

Available online at www.sciencedirect.com



Cognitive Psychology 54 (2007) 185-217

Cognitive Psychology

www.elsevier.com/locate/cogpsych

Information-processing modules and their relative modality specificity

John R. Anderson ^{a,*}, Yulin Qin ^a, Kwan-Jin Jung ^b, Cameron S. Carter ^c

^a Psychology Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA ^b Brain Imaging Research Center, University of Pittsburgh and Carnegie Mellon University, Pittsburgh, PA 15203, USA ^c Departments of Psychiatry and Psychology, University of California at Davis, Sacramento, CA 95817, USA

> Accepted 25 June 2006 Available online 17 August 2006

Abstract

This research uses fMRI to understand the role of eight cortical regions in a relatively complex information-processing task. Modality of input (visual versus auditory) and modality of output (manual versus vocal) are manipulated. Two perceptual regions (auditory cortex and fusiform gyrus) only reflected perceptual encoding. Two motor regions were involved in information rehearsal as well as programming of overt actions. Two cortical regions (parietal and prefrontal) performed processing (retrieval and representational change) independent of input and output modality. The final two regions (anterior cingulate and caudate) were involved in control of cognition independent of modality of input or output and content of the material. An information-processing model, based on the ACT-R theory, is described that predicts the BOLD response in these regions. Different modules in the theory vary in the degree to which they are modality-specific and the degree to which they are involved in central versus peripheral cognitive processes. © 2006 Elsevier Inc. All rights reserved.

Keywords: Cognitive architecture; fMRI

0010-0285/\$ - see front matter © 2006 Elsevier Inc. All rights reserved. doi:10.1016/j.cogpsych.2006.06.003

^{*} Corresponding author. Fax: +1 412 268 2844. *E-mail address:* ja+@cmu.edu (J.R. Anderson).

1. Introduction

A significant issue in the understanding of human cognition is the degree to which human thought is independent of the conditions of input and output. The informationprocessing theories of the 1970s (e.g., Anderson & Bower, 1973; Kintsch, 1974; Norman & Rumelhart, 1975) advanced the position that information, after it is abstracted from the stimulus, proceeds through cognition independent of initial modality and ultimate response. At the time, this position was advanced in contrast to what were perceived as behaviorist theories that held that all thought was grounded in stimulus and response. A sophisticated development of this sort of position was that of Paivio (1971, 1986). Recently, this grounded position has been re-invigorated in cognitive psychology with the emphasis on embodied cognition. For instance, researchers like Barsalou (1999, 2005) and Glenberg (e.g., 1997; Glenberg and Kaschak, 2002) have emphasized the basis of cognition in perception and action (Barsalou more emphasizing perception and Glenberg more emphasizing action). According to a strong statement of this view, cognition does not utilize knowledge or process that is abstracted from perception and action.

Brain imaging data have been cited for both the modality-dependent and the modalityindependent positions. For instance, Barsalou, Simmons, Barbey, and Wilson (2003) cite imaging evidence that specific modalities are involved in thought processes. Similarly, Jonides, Lacey, and Nee (2005) use neural imaging data to argue that working memory is tied to particular modalities. On the other hand, a number of studies have shown regions that process information identically independent of modality of input (Buckner, Koutstaal, Schacter, & Rosen, 2000; and Le Clec'H et al., 2000; Schumacher et al., 1996). In a sign that things might not simply be one way or the other, Buckner and Wheeler (2001) argue that certain prefrontal regions involved in memory are sensitive to modality while others are not. Part of the issue is definitional. For instance, Schumacher et al. find memory regions that respond to verbal content but are insensitive to whether input is visual or auditory. Thus, we have regions that are specific to content of material (i.e., verbal) but not to modality of input. This seems neither the abstract propositional representation nor a representation dependent on a perceptual modality. With respect to language and modality-specificity, a number of studies of sentence processing have found regions that are sensitive to modality of input and other regions that are sensitive to comprehension difficulty but not modality (Bookheimer et al., 1997; Carpentier et al., 2001; Constable et al., 2004; Michael, Keller, Carpenter, & Just, 2001).

To make more headway on these issues, one needs to make some more specific theoretical commitments. This paper will use the modern ACT-R theory (Anderson, 2005; Anderson et al., 2004) and its connections to brain imaging data. The theory has its origins in the earlier generation of abstract information-processing theories, but the current ACT-R paints a much more complex picture of cognition. The theory identifies different cognitive modules that are relatively pure perceptual, relatively pure motor, relatively pure amodal, but also other modules that are not so clearly categorized. The theory makes predictions about the activation patterns that should occur in regions associated with these modules. We will present data from an experiment designed to exercise all the modules that have been identified by the theory. The outcome of this effort will not be a simple judgment on debate between the abstract versus embodied cognition, but hopefully a step towards a richer understanding of the mind, its realization in the brain, and its connection to the external world. We should emphasize from the outset that our goal is to better articulate the modality dependence of various aspects of cognition not to pit one theory against another. Indeed, we will be exploring issues that have not been addressed by existing theories. For instance, Barsalou's perceptual symbol hypothesis is about representation and really does not address process issues (Barsalou, personal communication).

1.1. The ACT-R theory and its brain mapping

This paper will use the connections that have been made between modules in the ACT-R theory and various brain regions. According to the ACT-R theory, cognition emerges from interaction of a number of modules. Each of these modules performs its operations largely independent of other modules but the module can interact with the overall cognitive system by placing information in an associated buffer. A central production system can respond to patterns in these buffers and take actions, which include sending requests to various modules. There are seven modules that have been specified in the ACT-R theory but no doubt there are more in the human mind. The production system itself can be considered an eighth module. Each of these eight modules has been associated with a brain region for purposes of brain imaging.

Two of the modules are more or less pure perceptual systems. There is an aural module, based on the EPIC (Meyer & Kieras, 1997) module that processes aural input. As this paper will show, activity in a region in the vicinity of the auditory cortex appears to reflect the processing of this module. ACT-R also has a visual module for object recognition. While the visual system is realized in many regions of the brain, we have found a region near the fusiform gyrus to be sensitive to kinds of operations we postulate, in line with other research (Grill-Spector, Knouf, & Kanwisher, 2004; McCandliss, Cohen, & Dehaene, 2003) that has shown this region plays a critical role in perceptual recognition.

Two other modules are motor systems—a manual system for programming hand movements and a vocal system for programming speech. The ACT-R implementations of these systems are also borrowed from the EPIC theory. As this paper will show, sensory and motor regions along the central sulcus reflect activity of these modules. The motor region reflecting the manual module is located at that part of the central sulcus that represents the hand and fingers. The motor region reflecting speech is below in the portion of the motor strip that is associated with the face and tongue. While there is some controversy about how refined the somatotopic representation of the human body is, it is generally conceded that such gross discriminations among major body parts are maintained (Sanes & Schieber, 2001). It appears that these regions are involved in more than external action. For instance, Anderson, Oin, Sohn, Stenger, and Carter (2003) found that the manual region was active during the rehearsal of responses over a delay interval when the output was going to be manual. In the research to be reported, the vocal region is also active in rehearsal of verbal output. Thus, these systems can be recruited to maintain representations of information. More generally these regions have been found to be involved in motor imagery (Grezes & Decety, 2001; Jeannerod, 1995).

Two modules are somewhat abstracted from specific input and output. One is an imaginal module that is responsible for holding spatial representations of problems such as states of solution for an equation (Anderson, 2005) or states of the problem in Tower of Hanoi (Anderson, Albert, & Fincham, 2005). We have associated a posterior parietal region with this module, consistent with research such as Dehaene, Piazza, Pinel, and Cohen (2002) Reichle, Carpenter, and Just (2000). The other is a retrieval module that is responsible for managing retrieval requests from declarative memory. We have found a prefrontal region to be sensitive to it, again consistent with other research (Buckner, Kelley, & Petersen, 1999; Cabeza, Dolcos, Graham, & Nyberg, 2002; Fletcher & Henson, 2001; Lepage, Ghaffar, Nyberg, & Tulving, 2000; Wagner, Maril, Bjork, & Schacter, 2001; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001).

While these regions appear largely independent of modality of input or output, they do not appear to be completely independent of the content of the material. One sign of this is the hemispheric specificity of these regions. The left prefrontal region tends to be more active for verbal material and the right prefrontal for pictorial material (Gabrieli, 2001). Again the left parietal region appears to be more active in cases when detailed, local information is critical and the right parietal region when more global information is critical (Robertson & Rafal, 2000). As almost of our research has involved the processing of detailed, symbolic information such as algebraic equations, we have always found strong left hemisphere activation although there is often weaker activation on the right displaying similar patterns. One unpublished studied in our laboratory (Kao et al., In preparation) involving geometry problems has produced greater activation in the regions on the right. Thus, it would appear that there is different hemispheric involvement for the imaginal and retrieval modules depending on content of the material.

Finally, there are two modules that appear to be closest to true amodal systems. One is a goal module that keeps track of ones intentions while performing a task and controls the information processing accordingly. This we associate with the anterior cingulate in line with the ideas of others about this region serving some sort of control function in information processing (Carter et al., 2000; D'Esposito et al., 1995; Posner & Dehaene, 1994). There is some controversy about whether different regions of the anterior cingulate are specific to different response modalities (e.g., Picard & Strick, 1996) but the most thorough meta-analyses seem to indicate that the anterior cingulate does not have regional response sensitivity (Barch et al., 2001). The second module is the production system itself, perhaps even more central to control since it determines the action that will be taken next. Identification of the basal ganglia as serving this control function is also consistent with the thinking of others (Ashby & Waldron, 2000; Hikosaka et al., 1999; Poldrack, Prabakharan, Seger, & Gabrieli, 1999; Saint-Cyr, Taylor, & Lang, 1988). We have found the head of the caudate to be a region whose activation seems to reflect production rule firing.

Over the course of our research we have identified these regions that seem to be sensitive to the activities of these ACT-R modules. It is important that we are able to use the same regions in study after study rather than using exploratory efforts that would identify (at least somewhat) different regions in each study. This provides a counter to claims (e.g., Uttal, 2001) that results do not replicate over experiments. This also avoids the many problems with spurious false effects in exploratory studies and provides unbiased estimates of region responses to test the predictions of the theory.

Our original efforts (Anderson et al., 2003) used an exploratory study to help with the identification of the parietal, prefrontal, and motor regions. Based on this study we identified the following regions, each 5 voxels wide, 5 voxels long, and 4 voxels high¹ that we have used in all subsequent studies:

¹ A voxel in our research is 3.125 mm long and wide and 3.2 mm high.

