

available at [www.sciencedirect.com](http://www.sciencedirect.com)[www.elsevier.com/locate/brainres](http://www.elsevier.com/locate/brainres)


---



---

**BRAIN  
RESEARCH**


---



---



---

**Research Report**

# Errors of mathematical processing: The relationship of accuracy to neural regions associated with retrieval or representation of the problem state

Susan M. Ravizza<sup>a,\*</sup>, John R. Anderson<sup>b</sup>, Cameron S. Carter<sup>c</sup>

<sup>a</sup>Department of Psychology, 285C Psychology Building, Michigan State University, East Lansing, MI 48824, USA

<sup>b</sup>Departments Computer Science and Psychology, Carnegie Mellon University, USA

<sup>c</sup>Imaging Research Center and Departments of Psychiatry and Behavioral Sciences and Psychology, University of California, Davis, USA

---

**ARTICLE INFO**
**Article history:**

Accepted 8 August 2008

Available online 23 August 2008

**Keywords:**

Prefrontal cortex

Parietal cortex

Problem-solving

fMRI

Anterior cingulate

**ABSTRACT**

Regions in the prefrontal and parietal cortices contribute to mathematical problem-solving through their roles in retrieval and mental representation, respectively. This fMRI study examined whether activity in these regions tracked with subsequent errors in solving algebraic equations. Whereas previous studies have used recognition paradigms (e.g., decide whether  $2+2=5$  is correct) to assess the relationship of neural functioning with performance, participants in this study were required to generate an answer themselves. For the prefrontal region that in previous studies has exhibited activity modulated by retrieval demands, we found that activity was greater when equations were solved correctly than when errors were committed. Good solvers also tended to exhibit more activity in this region than poor problem-solvers. This was not true for the region in the parietal cortex that has been associated with representing the number of transformations to the equation. This suggests that, in our adult sample, successful performance was related to retrieval abilities rather than to difficulty in representing or updating changes in the equation as it is being solved.

© 2008 Elsevier B.V. All rights reserved.

---



---

**1. Introduction**

As any math teacher will attest, students produce frequent errors when solving mathematical problems. Indeed, the number and types of errors can be an important source of information for teachers in determining where students need further instruction. A number of behavioral studies have closely examined the types of errors that occur when solving equations, and these data have been used to generate and test cognitive models of component processes in mathematical problem-solving (Anderson et al., 1996; Koedinger and Nathan,

2004). Recently, the field has expanded to include neuroimaging studies of mathematical processing in order to confirm that these components map onto separable neural regions (Anderson et al., 2003, 2004a,b; Danker and Anderson, 2007; Qin et al., 2004). However, very few studies have determined whether neural regions associated with mathematical processing are important for successful performance. The dearth of studies on the neural underpinnings of mathematical errors is surprising given that mathematics is one of the few subjects where the distinction between right and wrong answers is relatively clear-cut. The goal of this study is to assess whether

\* Corresponding author. Fax: +1 517 353 1652.

E-mail address: [ravizzas@msu.edu](mailto:ravizzas@msu.edu) (S.M. Ravizza).

errors in mathematical performance are associated with activity in neural regions known to instantiate processes required for problem-solving.

A unique aspect of this study is its focus on the active generation of solutions to algebraic equations rather than a more passive evaluation process. To date, recognition paradigms have primarily been used to assess neural contributions to successful performance in mathematical tasks (Menon et al., 2002; Rivera et al., 2002). Rather than generating answers, participants are given an equation such as  $2+2=5$  and asked to determine its correctness. Greater activity in the prefrontal cortex has been observed when judging incorrect than correct equations (Menon et al., 2002), however, it is difficult to determine whether these regions are critical for successfully solving equations or whether they reflect a more general process needed in evaluating incorrect solutions by others (e.g., sustained selective attention or novelty detection). By asking participants to generate answers, this study is specifically designed to assess the neural underpinnings of successful performance.

While imaging studies of mathematical errors are rare, much headway has been made in verifying that component processes of mathematical problem-solving map onto separable neural regions. One model, the ACT-R information-processing model (Anderson et al., 2004a,b; Anderson, 2007), posits several modules necessary for tasks such as mathematical reasoning, and these modules have been shown to have distinct neural correlates (see Anderson, 2005 for a review). Three of these modules have been extensively studied in behavioral and neuroimaging studies of abstract symbol manipulation (Anderson et al., 2003, 2004a,b) and algebraic equation solving (Anderson et al., 2003; Qin et al., 2004), and will also be the focus of this experiment.

One module in the ACT-R cognitive architecture corresponds to an imaginal buffer that maintains the current mental representation of a problem state such as an equation. Neuroimaging studies have demonstrated that activity of the inferior parietal cortex displays a parametric increase when the number of changes to the problem state increases (Anderson et al., 2003; Qin et al., 2004); that is, when the number of steps needed to solve an equation increases (thereby increasing the number of transformations to the problem state), the inferior parietal cortex is recruited to a greater degree. The lateral prefrontal cortex, in contrast, does not respond to the number of transformations but, instead, is engaged in conditions where retrieval demands are high. This region is thought to correspond to a buffer that holds cues that drive the retrieval process. For instance, it would hold the retrieval cues “7+5” in retrieving the arithmetic fact  $7+5=12$  in service of solving the equation  $x-5=7$ . Activity in the lateral prefrontal cortex increases as the time to retrieve an item increases or the number of items to be retrieved increases. A third module is responsible for motor output and is reflected in activity of the motor cortex. This region displays peak activity at the time of responding.

