

The change of the brain activation patterns as children learn algebra equation solving

Yulin Qin^{*†}, Cameron S. Carter[‡], Eli M. Silk^{*}, V. Andrew Stenger^{§¶}, Kate Fissell[§], Adam Goode^{*}, and John R. Anderson^{*}

^{*}Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213; [‡]Imaging Research Center, University of California at Davis, Sacramento, CA 95817; and Departments of [§]Psychiatry and [¶]Radiology, University of Pittsburgh Medical Center, University of Pittsburgh, Pittsburgh, PA 15260

Contributed by John R. Anderson, February 24, 2004

In a brain imaging study of children learning algebra, it is shown that the same regions are active in children solving equations as are active in experienced adults solving equations. As with adults, practice in symbol manipulation produces a reduced activation in prefrontal cortex area. However, unlike adults, practice seems also to produce a decrease in a parietal area that is holding an image of the equation. This finding suggests that adolescents' brain responses are more plastic and change more with practice. These results are integrated in a cognitive model that predicts both the behavioral and brain imaging results.

The study reported here integrates behavioral methods, functional brain imaging (functional MRI), and cognitive modeling to study how children learn to solve equations. In particular, the children were solving equations like the following: $7x + 1 = 29$. Past research with adult college students (1) modeled algebra equation solving by the interaction of three cognitive modules in the adaptive control of thought-rational (ACT-R) cognitive architecture (2, 3). There was an imaginal module that held a representation of the equation and performed imagined transformations on the equations. There was a retrieval module that retrieved algebraic rules and arithmetic facts in the solution of this equation. Finally, there was a manual module that programmed the output of the answer by the hand. A region in the left parietal cortex, which has been associated with imagery (4–6) and spatial processing (7) in other studies, was found to correspond to the imaginal module. A region in the left prefrontal cortex, which has been associated with retrieval in other studies (8–14), was found to correspond to the retrieval module. Finally, a region in the left motor and sensory cortices, which controls the right hand, was found to correspond to the manual module.

After having identified these regions in algebra equation solving, we performed a series of experiments to determine whether they were specifically involved in algebra or were also involved in nonmathematical information-processing tasks (1, 15, 16). Similar involvement of these regions was found in a nonmathematical isomorph of algebra (artificial algebra) (1). Subsequent research (15), in which college students practiced the isomorph, found a speed-up that could be accounted for entirely in terms of reduced retrieval time. This finding was reflected in reduced activation in the prefrontal region of interest. There was not a comparable reduction in either the motor or parietal region.

The present research addresses the question of whether the brain activation patterns observed from adults will be shown in children learning algebra. Specifically, do children who are just learning equation solving show activation of the same regions as in adults' algebra (1) and will their improvement be explained in terms of reduction just in the prefrontal retrieval region shown in adults' artificial algebra learning (15)? There is reason to suspect that we might see changes in activation in more regions because children are presumably in a more plastic stage of neural development. Past brain imaging studies of children have shown that, during adolescent years, the white matter volume keeps increasing (some local areas changing even rather rapidly) and

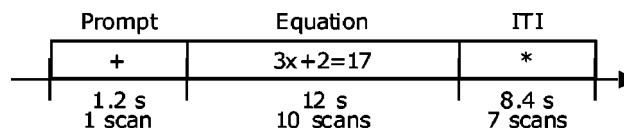


Fig. 1. The protocol of a trial. A trial lasted 21.6 s, with a red cross shown in the first 1.2 s as warning (the stimulus was visually shown on a black screen), then a 12-s period for the participants' solving the equation (which was showing in white characters) and keying the answer on the data glove, followed by 9.6 s for intertrial interval (ITI; showing a white star).

the gray matter volume in parietal and prefrontal cortex has begun to decrease, consistent with the findings from developmental neuroscience of myelination and synaptic pruning (17–23).