- 1. Prefrontal (Retrieval): Centered at Talairach coordinates $x = \pm 40$, y = 21, z = 21. This includes parts of Brodmann Areas 45 and 46 around the inferior frontal sulcus.
- 2. Parietal (Problem State or Imaginal): Centered at $x = \pm 23$, y = -64, z = 34. This includes parts of Brodmann Areas 7, 39, and 40 at the border of the intraparietal sulcus.
- 3. Motor1 (Manual): Centered at $x = \pm 37$, y = -25, z = 47. This includes parts of Brodmann Areas 2 and 4 at the central sulcus.

These regions were subsequently used in a number of studies (Anderson et al., 2005; Anderson, Qin, Stenger, & Carter, 2004; Qin, Anderson, Silk, Stenger, & Carter, 2004; Qin et al., 2003; Sohn, Goode, Stenger, Carter, & Anderson, 2003; Sohn et al., 2005). Exploratory analyses in most of these studies had identified an anterior cingulate region and Anderson (2005) reported the first effort to model it with ACT-R's goal module. This is a 5-voxel-long, 3-voxel-wide, and 4-voxel-high region:

4. Anterior Cingulate (Goal): Centered at $x = \pm 5$, y = 10, z = 38. This includes parts of Brodmann Areas 24 and 32.

Anderson (2005) also used a $4 \times 4 \times 4$ caudate region that had been first identified in Anderson et al. (2004). We have focused on this area because of the evidence that the basal ganglia served a production-rule-like function. Sometimes our imaging studies have found this region behaves in a way that can be predicted from production rule firing (e.g., Anderson, 2005).

5. Caudate (Procedural): Centered at $x = \pm 15$, y = 9, z = 2. This is a subcortical structure.

Three $5 \times 5 \times 4$ regions were identified for this paper because of its emphasis on assessing modality effects:

- 6. Fusiform Gyrus (Visual): Centered at $x = \pm 42$, y = -60, z = -8. This includes parts of Brodmann Area 37.
- 7. Auditory Cortex (Aural): Centered at $x = \pm 47$, y = -22, z = 4. This includes parts of Brodmann Areas 21, 22, and 42 in the region known as the auditory cortex. Note, however, this region excludes Brodmann Area 41, which is the primary auditory cortex.
- 8. Motor2 (Vocal): Centered at $x = \pm 44$, y = -12, z = 29. This includes parts of Brodmann Areas 2 and 4 at the central sulcus.

This completes the mapping of existing ACT-R modules onto brain regions. Of course, there are many brain regions not included above and cognitive functions not yet represented in ACT-R.

We need to state a few qualifying remarks about these brain-module associations. It might seem that we are claiming that the cortical regions listed perform the function associated with each module. While this is a plausible inference, it is not necessary to the logic of this approach. It is only necessary that the activity of the brain region reliably reflect a particular information-processing function. Even if the function is performed in that region, there is no reason to suppose that its activity will only reflect that function. Nonetheless, we have been fortunate that the regions seem to be rather pure indicators of their ascribed functions,² at least over the series of studies that we have performed. Finally, there is no claim that the ascribed function is restricted to a specific region. With respect to retrieval, for instance, the hippocampus gives a similar response (e.g., Wagner et al., 1998) but it appears not to give as strong a signal in our research. With respect to control, dorsolateral prefrontal structures probably play a role as well as the anterior cingulate. With respect to the caudate, we would expect to find a similar response in other structures connected to the basal ganglia, particularly the dorsal thalamus (and indeed we often do when we specifically check activity in this region).

1.2. The task

Our original interest in brain imaging came from our desire to better understand the problem solving involved in mathematical processing in algebra, and much of our research (e.g., Anderson et al., 2003; Oin et al., 2004) has looked at algebra equation solving. However, the problem with such naturalistic tasks is that multiple factors can be confounded in manipulations of task complexity. For instance, prefrontal and parietal activations tend to covary when one manipulates number of steps in equation solving, just as they tend to covary in a number of cognitive tasks (e.g., Acuna, Eliassen, Donoghue, & Sanes, 2002; Dehaene, Molko, Cohen, & Wilson, 2004; Marshuetz, Smith, Jonides, DeGutis, & Chenevert, 2000; Newman, Carpenter, Varma, & Just, 2003). Anderson et al. (2004) introduced a task that was better designed to separate two factors that are typically confounded in algebra equation solving. One is retrieval of relevant information such as arithmetic facts, which should affect the retrieval module, and the other is transformation of an internal representation (such as an equation) that should affect the imaginal module. The retrieval process is associated with the prefrontal region and the imaginal process with the parietal region. The task used in this paper will also attempt to pull these two regions apart. For other research separating these two regions see Sohn et al. (2003) and Sohn et al. (2005).

We wanted to have relatively separate measures of retrieval and problem transformation to separately assess the modality dependence of the retrieval and imaginal modules. Therefore, we continued use of the task in Anderson et al. (2004), modifying it just slightly so that it would be appropriate for pursuing the current question. In the first phase (outside the magnet), participants memorized information that they would use in the second phase of the experiment that took place in the magnet. The material to be memorized involved associations between two-letter words and two-digit numbers such as

 $AT \rightarrow 23$ and $BE \rightarrow 24.$

Then in the second phase of the experiment, participants either heard or saw permutations of the names "Dick," "Fred," and "Tom" paired with visual presentation of these two-letter words or two-digit numbers. Table 1 illustrates the various conditions of the experiment. Participants were told that the two-digit codes that they had learned were instructions for transforming the three-name sequences. Thus, "23" means that the second and third names should be switched. Applied to "Tom, Dick, Fred" it would produce "Tom, Fred, Dick." Some two-digit codes are "no-ops" such as "24" because one of the digits is greater than 3 and so in this case does not require a transformation. The

² Although we will see in this paper some overlap in the two motor regions.

Table 1

	No transformation	Yes transformation
No substitution	Stimulus: Tom Dick Fred Probe: 24 Response: Ring-index-middle	Stimulus: Tom Dick Fred Probe: 23 Response: Ring-middle-index
Yes substitution	Stimulus: Tom Dick Fred Probe: BE Response: Ring-index-middle	Stimulus: Tom Dick Fred Probe: AT Response: Ring-middle-index

Illustration of the four conditions of the experiment associations: AT is associated to 23; BE to 24; Dick to index; Fred to middle; Tom to ring

difference between no-op digit pairs and ones that require an operation is referred to as the transformation factor in Table 1 and throughout the paper. Participants can either be given the digit pair directly in which case no retrieval is required or be given a word from which they have to retrieve the digit pair. The requirement to perform this retrieval is referred to as the substitution factor in Table 1 and throughout the paper because it required the participant to substitute the digit for the word. Our expectation was that the transformation factor would more impact the parietal region for manipulating problem representation and that the substitution factor would more impact the prefrontal region for retrieving information.

Fig. 1 illustrates the procedure of the experiment as it was administered in the fMRI scanner. After a 1.5 s prompt, participants either heard or read names at the rate of one each half second. Then they either had a 4 s delay or not. The purpose of the delay was to manipulate the shape of the BOLD (Blood Oxygen Level Dependent) response as a test of our model. Then participants saw the digit or word instruction. They were instructed to perform the transformation mentally and to press the right thumb key when they were ready to give the answer. The time to press the thumb key is the most important behavioral measure reflecting the time to comprehend the instruction and plan the response. When the thumb key was pressed, they had to key in their letters quickly (if the output modality was manual) or say them quickly (if the output modality was vocal). The emphasis on speed of response meant that they had to have the response sequence preplanned. Anderson et al. (2004) showed that the insertion of the delay produced motor rehearsal of the responses and so we expected the insertion of a delay would increase activation of the motor region. This experiment will examine whether there are effects of delay in regions that do not have a motor function.

The major new manipulation in this study involved a manipulation of modality of input and output. As the input manipulation participants either saw the three names or heard



Fig. 1. The 28.5-s structure of an fMRI trial. In this instance, it shows that the instruction can either be a word (e.g., AT) or a number (e.g., 23).

them. Orthogonal to this they could either report the names by saying them or keying them. Those participants in the manual condition had been pretrained on a mapping of the names onto the index, middle, and ring fingers. While these factors would impact the activations associated with the predefined motor and sensory areas, the question was whether they would affect activation in the central regions.

There were five qualitative predictions for the outcome of the experiment:

- 1. Modality of input and output would not affect the behavior of the prefrontal, parietal, anterior cingulate, or caudate.
- 2. Delay would only affect the amount of activation in the motor areas and the control areas (cingulate and caudate) but not the perceptual areas or the parietal and prefrontal areas. Delay would affect the motor areas because they would be recruited to fill the delay and affect the control areas because extra goal states and production-rule firings are required to manage the rehearsal. The distribution of activity in other regions would be affected by the delay but the total activation would not change.
- 3. Transformation would affect the amount of parietal, anterior cingulate, and caudate activation but not prefrontal activation.
- 4. Substitution would affect the amount of prefrontal, anterior cingulate, and caudate activation but not parietal activation.
- 5. There would be stronger left lateralization of activation in the case of the prefrontal and parietal regions reflecting the symbolic nature of the task.

However, beyond these qualitative predictions the paper will try to fit the exact patterns of activation in these regions. The details of how predictions can be obtained will be presented after we report the experiment and assess the qualitative predictions.

2. Method

2.1. Participants

Fifty-two right-handed, healthy native English speakers, members of the CMU community (27 females) participated after informed consent. Their ages ranged from 18 to 33, with an average of 21.5. There were 13 participants in each group but because of problems, the final numbers were 12 in the visual-manual group, 11 in the visual-vocal group, 13 in the aural-manual group, and 13 in the aural-vocal group. One subject was not used in visual-manual group because of excessive head movement and 2 subjects were excluded from the visual-vocal group because of problems in recording their spoken answers. Table 2 provides some statistics on the four groups in terms of gender, age, and head movement during the course of the experiment. There are no significant differences in any of these statistics.

