

# What MEG can Reveal about Inference Making: the Case of if...then Sentences

Mathilde Bonnefond,<sup>1,2\*</sup> Ira Noveck,<sup>1</sup> Sylvain Baillet,<sup>3,4</sup> Anne Cheylus,<sup>1</sup>  
Claude Delpuech,<sup>5</sup> Olivier Bertrand,<sup>5</sup> Pierre Fourneret,<sup>1</sup>  
and Jean-Baptiste Van der Henst<sup>1</sup>

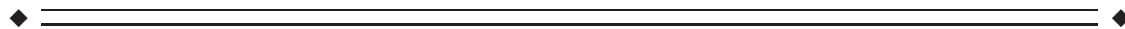
<sup>1</sup>Laboratoire sur le Langage, le Cerveau et la Cognition (L2C2), CNRS-University of Lyon 1,  
67 Boulevard Pinel, 69675 Bron, France

<sup>2</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen,  
6526 EN Nijmegen, Netherlands

<sup>3</sup>University of Paris UPMC, CNRS, Hôpital de La Salpêtrière, 75651 Paris

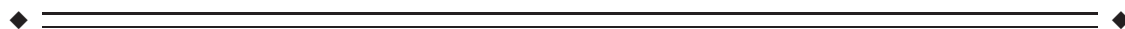
<sup>4</sup>Department of Neurology, Medical College of Wisconsin, Milwaukee, Wisconsin

<sup>5</sup>Centre de Recherche en Neurosciences de Lyon; INSERM U1028- CNRS UMR5292



**Abstract:** Characterizing the neural substrate of reasoning has been investigated with regularity over the last 10 years or so while relying on measures that come primarily from positron emission tomography and functional magnetic resonance imaging. To some extent, these techniques—as well as those from electroencephalography—have shown that time course is equally worthwhile for revealing the way reasoning processes work in the brain. In this work, we employ magnetoencephalography while investigating Modus Ponens (If P then Q; P//Therefore, Q) in order to simultaneously derive time course and the source of this fundamental logical inference. The present results show that conditional reasoning involves several successive cognitive processes, each of which engages a distinct cerebral network over the course of inference making, and as soon as a conditional sentence is processed. *Hum Brain Mapp* 34:684–697, 2013. © 2012 Wiley Periodicals, Inc.

**Key words:** conditional reasoning; MEG; dynamics



## INTRODUCTION

What cognitive mechanisms underlie our ability to perform deductive inferences? The emergence of positron emission tomography and functional magnetic resonance imaging (fMRI) has provided researchers with a way to

address this question by identifying the brain areas involved in various tasks of logical inference making. These include investigations of traditional Aristotelian syllogistic inferences (All As are Bs, All Bs are Cs, therefore All As are Cs), transitive inferences (A is taller than B; B is taller than; C; therefore A is taller than C) and propositional inferences (e.g., If P then Q; P; Therefore Q) [for a review, see Goel, 2007].

These experiments have allowed researchers to better characterize reasoning processes and have shown that inference making, far from being a unitary process based on a single cortical network, actually engages different brain regions depending on the type of reasoning task or argument. For instance, while some evidence indicates that propositional reasoning activates left frontal, and occasionally parietal, regions [Reverberi et al., 2007; Noveck et al.,

\*Correspondence to: Mathilde Bonnefond, PhD, Donders Institute for Brain, Cognition And Behaviour, Radboud University, Nijmegen, 6526 EN Nijmegen, Netherlands.  
E-mail: m.bonnefond@fcdonders.ru.nl

Received for publication 7 January 2011; Revised 29 June 2011; Accepted 18 August 2011

DOI: 10.1002/hbm.21465

Published online 16 January 2012 in Wiley Online Library (wileyonlinelibrary.com).

2004; Reveberri et al., 2010; Prado et al., 2010; see also Prado et al., in press, for a meta-analysis], it appears that transitive reasoning tasks primarily activate bilateral parietal and frontal areas, which are usually engaged by visuospatial tasks [Goel and Dolan, 2001; Goel et al., 2004; Knauff et al., 2003]. The findings across tasks also indicate that the activation of a reasoning network varies to some extent with the inference form [Prado et al., 2010; Prado et al. in press; though see Monti et al., 2007; Monti et al., 2009 which will be discussed later]. Taken together, these studies provide a major contribution to the understanding of the multiple mechanisms involved in reasoning (see also Houdé, 2007 for a broader view on reasoning).

Moreover, neuroimaging researchers of reasoning, while using fMRI, have strived to focus on the neural activity linked to various stages of reasoning [Fangmeier et al., 2006] or to the very moment that a task allows for premise-integration [Reverberri et al., 2007]. Fangmeier et al. [2006], who investigated transitive reasoning, showed that the processing of premises involves the occipito-temporal cortex, while the integration of the premises involves the anterior prefrontal cortex; this is followed by a validation-of-the-conclusion phase, which recruits the prefrontal cortex and the posterior parietal cortex (see also Fangmeier et al., 2009 for an account of this model for acoustic reasoning). Reverberri et al. [2007], who investigated propositional logical inferences (linked to conditionals and disjunctions), focused their work on the very moment that participants were in a position to carry out a logical inference. More specifically, Reverberri et al. [2007] showed that a left parietal and frontal network is implicated as soon as two integrable premises are available to a reasoner. For example, in (1) below, participants were presented three premises beginning with a categorical one and were able to generate an inference with the arrival of the third premise only (before being presented with a multiple choice of potential conclusions):

1. There is a square

If there is a hexagon then there is a triangle.  
If there is a square then there is a rhombus.

Since an inference can be produced as soon as two compatible premises are available [see Lea, 1995; Reverberri et al., 2009], this approach represented an advance over prior studies which had captured processing only at the conclusion level [e.g., see Noveck et al., 2004].

These and other more recent studies [see also Rodriguez-Moreno and Hirsch, 2009] demonstrate how the neuroimaging-of-reasoning literature has remained innovative and has, in fact, inspired the present work. Nevertheless, while these recent fMRI studies may cleverly isolate premise integration, they are not in the position to go much further, i.e., to finely distinguish among the potential cascade of events that a logical inference quickly unleashes. This is due, at least partly, to the fact that the latencies involved in

logical inference-making are on the order of several 100 ms while brain images are acquired with temporal resolution on the order of 2 to 3 seconds. This constraint is what has led some researchers to investigate neural correlates of inference making from a temporal perspective.

Bonnefond and Van der Henst [2009] used electrophysiological techniques in order to investigate brain processes that are instantaneously involved in the processing of conditional inference. Their findings showed that in the context of an If P then Q conditional, the minor premise P prompts a P3b component. In contrast, when the minor premise is R (mismatching premise), it results in a N2 component. Given that the P3b is typically associated with target detection (see Picton, 1992 for a review) and that the N2 is a signature of mismatch (see Folstein and Van Petten, 2008 for a review), the authors argued that their results indicate that the conditional premise If P then Q raises expectations regarding upcoming information. If the expectation is met (If P then Q; P), the participant has the means to produce an inference. They also found that inference generation is associated with a positive slow wave, a component traditionally linked to the immediate cognitive processes following the detection of a stimulus [Garcia-Larrea and Cezanne-Bert, 1998]. Moreover, the positive slow wave is followed by a contingent negative variation, which is known to be associated with expectations about an upcoming stimulus [Rohrbaugh et al., 1997]. The authors argued that the contingent negative variation is associated with the anticipation of the conclusion to be presented. Overall, the findings from Bonnefond and Van der Henst [2009] indicate that there is much relevant neural activity to be captured over the entire course of a conditional argument (from the moment a participant is in the position to make an inference) and that such a thorough approach can shed additional light on inference-making.

In this article, we take this sort of approach one step further. While continuing to focus on conditional arguments (and in particular Modus Ponens), which have arguably become the canonical object of study in the reasoning literature, we aim to advance our knowledge concerning the underlying cerebral processes involved in logical inference-making by using imaging techniques from magnetoencephalography (MEG). This technique, as is well known, has a temporal resolution that is extremely precise (on the order of milliseconds) while still allowing for source imaging.

