Trends in Neuroscience and Education 3 (2014) 50–62

Contents lists available at ScienceDirect

Trends in Neuroscience and Education

journal homepage: www.elsevier.com/locate/tine

Research Article

An fMRI investigation of instructional guidance in mathematical problem solving

Hee Seung Lee [a,*], Jon M. Fincham [b], Shawn Betts [b], John R. Anderson [b]

[a] Department of Education, Yonsei University, 50 Yonsei-ro, Seodaemu-gu, Seoul, Republic of Korea
[b] Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:
Received 5 July 2013
Received in revised form
9 January 2014
Accepted 15 January 2014

Keywords:
Instruction
Example
Mathematical problem solving
Prefrontal cortex

ABSTRACT

In this fMRI study, students learned to solve algebra-like problems in one of the four instructional conditions during behavioral session and solved transfer problems during imaging session. During learning, subjects were given explanatory or non-explanatory verbal instruction, and examples that illustrated the problem structure or the solution procedure. During transfer, participants solved problems that required complex graphical parsing and problems that required algebraic transformations. Explanatory instruction helped in the initial phase of learning, but this benefit disappeared in transfer. The example type had little effect on learning, but interacted with problem type in the transfer. Only for algebraic problems, the structural example led to better transfer than the procedural example. The imaging data revealed no effect of verbal instruction, but found that participants who had studied structural examples showed higher engagement in the prefrontal cortex and angular gyrus. Activity of the right rostrolateral prefrontal cortex in the initial transfer block predicted future mastery.

1. Introduction

Mathematical problem solving is a complex task that requires numerous cognitive functions including fact retrieval, calculation, and reasoning. Research on optimal instructional methods has focused on how students should learn (e.g., is it better to give students direct instruction or let them discover for themselves? [37,41,43,44,69]) and what students should learn (e.g., what should be emphasized and taught during instruction [51,59]). Considerable research has also examined the neural basis of mathematical knowledge to better understand the learning mechanisms that underlie mathematical problem solving (e.g., [4,15,29,35,49,50,52,53,61,72]). In the current research we want to examine the effects of various instructional methods on learning and transfer in mathematical problem solving and identify how these different instructional methods affect the activity of relevant brain regions.

1.1. Cortical areas involved in mathematical problem solving

Dehaene [24,25] proposed that number representations are distributed over several areas that code for a different aspect of numbers (known as "triple-code theory"). Expanding on this theory, Dehaene and his colleagues [26] further identified three parietal regions that are critical for number processing. The horizontal intraparietal sulcus (HIPS) supports numerical quantities or magnitude; the posterior superior parietal lobule (PSPL) supports the spatial and attentional aspects of number processing; and the left angular gyrus (AG) in connection with other left perisylvian language networks supports a verbal form of number processing. The left AG is also known to be reliably activated in a wide range of semantic tasks (e.g., [6,27,56,62]).

The models of equation solving [1,57] and mental multiplication [61] developed from the ACT-R theory [2,3] emphasize the role of the posterior parietal cortex (PPC) and lateral inferior prefrontal cortex (LIPFC) in mathematical problem solving. According to the ACT-R theory, the PPC is the imaginal module and its activity reflects maintenance and transformation of internal representations (e.g., operations on a mental representation of equations), whereas the LIPFC is the retrieval module and its activity reflects retrieval of declarative knowledge (e.g., arithmetic facts, task instructions). Various experiments support this view by showing that brain activity in the PPC correlates with problem complexity while LIPFC activity correlates with proficiency (for a review, see [2]).

Anderson et al. [4] and Wintermute et al. [72] examined cognitive and metacognitive functions of the five regions mentioned above when students extend their mathematical competence. Students learned to solve novel mathematical problems called pyramid problems. Participants solved problems that had a

* Corresponding author. Fax: +82 412 268 2844.
E-mail address: hslee00@yonsei.ac.kr (H.S. Lee).

http://dx.doi.org/10.1016/j.tine.2014.01.001
2211-9493

practiced algorithm (regular problems) and problems for which they had to invent solutions (exception problems). By observing profiles of brain activity patterns across different problem types and different periods of problem solving, two distinct networks – cognitive and metacognitive – were identified. Cognitive networks were more engaged during the solution than during the feedback period and showed little difference between regular and exception problems. In contrast, metacognitive networks were more engaged during the feedback than during the solution period and showed more activity for exception problems than for regular problems. In both Anderson et al. and Wintermute et al., the AG showed a strong metacognitive pattern while the PPC and PSPL showed cognitive patterns. The HIPS and LIPFC showed a mix of cognitive and metacognitive patterns.

Besides the five regions (HIPS, PSPL, AG, PPC, and LIPFC), a potentially interesting brain region for complex problem solving is the rostrolateral prefrontal cortex (RLPFC) [18,31]. Anderson et al. [4] and Wintermute et al. [72] found that this was another region that showed a metacognitive pattern of being more engaged for exception problems. A number of studies provide converging evidence that this region is indeed critical for metacognitive processes (for a review, see [55]). For instance, Christoff and Gabrieli [18] suggested that the RLPFC is involved in introspective evaluation of one's own thoughts and feeling. Numerous other studies implicate its metacognitive role in relational integration [9,10,19,71], cognitive branching [39,40], monitoring of memory retrieval [8], prospective memory [5,11], metacognitive awareness in perceptual decision-making [30], and shifting between external and internal processes [12].

1.2. *Designing instructional conditions: how students should learn and what students should learn*

One of the widely discussed issues in mathematics education is how students should learn for optimal learning – is it better for students to learn from direct instruction or is it better for them to be left to discover things on their own? Both positive and negative evidence exists for providing direct instruction (for a review, see [44]). For instance, several studies have shown that students who invented their own procedure developed better conceptual knowledge than those who were directly given instruction on solution procedures (e.g., [13,34,36]). In contrast, numerous studies have shown the superiority of giving direct instruction in various domains including science [16,65], mathematics [14,20,67], and procedure learning [60]. Lee et al. [46] proposed that the critical variable is not whether there is an instructional explanation, rather whether students appreciate the underlying structure of a problem. Such a problem structure is not often obvious to students learning problem solutions. Lee et al. found that, when hidden problem structure was revealed by non-verbal scaffolding devices (e.g., color-coding), students were successfully able to master a problem-solving skill even when there was no instructional explanation.

Several studies report that principle-based instruction was more effective than procedure-based instruction in mathematics education (e.g., [51,59]). When the underlying principles are not emphasized during instruction, students often fail to acquire conceptual understanding and simply focus on memorizing procedures. Solution steps will be learned and practiced without conceptual understanding, but learning without understanding can lead to poor transfer performance [33,63]. Therefore, good instruction should effectively convey underlying principles so that learners develop a good mental model of a domain. When instruction fails to emphasize hidden principles, students may look like successful learners in the directly taught domain, but not in a related but different domain. Students often rely too much on syntactic rules and do not know how to transform their learned procedures when problems are slightly changed [17,58].

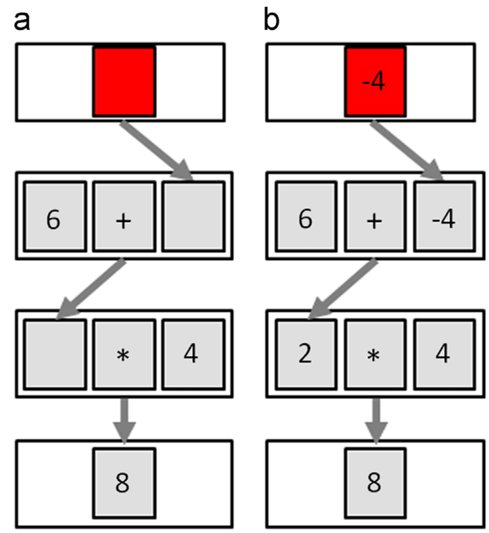

Fig. 1. An example of a propagate problem used in the learning session. (a) Original diagram of a propagate problem: Equivalent of $(6 + x) * 4 = 8$, (b) Completed problem.

2. Current experiment

The current study focused on the effects of different kinds of instructional guidance when students transfer to solving novel problems. We examined both transfer success and the engagement of different brain regions implicated in mathematical problem solving. Although participants were given different types of instructional guidance during the behavioral learning session, imaging data were collected while they solved transfer problems under the identical imaging transfer session. We developed a computer-based instructional system for teaching data-flow algebra problems like the one used by Lee et al. [45,46], which was originally adapted from Brunstein et al. [7].[1] By changing the representation of algebraic expressions to a visual representation of a data-flow diagram, we were able to test algebra learning anew in a college population. In data-flow representations (Fig. 1) a number flows from a top box through a set of arithmetic operations to a bottom box. If that number is unknown, the data-flow structure is equivalent to an algebraic equation with a variable. Fig. 1(a) and (b) shows an example problem and completed state of that problem, respectively. Fig. 1(a) is the data-flow equivalent of $(6+x)^*4=8$. The task is to determine what values to fill into the empty tiles in the boxes. For a linear structure like Fig. 1, the values can be determined by simply "propagating" the number up from the bottom, performing the arithmetic operations – for the problem of Fig. 1(a), this involves placing 2 in the empty tile above the bottom box (rewriting this as $2^*4=8$), then placing -4 in the tile above it (rewriting this as $6+-4=2$), and finally placing -4 in the top unknown box. In our previous studies [45], we observed that most students found propagating numbers easy and intuitive, and thus they did not need much assistance when learning to solve these "propagate problems". Such propagate problems were used to help familiarize students with the computerized instructional system used in this study.