Methods

Subjects. Ten normal pre-algebra students (expecting to take Algebra I the following year) who replied to an advertisement in a local Pittsburgh newspaper participated in this experiment [right-handed, native English speaker, 12–15 yr old (averaged 13.1), sixth to eighth grade, three females]. Participants were accompanied with their parents and were provided written informed consent in accordance with the Institutional Review Boards at the University of Pittsburgh and at Carnegie Mellon University.

Procedure. Children in this experiment solved three types of equations: 0-step equations (e.g., $1x + 0 = 4$), 1-step equations (e.g., $3x + 0 = 12$, $1x + 8 = 12$), and 2-step equations (e.g., $7x + 1 = 29$). So that the children's performance would be close to that of adults (see refs. 24 and 25), the solutions to all problems were the digits 2–5 [5–9 in adults (1)], which mapped onto the fingers of the right hand in a data glove. Also, in contrast to the adult problems, there were no borrowing operations in 2-step equations.

The paradigm of the learning procedure was very similar to that of adults' learning artificial algebra (15). The experiment lasted 5 days, with an event-related functional MRI (fMRI) scan on day 1 and day 5 and practice without scan from day 2 to day 4, plus a pre-scan training session on the day before day 1. In the pre-scan training session, the children were given a tutorial on algebra equation solving and key practice, and performed two blocks of the real tasks in an fMRI simulator. There were 16 trials per block (see Fig. 1 for the structure of a trial), 8 blocks for each scan day, and 10 blocks for each practice day. On the scan days, event-related fMRI data were collected by using a single-shot spiral acquisition on a GE 3T scanner, 1,200-ms repetition time (TR), 18 ms echo time (TE), 70° flip angle, 20-cm

Abbreviations: ACT-R, adaptive control of thought-rational; fMRI, functional MRI; ROI, regions of interest; BA, Brodmann's area; RT, reaction time; BOLD, blood oxygen level-dependent; rpe, ratio-of-practice-effect.

[†]To whom correspondence should be addressed. E-mail: yulingq@cmu.edu.

© 2004 by The National Academy of Sciences of the USA

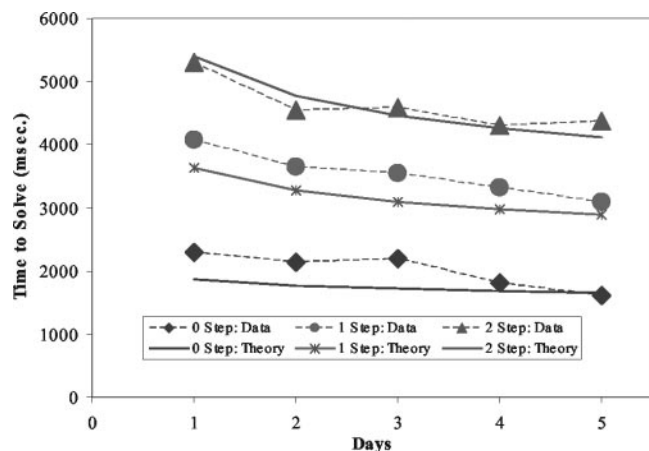


Fig. 2. Time to solve equations of differential complexity as a function of practice (observed data and the fit of an ACT-R model).

field of view (FOV), 21 axial slices each scan with 3.2-mm-thick, 64×64 matrix, and with AC-PC (anterior commissure-posterior commissure) on the second slice from the bottom.

fMRI Data Analyses. Before focusing on the three regions identified in past research, we performed an exploratory analysis to find out which brain regions varied significantly with condition. The regions of interest (ROIs) were selected according to the interaction in a 6-condition \times 18-scan ANOVA. The six conditions came from two levels of practice (day 1 and day 5) \times three levels of complexity conditions of the equations. To have a conservative test that dealt with nonindependence of scans, we used the Greenhouse–Geisser correction of assigning only five degrees of freedom to the numerator in the F -statistic for the 6-condition \times 18-scan interaction term. The interaction was examined in each voxel. To ignore the small particles, we selected regions that met the criteria of a minimum of 30 contiguous voxels with significant interaction at $P \leq 0.05$. According to Forman *et al.* (26), the probability of a false positive should be <0.05 . In the confirmatory analysis, the three ROIs are the same as in adults' artificial algebra (15): a left motor area [Brodmann's area (BA) 4/3], a left posterior parietal area (BA 39/40), and a left prefrontal area (BA 45/46). Each region was defined as 100 voxels (5 wide \times 5 long \times 4 deep), $\approx 16 \times 16 \times 13 \text{ mm}^3$, as shown in Fig. 3B.