2.2. Task and procedure

The trial structure is illustrated in Fig. 1. A trial began with a prompt, which was a red asterisk in the center of a rectangle (35 mm width, 11 mm height, with white edges and located in the center of the black screen). In the visual conditions, the names would be presented inside this rectangle. After 1.5 s, the stimulus was presented at the rate of .5 s per

	Condition	Condition					
	Aural		Visual				
	Manual	Vocal	Manual	Vocal			
Age	21.2	21.6	21.8	21.3			
Gender	4F–9M	6F-7M	8F-4M	7F–4M			
Mean shift (mm)							
x direction	-0.09	-0.07	0.05	-0.20			
y direction	2.16	2.24	2.01	1.82			
z direction	0.24	-0.31	-0.43	-0.53			

Table 2 Participant statistics

name. The name sequences were random permutations of three names: 'Tom,' 'Dick,' 'Fred.' These names were presented in the rectangle in the visual condition while in the auditory condition the rectangle was empty and the participant heard prerecorded names. Then, for a random half of the trials, before the instruction stage there was a 4 s delay during which a white word "DELAY" appeared inside the rectangle. In instructions stage, a white number or word representing an instruction was presented in the rectangle and the participants were instructed to apply the instruction to transform the sequence of names. Participants pressed the thumb key when they were ready to key in the final solution. After the participant pressed the thumb key, the rectangle became empty again. The thumb press provided a measure of the planning time. If the planning time exceeded 10.5 s, the trial was scored as incorrect. Once they pressed the key they were suppose to say the responses as fast as possible with no more than 1 s per response (key press or speech). This rapid responding was designed to prevent participants from postponing transformations until after they had pressed their thumbs. After the participants keyed or said the answer, the correct answer (correct sequence of the names) was presented auditorily or visually (depending on condition) at the speed of 500 ms per name (feedback stage). Then a white plus would be shown inside the rectangle for the remaining portion of the 28.5 s trial (19 scans), and the participants were told they could rest but that they should still keep their eyes open during this period.

2.3. Pre-scan practice

On the day before the scan day, there was a pre-scan session in which participants memorized 12 pairs of word-to-number correspondences (participants practiced until they could produce three consecutive trials without any errors), and practiced the actual task. In the manual condition they were given practice on finger-to-name mapping. In the vocal condition, they were encouraged to respond in normal speech and avoid head movements. Participants were not told the significance of the two-digits numbers (that they were instructions for string transformation) until after they had completed the memorization phase of the experiment. There were four blocks (12 trials per block) of practice at the actual task.

2.4. Parametric design

The design involved six factors each with two values. Two between-participant factors were manipulated—input and output modality. Three other factors were manipulated

within participants. Half of the trials involved a 4-s delay and half did not. Half of the trials presented two-letter words and half presented two-digit numbers as instructions. 2/3 of the trials required a transformation of the name sequence and 1/3 did not. Of the 2/3 of the trials requiring a transformation half required exchange of a pair of neighbors such as from 'Fred' 'Tom,' 'Dick' to 'Fred,' 'Dick,' 'Tom' (when, say, the instruction was "23") and half required exchange of first and third position (instruction was either "13" or "31"). We will also treat hemisphere as another factor in the statistical analysis but it is not one that is manipulated experimentally.

Participants were tested for 12 blocks, in which each of the within-participant conditions was tested once. The 12 within-participant conditions occurred in random order. Since there were not significant differences between the neighbor exchanges and the distant exchanges we collapsed over this and simply used the contrast of transformation versus no transformation. Over the experiment there were 12 observations for each of the four no-transformation conditions (the four conditions obtained by crossing delay and substitution) and 24 observations for each of the four transformation conditions for each participant. Since there are 49 participants and most of the critical analyses of the imaging data (see Fig. 4) turn out to involve looking at 2×2 effects for different regions, this means that there are approximately between 500 and 1000 observations contributing to each data point.

2.5. Event-related fMRI scan

Event-related fMRI data were collected by using a gradient echo-planar-image (EPI) acquisition on a Siemens 3T Allegra Scanner. The imaging parameters were TR = 1500 ms, TE = 30 ms, RF flip angle = 55° , FOV = 200 mm, matrix size = 64×64 (3.125×3.125 mm per pixel), slice thickness = 3.2 mm, slice gap = 0 mm, and 26 axial slices per scan with the AC-PC on the 20th slice from the superior. In addition, structural images of two-dimensional T1 spin echo were acquired at the same slice locations. The audio system was a Silent Scan Model SS-3100 from Avotec and confirmed to deliver the name easily recognizable despite the background noise of the EPI scan. To check for the correctness of the vocal responses, the data were analyzed as described in Jung, Prasad, Qin, and Anderson (2005).

EPI images were realigned to correct head motion using the algorithm AIR (Woods, Grafton, Holmes, Cherry, & Mazziotta, 1998) with 12 parameters. The EPI images were smoothed with a 3D Gaussian kernel of a 6 mm FWHM. The structural images were co-registered to our reference structural images³ and the resulting registration parameters were used to co-register the corresponding EPI images. The group data were analyzed using the NIS system⁴ according to the event-related structure of Fig. 1.

³ Our Talairached reference image is availabe at MRI link available from our website act-r.psy.cmu.edu. This is the reference brain used for all studies from our laboratory.

⁴ (http://kraepelin.wpic.pitt.edu/nis/index.html).

3. Results

3.1. Behavioral results

Fig. 2 shows the mean latencies from the presentation of the instruction to the thumb press (the measure of planning time) as a function of condition. These means are for correct trials only. Modality of input had no significant effect on the time for this key press (auditory = 2.90 sec and visual = 2.70 s; F(1,45) = 0.91, MSE = 4.58). All the other factors are displayed in Fig. 2 and are significant—output modality (F(1,45) = 4.05, p = .05; MSE = 4.58), delay (F(1,45) = 20.48, p < .0001; MSE = 0.113), transformation(F(1,45) = 248.72, p < .0001; MSE = 0.977) and substitution (F(1, 45) = 291.65; $p \le .0001$; MSE = 0.561). There is only one significant interaction and that is between delay and output modality (F(1,45) = 21.00, p < .0001; MSE = 0.113). The effects of substitution and transformation are large (both over a second) and do not interact with anything else. Since the effects of transformation and of substitution are large and similar in size, if there is an effect of one factor and not the other in the imaging research it will not be simply because one factor was a more potent manipulation. The effects of output modality and delay are a bit more subtle. Participants responded equally quickly in the vocal condition for both delay and no delay, but they were .3 s slower than the vocal conditions in the delay manual condition and .6 s slower in the no-delay manual condition. It appears that they found it somewhat difficult to convert from the names to finger presses, particularly when they did not have the advantage of a 4 s delay in which they could work out the mapping.

The subsequent finger presses were recorded in the case of manual output. The only significant effect is that participants got faster as they key (.392 s for first, .318 for second,



Fig. 2. Mean latencies (on correct trials) from the presentation of the instruction to the thumb press as a function of delay, modality and whether a transformation or a substitution was required. Also displayed are the predictions of the approximate model.

and .300 for third, F(2,48) = 39.91, p < .0001; MSE = 0.012). These key times are very fast compared to the latency associated with the first thumb press indicating that participants did comply with instructions to plan their response before pressing the thumb key.

We also performed an analysis of accuracy and there were three main effects—modality of output (81.9% correct for manual and 91.6% correct for vocal, F(1,45) = 28.31, p < .0001; MSE = 0.030), transformation (89.5% correct for no transformation and 83.8% correct for transformation, F(1,25) = 20.22, p < .0001; MSE = 0.011), and substitution (90.0% correct for no substitution and 83.3% correct for substitution, F(1,45) = 26.69, p < .0001; MSE = 0.017). There was a significant interaction between input modality and output modality (F(1,45) = 14.44, p < .0005; MSE = 0.030). When the output was manual participants were 7% more accurate when the input modality was visual but when the output was vocal they were 7% more accurate when the input modality was aural. Except for this interaction the accuracy and latency effects largely lined up and the overall correlation across conditions was r = -.645.

In general, the behavioral analyses indicate that the cognitive factors of substitution and transformation had strong, robust effects and they should be potent factors in the imaging analyses below.

3.2. Imaging data: confirmatory analyses

All analyses of imaging data were for correct trials only. We will first report effects for the predefined regions. Then we will report a series of exploratory analyses to determine what other regions might be showing systematic effects. For each of the predefined regions we will perform a $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2$ analysis of variance where the within-participants factors will be hemisphere, delay, transformation, and substitution and the between-participants factors are modality of input and output. The dependent measure is an estimate of the magnitude of the BOLD response for a trial. Each trial consists of 19 scans and the BOLD response varies across these scans. The dependent measure will be the degree to which the BOLD responses on scans 3–14 are above the baseline set by the BOLD responses on scans 1 and 2. For each of these 12 scans we measured their percentage above the baseline defined by the first two scans and our dependent measure is the sum of these 12 percentages. The later scans 15–19 are excluded from this analysis because the BOLD response tends to dip below baseline at the end (an issue discussed later). Fig. 4 displays the BOLD responses for all the scans and the predictions of a model to be described in a later section.

Table 3 summarizes all the analyses—reporting in part (a) for each of the eight predefined regions which factors were statistically significant and in part (b) the size and direction of the main effects. Fig. 3 attempts to illustrate these results showing the magnitudes of the *F*-values for the two "peripheral" factors of input and output modality and the two "central" factors of transformation and substitution. While there are details that will be discussed below, it is apparent that the factors of input and output modality only affected the perceptual and motor regions and, with one weak exception the cognitive factors only affected the cognitive regions.

3.3. Fusiform gyrus (visual) region

There were two significant main effects in the fusiform region which reflects visual processing—input modality (F(1,45) = 4.44, p < .05; MSE = 0.00531) and laterality

Table 3 Summary of the effects of the six factors in the eight regions

	Factor							
Region	Input	Output	Laterality	Delay	Transform	Substitute		
(a) Significance								
Fusiform gyrus	<.05	>.25	<.01	>.05	>.25	>.50		
Auditory cortex	<.0001	<.05	>.05	<.05	<.05	>.5		
Motor1	>.5	<.005	<.0001	<.001	>.5	>.5		
(manual)								
Motor2	>.05	<.01	>.1	<.0001	>.25	>.1		
(vocal)								
Parietal	>.25	>.25	<.05	<.0005	<.005	>.05		
Prefrontal	>.1	>.5	<.05	<.005	>.5	<.05		
Anterior	>.5	>.5	<.0001	<.0001	<.05	<.0005		
cingulate								
Caudate	>.5	>.5	<.01	<.001	>.5	>.25		
Region	Visual – Aural	Manual – Vocal	Left - Right	Delay - No Delay	Transform - Not	Substitute - Not		
	(%)	(%)	(%)	(%)	(%)	(%)		
(b) Effect size								
Fusiform gyrus	1.16	-0.69	0.58	0.39	-0.24	0.13		
Auditory cortex	-2.88	-1.28	-0.35	0.44	-0.40	-0.10		
Motor1	0.27	2.46	2.28	0.88	0.08	0.04		
(manual)								
Motor2	1.03	-1.81	0.27	0.80	-0.11	0.19		
(vocal)								
Parietal	0.92	0.26	1.14	0.62	0.47	0.37		
Prefrontal	1.00	0.27	0.66	0.50	0.04	0.53		
Anterior	0.17	-0.17	1.86	0.85	0.36	0.61		
cingulate								
Caudate	-0.07	-0.13	-0.31	0.46	0.04	0.10		



Fig. 3. A display of the *F*-values for the main effects of input modality, output modality, transformation, and substitution for the 8 predefined brain regions.