However, when problems cannot be solved by this simple propagation strategy, most students have difficulty inducing a correct procedure for determining the values to fill into the empty tiles of the diagram. Fig. 2(a) shows an example of such a problem

[1] The entire experimental materials and examples of transfer problems are available at https://www.dropbox.com/sh/utw5fuubt4cgxz1/qCDc4Dn4gw.

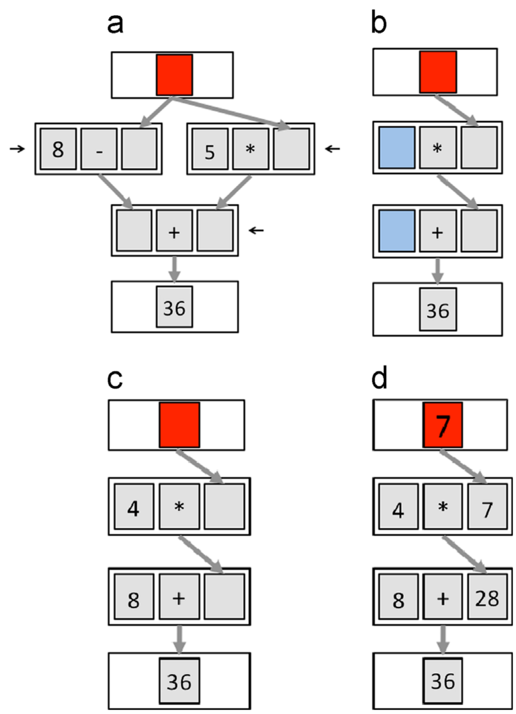

Fig. 2. An example of a linearize problem used in the learning session. (a) Original diagram of a linearize problem: Equivalent of (8 − x) + (5*x) = 36, the arrows indicate boxes to select for linearization; (b) Transformed diagram: Equivalent of (8 − x) + (5*x) → a + bx, students have to fill numbers into the blue tiles; (c) Results of linearization: Equivalent of 8 + 4x = 36, after filling in two blue tiles correctly, students can fill the rest of the empty tiles using propagation; and (d) Completed problem. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

used in this study. In this example, the diagram is equivalent to the algebraic expression, $(8-x)+(5^*x)=36$, where an unknown number in the top box flows down into the boxes below and the resultant value becomes 36. Different from propagation problems (e.g., Fig. 1), in cases like 2(a), two paths converge in a single result; thus a simple propagation procedure is not possible. The way to solve this problem within the rules of the system is to transform the diagram in Fig. 2(a) into the form in Fig. 2(b), where the data-flow diagram has a linear structure of $a+bx$, and to determine two numbers to fill into the two empty tiles as shown in Fig. 2(c). Because such a linear structure is not obvious to subjects who are not familiar with data-flow diagrams, this transformation step causes a major conceptual hurdle in this study. We call this transformation step *linearization* and will focus on this most difficult part of problem solving.

To correctly transform the diagram, students have to select the appropriate set of rectangular boxes to "linearize". In Fig. 2(a), students have to select the three rectangular boxes that are indicated by the arrows. After selecting these three boxes, clicking a linearize button transforms the data-flow diagram into the intermediate state in Fig. 2(b). The transformed diagram has a structure of $a+bx$. This linear structure is equivalent to the simplified expression of the original diagram when performing collection of like terms. Thus students need to find the coefficient that multiplies the top unknown value and the constant that adds to the first product. In this example, the selected boxes for linearization are simplified into the equivalent of $(8-x)+(5^*x)=8+4x$; thus a is 8 and b is 4. The coefficient 4 goes into the box next to the "*" operator and the constant 8 goes into the box next to the "+" operator. After filling in these two values, as shown in Fig. 2(c), now the problem becomes a propagate problem where students can apply the simple propagating strategy to fill in the rest of the empty tiles as shown in Fig. 2(d).

In this study, participants learned to solve these linearize problems under one of the four instructional conditions. We manipulated the type of verbal instruction (explanatory vs. non-explanatory) students were given when they requested a hint in the learning session. Crossed with this, we also provided two different kinds of instructional examples on students' help request. The examples illustrated either the structure behind the solution (structural example) or the solution procedure for calculating the answer (procedural example). After learning to solve problems under one of these four conditions, all participants were transferred to an identical transfer condition where they had to solve two types of more challenging problems.

3. Methods

3.1. Participants

Forty-eight graduate and undergraduate students (22 male and 26 female, $M=21.92$ years, $SD=2.6$) from Carnegie Mellon University participated in this study. All participants were right-handed and had normal or corrected-to-normal vision. Each participant was randomly assigned to one of the four learning conditions. Participants received $10/h plus a performance-based bonus.

3.2. Design and materials

A 2×2 between-subjects design was employed to test the effects of different types of verbal instruction and instructional examples. First, the type of verbal instruction was manipulated by providing explanatory or non-explanatory verbal instruction. In the **Explanatory** condition, hint texts provided explanations of why various actions were necessary and how intermediate answers were computed. In the **Non-explanatory** condition, the textual message just specified the actions required in the interface such as "click a box" and "enter a number." Thus in this condition, subjects have to discover problem solving rules on their own. This manipulation was applied during the entire learning session. Example verbal instructions are shown below the diagrams in Fig. 3 for the explanatory and non-explanatory conditions.

Crossed with the explanatory versus non-explanatory manipulation of verbal instruction, two different types of examples (structural vs. procedural) were constructed to illustrate different aspects of the problem solution. These examples were available whenever students requested a hint. Fig. 3 shows an example of each instructional example condition. In the **Structural example** condition shown in Fig. 3(a), when students requested a hint, algebraic expressions were directly drawn onto the diagram students were solving, to illustrate how the diagram is equivalent to an algebraic expression. The example reveals how the components of the diagram come together to be the equivalent of an algebraic expression on which one can perform collection of like terms. This type of example clearly shows that the data-flow diagram is just an equivalent of algebraic expression; thus, students would have a basis for understanding the values that they are entering for coefficient and constant. However, if students were to actually follow these steps mentally in solving a diagram they would face large working memory demands (especially when problems are large) because they would have to simultaneously calculate both the coefficient and constant terms.

In the **Procedural example** condition illustrated in Fig. 3(b), we illustrated the steps that proficient problem solvers in this system reported using in our previous behavioral studies [45,46]. When students requested a hint, the system first showed the calculation of the coefficient terms as shown in the left part of Fig. 3(b) and then the calculation of the constant as shown in the right part of

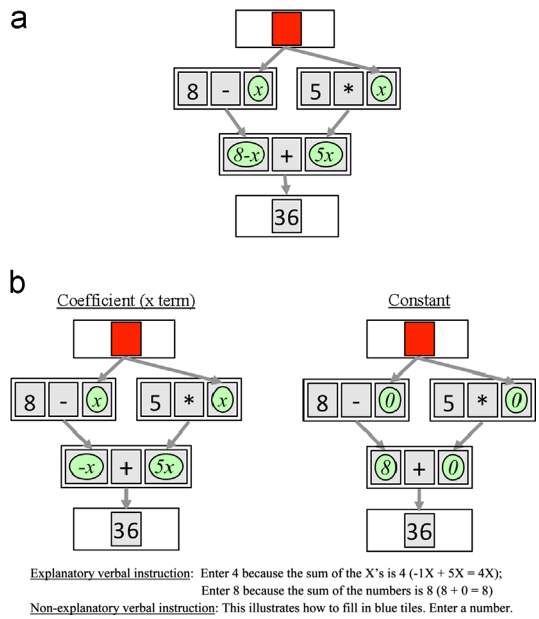

Fig. 3. Example illustration for (a) structural and (b) procedural example condition. In the (a) structural condition, both coefficient and constant terms are computed simultaneously using equivalent algebraic expressions, whereas in the (b) procedural condition, solution steps for the coefficient term (left) is illustrated first and then constant term (right) second. Depending on the verbal instructional conditions, different hint texts were provided and example texts are shown below the figures for explanatory and non-explanatory conditions. Hint texts were identical for structural and procedural example condition. Depending on which box is being filled in participants in the explanatory condition see the instruction for 4 or 8. In the non-explanatory condition they see the same verbal information for both boxes.

Table 1
Examples of learning problems and transfer problems. Each example shows the type of required transformations. There was only one type of linearize problem in the learning session and there were four subtypes of graphic and algebraic problems in the transfer session.

		Example problem and required transformation	Correct answers
Learning session	Propagate	$2–x+12=8$ (No transformations required)	$x=6$
	Linearize	$(8–x)+(5*x)\rightarrow a+bx$	$a=8$, $b=4$
Transfer session	**Graphic problems**		
	Type 1	$(((8–x)+x)+x)+(5*x)\rightarrow (a+bx)+x+(5*x)$	$a=8$, $b=0$
	Type 2	$(((8–x)+x)+x)+(5*x)\rightarrow (a+bx)+(5*x)$	$a=8$, $b=1$
	Type 3	$((8–x)+5)+(8–x)\rightarrow a+bx'$ $(x'=8-x)$	$a=5$, $b=2$
	Type 4	$((8–x)+5)+(8–x)\rightarrow a+bx$ $(x'=8-x)$	$a=21$, $b=-2$
	Algebraic problems		
	Type 1	$(8–x)+(5*x)\rightarrow(x+a)*b$	$a=2$, $b=4$
	Type 2	$(8–x)+(5*x)\rightarrow -a+bx$	$a=-8$, $b=4$
	Type 3	$(8–x)+(5*x)\rightarrow a–bx$	$a=8$, $b=-4$
	Type 4	$(8–x)+(5*x)\rightarrow 4*(a+bx)$	$a=2$, $b=1$

Fig. 3(b). By focusing on just one component, working memory load is reduced. Reducing cognitive load is known to enhance learning [66,68]. However, a potential problem of directly teaching how to calculate the components is that the equivalent algebraic transformation is never illustrated.