In this experiment, we present stimuli that resonate with distinct sensory (i.e., visual and auditory) modalities in order to optimally exploit MEG technology. More precisely, the conditional reasoning paradigm is designed with (1) a conditional statement (i.e., a major premise, such as If there is a square then there is a low sound); (2) an actual geometrical shape (i.e., a visual stimulus), which serves as a minor premise that matches or mismatches the antecedent of the conditional statement (e.g., a square or perhaps a nonsquare—a circle) and, finally; (3) a sound (i.e., an auditory stimulus) which serves as a

conclusion that either matches or mismatches the expected outcome. The example in (2) shows a conditional argument with a matching minor premise that allows for a *Modus Ponens* inference to go through:

2. Major premise: If there is a square then there is a low sound

Minor premise: 

Conclusion:  (low sound)

Our ultimate goal is to distinguish preliminary steps of inference-making from later steps. That is, while visual stimuli ought to recruit visual areas early (e.g., with the detection of the visual shape corresponding to the minor premise), auditory stimuli ought to recruit auditory areas relatively late (e.g., when the anticipated sound is inferred from the combination of the two premises; see the predictions later). Had the task involved only visual stimuli, it would have been harder to discriminate the beginning and the end of inference since in both cases visual areas would have been recruited. Therefore, such a procedure ought to provoke activity in specific areas of the brain and at distinct points in time; the upshot is that this investigation increases the degree of confidence one can have with respect to the link between cognitive processes and their neural activation [see Poldrack 2006], since it offers two complementary dimensions to capture these processes, namely time and (brain) space. This approach also allows us to bring into play (and to potentially capture on-line) a variety of supporting mechanisms that are presumed to be critical to reasoning.

Equipped with this experimental set-up, we posit the following four predictions. First, if the conditional premise yields specific expectations regarding upcoming information as indicated by Bonnefond and Van der Henst [2009], then a minor premise that matches the antecedent of the conditional should activate the network associated with the magnetic counterpart of the P3b component [in particular the parietal and posterior cingulate cortices, see Linden, 2005]; further, these activations should be observed in the wake of geometrical shape identification [i.e., in the visual ventral stream]. Second, a minor premise that does not match the antecedent of the conditional premise should elicit activity in the areas associated with the magnetic counterpart of the N2 component, i.e. the anterior cingulate cortex (ACC) [Kerns et al., 2004]. Third, assuming that inference-making can be captured at the very moment of premise integration, as posited by Lea [1995] and Reverberi et al. [2007], activity should be unique to the condition where the minor premise matches the antecedent of the conditional statement; furthermore, this should be detectable after the visual shape has been recognized as a match for the antecedent of the major premise (see below for detailed predictions about this cognitive step). Finally, following the (presumably) successful inte-

gration of the two premises, one should come upon evidence of activity in the auditory cortex even before the auditory conclusion is presented. Any sustained activity in this area is arguably linked to auditory working memory [see Pasternak and Greenlee, 2005].

Overall, by confidently isolating the activity linked to each step in a conditional reasoning argument, one can ask a more general question. What is the network that generally supports the inferential processes in conditional reasoning? There are currently two positions on that. One comes from the literature presented earlier, which reports that a left frontal area—that overlaps with language processing—is generally responsible for propositional reasoning [Noveck et al., 2004; Reverberi et al., 2007, 2010; Prado et al., 2010; Prado et al., in press.]. The other comes from Monti et al., [2007, 2009], who claim that the left rostrolateral prefrontal cortex (RLPFC) and the medial superior frontal gyrus (mSFG) are “core” regions of deduction because these areas become active uniquely when participants are required to carry out logical transformations (on complexly worded conditional statements). Their studies aim to show that deductive reasoning is independent of language. The overall objectives of the study remain twofold: determining the sequence of cognitive processes involved in this conditional reasoning task and localizing the regions associated with the integration of the premises.

## METHOD

### Participants

Sixteen healthy right-handed volunteers (5 males and 11 females, aged 20-30 years, mean: 23 years) with no history of neurological or psychiatric disorders participated in the study. All participants gave written informed consent and received compensation for their participation. Procedures were approved by the local ethics committee (CPP of Lyon, France). Four participants were excluded from the analyses due to excessive eye or head movements.

### Stimulus material and Paradigm design

Participants were seated upright in a magnetically shielded room. They were instructed to sit still and look at a screen placed about 80 cm in front of them. The visual stimuli consisted of 12 geometrical forms (arrow, asterisk, circle, cross, diamond, ellipse, moon, parallelogram, rectangle, square, star, triangle). The auditory stimuli consisted of three different tone sounds: low (300 Hz), medium (1,000 Hz), and high (1,300 Hz) pitch tones presented binaurally via air-conducting tubes with foam ear-tips. The experimental task was preceded by a training session that consisted in familiarizing the participants with these three different tone sounds. The loudness of the sounds was 45 db above the detection threshold of each

participant. Total number of trials was 475 (i.e. 460 + 15 fillers).

The paradigm consisted of two test conditions and one control condition. The two test conditions explicitly involved a conditional (if-then) premise such as “if there is a square then there is a low sound” followed by a premise which matched or not the antecedent of the conditional premise (Matching and Mismatching conditions, respectively) while the control condition simply said “There will be a shape and a sound.” It is important to note that the three conditions were not distributed equally. A behavioral pilot study lead us to discover that the task was perceived as being easier when containing more trials with premises and conclusions that respectively matched the antecedent and the consequent of the conditional statement. We presented 5 trials for each conditional sentence (see experimental procedures). In particular, we observed that when a matching trial arose after several mismatching trials (two or more) the rate of correct answers decreased (83%, vs. 98% when the matching trial arises first or after a single mismatching trial). Moreover, when asked to indicate what aspects of the task contributed to difficulty, participants reported that there were too many mismatching items. It might be the case that mismatching premises hamper the retention of the conditional in working memory as they introduced figures that are unrelated to it. Participants also expressed the feeling that mismatching items were more numerous than matching items. We thus presented more trials that contained matching, rather than mismatching, premises and conclusions (61% of the premises in the test conditions were matching premises; 61% of the conclusion following the matching premise were matching conclusions).<sup>1</sup>

In the matching condition (200 trials), the minor premise (i.e. a visual shape) matched the antecedent of the major premise and could be integrated with the major premise in order to infer which sound could be expected in the conclusion. In the mismatching condition (130 trials), the visual shape did not match the antecedent of the major premise. Thus, the expectations raised by the major premise were not satisfied and participants could not make a determinate inference regarding the consequent of the conditional statement. They were instructed to respond “indeterminate” when the sound appeared. The two conditions are represented in Figure 1.

As indicated above, the control condition (130 trials) presented participants with the same kind of visual and auditory stimuli as in the experimental conditions but they were not required to combine two premises in order to

perform an inference. In lieu of a conditional statement, they were shown the sentence: “There will be a shape and then a sound.” The 130 trials did involve a shape/sound pair and 15 filler trials were introduced with no visual stimuli. The experimental design is illustrated in Figure 1.

