Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS

- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

Symbolizing Algebraic Story Problems: Are Diagrams Helpful?

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
Bartel, Anna N.	University of Wisconsin – Madison	
Alibali, Martha W.	University of Wisconsin – Madison	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) 2021

Check type of content being submitted and complete one of the following in the box below:

- [x] If article: Name of journal, volume, and issue number if available
- [ ] If paper: Name of conference, date of conference, and place of conference
- [ ] If book chapter: Title of book, page range, publisher name and location
- [ ] If book: Publisher name and location
- [ ] If dissertation: Name of institution, type of degree, and department granting degree

Applied Cognitive Psychology, v35 n6

DOI or URL to published work (if available)

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** Institute of Education Scier through **[Grant number]** R305B15000 to **Institution]** University of W.The opinions expressed are those of the authors and do not represent views of the **[Office name]** U.S. Department of Educa or the U.S. Department of Education.

Symbolizing Algebraic Story Problems: Are Diagrams Helpful?

Diagrams and Symbolizing Story Problems

Anna N. Bartel and Martha W. Alibali
Department of Psychology, University of Wisconsin – Madison

Correspondence
Anna N. Bartel, Department of Psychology, University of Wisconsin – Madison,
1202 West Johnson Street, Madison, WI 53706-1611.
Email: anbartel@wisc.edu

The authors have no conflicts of interest to declare. The co-author has seen and agrees with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

Pre-registration page: https://osf.io/z2ctk
Data and materials: https://osf.io/k6gj7/?view_only=f1508dd2a4474fb197a454c26651d1f6

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Award #R305B150003 to the University of Wisconsin Madison. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education. The authors thank Vanesa Meneses, Helen Huang and Xiaoning Zhang for their help with data collection and coding.

Title: Symbolizing Algebraic Story Problems: Are Diagrams Helpful?

Abstract: Many people find that solving story problems is challenging. A key source of the difficulty is symbolizing problem situations. In two experiments, we tested whether diagrams promoted successful symbolization of two-operator story problems, and we tested whether diagrams that depicted the integration of operations were more beneficial than diagrams that depicted the two operations separately. Findings in the two experiments were similar but not identical. Overall, participants were more likely to generate conceptually correct equations—though not always in integrated form—in the presence of diagrams, and this pattern was driven by the beneficial effects of diagrams for participants with lower visuo-spatial abilities. Further, participants with more negative attitudes towards mathematics were more likely to generate integrated equations with diagrams that depicted the integration of operations. Thus, the effects of diagrams varied across subgroups of participants.

Keywords: diagrams, algebra, story problems, symbolization, individual differences, text and graphic comprehension

Symbolizing Algebraic Story Problems: Are Diagrams Helpful?

Students frequently encounter story problems in their mathematics instruction. Such problems are thought to help students build problem-solving skills and appreciate real-world applications of mathematics. Yet many students find story problems challenging (Reed, 1999). One key reason is the need to *symbolize* problem situations. Symbolization typically involves "translating" from a more grounded representation (e.g., a real-world situation described in words) to a symbolic equation. Given the crucial role of symbolization in mathematical problem solving, it is important for both theoretical and practical reasons to identify ways to support students' symbolization.

In this research, we consider whether *diagrams* can support people in symbolizing problem situations. Specifically, we ask whether diagrams that highlight mathematical relations can help people accurately symbolize story problems with multi-operation equations. Past research has focused on diagrams as a potential support for problem solving (e.g., Cooper et al., 2018), but to our knowledge, no research has focused specifically on diagrams and symbolization.

There are inconsistent findings regarding whether diagrams effectively support mathematical problem solving. Some studies have reported that diagrams are beneficial (e.g., Hegarty & Kozhevnikov, 1999; Hembree, 1992), whereas others have reported that diagrams can be detrimental (e.g., Berends & van Lieshout, 2009; Lee et al., 2009). These inconsistencies may arise because different types of diagrams support problem solving in different ways. In this research, we consider the effects of two different types of diagrams on symbolization of algebraic story problems.

Solving and Symbolizing Story Problems

Simple story problems can often be solved accurately using informal approaches (see, e.g., Gvozdic & Sander, 2020; Koedinger & Nathan, 2004); however, informal approaches often fail when applied to more complex story problems (Koedinger et al., 2008). Formal approaches typically require several steps, including identifying quantitative elements in the text and accurately representing the relations among those elements in symbolic form.

When asked to symbolize complex story problems, people often have difficulties (e.g., Mayer, 1982; Nathan et al., 1992). Heffernan and Koedinger (1997) reported that symbolizing a two-operator story problem is more difficult than symbolizing both operations separately, a pattern they termed the "composition effect." They argued that integrating two operations into single structure requires substituting a sub-equation into another equation, a process that is difficult and error-prone.

If difficulties in symbolization arise due to the challenges of integrating mathematical relations, then external supports for such integration should be beneficial. One potential form of support is diagrams that depict the operations and their integration.

Combining Text and Visual Representations

A large body of research showcases the benefits of combining text and visual representations for performance and learning (e.g., Butcher, 2006; Carney & Levin, 2002; Chu et al., 2017; Mayer & Moreno, 2003; Moreno, 2005; Plass et al., 2014). These benefits are the empirical basis for the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2005), which specifies the conditions under which people learn better from text and visuals combined than from text alone. The CTML is based on the assumption that verbal and visual information are processed in different subsystems of a limited-capacity system (see, e.g., Chandler & Sweller, 1991; Paas et al., 2003). According to the CTML, people perform best when relevant information is presented in both text and visuals, because they use generative processing to draw connections across modalities. Importantly, the CTML holds that the scale of the benefits of visual representations should depend on the extent to which they elicit generative processing.

An alternative theory, the Integrative model of Text and Picture Comprehension (ITPC), holds that people construct a unified mental model based on both text and visual representations (Schnotz, 2002, 2014; Schnotz & Bannert, 2003; see also Glenberg & Langston, 1992). This mental model is continuously updated as information is encoded and processed; therefore, its content depends on how people perceive and understand the relations between the text and the visual representations. From this perspective, visual representations are not always beneficial because different combinations of text and visual representations lead people to construct differing mental models, and some mental models are less effective at supporting performance.

In solving complex story problems, people may attend to certain information in the text based on their processing of the visual representations. In line with this view, we propose that diagrams that depict different elements of the problem structure should lead people to focus on different aspects of the text, and consequently to symbolize the problems differently. In brief, we hypothesize that, if integrating operations is a key source of people's difficulties in symbolizing story problems, then diagrams that directly depict this integration should support accurate symbolization more than diagrams that depict the operations separately.

Individual Differences in Use of Visual Representations

Certain individual differences may also influence how much people rely on visual representations and whether they benefit from them. People with weak mathematics abilities find symbolizing story problems to be especially challenging (Montague et al., 1991), and as such, they may find diagrams particularly beneficial. Indeed, Booth and Koedinger (2012) found that low-ability eighth graders performed better on story problems with diagrams than on comparable problems presented in text alone. However, other work has shown that visual representations can actually be detrimental for people with lower mathematics ability (e.g., Cooper et al., 2018).

One key to integrating these mixed findings may be whether the visual representations include extraneous features, sometimes called “seductive details” (see Rey, 2012). Such details may add processing load that is particularly challenging for people with lower mathematics abilities, but that is readily handled by those with higher abilities. In the present study, we used diagrams without extraneous features, and we investigated whether such diagrams were beneficial for participants across a range of mathematics abilities.

The effects of visual representations may also depend on attitudes towards mathematics. Cooper and colleagues (2018) found that people who valued mathematics more highly benefited more from diagrams than people who valued mathematics less highly. In the present study, we investigate whether participants with varying attitudes towards mathematics reap differential benefits from diagrams for story problems.