In the transfer session, instead of solving a complete form of data-flow diagrams (i.e., stages of Figs. 2(a) through (d)), participants were asked to perform only the linearization step. An original data-flow diagram and transformed diagram were simultaneously shown and participants had to determine what values to enter on the transformed diagrams. There were two types of transfer problems: graphic and algebraic problems. The **Graphic** problems required dealing with a novel complexity in the diagram structure but the basic procedure for determining the answer was unchanged. In contrast, **Algebraic** problems required algebraic transformations after collecting coefficient and constant terms. Table 1 shows examples of graphic and algebraic transfer problems. For linearize problems used in the learning session, the transformed diagram always has a structure of $a+bx$; thus subjects solve problems by collecting coefficient and constant terms from the original diagram. For graphic transfer problems, transformed diagrams also have a structure of $a+bx$, but original diagrams have a more complex structure. Only some portions of the original diagram have been transformed or the top unknown number was algebraic expression (e.g., $8-x$), rather than a single unknown number (e.g., x). This requires parsing of graphic complexity, and subjects have to figure out which boxes to include or exclude for computations. On the other hand, algebraic problems require an algebraic transformation after performing collection of like terms. Because the transformed diagrams have a structure (e.g., $a-bx$) other than $a+bx$, subjects have to understand the structure of the linearized diagram they have not practiced in the learning session.

Fig. 4 shows an example of graphic(a) and algebraic(b) problems. In Fig. 4(a), only two boxes (indicated by arrows) have been transformed whereas other boxes remain unchanged. This requires that students attend to which boxes were included or excluded for linearization because that in turn changes what to include or exclude when collecting coefficient and constant terms. This and the other graphic problems required a minimal change to the learned procedures – as long as students identified which boxes to include for computation they should solve the problem correctly. Here, the original diagram is equivalent to $(((8-x)+x)+x)+(5^*x)$ and the transformed diagram is equivalent to $(a+bx)+x+(5^*x)$. Because $(8–x)+x$ can be simplified into $8+0x$, the answers are 0 and 8 to this problem. There were four subtypes of graphic problems that involved various forms of graphical parsing. Fig. 4(a) shows only one subtype and the other three subtypes are shown in Table 1 (examples at website in footnote 1).

In contrast, to correctly solve the algebraic problems (Fig. 4b), after collection of coefficient and constant terms participants have to make an algebraic transformation based on the changed number or operator on the linearized diagram. For instance, in Fig. 4(b), the box with the "+" operator appears above the box with the "*" operator in the linearized diagram. This is different from the linearized diagrams participants solved during the learning session, in which the box with the "*" operator always appeared above the box with the "+" operator (see Fig. 2(c) for an example). If participants simply memorized that the coefficient goes into the top box and the constant goes into the bottom box, this problem would not be correctly solved. A further algebraic transformation is necessary. For instance, the original diagram of Fig. 4(b) is equivalent to $(8–x)+(5*x)$ and the transformed diagram is equivalent to $(a+x)*b$. The original expression $(8–x)+(5*x)$ can be simplified to $8+4x$, but to enter something into the transformed diagram participants have to further simplify this into $(2+x)*4$. To remove the possibility that participants might fail to correctly solve problems simply because they did not notice the structural change in the transformed diagram, we highlighted the changed parts using a different color. In this example, the operators, both "+" and "*" were highlighted. These problems can be correctly solved only if students appreciate how the two data-flow diagrams are equivalent. There were four subtypes of algebraic problems that involved various forms of algebraic transformations. Fig. 4(b) shows only one subtype and the other three subtypes are shown in Table 1 (examples at website in footnote 1).

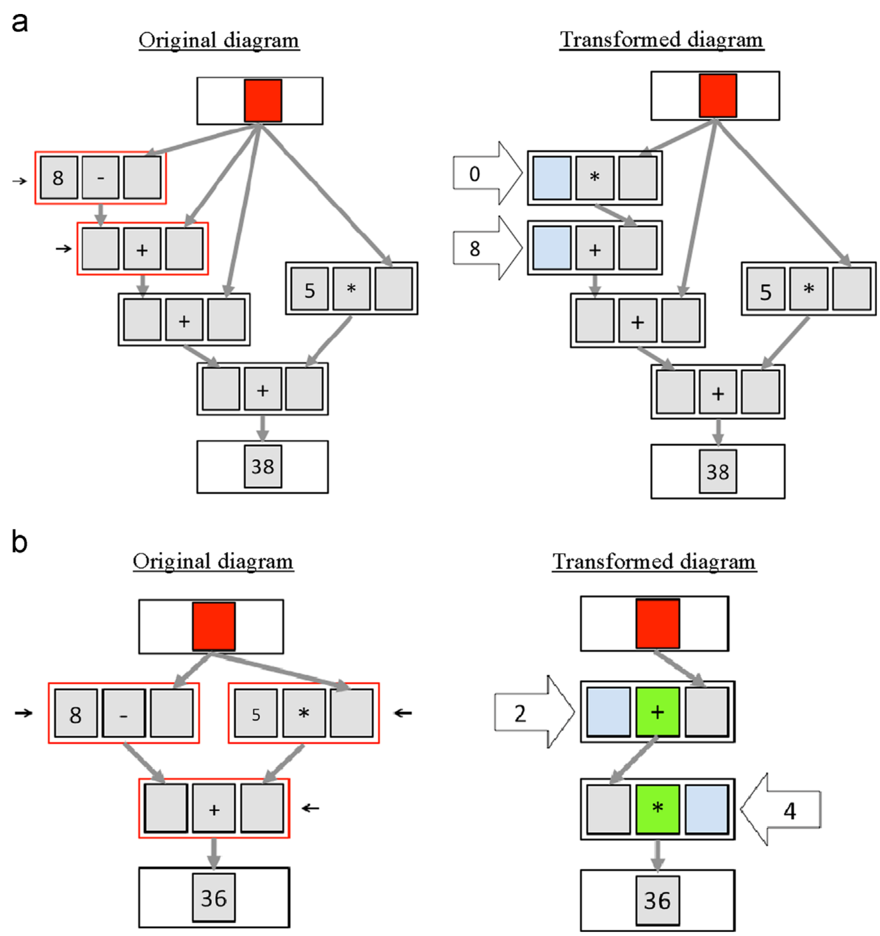

Fig. 4. An example of (a) graphic and (b) algebraic problems used in the transfer session. The left diagram shows the state before linearization and the right diagram shows the state after linearization. The correct answers are given in the arrow in this example. (a) is equivalent to transforming $(((8-x)+x)+x)+(5^*x)\rightarrow(a+bx)+x+(5^*x)$ where $a=8$, $b=0$ and (b) is equivalent to transforming $(8-x)+(5^*x)\rightarrow(a+x)^*\text{b}$ where $a=2$, $b=4$.

3.3. *Procedure-behavioral training session*

The study consisted of two sessions: a behavioral learning session and an imaging transfer session. Fig. 5(a) shows the schematic representation of the study procedure. Each session lasted 2 h and there were 1 or 2 days between the two sessions. The learning session consisted of two problem sections, one with 20 propagate and the other with 40 simple linearize problems. Each problem section followed after 1 worked-example and 1 guided problem that automatically showed a hint for every step without requests. Hints consisted of verbal instructions for each step and instructional examples that were designed for the corresponding instructional conditions. If a participant failed to perform a correct step of problem solving for 1 min, a hint automatically appeared for the corresponding step the participant was performing. After another 1 min was reached, that step was automatically solved by the system. This time limit feature was intended to reduce excessive floundering. To help reflection activity for the already solved problems, a corresponding hint stayed there for another 1 min or until a student tried to move on to the next step.

3.4. *Procedure – fMRI session*

The transfer session was identical for all instructional conditions. No hints or examples were provided in any conditions. Because the transfer task format (i.e., just filling in the two values) was quite different from the learning task, participants solved some practice problems to familiarize them with this new format during structural image acquisition. These practice problems involved the new format, but the same kinds of problem structure as they had practiced in the learning session; the problems did not involve the graphic or algebraic transfer problems. This was followed by 8 blocks of transfer problems and each block consisted of 8 problems (4 subtypes of graphic and 4 subtypes of algebraic problems). Each subtype of transfer problem appeared once in each block and the order of the problems was randomized for each block. Fig. 5(b) shows an illustration of events for each trial. A trial started with a 3-s fixation and then was followed by a problem that stayed on the screen until the participants submitted their answers or until 1 min was reached. Participants entered their answers via a numerical keypad displayed on the screen by using a mouse. Their response was followed by a feedback page, which showed one of the messages, “correct”, “incorrect”, or “time's up” on the top of the screen. Also, participants' own answers and correct answers were presented simultaneously so that they could compare them. The feedback was presented for 2 s for correct and 10 s for incorrect responses. There was no explanation provided in any conditions. After the feedback, there was another 3 s of fixation followed by a repetition-detection task for 12 s. In the repetition-detection task, letters appeared on the screen at a rate of 1/1.25 s. Participants were told to click a match button whenever the same letter appeared twice consecutively. This was intended to distract participants from thinking about the problem and return activation to a constant baseline.