### Experimental Procedures

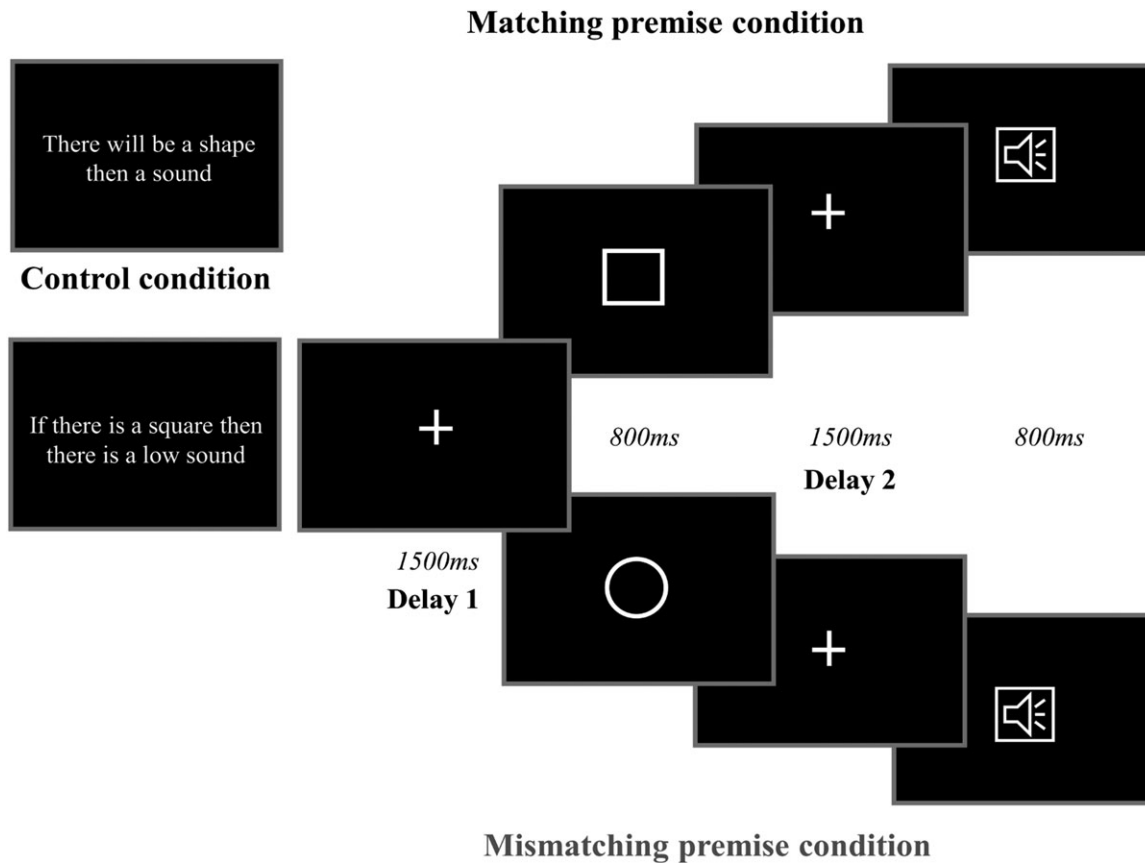
Stimuli were presented with Presentation 10.2 software (Neurobehavioral Systems, <http://www.neurobs.com/>). In the experimental conditions (i.e., the matching and mismatching conditions), a conditional premise was presented initially before being followed by five shape/sound pairs (i.e., If there is a square then there is a low sound was used once before five different shape/sound pairs were presented, representing at least one of each of the experimental conditions). Hence participants had to remember a single conditional premise throughout five trials. The type of premises and conclusions administered for the same P1 were randomized across trials but there was always at least one mismatching premise and one matching premise (the latter being followed either by a matching or a mismatching conclusion). Such a procedure allowed us to reduce the duration of a rather long experimental session. In the control condition, a neutral sentence was initially presented (i.e., there will be a shape and a sound) and was also followed by five shape/sound pairs.

Once participants had read a sentence, they had to press a key with their left hand to see the cross of the first delay (1,500 ms) followed by the visual shape (i.e., the minor premise). The shape was presented for 800 ms and after a 1,500 ms delay (delay 2), an auditory stimulus (i.e. the conclusion) was presented for 800 ms. Between the five shape/sound pairs, a red cross appeared and participants had to press a button to see the next shape/sound pair associated with the sentence he read the first time. In the experimental conditions, participants were instructed to decide whether the conclusion was valid, invalid, or indeterminate (i.e., they could not decide whether the premise is valid or inconsistent) by pressing a response key with their right hand. In the control condition, participants were instructed to respond only when a visual shape appeared before a sound. Some filler trials were introduced with no visual stimuli (i.e., the fixation cross remained of the screen until the sound was heard) and participants were asked not to respond to these fillers. The 460 trials involving 66 conditional statements (each of them followed by five shape/sound pairs) and 26 neutral statements were presented in four sessions of around 15 min each.

### MEG Recording

MEG was recorded using a whole-head system (CTF Inc., Vancouver, Canada) comprising of 275 first-order magnetic gradiometers. The signals were sampled at 600 Hz with an anti-aliasing filter at 150 Hz. Recording epochs

<sup>1</sup>In fact, despite our best efforts, debriefing questionnaires would later reveal that participants felt that they dealt with more mismatching premises/conclusions than matching premises/conclusions. This may result from the fact that the proportion of pure matching arguments (i.e. those which include a matching premise and a matching conclusion) is less than 50% and that participants were particularly attentive to the pure matching vs. other arguments contrast.



**Figure 1.**  
Experimental procedure. Timing of a sample trial. 127 × 97 mm (300 × 300 DPI).

lasted from 100 ms before trial onset to 6,000 ms after trial onset. The participants' head position was determined with head coils fixated at the nasion and the preauricular points at the beginning and end of each recording to ensure that head movements did not exceed 0.5 cm. Bipolar Ag-AgCl electrodes (band pass = 0.05-150 Hz) were used to record the electro-oculogram monitoring both horizontal and vertical eye movements and the electrocardiogram detecting the subjects' cardiac activity. These recordings were acquired in continuous mode digitized at 600 Hz, and stored for offline analysis.

### MEG Scalp Data Analysis

#### Data preprocessing

Trials were rejected from any further analysis if (a) the segment of interest contained identified eye-blinks or muscle artefacts, or if; (b) the standard deviation of any MEG recording channel within a sliding 200-ms time window exceeded 1,500 fT. The MEG signals were low-pass filtered off-line below 25 Hz. The trials with incorrect responses or with reaction times of more than three standard deviations

from the mean were also excluded from the analysis. This resulted in 4.5% of trials being removed from the dataset across subjects on average.

#### Event related field analyses

Event related field (ERF) analyses were conducted using ELAN-Pack software developed at INSERM U821 (Lyon, France). Such analyses consisted in averaging, for each block and each condition the MEG segments time locked to the onset of the visual stimulus in each trial over a 2,450 ms period, including a 150 ms prestimulus interval. The baseline correction was calculated for all conditions from the 150 ms prestimulus interval of the control condition (crossbaseline) to have the same baseline for each condition. We choose such a baseline because we presume that the prestimulus window of the control condition would elicit a lower level of expectation than the prestimulus window of the experimental condition. Indeed, the anticipation of a piece of information is known to generate specific waves [see Brunia, 1999]. However, we also analyzed the data with the pre-stimulus window of each condition as a baseline and the findings turned out similar.

To analyze the effects of condition in the MEG recordings, the root mean square (RMS) of the amplitude was calculated on every sensor within seven time windows centered around the seven main ERF components observed, that is 100–140, 140–180, 200–250, 270–320, 320–370, 370–500, and 550–2,100 ms (i.e., sustained activity during the second delay, i.e. the delay between the shape and the sound). The use of time windows, instead of each time point, was done to reduce calculus time. The first time point of each time window was reinitiated as zero and the value of each time point of the window was calculated relative to the value of this first point. This was done to negate effects from the previous time window. The mismatching and the control conditions were contrasted to assess the effect of mismatching between the minor premise and the antecedent of the conditional statement. The matching premise and the control conditions were contrasted to assess the effect of inference for the different time windows. For each participant, the same number of trials in the matching condition and in the control condition was randomly selected.

### MRI data

Individual T1-weighted MRI anatomical volume data was acquired for each subject's head with 1 mm isotropic voxel dimension with a 1.5T MRI scanner (Siemens Sonata Maestro Class) equipped with a standard quadrature head coil. Three markers visible in MRI images were positioned at the same locations as those of the MEG head position indicators. These landmarks were subsequently used for accurate registration between the MEG and the MRI reference frames through rigid-body affine transformation.

### MEG source imaging

MEG source imaging was performed using the BrainStorm software package (<http://neuroimage.usc.edu/brainstorm>). MEG forward modeling was achieved using the overlapping-sphere approach [Huang et al., 1999] based on the individual scalp geometry as identified from the segmentation of head tissues performed using the automatic segmentation pipeline of BrainVisa (<http://brainvisa.info>). Head movements inside the MEG helmet between acquisition runs were compensated by interpolating the MEG recordings onto a common sensor array, which position was the average position of the rigid MEG sensor helmet across the entire set of acquisition runs [Senot et al., 2008]. Source estimation was performed in each subject using a depth-weighted minimum norm imaging model of MEG generators consisting of 10,000 elementary current dipoles distributed over the individual cortical surface [Baillet et al., 2001]. Elementary current dipoles orientations were constrained perpendicularly to the cortical surface. Source amplitudes were estimated for each condition and for each subject on the 2,450 ms time window.

