

The neural basis of conditional reasoning with arbitrary content

Ira A. Noveck

Institut des Sciences Cognitives

Lyon, France

Vinod Goel & Kathleen W. Smith

York University

Toronto, Canada

Abstract

Behavioral predictions about reasoning have usually contrasted two accounts, Mental Logic and Mental Models. Neuroimaging techniques have been providing new measures that transcend this debate. We tested a hypothesis from Goel and Dolan (2003) that predicts neural activity predominantly in a left parietal-frontal system when participants reason with arbitrary (non-meaningful) materials. In an event-related fMRI investigation, we employed propositional syllogisms, the majority of which involved conditional reasoning. While investigating conditional reasoning generally, we ultimately focused on the neural activity linked to the two valid conditional forms -- *Modus Ponens* (*If p then q; p//q*) and *Modus Tollens* (*If p then q; not-q//not-p*). Consistent with Goel and Dolan (2003), we find a left lateralized parietal frontal network for both inference forms with increasing activation when reasoning becomes more challenging by way of Modus Tollens. These findings show that the previous findings with more complex Aristotelean syllogisms are robust and lend doubt to accounts of reasoning that accord primary inferential processes uniquely to either the right hemisphere or to language areas.

The neural basis of conditional reasoning with arbitrary content

The cognitive literature on deductive reasoning conventionally pits two approaches against one another. One approach supposes that people have the innate ability to carry out some basic inference forms like *Modus Ponens* (if p then q ; $p // \therefore q$), *disjunction elimination* (p or q ; not- $p // \therefore q$), as well as *and-elimination* (p and $q // \therefore p$), along with a built-in program to integrate these rules (Braine, 1990; Rips, 1994; Osherson, 1975). Often referred to as the *Mental Logic* approach or the *syntactic* approach to reasoning (Braine & O'Brien, 1998; also see the Natural Deductive System, Rips, 1994; Rips, 1983), it is reasonable to assume that mental logic's neuropsychological correlates would be linked to the left-hemisphere and its language processing regions (see Osherson et al., 1998).¹ The other approach, advocated by Mental Model theory (Johnson-Laird, 1983; Johnson Laird, Byrne, & Schaeken, 1992), assumes that reasoning is based on the truth conditions of a given statement. According to this account, a reasoner forms a mental model based on the premises of an argument, arrives at a putative conclusion by discovering a new, informative proposition that is true in the model, and then looks for counter-examples. In the event no such counterexamples are found, then the conclusion is stated as a valid consequence of the premises. This approach originally assumed – given the spatial nature of models – that reasoning is a right hemisphere activity (Johnson-Laird, 1994).

Neuroimaging techniques have been providing new measures that transcend this debate (Goel, Gold et al., 1998, 1997; Osherson, Perani et al., 1998; Goel, Buchel et al., 2000; Goel and Dolan, 2001; Parsons and Osherson, 2001; Knauff, Mulack et al., 2002; Acuna,

¹ To our knowledge, the leading advocates of the mental logic approach have never linked its account to a specific neurological mechanism or brain region.

Eliassen et al., 2002; Houde, Zago et al., 2001; Houde, Zago et al., 2000; Goel and Dolan, 2003). The work by Goel & colleagues has been particularly relevant. Working mostly with quantifiers and syllogistic tasks, Goel and Dolan's experiments have led to two kinds of novel findings. One is that there is a neural effect based on content. A parietal-frontal pathway is activated when participants reason with arbitrary materials whereas a temporal-frontal system, linked to language areas, is activated when participants reason with syllogisms using realistic statements (which resonate with one's beliefs). For example, a syllogism that presents an argument using arbitrary materials like *No A are B* gives rise to activity in the parietal-frontal pathway whereas a formally equivalent syllogism using materials like *No elephants are reptiles* differentially activates a temporal-frontal system. The second finding from their work concerns the right hemisphere, which appears to be recruited in the resolution of conflict, e.g. when a syllogism's valid conclusion is not in itself believable (consider a valid syllogism whose conclusion is *Some serial killers are not mean*). This has led to a novel account that views reasoning as a dual system (Goel and Dolan, 2003). One system engages a frontal-parietal system (BA 6, 7, 44) when a reasoning task uses arbitrary materials whereas the other system works with semantically meaningful materials that activates a left hemisphere ventral network involving temporal (BA 21/22) and frontal lobe regions (BA 44, 8, 9). The right prefrontal cortex is recruited in cases where there is a conflict. It thus appears that reasoning activity is exclusively linked neither to language areas nor to visuo-spatial areas.

While Goel and Dolan's account has been valuable to the reasoning community, it could benefit from further testing. Most of Goel's findings are based on classic Aristotelian syllogisms, which use universal and existential quantifiers and are often quite challenging. Although these are used to some extent in the reasoning literature, one can make the argument that these problems are not ecologically valid (e.g. see Geurts, 2003). Moreover, the most

extensive accounts of reasoning are usually built around more fundamental sentential connectives like conditionals, which explains why most of the heated debates between the mental logic and mental model theories took place over propositional reasoning (O'Brien, Braine, & Yang, 1994; Johnson Laird, Byrne, & Schaeken, 1994; Bonatti, 1994; Johnson Laird & Byrne, 2002). This implies that a more severe test of Goel's claims is called for -- one that presents items using maximally simple propositional reasoning problems. Arguably, the most simple among these involves conditional (*if-then*) statements.

There are a few imaging studies whose tasks turn on the connective *if-then* (Houdé et al., 2000 ; Houdé et al. 2002; Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002; Parsons & Osherson, 2001). However, it is difficult to draw strong conclusions from these studies either because a) the tasks investigated require much more than fundamental conditional inference-making or b) the tasks employ less than optimal experimental conditions. To give an example of (a), consider how Houdé trained participants to find falsifying cases of conditional statements like *If there is not a blue triangle then there is a red circle* (which requires a total mismatch of objects, i.e. an object that is not a blue triangle and an object that is not a red circle). Successful performance on this difficult task has arguably little to do with Modus Ponens or Modus Tollens per se but with getting beyond perceptual "traps" and knowing how to falsify conditionals. To give an example of (b), consider a conditional reasoning paradigm (Knauff et al., 2002) whose "conditional problems" presented both Modus Ponens and Modus Tollens arguments while employing realistic materials (*If a man is in love then he likes pizza*) as part of a single condition; these were compared to performance on relational reasoning problems employing arbitrary colored shapes (*the red rectangle is to the left of the green rectangle*). Important distinctions among conditional inferences are lost in this comparison and (with the benefit of hindsight) one can see that the neural content effect does not render

interpretable the experiment's comparison between spatial reasoning and conditional reasoning.