Results

Latency and Accuracy Results. The average accuracy of the participants' behavior was high (91.3% in day 1 and 93.3% in day 5) and close to that of adult algebra (93.6%). The average reaction

time of correct trials (RT) showed significant differences among the three equation types on both day 1 and day 5, and the difference of RT between these two days was also significant [on day 1, $F(2,27) = 12.68$, $P < 0.0005$; on day 5, $F(2,27) = 21.50$, $P < 0.0001$; comparing day 1 and day 5, $F(1,54) = 8.0$, $P < 0.01$]. Fig. 2 shows the decrease in RT across the five days along with the fit of an ACT-R model that will be described.

Event-Related fMRI Findings. Exploratory analysis. With the very conservative criteria mentioned above, nine ROIs were selected according to the interaction in a 6-condition \times 18-scan ANOVA. Table 1 and Fig. 3A indicate these regions. Fig. 4 illustrates the overall blood oxygen level-dependent (BOLD) functions obtained from each of these nine regions, averaged over complexity conditions and practice. These functions are plotted as percent increase above the baseline defined by the average activation of the first two and last two scans. ROI 1 corresponds to the anterior cingulate gyrus, and it has been found to yield effects in other of our experiments. ROIs 2, 3, and 4 correspond to the motor, parietal, and prefrontal regions found in prior research (1, 15, 16). ROIs 5 and 6 are the right and left supramarginal gyrus and yield negative functions, consistent with Gusnard and Raichle (27). We have found them to yield negative responses in our studies of algebra equation solving in adults. ROI 7 is in the left occipital, ROI 8 is the left head of caudate nucleus (extends to thalamus), and ROI 9 is the left putamen. Both ROIs 8 and 9 are areas in the basal ganglia.

Table 1 indicates which regions show the effects of condition and practice in terms of t values. The t values are positively signed if the 2-transformation condition yielded a larger BOLD function than the 0-transformation condition ($df = 18$) and if day 1 yielded a larger BOLD function than day 5 ($df = 9$). It can be seen that six regions yielded a larger BOLD function in the more complex conditions. With respect to the motor region, the effect that led to its selection in the condition-by-practice interaction is that its peak is delayed in more complex condition but the average BOLD response remains the same. The left and right supramarginal gyri yield negative t values, consistent with their overall negativity.

Only three of the six regions with positive BOLD functions yielded effects of practice. Two of these, the left parietal (ROI 3) and left prefrontal (ROI 4), showed lower activation on day 5 than day 1. The effect for the prefrontal region replicates our results with adults, but the effect for the parietal region is different from the adult population. The left putamen (ROI 9) shows a greater BOLD function on day 5 than day 1. The adults' data (15) did not show this trend in this area.

Confirmatory analysis. Fig. 3B shows the regions predefined on the basis of adult results (1, 15, 16). Their overlap with the exploratory regions is quite striking. However, working with predefined regions rather than exploratory regions has the advan-

Table 1. ROI, locations of centroid, and significances in exploratory analysis

ROI	Region of interest	BA	Voxel count	Center Talairach coordinate, mm (x, y, z)	Mean percent	t value (complexity)	t value (practice)
1	Anterior cingulate	32/24	150	−3, 13, 39	0.100	9.26**	1.75
2	Left motor	4/3/2	45	−38, −22, 45	0.084	0.13	1.81
3	Left parietal	40/39	80	−27, −64, 35	0.075	7.87**	4.18**
4	Left prefrontal	9/46/45/44	223	−40, 14, 23	0.069	10.14**	5.97**
5	Right supra-marginal gyrus	40/39/22	133	42, −54, 25	−0.053	−4.90**	−2.71*
6	Left supra-marginal gyrus	40/39/22	49	−47, −61, 21	−0.080	−3.19*	−4.20**
7	Left occipital	31/17/18/19	42	−23, −73, 13	0.012	2.94*	0.65
8	Left head of caudate nuclei	Extends to thalamus	33	−7, 4, 5	0.043	6.17**	0.64
9	Left putamen		37	−26, −9, 5	0.026	2.73*	−2.48*

*, $P \leq 0.05$; **, $P \leq 0.01$.

