

# Investigating Scaffolds for Sense Making in Fraction Addition and Comparison

Eliane Stampfer Wiese (stampfer@cs.cmu.edu)

Kenneth R. Koedinger (koedinger@cmu.edu)

Human-Computer Interaction Institute, 5000 Forbes Avenue

Pittsburgh, PA 15213 USA

## Abstract

What types of scaffolds support sense making in mathematics? Prior work has shown that grounded representations such as diagrams can support sense making and enhance student performance relative to analogous tasks presented with more abstract, symbolic representations. For grounded representations to support students' learning of symbolic representations, students' sense making must be maintained when both grounded and symbolic representations are presented together. This study investigates why students sometimes fail to coordinate these representations, in particular, why performance is high with fraction diagrams alone, but decreases when fraction symbols are included. Results indicate that symbols trigger incorrect transfer from whole-number procedures, and that students lack the qualitative reasoning that the diagrams are intended to tap. Specifically, students do not find it obvious that the sum of two positive symbolic fractions is larger than its two addends. Qualitative inference rules such as this one appear important in mediating the sense making process in the context of tempting misconceptions even when otherwise-supportive grounded representations are available.

**Keywords:** fractions; sense making; diagrams.

## Scaffolds for Sense-Making

A common lament in mathematics education is that students often execute procedures without understanding: without connecting procedural steps to their underlying concepts and without connecting symbols to their 'real-world' referents (Schoenfeld, 1988). One way to reinforce the conceptual foundation for procedures is to use visual representations, such as strip diagrams. These diagrams are not intended to help student execute procedures, but instead support them in thinking about the problems qualitatively (e.g., which amounts are bigger? Which operation is appropriate?). Visual representations are thought to help students apply their conceptual reasoning (Beckmann, 2004), and are recommended by an Institute of Education Sciences Practice Guide (Woodward et al., 2012). However, diagrams may not be intuitive for novices, and their presence can decrease problem-solving performance for students who have difficulty interpreting them (Booth & Koedinger, 2011).

Prior work on diagrams, including our own, indicates that their utility as sense making supports is sensitive to context, including the problem type and the students' prior knowledge (Stampfer & Koedinger, 2013). Booth and Koedinger (2011) hypothesized that several factors could contribute to students' misinterpretation of diagrams or their difficulty mapping between diagrams and problem statements, including a lack of domain knowledge and still-

developing formal reasoning. This paper, a follow-up study to our 2013 work, investigates alternate hypotheses for our results and examines students' domain knowledge and reasoning skills with addition. We briefly review our earlier findings and then present the current study.

## Fraction Bars and Fraction Symbols

Stampfer and Koedinger (2013) examined *fraction bars*, which represent the proper fraction  $n/d$  with a rectangle divided into  $d$  equal pieces,  $n$  of them shaded. We used four combinations of fraction bars and symbols to examine how the presence of symbols affected students' use of the fraction bars. Each scaffold type targeted one piece of knowledge: 1) equal areas represent equal amounts; 2) the bars represent fractions; 3) if two shaded areas are equal, the fractions they represent are equal. Figures 1-4 show sample comparison and addition problems for each scaffold type (and the Numbers-Only baseline). Comparison problems presented two fractions and asked if the first was bigger than, equivalent to, or smaller than the second. Addition problems presented equations and asked if they were true or false. False equations used the incorrect strategy of adding the numerators and denominators to obtain the sum. The study was within subject: each 5<sup>th</sup> grade student solved two problems of each scaffold type for each task.