### Cross-Subject Anatomical Coregistration

Individual source maps were geometrically registered to the Colin27 brain template from the Montreal Neurological Institute (MNI). This process was achieved by spatially normalizing the individual cortical surface tessellation from each subject to the Colin27 brain template and by linearly interpolating the individual source maps from the cortical distribution on surface vertices to the image volume.

### Spatial and Temporal Smoothing, Grand Averaging

Data from each subject were spatially smoothed, using a gaussian kernel (FWHM 11.8 mm), and temporally smoothed (moving average temporal window: 25 ms). The individual source amplitude maps were subsequently normalized with respect to baseline and averaged across subjects to yield the group's grand average of MEG source imaging models.

### Statistical Analysis

#### At the channel level

With the RMS data obtained in each time-window, we used non-parametric permutation tests between conditions based on t-statistics for 4,095 permutations. The statistic of global maximum across channels was used to control for the family-wise error rate across the entire set of MEG channels (i.e., this allows to correct the results for multiple comparisons).

#### At the source level

For each time window for which there was a significant difference at the channel level, we used the same statistic procedure to test for amplitude effects at the cortical source level [Pantazi et al., 2005]. Thresholding on the size of the effects was applied: clusters of at least 10 cortical vertices in the distributed sources model were considered. We used the same technique than at the source level to correct the results for multiple comparisons but across the vertices of the brain space instead of the MEG channels. The statistic of global maximum across vertices was used to control for the family-wise error rate across the entire set of vertices. The brain structures found being significantly activated were reported and identified according to a probabilistic atlas of the human brain [Hammers et al., 2003], with coordinates defined in the MNI coordinate system. We did not report Brodman areas associated with the stereotaxic coordinates since the accuracy of the labelling is controversial [Devlin and Poldrack, 2007].

## RESULTS

### Behavioral Results

Arcsine transformations were carried out on the rate of correct answers before analysis [Howell, 1997]. A log

**TABLE I. Brain areas activated in the Matching premise condition**

Anatomical location	No. of vertices in cluster	Vertex-level Z value	MNI coordinates		
			x	y	z
120 ms					
L. lateral occipital	90	12	-13	-105	-3
R. lateral occipital	80	16	15	-104	0
160 ms					
L. posterior temporal lobe	60	15	-30	-62	-19
R. posterior temporal lobe	66	14	41	-63	-20
220 ms					
R. postcentral parietal gyrus	88	12	65	-20	38
R. medial frontal gyrus	34	14	44	7	37
340 ms					
L. parietal	20	8	-38	-75	43
R. posterior temporal/parietal	43	10	58	-60	11

transformation was applied to the reaction time data. The rate of correct answers did not significantly differ across conditions (Matching condition: conclusion true 96% false conclusion 94%; mismatching condition: 100%; control condition 100%). Student's t tests revealed that reaction times did not significantly differ between the Mismatching condition (mean = 326 ms) and the control condition [mean = 333 ms;  $t(11) = 1.02$ ;  $P = ns$ ]. However, response times to the true conclusion of the Matching condition (mean = 421 ms) were faster than the response times to the false conclusion [mean = 527 ms;  $t(11) = 7.5$ ;  $P < 0.001$ ], but both were slower than response times of the control condition [ $t(11) = 7.15$ ;  $P < 0.001$ ].

### ERF Results

#### Common components

In all conditions, we observed four common components: the M100 component, at about 120 ms after the visual stimulus onset (minor premise), the M150 component at about 160 ms, the M200 component at about 225 ms and the M300 component at about 340 ms. The M200 component was slightly more pronounced in the experimental conditions than in the Control condition and the M300 was slightly more pronounced in the Matching condition. However, neither of these differences was significant (see Table I for a report of areas with a Z-score  $>6$  at each latency in the matching condition). As reported in Table I, the results confirm that the ventral stream is activated early during the processing of the visual shape.

In order to isolate the processing demands of the matching and the mismatching conditions respectively, each of these conditions was first compared with the control condi-

tion. Indeed, had we chosen to only compare the two test conditions to each other, it would have been difficult to separate out the effects raised by one condition from those raised by the other. Indeed, because each of the conditions comes with its own processing demands they are not neutral to each other. In particular, while the mismatching condition introduces a clash between the expectations raised by the major premise and the actual minor premise, the matching condition raises concordance between the two. It is thus only in a second stage that we directly compare matching and mismatching conditions. This comparison is performed in order to connect our data with the existing fMRI literature which often uses such a contrast.

#### Mismatching Effect (Figs. 2 and 3): Mismatching Condition vs. Control Condition

About 300 ms after stimulus onset, we observed a larger component (M290) in the Mismatching condition than in the control condition including two clusters, one over right central sensors and a second one over left frontal sensors. Permutation t tests performed on the RMS data obtained for the 250-320 ms time window revealed that the amplitude of this component was significantly larger in the mismatching condition than in the control condition (Fig. 2).

Permutation t test performed at the source level revealed a significant activation of the bilateral ACC (left:  $x = -8$ ;  $y = 35$ ;  $z = 21$ ; right:  $x = 5$ ;  $y = 36$ ;  $z = 21$ ; see Figure 3 and Table II for a complete report of significantly activated areas).

#### Matching effect (Figs. 2 and 4): Matching Condition vs. Control Condition

We observed three larger ERF components in the Matching condition when compared to the Control condition. First there was an M400 component which occurred at about 400 ms after the stimulus onset. This component includes several clusters, one over left temporal sensors (also over left/middle frontal/parietal sensors) and the other one over right temporal/frontal sensors. Second, we observed an M450 component which occurred at about 450 ms over middle frontal sensors. Given that these two components partially share the same sensors we considered them as a complex. Permutation t-tests performed on the RMS data of the 370-500 ms time window revealed significantly larger amplitude of this component in the Matching condition than in the Control condition. Third, we observed slow waves that started at about 550 ms after the stimulus onset and lasted until the end of the second delay (550-2,300 ms time window). They appeared over right and left temporal sensors and over right and left parietal sensors (also over frontal sensors). Permutation t-test showed that these components were significantly larger in the Matching condition than in the control condition, though only over the right hemisphere.