With the aim of more finely testing Goel and Dolan's hypothesis concerning the left parietal-frontal system, we use arbitrary content while investigating propositional reasoning (with a particular focus on conditionals to be described in the next section below). If Goel and Dolan's account is correct, one would predict activity predominantly in a left parietal-frontal network even for these fundamental kinds of inferences. Such a finding could put to rest the alternatives considered earlier that predict: 1) that language areas are necessary and sufficient for reasoning² or; 2) that right hemisphere spatial areas are necessary and sufficient for reasoning.

Experiment

Here we begin by presenting a brief description of conditional reasoning and the typical performance it prompts in cognitive studies. This is followed by a description of the experiment in detail. Conditional statements like *If A then C* are the most studied of the connectives (Johnson-Laird & Byrne, 2002). The first portion following the *If* statement is known as the *antecedent* and the portion following *then* is known as the *consequent*. As described earlier, there are two valid inference forms related to the conditional. One is called *Modus Ponens* and the other *Modus Tollens*. *Modus Ponens* is straightforward. If someone expresses *If A then C* and then confirms that *A* is the case, that justifies the conclusion *C*. *Modus Tollens* requires more effort to carry out: When the major premise is *If A then C* and the minor premise is *not-C*, this justifies the conclusion *not-A*. To put it in concrete terms, when one is informed that *If it is raining then the sidewalk is wet* along with the information that the *sidewalk is not wet*, one can logically conclude that *it is not raining*. *Modus Tollens* can be considered a 4 step process. First, one supposes the antecedent of the major premise's

² We define language regions here as Broca's and Wernicke's areas.

conditional (*It is raining*). Second, the supposition along with the conditional leads to one sort of conclusion -- *the sidewalk is wet*. Third, the supposed conclusion (from Step 2) and the given information (that *the sidewalk is not wet*) leads to a contradiction. Fourth, this contradiction warrants the elimination of the supposition (leading to *it is not the case that it is raining*). This kind of strategy is referred to as *Reductio Ad Absurdum*.

Although the paper focuses on the valid forms just described, no paper on conditional reasoning would be complete without presenting the two other – non-valid – inference forms related to the conditional. One is called the *Denial of the Antecedent*. As the name suggests, this arises when the antecedent of a conditional is negated, as in *If A then C* along with *not-A*, which often leads reasoners to conclude *not-C*. Consider *If it is raining then the sidewalk is wet* plus the fact that *it is not raining*; these premises tempt many people to accept the conclusion *the sidewalk is not wet*. However, this conclusion is not justified by formal logic (consider how a sunny day does not preclude a wet sidewalk). In the event that a participant is presented with these two kinds of premises, a logically appropriate response is to say *Inconclusive* (Can't Tell). The other non-valid inference form, the *Affirmation of the Consequent*, where one is confronted with *if A then C* and the information *C*, is also *Inconclusive*. Nevertheless, this form too leads many reasoners to accept a non-valid conclusion, in this case *A* (consider *If it is raining then the sidewalk is wet* plus the fact that *the sidewalk is wet*; these two premises might well invite the conclusion *it is raining*). However, the consequent could be true (the sidewalk could be wet) for reasons not having to do with the antecedent (e.g. the street cleaner just came by, someone has watered a lawn next to the sidewalk). To complete the design, these latter two inference forms were included in our study, but we were focused on the two valid conditional inference forms, Modus Ponens and Modus Tollens.

Normally, reasoners evaluate simple Modus Ponens arguments correctly at rates that are at least 90% and that are higher than for Modus Tollens, which typically yields correct responses that are at rates of about 60% (Noveck & Politzer, 1998). So, one can see right away that not all logical inferences yield similar behavioral results. Both of the traditional cognitive accounts address this difference in performance; for Mental Logic, Modus Tollens requires several strategic steps and for mental models, a greater number of models (that arise after a *fleshing out* procedure renders explicit more information about the conditional). The rates of *correct* responses to the potential fallacies (i.e. to say that nothing follows) typically range between 20% and 70%.

Before introducing the experiments, we discuss our baseline condition, which was carried out with a propositional syllogism having a trivially true conclusion. For example, given:

1) If there is a black rectangle then there is a blue circle.

There is a red triangle.

The conclusion *There is a red triangle* is a true conclusion, albeit a trivial one, because if the premise is true so is the conclusion. This baseline problem amounts to a slight improvement over previous ones in the neuroimaging literature and for two reasons. Unlike tasks that provide irrelevant conclusions, which for (1) would provide a conclusion like *There is a green star*, the trivially true conclusion provides for a definitive *true* response like one would expect in a Modus Ponens problem.³ The other is that it forces participants to remain engaged throughout the problem, even as they see that the second premise does not produce the means for a logical inference. With this aim of seeking participants' engagement throughout a given reasoning item, the task to be described in the next section also provides similar problems

³ Note too that tasks that present irrelevant conclusions typically require a more academic *validity* judgment.

containing false conclusions, like *There is not a red triangle* for (1) above, which are equally trivial, though *false*.

Method

Participants. Sixteen right-handed normal participants (9 males and 7 females) with a mean age of 26.7 (SD 5.9) and mean education level of 18.9 (SD 2.8), volunteered to participate in the study. All participants gave informed consent and the study was approved by the York University Ethics Committee and the Ethics Committee of the Robarts Research Institute at the University of Western Ontario.

Stimuli. Twenty items of each of the four conditional arguments were presented along with 10 of the baseline condition. All of the conclusions were affirmative so as to make items across conditions as similar as possible. The materials used colored shapes. For example, one Modus Ponens item looked like (2) below:

2) If there is a black square then there is a yellow circle.

