

Cognitive Neuroscience of Deductive Reasoning

Vinod Goel

Department Of Psychology

York University, Toronto, Ontario M3J 1P3 Canada

***** Draft Nov. 27, 2003*****

Address for Correspondence:

Dr. Vinod Goel
Dept. of Psychology
York University
Toronto, Ont.
Canada, M3J 1P3

Fax: 416-736-5814

Email: vgoel@yorku.ca

Goel, V. (in press). Cognitive Neuroscience of Deductive Reasoning. In *Cambridge Handbook of Thinking & Reasoning*, Eds. K. Holyoak & R. Morrison. Cambridge University Press.

1.0 Introduction

It is 4pm and I hear the school bus pull up to the house. Soon there is the taunting of a 13-year-old boy followed by the exaggerated screams of an 8-year-old girl. My kids are home from school. Exasperated, I say to my son, “If you want dinner tonight, you better stop tormenting your sister.” Given he doesn’t want to go to bed hungry, he needs to draw the correct logical inference. Sure enough, peace is eventually restored. Notice that he was not explicitly told to stop tormenting his sister. Yet we are not surprised by his actions. His behavior is not a mystery (assuming he wants his dinner). It is just an example of the reasoning brain at work.

Reasoning is the cognitive activity of drawing inferences from given information. All reasoning involves the claim that one or more propositions (the premises) provide some grounds for accepting another proposition (the conclusion). The above example involves a deductive inference (see Evans, this volume). A key feature of deduction is that conclusions are contained within the premises and are logically independent of the content of the propositions. Deductive arguments can be evaluated for validity, a relationship between premises and conclusion involving the claim that the premises provide absolute grounds for accepting the conclusion (i.e. if the premises are true, then the conclusion *must* be true).

2.0 Psychological Theories of Deductive Reasoning

Two theories of deductive reasoning (mental logic and mental models) dominate the cognitive literature. They differ with respect to the competence knowledge they draw upon, the mental representations they postulate, the mechanisms they invoke, and the neuroanatomical predictions they make. Mental logic theories (Braine, 1978; Henle, 1962; Rips, 1994) postulate

that Reasoners have an underlying competence knowledge of the *inferential role* of the closed-form, or logical terms, of the language (e.g. ‘all’, ‘some’, ‘none’, ‘and’, etc.). The internal representation of arguments preserve the structural properties of the propositional strings in which the premises are stated. A mechanism of inference is applied to these representations to draw conclusions from premises. Essentially, the claim is that deductive reasoning is a rule governed process defined over syntactic strings.

By contrast, mental model theory (Johnson-Laird, this volume; Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) postulates that Reasoners have an underlying competence knowledge of the *meaning* of the closed-form, or logical terms, of the language (e.g. ‘all’, ‘some’, ‘none’, ‘and’, etc.)¹ and use this knowledge to construct and search alternative scenarios.² The internal representation of arguments preserve the structural properties of the world (e.g. spatial relations) that the propositional string are about rather than the structural properties of the propositional strings themselves. The basic claim is that deductive reasoning is a process requiring spatial manipulation and search.

A third alternative is provided by dual mechanism theories. At a very crude level, dual mechanism theories make a distinction between formal, deliberate, rule-based processes and implicit, unschooled, automatic processes. However, dual mechanism theories come in various flavours that differ on the exact nature and properties of these two systems. Theories differentially emphasize explicit and implicit processes (Evans & Over, 1996), conscious and preconscious processes (Stanovich & West, 2000), formal and heuristic processes (Newell &

¹ Whether there is any substantive difference between “knowing the inferential role” and “knowing the meaning” of the closed-form terms, and thus the two theories is a moot point, debated in the literature.

² See Newell (1980b) for a discussion of the relationship between search and inference.

Simon, 1972), and associative and rule based processes (Goel, 1995; Sloman, 1996). The relationship among these proposals has yet to be clarified.

3.0 Relevance and Role of Neurophysiological Data

The reader will note that these are strictly cognitive theories, uninformed by knowledge of the brain. This is not an oversight. Until recently the central domains of human reasoning & problem solving have largely been cognitive & computational enterprises, with little input from neuroscience. In fact an argument advanced by cognitive scientists – based on the independence of computational processes and the mechanism in which they are realized (i.e. the brain) – has lead many to question the relevance of neuropsychological evidence for cognitive theories.

The “independence of computational level” argument is a general argument against the necessity of appealing to neurophysiology to capture the generalizations necessary to explain human mental life. The general idea is that liberation from neurophysiology is one of the great virtues of the cognitive/computational revolution. It gives us the best of both worlds. It allows us to use an intentional/semantic vocabulary in our psychological theories, and if this vocabulary meets certain (computational) constraints, we get a guarantee (via the Church-Turing hypothesis) that some mechanism will be able to instantiate the postulated process.³ Beyond this we don't have to worry about the physical. The psychological vocabulary will map onto the computational vocabulary, and it is after all, cognitive/computational structure, not physical structure, that captures the psychologically interesting generalizations.

³The Church-Turing hypothesis makes the conjecture that all computable functions belong to the class of functions computable by a Turing Machine. So if we constrain the class of functions called for by our psychological theories to the class of computable functions, then there will be some Turing Machine that can compute the function.

The argument can be articulated as follows:

- (P1) There are good reasons to believe that the laws of psychology need to be stated in intentional vocabulary (Fodor, 1975; Pylyshyn, 1984).
- (P2) Computation (sort of) gives us such a vocabulary (Cummins, 1989; Fodor, 1975; Goel, 1991, 1995; Newell, 1980a; Pylyshyn, 1984).
- (P3) Our theory construction is motivated by computational concepts/considerations and constrained by behavioural data.
- (P4) Computational processes are specified independently of physics and can be realized in any physical system.
- (C1) Therefore, there is no way, in principle, that neurological data can constrain our computational/cognitive theories.

A closer examination will reveal at least two flaws in the argument. First, premise P4 is not strictly true. Computational processes cannot be realized in any and every system (Giunti, 1997; Goel, 1991, 1992, 1995). If it was true, then computational explanations would be vacuous (Searle, 1990) and our problems much more serious. Now, it is true that computational processes can be realized in multiple systems, but that is far removed from universal realizability. The former gives computational theorizing much of its power; the latter drains computational explanations of much of their substantive content.

Second, the conclusion C1 depends on what “computational/cognitive theories” will be theories of. It is true that the organization of a computing mechanism (for example, whether a

Turing Machine has one heads or two) *is irrelevant* when we are interested in specifying what function is being computed and are concerned only with the mappings of inputs to outputs. This is a typical concern for mathematicians and logicians. If cognitive theories will only enumerate the functions being computed, then the argument would seem to hold. However, cognitive scientists (and often computer scientists) have little interest in computation under the aspect of functions. Our primary concern is with the procedures, which compute the functions (Marr, 1982). Real-time computation is a function of architectural considerations and resource availability and allocation. And it is real-time computation – the study of the behavioural consequences of different resource allocation and organization models – that must be of interest to cognitive science (Newell, 1980a; Newell & Simon, 1976), because it is only with respect to specific architectures that algorithms can be specified and compared (to the extent that they can be). If we are interested in the computational architecture of the mind – and we clearly are (Newell, 1990; Pylyshyn, 1984) – then the constraints provided by the mechanism which realizes the computational process become very relevant. Presumably neuroscience is where we will learn about the architectural constraints imposed on the human cognitive/ computational system. As such it can hardly be ignored.