We also consider visuospatial ability as a potential moderator of the effects of diagrams. Some studies have suggested that people with lower visuospatial ability derive greater benefits from visual representations than people with stronger visuospatial abilities (e.g., Vekiri, 2001). However, other studies have shown the opposite (e.g., Gyselinck et al., 2002), and a meta-analysis revealed a moderately stronger advantage of visuals for people with high spatial ability than those with low spatial ability (Höffler, 2010). In interpreting these findings, it is important to bear in mind that not all visual representations are “created equal”. Complex visual representations may require strong visuospatial abilities for successful interpretation, while simpler visual representations may be easily accessible, regardless of level of visuospatial abilities. In the present study, we investigated whether diagrams that highlight problem structure are differentially beneficial for participants with lower and higher visuospatial abilities.

Research Questions and Hypotheses

In brief, this study investigated whether diagrams help people successfully symbolize multi-operation algebraic story problems, and whether diagrams that directly depict the integration of operations support more successful symbolization than diagrams that do not depict that integration. To address these questions, we asked participants to symbolize problems either with no diagrams, with diagrams that depicted the integration of operations, or with diagrams that depicted each operation separately.

Building on past work, we hypothesized that, overall, participants who received problems with diagrams would perform better than participants who received problems without diagrams. Further, based on prior work suggesting that texts with different visual representations lead people to construct differing mental models (e.g., Schnotz, 2014), we hypothesized that participants who received diagrams that directly depicted the integration of operations would perform better than participants who received diagrams that depicted the operations separately.

We also examined whether the effects of diagrams were moderated by individual differences in attitudes towards mathematics, visuospatial abilities, and mathematics abilities.

We addressed these research questions in an experiment with undergraduate participants. We chose this participant sample given past work showing that undergraduates continue have challenges with algebraic story problems (e.g., Koedinger et al., 2008). We subsequently conducted a replication experiment to check on the robustness of our findings. The replication experiment also included additional items so we could examine whether the benefits of diagrams transferred to later items without diagrams.

Experiment 1

Method

Participants

Participants were 121 undergraduates (55% female, 45% male; M_{age} = 19.0 years, *SD* = 1.8) from large, public university in the Midwestern United States. Participants volunteered to participate in exchange for extra credit in introductory psychology. Per self-report, the sample was 71% White, 9% Asian, 6% Black or African-American, 5% Hispanic or Latinx, 1% Native American, and 8% multiracial. Most participants (92%) reported receiving their primary and secondary education in the US, and most (83%) reported being native English speakers.

Design

Participants were randomly assigned to one of three conditions (no diagram, discrete diagram, integrated diagram) in a between-subjects design. We also measured three continuous individual difference factors: mathematics ability, visuospatial ability, and attitudes towards mathematics.

Materials

Each participant solved eight story problems, each of which involved two operations (division-subtraction or division-addition). All problems used division so as to ensure that problems were sufficiently difficult to yield variability in performance. Two were very simple, result-unknown problems, in which the solution was the final quantity. The target problems for analysis were six start-unknown problems, in which the solution was one of the starting quantities.

The text of each problem was identical across conditions, and it included sufficient information for participants to symbolize and solve the problem. Each problem used a different cover story; all involved number variables (rather than weight or cost variables; Landy et al., 2014). The order of the start-unknown problems was randomized; the two result-unknown problems were always the third and sixth problems. Operation set (division-addition or division-subtraction) was randomized within these constraints.

We used two types of diagrams (see Figure 1). The *integrated* diagrams directly depicted the integration of the operations, whereas the *discrete* diagrams depicted the operations separately.

---Insert Figure 1---

Both types of diagrams are adaptations of "tape diagrams" (Chu et al., 2017; Murata, 2008), which represent quantities using rectangles that look like pieces of tape. We adapted the tape diagrams to represent the relations between operations more saliently. The diagrams contained no extraneous features.

All study materials, data files and analysis scripts are available on OSF.

Individual Difference Measures

Visuospatial ability was assessed using the Paper Folding Test (Ekstrom et al., 1976). Each item on this test presents a series of drawings of a square sheet of paper that is folded two or three times. The last drawing shows a hole punched in the folded paper. Participants select one of five drawings that depicts how the paper would look if unfolded. The test has an internal consistency of $\alpha = .84$, and it demonstrates convergent validity with measures of spatial orientation ability and object manipulation ability (Kozhevnikov & Hegarty, 2001).

Attitudes towards mathematics were assessed using an abridged version of Tapia and Marsh's (2004) Attitudes Toward Mathematics Inventory (ATMI), which includes four subscales: (1) self-confidence, defined as belief that one is good or bad at mathematics; (2) value, defined as belief that mathematics is useful or useless; (3) enjoyment, defined as liking or disliking of mathematics; and (4) motivation, defined as the tendency to engage in or avoid mathematical activities. The ATMI has an internal consistency of $\alpha = .96$ (Majeed et al., 2013).

*Mathematics abilit*y was assessed via self-reports of ACT or SAT mathematics subtest scores. For analysis, these scores were converted to percentiles based on the College Board's and ACT's percentile ranks for college-bound seniors. For participants who did not provide scores ($N = 8$), data were imputed using Multivariate Imputation by Chained Equations (MICE; Van Buuren & Oudshoorn, 2000), and the updated data set was made by 10 Gibbs sampling iterations.

Correlations among the individual difference measures are presented in Table 1.

---Insert Table 1---

Procedure

Participants took part in groups of up to 15 in a campus computer lab. Upon arrival, each participant was given the consent document and a booklet of problems. The experimenter read the consent document and instructions aloud. After providing consent, participants were instructed to open their booklets and begin. The first three pages of the booklet included three one-operator practice problems, which were used to familiarize participants with the diagrams. All participants, regardless of condition, received diagrams with these practice problems. These diagrams did not depict the integration of operations because the practice problems each involved only one operation.

After the practice problems, participants solved the experimental story problems. These problems were presented one per page, with the problem in the upper left and, for participants in the diagram conditions, a diagram in the upper right. Underneath the story problem, there were instructions for participants to (1) write an equation for the story problem, and (2) solve the equation and show their work. At the bottom of each page, participants were asked to rate how confident they felt about whether they solved the problem correctly. There were no time constraints.

After completing the problems, participants completed the Paper Folding Test and the Attitudes Toward Mathematics Inventory on a computer. Finally, participants completed the demographic form, which requested information about SAT or ACT math subtest scores. In all, the session lasted about 35 minutes.

Results and Discussion

Overall, participants were fairly accurate at solving the problems (M = 82% correct, SD = 23%) and highly confident in their solutions (M = 6.21, SD = 1.11 on a 7-point scale). Because our main focus is on symbolization, we do not further consider accuracy or confidence.

Planned Analyses

Our primary research question was whether participants' generation of correct, integrated equations varied as a function of condition (control, discrete diagram, integrated diagram) and the targeted individual difference factors (mathematics ability, visuospatial ability, and attitudes toward mathematics). We used the *lme4* package in *R* (Bates et al., 2015) to construct a generalized linear mixed-effects model, using the binomial family and the logit link function. The model included two orthogonal contrasts for condition: (1) diagram vs. no diagram (i.e., the diagram-general or DG contrast; coded -.67, .33, .33 for no diagram, discrete diagram, and integrated diagram, respectively), and (2) discrete vs. integrated diagram (i.e., the diagram-specific or DS contrast; coded 0, -.5, -.5, for no diagram, discrete diagram, and integrated diagram, respectively). The model also included operation set (division-addition [coded as -.5] or division-subtraction [coded as .5]) and the interactions of each of the contrasts with each of the individual difference factors (mean-centered).