There is a black square.

There is a yellow circle.

Six different colors (red, blue, green, yellow, black, and white) and six different shapes (square, circle, triangle, cross, star, rectangle) were used. No color or shape was used more than once in a given conditional sentence. These materials were employed (as opposed to letters) in order to provide reading and reaction times that correspond well with imaging timing. Efforts were made to keep the length of each item equivalent. The mean number of characters per syllogism was 104.8, ranging from 99 to 111.

The study included several forms of distractors. So as not to create expectations among participants, negative conclusions were presented in 3 other items for each of the conditional arguments. For example, if the Modus Ponens item above (in 2) were converted so as to have a negative conclusion, its third line would read *There is not a yellow circle* (of

course, this is just an example as colored shapes change for each item). The same holds for the other experimental conditions. In the baseline condition, however, there were 10 equivalent problems presenting a negative conclusion inconsistent with the second premise. Besides the 112 total items described that constitute this study, there were an additional 61 filler items using the same arbitrary materials, though concerning disjunctions. Finally, there were 19 rest periods, which were presented at the end of each of two experimental sessions (10 in one session and 9 in the other). These showed a series of X's in place of each of the lines where the syllogisms had been. Participants were asked to provide a response (any response) upon arrival of the line corresponding with the conclusion.

Training. Before going into the scanner, participants were given a training session in which the form of a problem was described (by saying that it contains two premises and a conclusion) along with its materials (they were told that the information was the basis for determining the presence or absence of colored shapes on a hidden blackboard). Participants were told to assume that the two premises are true and that they have to determine whether the provided conclusion logically follows from them. They were further told that if the conclusion follows, they are to respond *true*, if the conclusion is false then to respond *false* and if they can't tell to choose *inconclusive*. This was followed by seven examples that covered the three possible response options. The *inconclusive* option was exemplified by a problem containing the form, *not both p and q; not-p//q*. An example of a baseline problem was also included in order to demonstrate that sometimes a conclusion is trivial. The training and each of the seven examples were administered by the experimenter personally. In the event of a question or a misunderstanding, the experimenter aimed to resolve it without giving formal logical instruction.

Stimuli presentation. The stimuli were presented in one of two random orders. Each of these was broken up into two blocks so as to provide a break in the middle. Half of the

participants began with the first block in a given order and half with the second block. Thus, there were, in effect, 4 different presentations of the materials. Each ended with the presentation of a series of rest sessions.

A given reasoning item was presented in the following way. The beginning of the trial was signaled by a “*” in the left hand corner where the beginning of the first premise was to appear. The sentences then appeared on the screen one line at a time, with the first premise appearing at 500 msec, the second at 4000 msec, and the last sentence at 6500 msec. All three of the sentences remained on the screen until the end of the trial (10000 msec). Participants responded by pressing one of three buttons on a keypad after the appearance of the last sentence (see Figure 1 for a representative description). Participants were instructed to respond naturally and efficiently so as to be prepared to read the next trial. They were also told to let the trial pass and to focus on the upcoming problem in the event they could not respond quickly enough.

Insert Figure 1 about here

fMRI Scanning Techniques. A 4 Tesla Oxford Magnet Technologies magnet with a Siemens Sonata gradient coil was used to acquire T1 anatomical volume images (0.742 x 0.742 x 2.73 mm voxels) and twenty-two T2* -weighted interleaved multi-shot contiguous echoplanar images (3 x 3 x 6-mm voxels,), sensitive to blood oxygenation level dependent (BOLD) contrast. The images were acquired axially and positioned to cover the whole brain. Data were recorded during a single acquisition period. A total of 652 volume images were acquired over two sessions, (326 vol per session) with a repetition time (TR) of 3 s/vol. The first six volumes in each session were discarded to allow for T1 equilibration effects (leaving 320 volumes per session).

Trials from all conditions were randomly presented in a single-event design with a SOA of 10,000 msec. Each of two sessions consisted of 96 events (including rest trials) for a total of 192 events. Each session lasted 16.1 minutes. The scanner was synchronized with the presentation of every 3rd trial.

Data Analysis. Data were analyzed using Statistical Parametric Mapping, SPM 99 (Friston, Holmes et al., 1995). All volumes were spatially realigned to the first volume. Head movement was less than 2 mm in all cases. A mean image created from the realigned volumes was spatially normalized to the Montreal Neurological Institute brain template (Evans, Collins et al., 1993) using nonlinear basis functions (Ashburner and Friston, 1999). The derived spatial transformation was then applied to the realigned T2* volumes, which were finally spatially smoothed with a 12 mm FWHM isotropic Gaussian kernel (in order to make comparisons across subjects and to permit application of random field theory for corrected statistical inference (Worsley and Friston, 1995). The resulting time series across each voxel were high-pass filtered with a cut-off of 120 sec, using cosine functions to remove section-specific low frequency drifts in the BOLD signal. Global means were normalized by proportional scaling to a Grand Mean of 100.

Condition effects at each voxel were estimated according to the general linear model and regionally specific effects were compared using linear contrasts. Each contrast produced a statistical parametric map of the t-statistic for each voxel, which was subsequently transformed to a unit normal Z-distribution. The activations reported survived a voxel-level intensity threshold of $p < .05$ using a random effect model (corrected for multiple comparisons using False Discovery Rate) (Genovese, Lazar et al., 2002). Given the principle of anatomical symmetry, where an anatomical structure is significantly active in one hemisphere, we report any activation in the corresponding structure in the other hemisphere, even if it does not survive correction for multiple comparisons. The BOLD signal was

modeled as a canonical hemodynamic response function (hrf) with a time derivative at the presentation of the third sentence. The presentation of the two premises as well as motor responses were modeled out of the analysis.

Results

Behavioral scores were in keeping with expectations (see Table 1). Modus Ponens items prompted higher rates of correct performance (93%) than Modus Tollens (77%), leading to a significant difference, $p < .005$, one-tailed t-test). Rates of correct performance on the baseline condition (87%) was not statistically different from modus ponens, ($p > .15$, two-tailed) or Modus Tollens ($p > .15$, two-tailed).⁴ The two conditions that require an *inconclusive* response, the Denial of the Antecedent and the Affirmation of the Consequent conditions, prompted rates of correct responses (31% and 29%, respectively) that were in line with findings in the literature. Only correct responses were included in the imaging analyses.