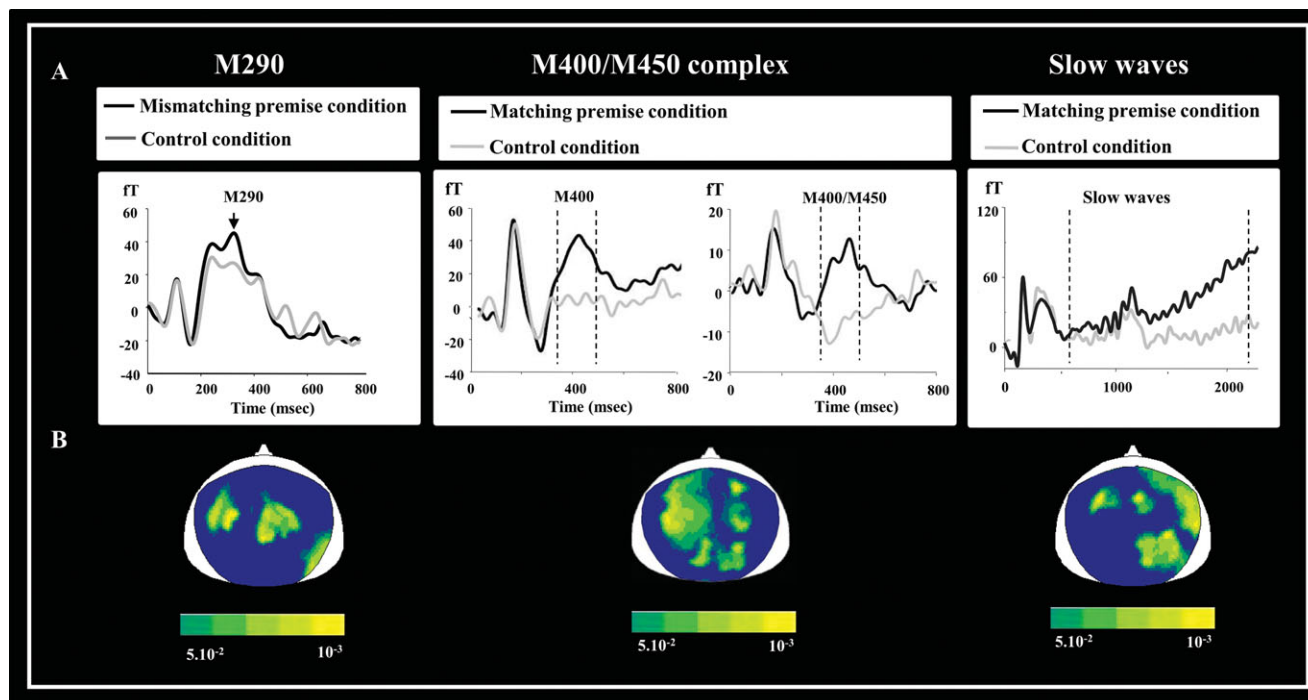


Figure 2.

A—(left) ERF profile of the mismatching premise condition and the control condition on MRC23. (middle) ERF profile of the matching premise and of the control conditions on MLT13 and MLF45. (right) ERF profile of the matching premise and of the control on MRP57. B—(left) p-map at the scalp level of the

M290 component. (middle) p-map at the scalp level of the M400/450 component. (right) p-map at the scalp level of the slow waves component.  $87 \times 46$  mm ( $300 \times 300$  DPI). [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

Permutations t tests at the source level between matching and control conditions showed that the M400/M450 complex was partly explained by a distributed network (see Fig. 4A) involving four areas: the bilateral superior parietal lobe (i.e., SPL; left :  $x = -14$ ;  $y = -80$ ;  $z = 46$ ; right:  $x = 15$ ;  $y = -75$ ;  $z = 35$ ), bilateral SFG (left :  $x = -24$ ;  $y = 22$ ;  $z = 62$ ; right:  $x = 29$ ;  $y = -3$ ;  $z = 70$ ), the right superior temporal lobe (i.e. STL:  $x = 54$ ;  $y = 53$ ;  $z = 14$ ) and the right inferior/medial frontal gyrus ( $x = 46$ ;  $y = 25$ ;  $z = 24$ ). The two last areas functionally correspond to the superior temporal sulcus and the inferior part of the dorsolateral prefrontal cortex (DLPFC) respectively (see Table II for a complete report of significantly activated areas in this time window).

Permutations t tests at the source level between these conditions in the time window of the slow waves revealed a source in the right STL ( $x = 60$ ;  $y = -18$ ;  $z = 9$ ) and in the bilateral superior frontal gyrus (left:  $x = -27$ ;  $y = 44$ ;  $z = 44$ ; right:  $x = 20$ ;  $y = 24$ ;  $z = 43$ ; see Fig. 4B). The former area corresponds to the auditory cortex and the latter to the superior part of the DLPFC (see Table II for a complete report of significantly activated areas in this time-window).

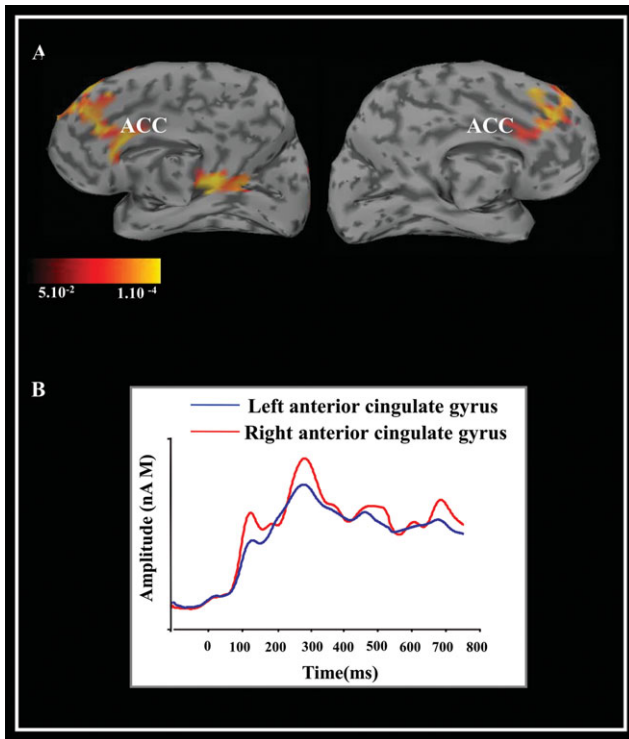
Time courses of activity within these areas revealed that the activity of the bilateral SPL, superior frontal gyrus (and also right superior temporal gyrus) peaked at 400 ms

while the right inferior/medial frontal cortex (including the DLPFC) peaked at 450 ms. The superior temporal gyrus (auditory cortex) and the bilateral superior frontal gyrus showed sustained activity during the delay (see Fig. 4C, for the clarity of the figure, only right hemisphere areas are reported).

### Matching Condition vs. Mismatching Condition

In order to compare our results to those obtained in studies using the Mismatching condition as a control, we also analyzed the contrast between the Matching and Mismatching conditions in the two time windows mentioned above. In the 370-500 ms time window (Fig. 4D), permutation t-tests at the source level revealed a similar network than the one observed for the contrast between Matching and Control conditions except for the right superior frontal gyrus (see Table II for a report of the observed activation). It is also important to note that we did not observe a bilateral frontal activation in this contrast but only the left superior frontal gyrus ( $x = -27$ ;  $y = 17$ ;  $z = 61$ ). In the 550-2,300 ms time window, permutation t-tests at the source level between these conditions revealed the same network as the one observed for the contrast between Matching and the control condition (see table II).





**Figure 3.**

A—p-map at the source level (time window = 270-320 ms). ACC = anterior cingulate cortex. B—Time course of activation of the left and right ACC.  $192 \times 225$  mm ( $300 \times 300$  DPI). [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

## DISCUSSION

In this paper, we investigated (1) the sequence of cognitive steps involved in this fundamental inference and (2) more specifically, the brain regions related to the integration of the premises in conditional reasoning. To achieve these goals, we presented a major conditional premise (e.g., If there is a square then there is a low sound) and then analyzed the processing of a minor premise (a shape) that either matches (matching condition) or mismatches (Mismatching condition) the antecedent of the major premise (e.g., a square or a circle, respectively). With respect to the potential expectations raised by the conditional premise, we expected the Matching condition to yield the magnetic counterpart of the P3b component (M300) and the Mismatching condition to yield the magnetic counterpart of the N2 component (M200), known to be produced by the ACC, since they are both known to reflect satisfaction and violation of expectations respectively [for reviews, see Folstein and Van Petten, 2008; Picton, 1992]. With respect to actual inference-making, we also expected only the Matching condition to yield multiple brain activities

linked to distinct moments in inference processing; e.g., this condition leads to the generation of a conclusion which then has to be maintained in working memory.