Our initial model included by-item random intercepts and slopes for the interactions of each of the condition contrasts with mathematics ability, mathematics attitudes, and visuospatial ability. We also included by-subject random intercepts and slopes for operation set. This full model did not converge, so we followed the recommendations of Barr et al. (2013) to achieve convergence, first increasing the number of iterations, then removing lower-order random effects, then removing covariances among random effects, and finally fixing random intercepts to zero.[1] The final model and associated statistics are presented in Table 2. To evaluate simple effects for significant interactions, we estimated the effect of condition at one standard deviation below and one standard deviation above the mean of the relevant individual difference measure.

---Insert Table 2---

We hypothesized that participants would perform better with diagrams than without. As seen in Table 2, the *diagram vs. no diagram* contrast was not significant as a main effect (Figure 1, Appendix 1); however, it interacted significantly with each of the three individual difference factors (Figures 2 – 4, Appendix 1). Tests of simple effects revealed that participants with lower math ability performed better with diagrams than without, $b = 1.07$, $\chi2(1) = 10.12$, $p = .001$, $OR = 2.91$, but the reverse was true for participants with higher math ability, $b = -.83$, $\chi2(1) = 6.10$, $p = .01$, $OR = .43$. Similarly, participants with lower visuospatial ability performed better with diagrams than without, $b = .83$, $\chi2(1) = 9.23$, $p = .002$, $OR = 2.29$, but the reverse was true for participants with higher visuospatial ability, *b* = -.65, $\chi 2(1) = 4.30$, *p* = .04, *OR* = .52. Finally, participants with more positive attitudes towards mathematics performed better with diagrams than without, *b* = .93, $\chi 2(1) = 8.64$, *p* = .003, *OR* = 2.53, but the reverse was true for participants with more negative attitudes, *b* = -.75, $\chi 2(1) = 5.32$, *p* = .02, *OR* = .47. To foreshadow the results of Experiment 2, however, we did not find identical patterns, so we do not discuss these findings in detail here. In a later section, we analyze data from both experiments together.

We further predicted that participants would be more likely to generate correct, integrated equations with diagrams that depicted integrated operations than with diagrams that depicted the operations separately. Indeed, participants in the integrated diagram condition were more likely to generated integrated equations than participants in the discrete diagram condition (see Figure 1, Appendix 1). However, this contrast was qualified by a significant interaction with visuospatial ability (see Figure 3, Appendix 1). Participants with lower visuospatial ability performed similarly with both diagrams, *b* = -.10, $\chi 2(1) = 0.11$, *p* = .73; however, participants with higher visuospatial ability performed better with the integrated diagram than the discrete diagram, *b* = 1.03, $\chi 2(1) = 9.88$, *p* = .001, *OR* = 2.77. It is worth noting that, for participants with higher visuospatial ability, performance in the discrete diagram condition was actually *lower* than in the no-diagram condition. The pattern suggests that participants with higher visuospatial ability in the discrete diagram condition may have focused specifically on the *separateness* of the operations, which was saliently displayed in the discrete diagrams, so they were *less* likely to generate integrated equations. Again, however, we did not find identical patterns in Experiment 2, so we do not dwell on these findings here.

Experiment 2

Experiment 2 was a preregistered replication of Experiment 1 (OSF preregistration), and it also addressed one additional research question. Experiment 1 had revealed beneficial effects of diagrams for some subgroups of participants. However, the findings could not address whether exposure to diagrams would support performance on subsequent problems without diagrams. To address this question, in Experiment 2, we included three transfer items without diagrams at the end of the problem set.

Method

Participants

Participants were 123 undergraduates (55% female, 43% male, 2% unreported; M_{age} = 18.5 years, SD = 1.99) drawn from the same participant pool as Experiment 1. The data were collected in a different semester and at a different time within the semester (Experiment 1: late Spring, Experiment 2: early Fall). Per self-report, the sample was 67% White, 22% Asian, 1% Black or African American, 3% Hispanic or Latinx, 1% some other race or ethnicity, 4% multiracial, and 2% unreported. Most participants (84%) received their primary and secondary education in the US and most (83%) were native English speakers. The sample did not differ from that for Experiment 1 in percent female, $\chi^2(1, N = 242) = 0.07$, $p = .80$, percent native English speakers, $\chi^2(1, N = 241) = 0.04$, $p = .84$, or percent US-educated, $\chi^2(1, N = 242) = 3.17$, $p = .07$. Participants in Experiment 2 were slightly younger, on average, than participants in Experiment 1 ($M = 18.65$, $SD = 1.32$, vs. $M = 19.22$, $SD = 1.86$), which is unsurprising given the timing of data collection. For participants who did not provide ACT or SAT scores ($N = 7$), missing data were imputed, as in Experiment 1. Correlations among the individual difference measures are presented in Table 1.

Design and Materials

We used the same design and materials as Experiment 1, with one exception: after the initial set of story problems, but before the individual difference measures, participants completed a brief distractor task to clear working memory, and they then completed three transfer items that did not include diagrams. The task to clear working memory was a letter marking task that lasted approximately two minutes. Participants were given an array of 746 letters and asked to circle all the *a*s. The transfer items were two two-operator story problems and one three-operator problem.

Procedure

The procedure was identical to that of Experiment 1 except that, after finishing the first booklet of problems, participants were given a second booklet that included the letter marking task and three transfer problems without diagrams. After completing this booklet, participants completed the individual difference measures on a computer. The study lasted approximately 45 minutes.

Results and Discussion

Preregistered Replication Analysis

As in Experiment 1, participants solved most problems correctly (M = 84%, SD = 24%) and they were highly confident in their solutions (M = 6.44, SD = .91 on a 7-point scale). Because our main focus is on symbolization, we do not further consider accuracy or confidence.

Using the same analytic approach as for Experiment 1, we examined whether participants' generation of correct, integrated equations varied as a function of diagram condition and the individual difference factors. Table 3 presents model statistics for the initial set of story problems.

---Insert Table 3---

We hypothesized that participants would perform better with diagrams than without. As seen in Table 2, the *diagram* vs. *no diagram* contrast was significant as a main effect (Figure 1, Appendix 2), unlike in Experiment 1. Also unlike Experiment 1, the interactions of this contrast with visuospatial ability and attitudes toward mathematics were non-significant. The interaction with mathematics ability was significant (see Figure 2, Appendix 2); however, the shape of the interaction differed from that observed in Experiment 1. In this experiment, participants with *higher* mathematics ability performed better with diagrams than without, $b = 1.55$, $\chi2(1) = 37.54$, $p < .001$, $OR = 4.71$, and participants with lower mathematics ability performed similarly with diagrams and without, $b = -.25$, $\chi2(1) = 0.93$, $p = .33$. We do not discuss these findings in detail here, in favor of an exploratory analysis that combines both samples, presented below.

We next considered the *integrated diagram* vs. *discrete diagram* contrast. Unlike Experiment 1, this contrast was not significant, either as a main effect or in interaction with any of the individual difference factors.

Preregistered Analysis: Transfer Problems

We analyzed the transfer problems to examine whether the effects of diagrams transferred to subsequent items without diagrams. We first considered the two-operator transfer problems; model statistics are presented in Table 4.

---Insert Table 4---

We hypothesized that participants who had experience with diagrams would perform better on these transfer problems than participants who had not had such experience. The diagram-general contrast was not significant as a main effect; however, it interacted significantly with attitudes towards mathematics. For participants with more positive attitudes towards mathematics, those who had seen diagrams on the initial set of problems did indeed perform better on the two-operator transfer problems those in the control condition, *b* = 1.15, $\chi2(1)$ = 5.92, *p* = .01, *OR* = 3.16; however, for participants with more negative attitudes, those who had previously seen diagrams and those who had not performed similarly, *b* = -.52, $\chi2(1)$ = 1.05, *p* = .31 (see Figure 3, Appendix 2). The diagram-specific contrast was also not significant, either as a main effect or in interaction with any of the individual difference factors.