Given that a number of neuroimaging studies have now provided converging evidence for the viability of these buffers and their neural instantiations, our current goal is to link the computations provided by these regions with success in solving algebra equations. One prediction might be that both

retrieval efficiency and visuospatial representational abilities would be associated with greater accuracy. If so, regions undertaking those processes should be engaged to a greater degree when equations are solved correctly than when errors are committed. Moreover, individuals with higher accuracy rates should also display more activity in those regions. To foreshadow our results, we found that activity of the lateral prefrontal cortex was greater during correct trials and tended to be greater for better-performing participants. In contrast, inferior parietal cortex activity was not modulated by accuracy. Thus, our results imply that errors are linked more to inefficient retrieval than to difficulty in representing the problem state in our adult sample.

In addition to examining the effects of accuracy, we also assessed the complexity of mental representation by varying the size of the operands. Several studies have found activity in a region of the inferior parietal cortex to be modulated by the demands of mental representation as assessed by varying the number of transformations. However, the visuospatial requirements in representing large numbers may also tax the imaginal buffer. Other studies have found greater parietal activity for larger numbers than smaller numbers (Stanesco-Cosson et al., 2000; Piazza et al., 2004), however, it is unclear whether the same region of the parietal cortex is responsive to both the number of transformations of the problem state and number size. By using a different manipulation to vary the difficulty of mental representation, it is possible to verify that this region corresponds to an imaginal buffer rather than being important for a process specific to visuospatial transformations.

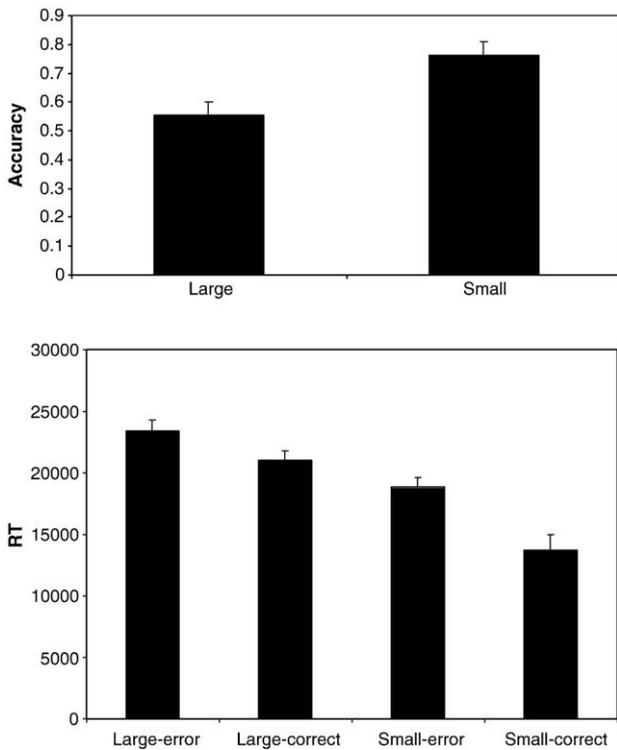
While parietal, prefrontal, and motor contributions to mathematical problem-solving have been studied intensively, less well-characterized is the role of the anterior cingulate (Anderson, 2005). In the ACT-R model, the anterior cingulate is hypothesized to contribute to the control of goal-oriented behavior. For example, there is much evidence supporting anterior cingulate involvement in the detection of response conflict and the ability to overcome prepotent responses (MacDonald, et al., 2000; vanVeen and Carter, 2002; Kerns et al., 2004). Less support has been acquired for a role of this region in more stimulus-based conflict (Van Veen and Carter, 2005). The slow-event related design in our experiment will allow us to assess whether the anterior cingulate is recruited when participants are mentally solving equations and are not making overt responses.

---

## 2. Results

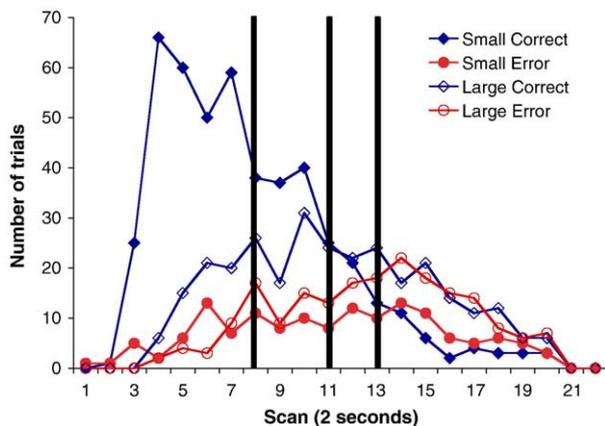
### 2.1. Behavioral

Null and incomplete responses were discarded, so only errors of commission were analyzed. The data can be classified into four categories according to whether the problem was large or small and whether the problem was answered correctly or not (Fig. 1). Averaging accuracies over subjects, they were 76.3% accurate for small problems and 55.5% for large problems. This was a highly significant difference ( $t(14)=7.81$ ,  $p<.001$ ). With respect to latencies, the averages of the subject means are 13.75 s. for small correct, 18.89 s. for small error, 21.05 s. for large correct, and 23.4 s. for large error. An ANOVA performed