Fig. 4. Observed (dotted lines connecting points) BOLD responses and predictions (solid lines) for the eight predefined regions. The data and predictions are plotted as a function of the mean time of each scan. (a) Effects of input modality and delay of the left fusiform gyrus. (b) Effects of input modality and delay on the left and right auditory cortex. (c) Effects of output modality and delay on the left motor area that is associated with the right hand. (d) Effects of output modality and delay on the left and right motor areas that are associated with the face and tongue. (e) Effects of transformation and delay on the left parietal region (f) Effects of substitution and delay on the left prefrontal region. (g) Effects of substitution and delay on the left anterior cingulate. (h) Effects of substitution and delay on the right caudate.

(F(1, 45) = 8.08; p < .01; MSE = 0.00074). The left fusiform gyrus shows the stronger result reflecting that the left region responds more to symbolic input (McCandliss et al., 2003). Fig. 4a displays the BOLD response for the four combinations of the input and delay factors for the left region because it gives the stronger response. The response is stronger for visual input but there is a response in the aural condition. Even in aural condition, the initial warning, the instruction prompt, the delay warning, and the final rest signal are visually presented and so it is not surprising that there is a response in this condition. Note in this region that the delay factor results in a stretching of the BOLD response since some of the input comes after the delay. However, there is no significant effect of delay on area under the curve implying that there is approximately the same amount of total processing in both the immediate and the delay condition. Also relevant to the issues of this paper, there were no significant effects of substitution or transformation on this region and so this region is not involved in supporting these cognitive activities.

3.4. Auditory cortex (aural) region

There were four significant main effects on the auditory cortex which reflects aural processing—input modality (F(1,45) = 25.72, p < .0001; MSE = 0.00641), output modality (F(1,45) = 4.78,p < .05;MSE = 0.00641), delay (F(1,45) = 5.57,p < .0001;MSE = 0.00070), and transformation (F(1, 45) = 7.19; p < .05; MSE = 0.00043). The output effect is for there to be a greater response in the condition where participants give an vocal response, as they hear themselves. There is no response in this region in the visual-motor condition where participants see the material and do not hear their responses. In terms of area under the curve (sum of the BOLD response for scans 3-14), the values are .26% for the visual-manual condition, 1.59% in the visual-vocal condition, 3.24% in the aural-manual condition, and 4.44% in the aural-vocal condition. Thus, the effect of input and output are additive. Since there are no laterality effects in this region, Fig. 4b displays the BOLD response averaged over the left and right regions for the four combinations of the input and delay factors. While there is an effect of delay on area under the curves in scans 3-14, this is because the two no-delay conditions that dip below zero before scan 14. If we only look at the area above zero in the aural input condition, there is no difference -4.30% in the no delay condition and 4.31% in the delay condition. Thus, like the visual region, this region does not seem to be involved in bridging the delay in contrast to the motor regions. This is the only perceptual or motor region to show a significant effect of a cognitive factor—in this case the transformation factor. However, the result is lower activation in the transformation condition—which is an anomalous result. We suspect that this may not be a real result (it is only significant at the .05 level). In any case this is not a result that suggests that the auditory area is involved in performing the transformation.

3.5. Motor1 (manual) region

There were three significant main effects on the motor1 region which reflect manual programming—output modality (F(1,45) = 14.74, p < .0005; MSE = 0.00872), laterality (F(1,45) = 34.12, p < .0001; MSE = 0.00283), and delay (F(1, 45) = 14.93; p < .0005; MSE = 0.00090). The left region is more active—a result that is not particularly surprising as the participants were responding with their right hands. There are interactions are between delay and laterality and between output modality and laterality, reflecting the fact that the effects only showed in the left hemisphere. Therefore, Fig. 4c displays the BOLD response for the four combinations of the other two factors in the left motor1 region. As expected there is a much stronger response when the output modality is manual. The residual response when the modality is vocal may just reflect the difficulty in separating the regions for vocal control from those for motor control. Also, participants still issue a thumb press in the vocal condition. The delay factor has an effect on the shape of the response with the peak delayed by the four seconds.

There are a couple of noteworthy aspects of the motor response which suggests, as in Anderson et al. (2004), that this region was involved in more than just response generation. First, note that it is rising in advance of the response generation. Second, note that the area under the curve is greater in the delay condition. In this case, the difference is not due to the late negativity as it was for the auditory regions—if we just look at the area above zero it is 6.05% in the manual delay condition and 5.05% in the manual no delay condition. Anderson et al. (2004) suggested this region was involved in response rehearsal and these results are consistent with that suggestion. Relevant to the issues of this paper, there were no significant effects of substitution or transformation on this region and so this motor region in is not involved in supporting these cognitive activities.

3.6. Motor2 (vocal) region

There were two significant main effects on the motor2 region which reflects vocal programming—output modality (F(1,45) = 9.51, p < .005; MSE = 0.00707) and delay (F(1,45) = 23.33; p < .0001; MSE = 0.00055). Since there are no significant laterality effects, Fig. 4d displays the BOLD response for the four combinations of output and delay for the average of the left and right regions. The effects are similar to the motor1 region except that this region is more responsive to vocal output. There is a weak response to manual output that may reflect lack of separation of the two regions or perhaps verbal rehearsal even in the manual condition. As the motor1 region, the early rise and effects of delay suggest that this region is involved in rehearsal of the response. Again, relevant to the issues of this paper, there were no significant effects of substitution or transformation on this region.

3.7. Parietal (imaginal) region

There were three significant main effects on the parietal region which is hypothesized to reflect changes in problem representation—laterality (F(1,45) = 4.38, p < .05;MSE = 0.00287), delay (F(1,45) = 24.65, p < .0005; MSE = 0.00049), and transformation (F(1, 45) = 9.70; p < .005; MSE = 0.00043). The effect is stronger in the left parietal region as we have found in other studies. Therefore, Fig. 4e displays the BOLD response for the four combinations of the other two factors in the left parietal region. The factors of delay and output modality yield a significant interaction (F(1, 45) = 6.34; p < .05;MSE = 0.00049). The area under the curve is significantly less in the case of vocal-no-delay (2.50%) than the other conditions (vocal-delay 3.52%, manual-no-delay 3.58%, manual-delay 3.79%). This is another case where a purported central region seems to be affected by modality. We will try to place an interpretation on the interaction after reviewing the prefrontal region where a similar interaction occurred. Otherwise, this region behaved as predicted. In particular, it showed effects of transformation and laterality and did not show effects of input modality or substitution.

3.8. Prefrontal (retrieval) region

There were three significant main effects on the prefrontal region which reflects retrieval operations—laterality (F(1,45) = 4.10, p < .05; MSE = 0.00137), delay (F(1,45) = 10.39, p < .005; MSE = 0.00047), and substitution(F(1, 45) = 5.94; p < .05; MSE = 0.00091). The effect is stronger in the left prefrontal region as we have found in other studies. Therefore, Fig. 4f displays the BOLD response for the four combinations of the other two factors in the left prefrontal region. As was the case for the parietal region, the factors of delay and output modality yield a significant interaction (F(1, 45) = 8.64; p < .01; MSE = 0.00047). Again, the area under the curve is significantly less in the case of vocal-no-delay (1.28%) than the other conditions (vocal-delay 2.22%, manual-no-delay 2.10%, manual-delay 2.12%). The fact that the parietal effect replicates in the prefrontal region even more significantly suggests that this unexpected result is not a spurious effect. The prefrontal region does confirm the predictions that there are main effects of substitution and laterality and not main effects of transformation, modality of input, and modality of output.

Given that one purpose of this experiment was to confirm a difference between the parietal and prefrontal regions we did contrasts to compare the magnitude of the transformation and substitution effects in the two regions. The effect of transformation is significant greater in the parietal (t(48) = 1.85; p < .05; MSE = .00034) but the effect of substitution is only marginally greater in the prefrontal (t(48) = 1.32; p < .10; MSE = .00032).

With respect to the unexpected interactions between delay and modality of output for the prefrontal and parietal, we think that this is related to the fact that generating the response is particularly simple in the vocal no-delay condition where one can basically repeat back what one has just heard. If there is a delay it is necessary to maintain the response over the delay and in the model to be described below, we are going to assume a role for the parietal and prefrontal in bridging the delay. This leaves unexplained why these regions also respond more strongly in the manual no-delay condition and we think this is related to the fact that in this condition the response code (key presses) is arbitrarily related to the input (names). However, this is not an effect that we will be modeling.

3.9. Cingulate (goal) region

While there was not a significant effect of either input or output modality in anterior cingulate, there were significant main effects of the other four factors—laterality (F(1,45) = 42.91, p < .0001; MSE = 0.00129), delay (F(1,45) = 24.48, p < .0001; MSE = 0.00058), transformation (F(1,45) = 5.97, p < .05; MSE = 0.00039), and substitution (F(1, 45) = 17.88; p < .0001; MSE = 0.00043). There is a significant interaction between input modality and laterality (F(1, 45) = 15.28; p < .0005; MSE = 0.00129) such visual input produced a stronger response than aural input on the right (3.17 versus 2.39%) while auditory input produced a stronger response than visual input on the left (5.09 versus 3.86%). We have no explanation of this effect. For both forms of input the response is stronger on the left side, an effect that may be related to morphological differences between left and right anterior cingulate (Yücel, Stuart, & Maruff, 2001). As a consequence, Fig. 4g displays the BOLD response only for the left region. The effect of input modality is only marginally significant in an analysis restricted just to the left side (F(1, 45) = 3.08; p < .1; MSE = 0.00474) and it does not even approach any significant interactions. In fact, there

are no significant interactions in the left cingulate. As the effects of delay and substitution are stronger than transformation, Fig. 4g displays the four combinations of these two factors. Except for the laterality result, the effects in this region are as predicted.

3.10. Caudate (procedural) region

There were two significant effects in this region—laterality (F(1,45) = 8.46, p < .01; MSE = 0.00022) and delay (F(1,45) = 17.88, p < .0001; MSE = 0.00023). In contrast to other regions, the right caudate gave a stronger response. We did not have any prior hypotheses about the laterality of response in this region nor a post hoc explanation of the result. Fig. 4h displays the BOLD response only for the right region, displaying the four combinations of delay and substitution. The null effects of input and output modality were predicted, as was the effect of delay. On the other hand, we also predicted effects of substitution and transformation on this region, which were not significant. While there might appear to be a weak effect of substitution in Fig. 4h, it did not approach statistical significance (F(1,45) = 0.57, MSE = 0.00032). However, Fig. 4h makes two other points apparent. First, the response in the caudate is weak, making it hard to get significant effects. Second, the effect predicted for substitution is quite small.