The results (Table 1) show striking differences between the scaffold types and between the addition and comparison tasks. On the comparison items, all three scaffold types with fraction bars yielded uniformly high performance. Low performance on the Numbers-Only control confirmed that students were not simply comparing the fractions with the symbols alone. Fraction addition showed a different pattern: performance was significantly different for all scaffold types, and decreased steadily as the salience of the numbers increased. These results show that the usefulness of fraction bars depends on the task. For fraction comparison, students display all three necessary skills for interpreting the diagrams, as evidenced by equally high performance on all scaffold types that included pictures. However, these skills are not uniformly shown with fraction addition. Without symbols, students can correctly interpret the fraction bars, including the multi-colored sum, as evidenced by their performance with the Pictures-Only scaffold. However, the presence of symbols seems to detract from students' use of the diagrams, such that performance on the Half Pictures and Numbers scaffold hovers around chance. Students' dismal performance (21%) with Numbers-Only indicates that adding the numerators and denominators is a tempting foil, as it draws on students' incorrect transfer from whole-

Compare, then circle the correct answer:

True or False:

Circle the correct answer: True False

a) is bigger than   
b) is equivalent to   
c) is smaller than

Figure 1: Pictures-Only. Area equals quantity.

Compare, then circle the correct answer:

True or False:

Circle the correct answer: True False

a) is bigger than   
b) is equivalent to   
c) is smaller than

Figure 2: Pictures and Numbers. Bars represent fractions.

Compare, then circle the correct answer:

True or False:

Circle the correct answer: True False

a)  $\frac{2}{3}$  is bigger than  $\frac{12}{18}$   
b)  $\frac{2}{3}$  is equivalent to  $\frac{12}{18}$   
c)  $\frac{2}{3}$  is smaller than  $\frac{12}{18}$

Figure 3: Half Pictures and Numbers. Mapping relationships from images to symbols.

Compare, then circle the correct answer

a)  $\frac{1}{3}$  is bigger than  $\frac{8}{19}$   
b)  $\frac{1}{3}$  is equivalent to  $\frac{8}{19}$   
c)  $\frac{1}{3}$  is smaller than  $\frac{8}{19}$

True or False:

$$\frac{2}{11} + \frac{1}{2} = \frac{15}{22}$$

Circle the correct answer: True False

Figure 4: Numbers-Only. Evaluating solved problems.

Table 1: Fall 5<sup>th</sup> Grade Percent Correct

| Scaffold           | Comparison | Addition |
|--------------------|------------|----------|
| Pictures Only      | .82        | .79      |
| Pictures & Numbers | .81        | .64      |
| Half Pics & Nums   | .83        | .46      |
| Numbers Only       | .50        | .21      |

number addition. However, our prior study did not have sufficient error-type data to confirm this suspicion.

Why did symbols decrease students' performance on the addition task? Students' performance with comparison

indicates that they can extract information equally from all three fraction-bar scaffold types. What prevents them from using this information with addition? We hypothesize a sense making process that demands recognition of two basic properties of positive-number addition for effective use of the fraction bars: 1) the magnitude of the sum equals the combined magnitudes of the addends; and 2) the sum is larger than each of the addends. With this knowledge, the incorrect symbolic addition equations should be easy to reject, since all propose a sum that is smaller than one of the addends. The present study examines whether students know the second, presumably more difficult, property.

The prior study also left other open questions. First, the 'true or false' options did not give any insight into students' reasoning. Second, the comparison items with non-equivalent fractions did not use foils based on possible misconceptions. Instead, they used fractions with similar magnitudes. Perhaps students have systematic misconceptions about equivalence, as they do with addition, but those misconceptions were simply not elicited. In the next section we describe the experimental design of the present study and then the test items that address these questions, with the results and discussions interleaved.

## Follow-Up Study

The present study addresses three questions about sense making support for fractions: 1) Is it obvious to students that the sum of two positive symbolic addends is larger than each addend individually? 2) When students do not recognize the correct sum of a fraction addition equation, is it due to incorrect whole-number transfer? 3) Are students tempted by systematic foils for fraction equivalence?

The present study was conducted with the same fifth-grade public-school students as the prior study. The prior study took place in the fall and the present study took place in the spring. Thus, students had about 5-6 months more classroom instruction in the present study than the previous one, explaining their higher scores on comparable tasks. 160 fifth-graders were given 20 minutes to complete the 34-item test forms, administered by their classroom teacher during the normal school day. The school tracked students into three achievement levels, which we refer to as High, Middle, and Low. To control for ordering effects, question order was determined randomly and half of the test forms were printed in reverse order. Within each class, students were randomly assigned to one of the four test forms, printed in either forward or reverse order. Two items were inadvertently left off the test forms of 19 students, and we account for this discrepancy in our analysis. Some research questions were addressed with between-subjects design, and some with within-subjects design. We discuss the particular research design for each set of items individually.