Next, we considered the more difficult, three-operator transfer problem; model statistics are presented in Table 5. Neither the diagram-general nor the diagram-specific contrast was significant, either as a main effect or in interaction with any of the individual difference factors.

---Insert Table 5---

Exploratory Analysis of Combined Data from Both Experiments

Given the differing findings across experiments, we examined whether the samples differed in any of the individual difference factors we assessed. Relative to participants in Experiment 1, participants in Experiment 2 had higher mathematics abilities (*M* = 90.25, *SD* = 10.68, vs. *M* = 87.26, *SD* = 12.77, *t*(242) = 1.99, *p* = .04) and more positive attitudes towards mathematics (*M* = 3.46, *SD* = .74, vs. *M* = 3.19, *SD* = .79, *t*(242) = 2.78, *p* = .005). The samples did not differ in visuospatial abilities (*M* = 12.21, *SD* = 3.79, vs. *M* = 11.68, *SD* = 3.83, *t*(242) = 1.09, *p* = .275).

In light of these differences, and in an effort to better understand the findings, we looked more closely at participants' error patterns. In some cases when participants failed to construct integrated equations, their responses indicated *incorrect symbolization of the conceptual structure of the problem*. As seen in Table 6, participants sometimes created symbolic equations with the wrong operations, created equations that involved distributing incorrectly, or misunderstood the problem structure entirely. In other cases, participants' errors revealed *correct symbolization of the conceptual structure of the problem*, despite not representing that structure in an integrated, two-operator equation. As shown in Table 6 for example, participants sometimes provided two separate one-operator equations, which, if combined, would yield a correct integrated equation, and they sometimes produced one-operator equations that reflected their having mentally performed one of the operations (rather than representing that operation in the equation).

---Insert Table 6---

Based on this framework, we recoded participants' responses into three categories: (1) correct, integrated equations, (2) conceptually correct but non-integrated equations, and (3) conceptually incorrect responses. We then analyzed the combined data from both experiments in two steps. First, we examined the likelihood that participants provided conceptually correct responses of any sort (i.e., correct, integrated equations *or* conceptually correct but non-integrated equations) as a function of experimental condition and the individual difference factors. This analysis reveals the role of diagrams in participants' *grasping the conceptual structure* of the problems. Second, we evaluated the likelihood that, when offering conceptually correct responses, participants offered integrated equations. This analysis reveals the role of diagrams in participants' *producing integrated equations*, given that they demonstrated understanding of the problems. For supplemental analyses of these dependent variables for each of the experiments separately, and for analyses of the original dependent variables for the combined data from both experiments, please see Appendix 3.

To examine the likelihood that participants provided conceptually correct responses, we used the same analytic approach as before, with likelihood of providing a conceptually correct response (i.e., a correct, integrated equation *or* a correct but non-integrated equation) as the dependent variable. Model statistics are presented in Table 7.

---Insert Table 7---

The diagram-general contrast yielded a significant main effect. As seen in Figure 2, participants in the diagram conditions had a higher probability of producing conceptually correct equations than participants who received no diagrams.

---Insert Figure 2---

This contrast was qualified by a significant interaction with visuospatial ability. To better understand the interaction, we evaluated simple main effects. As seen in Figure 3, participants with higher visuospatial abilities tended to provide conceptually correct responses, regardless of diagram presence, $b = -.01$, $\chi2(1) = .003$, $p = .95$, but participants with lower visuospatial abilities were more likely to provide conceptually correct responses when diagrams were present, $b = .66$, $\chi2(1) = 13.26$, $p < .001$, $OR = 1.93$. Thus, diagrams supported participants with lower visuospatial abilities in providing conceptually correct responses.

---Insert Figure 3---

We next consider the *discrete* vs. *integrated diagram* contrast. This contrast was not significant as a main effect (see Table 7); however, it interacted significantly with mathematics attitudes. As seen in Figure 4, participants with more positive attitudes were similarly likely to provide conceptually correct responses with both diagrams, $b = -.29$, $\chi2(1) = 1.03$, $p = .31$, but participants with more negative attitudes were more likely to provide conceptually correct responses when they received integrated diagrams than when they received discrete diagrams, $b = .77$, $\chi2(1) = 8.53$, $p = .003$, $OR = 2.16$. As seen in Figure 4, for participants with negative attitudes, performance in the discrete diagram condition was numerically *lower* than performance in the no-diagram condition, suggesting a possible negative effect of discrete diagrams for those with negative attitudes towards mathematics.

---Insert Figure 4---

We next considered the likelihood that participants *produced integrated equations*, given that they provided conceptually correct responses. For each participant, we calculated the proportion of conceptually correct responses on which they produced integrated equations, and we used a linear model to evaluate the likelihood of participants' producing integrated equations as a function of diagram condition and the individual difference measures. Conceptually incorrect responses (n = 146 in Experiment 1, n = 157 in Experiment 2) were excluded from this analysis. Nineteen participants (n = 14 from Experiment 1, n = 5 from Experiment 2) were excluded for not producing any conceptually correct responses.

Model statistics are presented in Table 8. There were no significant effects of either the *diagram vs. no diagram* or the *discrete vs. integrated diagram* contrast and no significant interactions involving either contrast. The sole significant effect was a main effect of mathematics ability, with participants with higher ability being more likely to produce integrated equations.

---Insert Table 8---

Taken together, the findings indicate that diagrams supported participants with lower visuospatial ability in grasping the conceptual structure of the story problems. Further, the integrated diagrams supported participants with negative attitudes in grasping the conceptual structure of the story problems more effectively than the discrete diagrams. However, given that participants grasped the conceptual structure of the problems, there was no evidence that diagrams specifically supported them in *generating integrated equations* to symbolize that structure.

General Discussion

In two experiments, we tested whether diagrams promoted successful symbolization of two-operator story problems, and we tested whether diagrams that depicted the integration of operations were more beneficial than diagrams that depicted the operations separately. In one experiment, we also examined whether the effects of diagrams persisted, once the diagrams were no longer present. In the following sections, we briefly summarize our main findings.

Did diagrams promote successful symbolization?

Our first major goal was to examine whether diagrams could scaffold participants' symbolization of complex story problems. Indeed, participants were more likely to generate conceptually correct equations—though not always in integrated form—in the presence of diagrams. Importantly, this pattern was driven by the beneficial effects of diagrams for participants with lower visuospatial abilities. Participants with higher visuospatial abilities did not need the support of diagrams; they tended to produce conceptually correct equations, whether or not a diagram was present.

The two separate experiments yielded mixed findings about whether diagrams specifically supported participants' generation of *integrated* equations. For this outcome measure, we observed beneficial effects for some subgroups in each experiment, and slightly detrimental effects for some subgroups in Experiment 1. In the combined analysis, we found no evidence that diagrams supported participants in generating integrated equations, given that participants grasped the conceptual structure of the problems. Thus, it seems that the primary locus of the effect of diagrams was on participants' generating *conceptually correct* equations—

which reveal understanding of the conceptual structure of the problems—rather than on generating *integrated* equations, *per se*.

The benefits of diagrams for symbolization for participants with lower visuospatial abilities were fairly robust, emerging both in the combined analysis and in Experiment 1. Given that visuospatial ability involves identifying and imagining visual and spatial relations among objects, our data suggest that diagrams supported participants with lower visuospatial abilities in understanding the spatial relations described in the problems (e.g., objects being divided into groups). For participants with lower visuospatial abilities in the control condition, it may have been challenging to imagine these spatial relations.