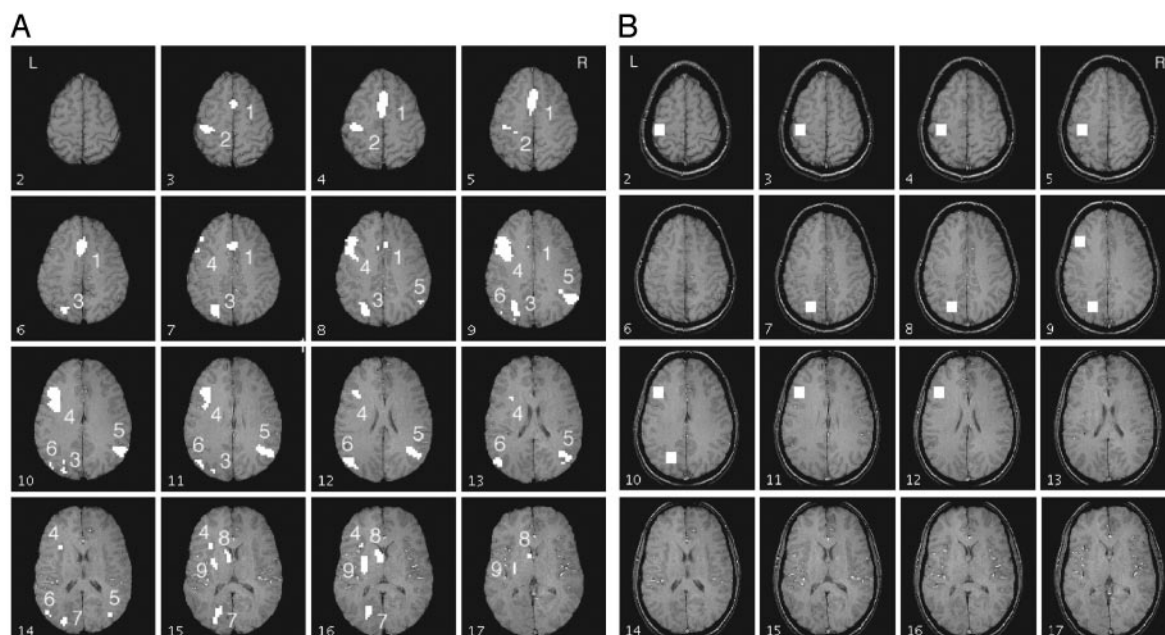


Fig. 3. (A) Activation map for 16 slices starting at the third slice from the top, showing areas with a significant interaction between scan and condition. Only regions with >30 contiguous voxels and $P \leq 0.05$, $df = 5$, are chosen. See Table 1 for identification of regions. The AC-PC (anterior commissure-posterior commissure) line is two slices below slice 17 in this figure. Left is the left side of the brain. (B) ROIs for confirmatory analysis. Each is of $100 (5 \times 5 \times 4)$ voxels ($16 \times 16 \times 13 \text{ mm}^3$). The Talairach coordinates of the center for the left prefrontal region (BA 45/46) are $x = -40$, $y = 21$, $z = 21$; of the left posterior parietal region (BA 39/40), $x = -23$, $y = -64$, $z = 34$; and of the left motor area (BA 4/3), $x = -37$, $y = -25$, $z = 47$.