3.11. Confirmatory analyses: Summary

These confirmatory analyses have involved looking at 6 factors in 8 regions. Given the number of tests involved, the effects have been remarkably consistent with the predictions. With respect to the four motor and perceptual areas, each was responsive to the perceptual or motor factor that should drive it. There was evidence that the motor regions were being used to bridge the delay. There were no cases where these regions responded more in dealing with transformation or substitution, although there was the anomalous case of the auditory area responding weaker in the case of transformation.

With respect to the parietal and prefrontal areas, they were both left lateralized as expected given the symbolic nature of the task. Neither region was affected by modality of input or output. As predicted, parietal responded only to transformation confirming its role in problem representation and prefrontal responded only to substitution confirming its role in retrieval. The unexpected result for these regions was the interaction between delay and modality of output. One way of characterizing this interaction is that there was an effect of delay only when the output modality was vocal. Consistent with the results in the manual condition, the earlier Anderson et al. (2004) experiment, which only used manual output, did not find an effect of delay in this region. This effect might be related to difficulty of maintaining or mapping the code. Except for this, there were no interactions in these regions involving input modality or output modality.

As predicted, the anterior cingulate and caudate showed no effects of input or output modality. The anterior cingulate responded to all of the factors of delay, transformation, and substitution—confirming unpublished results from Anderson et al. (2004). The caudate responded to delay as predicted but not significantly to transformation or substitution in contrast to prediction. Reanalyzing data from Anderson et al. (2004), the caudate had not responded differentially to transformation or substitution in that exper-

iment either. Other research on algebra (e.g., Anderson, 2005) has found the caudate does respond when these two factors are combined in an overall manipulation of complexity. In this experiment there is a weak response in the caudate and the predictions are only for weak effects.

3.12. Imaging data: Exploratory analyses

This section will report exploratory analyses to determine what other regions might be reliably responding to the factors in the experiment. For the five manipulated factors (input modality, output modality, delay, transformation, and substitution) we did condition-by-scan interaction looking for regions that responded differently in the two conditions. To have a conservative test that dealt with non-independence of successive scans we used the correction of assigning only 1 degree of freedom to the numerator in the *F*-statistic for the interaction term and required a minimum of 30 contiguous voxels with significant interaction at $p \leq 0.01$ (Forman et al., 1995). The analysis with delay was not informative as it tended to pick up huge regions that varied as a function of trial structure,

Table 4Results of exploratory analyses—see Fig. 5

Region of interest	Brodman area (s)	Voxel count	Stereotaxic coordinates (mm)			Condition 1	Condition 2	
			x	y z				
(a) Input by scan						Visual (%)	Aural (%)	
a1. Right superior parietal	40	36	38	-57	45	8.04	1.46	
a2.Right auditory cortex	22	509	47	-13	6	0.66	3.11	
a3. Left auditory cortex	22	437	-51	-19	2	0.79	3.61	
a4. Visual cortex	17	64	17	-83	11	-0.09	0.33	
(b) Output by Scan						Manual (%)	Vocal (%)	
b1. Left motor region	2, 4, 40	544	-39	-27	42	4.52	1.32	
b2. Right motor region	2, 40	48	38	-28	36	2.36	0.39	
b3. Right motor region	2, 4, 40	160	47	-10	15	0.37	3.63	
b4. Left motor region	4, 6, 43	70	-54	-8	14	1.53	4.94	
b5. Right inferior prefrontal	47	67	47	32	-4	-1.38	2.04	
b6. Medial frontal gyrus	24, 25, 32	254	-2	17	-6	0.67	-1.92	
(c) Transformation × scan						No trans (%)	Trans (%)	
c1. Anterior cingulate	6, 24, 32	255	$^{-2}$	5	43	3.71	4.20	
c2. Left motor	2	398	-42	-6	37	3.06	3.58	
c3. Left posterior parietal	40	156	-27	-51	41	4.17	4.82	
c4. Polar frontal	10, 32	97	$^{-2}$	44	$^{-2}$	-2.36	-3.59	
c5. Left auditory cortex	22, 42	352	-51	-18	9	3.04	2.72	
c6. Right auditory cortex	22, 42	157	50	-16	9	3.00	2.55	
(d) Substitution \times scan						No sub	Sub	
d1. Anterior cingulate	6, 24, 32	210	$^{-2}$	11	42	3.69	4.52	
d2. Left prefrontal	44	372	-45	7	26	2.42	3.38	
d3. Left posterior parietal	19, 39, 40	165	-27	-60	38	4.45	5.17	
d4. Right supramarginal gyrus	40	46	50	-31	30	1.02	0.36	
d5. Left auditory cortex	42	140	-54	-21	9	3.45	3.32	
d6. Right auditory cortex	22, 42	41	50	-13	9	3.01	3.03	
d7. Left inferior frontal	47	46	-32	20	-1	1.93	2.82	
d8. Polar frontal	10, 32	53	1	38	$^{-8}$	-2.91	-3.75	

but Table 4 and Fig. 5 present the results of the remaining analysis. The significant interactions that identified these regions can usually be characterized by a difference in the percent rise or fall in the BOLD response in scans 3–14 rose from the baseline set by scans 1 and 2. Thus, Table 4 reports for each region found for each condition the sum of the percentages for these 12 scans (the same dependent measure used in the confirmatory analysis). Fig. 5 locates the exact regions on the brain slices.

Table 4a and Fig. 5a show the four regions with significant input-by-scan interactions in the exploratory analysis. This analysis identified areas in the auditory cortex (a2 and a3) that were close to the predefined auditory regions and responded more strongly to auditory input. It failed to find regions in the fusiform gyrus despite the significant effects that were in the region. It did find a right superior parietal region (a1) that did respond more



Fig. 5. Activation map for 20 slices, starting with the second slice from the top. Only regions with more than 30 contiguous voxels and p < .01 are shown. See Table 4 for identification of regions. The AC-PC line is slice 19. (a) Input-by-scan interaction; (b) Output-by-scan interaction; (c) Transformation-by-scan interaction; and (d) Substitution-by-scan interaction. The left side of the brain is shown on the left side. The predefined regions are shown as black squares unless covered by exploratory regions. Each exploratory ROI is numbered (for coordination with Table 4) in the first slice at which it appears.

strongly in the case of visual input. It also found a region on the edge of the visual cortex (a4) that responded more strongly to auditory input.

Table 4b and Fig. 5b show the six regions with significant output-by-scan interactions in the exploratory analysis. This analysis identified a large left region (b1), overlapping with the predefined motor1 region, which displayed a stronger response in the case of a manual response. It also found a much smaller homologous right region (b2) that similarly responded. Lower, it similarly found right and left motor regions (b3 and b4), overlapping with the motor2 regions that responded more strongly in the case of a vocal response. It found a prefrontal region (b5) that is a right homolog of left Broca's area that responded more strongly to vocal output.⁵ Finally, it found a low medial frontal area (b6) that responded positive to manual output and negatively to vocal output.

Table 4c and Fig. 5c show the six regions with significant transformation-by-scan interactions in the exploratory analysis. It does find two regions that overlap with regions we predict will be sensitive to transformation. Region c1 overlaps with the predefined anterior cingulate and region c3 overlaps with the predefined parietal region. In addition it finds a left motor and prefrontal area (c2) that responds more strongly to transformation. Anderson et al. (2003) did find a very similar prefrontal region in their study of algebra. Region c4 is a polar frontal region that responds negatively and more so when there is a transformation. Such negative responding polar frontal regions are frequently obtained in this research (Gusnard & Raichle, 2001). Finally, regions c5 and c6 are auditory regions that are selected mainly because of the delayed auditory peak to feedback in the transformation condition, not a difference in overall magnitude of response. In fact, the response is slightly weaker in the transformation condition in these areas.

Table 4d and Fig. 5d show the eight regions found to show significant substitution-byscan interactions in the exploratory analysis. It does find two regions that overlap with regions predicted to be sensitive to substitution. Region d1 overlaps with the predefined anterior cingulate and region d2 overlaps with the predefined prefrontal region. There is also a significant left posterior parietal region (d3) that overlaps with the predefined region; the predefined left parietal region almost showed a significant effect of substitution. Thus, it appears that the parietal may have some involvement in the process of substitution as indicated in some of the more detailed ACT-R models (e.g. Anderson et al., 2004). Region d4 is the right supramarginal gyrus and it displays a weaker effect in the case of substitution. Regions d5 and d6 are in the auditory cortex and, as in the case of transformation, they are selected because of the delayed auditory peak to feedback in the substitution condition. Regions d7 is in the anterior left insula. Region d8 is the polar frontal region that is responding negatively as in the case of transformation.

By way of summary, the exploratory analyses are roughly consistent with the confirmatory results. Except for the failure to find an effect in the fusiform gyrus for the predefined input-by-scan analysis, there were regions that showed all of the predicted effects. Most of the other regions were interpretable within our analysis.

While such exploratory analyses can be important in identifying effects that might not be found in the predefined analysis, they have a number of weaknesses. One weakness, illustrated by the failure to find a significant effect in the fusiform gyrus, is that they are conservative to avoid false alarms with so many statistical tests. Thus, they fail to find real

⁵ A lowered threshold found a homologous left region.

effects. Another weakness is that, because the regions are selected for statistical significance, they result in biased estimates of the actual effect size in the regions. This is why we use the predefined regions for model fitting.

3.13. Fit of an ACT-R model to the BOLD responses

Fig. 4 already illustrated the ability of the ACT-R model to account for the exact shape of the BOLD responses obtained in the different regions. Fig. 6 illustrates the information-processing steps of this model for the 8 within-participant conditions of the experiment.⁶ This model is a somewhat simplified version of the one described in Anderson et al. (2004). One major simplification is that all ACT-R modules took time that could be measured in increments of one-quarter second as illustrated in the figure. As another approximation, we have ignored the slight effects of modality of output and delay on response time and have assumed that the timing is identical in all combinations of input and output modality. (Fig. 2 shows the timing predictions of the model.)

We will explain all the steps in the simplest condition (no delay, no transformation, no substitution) and then explain how the more complex conditions are elaborated. We will also initially describe the model for the visual-manual condition and then describe how things differ with choice of input or output modality.

Scan 1. Two hundred and fifty milliseconds are spent to encode the red asterisk warning of the next trial.

Scan 2. The three names are presented at a rate of one per half second. Each name requires an encoding plus a motor recoding into a finger response.

Scan 3. Two hundred and fifty milliseconds are spent to encode the instruction, which is a number like "24" that indicates no transformation. The model then rehearses the finger responses (to give them fast enough). Then the model hits the thumb key and begins to key out the response.