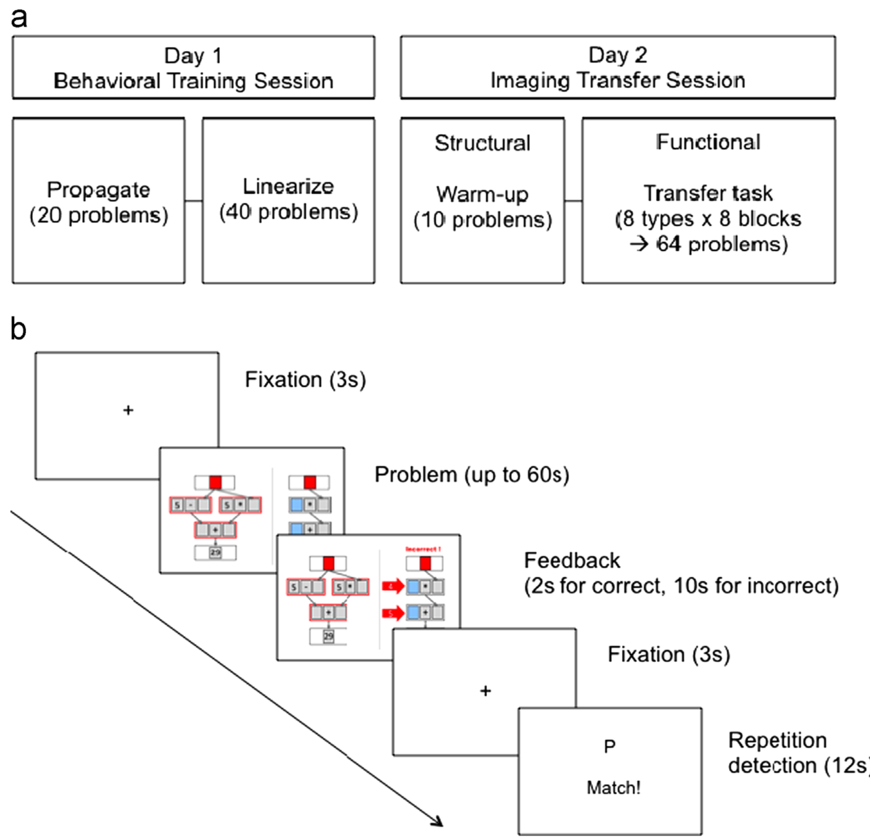

Fig. 5. An illustration of (a) the study procedure and (b) events for each trial during imaging transfer session.

Images were acquired using gradient-echo echo planar image (EPI) acquisition on a Siemens 3T Verio Scanner using a 32 channel RF head coil, with 2 s repetition time (TR), 30 ms echo time (TE), 79° flip angle, and 20 cm field of view (FOV). The experiment acquired 34 axial slices on each TR using a 3.2 mm thick, 64×64 matrix. This produces voxels that are 3.2 mm high and 3.125×3.125 mm^2. The anterior commissure-posterior commissure (AC-PC) line was on the 11th slice from the bottom scan slice. Acquired images were pre-processed and analyzed using AFNI [21,22]. Functional images were motion-corrected using 6-parameter 3D registration. All images were then slice-time centered at 1 s and co-registered to a common reference structural MRI by means of a 12-parameter 3D registration and smoothed with a 6 mm full-width-at-half-maximum 3D Gaussian filter to accommodate individual differences in anatomy.

4. Results

Out of 48 participants, data from 9 participants were excluded from the final analysis due to various reasons. One participant (explanatory/structural) did not show up for the transfer session. Seven participants (1 from explanatory/structural, 3 from non-explanatory/structural, 3 from non-explanatory/procedural) were excluded due to a poor learning performance (mean number of errors per problem was greater than 2). One further participant (explanatory/procedural) was removed due to a poor transfer performance (0% of correct responses in algebraic transfer problems). This resulted in a total of 39 participants (10 explanatory/structural, 9 explanatory/procedural, 10 non-explanatory/structural, and 10 non-explanatory/procedural).

4.1. Behavioral results

For the learning session, to identify initial and later performance differences among conditions, we focused on the first 16 linearize problems (blocks 1–2) and the last 16 problems (blocks 4–5). We measured the number of problems where participants "correctly" performed linearization steps without a single error or hint. A 2×2 analysis of variance (ANOVA) was performed to determine the effect of verbal instruction and instructional example on the initial and later periods of learning. There was a significant main effect of verbal instruction on the initial period of learning, $F(1, 35)=4.97$, $MSE=2624.45$, $p=0.032$; however, the effect of verbal instruction disappeared on the later period of learning, $F(1, 35)=0.55$, $MSE=4.25$, $p=0.462$. In the initial phase of learning, participants who were given explanatory verbal instruction ($M=68.75$, $SD=17.43$) solved around 16% more problems correctly than those who were just told about the interface and left to induce a rule for themselves ($M=52.19$, $SD=26.30$). This implies that the benefit of provision of explanatory verbal instruction (vs. non-explanatory instruction) existed only in the initial period of learning. There was neither a main effect of instructional example, $Fs<1$, nor an interaction effect, $Fs<1$ in both initial and later periods of learning.

For the transfer session, problem type (graphic vs. algebraic) was included as a within-subjects variable. A $2 \times 2 \times 2$ mixed ANOVA was performed on transfer performance as measured by the number of correctly solved problems.[2] Regarding the overall transfer performance, there were no main effects of verbal instruction, $F(1, 35)=$

[2] Transfer problems involved only linearization steps, and not propagation; the number of correctly solved problems was the same as the number of correctly performed linearization steps.

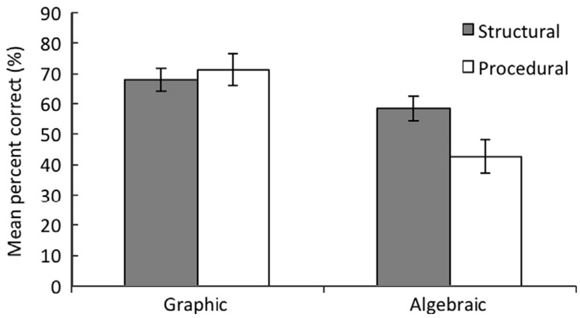

Fig. 6. Mean percentages of correctly solved problems in the transfer session. Error bars represent 1 standard error of mean.

Table 2
Locations of predefined regions.

ROI	Brodmann area (s)	Volume	Talairach coordinates (x, y, z)
PPC	7 and 39	12.8 mm. (high) by 15.6×15.6 mm^2	$\pm 23, -63, 40$
LIPFC	9 and 46	12.8 mm. (high) by 15.6×15.6 mm^2	$\pm 43, 23, 24$
PSPL	7	12.8 mm. (high) by 12.5×12.5 mm^2	$\pm 19, -68, 55$
HIPS	40	12.8 mm. (high) by 12.5×12.5 mm^2	$\pm 34, -49, 45$
AG	39	12.8 mm. (high) by 12.5×12.5 mm^2	$\pm 41, -65, 37$
RLPFC	10	12.8 mm. (high) by 205.078125 mm^2	$\pm 33, 47, 8$

2.82, $MSE=2045.15$, $p=0.102$, nor instructional example, $F(1, 35)=1.05$, $MSE=758.03$, $p=0.314$, nor an interaction effect between these two factors, $F<1$. However, there was a significant main effect of problem type, $F(1, 35)=39.31$, $MSE=7038.99$, $p<0.001$. Fig. 6 shows the mean percentages of correctly solved problems in the transfer session. Participants overall performed better on graphic problems ($M=69.68$, $SD=17.60$) than on algebraic problems ($M=50.96$, $SD=25.89$). More interestingly, the problem type interacted with the type of instructional example provided in the learning session, $F(1, 35)=9.78$, $MSE=1751.73$, $p=0.004$. For graphic problems, there was no mean difference between the two example groups, but for algebraic problems, the structural group ($M=58.44$, $SD=24.08$) solved more problems correctly than the procedural group ($M=43.09$, $SD=25.99$). Problem type did not interact with verbal instruction, $F(1, 35)=4.05$, $MSE=724.66$, $p>0.05$. There was no three-way interaction, $F(1, 35)=1.79$, $MSE=319.79$, $p=0.19$.

4.2. Imaging analysis – predefined regions

Because of their importance in past research, we chose the PPC and LIPFC from the ACT-R theory [2] and the PSPL, HIPS, and AG from Dehaene et al. [26] plus the RLPFC [4,72]. Table 2 shows the locations of these regions. For each region, we extracted an estimate of the engagement during the solution period and feedback period. A design matrix was constructed consisting of 8 regressors of interest. These were obtained by crossing problem type (graphic vs. algebraic) with correctness (correct vs. incorrect), and splitting each trial into two periods.[3] The trial periods consisted of the variable solution period from stimulus onset up until response completion, and the feedback period (2 s for correct response and 10 s for incorrect response). These 8 regressors were created by convolving the boxcar functions of the described variables with the standard SPM hemodynamic response function [32]. A baseline model consisted of an order-4 polynomial to remove general signal drift. This analysis yielded 8 beta weights for each participant. We performed two different sets of ANOVAs. In the first analysis, we will focus on the beta weights for the solution period only and in the second analysis we will look at the contrast between the solution and feedback periods for incorrect trials.