tage that our estimation of their response is not biased by the statistical selection criteria. Fig. 5 illustrates the BOLD functions obtained for the three regions illustrated in Fig. 3B for both days 1 and 5. As typical BOLD functions, they rise to a peak $\approx 5 \text{ s}$ after the events of interest. The BOLD functions for the motor area rise to similar peaks that are delayed with the emission of the

response. The BOLD functions for the parietal region rise to different heights to reflect the number of transformations of the problem state. The same is true for the BOLD functions for the prefrontal region except that in the 0-step condition they show almost no rise. We performed a series of statistical tests to confirm the significance of what is apparent in the BOLD functions. According to Anderson *et al.* (1), one can measure the hemodynamic demand in a condition by the total area of these BOLD functions above the baseline. Therefore, ANOVA were performed on measures of these areas for each of the three regions where the factors were equation complexity and day of practice. There were no significant effects for the motor region [$F(1,9) = 0.79$ for practice; $F(2,18) = 0.30$ for transformation], but both effects were significant for the parietal region [$F(1,9) = 15.43$, $P < 0.005$ for practice; $F(2,18) = 30.49$, $P < 0.0001$, for transformation] and the prefrontal regions [$F(1,9) = 21.10$, $P < 0.005$, for practice; $F(2,18) = 22.70$, $P < 0.0001$, for transformation]. Another analysis was performed of whether the peak times of the BOLD functions differed as a function of condition. The only significant effect was the effect of number of steps for the motor particle. Thus, the statistical tests confirm the trends that are most apparent in Fig. 5.

Note that, in adults' learning artificial algebra (15), the practice BOLD effect in the parietal region was not significant. This difference between children and adults in this region is not because children learned more than the adults. The practice effect in behavior (average RT of correct trials in day 1 – average RT of correct trials in day 5) of children was in fact less than that of the adults [mean(adults) = 1,432, SD = 551, mean(children) = 885, SD = 504, $t(16) = 2.19$, $P = 0.043$ for two-sample, two-tail t test]. If we define behavioral ratio-of-practice-effect (rpe) as [(average RT of correct trials in day 1) – (average RT of correct trials in day 5)]/(average RT of correct trials in day 1), this trend would be even stronger [mean(adults) = 0.356, SD = 0.115; mean(children) = 0.222, SD = 0.08; $t(16) = 2.88$, $P = 0.011$ for two-sample, two-tail t

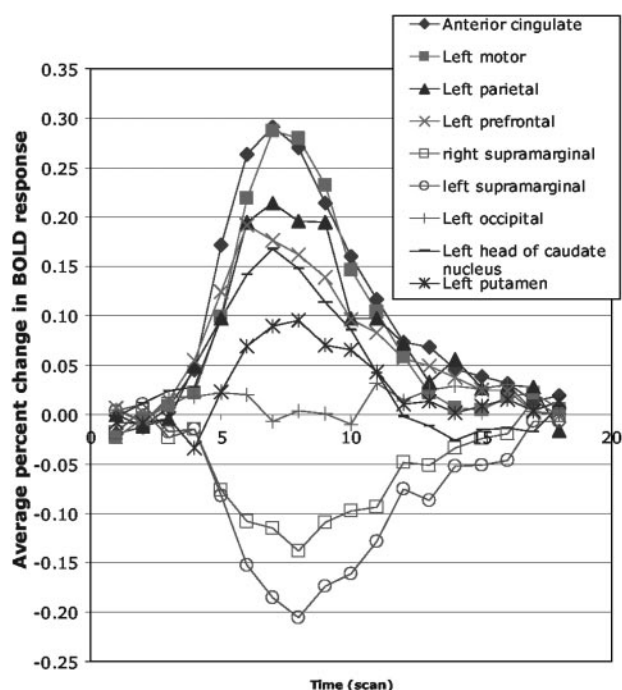


Fig. 4. The BOLD functions obtained from each of the ROIs selected by exploratory analysis, averaged over complexity conditions and practice.