## Comparing Addends and Sums

To see if students knew that the sum of two positive, symbolic addends was bigger than each addend alone, items presented a correct addition equation and asked if the sum

This addition is correct.

$$843,216,001 + 169,582,503,244 = 170,425,719,245$$

With that information only, answer these two questions:

1) 843,216,001 is bigger than 170,425,719,245

True      False      Can't tell from the information given

2) 169,582,503,244 is bigger than 170,425,719,245

True      False      Can't tell from the information given

Figure 5: Addend-Sum item with whole numbers

This addition is correct.

$$.617 + .083 = .7$$

With that information only, answer these two questions:

1) .617 is bigger than .7

True      False      Can't tell from the information given

2) .083 is bigger than .7

True      False      Can't tell from the information given

Figure 6: Addend-Sum item with decimals

This addition is correct:

$$\frac{24}{64} + \frac{39}{104} = \frac{6}{8}$$

With that information only, answer these two questions:

1)  $\frac{6}{8}$  is bigger than  $\frac{24}{64}$

True      False      Can't tell from the information given

2)  $\frac{6}{8}$  is bigger than  $\frac{39}{104}$

True      False      Can't tell from the information given

Figure 7: Addend-Sum item with fractions

was bigger than each added (or visa versa). Response options were True, False, and Can't tell from the information given. Items had either whole numbers, decimals, fractions, or variables (Figures 5-8). Items with variables had two presentations. We used a between-subjects design, assigning each student to one of the four number types, with 5 problems of that type. To control for students simply selecting true or false for all of the problems, 3 problems asked if the sum was bigger than each addend, and 2 problems asked if each addend was bigger than the sum. Students in the variables condition had 3 problems with shapes and 2 with people (Figure 8).

**Addend-Sum Results** Questions were scored 1 if correct and 0 otherwise. Table 2 shows overall percent correct for each number type. Figure 9 shows performance for each number type by tracking level. We ran an ANOVA on percent correct (dependent) with test form, tracking level, and question order (forward vs. reversed) as fixed factors. Question order was not significant and there were no

This addition is correct, but all of the numbers are covered.

All of the numbers are bigger than 0.

$$\square + \triangle = \text{Flag}$$

With that information only, answer these two questions:

1) The number covered by  is bigger than the number covered by 

True      False      Can't tell from the information given

2) The number covered by  is bigger than the number covered by 

True      False      Can't tell from the information given

Michelle and Joe are each thinking of a number. Those numbers are both bigger than 0. When you add those numbers together, you get the number that Sandra is thinking of.

With that information only, answer these two questions:

1) Michelle's number is bigger than Sandra's number

True      False      Can't tell from information given

2) Joe's number is bigger than Sandra's number

True      False      Can't tell from information given

Figure 8: Addend-Sum items with variables

significant interactions with order, so we re-ran the analysis using only test form and tracking level. There was a significant effect of form, tracking level, and a significant interaction (all  $p < .01$ ). Post-hoc Tukey tests showed significant differences between Variables and Whole Numbers and Variables and Decimals (both  $p < .01$ ), and Fractions and Whole Numbers ( $p = .022$ ).

Table 2: Percent Correct for Addend-Sum Items

| Whole | Decimals | Fractions | Variables |
|-------|----------|-----------|-----------|
| 79%   | 75%      | 61%       | 51%       |