We also evaluated whether the effects of diagrams varied depending on participants' mathematics abilities and attitudes towards mathematics. Regarding attitudes towards mathematics, we found some evidence in each experiment that diagrams were especially beneficial for participants with more positive attitudes. People with more positive attitudes may make greater efforts to map between text and diagrams, and they may consequently gain more from the presence of the diagrams. Our findings align with those of Cooper et al. (2018), who also found that the benefits of diagrams were greater for individuals with more positive attitudes towards mathematics.

Regarding mathematics ability, the data were inconsistent, with conflicting patterns across experiments, and no significant interaction of diagram condition and mathematics ability in the combined analysis. We suggest that due to their schematic nature, the diagrams we used were beneficial for many participants, but due to their complexity, for some participants, the diagrams presented an additional cognitive load. Future work is needed to understand the conditions under which diagrams serve as a scaffold versus as an additional source of challenge.

Our main finding—that diagrams supported participants with lower visuospatial abilities in grasping the conceptual structure of the problems—aligns with other findings in the literature suggesting that diagrams are generally beneficial for mathematical problem solving. Although most past studies did not include measures of visuospatial abilities, several studies have reported beneficial effects of diagrams on learning and problem solving. For example, Nagashima et al. (2020) compared two versions of an intelligent tutoring system for learning about equations, one that included diagrams to represent the equations and one that did not. After a set of lessons with the intelligent tutoring system, fifth-graders who used the version with diagrams had higher scores on a posttest of conceptual understanding. Along similar lines, Chu et al. (2017) found that diagrams supported middle-school students in correctly solving symbolic equations. Both of these previous studies used schematic diagrams, as we did in the current work. Thus, our findings extend prior literature indicating that schematic diagrams can support learning and problem solving, by showing that such diagrams also support correct symbolization, at least for some subgroups of participants.

Were diagrams that depicted the integration of operations more beneficial than diagrams that depicted the operations separately?

A second major goal of our work was to examine whether different types of diagrams had differential effects. We hypothesized that participants would be more successful at generating integrated equations in the presence of the integrated diagram, which directly depicted the integration of operations, than in the presence of the discrete diagram, which depicted the operations, but not their integration. This hypothesis received limited support. The predicted pattern held for participants with high visuospatial abilities in Experiment 1, but it was not replicated in Experiment 2. The combined analysis revealed that participants with more negative attitudes were more likely to generate conceptually correct equations in the presence of the integrated diagram. However, the data pattern suggested that this could be due to a possible negative effect of the discrete diagram, rather than a positive effect of the integrated diagram; this possibility requires further study. By saliently displaying two separate operations, the discrete diagram might have actually discouraged some participants from generating integrated equations. Overall, the findings suggest that diagrammatically scaffolding the integration of operations may have been valuable for those who were adept at gleaning information from visuospatial representations and for those with negative attitudes, who might otherwise not engage deeply with mathematics. However, on the whole, there was not strong evidence that the integrated diagrams supported participants in integrating operations when symbolizing story problems.

One possible explanation for these findings is that the complex, integrated diagrams may have been daunting, especially for participants with lower visuospatial abilities. Given their complexity, participants with lower visuospatial abilities may have engaged less with the integrated diagrams, so they did not yield benefits over the simpler diagrams. Another possibility is that the differences between the two diagrams were not sufficiently great to yield differential effects for most participants. In this regard, it is worth noting that both types of diagrams were schematic, and the modified tape diagrams were similarly unfamiliar to all participants.

Did the benefits of diagrams persist once diagrams were no longer present?

A third goal of our work was to evaluate whether any potential benefits of diagrams persisted, once the diagrams were no longer present. In Experiment 2, participants with more positive attitudes towards mathematics who had previously solved problems with diagrams performed better on *text-only* two-operator problems than participants in the no-diagram condition. However, overall, we found relatively minimal transfer of the beneficial effects of diagrams to subsequent items without diagrams. The data suggest that the benefits of diagrams are applicable, mainly when diagrams are present.

Theoretical Implications

Our findings are compatible with multiple accounts of how text and diagrams are integrated. From the perspective of the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2005), both the discrete and integrated diagrams likely evoked generative processing, and therefore both yielded benefits. According to the Integrative model of Text and Picture Comprehension (ITPC; Schnotz, 2002; 2014), both text and diagrams provide information that contributes to mental model construction. Relative to the no-diagram condition, participants in the diagram conditions had access to more information for building and refining mental models, yielding benefits for performance.

The limited evidence that integrated diagrams were more beneficial than discrete diagrams is also compatible with both the CTML and the ITPC. From the perspective of the CTML, the integrated diagrams may have evoked more generative processing than the discrete diagrams, at least for some participants. From the perspective of the ITPC, participants who received the integrated diagrams may have constructed mental models in which the integration of the operations was more salient, and this may have supported more effective symbolization.

The combined analysis suggested that, although both types of diagrams helped many participants grasp the conceptual structure of problems, neither diagram helped participants to generate *integrated* equations when symbolizing the problems. Why might this be the case? One possibility is that the skills involved in *producing* integrated equations are not skills that diagrams can support. To generate an integrated equation, people need syntactic knowledge of algebraic formalisms, and they need to be able to apply this knowledge "on demand." Diagrams can support understanding of story situations, but they cannot help people know, for example, how to express the integration of division and subtraction symbolically.

Our findings highlight the range of factors that contribute to people's abilities to symbolize and to their engagement with diagrams. Our analyses revealed effects involving all three of the individual difference factors that we considered, and there are surely other relevant factors that we did not consider. For example, participants' prior experience with tape diagrams may have influenced engagement with the diagrams, but we did not measure this experience.

How can we conceptualize the combined effects of many factors in affecting performance? We suggest that the benefits of diagrams depend on the dynamic interaction of individual, contextual, and diagram-specific factors (for related arguments regarding children's strategy use, see Alibali et al., 2019). These factors may accumulate—or cancel one another out—to influence the likelihood that a given individual relies on a given diagram in a given context. Some factors push people *towards* using diagrams (e.g., positive attitudes towards mathematics), and other factors push people *away* from using particular diagrams (e.g., the complexity of the integrated diagram for individuals with low visuospatial abilities). Thus, the cumulative likelihood or "risk" of benefiting from diagrams can be conceptualized as depending on an accumulation of factors across different levels of analysis (for other work using cumulative risk models, see, e.g., Price & Hyde, 2009).

This perspective highlights two classes of factors that we believe should be distinguished in future work: (1) factors that influence people's *tendency to engage* with diagrams, and (2) factors that influence the *nature of the information* people glean from diagrams. In the current experiments, these two aspects of diagram use are not easily separated. Future studies could address this distinction by using eye tracking to evaluate participants' engagement with diagrams or by including tasks that require participants to engage with the diagrams before they symbolize.

Educational Implications

In practical terms, this dynamic perspective suggests that educators may wish to implement activities that encourage or even require students to engage with diagrams, so that students who stand to benefit from diagrams can actually reap those benefits. For example, students could be asked to label elements of the diagrams that correspond with elements of the story problems.

Our findings further suggest that diagrams may be most useful in situations in which discerning mathematical structure is challenging—both in story problems and in the "real world". However, our findings did not show that brief experience with diagrams transferred to later performance on problems without diagrams. One subgroup—participants with more positive attitudes towards mathematics—did show some beneficial effects of experience with diagrams on transfer. On the whole, however, the data suggest that diagrams are most valuable when they are available during problem solving.