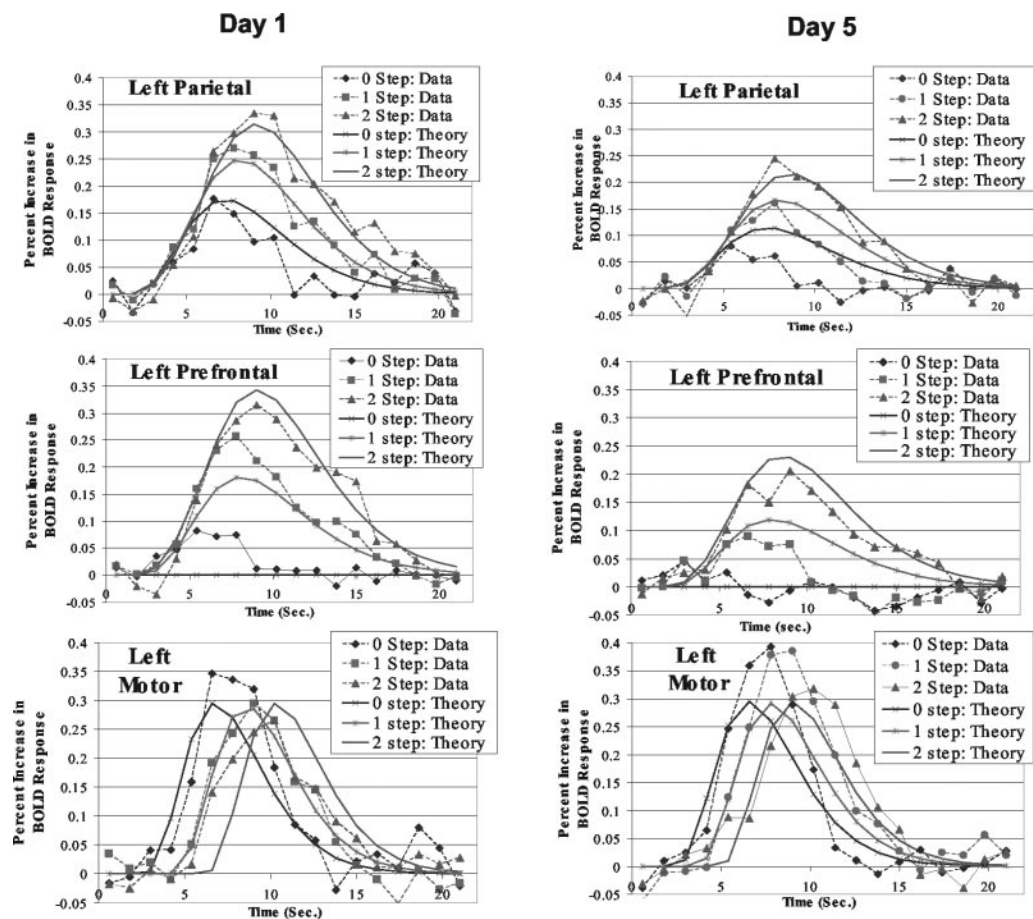


Fig. 5. The BOLD functions obtained for the three regions illustrated in Fig. 3B for both days 1 and 5 and the predictions of an ACT-R model of the task.

test). If we put child and adult data together and sort them based on their rpes, we can get three subsets (following refs. 24 and 25): (i) a matched group, formed by children and adults with very similar rpe (six children, mean = 0.278, SD = 0.048; five adults, mean = 0.289, SD = 0.062; $t(9) = -0.33$, $P = 0.749$ for two-sample, two-tail t test); (ii) the remaining four children with smaller rpe formed a nonmatched child group (mean = 0.139, SD = 0.02, comparing matched with nonmatched children, $t(8) = 5.32$, $P < 0.001$, for two-sample, two-tail t test); (iii) the remaining three adults with larger rpe formed nonmatched adult group (mean = 0.468, SD = 0.09, comparing matched with nonmatched adults, $t(6) = -3.30$, $P = 0.016$, for two-sample, two-tail t test). In matched group, children had significant practice BOLD effect in the parietal region [$F(1,5) = 7.74$, $P = 0.039$], but adults did not [$F(1,4) = 0.965$, $P = 0.382$]. In nonmatched groups, children still had significant practice BOLD effect in the parietal region [$F(1,3) = 26.12$, $P = 0.014$]. There was some practice effect shown in the nonmatched adult group but not significant [$F(1,2) = 7.11$, $P = 0.117$]. It seems that age and not performance is the major factor causing the significant practice effect in the parietal region.