For solution period we examined 2 between-subjects variables and 3 within-subjects variables. The between-subjects variables were verbal instruction (explanatory vs. non-explanatory) and instructional example (structural vs. procedural). The within-subjects variables were problem type (graphic vs. algebraic), correctness (correct vs. incorrect), and hemisphere (left vs. right). For each of the six predefined regions, we performed a series of $2 \times 2 \times 2 \times 2 \times 2$ mixed ANOVAs. We were interested in main effects but we did the more complex analysis to detect whether there were any complicating interactions. Although there were some interactions with hemisphere, in all cases these took the form of the effect being stronger in one hemisphere than in the other but not changing the direction of main effect. There were no other significant interactions. Table 3 shows the main effects found in the six predefined regions. We did not find any brain regions that showed a reliable effect of verbal instruction. In contrast, there were significant effects of the instructional example: the structural group (vs. procedural) showed significantly higher engagement in the LIPFC (structural $M=0.52\%$, procedural $M=0.26\%$) and the AG (structural $M=0.17\%$, procedural $M=-0.05\%$) regions. The other four regions showed non-significant effects in the same direction. Regarding the problem type, all of the regions but the PSPL showed significantly greater engagement for algebraic problems (vs. graphic). Regarding the effect of correctness, there was greater activation for correct trials in the PPC (correct $M=0.95\%$, incorrect $M=0.91\%$) and the PSPL (correct $M=2.16\%$, incorrect $M=2.04\%$). The HIPS showed a marginal effect in the same direction ($p=0.057$).

In order to look at the effect of period (solution vs. feedback) we also performed a separate $2 \times 2 \times 2 \times 2 \times 2$ mixed ANOVA by entering period as a within-subjects factor and excluding the correctness factor. We focused on incorrect trials only because the brief feedback period for correct trials yielded poor estimates of effects. The main effects are shown in the rightmost column of Table 3. The PPC, PSPL, and HIPS showed significantly greater engagement during the solution period than during the feedback period, whereas the LIPFC, AG and RLPFC regions showed the reverse pattern.

In addition to the ANOVA, we also looked at the mean percent change from baseline to examine the time courses of the six ROIs. For this analysis, we adopted an event-aligning method [61]. The event-aligning procedure enables us to consider the entire time course of varying lengths of trials. We used onset of the problem and onset of the feedback as behavioral markers and then aligned each scan from each trial interval to the mean length of the interval. Therefore, in case of the short trial, the same scan was copied to multiple positions, whereas in the case of the long trials, the middle scans were deleted so that all trails had the same length of the interval. The mean length of correct trials was 14 scans and the mean length of incorrect trials was 16 scans.

Fig. 7 presents these time courses for the six ROIs (averaged over left and right hemispheres). The correct responses (Fig. 7a) consisted of 14 solving scans, 1 feedback scan, and 6 post-feedback scans. The incorrect responses (Fig. 7b) consisted of 16

[3] We removed the cases where participants were timed out. There were a total of 148 timed out trials and they were less than 6 % of the total cases used for this study.

Table 3
Main effects of 2 × 2 × 2 × 2 × 2 ANOVA with beta weights of the solution period (F's with dfs of 1 and 35). The rightmost column reports a main effect of period with incorrect trials.

ROI	Verbal instruction [Exp > Non]	Instructional example [Struc > Proc]	Problem type [Alg > Graph]	Correctness [Cor > Inc]	Hemisphere [Left > Right]	Period [Sol > Feed]
PPC	(0.95)	1.96	4.56*	6.23*	13.15**	9.34**
LIPFC	0.03	5.69*	22.64***	(0.01)	0.78	(10.93)**
PSPL	(0.75)	0.72	0.3	17.56***	0.21	38.18***
HIPS	0.31	1.65	8.10**	3.86	34.37***	6.43*
AG	(1.24)	4.64*	12.45**	0.92	(3.33)	(49.61)***
RLPFC	0.14	2.65	26.77***	−0.06	2.74	(24.24)***

Values in parentheses indicate an opposite direction of the stated main effect.

* Significance level: $p < 0.05$.
** Significance level: $p < 0.01$.
*** Significance level: $p < 0.001$.

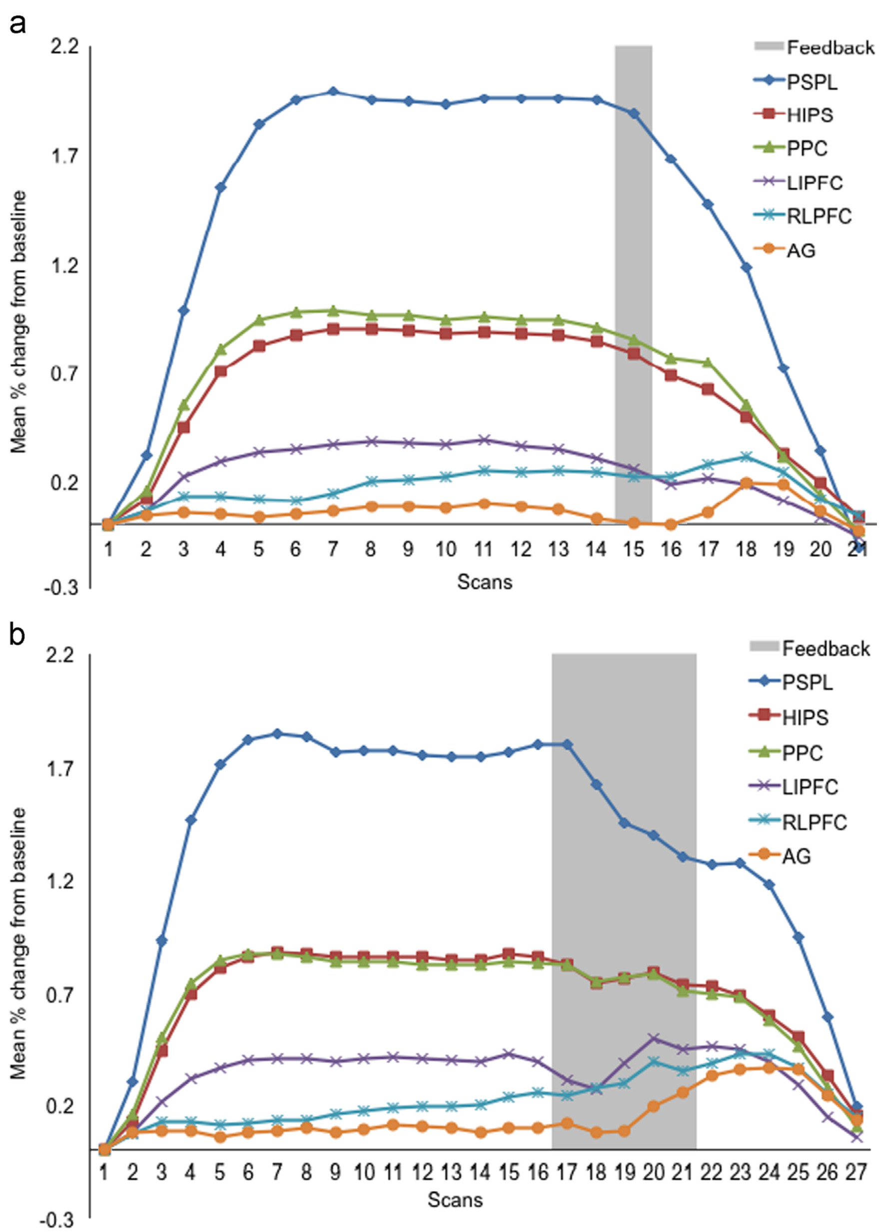

Fig. 7. Time courses for the six ROIs (averaged over hemisphere) of (a) correct responses and (b) incorrect responses. Each marker represents a 2-s TR.

solving scans, 5 feedback scans, and 6 post-feedback scans. The PPC, PSPL, and HIPS showed decreased engagement after the feedback regardless of correctness of the trials, whereas the LIPFC, AG and RLPFC regions showed increased engagement after the feedback, especially in case of incorrect trials. Fig. 8 shows how time courses differ between two example conditions

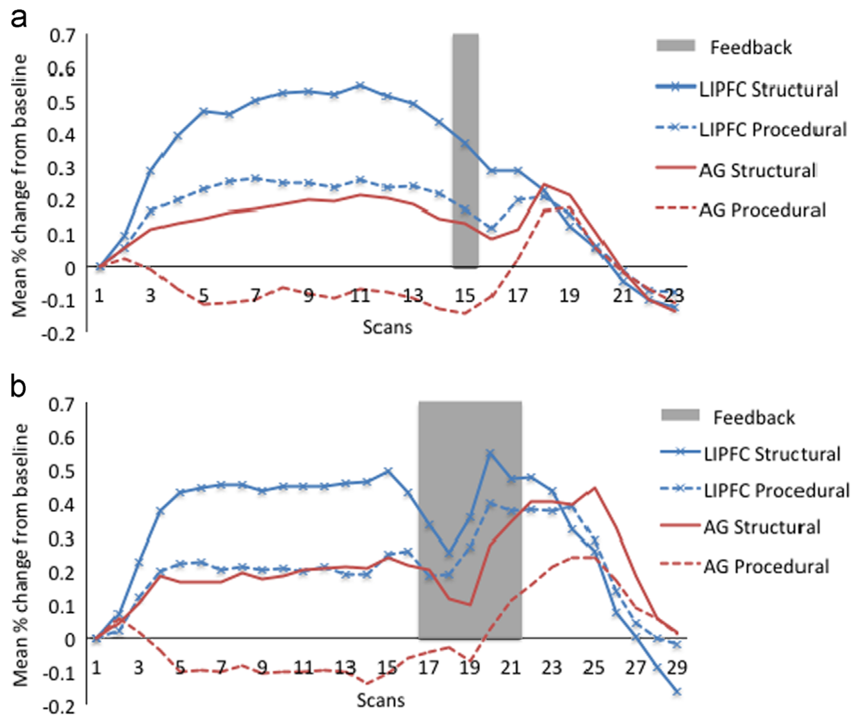

Fig. 8. Time courses for the LIPFC and AG regions (averaged over hemisphere) between structural and procedural example conditions in (a) correct responses and (b) incorrect responses. Each marker represents a 2-s TR.