**Fig. 1 – Accuracy and reaction times of participants solving algebraic equations. Reaction time corresponds to the moment that participants pressed the key indicating that they knew the answer to the equation.**

on these latencies confirmed that the effects of size and correctness are both highly significant ( $F(1,14) = 38.56, p < .0001$ , for accuracy;  $F(1,14) = 57.41, p < .0001$ , for size) and the interaction between the two approached, but did not reach, significance ( $F(1,14) = 4.2, p = .06$ ). Fig. 2 shows the distribution of timing of responses summed across participants in these four categories in terms of which scan the response was emitted. Note these means are measured from 1 s into the



**Fig. 2 – The number of responses given during each scan in each condition. Dark vertical lines denote the scan associated with the mean RT in each condition (i.e., scan 8 for small correct, 11 for small error and large correct, and 13 for large error).**

scanning sequence because the warning prompt occupies the first second. Therefore, the scans associated with these mean times are 8 for small correct, 11 for small error and large correct, and 13 for large error. These scans are indicated by vertical lines in Fig. 2.

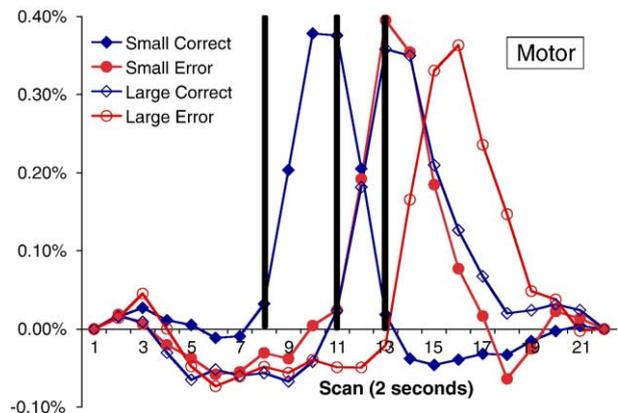
**2.2. Imaging results**

The wide distribution of times in Fig. 2 poses a challenge for aggregating the imaging data in a way to provide meaningful analysis. To deal with this, we developed a method (see Anderson and Qin, 2008) for warping data from an individual trial to the mean length of the response across trials. First, for each trial and each region of interest we calculated the percent change of the (Blood Oxygen Level Dependent) BOLD response on each scan with reference to the first scan on a trial. We then broke these change scores into two intervals. One interval was from the first scan until the scan of the response and the other from the first scan after the response until the last scan of the trial. We then warped these two intervals onto the mean lengths for these intervals in that condition. The following is the warping procedure for taking a scan sequence of length  $n$  and deriving a scan sequence of the mean length  $m$ . It depends on the relative sizes of  $m$  and  $n$ :

1. If  $n$  is greater than or equal to  $m$ , create a sequence of length  $m$  by taking  $m/2$  scans from the beginning and  $m/2$  from the end. If  $m$  is odd select one more from the beginning. This means just deleting the  $n-m$  scans in the middle.
2. If  $n$  is less than  $m$ , create a beginning sequence of length  $m/2$  by taking the first  $n/2$  scans and padding with the last scan in this first  $n/2$ . Construct the end similarly. If either  $n$  or  $m$  is odd, the extra scan is from the beginning.

This creates scan sequences that preserve the temporal structure of the beginning and end of the sequences and just represent the approximate average activity in their middle.

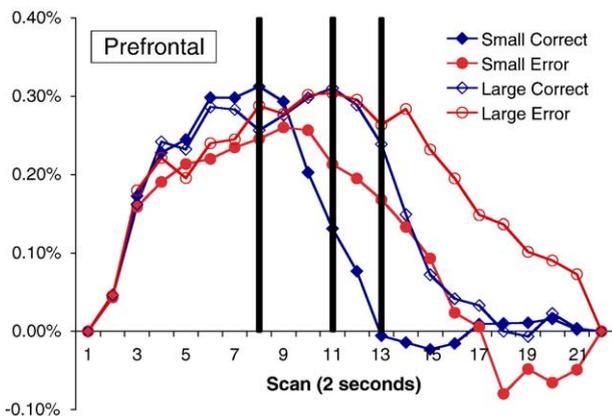
Fig. 3 shows the results we obtained for the predefined motor region. It shows nicely staggered motor responses



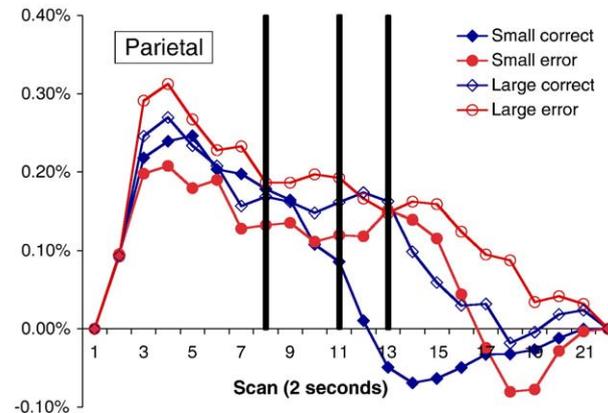
**Fig. 3 – Time course of activity in the left motor region of interest in each condition. Dark vertical lines denote the scan associated with the mean RT in each condition (i.e., scan 8 for small correct, 11 for small error and large correct, and 13 for large error).**