 Insert Table 1 about here

The last column of Table 1 represents the response reaction times to correct responses only. These too are not surprising to the extent that Modus Ponens is significantly faster than the other three conditional inferences (Barrouillet, Grosset & Lecas, 2000). Note too that the Modus Ponens condition is not significantly faster than the baseline condition. We will focus mostly on Modus Ponens and Modus Tollens because they are both valid forms (and thus comparable to one another) and because they both prompted a high rate of correct responses. The data from the two Inconclusive forms are based on fewer data points and are affected by

⁴ Errors to the baseline problems were attributable to two participants, whose responses reflect difficulty in accepting a trivial conclusion. When one eliminates these two participants from the analysis, rates of correct responses to the baseline increase to 99%.

strong individual tendencies (the number of correct responses for any individual participant ranged between 1 to 19 items for each of these two inference forms).

To isolate conditional inference making, we made three main comparisons. First, we compared Modus Ponens with baseline. This resulted in activation of the left superior parietal lobule (BA7), the left lingual gyrus (BA 19), and the left inferior temporal lobe (BA37). Second, we compared Modus Tollens with baseline.⁵ This resulted in activation in the left superior parietal lobe (BA 7), the left cingulate gyrus (BA 32), the left middle frontal gyrus (BA 6), and the left inferior prefrontal gyrus (BA 47). (The right inferior/middle prefrontal cortex was also activated at a reduced threshold of $p > .01$ [54, 24, 30; $Z=2.78$].) Finally, the inference forms, Modus Ponens and Modus Tollens, were directly compared to one another. An analysis of Modus Tollens minus Modus Ponens resulted in activation in the left inferior parietal lobe (BA 40), the left cingulate gyrus (BA 32), the left dorsal lateral prefrontal cortex (BA 9) and the inferior prefrontal cortex (BA 47). (Again, the right inferior prefrontal cortex was activated at a reduced threshold of $p > .01$ [51, 30, 27; $Z=2.66$].) A reverse analysis (Modus Ponens minus Modus Tollens) did not reveal any significant activation. These results are summarized in Table 2 and in Figure 2.⁶ We also compared the two *Inconclusive* inference forms to the baseline condition. Neither the *Denial of the Antecedent* condition nor *Affirmation of the Consequent* condition yielded significant results.

⁵ We chose to use the same, minimal baseline for investigating both Modus Ponens and Modus Tollens for two reasons. One is that it makes intuitive sense to use just one basic baseline for all inference forms and the other is that the alternative (a second baseline for investigating Modus Tollens from the problems having a negative conclusion inconsistent with the second premise) would render less informative the analysis of simple effects for MP and MT.

⁶ Two additional analyses were conducted. One carried out the same analyses that included only the 14 participants who correctly responded to the baselines (as instructed); these revealed highly similar findings. The other was based on recordings that began with the arrival of the second premise (to address the possibility that inference-making actually begins as soon as the second premise is read and integrated); here too, the results looked very much the same as reported.

Insert Table 2 about here

Insert Figure 3 about here

Discussion

This investigation aimed to test predictions from Goel and Dolan's dual system hypothesis which anticipates that reasoning with arbitrary materials ought to lead to left parietal-frontal activation. This was largely supported. The left parietal regions, especially the left superior parietal lobule (BA7), were activated in both Modus Ponens and Modus Tollens problems. The left lateralization of the parietal frontal network is consistent with the findings from Goel et al. (Goel, Buchel et al., 2000; Goel and Dolan, 2003) where the most stringent tests of "no-content" minus "content" reasoning activated the left parietal cortex. The present study, by revealing highly similar activations in this area with valid propositional forms, shows that these prior effects are robust.

How do the present findings relate to others on reasoning in the neuropsychological literature? Although there are not many patient studies on reasoning, the few that exist consistently implicate the left hemisphere in logical reasoning and inference-making (Gazzaniga and Smylie, 1984, Caramazza et al., 1976, Read, 1981). For example, Caramazza et al. (1976) administered to brain-damaged patients two-term problems such as the following: "Mike is taller than George. Who is taller?" They reported that left hemisphere patients were impaired in all forms of the problem but right hemisphere patients were only impaired when the form of the question was incongruent with the premise (e.g. who is shorter?). In a recent study using matched verbal and spatial reasoning tasks, Langdon and Warrington (2000) found that only left hemisphere patients failed the verbal section, while

both left and right hemisphere patients failed the spatial sections. They concluded by emphasizing the critical role of the left hemisphere in both verbal and spatial logical reasoning. These studies are very consistent with Goel et al.'s (2000) claim that the left hemisphere is necessary and sometimes sufficient for logical reasoning, while the right hemisphere is sometimes necessary, but not sufficient. One notes greater activation in the parietal frontal pathway when reasoning becomes more challenging by way of Modus Tollens. This is arguably due to the greater reasoning resources required to carry out this inference. Here we consider three possible contributions to this increased activity.

First, as we reported, right prefrontal activity becomes more apparent during Modus Tollens inference-making, although it is detectable only at reduced thresholds. This can be viewed as consistent with earlier findings (described in the Introduction) showing right prefrontal activity in cases that require some amount of conflict resolution (e.g. making appropriate validity judgments about non-believable conclusions) because Modus Tollens also involves a certain amount of conflict resolution. On the one hand, the very mention of the conditional in the major premise could make its antecedent appear possible while, on the other, the correct response requires an outright rejection of the same antecedent. It thus appears that right frontal areas can be recruited for the purposes of reconciling conflicting information even within the confines of a proper logical inference.

A second way Modus Tollens requires greater resources can be found in its series of *Reductio Ad Absurdum* steps (described earlier), where a reasoner makes a supposition, sees that it leads to a contradiction, and then rejects the supposition. This leads to four steps (three more than for Modus Ponens) involving the temporary storage of a supposition, the evaluation of a concomitant conclusion and the subsequent rejection of the supposition. All this requires working memory. Moreover, the increased activity in the ventral aspect of the frontal inferior

gyrus (BA 47) may be involved specifically in the process of selecting among plausible alternative suppositions (Goel & Dolan, 2000; Rogers et al., 1999).