Table 4
Regions identified from exploratory analysis using the graphic versus algebraic contrast: main effects of 2 × 2 × 2 ANOVA with beta weights of solution period (F's with *dfs* of 1 and 35). The rightmost column reports a main effect of period with incorrect trials.

Region		Brodmann area (s)	Coordinates (*x*, *y*, *z*)	Voxel Count	Verbal instruction [Exp > Non]	Instructional example [Struc > Proc]	Correctness [Cor > Inc]	Period [Sol > Feed]
(a) Algebraic > Graphic								
1	Medial/Superior Frontal Gyrus	8/32	−2, 21, 48	272	0.07	2.69	0.01	(50.68)***
2	L Middle Frontal Gyrus	6	−25, 3, 58	178	1.12	0.77	1.60	7.42*
3	L Precuneus	7	−6, −67, 45	173	1.05	1.65	6.71*	(0.64)
4	R Middle Frontal Gyrus	6	37, 6, 43	62	(0.72)	2.48	0.32	(36.49)***
5	L Inferior/Superior Parietal Lobule L Angular Gyrus L Supramarginal Gyrus	40/7/39	−39, −59, 45	545	(1.16)	3.37	2.73	(7.38)*
6	L Inferior/Middle/Superior Frontal Gyrus	10/46/45/9	−33, 35, 21	1005	0.24	2.77	(0.38)	(14.65)***
7	R Middle/Superior Frontal Gyrus	10/46/9	33, 32, 25	166	0.03	4.66*	0.00	(4.05)
8	R Inferior Parietal Lobule	40	46, −47, 45	16	0.01	4.32*	0.69	(22.98)***
9	L Insula L Putamen L Caudate	13	−24, 15, 10	180	0.11	3.89	0.09	(8.68)**
10	R Inferior/Middle/Superior Frontal Gyrus	10	33, 48, 4	49	(0.01)	4.63*	(0.03)	(33.03)***
11	R Caudate		14, 11, 12	33	(0.01)	6.13*	0.11	(2.18)
12	L Thalamus		−9, −10, 16	33	(0.21)	4.32*	1.19	(2.90)
13	R Insula	13	29, 22, 5	37	(0.05)	5.00*	(1.31)	(22.55)***
14	L Middle Temporal Gyrus	37	−53, −39, −1	20	(0.55)	(0.15)	1.05	(39.32)***
15	R Inferior Occipital/Fusiform Gyrus	17/18	28, −93, −6	16	(0.30)	2.70	4.94*	3.01
16	L Fusiform/Inferior Temporal Gyrus	37/20	−41, −57, −13	54	0.11	2.60	6.79*	(20.54)***
(b) Graphic > Algebraic								
17	L Cuneus/Middle Occipital Gyrus	18	−11,−96,14	21	(0.18)	0.14	4.50*	(56.11)***

Values in parentheses indicate an opposite direction of the stated main effect.

* Significance level: $p < 0.05$.
** Significance level: $p < 0.01$.
*** Significance level: $p < 0.001$.

in the LIPFC and AG regions. In both regions, the structural group showed higher engagement than the procedure group in both correct (Fig. 8a) and incorrect responses (Fig. 8b). It is particularly striking how the AG only shows a positive response during the solving phase only for the structural example condition.

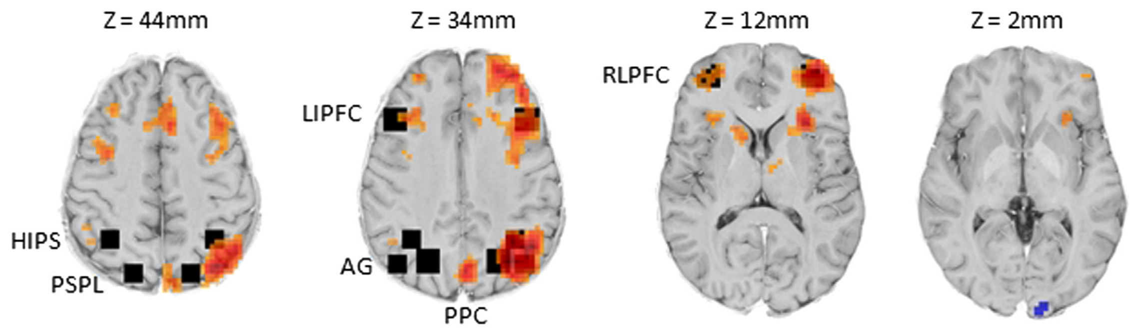

Fig. 9. Exploratory regions showing a significant effect of the contrast between activity during graphic problem solving and algebraic problem solving. The black squares show the six predefined regions in the experiment. The red regions show greater activity solving algebraic problems. The one blue region shows greater activity solving graphic problems. The *z* coordinates for a brain slice (radiological convention: image left = participant's right) is at $x=y=0$ in Talairach coordinates. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.3. Imaging analysis – exploratory regions

Three exploratory analyses were performed looking for significant effects of verbal instruction (explanatory vs. non-explanatory), instructional examples (structural vs. procedural), and problem type (graphic vs. algebraic) for the correct solution period. These analyses looked for regions of at least 13 contiguous voxels that showed a voxel-wise significance of 0.0001 for the difference between the contrast of the described variables. Using these values results in a brain-wise significance estimated to be less than 0.01 by simulation [21,22]. There were no significant effects of the two between-subjects variables of verbal instruction or instructional examples, but 17 regions emerged for the within-subjects variable of problem type (see Table 4).

Fig. 9 displays the regions showing a significant effect of problem type. The red regions show greater activity while solving algebraic problems whereas the one blue region shows greater activity while solving graphic problems. To illustrate how the exploratory regions overlap with our predefined regions, black squares were used to represent the six predefined regions in the experiment. Many of the 16 regions showing greater activity while solving algebraic problems overlapped with our predefined regions. This includes a posterior region (5) that overlaps with the left predefined PPC, HIPS, and AG and anterior regions (6, 7, and 10) that overlap with the left and right LIPFC and RLPFC.

For each of the 17 regions, Table 4 gives the effects of verbal instruction (explanatory vs. non-explanatory), and instructional example (structural vs. procedural) and the within-subjects variable was correctness (correct vs. incorrect). These factors were orthogonal to the problem type that was used to select the regions. Consistent with our predefined region analysis, there was no reliable effect of verbal instruction in any of the 17 regions. In contrast, 6 regions showed greater activity in the structural example condition than in the procedural example condition. This includes the region (7) that overlaps with the right predefined LIPFC and the region (10) that overlaps with the right predefined RLPFC. The region (5) that overlaps with our predefined AG had only a marginally significant effect of instructional example. The last column of Table 4 gives the effect of period. Consistent with our predefined region analysis, the region (5) that overlaps with the left AG and region (10) that overlaps with the right RLPFC showed greater activity during the feedback period than during the solution period.

4.4. Predicting future learning

Does engagement of different brain regions predict future learning? To examine whether engagement of certain brain regions predicted future performance, we looked at the relationship

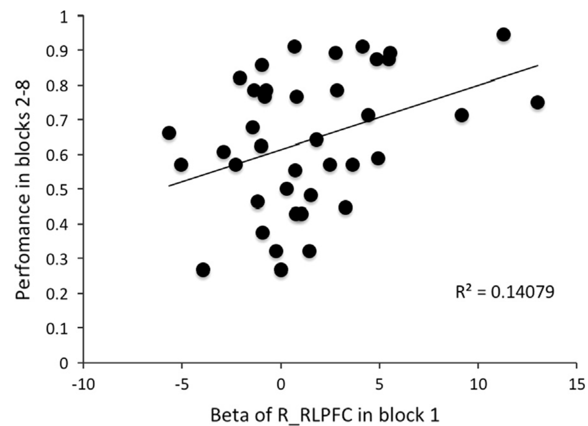

Fig. 10. Relationship between brain activity of the right RLPFC and the performance in blocks 2–8.

between the activity in the regions on the first block and performance on later blocks (blocks 2–8). The predictive variables were the beta weights obtained from the solution period for each of the predefined and exploratory brain regions for the first block. Because the accuracy in the first block was low (overall $M=34\%$, graphic $M=46\%$, algebraic $M=22\%$), we ignored correctness and problem type. The criterion variables were the proportion of correct responses for graphic problems, algebraic problems, and both problem types on blocks 2–8. Among the predefined regions only the right RLPFC showed a significantly positive correlation ($r=0.375$) with the future performance for both types of problems. Fig. 10 shows the relationship between brain activity in the right RLPFC during the first block and performance of the future blocks (2–8). Consistently, among the exploratory regions, only region 10 showed a significant correlation ($r=0.342$) and this region overlaps with the right predefined RLPFC.