beginning their rise with the scan of the response. As with all of our predefined regions, we performed a separate analysis of the area under the curve for the first 7 scans and the area under the curve for the remaining 15 scans (i.e. the sum of the first 7 points and the sum of the last 15 points). The first seven scans reflect that portion of the imaging data that has not yet been influenced by response generation in any condition. In the case of the motor area there are no significant effects here (size:  $F(1, 14)=.6$ , correctness:  $F(1,14)=1.23$ , interaction  $F(1,14)=1.44$ ,  $ps>.1$ ). Moreover, the overall area is not significantly different than zero ( $t(14)=-0.89$ ,  $p>.1$ ). This confirms what is apparent to the eye, which is that this region is basically dormant in the early interval. On the other hand, there is a highly significant rise after this point ( $t(14)=4.44$  for area under the curve,  $p<.001$ ) but the size of this rise also does not vary with condition (size:  $F(1, 14)=.89$ , correctness:  $F(1,14)=.01$ , interaction  $F(1,14)=1.11$ ,  $ps>.1$ ). These are more or less the results we would expect from the motor region. The fact that we obtained them indicates that there is no confounding of motor difficulty with condition and that our warping procedure does succeed in bringing out the structure of the BOLD response over such variable latencies.

Fig. 4 shows the results we obtain for the predefined prefrontal region. As a general qualitative statement, the results in these areas quickly rise and reach an asymptotic level about scan 7 and drop off with the emission of the response. This is the pattern we would expect if this region were principally responsible for retrieval of arithmetic facts. Unlike the results for the motor cortex, there are significant differences in areas under the curve for the first 7 scans and for later scans. For the interval before response generation, there is significantly greater area under the curve in case of correct compared to erroneous responses ( $F(1,14)=5.25$ ,  $p<.05$ ). This is an important result because it indicates that the activation in the prefrontal region is signaling error in advance of any response. The other effects were not significant in the first 7 scans (size:  $F(1,14)=.02$ , interaction  $F(1,14)=0.38$ ,  $ps>.1$ ). Both main effects were significant in the later scans (size:  $F(1, 14)=15.08$ ,  $p<.005$ , correctness:  $F(1,14)=10.00$ ,  $p<.01$ ) and there was no significant interaction ( $F(1,14)=1.78$ ,  $p>.1$ ). Both of the main effects for the later scans



**Fig. 4 – Time course of activity in the left prefrontal region of interest in each condition. Dark vertical lines denote the scan associated with the mean RT in each condition (i.e., scan 8 for small correct, 11 for small error and large correct, and 13 for large error).**

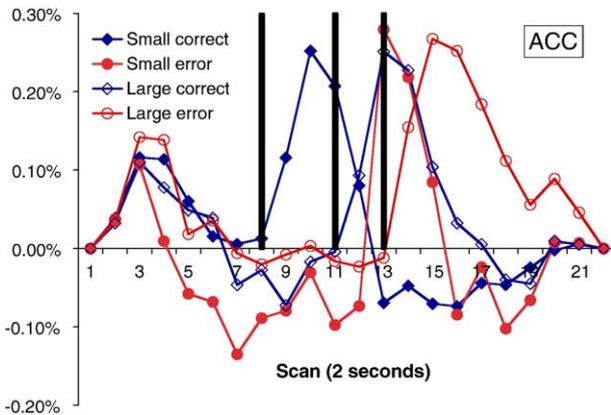


**Fig. 5 – Time course of activity in the left parietal region of interest in each condition. Dark vertical lines denote the scan associated with the mean RT in each condition (i.e., scan 8 for small correct, 11 for small error and large correct, and 13 for large error).**

reflect the fact that subjects take longer and are presumably doing more work when they make errors or face large problems. Note that area under the curve is greater for error trials in later scans, but smaller before a response is generated. It is as if the extra effort early avoids the need for extra work later. In terms of total area (summing all scans) there was a marginally significant tendency for greater area for errors ( $F(1,14)=3.58$ ,  $p<.1$ ).

Fig. 5 shows the results we obtained for the predefined parietal region. It shows a more rapid rise than the prefrontal area, reaching a peak about scan 4 and dropping off from scan 4 to scan 8 whereas the prefrontal region is continuing to rise over the interval. The rapid involvement of this region would be expected if it is involved in building up an internal representation of the problem. In terms of the first 7 scans there are no significant effects (size:  $F(1, 14)=1.75$ ,  $p>.1$ ; correctness:  $F(1,14)=.002$ ,  $p>.1$ ; interaction  $F(1,14)=4.19$ ,  $p=.06$ ). Moreover, these results suggest that this region of the parietal cortex is not responsive to number size in the early interval. Indeed, neither of the simple effects of number size were significant (large vs. small correct:  $t(14)=.05$ ,  $p=.96$ ; large vs. small error:  $t(14)=1.73$ ,  $p=.11$ ). However, for the interval after the response, both main effects were significant (size:  $F(1, 14)=8.66$ ,  $p<.05$ , correctness:  $F(1,14)=7.46$ ,  $p<.05$ ) and there was no significant interaction ( $F(1,14)=.024$ ,  $p>.1$ ). Activity in the late interval seems basically to reflect the duration of the trial. The parietal appears to behave as a region that is heavily involved in building up an initial problem representation and then continues to update that representation as long as the trial lasts.