ACT-R Modeling. Fig. 5 also shows the predictions of an ACT-R model of this task, which we will now describe, and Fig. 6 shows the operations of four ACT-R modules in solving the equation $7x + 1 = 29$. There are modules associated with visual encoding of the equation, mental transformation of the equation, retrieval of algebraic and arithmetic information, and keying of the response. No brain region was found corresponding to the visual module, presumably because an equation of the same visual

complexity is presented for the same duration in all conditions. The other three modules have the following correspondence to our predefined regions: imaginal module corresponds to left parietal, retrieval module to left prefrontal, and manual module to left motor.

The actual ACT-R module is a computer simulation of a production system that reads the equation, solves it, and keys the answer. In addition to the module activity illustrated in Fig. 6, there are varying number of production rule firings that coordinate the activity of the modules. Each of these production rule firings takes 0.05 s. These 50 ms rule firings, the visual encodings of elements from the equation (0.135 s per element), and the time to program and execute the key press (0.4 s) constitute the fixed components of the model. All of these fixed times are based on prior values in the ACT-R architecture. We assume that the imaginal and retrieval activities will speed up, corresponding to the decrease in the BOLD response for these modules in Fig. 5. The predictions of response times in Fig. 2 are simply a result of adding up these fixed and decreasing components. The following are the equations for the time for the three conditions of the experiment

$$\text{Condition 0 (Day)} = 1.26 + 1.0 \times \text{Transformation (Day)} \quad [1]$$

$$\begin{aligned} \text{Condition 1 (Day)} = & 1.47 + 2.0 \times \text{Transformation (Day)} + \\ & 1.5 \times \text{Retrieval (Day)} \end{aligned} \quad [2]$$

$$\begin{aligned} \text{Condition 2 (Day)} = & 1.69 + 3.0 \times \text{Transformation (Day)} + \\ & 3.0 \times \text{Retrieval (Day)}, \end{aligned} \quad [3]$$

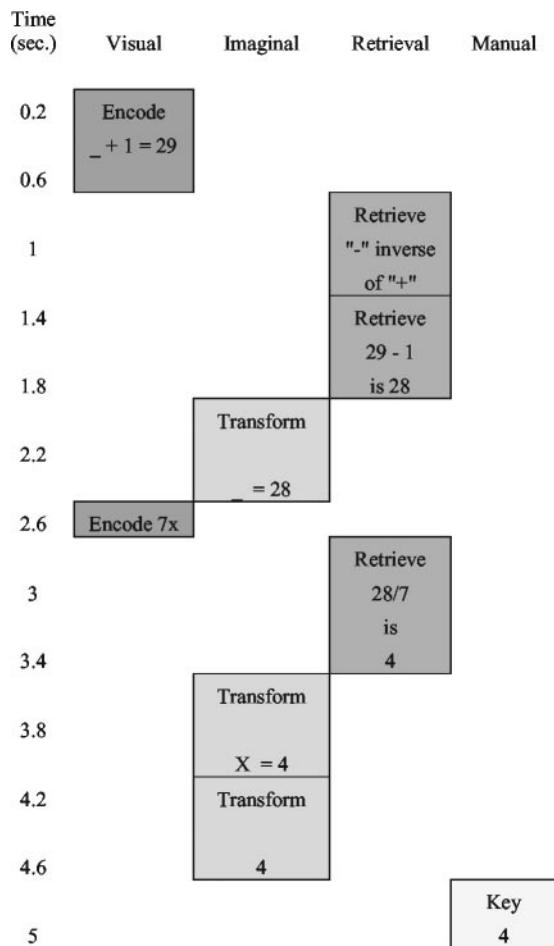


Fig. 6. The operations of four ACT-R modules in solving the equation $7x + 1 = 29$.