5. General discussions

This paper investigated the effect of various instructional methods on learning and transfer by examining both behavioral and imaging data. We varied the way students learned by manipulating the verbal instruction (explanatory vs. non-explanatory) and what aspects of the problems the instructional examples illustrated (structural vs. procedural). Students were then tested with two different types of transfer problems (graphic vs. algebraic). Both behavioral and imaging data showed that verbal instruction did not have an effect on transfer while instructional example had an effect. Students who learned with structural examples showed better transfer to algebraic problems than those who learned with procedural examples. These students also showed greater engagement in the LIPFC and AG regions. The results were largely consistent when we tested an effect of instructional conditions on the exploratory regions identified from the contrast of graphic versus algebraic problems. The exploratory right RLPFC was additionally identified as showing greater activity among the structural group than the procedural group. The structural examples used algebraic knowledge to illustrate the structure behind solutions whereas the procedural examples illustrated the sequence of operations for computing the answers. Greater dependency on algebra knowledge in the structural condition perhaps increased activity in the LIPFC and AG regions. Many studies have shown that the LIPFC is involved in advanced tasks such as algebra, geometry, or calculus (e.g., [42,54,57,64]). More specifically, this region appears to support retrieval of arithmetic and semantic facts in mathematical problem solving [23,48]. Also, the greater activity in the AG from the structural condition perhaps reflects the verbal processing of numbers [26]. Structural examples were likely to cause larger working memory loads by simultaneously illustrating the coefficient and constant terms while procedural examples illustrated the calculation of each term separately. Simultaneous computation of two terms perhaps increased the working memory load and a tendency for verbalization among the structural students.

The two kinds of instructional examples resulted in different levels of transfer performance depending on the type of test problems. Graphic problems did not necessarily require use of algebra knowledge; thus the procedural students were able to solve these problems as well as the structural students. In contrast, to correctly solve algebraic problems, it is very helpful to understand how the data-flow diagram is equivalent to an algebraic expression. Accordingly, the structural participants were better able to solve algebraic problems than the procedural participants. The PPC, LIPFC, HIPS, AG, and RLPFC showed greater activity while solving algebraic problems than graphic problems. This is perhaps because algebraic problems require more cognitive and metacognitive demands than graphic problems. Only one exploratory region showed greater engagement while solving graphic problems than algebraic problems. This region was part of the occipital cortex that overlaps with BA18. For graphic problems,

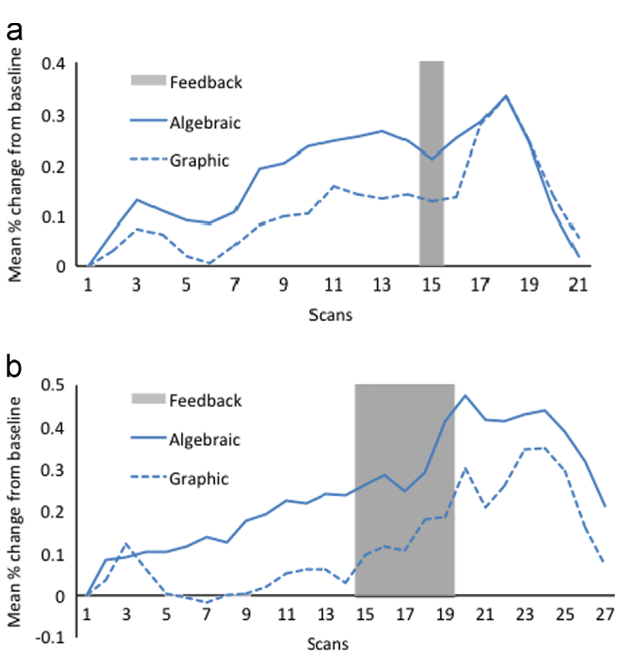

Fig. 11. Time courses for the right RLPFC of graphic and algebraic problem type in (a) correct responses and (b) incorrect responses. Each marker represents a 2-s TR.

parsing of the complex graphic structure is critical for determining which part of the diagram to include or exclude in the computation. This may have produced increased activity in this region.

We also looked at how our predefined and exploratory regions are involved during the solution and feedback periods to examine cognitive or metacognitive roles of these regions. According to the prior studies [4,72], cognitive networks were more engaged during the solution period than during the feedback period while metacognitive networks were more engaged during the feedback period. Overall, the results of our study were consistent with the previous studies. In our study, the LIPFC, AG, and RLPFC regions showed metacognitive patterns whereas the PPC, PSPL, and HIPS regions showed cognitive patterns. In both studies by Anderson et al. and Wintermute et al., the AG and RLPFC showed a strong metacognitive pattern whereas the parietal regions including the PPC and PSPL showed strong cognitive patterns. The LIPFC and HIPS regions showed cognitive patterns in Wintermute et al., but a mixed pattern in Anderson et al. Lastly, we looked at the relation between the brain activity during block 1 and behavioral performance for each problem type on blocks 2–8. The results showed that the right RLPFC predicted overall future performance.

Across the various analyses, we have repeatedly found that the right RLPFC was related to the transfer results; it showed an effect of instructional example (structural > procedural), an effect of problem type (algebraic > graphic), an effect of period (feedback > solution), and a positive correlation with future performance. Fig. 11 shows time courses of the right RLPFC for the correct responses (a) and incorrect responses (b). As in Fig. 7, we used the event aligning procedure. Regardless of the correctness of the trials, the algebraic problem type showed greater engagement than the graphic problem type. Also, the post-feedback engagement was greater than the pre-feedback engagement and this pattern was especially distinctive in case of incorrect trials (Fig. 11b) than correct trials (Fig. 11a). It appears that RLPFC is the critical region that mediates the effect between the example manipulation and the transfer outcome in our study. Mastery of the transfer problems used in this study requires understanding of multiple types of relations: between boxes and arrows (e.g., figuring out which parts are connected, propagating numbers from top to bottom), between diagrams before and after linearization, and between the data-flow diagrams and algebra knowledge. Our example manipulation affected how well participants could appreciate these relations. The structural example was designed to help understand these relations by focusing on the problem structure behind the solutions and this seemed to promote greater activation in the RLPFC area. The RLPFC has been shown to play a role in the integration of multiple relations (e.g., [19]; Wendelken and Bunge, 2010). The right RLPFC seems particularly specialized for processing visuospatial relations (e.g., [10]). The greater RLPFC activation in turn led to better transfer, especially for algebraic problems that required understanding of the relation between the diagrams and algebraic expressions. Therefore, perhaps higher engagement in this region implies better ability at relational reasoning in this domain. Desco et al. [28] similarly reported that math-gifted adolescents showed increased activation in the parietal and frontal regions and that this was associated with enhanced skills in visuospatial reasoning.

Putting both behavioral and fMRI results together, this study demonstrated that what was critical was understanding of what the examples illustrated, and not whether these examples were accompanied by verbal explanations. This finding is consistent with the previous study by Klahr and Nigam [38] who found that transfer performance depended on whether students mastered a skill and not whether there was direct instruction. Also this is consistent with the study by Lee et al. [45] who showed that successful learning depends on whether the hidden structure of a problem is revealed and not whether there is verbal instruction. All of these studies suggest that what students learned is more important than how they learned. Presence or absence of verbal instruction did not have an effect on transfer success or brain activity patterns. Rather, understanding of what the example illustrated had an effect on both transfer success and brain activity patterns. The structural example seemed to promote relational understanding, which was reflected in greater activation in the RLPFC. This increased activation in turn resulted in greater transfer performance. The structural students were better able to appreciate hidden problem structure by making a connection between the current domain and their algebra knowledge.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A100109 to Carnegie Mellon University. Preliminary reports of portions of this work were presented at the annual meetings of the Society for Research on Educational Effectiveness (Washington, DC, September, 2011). We thank Aryn Pyke and Trudy Buwalda for valuable comments on the paper. Correspondence may be directed to Hee Seung Lee, Department of Education, Yonsei University, 50 Yonsei-ro, Seodaemu-gu, Seoul, South Korea (hslee00@yonsei.ac.kr).