**Addend-Sum Discussion** Except for the High group, most students could not apply the addend-sum relationship to all four number types. This evidence supports the hypothesis that students' difficulty interpreting the fraction-addition diagrams arises from a gap in prior knowledge: they do not always recognize the significance of a proposed sum being smaller than one of the addends because they do not have a strong, fluent knowledge of the qualitative addend-sum relationship. Confusion may stem from addition with negative numbers (addition does not always make bigger) or fraction multiplication (even for positive numbers, operations do not always go in the same direction). Students can solve the whole number and decimal problems by directly comparing the numbers in each question without considering the addend-sum relationship. It is much harder to directly compare unlike-denominator fractions, and impossible for variables. This difference in strategy likely explains the significant differences between Variables and Wholes/Decimals. Performance by tracking level suggests how mastery of this relationship may develop. Whole-Number performance is about the same with all three tracks, likely reflecting a direct-comparison strategy and familiarity with whole numbers. Decimals performance is low for Low-track students (~50%), likely reflecting unfamiliarity with

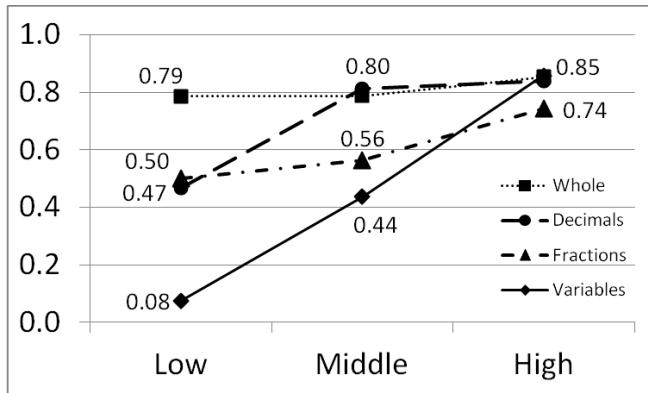


Figure 9: Percent correct on Addend-Sum items by form and class tracking. Points differing by < 3% were averaged.

decimal comparison, but rises to Whole-Number level with Middle-track students. With Variables, Low-track students perform just below chance, indicating that they do not understand how addends and sums relate in the abstract. This abstract understanding trails fraction performance for Middle- and Low-track students. Although this qualitative relationship is important for reasoning about addition, these results suggest that students may not fully grasp the addend-sum relationship until they have extensive practice adding numbers of many types. This finding is in line with theories that procedural and conceptual skills develop iteratively (Rittle-Johnson, Siegler, & Alibali, 2001).

### Fraction Addition Items and Equivalence Foils

Addition and comparison items used the same types of scaffolds as the prior study, so we do not give examples of all types. However, this time addition items offered three responses: the sum could be too small, correct, or too big. Comparison items with non-equivalent fractions aimed to assess the extent of three potential misconceptions: fractions with the same numerator are equivalent, regardless of denominator (e.g.,  $\frac{3}{4}$  and  $\frac{3}{16}$ ); squaring the numerator and denominator maintains equivalence (e.g.,  $\frac{2}{5}$  and  $\frac{4}{25}$ ); and adding the same number to the numerator and denominator maintains equivalence (e.g.,  $\frac{11}{12}$  and  $\frac{14}{15}$ ).

Compare, then circle the correct answer:

$\frac{7}{8}$

$\frac{9}{10}$

$\frac{16}{18}$

a)  $\frac{7}{8}$  is smaller than  $\frac{9}{10}$   
b)  $\frac{7}{8}$  is equivalent to  $\frac{9}{10}$   
c)  $\frac{7}{8}$  is bigger than  $\frac{9}{10}$

Hint:

$\frac{4}{6} + \frac{2}{9}$

$\frac{16}{18}$

Is this correct?  
 $\frac{4}{6} + \frac{2}{9} = \frac{16}{18}$

Circle the answer that goes in the blank:  
 $\frac{16}{18}$  is \_\_\_\_.  
a) too small  
b) correct  
c) too big

Figure 10: Half Pictures and Numbers, one-less foil (left) and correct addition (right)

We refer to the foil types that target these misconceptions as *same numerator*, *squaring*, and *one-less*, respectively (one-less refers to the addition misconception since each numerator is one less than its denominator). Figure 10 gives an example of the Half-Pictures-and-Numbers scaffold with the One-Less foil and correct addition. This study used a between-subject design for scaffold (with each test form using only one of the four scaffold types) and a within-subject design for task (each student did comparison and addition items). Tests included 6 addition items and 12 comparison items.