where Transformation(Day) is the time spent transforming the equation on a day and Retrieval(Day) is the time spent retrieving information. The intercepts (1.26, 1.47, and 1.69) reflect the fixed costs for that condition and the transformation and retrieval factors are multiplied by the number of transformations and retrievals for that condition. Assuming that transformation and retrieval follow power laws (28), they were set to be

$$\text{Transformation (Day)} = T \times \text{Day}^{-c} \quad [4]$$

$$\text{Retrieval (Day)} = R \times \text{Day}^{-c} \quad [5]$$

In fitting the latency data in Fig. 2, the parameters T and R were both estimated to be both 0.63 s and c was estimated to be 0.28. These estimations yield reasonable fits to the observed latencies (correlation coefficient = 0.986).

Having now set the times of the various components of processing, predictions can be made about the BOLD functions in the three regions of interests. See Anderson *et al.* (1) for the details of the methodology. Briefly, with the timing information (t) of the model, the BOLD response can be predicted by

$$CB(t) = M \int_0^t i(x) B\left(\frac{t-x}{s}\right) dx, \quad [6]$$

where M is the magnitude scale for response, s is the latency scale, $i(x)$ is 1 if the module is occupied at time x and 0 otherwise, and $B(t)$ is a gamma function

Table 2. Parameters and the quality of the BOLD function prediction

	Imaginal	Retrieval	Manual
Scale (s)	1.557	1.735	1.201
Exponent (a)	4.760	4.305	4.582
Magnitude:			
$M \Gamma(a+1)^*$	0.733	1.612	4.303
Correlation day 1	0.950	0.946	0.913
Correlation day 5	0.902	0.882	0.953

*This is a more meaningful measure because the height of the function is determined by the exponent as well as M .

$$B(t) = t^a e^{-t} \quad [7]$$

that describes the BOLD response to an event that varies according to time t since the event (29–31). The predictions shown in Fig. 5 were obtained from the module time course shown in Fig. 6 by using this methodology.

Whereas the exact shapes of the BOLD functions depend on the estimation of the parameters M , s , and a (Table 2), the relative areas under the curves and the peaks are parameter-free predictions of the timing of the modules. The differential timing of the response is controlling the shifting peaks in the motor region, but the actual BOLD function is not changing because the programming of the response is not changing. With respect to the parietal region, the BOLD response reflects the number of transformations: there are up to two to rework the equation and a final transformation to retrieve the response from the equation. With respect to the prefrontal regions, there are potentially three things to be retrieved (all three are illustrated in Fig. 6) depending on condition. The effect of condition on the magnitude of the BOLD response reflects the number of transformations or retrieval. The one qualitative difference between the parietal and the prefrontal region is that the parietal region shows a response even in the 0-step condition (because the answer has to be extracted from the equation) whereas the prefrontal region does not (because nothing needs to be retrieved in this condition). It is worth noting that the magnitude of reduction of the BOLD response in these two regions is predicted directly from fitting the latency speed-up and required no additional estimation.

Discussion

Accumulated brain imaging evidence has shown similar brain activation patterns of adolescents and adults in performing various high-level cognitive tasks as well as certain differences in some areas depending on the tasks (20, 32–37). The exploratory analysis of the current study shows that the active areas in children's algebra equation learning are similar to areas active in adults (1, 15). The results of the confirmatory analysis largely confirm the results obtained with adults (15). The one difference is that the parietal cortex showed a practice effect whereas this was not found with adults.

It should be noted that the ACT-R theory provides a more detailed account of what is behind the speed-up in the retrieval component and the decreased activation in the prefrontal region. According to a connectionist implementation of ACT-R (38), the process of retrieval of a memory is the process of selecting the correct memory trace in competition with other memory traces. As a memory trace is strengthened with practice, the signal corresponding to the target trace is enhanced and it takes the system fewer cycles to settle into a state corresponding to the target trace. However, there is no corresponding learning process in ACT-R corresponding to the speed with which the system can transition from one problem representation to an-

PNAS

PNAS

100

100

Qin et al. PNAS April 13, 2004 vol. 101 no. 15 5691