Scan 4. The last two keys of the response are issued. Seven hundred and fifty milliseconds later the feedback begins and the participant encodes the feedback which occurs at a rate of one name every half second. Each name is encoded and compared to the memory of the response.

Scan 5. The encoding and comparison continues. After encoding the feedback a white plus is encoded indicating the end of the trial.

The substitution conditions involve an extra 1.25 s during which the digit instruction denoted by the word is retrieved and the transformation conditions involve an extra 1.50 s during which two names in the sequence are re-ordered. Anderson et al. (2004) analyzes in more detail how ACT-R performs these computations but here we just treat them as unanalyzed operations. The delay involves inserting an extra 4 s after scan 2 and before the instruction. In line with the model in Anderson et al. (2004) the model rehearses the digits once during this interval. This produces the interaction in the model between modality of output and delay.

⁶ A spreadsheet version of this model can be obtained from the Models link of our website: act-r.psy.cmu.edu.

			No L	Delay		Delay			
		No Tranformation Transformation				No Tranformation Transformation			
		No Sub	Sub	No Sub	Sub	No Sub	Sub	No Sub	Sub
Scan	Time	1.0	~~~~		~~~~	1.0.040	~	1.0.040	~~~
	0.00	Encode-V	Encode-V	Encode-V	Encode-V	Encode-V	Encode-V	Encode-V	Encode-V
	0.25								
1	0.50								
1	0.75								
	1.00								
	1.25								
	1.50	Encode	Encode	Encode	Encode	Encode	Encode	Encode	Encode
	1.75	Recode	Recode	Recode	Recode	Recode	Recode	Recode	Recode
2	2.00	Encode	Encode	Encode	Encode	Encode	Encode	Encode	Encode
_	2.25	Recode	Recode	Recode	Recode	Recode	Recode	Recode	Recode
	2.50	Encode	Encode	Encode	Encode	Encode	Encode	Encode	Encode
	3.00	Encode-)/	Encode-)/	Encode-)/	Encode-)/	Encode-V/	Encode-V	Encode-V	Encode-V/
	3 25	Dehearse	Substitute	Transform	Substitute	Dehearse	Dehearce	Dehearse	Dehearse
	3 50	Rehearse	Substitute	Transform	Substitute	Rehearse	Rehearse	Rehearse	Rehearse
3	3.75	Rehearse	Substitute	Transform	Substitute	Rehearse	Rehearse	Rehearse	Rehearse
	4.00	Motor-M	Substitute	Transform	Substitute	nonouroo	nonouroo	nonouroo	nonouroo
	4.25	Motor	Substitute	Transform	Substitute				
	4.50	Motor	Rehearse	Transform	Transform				
	4.75	Motor	Rehearse	Rehearse	Transform				
4	5.00		Rehearse	Rehearse	Transform				
4	5.25		Motor-M	Rehearse	Transform				
	5.50		Motor	Motor-M	Transform				
	5.75	Encode	Motor	Motor	Transform				
	6.00	Compare	Motor	Motor	Rehearse				
	6.25	Encode		Motor	Rehearse				
5	6.50	Compare			Rehearse				
5	6.75	Encode			Motor-M				
	7.00	Compare	Encode	_	Motor	Encode-V	Encode-V	Encode-V	Encode-V
	7.25	Encode-V	Compare	Encode	Motor	Rehearse	Substitute	Transform	Substitute
	7.50		Encode	Compare	Motor	Rehearse	Substitute	Transform	Substitute
	7.75		Compare	Encode		Renearse	Substitute	Transform	Substitute
6	8.00		Encode	Compare		Motor-M	Substitute	Transform	Substitute
	8 50			Compare	Encode	Motor	Debearce	Transform	Transform
	8 75		LIICOUE-V	Encode-V	Compare	Motor	Dehearse	Deheorse	Transform
	9.00			LINCOUC V	Encode	MOLOI	Rehearse	Rehearse	Transform
	9.25				Compare		Motor-M	Rehearse	Transform
_	9.50				Encode		Motor	Motor-M	Transform
7	9.75				Compare	Encode	Motor	Motor	Transform
	10.00				Encode-V	Compare	Motor	Motor	Rehearse
	10.25					Encode		Motor	Rehearse
	10.50					Compare			Rehearse
	10.75					Encode			Motor-M
0	11.00					Compare	Encode		Motor
0	11.25					Encode-V	Compare	Encode	Motor
	11.50						Encode	Compare	Motor
	11.75						Compare	Encode	
	12.00						Encode	Compare	
	12.25						Compare	Encode	-
9	12.50						Encode-V	Compare	Encode
- E	12.75							Encode-V	Compare
	13.00								Encode
<u> </u>	13.25								Compare
	12 75								Compare
	14.00								Compare Encode V
10	14 25								Encoue-v
	14 50								
	14 75								
	14.75			12 S		I			

Fig. 6. Activity of the various modules in the eight conditions of the experiment.

The operations to encode the names will depend on the modality of presentation but there are always three visual items presented—the initial prompt, the instruction, and the final end-of-trial signal. Therefore, we have denoted these as "Encode-V" in Fig. 6 to indicate that they are always visual, independent of input condition. In addition, there is a visual "Delay" presented at the beginning of the delay. Similarly, most of the motor operations and rehearsal operations will depend on modality of output. However, participants always press their thumb to indicate the initiation of output (to give a reliable and constant measure of time). Therefore, this operation is denoted as "Motor-M" in Fig. 6 to indicate that it will always be manual, independent of output condition. Its timing determines the latency in Fig. 2 (measured from the presentation of the instruction).

As we will explain, the time course in Fig. 6 was used to generate predictions for the BOLD responses that we would see in the various regions. Therefore, it is important to know how the lengths of the processes were determined. The quarter second for the encode, recode, rehearse, and motor steps are based on established ACT-R parameters (Anderson et al., 2005) rounded to the quarter second. The compare time was set to be a quarter second to conform. The length of the substitute and transform operations were determined to fit the latency data in Fig. 2.

The activity in Fig. 6 generates the predictions for the BOLD response in the various regions. We will go through the details of how these predictions are obtained. The activities in Fig. 6 create demands at different points in time for processing in certain modules. For instance, whenever a visual encoding is required there will be a demand on the visual module. We represent this as a demand function, D(t) for the module at various points in time. D(t) is the proportion the module is active during the scan that includes t. Thus, since the visual module is required 1/6th of the time during the first scan to encode the prompt, D(t) will be 1/6 for the first 1.5 s of the trial. Fig. 7 displays the value of the demand functions during the first 15 scans of the trial for each module.⁷ Fig. 7 also illustrates how we dealt with the negativity at the end of the trials. We were able to model this by assuming that the demand function went slightly negative (-.1) for 7.5 s after the cue signaling trial end. This is illustrated in Fig. 7 by the negative values of the demand function at the end of the trial. This might be viewed as a "relaxation" response in the region of interest where the system allows itself to go into a resting mode during the inter-trial interval (ISI in Fig. 1).⁸

One can then get a predicted BOLD response, B(t), for a region by convolving the demand functions, D(x), in Fig. 7 with a standard hemodynamic function H(t):

$$B(t) = M \int_0^t D(x)H(t-x)dx,$$

where M scales the magnitude of response in this region. While this predicts a continuous BOLD function, it is measured only once a trial and so we calculate for each trial the value

 $^{^{7}}$ Fig. 7 presents demand functions corresponding to the predictions in Fig. 4. Just as the predictions in Fig. 4 are averaged over a number of conditions, the demand functions in Fig. 7 are averaged over a number of conditions.

⁸ We could be modeling just a standard poststimulus undershoot (e.g., Mandeville et al., 1999) but some aspects of the negativity in this experiment seem different from what is normally observed. In particular, the timing of the negative response seems the same in all regions and within a region the magnitude of the negative response does not appear to vary with the magnitude of the positive response.



Fig. 7. Model-based demand functions for the eight predefined regions examined in Fig. 4.

of this function at the midpoint of the trial. Thus, for instance, the predicted magnitude of the BOLD response for the first scan in the visual region will be

$$B(.75) = M \int_0^{.75} \frac{1}{6} H(.75 - x) dx$$

since the value of demand function is 1/6 and there are .75 s since the beginning of the trial to the midpoint. Similarly, the predicted magnitude for the second scan input will be

$$B(2.25) = M\left[\int_0^{1.5} \frac{1}{6}H(2.25 - x)dx + \int_{1.5}^{2.25} \frac{1}{2}H(2.25 - x)dx\right]$$

since the demand function takes the value 1/2 from 1.5 to 3.0 s.

In our past research, we have used a standard gamma function for the hemodynamic function, as is the custom (e.g., Boyton, Engel, Glover, & Heeger, 1996; Cohen, 1997; Dale & Buckner, 1997; Friston et al., 1998; Glover, 1999):

$$H(t) = \left(\frac{t}{s}\right)^b \mathrm{e}^{-t/s},$$

where s is the time scale and b is exponent. However, the characterization of the BOLD response is conceptually cleaner if the simple gamma function is replaced with the gamma function from statistics, which has the convenient property that it always integrates to 1:

$$H(t) = \frac{t^{a-1} \mathrm{e}^{-t/s}}{s^a \Gamma(a)},$$

where Γ is the gamma function (the factorial is a special case: for integer a, $\Gamma(a + 1) = a!$)

The mode of the statistical gamma function is (a - 1) * s, the mean is a * s, and its variance is $a * s^2$. One can convert from the first formulation to the second by setting b = a - 1 and changing the magnitude scale. However, this formulation makes magnitude values, M, more comparable across regions that might vary in their exponent and time scale.

All the regions except the auditory were fit using a constant value of 4 for the exponent a and a constant value of 1.54 s for s. For the auditory region a = 8 and s = .634. Slightly better fits are possible with non-integer values of a, but we preferred to have an integer value. The magnitude, M, parameters are: Visual 2.38%; Aural 2.22%; Motor1 3.16%; Motor2 2.22%; Parietal 1.79%; Prefrontal 1.15%; Cingulate 2.90%; and Caudate 0.56%.

Part (a) of Fig. 7 illustrates the demand function for the visual area. Note that the visual module responds to a demand for active processing of input and does not generate a demand just because visual information is on the screen. This distinguishes the fusiform area from the primary visual cortex, which presumably simply responds to visual input. In general, the modules in ACT-R only create demand when engaged in active processing.