We will summarize our results in the processing order in which they arrive. First of all, we observed a more pronounced M200-like component, stemming from the ACC, in the mismatching condition. The activation of the ACC in the mismatching condition is typically viewed as the signature of a perceptual conflict resulting from a violation of a participant's expectations [Kerns et al., 2004] and thus confirms the hypothesis that a conditional statement yields expectations regarding the upcoming minor premise. This result is in line with the results of a previous electroencephalography (EEG) study in which a pronounced frontocentral N2 component was observed when the minor premise mismatched with the antecedent of the major premise [i.e., If P then Q; R, see Bonnefond and Van der Henst, 2009]. Another complementary explanation that accounts for the presence of N2 could be that inhibitory control is particularly active in the mismatching premise condition. Indeed, the processing of the conditional statement could generate a strategy in which the participant focuses only on the items mentioned in such a way that when they process the mismatching minor premise, they have to provide extra effort in order to determine how to deal with this stimulus, i.e., they have to inhibit their earlier strategy [see Daurignac et al., 2006; Joliot et al., 2009]. Whatever the process behind these results, and as discussed in the article of Bonnefond and Van der Henst [2009], such findings reveal that one must be cautious in using mismatching cases as a baseline of a reasoning task in fMRI [Goel et al., 2000; Goel and Dolan, 2001, 2003; Qiu et al., 2007]. That said, we think that two mitigating remarks should be made regarding the M200-like component observed in the present experiment. One is that the paradigm was designed so that there were fewer trials in the Mismatching condition than in the Matching condition (much like in the fMRI studies cited above where there were fewer trials in the Control condition than in the reasoning conditions). Such a design can actually increase the conflict generated by the mismatching premise as it has been reported that rare nontarget stimuli elicit a greater N2 than frequent nontarget stimuli [Squires et al., 1975]. Second, this component appears later (300 ms after stimulus onset) than the well-known frontocentral N2 component in EEG studies [though some EEG studies report a late N2, see van Veen & Carter, 2002]. One explanation for this disparity could be that participants sometimes reported difficulty in detecting differences between shapes (in the debriefing questionnaire).

We did not observe a more pronounced M300 in the matching condition compared to the control condition and we will discuss this absence below. However, we did find that satisfied Modus Ponens inference making (If P then Q; P) allowed us to capture inference making both temporally and spatially in the brain. In short, we found that integration led first to activation in the parietofrontal network followed

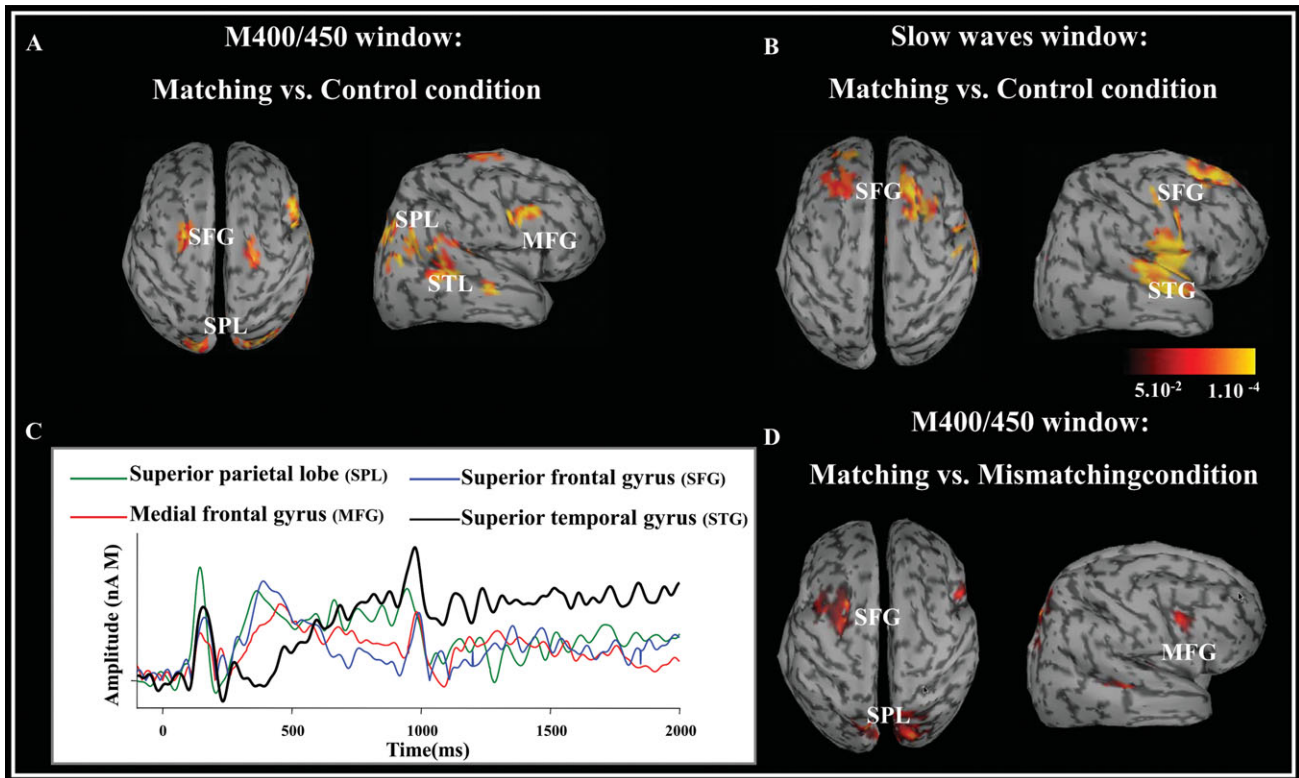
TABLE II. Brain areas activated across the different contrasts

Anatomical location	No. of vertices in cluster	Vertex-level <i>P</i> value	MNI coordinates		
			<i>x</i>	<i>y</i>	<i>z</i>
Mismatching premise condition > control condition (270–320 ms)					
L. anterior cingulate gyrus	15	<0.001	–8	35	21
L. superior frontal gyrus	18	<0.001	–1	55	41
L. posterior temporal lobe	12	<0.001	–70	–36	–5
R. anterior cingulate gyrus	24	<0.001	5	36	21
R. superior frontal gyrus	70	<0.001	6	45	37
R. posterior temporal lobe	10	<0.001	52	–57	11
R. posterior temporal lobe	10	<0.001	13	–41	2
Matching premise condition > control condition (370–500 ms)					
L. superior parietal lobe	60	<0.001	–14	–80	46
L. superior frontal gyrus	54	<0.01	–24	22	62
R. inferior/medial frontal gyrus	45	<0.001	46	25	24
R. superior frontal gyrus	45	<0.001	29	–3	70
R. superior temporal lobe	23	<0.001	54	–53	14
R. posterior/inferior temporal lobe	11	<0.001	56	–33	–4
R. superior parietal lobe	30	<0.001	15	–75	35
R. rest parietal	59	<0.001	54	–58	43
Matching premise condition > mismatching premise condition (370–500 ms)					
L. superior parietal lobe	25	<0.01	–8	–84	41
L. superior frontal gyrus	80	<0.01	–27	17	61
R. inferior/medial frontal gyrus	32	<0.01	54	19	36
R. superior temporal lobe	10	<0.05	53	–30	2
R. superior parietal lobe	34	<0.01	14	–84	46
Matching premise condition > control condition (Delay: 550–2,100 ms)					
L. superior frontal gyrus	20	<0.01	–27	44	44
L. posterior temporal lobe	29	<0.001	–60	–59	11
L. occipital cuneus	60	<0.001	–12	–68	13
L. inferior temporal lobe	36	<0.001	–52	–69	0
R. superior temporal gyrus	256	<0.001	60	–18	9
R. superior frontal gyrus	53	<0.001	20	24	43
Matching premise condition > mismatching premise condition (Delay: 550–2,100 ms)					
L. superior frontal gyrus	20	<0.05	–17	51	41
L. inferior temporal lobe	36	<0.001	–51	–72	1
R. superior temporal gyrus	230	<0.001	66	–14	11
R. superior frontal gyrus	44	<0.001	27	44	43

by an activation of the DLPFC and, finally, the recruitment of the right auditory cortex. We consider each of these later.