But this whole line of argument and counter argument makes an unwarranted assumption. It assumes that the only contribution that neuroscience can make is in terms of specifying mechanisms. However, a glance through any neuroscience text (for e.g. (Kandel, Schwartz, & Jessell, 1995)) will show that neuroscience is still far from making substantive contributions to our understanding of the computational architecture of the central nervous system. This is many years in the future.

There are, however, two more immediate contributions – localization and dissociation – that cognitive neuroscience can make to our understanding of cognitive processes, including reasoning.

(1) Localization of brain functions: It is now generally accepted that Franz Joseph Gall (Gall & Spurzheim, 1810-1819) was largely right and Karl Lashley (1929) largely wrong about the organization of the brain. There is a degree of modularity in its overall organization. Over the years neuropsychologists and neuroscientists have accumulated some knowledge of this organization. For example, we know some brain regions are involved in processing language while other regions process visual spatial information. Finding selective involvement of these regions in complex cognitive tasks – like reasoning – can help us differentiate between competing cognitive theories that make different claims about linguistic and visuo-spatial processes in the complex task (as do mental logic and mental model theories of reasoning).

(2) Dissociation of brain functions: Brain lesions result in selective impairment of behavior. Such selective impairments are called dissociations. A *single* dissociation occurs when we find a case of a lesion in region x resulting in a deficit of function a but not function b . If we find another case, in which a lesion in region y results in a deficit in function b but not in function a , then we have a *double* dissociation. Recurrent patterns of dissociation provide an indication of causal joints in the cognitive system invisible in uninterrupted normal behavioural measures (Shallice, 1988). Lesion studies identify systems *necessary* for the cognitive processes under consideration. Neuroimaging studies identify cortical regions *sufficient* for various

cognitive processes.⁴ Both are sources of knowledge regarding dissociation of cognitive functions.

The more important of these two contributions is the identification of dissociations and warrants further discussion. Cognitive theories are functional theories. Functional theories are notoriously under constrained. That is, they are “black box” theories. We usually use them when we do not know the underlying causal structure. This devalues the currency of functional distinctions. But if we can show that our functional distinctions map onto causally individuated neurophysiological structures, then we can have much greater confidence in the functional individuation.

By way of an example, suppose that we individuate the following three functions on the basis of behavioral data: (f1) raise left arm, (f2) raise left foot, (f3) wiggle right ear. If these functions can be mapped onto three causally differentiated structures in a one-to-one fashion, we would be justified in claiming to have discovered three distinct functions. If, however, all three of our behaviorally individuated functions map onto one causally differentiated structure, in a many-to-one fashion, we would say that our functional individuation was too fine grained and collapse the distinctions until we achieved a one-to-one mapping. That is, raising the left arm does not constitute a distinct function from raising the left foot and wiggling the right ear, but the conjunction of the three do constitute a single function. If we encountered the reverse situation, where one behavioral function mapped onto several causally distinct structures, we would conclude that our individuation was too coarse-grained and refine it until we achieved a one-to-one mapping. One final possibility is a many-to-many mapping between our functional

⁴ These are of course logical claims about neuroimaging and lesion studies. As in all empirical work there are a number of complicating factors, including the relationship between statistical significance (or insignificance) and reality of an observed effect.

individuation and casually individuated physiological structures. Here we would have a total cross-classification and would have to assume that our functional individuations (f1, f2, f3) are simply wrong and start over again.⁵

The most famous example of a dissociation comes from the domain of language. In the 1860's Paul Broca described patients with lesions to the left posterior inferior frontal lobe who had difficulties in the production of speech but were quite capable of speech comprehension. This is a case of a single dissociation. In the 1870's Carl Wernicke described two patients (with lesions to the posterior regions of the superior temporal gyrus) who had difficulty in speech comprehension, but were quite fluent in speech production. Jointly the two observations indicate a double dissociation and tell us something important about the causal independence of language production and comprehension systems. If this characterization is accurate (and there are now some questions about its accuracy) it tells us that any cognitive theory of speech production and comprehension needs to postulate two distinct functions/mechanisms.

4.0 Neuroanatomical Predictions of Cognitive Theories of Reasoning

Given that the relevance of neuroanatomical data to cognitive theories has not been fully appreciated, it is not surprising that there are few explicit neuroanatomical predictions made by these theories. The one exception is mental model theory. Johnson-Laird (1994) has predicted that if mental model theory is correct, then reasoning must occur in the right hemisphere. The rationale here presumably is that mental model theory offers a spatial hypothesis and anecdotal

⁵ Again, I am making a logical point, independent of the usual complexities of mapping behaviour onto causal mechanisms.

neuropsychological evidence suggests that spatial processing occurs in the right hemisphere. A more accurate prediction for mental model theory would be that the neural structures for visuo-spatial processing contribute the basic representational building-blocks used for logical reasoning (i.e. the visuo-spatial system is necessary and sufficient for reasoning). I will use the latter prediction.

By contrast, mental logic theory is a linguistic hypothesis (Rips, 1994) and needs to predict that the neuroanatomical mechanisms of language (syntactic) processing underwrite human reasoning processes (i.e. that the language (syntactic) system is both necessary and sufficient for deductive reasoning). Both mental model and mental logic theories make explicit localization predictions (i.e. whether linguistic or visuo-spatial systems are involved) and implicit dissociation predictions, specifically that the one system is necessary and sufficient for reasoning.

Dual mechanism theory needs to predict the involvement of two different brain systems in human reasoning, depending on which system is engaged (i.e. the formal, deliberate, rule-based system or the implicit, unschooled, automatic system). But, it is difficult to make a prediction about localization without further specification of the nature of the two systems. Nonetheless, dual mechanism theory makes a substantive prediction about a dissociation in the neural mechanisms underlying the two different forms of reasoning.

5.0 Functional Anatomy of Reasoning

My colleagues and I have been carrying out a series of studies to investigate the neural basis of logical reasoning (Goel, Buchel, Frith, & Dolan, 2000; Goel & Dolan, 2000, 2001; Goel & Dolan, 2003; Goel, Gold, Kapur, & Houle, 1997, 1998; Goel, Grafman, Sadato, & Hallet,

1995). Our initial goal was to address the hypotheses made by the cognitive theories of reasoning, and in particular, differentiate between mental logic and mental model theories. We have made some progress along these lines (though with surprising results) and have also provided insights into the role of prefrontal cortex (PFC) in logical reasoning.

5.1 Basic paradigm and strategy

We have been presenting subjects with syllogisms – each consisting of two premises and a conclusion (e.g. All dogs are pets; All poodles are dogs \square All poodles are pets) – while they undergo PET (Positron Emission Tomography) or fMRI (functional magnetic resonance imaging) brain scans, and asking them to exhibit knowledge of what logically follows from the premises by confirming/denying the given conclusion. Our strategy has been to (largely) stay with one type of argument (syllogisms), manipulate content (holding the logically relevant information constant), and see how the brain reacts. The specific content manipulations are described in the studies below.

Neuroimaging studies typically require a rest or baseline condition against which to compare the active condition. For our baseline tasks we used trials in which the first two sentences were related but the third sentence was unrelated (e.g. All dogs are pets; All poodles are dogs \square All fish are scaly). Stimuli were presented one sentence at a time, with each sentence staying up until the end of the trial. Trials appeared randomly in an event-related design (Figure 1). The task in all trials was the same. Subjects were required to determine whether the conclusion followed logically from the premises (i.e. whether the argument was valid). In baseline trials, where the first two sentences were related, subjects would begin to construct a representation of the problem, but when the third, unrelated, sentence appeared they would

immediately disengage the task and respond 'no'. In reasoning condition trials, where the three sentences constituted an argument, subjects would continue with the reasoning component of the task after the presentation of the 3rd sentence. The difference between completing the reasoning task and disengaging after the presentation of the 3rd sentence isolates the reasoning components of interest. The data were modeled at the presentation of the 3rd sentence. The presentation of the first two sentences and subjects' motor responses were modeled out. This basic design was used in each of the imaging studies discussed below.

figure 1

We chose to use syllogisms (which test knowledge of quantification & negation) for technical reasons. Imaging studies require multiple presentations of stimuli to register a reliable neural signal. Syllogisms come in 64 different forms and thus allow for multiple trial presentations with minimal or no repetition of form.