Fig. 6 shows the results we obtain for the predefined anterior cingulate region. This region shows an early rise when the strategy is selected and a late rise with the emission of the response. In terms of area under the curve, the main effects are not significant during the first 7 scans (size:  $F(1, 14)=1.52$ , Correctness:  $F(1,14)=.82$ ,  $ps>.1$ ) but the interaction term is ( $F(1,14)=6.44$ ,  $p<.05$ ). The interaction is driven by the sharp drop for small errors. Given that this is the condition with the fewest observations per subject one might be suspicious of this result. In terms of late area there are no

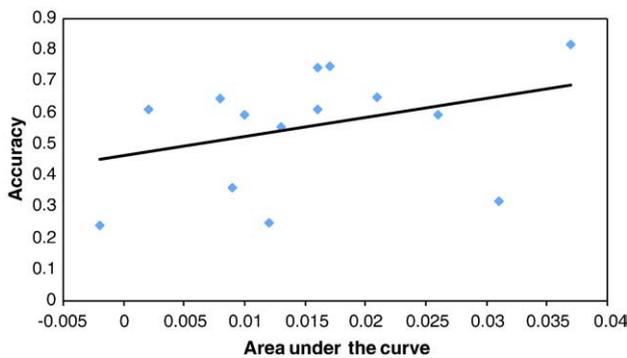


**Fig. 6 – Time course of activity in the anterior cingulate region of interest in each condition. Dark vertical lines denote the scan associated with the mean RT in each condition (i.e., scan 8 for small correct, 11 for small error and large correct, and 13 for large error).**

significant effects (size:  $F(1, 14)=2.78$ , correctness:  $F(1,14)=.64$ , interaction  $F(1,14)=1.89$ ,  $ps > .1$ ). It is noteworthy that in this experiment, unlike simpler experiments, there seems to be no connection between anterior cingulate activation and error.

**2.3. Correlations between performance and neural activity**

To assess whether our regions were directly involved in performance, we correlated area under the curve in the early stage with accuracy in the problem-solving stages across participants (see Fig. 7). Although a positive association was observed between activity in the left prefrontal cortex (correct trials only) and accuracy in the large-number condition ( $r(14)=.354$ ,  $p > .1$ ), it was not close to significance. However, this was mostly due to one outlier who had low accuracy despite greater activity in the prefrontal cortex. This outlier had the highest deviation score from the sample as a whole as indexed by the Mahalanobis distance score (5.64), and this participant also had the most influence on the correlation statistic according to a leverage analysis (.4). If this outlier is removed, the correlation is significant ( $r(13)=.603$ ,  $p < .05$ ). Accuracy was only weakly related to prefrontal activity in the small-number condition and was in the opposite direction ( $r(14)=-.173$ ,  $p > .1$ ).



**Fig. 7 – The relationship of performance and prefrontal activity for each participant.**

**Table 1 – Regions produced in the voxelwise analysis**

Contrast	BA	x	y	z	Large r	Small r
<i>Accuracy (Correct&gt;Error)</i>						
SMA	6	-7	-16	57	.19	.08
Left middle frontal gyrus	6/9	-56	6	40	.48*	.01
Left putamen/thalamus		-25	-2	4	.41	.00
Right lingual gyrus/vermis		11	-55	-3	.16	-.36
<i>Number size (Large&gt;Small)</i>						
Right middle temporal gyrus	21	63	-12	-11	.09	-.26
Right lingual gyrus/vermis		6	-66	-6	.00	-.28
<i>Number size (Small&gt;Large)</i>						
Left inferior frontal gyrus	45	-32	26	7	.23	-.34

\* Denotes significance at  $p < .05$  if outlier is excluded.

The parietal cortex and the anterior cingulate did not display a significant relationship between percent change in activity and accuracy.

**2.4. Voxel-wise analyses**

Several regions were produced in our whole-brain analysis of accuracy effects. Only regions where activity in one condition was above baseline are listed in Table 1. To determine whether these regions were associated with performance, we correlated area under the curve for the early interval (i.e. 0–10 s after stimulus onset). Of the regions showing greater activity for correct trials, activity in the left middle frontal gyrus tended to be related to performance ( $r(14)=.48$ ,  $p=.067$ ; if same outlier removed [see ROI analysis]:  $r(13)=.57$ ,  $p < .05$ ); that is, individuals who were more likely to solve equations accurately also displayed greater activity in this region. As in the pre-defined frontal region, this was only true in the large number condition (see Table 1 for r-values).

None of the regions displaying number size effects was related to performance.

**3. Discussion**

Evidence has accumulated for distinct contributions of the parietal and prefrontal cortices to mathematical problem-solving with the former subserving imaginal processing and the latter goal-directed retrieval (Anderson et al., 2003, 2004a, b; Danker and Anderson, 2007; Qin et al., 2004). This study extends these findings by assessing how the functioning of these modules contributes to successful performance when solving algebra equations. Errors might occur due to a failure of either of these buffers — either because of difficulty in numerical/spatial representation or because retrieval is impaired or inefficient. We provide evidence that a prefrontal region previously shown to be modulated by retrieval difficulty is also associated with accurate problem-solving. Activity in the left PFC was greater when equations were solved correctly than incorrectly, and better solvers tended to

display more activity in this region. In contrast, parietal cortex activity did not differ depending on whether participants provided a correct answer.