Limitations and Future Directions

This work involved a near-direct internal replication that yielded results that were not identical to the initial experiment. Because the two experiments were independent efforts to estimate effects (see Morehead et al., 2019), we do not view the inconsistencies as deeply problematic. Instead, we view our replication effort as valuable, for three main reasons. First, because of the inconsistent results, we identified important differences in two samples drawn from what was ostensibly the same population. We suspect that these differences are due to the types of students who choose to participate in research for extra credit at different points in the semester. Other researchers have also noted such differences (Richert & Ward, 1976; Grimm et al., 2016), suggesting caution for those who regularly use extra-credit participant pools. Second, our inconsistent results pushed us to dig deeper into our data and to develop a new outcome measure—conceptually correct responses—that yielded meaningful results. Third, the inconsistent results reveal important information about the robustness and generalizability of our findings—information that other scholars should take into account, should they choose to build upon this work.

It is also important to consider constraints on the generality of our results. Both samples were made up of undergraduates with fairly strong mathematics abilities. Younger participants or participants with less mathematical experience might show different patterns of benefits from diagrams. In addition, the specific type of tape diagrams that we used was novel for all participants, as we created it for this study. Some participants may have viewed these diagrams, especially the discrete versions, as disjointed or artificial, and this may have created barriers for structure mapping and interpretation. The findings might not generalize to more familiar diagram formats, which might require less effort to interpret.

More generally, our findings show that the effects of diagrams were not uniform for all subgroups of participants. Specifically, the findings suggest that some subgroups of participants did not engage deeply with the diagrams. Future work should more closely examine how participants attend to diagrams and map between diagrams and text.

There may also be other individual difference factors that influence how people engage with diagrams. For example, participants may vary in their understanding of tape diagrams as representations of information, and this understanding may influence their attention to the diagrams and their efforts to integrate diagrams and text. Future work should assess participants' knowledge of conventions for diagrams and how this knowledge relates to their use of diagrams in problem solving.

Conclusion

In sum, understanding how diagrams influence symbolization is a complicated endeavor, because performance depends on many factors, including visuospatial ability and attitudes towards mathematics. In designing materials to support performance, it is important to consider, not only how a provided visual representation may help, but also *whom* the visual representation may help. Although diagrams were not uniformly beneficial, our findings underscore the value of diagrams in supporting mathematics performance for many learners, particularly for the challenging task of symbolizing integrated equations in complex story problems.

References

Alibali, M. W., Brown, S. A., & Menendez, D. (2019). Understanding strategy change: Contextual, individual, and metacognitive factors. *Advances in Child Development and Behavior*, *56*, 227-256. https://doi.org/10.1016/bs.acdb.2018.11.004

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package 'lme4'. *Convergence*, *12*, 2.

Berends, I. E., & Van Lieshout, E. C. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction, 19*, 345-353. https://doi.org/10.1016/j.learninstruc.2008.06.012

Booth, J. L., & Koedinger, K. R. (2012). Are diagrams always helpful tools? Developmental and individual differences in the effect of presentation format on student problem solving. *British Journal of Educational Psychology*, *82*, 492-511. https://doi.org/10.1111/j.2044-8279.2011.02041.x

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*, 182. https://doi.org/10.1037/0022-0663.98.1.182

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*, 5-26. https://doi.org/10.1023/a:1013176309260

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293-332. https://doi.org/10.1207/s1532690xci0804_2

Chu, J., Rittle-Johnson, B., & Fyfe, E. R. (2017). Diagrams benefit symbolic problem-solving. *British Journal of Educational Psychology*, *87*, 273-287. https://doi.org/10.1111/bjep.12149

Cooper, J. L., Sidney, P. G., & Alibali, M. W. (2018). Who benefits from diagrams andillustrations in math problems? Ability and attitudes matter. *Applied Cognitive Psychology*, *32*, 24-38. https://doi.org/10.1002/acp.3371

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational testing service.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, *31*, 129-151. https://doi.org/10.1016/0749-596x(92)90008-l

Grimm, L. R., Markman, A. B., & Maddox, W. T. (2012). End-of-semester syndrome: How situational regulatory fit affects test performance over an academic semester. *Basic and Applied Social Psychology*, *34*, 376-385. https://doi.org/10.1080/01973533.2012.693427

Gvozdic, K., & Sander, E. (2020). Learning to be an opportunistic word problem solver: going beyond informal solving strategies. *ZDM*, 1-13. https://doi.org/10.1007/s11858-019-01114-z

Gyselinck, V., Cornoldi, C., Dubois, V., De Beni, R., & Ehrlich, M. F. (2002). Visuospatial memory and phonological loop in learning from multimedia. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 16(6), 665-685. https://doi.org/10.1002/acp.823

Heffernan, N., & Koedinger, K. R. (1997). The composition effect in symbolizing: The role of symbol production versus text comprehension. In M. G. Shafto, & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 307–312). Hillsdale, NJ: Erlbaum.

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91, 684. https://doi.org/10.1037/0022-0663.91.4.684

Hembree, R. (1992). Experiments and relational studies in problem solving: A meta-analysis. *Journal for Research in Mathematics Education*, 242-273. https://doi.org/10.2307/749120

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—a meta-analytic review. *Educational Psychology Review*, *22*, 245-269. https://doi.org/10.1007/s10648-010-9126-7

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science*, *32*, 366-397. https://doi.org/10.1080/03640210701863933

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences*, *13*, 129-164. https://doi.org/10.1207/s15327809jls1302_1

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, *29*, 745-756. https://doi.org/10.3758/bf03200477

Landy, D., Brookes, D., & Smout, R. (2014). Abstract numeric relations and the visual structure of algebra. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1404. https://doi.org/10.1037/a0036823

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, *101*, 373-387. https://doi.org/10.1037/a0013843

Majeed, A. A., Darmawan, I. G. N. & Lynch, P. (2013). A confirmatory factor analysis of attitudes toward mathematics inventory (ATMI). *The Mathematics Educator, 15*, 121-135.

Mayer, R. E. (1982). Memory for algebra story problems. *Journal of Educational Psychology*, *74*, 199-216. https://doi.org/10.1037/0022-0663.74.2.199

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*, 43-52. https://doi.org/10.1207/s15326985ep3801_6

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. Cambridge university press. https://doi.org/10.1017/cbo9781139547369.002

Montague, M., Bos, C. S., & Doucette, M. (1991). Affective, cognitive, metacognitive attributes of eighth-grade mathematical problem solvers. *Learning Disabilities Research & Practice, 6*(3), 145–151.

Morehead, K., Dunlosky, J., & Rawson, K. A. (2019). How much mightier is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer (2014). *Educational Psychology Review*, *31*, 753-780. https://doi.org/10.1007/s10648-019-09468-2 .

Moreno, R. (2005). Instructional technology: Promise and pitfalls. *Technology-based education: Bringing researchers and practitioners together*, 1-19.

Murata, A. (2008). Mathematics teaching and learning as a mediating process: The case of tape diagrams. *Mathematical Thinking and Learning*, *10*, 374-406. https://doi.org/10.1080/10986060802291642

Nagashima, T., Bartel, A.N., Silla, E. M., Vest, N. A., Alibali, M. W. & Aleven, V. (2020). Enhancing conceptual knowledge in early algebra through scaffolding diagrammatic self-explanation. In *Proceedings of the 14th International Conference of the Learning Sciences*.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition & Instruction*, *9*, 329-389. https://doi.org/10.1207/s1532690xci0904_2

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*, 1-4. https://doi.org/10.1207/s15326985ep3801_1

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction*, *29*, 128-140. https://doi.org/10.1016/j.learninstruc.2013.02.006

Price, M. N., & Hyde, J. S. (2009). When two isn't better than one: Predictors of early sexual activity in adolescence using a cumulative risk model. *Journal of Youth and Adolescence*, *38*(8), 1059-1071. https://doi.org/10.1007/s10964-008-9351-2

Reed, S. K. (1999). *Word problems: Research and curriculum reform*. Mahwah, NJ: Lawrence Erlbaum. https://doi.org/10.4324/9781410604484

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7*(3), 216–237. https://doi.org/10.1016/j.edurev.2012.05.003

Richert, A. J. & Ward, E. F. (1976). Experimental performance and self-evaluation of subjects sampled early, middle, and late in an academic term: Sex and task. *Psychological Reports, 39*, 135-142. https://doi.org/10.2466/pr0.1976.39.1.135

Schnotz, W. (2002). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, *14*, 101-120. https://doi.org/10.1023/a:1013136727916

Schnotz, W. (2014). The integrated model of text and graphics comprehension. *The Cambridge handbook of multimedia learning*, 73-103. https://doi.org/10.1017/cbo9781139547369.006

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*, 141-156. https://doi.org/10.1016/s0959-4752(02)00017-8

Tapia, M., & Marsh, G. E. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly*, *8*(2), 16-22.