We chose to manipulate content because logically, the content of an argument is irrelevant to the determination of its validity. For example, the argument

All men are mortal;

Socrates is a man;

□ Socrates is mortal

is valid by virtue of the fact it has the following form

All A are B;

C is A;

□ C is B

It remains valid irrespective of whether it is about Socrates or elephants. Validity is a function of the logical structure of the argument as opposed to the content of the sentences.

However, it is well known that the semantic contents of arguments affect people's validity judgments. In a classic study Wilkins (1928) showed that subjects performed better on syllogisms containing sentences with familiar semantic content (e.g. "All apples are red") than on syllogisms lacking semantic content (e.g. "All A are B"). When the semantic content of syllogisms was incongruent with beliefs (e.g. "All apples are poisonous"), performance suffered even more. These results have been explored and extended in the more recent literature (Cherubini, Garnham, Oakhill, & Morley, 1998; Evans, Barston, & Pollard, 1983; Oakhill & Garnham, 1993; Oakhill, Johnson-Laird, & Garnham, 1989). The effect is very robust and has challenged cognitive theories of reasoning.

We discuss our key findings below. They include: (i) a dissociation between a frontal-temporal system and a parietal system as a function of the familiarity of the content of the reasoning material; (ii) asymmetrical involvement of right & left PFC, with the left PFC being necessary and sometimes sufficient, and right PFC being sometimes necessary (in unfamiliar, incoherent, conflicting situations), but not sufficient; and (iii) clarifying roles of right PFC and ventral medial prefrontal cortex (VMPFC) in belief-logic conflict resolution.

5.2 Basic Findings

5.2.1 Dissociable neural networks

In Goel et al. (2000) we scanned 11 right-handed normal subjects using event-related fMRI, to measure task related neural activity, while they engaged in syllogistic reasoning. The study was designed to manipulate the presence of content in logical reasoning. Half of the arguments contained content sentences such as:

All dogs are pets;

All poodles are dogs □

All poodles are pets

while the other half contained “no content” versions of these sentences such as:

All P are B;

All C are P

□ All C are B

The logically relevant information in both conditions was identical. Half of the arguments were valid, and the other half were invalid. For our baseline tasks we used trials in which the first two sentences were related but the third sentence was unrelated, as in the following examples:

All dogs are pets;

All poodles are dogs □

All fish are scaly

and

All P are B;

All C are P

□ All N are D

If mental model theory is correct all reasoning trials should activate a visuo-spatial system (perhaps parietal cortex). If mental logic theory is correct we would expect activation of the language system (left frontal and temporal lobe regions). Dual mechanism theory predicts engagement of two distinct (but unspecified) neural systems, depending on whether subjects respond in a “schooled”, formal manner, or an intuitive, implicit manner. What we actually found was that the main effect of reasoning implicated large areas of the brain (Figure 2) including regions predicted by both mental model and mental logic theories.

figure 2 approx here

However, closer examination revealed this to be a composite activation consisting of two dissociable neural systems. The content reasoning trials compared to no-content reasoning trials revealed activation in left middle / superior temporal lobe (BA 21/22), left temporal pole (BA 21/38), and left inferior frontal lobe (BA 47) (Figure 3a). This is essentially a language and

memory system. A similar network has been activated in previous studies of deductive reasoning using contentful sentences (Goel et al., 1997, 1998).

figure 3 approx here

The reverse comparison of no-content reasoning trials vs. content reasoning trials resulted in activation of bilateral occipital (BA 19), bilateral superior and inferior parietal lobes (BA 7), and bilateral dorsal (BA 6) and inferior (BA 44) frontal lobes (Figure 3b). This pattern of activation is known to be involved in the internal representation and manipulation of spatial information (Jonides, Smith, Koeppe, Awh, & Minoshima, 1993; Kosslyn, Koenig, Cave, Tang, & Gabrieli, 1989) and is very similar to that reported for transitive inference involving geometrical shapes (Acuna, Eliassen, Donoghue, & Sanes, 2002), and certain types of mathematical reasoning involving approximation of numerical quantities (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999).

It is possible to argue that the patterns of activation revealed by the direct comparison of content and no-content conditions are just a function of the presence or absence of content words, rather than being indicative of different reasoning mechanisms. To exclude this possibility we examined the content (content, no content) by task (reasoning, baseline) interaction. The modulation of reasoning, by the addition of content [(content reasoning – content baseline) - (no-content reasoning – no-content baseline)] revealed activation in Wernicke's area. The reverse interaction, which examined the effect of the absence of semantic

content, [(no-content reasoning – no-content baseline) - (content reasoning - content baseline)] activated left parietal cortex. This interaction analysis eliminates the above possibility and confirms the involvement of these two systems in the reasoning process.

Contrary to both, mental logic theories that predict the language (syntactic) system is necessary and sufficient for deductive reasoning, and mental model theories that predict the visuo-spatial system is necessary and sufficient for logical reasoning, Goel et al. (2000) found evidence for the engagement of both systems. The presence of semantic content engages the language and long-term memory systems in the reasoning process. The absence of semantic content engages the visuo-spatial system in the identical reasoning task. Before discussing the implications of these results for cognitive theories, let us consider some additional issues and data.

The Goel et al. (2000) study raises several interesting questions, one of which has to do with the involvement of a parietal visual-spatial system in the no-content or abstract syllogism condition. A second question has to do with the exact property of the stimuli that leads to the modulation of neural activity between frontal-temporal and parietal systems. Pursuing the first question lead to a clarification of the second question.

The first question is whether argument forms involving three-term spatial relations such as:

The apples are in the barrel;

The barrel is in the barn

□ The apples are in the barn

and

A are in B;

B is in C

□ A are in C

are sufficient to engage the parietal system irrespective of the presence of content? One rationale for thinking this might be the case is subjects' reported phenomenological experience of using a visuo-spatial strategy during these tasks. Secondly, neuroimaging studies have also shown the involvement of the parietal system in the encoding of relational spatial information (Laeng, 1994; Mellet et al., 1996). To address this question we carried out another fMRI study, this time using three-term relational arguments like the above (Goel & Dolan, 2001).

Goel and Dolan (2001) found that reasoning about abstract and concrete three-term relations, as in the above examples, recruited a bilateral parietal-occipital system, with greater involvement of parietal and occipital lobes in the abstract condition compared to the concrete condition. There was an absence of the two dissociable networks for concrete and abstract reasoning reported in the first study. In particular, the temporal lobe (BA 21/22) activation evident in concrete syllogistic reasoning in the first study, was conspicuously absent in this study. The lack of temporal lobe (BA 21/22) activation in Goel and Dolan (2001) might be explained by analysing the nature of the content used in the two studies. The concrete sentences in Goel et al. (2000), were of the form "All apples are poisonous" whereas the concrete sentences in Goel & Dolan (2001) were of the form "John is to the right of Mary". The former sentence types predicate known properties to known objects. We have beliefs about whether "all apples

are poisonous.” By contrast, the latter sentence types do not allow for such beliefs.⁶ This leaves open the interesting possibility that involvement of BA 21/22 in reasoning may be specific to content processing involving belief networks rather than just concrete contents.