**Picture and Symbol Scaffold Results** 19 tests inadvertently had 11 comparison items instead of 12, so we used percent correct instead of raw scores in all analyses. An ANOVA with task (comparison and addition) as a repeated measure, and with scaffold type, tracking level, and question order (forward vs. reversed) as fixed factors showed that question order was not significant and had no significant interactions, so we re-ran the analysis without it. We found a significant effect of task ( $p < .01$ ) but no significant task by scaffold interaction. For each task (comparison and addition) we ran an ANOVA on percent correct (dependent) with scaffold type and tracking level as fixed factors. For comparison items, there was a significant effect of scaffold and class level (both  $p < .01$ ), with a scaffold by class level interaction ( $p = .013$ ). Post-hoc Tukey tests showed significant differences between Numbers Only and all other scaffold types (all  $p < .001$ ) but no other significant differences. For addition items, scaffold and class were again significant, with a marginal interaction ( $p = .058$ ). Post-hoc Tukey tests showed significant differences between Numbers Only and all other scaffold types (all  $p < .015$ ); Pictures Only and Pictures and Numbers ( $p < .01$ ); and a marginal difference between Pictures Only and Half-Pictures-and-Numbers ( $p = .087$ ). Since those tests revealed no differences between the two scaffold types with both representations and did reveal differences between them and the scaffold types with one representation, we collapse those two scaffold types for further analysis. An ANOVA with the three scaffold groups and class level as fixed factors showed significant main effects ( $p < .01$ ) and a significant interaction ( $p = .031$ ). Post-hoc tests show significant differences between all three scaffold groups (all  $p < .01$ ). Figure 11 shows performance for the three groups.

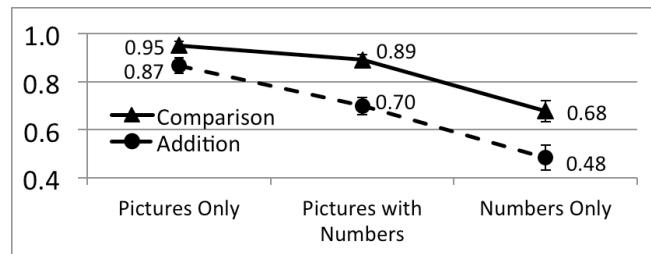


Figure 11: Percent correct on comparison and addition items by scaffold type, with standard error bars

**Picture and Symbol Scaffold Discussion** These results are consistent with the progression of performance on these scaffolds from 5<sup>th</sup> through 7<sup>th</sup> grade (Wiese & Koedinger, 2014). Like 6<sup>th</sup> graders, for the spring 5<sup>th</sup> graders addition is harder than comparison, but the scaffolds affect the difficulty of both tasks in the same way. Also, their pattern of differences in addition scores between scaffold types is closer to that of 6<sup>th</sup> graders (in the fall, all differences were significant). Finally, the comparison results were replicated with systematic equivalence foils.

**Error Analysis** Incorrect transfer from whole numbers is demonstrated when students say the strategy of adding both numerators and denominators (add-both) is correct, or when they say the correct answer is too big (since the numerator and denominator are both larger than the corresponding result of add-both). This error can also occur with the Pictures-Only scaffold if students count the number of shaded segments instead of comparing the overall sizes of the shaded amounts. Figure 12 shows the rate of correct responses and three error types (whole number error, other, and blank) for all of the addition items by scaffold type. Each student had 6 questions, and that many opportunities to make this error. Figure 13 shows how many whole-number errors each student made within each scaffold type. For example, 20% of students in the Pictures condition made 1 or 2 whole-number errors, while 20% of students in

the Numbers condition made 5 or 6 whole-number errors. The majority of errors are consistent with whole-number thinking. These errors are most pronounced with Numbers-Only, but are mitigated by the diagrams, suggesting that the fraction symbols trigger this misconception. Together with the Addend-Sum results, mediocre performance on the Pictures with Numbers scaffolds (70%) suggests that the diagrams do not help some students tap their conceptual, qualitative understanding of addition with numbers because that qualitative understanding is not fully in place. Therefore, combining diagrams with numbers improves performance relative to numbers-only, but does not make the answers obvious for all students.