From the point at which the visual stimulus is presented (the minor premise), the matching vs. control contrast revealed a specific network involving parietofrontal areas with activity peaking at 400 ms, that we associate with inference generation. Moreover, at this point, and with the same contrast, one can also notice the activation of the posterior part of the superior temporal gyrus, which is linked to the superior temporal sulcus. Such a structure has been associated with audiovisual integration in numerous imaging studies [for a review, see Calvert, 2001]. Such integration is of particular interest here, since the task involves both visual and auditory stimuli whose links are enunciated by a conditional statement. In order to better relate our data to the existing fMRI literature in reasoning, we also performed the more typical matching vs. mismatching comparison. This contrast shows a pattern of

results which is similar to that obtained with the matching vs. control contrast but also reveals more activity in an area located in the left frontal cortex as compared to the right frontal cortex. This area is close to the one reported in several studies investigating propositional reasoning [Prado et al., 2010; Reverberi et al., 2010; Reverberi et al., 2007; Prado et al. in press] but is more centrally located. This difference may simply result from the lower spatial resolution of MEG as compared to fMRI and from the type of material used. It is also important to point out that one does not find evidence of activity, at this very critical stage of processing, in the left RLPFC and the mSFG, the areas that Monti et al. [2007, 2009] consider “core” regions of deduction. However, it should be noted that Monti et al. did not compare an inferential condition to a noninferential condition but contrasted two conditions that differed with respect to inferential difficulty: easy (Modus Ponens) vs. difficult (Modus Tollens: If P then Q; Not-Q;



**Figure 4.**

A—p-map at the source level (time window = 370-500 ms). SFG: superior frontal gyrus. SPL: superior parietal lobe. STL: superior temporal lobe. MFG: medial frontal gyrus. B—p-map at the source level (time window = 550-2,100 ms). SFG: superior frontal gyrus. PTL: posterior temporal lobe. C—Time course of activation of four regions of interest (SPL, SFG, MFG, and STG in the right

hemisphere). D—p-map at the source level of the contrast between matching and mismatching conditions (time window = 370-500 ms). SFG: superior frontal gyrus (from the M400-450 window). SPL: superior parietal lobe. MFG: medial frontal gyrus. 99 × 59 mm (300 × 300 DPI). [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

therefore not P). It might then be the case that the RLPFC and the mSFG are specifically involved when a task reaches a certain level of difficulty but not when an elementary inference is triggered. Thus Monti et al. might have identified regions involved in reasoning complexity but not necessarily core regions *per se*.

Following the parietofrontal activation but before the activation of the auditory cortex, we found activation of the right DLPFC with an activity peaking at 450 ms. This activation could be associated with the encoding of the generated conclusion. Indeed, this structure was also found to be activated during sound encoding by Opitz et al. [Opitz et al., 2000]. In their experiment, participants were required to assess the loudness of a sound and were then tested in a recognition task. Moreover, the timing of this activation corresponds well with such a function (i.e. it arrives after the integration of the premises but before the maintenance of the conclusion generated in working memory). However, given the weak specificity of this area, one must be cautious with such an interpretation.

Finally, the activation of the auditory cortex during the delay (i.e. 550-2,100 ms) arguably reflects the maintenance of the inferred conclusion in working memory until it can be compared to the presented conclusion. We draw this conclusion because we found that the right auditory cortex was significantly more active in the Matching condition than in the Control condition in which a sound was also expected. Several studies reveal the implication of the auditory cortex in working memory tasks (see Pasternak and Greenlee, 2005 for a review) and some studies show that the right temporal cortex is specifically implicated in tonal working memory tasks [Lancelot et al., 2003a; Lancelot et al., 2003b; Samson and Zatorre, 1988; Zatorre and Samson, 1991]. Taken together, these findings indicate that the inferential process is not an atomic process but is one composed of several phases.

We now turn to the absence of the M300 which is not in line with the EEG study [Bonnefond and Van der Henst, 2009], in which the authors found that a P3b component arose in their inferential condition. We highlight three factors that can account for such an absence. The first factor

relates to the difference between the control conditions in the two studies. In Bonnefond and Van der Henst [2009], the control condition consisted of a minor premise which was provided before the conditional statement: *P*; If *P* then *Q*. Hence, this premise did not occur in a context eliciting specific expectations regarding what the minor premise should be. This distinctly differs from the inferential condition, which consisted of a minor premise presented after a conditional statement: If *P* then *Q*; *P*. In the comparable contrast of the present task, the expectation of seeing a shape was high in the context of both conditions (context of the inferential condition: If specific geometrical shape then specific sound vs. context of the control condition: There will be a shape then a sound). In the control condition here, there is always an expectation to see the shape at the moment that it is presented. Arguably, the conditional is strong at raising expectations for the minor premise when compared to a case where there are no prior expectations [Bonnefond and Van der Henst, 2009] but may be not strong enough to distinguish itself from an instruction that tells participants to await some generic shape.<sup>2</sup>

This analysis underlines the notion that expectations are arguably not specific to inference making. By comparing an inferential condition to a control condition when both are likely to elicit expectations, as we did here, one reduces the possibility of observing expectations associated with inference making. However, one also increases the probability of delineating specific processes raised by such a mechanism. More generally, the choice of a control condition is guided by the stance one takes in investigating a cognitive process. On the one hand, if one aims at describing the highest number of processes involved in a cognitive task one should design a control condition that requires as little effort as possible so that nonspecific but relevant processing may be captured.<sup>3</sup> On the other hand, if one aims at capturing what is very specific to a cognitive mechanism one should design a control condition that minimally differs from the mechanism under investigation. In the present case a control condition that would have raised the same type of expectations as the inferential condition ought to have included a sentential context such as “there will be a square and then a low sound.” However the absence of difference, with respect to the magnetic counterpart of the P3b (i.e., M300), between the control condition and the inferential condition may already indicate these two conditions elicit a similar level of expectation.

We now turn to the second reason that may account for this absence of difference. It concerns the recording technique and the difference between MEG and EEG. What is observed through one technique cannot necessarily be

observed through the other since MEG, while having a better spatial resolution, is less sensitive to deep and radial sources. It should be noted that a difference between the M300 and the P3b was reported in the results from Croize et al. [2004] who used both EEG and MEG techniques to explore the neural correlates of working memory. In their task, participants were required to make symmetry judgements across two conditions. In the “simultaneous” condition, a single “eight-ray” figure contained dots that were either on opposite sides of each other or not as participants were required to determine whether or not the dots were symmetrical to each other. In the “delayed” condition, two such figures were presented in sequence (3 seconds apart) and the participant was required to determine whether the dot in the second figure was in a position symmetrically opposed to a dot presented in the previous figure. They found that the P3b was more pronounced in the “delayed comparison” task than it was in the “simultaneous” task. However, they did not find any difference in the M300 window across these two tasks. In light of these data and our own, the sources of the EEG P3b may have a signature that is too weak in MEG (such a phenomenon is well-known for quasiradial and deep sources).

Finally, a third factor relates to the frequency of stimuli across conditions. As indicated earlier the current design resulted in more matching trials than mismatching trials (61% vs. 39%). However, the P3b amplitude is inversely related to the frequency of target stimuli and in many experimental settings that elicit a P3b, the target stimuli are relatively rare (such as the oddball paradigm; see Picton, 1992, for review). Hence, the relatively high number of matching trials in the present design (versus 50% in Bonnefond and Van der Henst, 2009) may account for the absence of the magnetic counterpart of the P3b.

To summarize, it is difficult to say that there is one single specific network that is dedicated to valid inference making. Rather, it appears that it is much wider and it includes the generation of a conclusion, its encoding and its maintenance in working memory. Some of these reported areas of activation overlap with previous findings in the neuroimaging-of-reasoning literature, e.g., the activation of parietofrontal network.

It is important to note that our observations are based solely on Modus Ponens and the advantage this carries. Although ubiquitous, this inference has received less empirical attention than other types of inferences. The reason is that cognitive scientists have traditionally investigated reasoning by focusing on difficulty [see Johnson-Laird & Byrne, 1991; Brain and O’Brien, 1998] because such a variable can adequately test contrastive predictions. As a consequence, a fair amount of typical reasoning tasks are quite challenging for people who are not trained in logic. When included in reasoning experiments, Modus Ponens is then often used as a control for other more complex inferences (see Monti et al., 2007, for instance) and is not investigated in itself [but see Noveck et al., 2004]. However, neuroimaging offers a direct way to approach

<sup>2</sup>However, it must be noted that the presence of a M200 component in case of mismatching does indicate that the conditional statement elicits expectations about a specific shape.