Part (b) of Fig. 7 illustrates the demand function for the aural area. It is involved when encoding or recoding an auditory stimulus and when outputting vocally—in the later case reflecting hearing ones own speech. The aural region was the one area that could not be fit with the same *a* and *s* parameters as the other regions. For this region a = 8 and s = .634. This results in a hemodynamic function, H(t) with almost identical times to peak (4.44 versus 4.60 s for the other regions) but a substantially reduced standard deviation (1.79 versus 3.07 s). This difference is apparent empirically in the much steeper BOLD responses obtained in this region. One reason for this is the fact that the there is little variability in time at which auditory input is processed. Presentation of auditory input is under experiment control, not participant control. On the other hand, the timing of the non-perceptual modules will vary from trial to trail around the means in Fig. 6. While the timing of the visual input is externally determined, participants have more freedom in when they will attend to it. The greater variability in the timing of events maps onto greater variability

210

ity of the hemodynamic function. In a study, where we collected hemodynamic responses associated with fairly tightly timed button presses (Kao et al., In preparation) we found a hemodynamic function for the motorl region very much like the one estimated here for the auditory region. The values in the Kao et al., paper were a = 7.8 and s = .8. The parameters from this experiment are very similar to the parameters estimated in Anderson (2005) that fit similarly variable times—that paper estimated a = 4 and had estimates of s about 1.5.

Parts (c) and (d) of Fig. 7 show the demand functions for the motor and vocal areas. These regions are close by and so we estimated .36 of the hemodynamic demand felt by the manual area was felt by the vocal area and vice versa. In addition, there is still one key press in the vocal condition contributing to the demand function for motor1 manual area.

Parts (e) and (f) of Fig. 7 show the demand functions for the parietal and prefrontal areas. Both are involved in rehearsal as indicated by the significant effect of delay. We have generalized this to involve these regions wherever there is recoding, rehearsal, or comparison in Fig. 6. They differ, however, in the fact that the parietal is involved in transformation and the prefrontal in substitution. Thus, the predicted BOLD responses in parts (e) and (f) of Fig. 7 are similar except that the parietal shows an effect of transformation and the prefrontal an effect of substitution.

Parts (g) and (h) of Fig. 7 show the demand functions for the cingulate and caudate areas. While the ACT-R model expects an effect of both substitution and transformation for these regions, Fig. 7 breaks the data out as a function of substitution because this was the factor that showed the larger effects in the data. Fig. 6 does not explicitly show the cingulate tracking changes in control states or the caudate tracking production rules, but these activities can be calculated from properties of that figure. The cingulate is engaged whenever there is a change of the control state to enable the next major stage of the task-encoding, bridging delay, substituting, transforming, rehearsing, responding, and checking the answer. We estimated that each of these control changes produced .5 s (two of the quarter second units in Fig. 6) of activity in the cingulate. The caudate, tracking production rule firing, is just a finer grain version of the cingulate—a production fires whenever there is a change in Fig. 6 and we assume this occupies the caudate for a quarter of a second.⁹ Most activities in Fig. 6 just occupy a quarter second unit but the substitution and transformation units occupy more. The periods of substitution and transformation are periods of relatively little production firing and long waits until the retrieval module completes substitution or the imaginal module completes the transformation. A difference between the predictions for the cingulate and caudate is that we expect a relatively large effect of substitution and transformation on the cingulate because they require one more control state change among few such changes, but a relatively small effect on the caudate because substitution and transformation require just one more production firing among many.

Fig. 7 specifies our assumptions about the role of each of these regions in the performance of the task. To review these assumptions are:

 $^{^{9}}$ In ACT-R the actual time per production firing is .05 s but in a more detailed ACT-R model might have multiple productions firing in a quarter second interval. On the other hand, the large estimate of a production firing time might explain why our magnitude estimate (0.56%) is the lowest for the caudate.

- (1) The perceptual areas respond strongly to the perceptual demands of the task but not the cognitive or motor demands of the task.
- (2) The motor areas respond strongly to the motor demands of the task and are involved in rehearsal but not the perceptual or other cognitive demands of the task.
- (3) The parietal area responds to the need to manipulate problem representations and the prefrontal region to the need to perform memory operations. Neither responds to the perceptual and motor demands. The regions are also involved in rehearsing the material.
- (4) The cingulate and caudate are only involved in the cognitive demands of the task and do not respond in any way to the perceptual or motor demands.

The previous statistical analyses supported these conclusions in terms of which factors had effects and which did not (e.g., see Fig. 3). This modeling enterprise tests how well a theory embodying these assumptions can come up with a rigorous account of the BOLD response in these regions. This raises the question of how to measure the match between the data and the fits based on the theory. Fig. 4 reported the correlations between the data and the theoretical fits and in all cases the correlations are high, confirming what is apparent from visual inspection. However, the match is not perfect and there is the question of whether to attribute the differences to random noise or systematic deviation between data and theory. Anderson et al. (2003) suggested using chi-square statistics, which measure the degree of mismatch against the noise in the data. They are calculated as

$$\chi^2 = \frac{\sum_i (\hat{X}_i - \overline{X}_i)^2}{S_{\overline{X}}^2},$$

where the numerator is the squared deviation between prediction and data and the denominator is estimated from the interaction between conditions and participants. We calculated this for the 8 fits in Fig. 4. In this case, the statistic can be treated as having 75 degrees of freedom, calculated as the 78 data points minus the three parameters estimated for the BOLD function.¹⁰ As Anderson (2005) discusses, the chi-square statistic is not a perfect measure of fit but it does serve to alert us to when we should be taking deviations seriously. With 75 degrees of freedom a χ^2 value greater than 96 is significant. Four of our regions show significant deviations by this measure:¹¹

Motor1: The chi-square is 125.8. The principle problem is that it predicts the manual conditions will peak a little sooner than they do while it predicts the vocal conditions will peak a little later than they do.

Parietal. The χ^2 is 131.9. The principle problem is that the model over-predicts the effect of transformations.

Cingulate. The χ^2 is 119.9. The principle problem is that the model predicts a difference between the delay and no-delay earleir than it appears in the data.

¹⁰ Actually, this might be viewed as having more degrees of freedom because the fits are constrained to share the same a and s parameters except for the aural region.

¹¹ The χ^2 for the other regions are 85.9 for the visual, 58.1 for the aural, 73.3 for the vocal, and 45.0 for the prefrontal.

Caudate. The χ^2 is 243.8. The principle problem is that the model predicts the response will go down to baseline more slowly than observed.

It would be possible to propose more elaborate models that "fixed" these problems, but the assumptions of such models would be ad hoc, just motivated by the deviations. In our view, the overall match up between theory and data is compelling, but these statistical tests serve to alert us to the places where something else is happening. These points of deviation may prove to be significant if they can be given motivated theoretical interpretations.

4. Conclusions

With respect to the issue of the modality-specificity of cognition, this research has yielded an answer more articulated than simply that all of central cognition is based on perceptual-motor processes or that all of central cognition is abstracted away from the conditions of input and output. There seemed four levels of involvement of brain regions in peripheral processes versus central at least in these experiments. The perceptual regions (auditory cortex, fusiform gyrus) seemed only to respond to perceptual factors. The motor regions seemed to be involved in rehearsal as well as output. Two of the cortical regions (left parietal, left prefrontal) seemed to be involved in central functions but their lateralization suggested that they were seemed tied to the verbal nature of the material. The final two regions (caudate, cingulate) seemed to be completely abstracted away from conditions of input, output, or verbal content and serve simply general control functions. While these eight regions hardly exhaust the brain, they proved to be a sufficient number to establish all possibilities occur on the embodied-versus-abstract continuum.

Much of the research emphasizing perceptual-motor aspects of cognition has been concerned with the nature of concepts, particularly in language. Our experimental task is quite artificial and has no obvious connection with the conceptual structure of language. Elsewhere (Anderson et al., 2003) we have developed the connections between this task and the symbolic manipulations that are part of high-school algebra, but algebra is itself an abstract domain. Part of the difficulty of algebra for children is the further abstraction it requires from their previous mathematical activities. So, the research in this current paper does not really challenge the conclusions from much of the research emphasizing the perceptual-motor character of thought. However, it suggests there are limitations to these conclusions as we go to some more abstract human intellectual activities.

This research adds to the growing consensus about the function of these eight brain regions. In the case of the perceptual and motor areas, their associated perceptual and motor functions are hardly news. However, it is still newsworthy to offer further evidence that the parietal seems to play a representational function, the prefrontal a retrieval function, the cingulate a control function, and the caudate a procedural function. It is also newsworthy that the perceptual and motor areas play relatively little role in the cognitive aspects of this task.

Some qualifications need to be made to make it clear that we are not proposing a oneto-one mapping between these eight regions and the eight functions. First, other regions also serve these functions. Many areas are involved in vision and the fusiform gyrus has just proven to be the most useful to monitor. Similarly, many regions have been shown to be involved in retrieval, particularly the hippocampus. The prefrontal region is just the easiest to identify and seems to afford the best signal-to-noise ratio. Equally, we are not claiming these regions only serve one function. This paper has found some evidence for multiple functions. For instance, the motor regions are involved in rehearsal as well as external action.

There is evidence that the anterior cingulate is involved in other functions besides abstract control such as error detection (Falkenstein, Hohnbein, & Hoorman, 1995; Dehaene, Posner, & Tucker, 1994) and monitoring of response conflict (Carter et al., 2000; MacDonald, Cohen, Stenger, & Carter, 2000). However, the anterior cingulate is a relatively large region. Van Veen and Carter (2005) found evidence that response conflict was handled by a more anterior and ventral region than a region which responded to what they called semantic conflict. Their semantic conflict is not unlike our need to shift control from one line of processing to another. Interestingly, the center of their semantic conflict region is quite close to our anterior cingulate region (indeed our region is slightly further posterior and further away from their response conflict region). Thus, it seems that, while ACC does respond to response conflict, it is also involved in control more generally (see also Badre & Wagner, 2004) as the ACT-R theory proposes.

This research has also shown that an information-processing model can be used to make predictions about the BOLD responses in these regions. We have stepped back a bit from the details of an ACT-R model (unlike Anderson, 2005) in part to make it somewhat clearer how to map a generic information-processing model onto the BOLD response. FMRI data are often criticized because of their poor temporal resolution. However, when one begins to look at relatively complex tasks such as the one in this paper, the task unfolds over a sufficiently long period that one can begin to perceive some of the temporal structure of the task in the BOLD response. This is facilitated if one can introduce delays into the task performance as we did here. Another virtue of the relatively complex tasks like the current is that it tends to drive relatively strong responses, usually yielding good signal-to-noise ratios in the BOLD responses.

Acknowledgments

We thank Larry Barsalou and Jennfer Ferris for their comments on this paper and Jennifer Ferris and Yvonne Kao for their help in the preparation of Fig. 5. This research was supported by NIMH award MH068243 to Anderson, NSF ROLE Grant REC-0087396 to Anderson and Carter, and NIMH career award KO2 MH01306 to Carter.