Vekiri, I. (2001). An investigation of the role of graphical design and student characteristics in scientific reasoning with weather maps. Doctoral dissertation, University of Michigan.

Van Buuren, S. & Oudshoorn C.G.M. (2000), Multivariate imputation by chained equations:

MICE V1.0 User's manual. Report PG/VGZ/00.038. Leiden, TNO Preventie en Gezondheid.

Data sharing statement:

The data that support the findings of this study are openly available through OSF at https://osf.io/k6gj7/?view_only=f1508dd2a4474fb197a454c26651d1f6

Endnote

[1] Although previously a common practice, it is not recommended to fix random slopes to zero because simulations have shown that assuming that all subjects have the same average score on the outcome (removal of random intercept) leads to less biased parameter estimates than removing the random slopes for the predictors for which *a priori* predictions were made (Barr et al., 2013).

Tables

Table 1. Correlations among the individual difference factors in Experiment 1 (above the diagonal) and Experiment 2 (below the diagonal)

	1	2	3
1 Visuospatial ability	-	0.44 ***	.31 **
2 Attitudes towards math	0.40 **	-	.61 ***
3 ACT/SAT percentile score	0.29 ***	0.54 ***	-

* p < .05, ** p < .01, *** p < .001

Table 2. Final model statistics for Experiment 1, with probability of correct integration as the dependent variable

Predictor	***B*** **(SE)**	**OR**	χ^2	***p*** **value**
Diagram general (DG) contrast	.08 (.19)	1.08	.21	.64
Diagram specific (DS) contrast	.46 (.21)	1.58	4.43	.03*
Math Ability	.02 (.01)	1.02	5.65	.01*
Visuospatial Ability	.07 (.03)	1.07	8.13	.004*
Math Attitudes	.52 (.14)	1.68	12.35	< .001***
DG*Math Ability	-.07 (.02)	.93	5.10	< .001***
DG*Visuospatial Ability	-.19 (.05)	.83	9.45	< .001***
DG*Math Attitudes	1.06 (.32)	2.87	6.82	.001**
DS*Math Ability	-.02 (.01)	.98	.06	.23
DS*Visuospatial Ability	.15 (.06)	1.16	5.10	.01*
DS*Math Attitudes	-.39 (.35)	.68	2.83	.27
Operation Set	-.17 (.16)	.84	1.02	.31

* p < .05, ** p < .01, *** p < .001

Table 3. Final model statistics for Experiment 2, with probability of correct integration as the dependent variable

Predictor	***B*** **(SE)**	**OR**	χ^2	***p*** **value**
Diagram general (DG) contrast	.69 (.16)	1.99	15.94	<.001***
Diagram specific (DS) contrast	-.30 (.21)	.74	1.82	.15
Math Ability	.05 (.01)	1.05	23.33	<.001***
Visuospatial Ability	.06 (.02)	1.06	5.82	.01*
Math Attitudes	.08 (.13)	1.08	.12	.55
DG*Math Ability	.09 (.02)	1.09	23.34	<.001***
DG*Visuospatial Ability	-.03 (.04)	.97	.51	.51
DG*Math Attitudes	-.08 (.25)	.92	.21	.74
DS*Math Ability	.01 (.03)	1.01	.0013	.73
DS*Visuospatial Ability	.05 (.06)	1.05	.85	.40
DS*Math Attitudes	-.13 (.25)	.87	.008	.69
Operation Set	-.15 (.16)	.86	.93	.34

* p < .05, ** p < .01, *** p < .001

Table 4. Final model statistics for the two two-operator transfer items in Experiment 2 where the dependent variable was probability of correct integration

Predictor	***B*** **(SE)**	**OR**	χ^2	***p*** **value**
Diagram general (DG) contrast	.31 (16)	1.36	.90	.34
Diagram specific (DS) contrast	.005 (.41)	1.005	.0002	.99
Math Ability	.05 (.02)	1.05	6.58	.01*
Visuospatial Ability	.11 (.05)	1.12	5.85	.02*
Math Attitudes	.44 (.24)	1.55	3.32	.06
DG*Math Ability	.06 (.03)	1.06	2.86	.09
DG*Visuospatial Ability	-.14 (.09)	.86	2.10	.14
DG*Math Attitudes	1.11 (.49)	3.03	5.28	.02*
DS*Math Ability	-.02 (.05)	.98	.22	.64
DS*Visuospatial Ability	-.03 (.12)	.97	.06	.80
DS*Math Attitudes	-.03 (.63)	.97	.004	.95
Operation Set	-.005 (.30)	.99	.0003	.98

* p < .05, ** p < .01, *** p < .001

Table 5. Final model statistics for three-operator transfer item in Experiment 2 where the dependent variable was probability of correct integration

Predictor	***B*** **(SE)**	**OR**	χ^2	***p*** **value**
Diagram general (DG) contrast	-.08 (.46)	.92	.04	.85
Diagram specific (DS) contrast	.26 (.53)	1.30	.24	.62
Math Ability	.03 (.02)	1.03	1.06	.30
Visuospatial Ability	.28 (.07)	1.32	20.26	<.001***
Math Attitudes	.24 (.35)	1.27	.46	.50
DG*Math Ability	.03 (.05)	1.03	.37	.55
DG*Visuospatial Ability	-.05 (.17)	.95	.11	.73
DG*Math Attitudes	1.10 (.72)	3.00	2.33	.12
DS*Math Ability	.006 (.08)	1.006	.005	.94
DS*Visuospatial Ability	.06 (.17)	1.06	.10	.75
DS*Math Attitudes	.22 (.90)	1.25	.06	.80

* p < .05, ** p < .01, *** p < .001

Table 6. Error analysis codes with an example problem

Example problem: Alex has a garden with a certain number of carrots growing in it. He picks all the carrots and puts an equal number of carrots in each of four baskets. He then decides to put three store-bought carrots in each basket. In the end, every basket has eleven carrots. How many carrots were in Alex's garden to begin with?