Finally, another possibility is that Modus Tollens arrives by way of a bias so that the negation of the consequent also readily negates the antecedent without the intermediate steps that justify this conclusion (e.g. Evans, 1983). Perhaps the bias, which is envisaged as a low-cost maneuver that bypasses heavy processing, is associated with the increased activity. The present methods do not allow us to distinguish carefully between these two latter accounts, but the data are arguably more favorable to the Reductio account. That Modus Tollens prompts some amount of right frontal activity on the sort often found when reasoners deal with potentially inconsistent information does more readily indicate that Tollens is the result of an effort to carefully work out the correct response.

The fact that the Denial of the Antecedent and Affirmation of the Consequent conditions do not provide any noticeable activity relative to the baseline condition is noteworthy. One possibility is that the low rates of correct performance (roughly 30% in each condition) block us from getting a critical mass of data to analyze (because only correct responses were included); a mean rate of six correct trials per participant is not sufficient for a reliable signal. Another possibility is that participants who respond correctly to this problem do so because a correct response is an indication that no further processing is required on the critical concluding sentence; in effect, there is no more reasoning to do once it is recognized that the two premises prompt no valid inferences. All that is left to do is to press the “Inconclusive” button.

We now compare the present findings to other studies of reasoning in the functional neuroimaging literature. Aside from the work spearheaded by Goel and his colleagues, there are two other sets of (PET) studies from Houdé and colleagues (Houdé, Zago et al., 2001; Houdé, Zago et al., 2000) and Parsons and colleagues (Osherson et al, 1998; Parsons and

Osherson, 2001), whose results are relevant. We begin with Houdé and colleagues' who find a shift from a bilateral posterior area to a left-frontal network after training on the conditional falsification task described earlier (where participants have to find counterexamples to *if p then q*). To the extent that Houdé and colleagues find an increase in reasoning activity in the left hemisphere, their data are largely consistent with our findings here showing that reasoning activity increases in the frontal regions with the more complex Modus Tollens inference. However, the principal areas of activity in their reported shift do not correspond with those we report here. This is probably due to the task in question. As highlighted in the Introduction, Houdé et al's task requires much more than conditional reasoning alone. Participants are not only required to represent a complex conditional rule with a negative antecedent (e.g. *if there is not a red square on the left, then there is a yellow circle on the right*), they are required to find counterintuitive counterexamples to it. One can see that such a task is much more effortful and complex than the one described here which aims to maximally isolate activation to logical inference-making. For example, in our Modus Ponens problem, a participant simply confirms that a conclusion follows from the premises and in our Modus Tollens problem, the participant has to disconfirm the conclusion.

Parsons and colleagues report that deduction prompts activity in right hemisphere homologues of left language areas in the middle temporal lobe, inferior frontal areas and basal ganglia. Moreover, they find no activity in the spatial-parietal areas. We see three possible reasons why their findings diverge from our own. One is that the tasks in their study are quite different from ours. In their *deductive* condition Parsons & Osherson (2002) asked participants to determine whether a provided conclusion followed logically in arguments like the following:

(3) If his home is mortgaged then he is married.

It is not true that he is both married and has children.

His home is not mortgaged.

To determine why this argument does not provide a logical conclusion, one has to suppose that “he” is either married without children, is not married and has children, or is neither married nor has children. In the two cases in which he is not married, it follows that his home is not mortgaged (by way of Modus Tollens). In the remaining case (where he is married), one cannot draw any valid conclusions. Taking all the possible cases together, the correct response is to say something along the lines of *inconclusive*. In fact, each of the study’s six provided arguments are like the one in (3) and do not provide the means for drawing a valid conclusion. If participants provided a definitive response, it was based on non-valid steps or invited inferences. In short, the tasks in Parson and Osherson’s study do not localize logical reasoning per se, but allow for extensive reasoning about ultimately inconclusive outcomes. Unlike in Parsons and Osherson, our approach has been to isolate, as cleanly as possible, the inference-making activity related to valid conditional reasoning.

Secondly, Parson and Osherson (2002) do not specify what conditions were activated in the simple comparison of deduction minus sentence comprehension or rest baselines. They only state that the language areas were not activated. The comparison that they do report is deduction vs. probabilistic reasoning and there is no reason to believe that this activation should correspond to the deduction vs. baseline activation in our study.

The third reason concerns the PET methodology used in both groups of studies. The Pet methodology results in important limitations because a) it is not possible to separate reading of the stimuli from the reasoning task, b) it is impossible to distinguish between correct and incorrect trials, and c) the number of trials that can be presented per injection is limited by the half life of the isotope (approximately 40 sec for H₂O studies). The Parsons and Osherson results are based on 6 trials per injection. Our single event fMRI design does not suffer from these limitations. Using a single sentence presentation format and modeling

the response with a canonical hrf function at the presentation of the conclusion allows us to largely parcel out reading from reasoning. The design also allows us to separate correct from incorrect trials. Lastly we were able to present 20 trials of each argument type to ensure an adequate, stable BOLD signal response.

One question for future work is why one is more likely to find left inferior parietal activity with arbitrary materials than with meaningful content. One possibility is that arbitrary materials render salient the logical procedures that are needed to respond in experiments like ours. When materials become more meaningful (i.e. they more closely resemble conversational exchanges) they prompt a listener to engage in a wider array of inferences, making the logical inference appear less prominent. The parietal activity identified here can be said to work as part of a more general purpose system; one whose importance is potentially superseded as more specific information becomes available.

In summary, this work provides the strongest evidence yet that fundamental inference forms, like those found in conditionals, activate left parietal regions when the reasoning materials are arbitrary. This includes the case of Modus Ponens, which is among the most basic and easily produced of logical inferences. This lends doubt to accounts of reasoning that accord primary inferential processes to other areas in the brain.