**Equivalence Foil Results** All test forms included questions with all foil types: 3 equivalent, 5 same numerator, 2 squaring, and 2 one-less (19 students only had one). Table 3 shows scores by foil and scaffold type. An ANOVA on percent correct with foil type as a repeated measure and scaffold and class level as fixed factors showed a significant effect of foil ( $p<.01$ ) and a significant foil by scaffold interaction ( $p=.023$ ). We then ran ANOVAs for each scaffold type separately (individual question score as dependent, foil type and class level as fixed factors and student as random factor). For Pictures Only and Pictures with Numbers, foil type was significant ( $p<.01$ ) and post-hoc Tukey tests showed the One-Less foil was different from all the others (all  $p<.02$ ). For Numbers Only, there was no significant effect of foil type.

**Error Analysis and Discussion** Since the three scaffold types with pictures had similar results, we combine them for the error analysis. For each equivalence foil, the three error types are: mistaken for equivalent, wrong direction of inequality, and blank. Figure 14 shows the percentage of responses for each equivalence foil for scaffolds with pictures. Figure 15 shows the same for Numbers Only.

Without diagrams, all four comparison items are similarly difficult, giving no evidence for consistent misconceptions. With diagrams, performance is high on all but the One-Less foil. The error analysis shows that the most popular incorrect response on the One-Less problem is that the fractions are equivalent. This error pattern is not repeated for the other foil types with pictures or for any foil type with numbers only (for numbers only the equivalence error occurs at a similar rate in terms of overall responses, but not as a percentage of erroneous responses). Perhaps students who do not look closely at the pictures are fooled by the small ( $< 3\%$ ) size difference of the One-Less pairs. That

Table 3: Percent Correct by Foil and Scaffold.

| Scaffold      | Equivalent | Same Numerator | Squaring | One Less |
|---------------|------------|----------------|----------|----------|
| Pictures Only | .97        | .98            | .95      | .86      |
| Pics w. Numbs | .91        | .91            | .93      | .76      |
| Numbers       | .74        | .69            | .61      | .61      |

Figure 12: Responses to the addition items, by scaffold type.

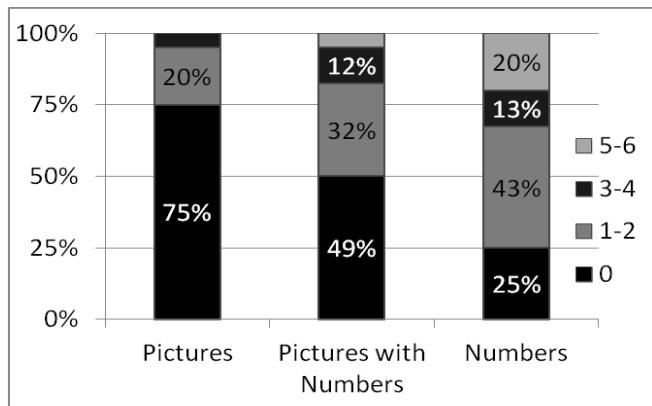


Figure 13: Whole-number errors made by each student, by scaffold type.

difference is much smaller than the ~7% average difference between non-equivalent fractions in the prior study. Alternatively, perhaps students noticed the discrepancy and decided it did not matter because 1) it was close enough; and 2) adding the same number to the numerator and denominator seems similar to the correct procedure. Although pictures improved performance overall, this result is one example of their potential drawbacks: depending on their scale they may appear to show untrue relationships, and could possibly reinforce misconceptions.