Error Type	**Description**	**Example**
	Correct Symbolization of the Conceptual Structure (indicates participant symbolized the mathematical relations in the problem)	
Integrated equation	Participant generates a single equation with both operations	situation model: n/4 + 3 = 11 solution model: (11 – 3) * 4 = n
One step solved mentally	Participant solves one step mentally, and produced a one-operator equation from a two-operator story problem	n = 8(4)
Two equations	Participant generates two separate one-operator equations	11 – 3 = x; x * 4 = n
	Incorrect Symbolization of the Conceptual Structure (indicates participant did not symbolize the mathematical relations in the problem)	
Wrong operation	Participant generates an equation that generally has one incorrect operation	4n + 3 = 11
Distributed incorrectly	Participant creates an equation that incorrectly distributes an operation	4n + 3(4) = 11(4)
Misunderstand situation	Participant generates an incorrect solution path from the story problem	4 (n - 3) = 11

Table 7. Final model statistics for the exploratory analysis of combined data from Experiment 1 and 2 where the dependent variable was probability of correctly symbolizing the conceptual structure of the story problem

Predictor	*B* (SE)	OR	χ^2	*p* value
Diagram general (DG) contrast	.32 (.14)	1.38	4.99	.02*
Diagram specific (DS) contrast	.24 (.17)	1.27	1.87	.17
Math Ability	.02 (.006)	1.02	13.54	<.001***
Visuospatial Ability	.09 (.02)	1.09	24.05	<.001***
Math Attitudes	-.05 (.11)	.95	.21	.65
DG*Math Ability	.004 (.01)	1.004	.22	.64
DG*Visuospatial Ability	-.09 (.03)	.91	4.92	.02*
DG*Math Attitudes	.34 (.22)	1.40	2.34	.13
DS*Math Ability	.02 (.01)	1.02	.97	.32
DS*Visuospatial Ability	.07 (.04)	1.07	2.16	.14
DS*Math Attitudes	-.70 (.28)	.49	6.04	.01*
Operation Set	-.006 (.13)	.99	.002	.96

* p < .05, ** p < .01, *** p < .001

Table 8. Final model statistics for the exploratory analysis of combined data from Experiment 1 and 2 where the dependent variable was the percent correct integration given a conceptually correct response was provided.

Predictor	*B* (SE)	*F* value	*p* value
Diagram general (DG) contrast	.03 (.05)	.40	.53
Diagram specific (DS) contrast	-.01 (.06)	.04	.83
Math Ability	.005 (.002)	4.21	.04*
Visuospatial Ability	.01 (.006)	3.78	.05
Math Attitudes	.06 (.04)	2.28	.13
DG*Math Ability	-.0006 (.004)	.02	.90
DG*Visuospatial Ability	-.02 (.01)	2.01	.16
DG*Math Attitudes	.002 (.08)	.0006	.98
DS*Math Ability	.003 (.007)	.17	.68
DS*Visuospatial Ability	-.01 (.02)	.37	.54
DS*Math Attitudes	-.04 (.09)	.20	.65

Figure Captions.

Figure 1. Sample story problem and corresponding diagrams. Two-operator story problems were paired with either (A) discrete diagrams, (B) integrated diagrams, or were presented with no diagrams (not shown).

Figure 2. Main effect of the diagram-general contrast of *diagram vs. no diagram* on the probability of correctly symbolizing the conceptual structure of the story problem

Figure 3. Interaction of visuospatial ability and the diagram condition on the probability of correctly symbolizing the conceptual structure of the story problem.

Figure 4. Interaction of attitudes towards mathematics and the diagram-specific contrast of *integrated diagram* vs. *discrete diagram* on the probability of correctly symbolizing the conceptual structure of the story problem.

Appendix 1.
Results from Experiment 1

Figure Captions.

-Insert Fig 1 in Appendix 1-

Figure 1. Main effect of the diagram-specific contrast of *integrated diagram* vs. *discrete diagram* on the probability of correct integration.

-Insert Fig 2 in Appendix 1-

Figure 2. Interaction of math ability and the diagram-general contrast of *diagram* vs. *no diagram* on the probability of correct integration.

-Insert Fig 3 in Appendix 1-

Figure 3. Interaction of visuospatial ability and both the diagram-general contrast of *diagram* vs. *no diagram*, and the diagram-specific contrast of *integrated diagram* vs. *discrete diagram* on the probability of correct integration.

-Insert Fig 4 in Appendix 1-

Figure 4. Interaction of attitudes towards mathematics and the diagram-general contrast of *diagram* vs. *no diagram* on the probability of correct integration.

Appendix 2.
Results from Experiment 2

-Insert Fig 1 in Appendix 2-

Figure 1. Main effect of the diagram-general contrast of *diagram* vs. *no diagram* on the probability of correct integration.

-Insert Fig 2 in Appendix 2-

Figure 2. Interaction of math ability and the diagram-general contrast of *diagram* vs. *no diagram* on the probability of correct integration.

-Insert Fig 3 in Appendix 2-

Figure 3. Interaction of attitudes toward mathematics and the diagram-general contrast of *diagram* vs. *no diagram* for 2-operation transfer items on the probability of correct integration.

Appendix 3.
Supplementary Analyses

Table 1. Final model statistics for the combined data from Experiment 1 and Experiment 2, where the dependent variable was probability of correctly symbolizing the conceptual structure of the story problem

Predictor	*B* (SE)	OR	χ^2	*p* value
Diagram general (DG) contrast	.32 (.12)	1.38	6.66	.009**
Diagram specific (DS) contrast	.09 (.14)	1.09	.38	.54
Math Ability	.02 (.006)	1.02	13.90	<.001***
Visuospatial Ability	.07 (.02)	1.07	20.00	<.001***
Math Attitudes	.25 (.09)	1.28	6.83	.008**
DG*Math Ability	.002 (.01)	1.002	.05	.82
DG*Visuospatial Ability	-.13 (.04)	.87	12.53	<.001***
DG*Math Attitudes	.58 (.20)	1.79	8.84	.003**
DS*Math Ability	-.0008 (.02)	.99	.003	.96
DS*Visuospatial Ability	.04 (.04)	1.04	1.01	.31
DS*Math Attitudes	-.35 (.23)	.70	2.20	.13
Operation Set	-.16 (.11)	.85	2.04	.15

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 2. Final model statistics for Experiment 1, where the dependent variable was probability of correctly symbolizing the conceptual structure of the story problem

Predictor	*B* (SE)	OR	χ^2	*p* value
Diagram general (DG) contrast	-.06 (.23)	.94	.06	.81
Diagram specific (DS) contrast	.68 (.26)	1.97	6.77	.009**
Math Ability	.04 (.01)	1.04	13.03	<.001***
Visuospatial Ability	.13 (.03)	1.13	20.61	<.001***
Math Attitudes	-.14 (.18)	.86	.62	.42
DG*Math Ability	-.05 (.02)	.95	4.50	.03*
DG*Visuospatial Ability	-.11 (.07)	.89	2.94	.08
DG*Math Attitudes	.51 (.40)	1.66	1.71	.19
DS*Math Ability	.02 (.02)	1.02	.50	.47
DS*Visuospatial Ability	.16 (.07)	1.17	5.91	.01*
DS*Math Attitudes	-.96 (.43)	.38	5.03	.02*
Operation Set	-.08 (.19)	.92	.21	.65

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 3. Final model statistics for Experiment 2, where the dependent variable was probability of correctly symbolizing the conceptual structure of the story problem

Predictor	***B* (SE)**	**OR**	χ^2	***p* value**
Diagram general (DG) contrast	.69 (.20)	1.99	12.24	<.001***
Diagram specific (DS) contrast	-.06 (.25)	.94	.06	.80
Math Ability	.03 (.01)	1.03	4.15	.04
Visuospatial Ability	.05 (.03)	1.05	3.19	.07
Math Attitudes	-.001 (.16)	.99	.0001	.99
DG*Math Ability	.05 (.02)	1.05	5.36	.02*
DG*Visuospatial Ability	-.04 (.05)	.96	.61	.43
DG*Math Attitudes	.19 (.30)	1.20	.40	.53
DS*Math Ability	.01 (.04)	1.01	.11	.74
DS*Visuospatial Ability	.01 (.07)	1.01	.02	.89
DS*Math Attitudes	-.18 (.42)	.83	.18	.67
Operation Set	.07 (.19)	1.07	.15	.70

* $p < .05$, ** $p < .01$, *** $p < .001$