This hypothesis was tested in Goel & Dolan (2003) where subjects were presented with arguments such as:

No reptiles are hairy;

Some elephants are hairy

□ No elephants are reptiles

containing sentences that subjects could be expected to have beliefs about, and belief-neutral arguments such as:

No codes are highly complex;

Some quipu are highly complex

□ No quipu are codes

containing sentences that subjects may not have beliefs about (because they may not know the meaning of one or more key terms). The referential terms in the two conditions were counterbalanced for abstract and concrete categories.

⁶ It is possible to generate relational sentences one can have beliefs about, for e.g. “London is north of Rome” or “Granite is harder than diamonds”.

The results of this study replicated and clarified the results of Goel et al. (2000). Modulation of the reasoning task by absence of belief [(belief-neutral reasoning – belief-neutral baseline) - (belief-laden reasoning – belief-laden baseline)] revealed activation in the left superior parietal lobe (BA 7) unique to the belief-neutral condition. The reverse modulation [(belief-laden reasoning – belief-laden baseline) - (belief-neutral reasoning – belief-neutral baseline)] revealed activation of anterior left middle temporal gyrus (BA 21) activation unique to the belief-bias condition. These results confirm that a critical (sufficient) factor in the modulation of activity between these two neural systems is the presence of familiar or belief-laden content in the reasoning processes.

5.2.2 Generalization of Dissociation to Transitive Reasoning

We have demonstrated dual pathways for reasoning about categorical syllogisms. The question arises whether the results generalize to other forms of logical reasoning, particularly three-term spatial relations, where one might think the visuo-spatial system may be sufficient. To answer this question Goel et al. (in press) studied 14 volunteers using event-related fMRI, as they reasoned about landmarks in familiar and unfamiliar environments.

Half the arguments contained sentences such as

“Paris is south of London;

London is south of Edinburgh;

□ Paris is south of Edinburgh”

describing environments that subjects would be familiar with (as confirmed by a post-scan questionnaire), while the other half contained sentences such as

“The AI lab is south of the Roth Centre;

Roth Centre is south of Cedar Hall;

□ AI lab is south of Cedar Hall”

that subjects could not be familiar with because they describe a fictional unknown environment.

Our main finding was a task (reasoning and baseline) by spatial content (familiar and unfamiliar) interaction. Modulation of reasoning towards unfamiliar landmarks resulted in bilateral activation of superior & inferior parietal lobule (BA 7, 40), dorsal superior frontal cortex (BA 6), and right superior and middle frontal gyri (BA 8), regions widely implicated in visuo-spatial processing. By contrast, modulation of the reasoning task towards familiar landmarks, engaged right inferior/orbital frontal gyrus (BA 11/47), bilateral occipital (BA 18, 19) and temporal lobes. The temporal lobe activation included right inferior temporal gyrus (BA 37), and posterior hippocampus and parahippocampal gyrus, regions implicated in spatial memory and navigation tasks. These results provide support for the generalization of our dual mechanism account to transitive reasoning and highlight the importance of the hippocampal system in *reasoning* about landmarks in familiar spatial environments.

5.2.3 Evidence for Dissociation from Patient Data

If we are correct that reasoning involving familiar situations engages a frontal-temporal lobe system, while formally identical reasoning tasks involving unfamiliar situations recruit a

frontal-parietal visuo-spatial network – with greater frontal lobe involvement in the former than the latter – then frontal lobe lesion patients should be more impaired on reasoning about familiar situations than on unfamiliar situations. To test this hypothesis Goel et al. (in review) administered the Wason 4-Card Selection Task (Wason, 1966) to 19 frontal lobe patients and 19 age- and education-matched normal controls.

Wason 4-Card Selection Task (WST) (Wason, 1966) is the most widely used task to explore the role of content in reasoning. In this task subjects are shown four cards. They can see what is printed on one side of each card, but not the other side. They are given a rule of the form if p then q (e.g. "If a card has a vowel on one side, it must have an even number on the other side.") and asked which cards they must turn over in order to verify the rule. The visible values on the cards correspond to the p , $not-p$, q , and $not-q$ cases of the rule. According to standard propositional logic, the correct choices are p (to verify q is on the other side) and $not-q$ (to verify p is not on the other side). Given an arbitrary rule like the above, typically less than 25% of normal subjects will turn over both the p and the $not-q$ cards. However, the introduction of familiar, meaningful content in a rule (e.g. "If anyone is drinking beer, then that person must be over 21 years old.") greatly facilitates performance (Cheng & Holyoak, 1985; Cosmides, 1989; Cox & Griggs, 1982; Gigerenzer & Hug, 1992; Griggs & Cox, 1982; Wason & Shapiro, 1971).

Specifically we manipulated the social knowledge involved in the task in the form of "permission schemas (Cheng & Holyoak, 1985). Subjects performed the task with an arbitrary rule condition, ("If a card has an "A" on one side, then it must have a "4" on the other side."), an abstract permission condition, ("If one is to take action "A", then one must first satisfy precondition "P".), and a concrete permission condition, ("If a person is to drink alcohol, he or she must be at least 21.").

The principal findings were that, in the “purely logical” (arbitrary rule) condition, frontal lobe patients performed just as well (or just as poorly) as normal controls. However, patient performance did not improve with the introduction of social knowledge in the form of abstract or concrete permission schemas, as did normal control performance. Furthermore, there was no significant correlation between volume loss, IQ scores, memory scores, or years of education and performance in the abstract or concrete permission schema conditions. Thus the failure of patients to benefit from social knowledge cannot be explained in terms of volume loss, IQ scores, memory scores, or years of education.

Consistent with the neuroimaging data, our interpretation is that the arbitrary rule condition of the WST involves greater activation of the parietal lobe system, while the permission schema trials result in greater engagement of a frontal-temporal lobe system. The normal controls have both mechanisms intact and can take advantage of social knowledge cues to facilitate the reasoning process. The patients’ parietal system is intact, hence their performance on the arbitrary rule trial is the same as the normal controls. Their frontal lobe system is disrupted, preventing them from taking advantage of social knowledge cues in the permission schema trials.⁷

5.2.4 Hemispheric Asymmetry

Our imaging studies have also revealed an asymmetry in frontal lobe involvement in logical reasoning. Reasoning about belief-laden material (e.g. All dogs are pets; All poodles are dogs; All poodles are pets) activates left prefrontal cortex (Figure 4a), while reasoning about

⁷ See also Chapter 17 (this volume), for further discussion of disrupted thinking in patient populations.

belief-neutral material (e.g. All A are B; All C are A; All C are B) activates bilateral prefrontal cortex (Figure 4b) (Goel et al., 2000; Goel & Dolan, 2003). This asymmetry shows up consistently in patient data.

Caramazza et al. (1976) administered two-term problems such as the following: “Mike is taller than George. Who is taller?” to brain-damaged patients. 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?). Read (1981) tested temporal lobectomy patients on three-term relational problems with semantic content (e.g. "George is taller than Mary. Mary is taller than Carol. Who is tallest?"). Subjects were told that using a mental imagery strategy would help them to solve these problems. He reported that left temporal lobectomy patient performance was more impaired than right temporal lobectomy patient performance. In a more recent study using matched verbal and spatial reasoning tasks Langdon and Warrington (2000) found that only left hemisphere patients failed the verbal section, 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.