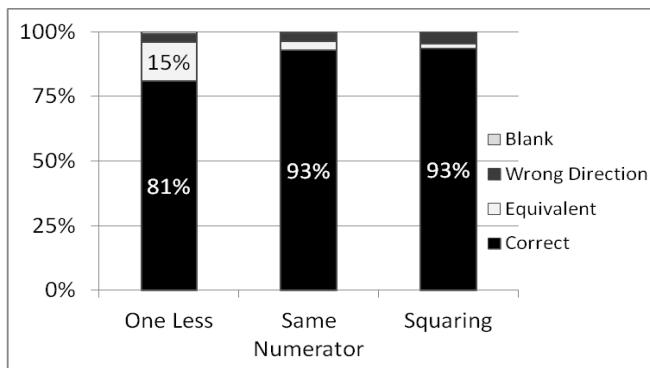


Figure 14: Responses to foils for all scaffolds with pictures.

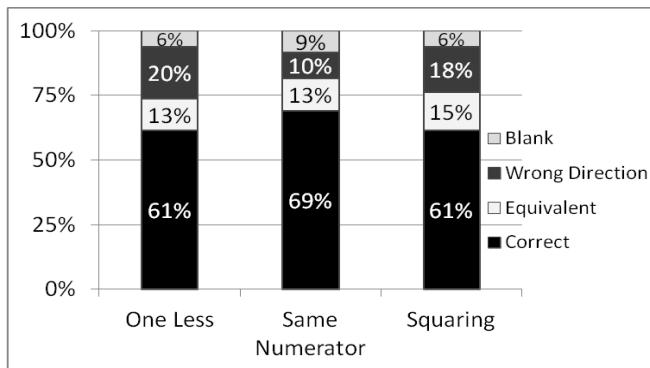


Figure 15: Responses to foil items for Numbers condition

## Conclusion

Diagrams are thought to aid sense making by helping students apply conceptual (often qualitative) reasoning to a problem (e.g., which amounts are equal? What operation is needed?). This study provides evidence for diagrams' overall sense-making support, but also offers an explanation for why students do not always use diagrams effectively: they may lack that conceptual, qualitative reasoning that diagrams are intended to tap. This prior knowledge (e.g., that the sum of two positive addends is larger than each addend) may be obvious to adults but not to students. Further, students may be able to apply this knowledge in some contexts (e.g., with diagrams alone) but not others (e.g., the addend-sum items with fraction symbols). These findings, that students do not always know how to use visual representations, support the IES Practice Guide

recommendations that students be taught these skills explicitly (Woodward et al., 2012). However, the current recommended instruction focuses on mapping between a story problem and a visual representation, and then the visual representation and symbols. The results from this study suggest that students may also benefit from instruction on what type of qualitative reasoning is relevant to the problem, and how to apply that reasoning.

More generally, it seems more caution is needed in applying expert intuitions about sources of support for student sense making. While qualitative inferences may support sense making with quantitative problems, that qualitative reasoning itself may develop slowly through quantitative experience. That is, students may not apply the general relationships between addends and sums, or multiplicands and products, etc., until after they have extensive practice with those operations or equations.

## Acknowledgements

This work was supported in part by Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education (# R305B090023), and the PSLC which is funded by NSF award SBE-0836012.

## References

Beckmann, S. (2004). Solving Algebra and Other Story Problems with Simple Diagrams: a Method Demonstrated in Grade 4–6 Texts Used in Singapore. *The Mathematics Educator*, 14(1), 42–46.

Booth, J. L., & Koedinger, K. R. (2011). 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(3), 492–511.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing Conceptual Understanding and Procedural Skill in Mathematics: An Iterative Process. *Journal of Educational Psychology*, 93(2), 346–362.

Schoenfeld, A. H. (1988). When Good Teaching Leads to Bad Results: The Disasters of “Well Taught” Mathematics Courses. *Educational Psychologist*, 23(2), 1–22.

Stampfer, E., & Koedinger, K. R. (2013). When seeing isn't believing: Influences of prior conceptions and misconceptions. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1384–1389). Berlin, Germany: Cognitive Science Society.

Wiese, E. S., & Koedinger, K. R. (2014). How useful are fraction bars for understanding fraction equivalence and addition? A difficulty factors assessment with 5th, 6th, and 7th graders. *Presented at the Spring 2014 Meeting of the Society for Research on Educational Effectiveness*.

Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A., ... Ogbuehi, P. (2012). *Improving Mathematical Problem Solving in Grades 4 Through 8*.