References

- Acuna, B. D., Eliassen, J. C., Donoghue, J. P., & Sanes, J. N. (2002). Frontal and parietal lobe activation during transitive inference in humans. *Cerebral Cortex*, 12, 1312–1321.
- Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. Cognitive Science, 29, 313–342.
- Anderson, J. R., Albert, M. V., & Fincham, J. M. (2005). Tracing problem solving in real time: fMRI analysis of the subject-paced tower of hanoi. *Journal of Cognitive Neuroscience*, 17, 1261–1274.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review*, 111, 1036–1060.
- Anderson, J. R., & Bower, G. H. (1973). Human associative memory. Washington, DC: Winston.
- Anderson, J. R., Qin, Y., Sohn, M-H., Stenger, V. A., & Carter, C. S. (2003). An information-processing model of the BOLD response in symbol manipulation tasks. *Psychonomic Bulletin & Review*, 10, 241–261.
- Anderson, J. R., Qin, Y., Stenger, V. A., & Carter, C. S. (2004). The relationship of three cortical regions to an information-processing model. *Journal of Cognitive Neuroscience*, 16, 637–653.

- Ashby, F. G., & Waldron, E. M. (2000). The neuropsychological bases of category learning. Current Directions in Psychological Science, 9, 10–14.
- Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, 41, 473–487.
- Barsalou, L. W. (1999). Perceptual symbol systems. Behavioral and Brain Sciences, 22, 577-609.
- Barsalou, L. W. (2005). Abstraction as dynamic interpretation in perceptual symbol systems. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories* (pp. 389–431). Carnegie Symposium Series. Majwah, NJ: Erlbaum.
- Barsalou, L. W., Simmons, W. K., Barbey, A., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7, 84–91.
- Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T. E., Ollinger, J. M., & Snyder, A. Z. (2001). Anterior cingulate and response conflict: effects of response modality and processing domain. *Cerebral Cortex*, 11, 837–848.
- Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T., Malow, B. A., Gaillard, W. D., Sato, S., et al. (1997). A direct comparison of PET activation and electrocortical stimulation mapping for language localization. *Neurology*, 48, 1056–1065.
- Boyton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. Journal of Neuroscience, 16, 4207–4221.
- Buckner, R. L., Kelley, W. M., & Petersen, S. E. (1999). Frontal cortex contributes to human memory formation. *Nature Neuroscience*, 2, 311–314.
- Buckner, R. L., Koutstaal, W., Schacter, D. L., & Rosen, B. R. (2000). fMRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain*, 123, 620–640.
- Buckner, R. L., & Wheeler, M. E. (2001). The cognitive neuroscience of remembering. Nature Reviews Neuroscience, 2, 624–634.
- Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage*, 16, 317–330.
- Carpentier, A. C., Pugh, K. R., Westerveld, M., Studholme, C., Skrinjar, O., Thompson, J. L., et al. (2001). Functional MRI of language processing: dependence on input modality and temporal lobe epilepsy. *Epilepsia*, 42, 1241–1254.
- Carter, C. S., MacDonald, A. M., Botvinick Ross, L. L., Stenger, V. A., Noll, D., et al. (2000). Parsing executive processes: strategic versus evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 1944–1948.
- Cohen, M. S. (1997). Parametric analysis of fMRI data using linear systems methods. *NeuroImage*, 6, 93-103.
- Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., et al. (2004). Sentence complexity and input modality effects in sentence comprehension: an fMRI study. *Neuroimage*, 22, 11–21.
- Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. Human Brain Mapping, 5, 329–340.
- Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. Current Opinion Neurobiology, 14, 218–224.
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2002). Three parietal circuits for number processing. Cognitive Neuropsychology, 20, 487–506.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, 5, 303–305.
- D'Esposito, M., Piazza, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive of working memory. *Nature*, 378, 279–281.
- Falkenstein, M., Hohnbein, J., & Hoorman, J. (1995). Event related potential correlates of errors in reaction tasks. In G. Karmos, M. Molnar, V. Csepe, I. Czigler, & J. E. Desmedt (Eds.), *Perspectives of event-related potentials research* (pp. 287–296). Amsterdam: Elevier Science B.V..
- Fletcher, P. C., & Henson, R. N. A. (2001). Frontal lobes and human memory: insights from functional neuroimaging. *Brain*, 124, 849–881.
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance Imaging–Method*, 33, 636–647.
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: characterizing differential responses. *NeuroImage*, 7, 30–40.

- Gabrieli, J. D. E. (2001). Functional neuroimaging of episodic memory. In R. Cabeza & A. Lingstone (Eds.), Handbook of functional neuroimaging of cognition (pp. 253–292). Cambridge, MA: MIT Press.
- Glenberg, A. M. (1997). What memory is for. Behavioral and Brain Sciences, 20, 1-55.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. Psychonomic Bulletin & Review, 9, 558–565.
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage, 9, 416-429.
- Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, 12, 1–19.
- Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, 7, 555–562.
- Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. Nature Review Neuroscience, 2, 685–694.
- Hikosaka, O., Nakahara, H., Rand, M. K., Sakai, K., Lu, Z., Nakamura, K., et al. (1999). Parallel neural networks for learning sequential procedures. *Trends in Neuroscience*, 22, 464–471.
- Jeannerod, M. (1995). Mental imagery in the motor context. Neuropsychologia, 33, 1419–1432.
- Jonides, J., Lacey, S. C., & Nee, D. E. (2005). Processes of working memory in mind and brain. Current Directions in Psychological Science, 14, 1–5.
- Jung, K-J., Prasad, P., Qin, Y., & Anderson, J. R. (2005). Extraction of overt verbal response from acoustic noise from the scanner in fMRI by use of segmented active noise cancellation. *Magnetic Resonance in Medicine*, 53, 739–744.
- Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, NJ: Erlbaum.
- Kao, Y., Fincham, J., Jung, K-J, Qin, Y., Anderson, J.R. (In preparation). An information-processing approach to deconstructing a geometry proof task.
- Le Clec'H, G., Dehaene, S., Cohen, L., Mehler, J., Dupoux, E., Poline, J. B., et al. (2000). Distinct cortical areas for names of numbers and body parts independent of language and input modality. *NeuroImage*, 12, 381–391.
- Lepage, M., Ghaffar, O., Nyberg, L., & Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. Proceedings of National Academy of Sciences Unite States of America, 97, 506–511.
- MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288, 1835–1838.
- Mandeville, J. B., Marota, J. J., Ayata, C., Moskowitz, M. A., Weisskoff, R. M., & Rosen, B. R. (1999). MRI measurement of the temporal evolution of relative CMRO(2) during rat forepaw stimulation. *Magnetic Resonance in Medicine*, 42, 944–951.
- Marshuetz, C., Smith, E. E., Jonides, J., DeGutis, J., & Chenevert, T. L. (2000). Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *Journal of Cognitive Neuroscience*, 12, 130–144.
- Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance. Part 1. Basic mechanisms. *Psychological Review*, 104, 2–65.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7, 293–299.
- Michael, E. B., Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). fMRI investigation of sentence comprehension by eye and ear: modality fingerprints on cognitive processes. *Human Brain Mapping.*, 13, 239–252.
- Newman, S. D., Carpenter, P. A., Varma, S., & Just, M. A. (2003). Frontal and parietal participation in problemsolving in the tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia*, 41, 1668–1682.
- Norman, D. A., & Rumelhart, D. E. (1975). Explorations in cognition. New York: W.H. Freeman.
- Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart, and Winston.
- Paivio, A. (1986). Mental representations: A dual coding approach. New York: Oxford University Press.
- Picard, N., & Strick, P. L. (1996). Medial wall motor areas: a review of their location and functional activation. *Cerebral Cortex*, 6, 342–353.
- Poldrack, R. A., Prabakharan, V., Seger, C., & Gabrieli, J. D. E. (1999). Striatal activation during cognitive skill learning. *Neuropsychology*, 13, 564–574.
- Posner, M. I., & Dehaene, S. (1994). Attentional networks. Trends in Neurosciences, 17, 75-79.
- Qin, Y., Anderson, J. R., Silk, E., Stenger, V. A., & Carter, C. S. (2004). The change of the brain activation patterns along with the children's practice in algebra equation solving. *Proceedings of National Academy of Sciences United States of America*, 101, 5686–5691.

- Qin, Y., Sohn, M.-H., Anderson, J. R., Stenger, V. A., Fissell, K., Goode, A., et al. (2003). Predicting the practice effects on the blood oxygenation level-dependent (BOLD) function of fMRI in a symbolic manipulation task. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 4951–4956.
- Reichle, E. D., Carpenter, P. A., & Just, M. A. (2000). The neural basis of strategy and skill in sentence-picture verification. *Cognitive Psychology*, 40, 261–295.
- Robertson, L. C., & Rafal, R. (2000). Disorders of visual attention. In M. Gazzaniga (Ed.), The new cognitive neuroscience (2nd ed., pp. 633–650). Cambridge: MIT Press.
- Saint-Cyr, J. A., Taylor, A. E., & Lang, A. E. (1988). Procedural learning and neostriatal dysfunction in man. *Brain*, 111, 941–959.
- Sanes, J. N., & Schieber, M. H. (2001). Orderly somatotopy in primary motor cortex: Does it exist? *NeuroImage*, 13, 968–974.
- Schumacher, E. H., Lauber, E., Awh, E., Jonides, J., Smith, E. E., & Koeppe, R. A. (1996). PET evidence for an amodal verbal working memory system. *NeuroImage*, 3, 79–88.
- Sohn, M.-H., Goode, A., Stenger, V. A., Carter, C. S., & Anderson, J. R. (2003). Competition and representation during memory retrieval: roles of the prefrontal cortex and the posterior parietal cortex. *Proceedings of National Academy of Sciences United States of America*, 100, 7412–7417.
- Sohn, M.-H., Goode, A., Stenger, V. A., Jung, K.-J., Carter, C. S., & Anderson, J. R. (2005). An informationprocessing model of three cortical regions: evidence in episodic memory retrieval. *NeuroImage*, 25, 21–33.
- Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press.
- Van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: a functional MRI study. *NeuroImage*, 27, 497–504.
- Wagner, A. D., Maril, A., Bjork, R. A., & Schacter, D. L. (2001). Prefrontal contributions to executive control: fMRI evidence for functional distinctions within lateral prefrontal cortex. *NeuroImage*, 14, 1337–1347.
- Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31, 329–338.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281, 1188–1191.
- Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., & Mazziotta, J. C. (1998). Automated image registration: I. General methods and intraparticipant, intramodality validation. *Journal of Computer Assisted Tomography*, 22, 139.
- Yücel, M., Stuart, G. W., Maruff, P., et al. (2001). Hemispheric and gender-related differences in gross morphology of the anterior cingulate/paracingulate cortex in normal volunteers: an IRM morphometric study. *Cerebral Cortex*, 11, 17–25.