Acknowledgements

This work was supported by grants from the CNRS (ATIPE, 1999) to the first author and by a McDonnell-Pew Program in Cognitive Neuroscience Award to VG. Correspondence concerning this article should be addressed to Ira Noveck, Institut des Sciences Cognitives, 67 Blvd. Pinel, 69675 Bron, France (email: noveck@isc.cnrs.fr) or to Vinod Goel, Department of Psychology, York University, Toronto, Ontario M3J 1P3 Canada (vgoel@yorku.ca).

References

- ACUNA BD, ELIASSEN JC, DONOGHUE JP, and SANES JN. Frontal and parietal lobe activation during transitive inference in humans. *Cereb Cortex*, 12: 1312-21, 2002.
- ASHBURNER J and FRISTON KJ. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7: 254-66, 1999.
- BARROUILLET, P. GROSSET, N., & LECAS, J. Conditional reasoning by mental models: chronometric and developmental evidence. *Cognition*, 75, 237-266, 2000.
- BONNATI, L. Propositional reasoning by model? *Psychological Review*, 101(4), 725-733, 1994.
- BRAINE, M. D. S., & O'BRIEN, D. P. *Mental logic*. Mahway, NJ: Erlbaum, 1998.
- CARAMAZZA A, GORDON J, ZURIF EB, and DELUCA D. Right-hemispheric damage and verbal problem solving behavior. *Brain Lang*, 3: 41-6, 1976.
- CORBETTA M, MIEZIN FM, DOBMEYER S, SHULMAN GL, and PETERSEN SE. Selective and divided attention during visual discriminations of shape, color, and speed: Functional anatomy by positron emission tomography. *J Neurosci*, 11: 2383-402, 1991.
- CORBETTA M, MIEZIN FM, SHULMAN GL, and PETERSEN SE. A pet study of visuospatial attention. *J Neurosci*, 13: 1202-26, 1993.
- EVANS AC, COLLINS DL, MILLS SR, BROWN ED, KELLY RL, and PETERS TM. 3d statistical neuroanatomical models from 305 mri volumes. *Proc. IEEE-Nuclear Science Symposium and Medical Imaging Conference*: 1813-1817, 1993.
- EVANS, J. S. B. Linguistic determinants of bias in conditional reasoning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 35a(4), 635-644, 1983.
- FRISTON K, HOLMES A, WORSLEY K, POLINE J-B, FRITH C, and FRACKOWIAK R. Statistical parametric maps in functional imaging: A general approach. *Human Brain Mapping*, 2: 189-210, 1995.
- GAZZANIGA MS and SMYLYE CS. Dissociation of language and cognition. *Brain*, 107: 145-153, 1984.
- GENOVESE CR, LAZAR NA, and NICHOLS T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15: 870-8, 2002.
- GEURTS, B. Reasoning with quantifiers. *Cognition*, 86(3), 223-225, 2003.
- GOEL V, BUCHEL C, FRITH C, and DOLAN RJ. Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage*, 12: 504-514, 2000.
- GOEL V and DOLAN RJ Anatomical Segregation of Component Processes in an Inductive Inference Task. *Journal of Cognitive Neuroscience*, 12(1), 1-10, 2000.
- GOEL V and DOLAN RJ. Functional neuroanatomy of three-term relational reasoning. *Neuropsychologia*, 39: 901-909, 2001.
- GOEL V and DOLAN RJ. Explaining modulation of reasoning by belief. *Cognition*, 87: B11-22, 2003.
- GOEL V, GOLD B, KAPUR S, and HOULE S. The seats of reason: A localization study of deductive & inductive reasoning using pet (o15) blood flow technique. *Neuroreport*, 8: 1305-1310, 1997.
- GOEL V, GOLD B, KAPUR S, and HOULE S. Neuroanatomical correlates of human reasoning. *Journal of Cognitive Neuroscience*, 10: 293-302, 1998.
- HOUDE O, ZAGO L, CRIVELLO F, MOUTIER S, PINEAU A, MAZOYER B, and TZOURIO-MAZOYER N. Access to deductive logic depends on a right ventromedial prefrontal area devoted to emotion and feeling: Evidence from a training paradigm. *Neuroimage*, 14: 1486-92, 2001.

- HOUDE O, ZAGO L, MELLET E, MOUTIER S, PINEAU A, MAZOYER B, and TZOURIO-MAZOYER N. Shifting from the perceptual brain to the logical brain: The neural impact of cognitive inhibition training. *J Cogn Neurosci*, 12: 721-8, 2000.
- JOHNSON LAIRD, P. N., BYRNE, R. M., & SCHAEKEN, W. Propositional reasoning by model. *Psychological Review*, 99(3), 418-439, 1992.
- JOHNSON LAIRD, P. N., & BYRNE, R. M. J. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109(4), 646-678, 2002.
- JOHNSON LAIRD, P. N., BYRNE, R. M. J., & SCHAEKEN, W. Why models rather than rules give a better account of propositional reasoning: A reply to Bonatti and to O'Brien, Braine, and Yang. *Psychological Review*, 101(4), 734-739, 1994.
- JOHNSON-LAIRD, P. N. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press, 1983..
- JONIDES J, SCHUMACHER EH, SMITH EE, KOEPPE RA, AWH E, REUTER-LORENZ PA, MARSHUETZ C, and WILLIS CR. The role of parietal cortex in verbal working memory. *J Neurosci*, 18: 5026-34, 1998.
- KNAUFF M, MULACK T, KASSUBEK J, SALIH HR, and GREENLEE MW. Spatial imagery in deductive reasoning: A functional mri study. *Brain Res Cogn Brain Res*, 13: 203-12, 2002.
- LANGDON D and WARRINGTON EK. The role of the left hemisphere in verbal and spatial reasoning tasks. *Cortex*, 36: 691-702, 2000.
- MELLET E, TZOURIO N, CRIVELLO F, JOLIOT M, DENIS M, and MAZOYER D. Functional anatomy of spatial mental imagery generated from verbal instructions. *The Journal of Neuroscience*, 16: 6504-12, 1996.
- NOBRE AC, SEBESTYEN GN, GITELMAN DR, MESULAM MM, FRACKOWIAK RS, and FRITH CD. Functional localization of the system for visuospatial attention using positron emission tomography. *Brain*, 120: 515-33, 1997.
- NOVECK, I. A., & POLITZER, G. Leveling the playing field: Investigating competing claims concerning relative inference difficulty. *Braine, Martin D. S. (Ed); O'Brien, David P. (Ed); Mental logic. (pp. 367 384). Mahwah, NJ: Lawrence Erlbaum., 1998.*
- O'BRIEN, D. P., BRAINE, M. D. S., & YANG, Y. Propositional reasoning by mental models? Simple to refute in principle and in practice. *Psychological Review*, 101(4), 711-724, 1994.
- OSHERSON, D. N. Models of logical thinking. In R. Falmagne (Ed.), *Reasoning : Representation and process in children and adults (pp. 81-91)*. Hillsdale, NJ: Lawrence Erlbaum, 1975.
- OSHERSON D, PERANI D, CAPPAS S, SCHNUR T, GRASSI F, and FAZIO F. Distinct brain loci in deductive versus probabilistic reasoning. *Neuropsychologia*, 36: 369-76, 1998.
- PARSONS LM and OSHERSON D. New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cereb Cortex*, 11: 954-65, 2001.
- READ DE. Solving deductive-reasoning problems after unilateral temporal lobectomy. *Brain and Language*, 12: 116-127, 1981.
- RIPS, L. *The Psychology of Proof: Deduction in human thinking*. Cambridge, MA: MIT Press, 1994.
- RIPS, L. J. Cognitive processes in propositional reasoning. *Psychological Review*, 90(1), 38-71, 1983.
- ROGERS, R. D., OWEN, A. M., MIDDLETON, H. C., WILLIAMS, E. J., PICKARD, J. D., SAHAKIAN, B. J., & ROBBINS, T. W.. Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *J Neurosci*, 19(20), 9029-9038, 1999.