In the WCT patient study discussed above (Goel et al., in review), not only was it the case that frontal lobe patients failed to benefit from the introduction of familiar content into the task, the result was driven by the poor performance of left hemisphere patients. There was no difference in performance between right hemisphere patients and normal controls, only between left hemisphere patients and controls. These data show that the LH is necessary and often sufficient for reasoning, while the RH is sometimes necessary, but not sufficient. (This is of course contrary to the Johnson-Laird (1994) prediction for mental model theory, but as noted

above, we chose to modify in this prediction to make it consistent with neuropsychological data.)

5.2.5 Dealing with Belief-Logic Conflicts

Although from a strictly logical point of view deduction is a closed system, we have already mentioned above that beliefs about the conclusion of an argument influence people's validity judgments (Wilkins, 1928). When arguments have a familiar content it will be the case that the truth value (or believability) of a given conclusion will be either consistent or inconsistent with the logical judgment. Subjects perform better on syllogistic reasoning tasks when the truth value of a conclusion (true or false) coincides with the logical relationship between premises and conclusion (valid or invalid) (Evans et al., 1983). Such trials are facilitatory to the logical task and consist of valid arguments with believable conclusions (e.g. Some children are not Canadians; All children are people; □ Some people are not Canadians) and invalid arguments with unbelievable conclusions (e.g. Some violinists are not mutes; No opera singers are violinists; □ Some opera singers are mutes). Where the logical conclusion is inconsistent with subjects' beliefs about the world, the beliefs are inhibitory to the logical task, and decrease accuracy (Evans et al., 1983). Inhibitory belief trials consist of valid arguments with unbelievable conclusions (e.g. No harmful substances are natural; All poisons are natural; □ No poisons are harmful) and invalid arguments with believable conclusions (e.g. All calculators are machines; All computers are calculators; □ Some machines are not computers). Performance on arguments that are belief-neutral usually falls between these two extremes (Evans, Handley, & Harper, 2001).

Goel et al. (2000) noted that when logical arguments result in a belief-logic conflict, the nature of the reasoning process is changed by the recruitment of the right lateral prefrontal cortex

(Figure 4c). Goel and Dolan (2003) further noted that within the inhibitory belief trials, a comparison of correct items with incorrect items (correct inhibitory belief trials - incorrect inhibitory belief trials) revealed activation of right inferior prefrontal cortex (Figure 5a). The reverse comparison of incorrect response trials with the correct response trials (incorrect inhibitory belief trials - correct inhibitory belief trials) revealed activation of VMPFC (Figure 5b).

Within the inhibitory belief trials the prepotent response is associated with belief-bias. Correct responses (in inhibitory trials) indicate that subjects detected the conflict between their beliefs and the logical inference, inhibited the prepotent response associated with the belief-bias, and engaged the reasoning mechanism. Incorrect responses in such trials indicate that subjects failed to detect the conflict between their beliefs and the logical inference and/or inhibit the prepotent response associated with the belief-bias. Their response is biased by their beliefs. The involvement of right prefrontal cortex in correct response trials is critical in detecting and/or resolving the conflict between belief and logic. Such a role of the right lateral prefrontal cortex was also noted in (Goel et al., 2000), and in a study of maintenance of an intention in the face of conflict between action and sensory feedback (Fink et al., 1999). A similar phenomenon has been noted in the Caramazza et al. (1976) study mentioned above where right hemisphere patients were only impaired when there was an incongruity in the form of the question and the premises. By contrast, the activation of VMPFC in incorrect trials highlights its role in non-logical, belief-based responses.

6.0 Consequences for Cognitive Theories of Reasoning

We now briefly address the question of how these data map onto the cognitive theories of reasoning, with which we began our discussion. This is a complex question because the data do not fit neatly with any of the three theories. First and foremost, we show a dissociation in mechanisms involved in belief-neutral and belief-laden reasoning. The two systems we have identified are roughly the language system and the visuo-spatial system, which is what mental logic theory and mental model theory respectively predict. However, neither theory anticipates this dissociation. Each theory predicts that the system it postulates is necessary and sufficient for reasoning. This implies that the neuroanatomical data cross-classifies these cognitive theories. A further complication is that mental logic theory implicates the syntactic component of language in logical reasoning. Our studies activate both the syntactic and semantic systems and components of long-term memory.

Our results do seem compatible with some form of dual mechanism theory, which explicitly predicts a dissociation. However, as noted above, this theory comes in various flavours and some advocates may not be keen to accept our conclusions. The distinction that our results point to is between reasoning with familiar, conceptually coherent material vs. unfamiliar, nonconceptual or incoherent material. The former engages a left frontal-temporal system (language and long-term memory) while the latter engages a bilateral parietal (visuo-spatial) system. Given the primacy of belief-bias over effortful thinking (Sloman, 1996) we believe that the frontal-temporal system is more “basic,” and effortlessly engaged. It has temporal priority. By contrast, the parietal system is effortfully engaged when the frontal-temporal route is blocked due to a lack of familiar content, or when a conflict is detected between the logical response and belief-bias. This is very consistent with the dual mechanism account developed by Newell & Simon (1972) for the domain of problem solving. On this formulation our frontal-temporal

system corresponds to the “heuristic” system while the parietal system corresponds to the “universal” system. Reasoning about familiar situations automatically utilizes situation-specific heuristics, which are based on background knowledge and experience. Where no such heuristics are available (as in reasoning about unfamiliar situations), universal (formal) methods must be used to solve the problem. In the case of syllogistic reasoning this may well involve a visuo-spatial system.

Our results go beyond addressing cognitive theories of reasoning and provide new insight into the role of the prefrontal cortex in human reasoning. In particular, the involvement of the prefrontal cortex in logical reasoning is selective and asymmetric. Its engagement is greater in reasoning about familiar, content-rich situations than unfamiliar, content-sparse situations. The left prefrontal cortex is necessary and often sufficient for reasoning. The right prefrontal cortex is sometimes necessary, but not sufficient for reasoning. It is engaged in the absence of conceptual content and in the face of conflicting or conceptually incoherent content (as in the belief-logic conflicts discussed above). Finally, the VMPFC is engaged by non-logical, belief-biased responses.

7.0 Current Issues & Future Directions

While some progress has been made over the past 8 years, the cognitive neuroscience of reasoning is in its infancy. The next decade should be an exciting time of rapid development. There are a number of issues that we see as particularly compelling for further investigation. The first is generalizability of the results. Will the results regarding syllogisms, which are quite difficult, generalize to basic low-level inferences such as modus ponens and modus tollens?

Second, all the imaging studies to date have utilized a paradigm involving the recognition of a given conclusion as valid or invalid. It remains to be seen whether the generation of a conclusion would involve the same mechanisms. Third, given the involvement of visuo-spatial processing systems in much of reasoning, and the postulated differences between males and females in processing spatial information (Jones, Braithwaite, & Healy, 2003), one might expect neural-level differences in reasoning between the sexes. Fourth, the issue of task difficulty has not been explored. As reasoning trials become more difficult, are additional neural resources recruited, or are the same structures activated more intensely? Fifth, what is the effect of learning/training on the neural mechanisms underlying reasoning? Sixth, most imaging studies to date have focused on deduction. While deduction is interesting, much of human reasoning actually involves induction. The relationship between the two at the neural level is still an open question. Finally, reasoning does not occur in a vacuum. Returning to the example of my children, with which I began, if I say to my son “If you want dinner tonight, you better stop tormenting your sister” in a calm, unconcerned voice, it usually has an effect. However, if I state the same proposition in an angry, threatening voice, the impact is much more complete and immediate. Given that the logic of the inference is identical in the two cases, the emotions introduced into the situation through the modulation of my voice, are contributing to the impact of the proposition. In fact, emotions can be introduced into the reasoning process in at least one of three ways, (i) in the content/substance of the reasoning task; (ii) in the presentation of the content of the reasoning task (as in voice intonation); and (iii) in the preexisting mood of the reasoning agent. We are currently channeling much of our research efforts to understanding the neural basis of the interaction between emotions and rational thought.

