and metadeduction. Neuropsychologia 47 (1107), 1116. https://doi.org/10.1016/j.neuropsychologia.2009.01.004.
Ridderinkhof, K. Richard, Ullsperger, Markus, Crone, Eveline A., Nieuwenhuis, Sander, 2004. The role of the medial frontal cortex in cognitive control. Science 306 (5695), 443–447.
Rips, L.J., 2001. Two kinds of reasoning. Psychol. Sci. 12 (2), 129–134.
Rotello, C.M., Heit, E., 2009. Modeling the effects of argument length and validity on inductive and deductive reasoning. J. Exp. Psychol. Learn. Mem. Cogn. 35 (5), 1317.
Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A.R., Schulz, J.B., Eickhoff, S.B., 2012. Modelling neural correlates of working memory: a coordinate-based meta-analysis. Neuroimage 60 (1), 830–846. https://doi.org/10.1016/j.neuroimage.2011.11.050.
Sternberg, S., 2001. Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. Acta Psychol. (Amst) 106, 147–246.S.
Verschueren, N., Schaeken, W., D'Ydewalle, G., 2005. A dual-process specification of causal conditional reasoning. Think. Reason. 11 (3), 239–278. https://doi.org/10.1080/13546780442000178.
Wilke, M., Schmithorst, V.J., Holland, S.K., 2002. Assessment of spatial normalization of whole-brain magnetic resonance images in children. Hum. Brain Mapp. 17 (1), 48–60.