References

[1] Anderson JR. Human symbol manipulation within an integrated cognitive architecture. Cognit. Sci. 2005;29:313–42.
[2] Anderson JR. How Can the Human Mind Occur in the Physical Universe? New York: Oxford University Press; 2007.
[3] Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y. An integrated theory of mind. Psychol. Rev. 2004;111:1036–60.
[4] Anderson JR, Betts S, Ferris JL, Fincham JM. Cognitive and metacognitive activity in mathematical problem solving: prefrontal and parietal patterns. Cogniti. Affective Behav. Neurosci. 2011;11:52–67.
[5] Benoit RG, Gilbert SJ, Frith CD, Burgess PW. Rostral prefrontal cortex and the focus of attention in prospective memory. Cereb. Cortex 2012;22(8):1876–86.
[6] Binder JR, Desai RH, Graves WW, Conant LL. Whereis the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies Cereb. Cortex 2009;19:2767–96.
[7] Brunstein A, Betts S, Anderson JR. Practice enables successful learning under minimal guidance. J. Educ. Psychol. 2009;101:790–802.
[8] Buckner RL. Functional-anatomic correlates of control processes in memory. J. Neurosci. 2003;23:3999–4004.
[9] Bunge SA, Wendelken C, Badre D, Wagner AD. Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. Cereb. Cortex 2005;15(3):239–49.
[10] Bunge SA, Helskog HE, Wendelken C. Left, but not right, rostrolateral prefrontal cortex meets a stringent test of the relational integration hypothesis. NeuroImage 2009;46(1):338–42.
[11] Burgess PW, Quayle A, Frith CD. Brain regions involved in prospective memory as determined by positron emission tomography. Neuropsychologia 2001; 39:545–55.
[12] Burgess PW, Gilbert SJ, Okuda J, Simons JS. Rostral prefrontal brain regions (area 10): a gateway between inner thought and the external world? In: Prinz W, Sebanz N, editors. Disorders of Volition. Cambridge: MIT Press; 2006. p. 373–96.
[13] Carpenter TP, Franke ML, Jacobs VR, Fennema E, Empson SB. A longitudinal study of invention and understanding in children's multidigit addition and subtraction. J. Res. Math. Educ. 1998;29:3–20.
[14] Carroll WM. Using worked examples as an instructional support in the Algebra classroom. J. Educ. Psychol. 1994;86:360–7.
[15] Castelli F, Glaser DE, Butterworth B. Discrete and analogue quantity processing in the parietal lobe: a functional MRI study. Proc. Natl. Acad. Sci., USA 2006; 103:4693–8.
[16] Chen Z, Klahr D. All other things being equal: children's acquisition of the control of variables strategy. Child Dev. 1999;70:1098–120.
[17] Chi MTH, Bassok M. Learning from examples via self-explanations. In: Resnick LB, editor. Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser. Hillsdale, NJ: Erlbaum; 1989. p. 251–82.
[18] Christoff K, Gabrieli JDE. The frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. Psychobiology 2000;28:168–86.
[19] Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, Gabrieli JDE. Rostrolateral prefrontal cortex involvement in relational integration during reasoning. Neruoimage 2001;14:1136–49.
[20] Cooper G, Sweller J. The effects of schema acquisition and rule automation on mathematical problem-solving transfer. J. Educ. Psychol. 1987;79:347–62.
[21] Cox R. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 1996;29:162–73.
[22] Cox R, Hyde JS. Software tools for analysis and visualization of fMRI data. NMR Biomed. 1997;10:171–8.
[23] Danker J, Anderson JR. The role of prefrontal and posterior parietal cortex in algebra problem solving: a case of using cognitive modeling to inform neuroimaging data. Neuroimage 2007;35:1365–77.
[24] Dehaene S. The Number Sense: How the Mind Creates Mathematics. New York: Oxford University Press; 1997.
[25] Dehaene S, Cohen L. Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic. Cortex 1997; 33:219–50.
[26] Dehaene S, Piazza M, Pinel P, Cohen L. Three parietal circuits for number processing. Cognit. Neuropsychol. 2003;20:487–506.
[27] Desai RH, Binder JR, Conant LL, Mano QR, Seidenberg MS. The neural career of sensorimotor metaphors. J. Cognit. Neurosci. 2011;23(9):2376–86.
[28] Desco M, Navas-Sanchez FJ, Sanchez-Gonzalez J, Reig S, Robles O, Franco C, Guzman-De-Villoria JA, Garcia-Barreno P, Arango C. Mathematically gifted adolescents use more extensive and more bilateral areas of the fronto-parietal network than controls during executive functioning and fluid reasoning tasks. Neuroimage 2011;57:281–92.
[29] Eger E, Sterzer P, Russ MO, Giraud AL, Kleinschmidt A. A supramodal number representation in human intraparietal cortex. Neuron 2003;37:719–25.
[30] Fleming SM, Huijgen J, Dolan RJ. Prefrontal contributions to metacognition in perceptual decision-making. J. Neurosci. 2012;32(18):6117–225.
[31] Fletcher PC, Henson RNA. Frontal lobes and human memory: Insights from functional neuroimaging. Brain 2001;124:849–81.
[32] Friston KJ. Statistical Parametric Mapping: The Analysis of Functional Brain. London: Academic; 2006.
[33] Hiebert J, Lefevre P. Conceptual and procedural knowledge in mathematics: an introductory analysis. In: Hiebert J, editor. Conceptual and Procedural Knowledge: the Case of Mathematics. Hillsdale, NJ: Erlbaum; 1986. p. 1–27.
[34] Hiebert J, Wearne D. Instruction, understanding, and skill in multidigit addition and subtraction. Cognit. Instr. 1996;14:251–83.
[35] Issacs EB, Edmonds CJ, Lucas A, Gadian DG. Calculation difficulties in children of very low birth weight: a neural correlate. Brain 2001;124:1701–7.
[36] Kamii C, Dominick A. The harmful effects of algorithms in Grades 1–4. In: Morrow LJ, Kenney MJ, editors. The Teaching and Learning of Algorithms in School Mathematics: 1998 Yearbook. Reston, VA: National Council of Teachers of Mathematics; 1998. p. 130–40.
[37] Kirschner PA, Sweller J, Clark RE. Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educ. Psychol. 2006;41:75–86.
[38] Klahr D, Nigam M. The equivalence of learning paths in early science instruction. Psychol. Sci. 2004;15:661–7.
[39] Koechlin E, Basso G, Pietrini P, Panzer S, Grafman J. The role of the anterior prefrontal cortex in human cogntion. Nature 1999;399:148–51.
[40] Koechlin E, Corrado G, Pietrini P, Grafman J. Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. Proc. Natl. Acad. Sci. USA 2000;97:7651–6.
[41] Koedinger KR, Aleven V. Exploring the assistance dilemma in experiments with cognitive tutors. Educ. Psychol. Rev. 2007;19:239–44.
[42] Krueger F, Spampinato MV, Pardini M, Pajevic S, Wood JN, Weiss GH, Landgraf S, Grafman J. Integral calculus problem solving: an fMRI investigation. Neuroreport 2008;19:1095–9.
[43] Kuhn D. Is direct instruction an answer to the right question? Educ. Psychol. 2007;42:109–13.
[44] Lee HS, Anderson JR. Student learning: what has instruction got to do with It? Ann. Rev. Psychol. 2013;64:445–69.
[45] Lee, H.S., Betts, S., & Anderson, J.R. (submitted-(a)). Revealing hidden structure by verbal instruction versus nonverbal scaffolding methods.
[46] Lee, H.S., Betts, S., & Anderson, J.R. (submitted-(b)). Do similar examples help learning? – Not in right context.
[48] Menon V, Rivera SM, White CD, Glover GH, Reiss AL. Dissociating prefrontal and parietal cortex activation during arithmetic processing. Neuroimage 2000;12:357–65.
[49] Molko N, Cachia A, Riviere D, Mangin JF, Bruandet M, Le Bihan D, Cohen L, Dehaene S. Functional and structural alterations of the intraparietal sulcus in a developmental dyscalculia of genetic origin. Neuron 2003;40:847–58.
[50] Naccache L, Dehaene S. The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. Cereb. Cortex 2001;11:966–74.
[51] Perry M. Learning and transfer: instructional conditions and conceptual change. Cognit. Dev. 1991;6:449–68.
[52] Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S. Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron 2004;44:547–55.
[53] Pinel P, Piazza M, Le Bihan D, Dehaene S. Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. Neuron 2004;41:983–93.
[54] Qin Y, Anderson JR, Silk E, Stenger VA, Carter CS. The change of the brain activation patterns along with the children's practice in algebra equation solving. Proc. Natl. Acad. Sci. USA 2004;101:5686–91.
[55] Ramnani N, Owen AM. Anterior prefrontal cortex. Insights into function from anatomy and neuroimaging. Nat. Rev. Neurosci. 2004;5:184–94.

[56] Rapp AM, Mutschler DE, Erb M. Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies Neuroimage 2012;63:600–10.
[57] Ravizza SM, Anderson JR, Carter CS. Errors of mathematical processing: the relationship of accuracy to neural regions associated with retrieval or representation of the problem state. Brain Res. 2008;1238:118–26.
[58] Reed SK, Dempster A, Ettinger M. Usefulness of analogous solutions for solving algebra word problems. J. Exp. Psychol.: Learning, Mem. Cognit. 1985;11:106–25.
[59] Rittle-Johnson B, Alibali MW. Conceptual and procedural knowledge of mathematics: does one lead to the other? J. Educ. Psychol. 1999;91:175–89.
[60] Rittle-Johnson B, Siegler RS, Alibali MW. Developing conceptual understanding and procedural skill in mathematics: an iterative process. J. Educ. Psychol. 2001;93:346–62.
[61] Rosenberg-Lee M, Lovett M, Anderson JR. Neural correlates of arithmetic calculation strategies. Cognit. Affect. Behav. Neurosci. 2009;9:270–85.
[62] Seghier ML, Fagan E, Price CJ. Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. J. Neurosci. 2010;30:16809–17.
[63] Skemp RR. Relational understanding and instrumental understanding. Arith. Teach. 1976;26:9–15.
[64] Sohn MH, Goode A, Koedinger KR, Stenger VA, Carter CS, Anderson JR. Behavioral equivalence does not necessarily imply neural equivalence: evidence in mathematical problem solving. Nat. Neurosci. 2004;7:1193–4.
[65] Strand-Cary M, Klahr D. Developing elementary science skills: Instructional effectiveness and path independence. Cognit. Dev. 2008;23:488–511.
[66] Sweller J. Cognitive load during problem solving. Cognit. Sci. 1988;12:257–85.
[67] Sweller J, Cooper GA. The use of worked examples as a substitute for problem solving in learning algebra. Cognit. Instr. 1985;2:59–89.
[68] Sweller J, Van Merriënboer JJG, Paas FGWC. Cognitive architecture and instructional design. Educ. Psychol. Rev. 1998;10:251–96.
[69] Tobias S, Duffy TM, editors. Constructivist Instruction: Success or Failure. New York, NY: Routledge; 2009.
[70] Wendelken C, Bunge SA. Transitive Inference: distinct Contributions of Rostrolateral Prefrontal Cortex and the Hippocampus. Journal of Cognitive Neuroscience 2010;22(5):837–47.
[71] Wendelken C, Nakhabenko D, Donohue SE, Carter CS, Bunge SA. “Brain is to thought as stomach is to??”: Investigating the role of rostrolateral prefrontal cortex in relational reasoning J. Cognit. Neurosci. 2008;20:682–93.
[72] Wintermute S, Betts SA, Ferris JL, Fincham JM, Anderson JR. Brain networks supporting execution of mathematical skills versus acquisition of new mathematical competence. Pub. Libr. Sci. (PLOS) ONE 2012;7:e50154.