- SMITH EE and JONIDES J. Neuroimaging analyses of human working memory. *Proc Natl Acad Sci U S A*, 95: 12061-8, 1998.
- WORSLEY KJ and FRISTON KJ. Analysis of fmri time-series revisited - again. *Neuroimage*, 2: 173-181, 1995.

Table 1

Summary of conditional reasoning performance for the four conditional forms as well as the novel baseline condition.

Inference form	Provided Premises	Provided conclusion	Correct response	Percentage correct	Mean Response Time
Modus ponens	If P then Q. P.	Q.	True	94	1260
Denial of Antecedent	If P then Q. Not-P.	Q.	Inconclusive	31	1844
Affirmation of Consequent	If P then Q. Q.	P.	Inconclusive	29	1605
Modus Tollens	If P then Q. Not-Q.	P.	False	77	1542
Baseline	If P then Q. R.	R.	True	87	1230

Notes. Mean response times are based on correct responses only.

Table 2

Coordinates and Z-scores for regions showing differential activation during Modus Ponens reasoning versus baseline, Modus Tollens versus baseline, and Modus Tollens versus modus ponens.

Location	Brodmann area	MNI Coordinates			Z-score
		x	y	z	
Modus Ponens – baseline					
Left superior parietal	7	-21	-75	30	3.78
	7	-15	-72	27	3.46
Left lingual gyrus	19	-21	-72	3	3.43
Left inferior temporal lobe	37	-48	-63	-3	3.30
Modus Tollens – baseline					
Left superior parietal lobule	7	-33	-51	30	3.63
	7	-15	-51	51	3.23
	7	-24	-75	27	3.48
Left cingulate gyrus	32	-6	12	45	3.89
Left middle frontal gyrus	6	-36	0	45	3.53
Left inferior prefrontal cortex	47	-54	15	-9	3.64
Modus Tollens – Modus ponens					
Left inferior parietal lobe	40	-39	-54	39	3.80
Left cingulate gyrus	32	-3	18	42	3.70
Left dorsal lateral prefrontal cortex	9	-45	15	39	3.56
Left inferior prefrontal cortex	47	-51	24	-9	3.15

Figure 1.

A graphic description of experimental presentation.

Figure 2.

First row (a) and (b): Modus Ponens vs. baseline activated left superior parietal lobule (BA 7) (-21, -75, 30; $Z= 3.78$ & -15, 72, -27; $Z=3.46$), left lingual gyrus (BA 19) (-21, -72, 3; $Z= 3.43$), and left inferior temporal lobe (BA 37) (-48, -63, -3; $Z= 3.30$)

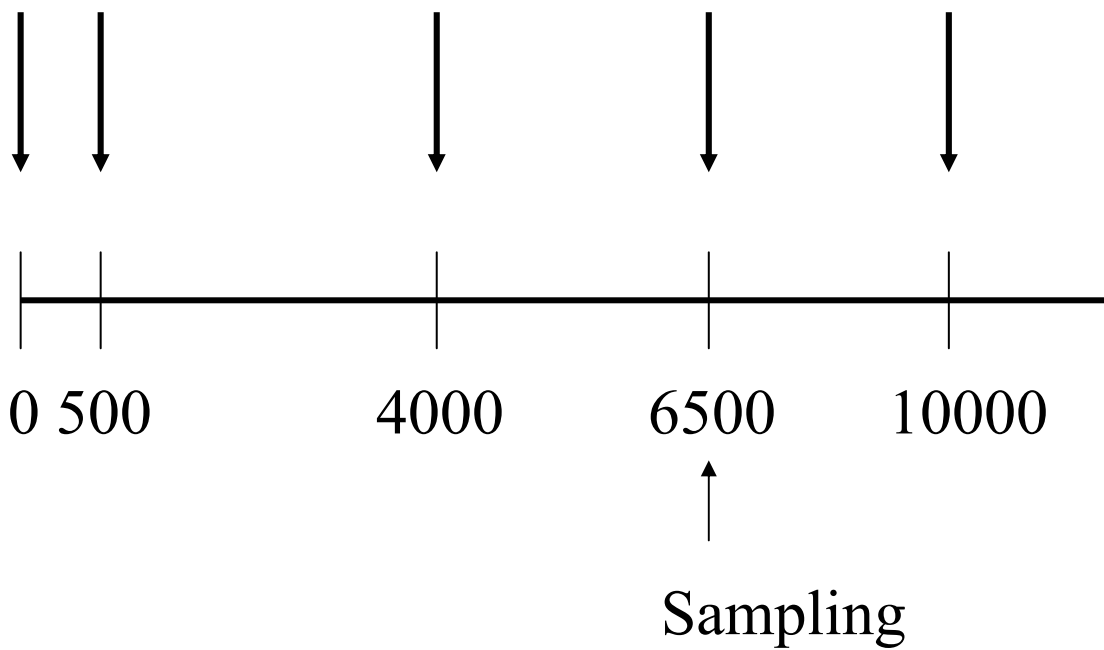
Second row (c) and (d): Modus Tollens vs. baseline activated left superior parietal lobule (BA 7) (-33, -51, 30; $Z= 3.63$ & -15, -51, 51; $Z= 3.23$ & -24, -75, 27; $Z= 3.48$), left cingulate gyrus (BA 32) (-6, 12, 45; $Z= 3.89$), left middle frontal gyrus (BA 6) (-36, 0, 45; $Z= 3.53$) and left inferior prefrontal cortex (BA 47) (-54, 15, -9; $Z= 3.64$).