The tight coupling of performance level with prefrontal cortex functioning is consistent with a large body of research examining performance in a diverse set of processes such as response selection, language comprehension, and memory (MacDonald et al., 2000; Kerns et al., 2004, Cohen et al., 1994). Prefrontal cortex is thought to guide goal-directed behavior by biasing or modulating the pattern of activity in more posterior motor, sensory, semantic, and attentional regions (Miller and Cohen, 1991). Applied to mathematical performance, we propose that the prefrontal cortex guides the retrieval of task-relevant information for solving the equation. For example, the prefrontal cortex would direct whether the result of  $5+2$  or  $5*2$  should be retrieved depending on the context of the equation. When the prefrontal cortex fails to guide retrieval appropriately, our results demonstrate that errors occur.

Interestingly, another site in the prefrontal cortex was associated with successful performance. Similar to our pre-specified region of interest in the inferior frontal gyrus, a more superior and posterior site in the middle frontal gyrus displayed greater activity for correct than error trials and tended to be more active for better problem-solvers. Note, however, that for both these sites, individual accuracy scores and neural activity were associated only in the large-number condition. Although, these prefrontal regions may show some specificity in terms of responding to numerical magnitude, these data reflect the critical link between prefrontal cortex functioning and successful problem-solving.

One might predict that successful performance would also rely on visuospatial abilities, however, parietal cortex activity was not related to accuracy. This region seemed to basically reflect time on task — showing greater activation when errors were made and greater activation for large equations. However, this greater activation occurred late in the processing. There was no relationship between activity in the early scans and either accuracy or size. Early on, it seems the parietal cortex is equally involved in the analysis of all problems and continues this involvement into later scans until the problem is solved. This results in greater activation for conditions with longer latencies.

Our results are consistent with a previous study of error detection in mathematical processing. Using a verification task, Menon et al. (2002) reported effects of accuracy in the prefrontal cortex whereas parietal regions were not affected by the accuracy of the equation. Instead, parietal cortex was modulated by the number of operands in the equation, consistent with our previous work reporting that activity in this region is affected by the number of mathematical steps that are required to solve the equation. Taken together, these results would imply that students may be better served practicing equations with varying levels of retrieval demands rather than working with large quantities per se in order to master algebra.

While suggestive of potential strategies for educating adult algebra students, it is unclear whether a different pattern would be observed for adolescents. Adults may be better at representing large quantities because of their greater experience in non-algebraic tasks such as balancing a checkbook. In contrast, adolescents' imaginal abilities may not be fully developed and, if so, an association between accuracy and

activity in the parietal cortex could be observed. This would be an important question to answer in future investigations.

While we have previously found our pre-specified region in the parietal cortex to be important in representing transformations of the problem state, this region showed little evidence of representing numerical magnitude. Previous studies (Stanesco-Cosson et al., 2000; Piazza et al., 2004) have shown that larger numbers engage the parietal cortex to a greater degree than smaller numbers, however, the region showing this effect lies in the intraparietal sulcus. In contrast, our prespecified region of interest is located in a more ventral site in the precuneus (Brodmann Area 7) more than a centimeter away from the region of the intraparietal sulcus showing number size effects. This result suggests that the visuospatial representations needed to track changes in the state of the equation are separate from a representation of number magnitude. There are a number of potential differences between the imaginal requirements of the two processes. For example, transformations of the problem state require that representations be frequently updated while magnitude is an unchanging characteristic of a number. In addition, number size may induce a mental representation of a number line while equation transformations may not rely on representations that are characterized by distance. Adjudicating between these hypotheses is beyond the scope of this paper, but our results do suggest that there are separate systems for the mental representations of magnitude and problem state.

Surprisingly, given its importance in studies of cognitive control, the anterior cingulate showed neither accuracy nor number size effects. However, the region was active in the problem-solving stage when no overt response was given as well as being active while participants were entering their response. The anterior cingulate was active even in the large number condition almost 10 s in advance of the first motor response. Thus, the anterior cingulate was active when there was no competition between overt responses. This suggests that this region may be important for goal selection in a similar way to its hypothesized role in response selection. For example, it may be important for detecting conflict between problem-solving strategies or particular operations such as whether to add or subtract.

The parietal and prefrontal sites we examined have shown robust activity across several studies indicating their importance in visuospatial processing and retrieval. The goal of this study was to assess the contributions of these computations to performance in mathematical problem-solving. We have shown that while both modules are important for solving algebra equations, activity of the prefrontal cortex is more strongly associated with solving those equations correctly.

## 4. Experimental procedures

### 4.1. Subjects

Fifteen subjects (11 females, 4 males) with an average age of 24.6 years participated in this study for monetary compensation. All participants were right-handed with normal or corrected-to-normal vision. Participants were pre-screened in a separate session outside the scanner to ensure that they made an adequate number of errors (25%–75%). Out of all the

participants who were pre-screened, 32% had an error rate that fell within our range.

#### 4.2. Stimuli

Participants had to solve two-step algebra equations that required either multiplication or division in one step and addition or subtraction in the other step (e.g.,  $14 + (7 * x) = 35$ ). A list of 214 equations was created where the parentheses enclosed either the first two numbers or the last two numbers on the left side of the equation in equal proportions. The number to solve for could be either in the first position (25%), second position (50%), or third position (25%). The positions of the parentheses and the “x” were manipulated in order to prevent learning over the course of the experiment thereby ensuring enough trials to be able to examine errors.

In addition to accuracy, the neural substrates of number size were also of interest. In the small-number condition, the operands on the left side of the equation were all one digit numbers (e.g.,  $(x+4)*5=30$ ) or required well-learned knowledge of the multiplication table (e.g.,  $1 \times 1$  to  $12 \times 12$ ). The large number condition contained at least one two-digit number and more advanced multiplication/division (e.g.,  $19 + (13 * x) = 71$ ).