Acknowledgments

VG is supported by a McDonnell-Pew Program in Cognitive Neuroscience Award, NSERC & CIHR grants, and a Premier's Research Excellence Award.

References

- Acuna, B. D., Eliassen, J. C., Donoghue, J. P., & Sanes, J. N. (2002). Frontal and Parietal Lobe Activation during Transitive Inference in Humans. *Cereb Cortex*, *12*(12), 1312-1321.
- Braine, M. D. S. (1978). On the Relation Between the Natural Logic of Reasoning and Standard Logic. *Psychological Review*, *85*(1), 1-21.
- Caramazza, A., Gordon, J., Zurif, E. B., & DeLuca, D. (1976). Right-hemispheric damage and verbal problem solving behavior. *Brain Lang*, *3*(1), 41-46.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognit Psychol*, *17*(4), 391-416.
- Cherubini, P., Garnham, A., Oakhill, J., & Morley, E. (1998). Can any ostrich fly?: some new data on belief bias in syllogistic reasoning. *Cognition*, *69*(2), 179-218.
- Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task [see comments]. *Cognition*, *31*(3), 187-276.
- Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. *Mem Cognit*, *10*(5), 496-502.
- Cummins, R. (1989). *Meaning and Mental Representation*. Cambridge, Massachusetts: The MIT Press.
- Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science*, *284*(5416), 970-974.
- Evans, J. S., Handley, S. J., & Harper, C. N. (2001). Necessity, possibility and belief: a study of syllogistic reasoning. *Q J Exp Psychol A*, *54*(3), 935-958.
- Evans, J. S. B. T., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, *11*, 295-306.
- Evans, J. S. B. T., & Over, D. E. (1996). *Rationality and Reasoning*. NY: Psychology Press.
- Fink, G. R., Marshall, J. C., Halligan, P. W., Frith, C. D., Driver, J., Frackowiak, R. S., & Dolan, R. J. (1999). The neural consequences of conflict between intention and the senses. *Brain*, *122*(Pt 3), 497-512.
- Fodor, J. A. (1975). *The Language of Thought*. Cambridge, Massachusetts: Harvard University Press.
- Gall, F. J., & Spurzheim, J. (1810-1819). *Anatomie et Physiologie du Systeme Nerveux en General, et der Cerveauen Particulier*. Paris: F. Schoell.
- Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: social contracts, cheating, and perspective change. *Cognition*, *43*(2), 127-171.
- Giunti, M. (1997). *Computation, dynamics, and cognition*. New York: Oxford University Press.

- Goel, V. (1991). Notationality and the Information Processing Mind. *Minds and Machines*, 1(2), 129-165.
- Goel, V. (1992). *Are Computational Explanations Vacuous?* Paper presented at the Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, Bloomington, Indiana.
- Goel, V. (1995). *Sketches of Thought*. Cambridge, MA: MIT Press.
- Goel, V., Buchel, C., Frith, C., & Dolan, R. J. (2000). Dissociation of Mechanisms Underlying Syllogistic Reasoning. *NeuroImage*, 12(5), 504-514.
- Goel, V., & Dolan, R. J. (2000). Anatomical Segregation of Component Processes in an Inductive Inference Task. *Journal of Cognitive Neuroscience*, 12(1), 1-10.
- Goel, V., & Dolan, R. J. (2001). Functional Neuroanatomy of Three-Term Relational Reasoning. *Neuropsychologia*, 39(9), 901-909.
- Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition*, 87(1), B11-22.
- Goel, V., Gold, B., Kapur, S., & Houle, S. (1997). The Seats of Reason: A Localization Study of Deductive & Inductive Reasoning using PET (O15) Blood Flow Technique. *NeuroReport*, 8(5), 1305-1310.
- Goel, V., Gold, B., Kapur, S., & Houle, S. (1998). Neuroanatomical Correlates of Human Reasoning. *Journal of Cognitive Neuroscience*, 10(3), 293-302.
- Goel, V., Grafman, J., Sadato, N., & Hallett, M. (1995). Modelling Other Minds. *NeuroReport*, 6(13), 1741-1746.
- Goel, V., Makale, M., & Grafman, J. (in press). The Hippocampus Formation Mediates Logical Reasoning about Familiar Spatial Landmarks. *Journal of Cognitive Neuroscience*.
- Goel, V., Shuren, J., Sheesley, L., & Grafman, J. (in review). Asymmetrical Involvement of Frontal Lobes in Social Reasoning.
- Griggs, R. A., & Cox, J. R. (1982). The Elusive Thematic-Materials Effect in Wason's Selection Task. *British Journal of Psychology*, 73, 407-420.
- Henle, M. (1962). On the Relation Between Logic and Thinking. *Psychological Review*, 69(4), 366-378.
- Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, Mass.: Harvard University Press.
- Johnson-Laird, P. N. (1994). Mental Models, Deductive Reasoning, and the Brain. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 999-1008). Cambridge, MA: MIT Press.
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Jones, C. M., Braithwaite, V. A., & Healy, S. D. (2003). The evolution of sex differences in spatial ability. *Behav Neurosci*, 117(3), 403-411.
- Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., & Minoshima, S. (1993). Spatial Working Memory in Humans as Revealed by PET. *Nature*, 363, 623-625.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (Eds.). (1995). *Essentials of Neural Science and Behaviour*. Norwalk, Connecticut: Appleton & Lange.
- Kosslyn, S. M., Koenig, O., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for Two Types of Spatial Representations: Hemispheric Specialization for Categorical and Coordinate Relations. *Journal of Experimental Psychology: Human Perception and Performance*, 15(4), 723-735.

- Laeng, B. (1994). Lateralization of Categorical & Coordinate Spatial Functions: A Study of Unilateral Stroke Patients. *Journal of Cognitive Neuroscience*, 6(3), 189-203.
- Langdon, D., & Warrington, E. K. (2000). The role of the left hemisphere in verbal and spatial reasoning tasks. *Cortex*, 36(5), 691-702.
- Lashley, K. S. (1929). *Brain mechanisms and intelligence : a quantitative study of injuries to the brain*. Chicago: University of Chicago Press.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman.
- Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., & Mazoyer, D. (1996). Functional Anatomy of Spatial Mental Imagery Generated from Verbal Instructions. *The Journal of Neuroscience*, 16(20), 6504-6512.
- Newell, A. (1980a). Physical Symbol Systems. *Cognitive Science*, 4, 135-183.
- Newell, A. (1980b). Reasoning, Problem Solving, and Decision Processes: The Problem Space as a Fundamental Category. In R. S. Nickerson (Ed.), *Attention and Performance VIII*. Hillsdale, N.J.: Lawrence Erlbaum.
- Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.
- Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall.
- Newell, A., & Simon, H. A. (1976). Computer Science as Empirical Inquiry: Symbols and Search. *Communications of the ACM*, 19(March), 113-126.
- Oakhill, J., & Garnham, A. (1993). On theories of belief bias in syllogistic reasoning. *Cognition*, 46(1), 87-92; discussion 93-87.
- Oakhill, J., Johnson-Laird, P. N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition*, 31(2), 117-140.
- Pylyshyn, Z. W. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, Massachusetts: MIT Press.
- Read, D. E. (1981). Solving Deductive-Reasoning Problems After Unilateral Temporal Lobectomy. *Brain and Language*, 12, 116-127.
- Rips, L. J. (1994). *The Psychology of Proof: Deductive Reasoning in Human Thinking*. Cambridge, MA: MIT Press.
- Searle, J. R. (1990). *Is the Brain a Digital Computer?* Paper presented at the Sixty-fourth Annual Pacific Division Meeting for the American Philosophical Association, Los Angeles, CA, March 30, 1990.
- Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.
- Sloman, S. A. (1996). The Empirical Case For Two Systems Of Reasoning. *Psychological Bulletin*, 119(1), 3-22.
- Stanovich, K. E., & West, R. F. (2000). Individual Differences in Reasoning: Implications for the Rationality Debate. *Behavioral & Brain Sciences*, 22, 645-665.
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New Horizons in Psychology* (pp. 135-151). Harmondsworth: Penguin.
- Wason, P. C., & Shapiro, D. A. (1971). Natural and Contrived Experience in a Reasoning Problem. *Quarterly Journal of Experimental Psychology*, 23, 63-71.
- Wilkins, M. C. (1928). The effect of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology*, 16(102).