Third row (e) and (f): Modus Tollens vs Modus Ponens activated left inferior parietal lobule (BA 40) (-39, -54, 39; $Z= 3.80$), left cingulate gyrus (BA 32) (-3, 18, 42; $Z= 3.70$), left dorsal lateral (BA 9) (-45, 15, 39; $Z= 3.56$), and inferior (BA 47) (-51, 24, -9; $Z= 3.15$) prefrontal cortex.

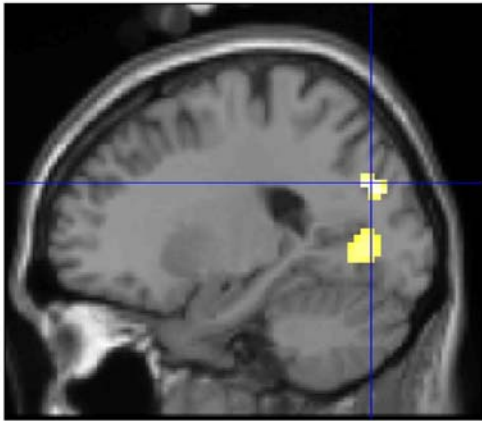
Event-related stimuli presentation presentation

* Premise 1 Premise 1 Premise 1 *

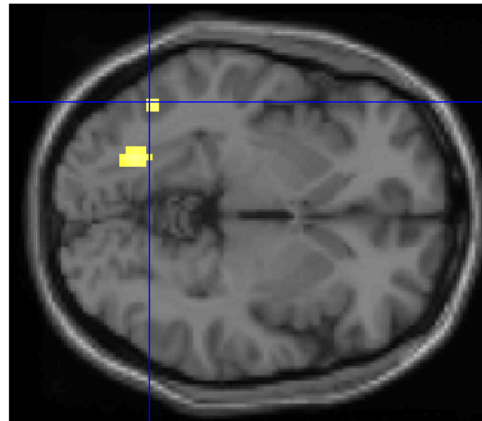
Premise 2 Premise 2 Conclusion



Modus Ponens - Baseline

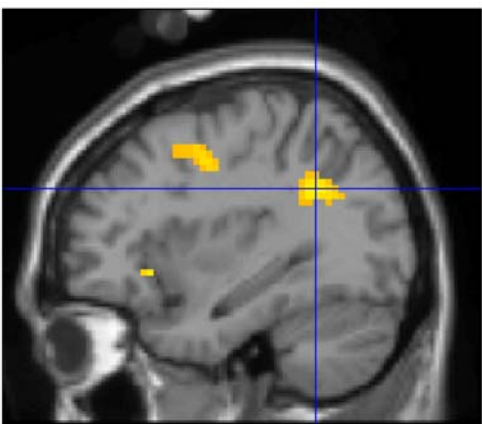


a

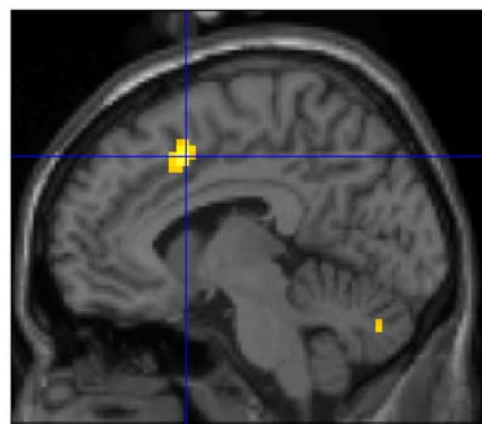


b

Modus Tollens - Baseline

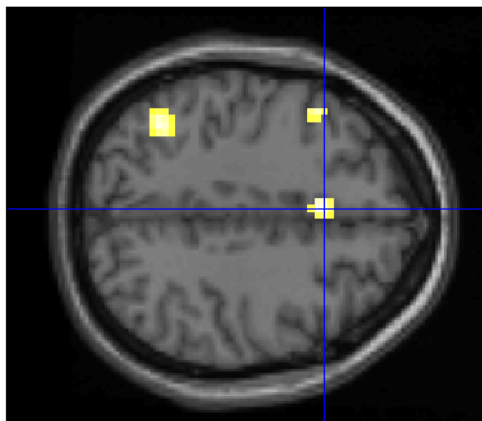


c

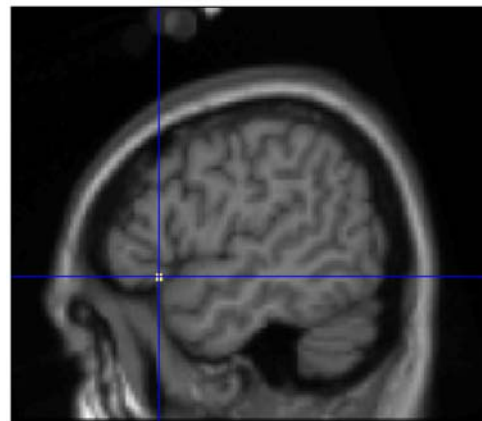


d

Modus Tollens - Modus Ponens



e



f