<sup>3</sup>This approach is especially relevant when the description of processes is improved by access to the temporal profile of neural activity, (as it is the case with MEG; see Introduction).

inference beyond the manipulation of difficulty. While behavioral measures may hardly describe the mechanisms underlying Modus Ponens when comparing such inference to a control (i.e. non-inferential) condition<sup>4</sup> this is obviously not the case for neuroimaging. Of course what we observed for Modus Ponens may be specific to that inference and the generalization of our results will depend on the investigation of a broader range of inferences. There are actually a number of factors such as content, logical form, and complexity, which are likely to affect the neural profile we reported for Modus Ponens. For instance, it could be the case that more complicated inferences, such as backward inferences (i.e., affirmation of the consequent: If P than Q; Q therefore P) or inferences including a negation (i.e., or-elimination: P or Q; Not-P; Therefore Q) would involve the areas reported by Monti et al. [2007]. However, the present investigation demonstrates that MEG techniques provide for a promising and exciting avenue of research into the neuroimaging of reasoning.

### ACKNOWLEDGMENTS

We thank Françoise Lecaigard, Jerome Prado, Coralie Chevallier, and Guillaume Sescousse for helpful discussions. This work has been supported by Neuroreasoning, a French National Research Agency grant (ANR-07) awarded to Ira Noveck and Jean-Baptiste Van der Henst.

### REFERENCES

- Baillet S, Mosher J, Leahy R (2001): Electromagnetic brain mapping. *IEEE Signal Process Mag* 18:14–30.
- Bonnefond M, Van der Henst JB (2009): What's behind an inference? An EEG study with conditional arguments. *Neuropsychologia* 47:3125–3133.
- Braine M, O'Brien D (1998): *Mental Logic*. Mahwah, NJ, Lawrence Erlbaum.
- Brunia CH (1999): Neural aspects of anticipatory behavior. *Acta Psychol (Amst)* 101:213–242.
- Calvert GA (2001): Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cereb Cortex* 11:1110–1123.
- Croize AC, Ragot R, Garnero L, Ducorps A, Pelegrini-Issac M, Dauchot K, Benali H, Burnod Y (2004): Dynamics of parieto-frontal networks underlying visuospatial short-term memory encoding. *Neuroimage* 23:787–799.
- Daurignac E, Houde O, Jouvent R (2006): Negative priming in a numerical Piaget-like task as evidenced by ERP. *J Cogn Neurosci* 18:730–736.
- Devlin JT, Poldrack RA (2007): In praise of tedious anatomy. *Neuroimage* 37:1033–1041; discussion 1050–1038.
- Fangmeier T, Knauff M, Ruff CC, Sloutsky V (2006): FMRI evidence for a three-stage model of deductive reasoning. *J Cogn Neurosci* 18:320–334.
- Fangmeier T, Knauff M (2009): Neural correlates of acoustic reasoning. *Brain Res* 1249:181–190.
- Folstein JR, Van Petten C (2008): Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology* 45:152–170.
- Garcia-Larrea L, Cezanne-Bert G (1998): P3, positive slow wave and working memory load: a study on the functional correlates of slow wave activity. *Electroencephalogr Clin Neurophysiol* 108:260–273.
- Goel V (2007): Anatomy of deductive reasoning. *Trends Cogn Sci* 11:435–441.
- Goel V, Buchel C, Frith C, Dolan RJ (2000): Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage* 12: 504–514.
- Goel V, Dolan RJ (2001): Functional neuroanatomy of three-term relational reasoning. *Neuropsychologia* 39:901–909.
- Goel V, Dolan RJ (2003): Explaining modulation of reasoning by belief. *Cognition* 87:B11–B22.
- Goel V, Makale M, Grafman J (2004): The hippocampal system mediates logical reasoning about familiar spatial environments. *J Cogn Neurosci* 16:654–664.
- Hammers A, Allom R, Koeppe MJ, Free SL, Myers R, Lemieux L, Mitchell TN, Brooks DJ, Duncan JS (2003): Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* 19:224–247.
- Houde O (2007): First insights on neuropedagogy of reasoning. *Think Reasoning* 13:81–89.
- Howell DC (1997): *Statistical method for psychology*. Belmont, CA: Duxbury Press.
- Huang MX, Mosher JC, Leahy RM (1999): A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Phys Med Biol* 44:423–440.
- Joliot M, Leroux G, Dubal S, Tzourio-Mazoyer N, Houde O, Mazoyer B, Petit L (2009): Cognitive inhibition of number/length interference in a Piaget-like task: evidence by combining ERP and MEG. *Clin Neurophysiol* 120:1501–1513.
- Johnson-Laird PN, Byrne RMJ (1991): *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- Kerns JG, Cohen JD, MacDonald AW III, Cho RY, Stenger VA, Carter CS (2004): Anterior cingulate conflict monitoring and adjustments in control. *Science* 303:1023–1026.
- Knauff M, Fangmeier T, Ruff CC, Johnson-Laird PN (2003): Reasoning, models, and images: Behavioral measures and cortical activity. *J Cogn Neurosci* 15:559–573.
- Lancelot C, Ahad P, Noulhiane M, Hasboun D, Baulac M, Samson S (2003a): Spatial and non-spatial auditory short-term memory in patients with temporal-lobe lesion. *Neuroreport* 14: 2203–2207.
- Lancelot C, Samson S, Ahad P, Baulac M (2003b): Effect of unilateral temporal lobe resection on short-term memory for auditory object and sound location. *Ann N Y Acad Sci* 999:377–380.
- Lea R (1995): On-line evidence for elaborative logical inferences in text. *J Exp Psychol Learn* 21:1469–1482.
- Linden DE (2005): The p300: Where in the brain is it produced and what does it tell us? *Neuroscientist* 11:563–576.
- 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 Natl Acad Sci U S A* 106:12554–12559.
- Noveck IA, Goel V, Smith KW (2004): The neural basis of conditional reasoning with arbitrary content. *Cortex* 40:613–622.

<sup>4</sup>Such a comparison can however show that elementary inferences are actually made online.

- Opitz B, Mecklinger A, Friederici AD (2000): Functional asymmetry of human prefrontal cortex: encoding and retrieval of verbally and nonverbally coded information. *Learn Mem* 7:85–96.
- Pantazis D, Nichols TE, Baillet S, Leahy RM (2005): A comparison of random field theory and permutation methods for the statistical analysis of MEG data. *Neuroimage* 25:383–394.
- Pasternak T, Greenlee MW (2005): Working memory in primate sensory systems. *Nat Rev Neurosci* 6:97–107.
- Picton TW (1992): The P300 wave of the human event-related potential. *J Clin Neurophysiol* 9:456–479.
- Poldrack RA (2006): Can cognitive processes be inferred from neuroimaging data? *Trends Cogn Sci* 10:59–63.
- Prado J, Van Der Henst JB, 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 JR. The brain network for deductive reasoning: A quantitative meta-analysis of 28 neuroimaging studies. *J Cognitive Neurosci* (in press).
- Qiu J, Li H, Huang X, Zhang F, Chen A, Luo Y, Zhang Q, Yuan H (2007): The neural basis of conditional reasoning: An event-related potential study. *Neuropsychologia* 45:1533–1539.
- Reverberi C, Cherubini P, Frackowiak RS, 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 RS, 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 metaduction. *Neuropsychologia* 47:1107–1116.
- Rodriguez-Moreno D, Hirsch J (2009): The dynamics of deductive reasoning: An fMRI investigation. *Neuropsychologia* 47:949–961.
- Rohrbaugh JW, Dunham DN, Stewart PA, Bauer LO, Kuperman S, O’Connor SJ, Porjesz B, Begleiter H. (1997): Slow brain potentials in a visual-spatial memory task: topographic distribution and inter-laboratory consistency. *Int J Psychophysiol* 25:111–122.
- Samson S, Zatorre RJ (1988): Melodic and harmonic discrimination following unilateral cerebral excision. *Brain Cogn* 7:348–360.
- Senot P, Baillet S, Renault B, Berthoz A (2008): Cortical dynamics of anticipatory mechanisms in interception: A neuromagnetic study. *J Cogn Neurosci* 20:1827–1838.
- Squires NK, Squires KC, Hillyard SA (1975): Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr Clin Neurophysiol* 38:387–401.
- van Veen V, Carter CS (2002): The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiol Behav* 77:477–482.
- Zatorre RJ, Samson S (1991): Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114:2403–2417.