#### 4.3. Procedure

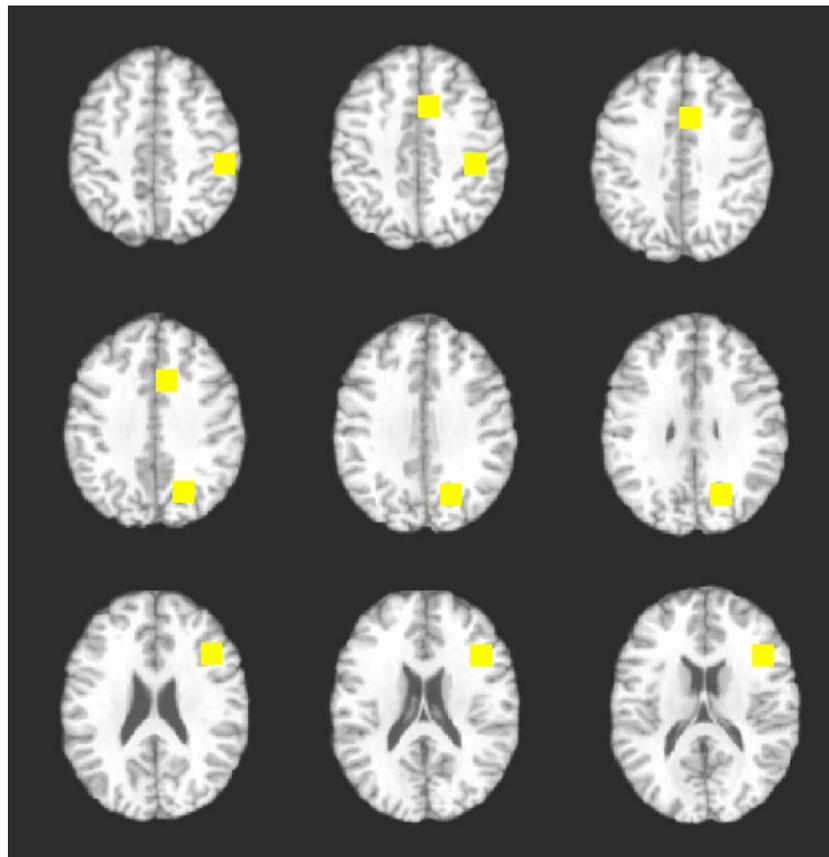
A trial began with an alerting stimulus (i.e., !) for 1 s followed by an equation which stayed on the screen until the

participant indicated they knew the answer or 39 s had elapsed. When participants felt that they had solved the equation, they pressed a key under the thumb of their right hand which then brought up a “counter”, set at 00, that was displayed in the center of the screen. Participants pressed a key under their index finger to increment the tens column and their middle finger to increment the ones column. Both the tens and ones column could be set back to 0 by the participant if they made an input error. Once they had entered their answer, they pressed a key under their thumb to indicate they were done. A fixation stimulus was then presented for the amount of time needed to ensure a trial duration of 44 s.

Participants completed 6–9 blocks of 12 trials each and equations were chosen at random from the stimulus list. Behavioral data were acquired using the E-prime software package interfaced with a response box and a color LCD projector.

#### 4.4. fMRI acquisition

Data were acquired on a 3.0 T Siemens Trio magnetic resonance imaging scanner located in the UC Davis Imaging Research Center. We used an echo planar imaging protocol to acquire 36, 3.4 mm axial slices every 2 s (TE 25 ms, 90° flip angle, 22 cm FOV, 3.4375 mm  $\times$  3.4375 mm resolution). T2-weighted in-plane scans were acquired in each scanning session as well, using the same slice thickness as the functional scans but with an inplane resolution of .859 mm  $\times$  .859 mm.



**Fig. 8** – Pre-specified regions of interest in the anterior cingulate, left motor, prefrontal, and parietal cortex based on our previous work (Anderson, 2005).

Motion correction was undertaken using an automated image recognition (AIR) program which implements two iterations of a six-parameter, rigid body realignment, to co-register each image with the first functional image acquired in a session. Data from participants moving more than an average of 3 mm in any direction or rotating their heads more than 2° across the session were excluded from the experiment. Images were spatially smoothed with an 8 mm FWHM Gaussian filter and an additive baseline correction was applied to each voxel-wise time course independently. Intensities in each voxel were detrended with a simple linear regression to remove intensity changes due to scanner drift. Structural and functional images for each participant were cross-registered to the MNI common reference brain and the Talairach atlas system using AFNI.

#### 4.5. fMRI data analysis

Statistical analyses were implemented using the NIS software suite developed by the University of Pittsburgh and Princeton University (<http://kraepelin.wpic.pitt.edu/nis>). We excluded trials where participants did not respond or were not finished incrementing the response counter. An average of 19% of the trials had to be discarded due to null or incomplete responses (range: 5%–34%). In the large-number condition, participants had an average of 18 correct responses (range 6–32) and 14 erroneous responses (range 6–28) which was sufficient for event-related fMRI analyses. Participants made an average of 31 (range 18–45) correct responses in the small-number condition, however, they committed very few errors (average of 10; range 3–30). Despite the low error rate in the small-number condition, we did assess interaction effects of number size by accuracy so that we could report the complete data set; however, interaction effects must be interpreted with some caution. Note that our effects of interest mostly lay in main effects of accuracy and number size.