Figures

Figure 1. Stimuli presentation: Stimuli from all conditions were presented randomly in an event-related design. An “*” indicated the start of a trial at 0 seconds. The sentences appeared on the screen one at a time with the first sentence appearing at 500 ms, the second at 3500 ms, and the last sentence at 6500 ms. The length of trials varied from 10.25-14.35 seconds, leaving subjects 3.75 to 7.85 seconds to respond.

Figure 2. Main effect of reasoning [(content reasoning + no content reasoning) – (content preparation + no content preparation)] revealed activation of bilateral cerebellum (R > L), bilateral fusiform gyrus, left superior parietal lobe, left middle temporal gyrus, bilateral inferior frontal gyrus, bilateral basal ganglia nuclei (centered around the accumbens, caudate nucleus, and putamen), and brain stem.

Figure 3. (a) The content reasoning - no-content reasoning comparison revealed activation of the left middle / superior temporal lobe (BA 21/22), the left inferior frontal lobe (BA 47), and bilateral (BA 17) and lingual gyri (BA 18). (b) The no-content reasoning - content reasoning comparison revealed activation of (a) bilateral occipital (BA 18, 19) and (c) bilateral superior and inferior parietal lobes (BA 7, 40), bilateral precentral gyrus (BA 6), and bilateral middle frontal gyrus (BA 6).

Figure 4. (a) Reasoning involving familiar conceptual content activates left inferior prefrontal cortex. (b) Reasoning involving unfamiliar content activates bilateral prefrontal cortex. (c) Right prefrontal cortex mediates belief-logic conflict detection and/or resolution.

Figure 5. (a) Correct inhibitory trials activate right prefrontal cortex. (b) Incorrect inhibitory trials activate VMPFC cortex.

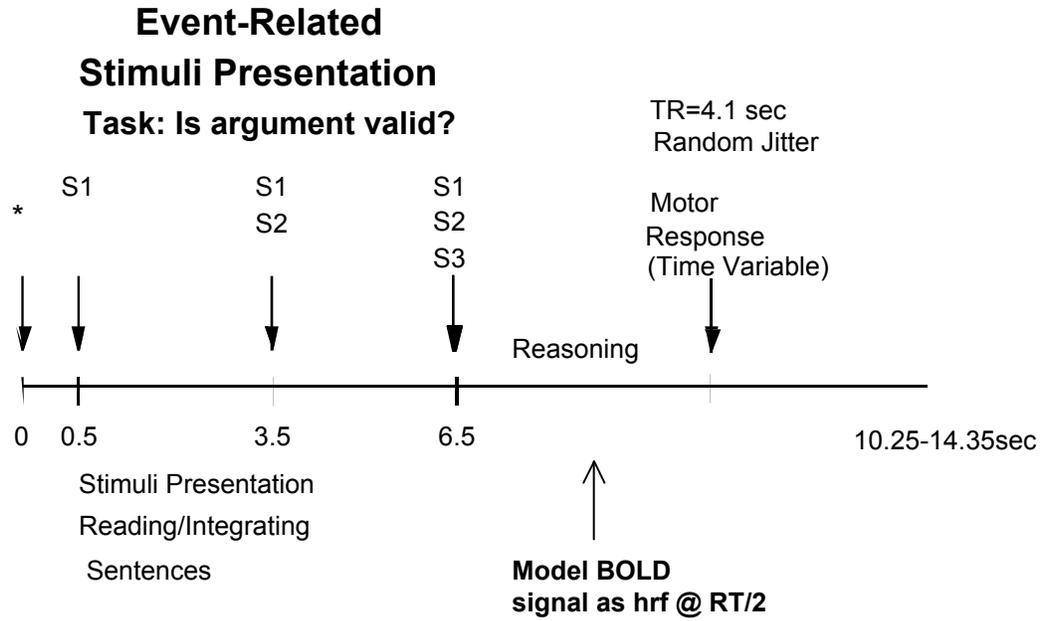


Figure 1

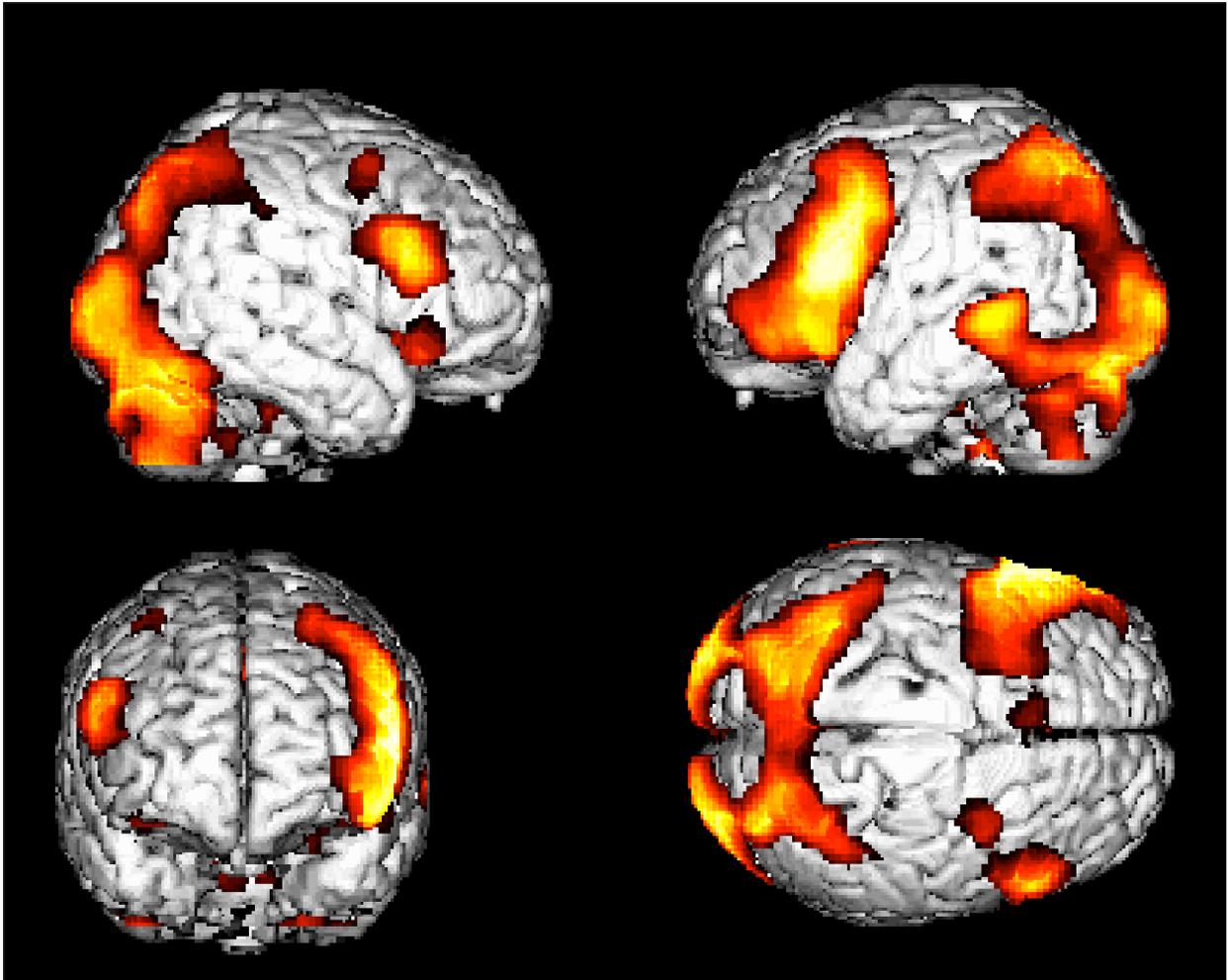
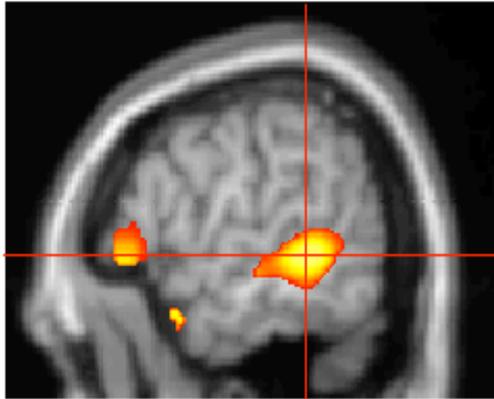
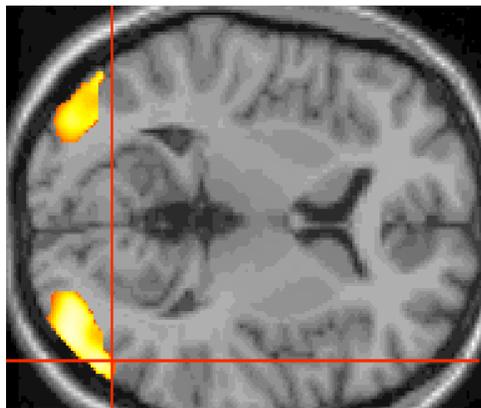


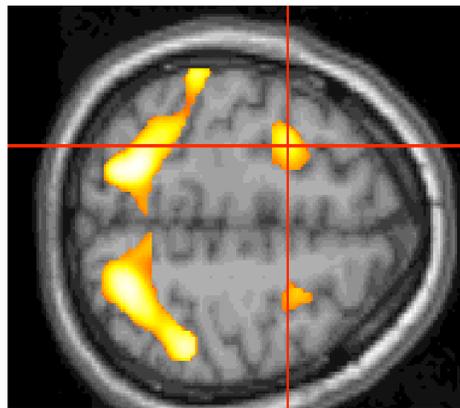
Figure 2



a

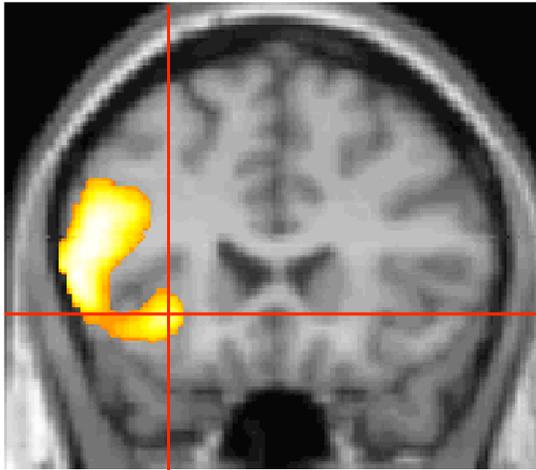


b

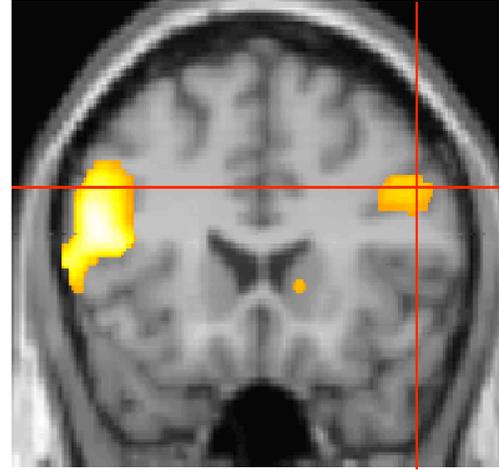


c

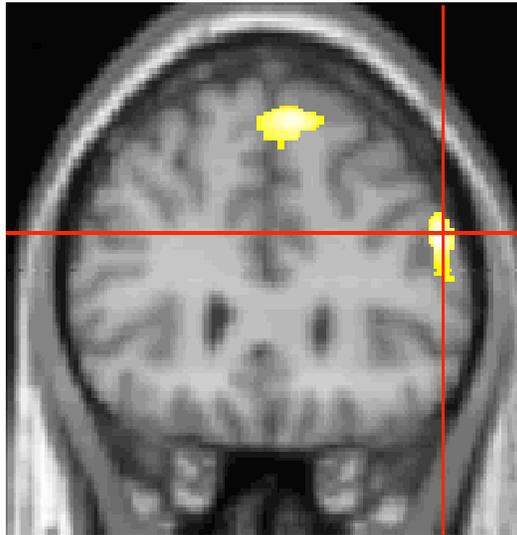
Figure 3



a



b



c

Figure 4

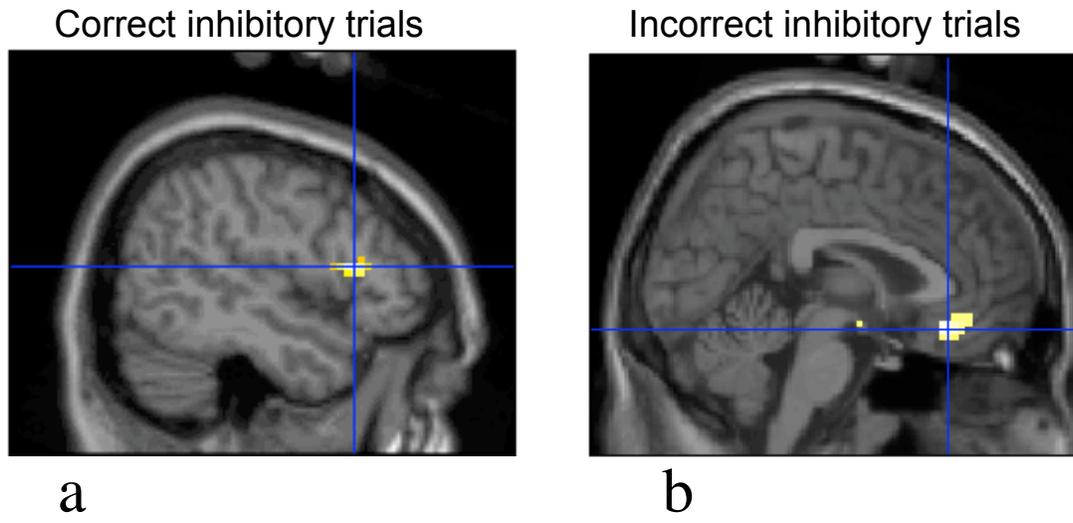


Figure 5