Mean intensities of each voxel in each condition were averaged across participants. To assess the activity of neural regions associated with component process of the ACT-R model, we examined activity in regions centered on the peak coordinates reported in our previous studies (Fig. 8). Note that these regions of interest were all in the left hemisphere as we have found that mathematical problem-solving displays a preference for engaging this hemisphere (Anderson, 2005). A left prefrontal region was centered at  $x=-40$ ,  $y=21$ ,  $z=21$  (all coordinates refer to Talairach space) and activity was averaged over all voxels within this block ( $16 \times 16 \times 10$  mm) in each condition. Identically-sized regions of interest were used to assess parietal cortex ( $-23$ ,  $-64$ ,  $34$ ), anterior cingulate ( $-5$ ,  $10$ ,  $38$ ), and left motor cortex ( $27$ ,  $-25$ ,  $47$ ) activity. Given our focus on a small number of pre-specified regions of interest, a threshold of  $p < .05$  uncorrected was used.

A voxel-wise analysis was also performed to assess whether regions of interest outside our prespecified areas displayed accuracy or number size effects in the interval before a response was generated. Separate 4 (time points 0–10)  $\times$  2 (Correct/Incorrect or Small/Large) ANOVAS were performed on the signal intensity in each voxel using subject as a random factor. Using Monte-Carlo sampling (AfniAlphaSim) to test for the incidence of false positives, we implemented a threshold of

at least  $p < .001$  and a cluster size of 27 voxels to ensure that our rate of family-wise error was less than .05.

#### Acknowledgments

We would like to thank KC Keur Moua for her assistance. This research was funded by NSF ROLE grant REC-0087396NSF.

#### REFERENCES

- Anderson, J.R., 2005. Human symbol manipulation within an integrated cognitive architecture. *Cog. Sci.* 29 (3), 313–341.
- Anderson, J.R., 2007. How Can the Human Mind Occur in the Physical Universe? Oxford University Press, New York.
- Anderson, J.R., Qin, Y., 2008. Using brain imaging to extract the structure of complex events at the rational time band. *J. Cog. Neurosci.* 20 (9), 1624–1636.
- Anderson, J.R., Reder, L.M., Lebiere, C., 1996. Working memory: activation limitations on retrieval. *Cog. Psychol.* 30, 221–256.
- 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 Bull. Rev.* 10, 241–261.
- Anderson, J.R., Qin, Y., Stenger, V.A., Carter, C.S., 2004a. The relationship of three cortical regions to an information-processing model. *J. Cog. Neurosci.* 16 (4), 637–653.
- Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y., 2004b. An integrated theory of Mind. *Psychol. Rev.* 111, 1036–1060.
- Cohen, J.D., Forman, S.D., Braver, T.S., Casey, B.J., Servan-Schreiber, D., Noll, D.C., 1994. Activation of prefrontal cortex in a non-spatial working memory task with functional MRI. *Hum. Brain Map* 1, 293–304.
- Danker, J.F., Anderson, J.R., 2007. The roles of prefrontal and posterior parietal cortex in algebra problem-solving: A case of using cognitive modeling to inform neuroimaging data. *Neuroimage* 35, 1365–1377.
- Kerns, J.G., Cohen, J.D., Stenger, V.A., Carter, C.S., 2004. Prefrontal cortex guides context-appropriate responding during language production. *Neuron* 43 (2), 283–291.
- Koedinger, K.R., Nathan, M.J., 2004. The real story behind story problems: Effects of representations on quantitative reasoning. *J. Learn. Sci.* 13 (2), 129–164.
- MacDonald, A.W., Cohen, J.D., Stenger, V.A., Carter, C.S., 2000. Dissociating the role of dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science* 288, 1835–1837.
- Menon, V., Mackenzie, K., Rivera, S.M., Reiss, A.L., 2002. Prefrontal cortex involvement in processing incorrect arithmetic equations: evidence from event-related fMRI. *Hum. Brain Map.* 16 (3), 119–130.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal function. *Ann. Rev. Neuro.* 24, 167–202.
- Piazza, M., Izard, V., Pinel, P., LeBihan, D., Dehaene, S., 2004. Tuning curves for approximate numerosity in the human parietal cortex. *Neuron* 44 (3), 547–555.
- Qin, Y., Carter, C.S., Silk, E., Stenger, V.A., Fissell, K., Goode, A., Anderson, J.R., 2004. The change of the brain activation patterns as children learn algebra equation solving. *Proc. Natl. Acad. Sci. U. S. A.* 101 (15), 5686–5691.
- Rivera, S.M., Menon, V., White, C.D., Glaser, B., Glover, G., Reiss, A.L., 2002. Functional brain activation during arithmetic processing in females with fragile X syndrome is related to FMR-1 protein expression. *Hum. Brain Map.* 16 (4), 206–218.

- Stanescu-Cosson, R., Pinel, P., van De Moortele, P.F., Le Bihan, D., Cohen, L., Dehaene, S., 2000. Understanding dissociations in dyscalculia: a brain imaging study of the impact of number size on the cerebral networks for exact and approximate calculation. *Brain* 123 (11), 2240–2255.
- Van Veen, V., Carter, C.S., 2002. The Timing of Action-Monitoring Processes in the Anterior Cingulate Cortex. *J. Cog. Neurosci.* 14 (4), 593–602.
- 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.