In Papafragou, A., Grodner, D., Mirman, D., & Trueswell, J.C. (Eds.), Proceedings of the 38th Annual Conference of the Cognitive Science Society (954-959). Austin, TX: Cognitive Science Society. Poster presented at CogSci, Philadelphia, PA. (August 2016)

Benefits for Grounded Feedback over Correctness in a Fraction Addition Tutor

Eliane Stampfer Wiese (eliane.wiese@berkeley.edu)
Graduate School of Education, 4407 Tolman Hall
Berkeley, CA 94720 USA

Rony Patel (rbpatel@andrew.cmu.edu)
Kenneth R. Koedinger (krk@cs.cmu.edu)
Department of Psychology, 342c Baker Hall
Pittsburgh, PA 15213 USA

Abstract

Do students activate conceptual and procedural knowledge simultaneously when learning fraction addition? In grounded feedback, student actions on a target, to-be-learned representation are reflected in a more familiar feedback representation to promote conceptual learning within procedural practice. An experiment with 163 4th and 5th graders shows improved learning with a grounded feedback tutor over a symbols-only control with step-level right/wrong feedback. Learning with grounding also transferred to symbols-only assessment items, providing some support for the simultaneous activation view.

Keywords: fraction addition; simultaneous activation; magnitude representation

Introduction

Math and science are often communicated with abstract symbols. Learning these domains involves fluently using these symbols and correctly applying conceptual principles to them. How might a second representation provide grounding for learning a symbolic representation? *Grounded feedback* is based on the common characteristics of tutor designs that were previously shown to be successful: students manipulate a to-be-learned representation, while a linked representation reflects those inputs in a more accessible form (Mathan & Koedinger, 2005; Nathan, 1998). This feedback aims to take the less-familiar representation that the student is learning and ground it both in another representation and in the student's prior knowledge. We hypothesize that grounded feedback allows the student to apply her prior conceptual knowledge to the more-familiar feedback representation and then decide if her work with the to-be-learned representation is correct. This hypothesis follows from Ohlsson's theory of learning from performance errors: learners identify errors when there is a discrepancy between what the learner expects and what actually happens (Ohlsson, 1996). Grounded feedback provides the context in which the discrepancy can occur. For example, a learner may guess that 1/10 is larger than 1/4 because 10 is bigger than 4. Comparing two equal-sized rectangles, one with 1/10 and one with 1/4 shaded should alert the learner to his error: he expects 1/10 to have more shaded, but sees that it has less. The more accessible rectangle representation serves to disambiguate the meaning of the symbolic representation (Ainsworth, 1999). Importantly, in grounded feedback students do not directly manipulate the more accessible representation. Transfer between symbolic and non-symbolic representations is difficult for students (Uttal et al., 2013), likely because the cognitive demands of working in each type of representation are different (Sarama & Clements, 2009). Therefore, while grounded feedback includes an accessible representation to facilitate sense making and self-evaluation, having students act directly on the to-be-learned representation encourages transfer. Specifically, this grounded feedback tutor teaches fraction addition, and students act directly on the symbolic fractions.

Solving a fraction addition problem correctly may involve at least two steps: rejecting incorrect strategies and using the correct strategy. The most common incorrect strategy for fraction addition problems is the independent whole number strategy (Ni & Zhou, 2005). To execute this strategy, students independently add the numerators and denominators of the addends to get a final answer. For example, when adding 1/2 and 1/3, students incorrectly executing the independent whole number strategy would get the answer 2/5. 2/5 is less than the expected sum. In fact, it's even less than one of the addends (1/2). 34% of fraction addition and fraction subtraction problems resulted in errors due to the use of the independent whole number strategy for 6th and 8th graders (Siegler, Thompson, & Schneider, 2011). Grounded feedback can help students realize why the independent whole number strategy is incorrect (i.e., it results in a magnitude-incongruent answer) and why the correct strategy is correct (i.e., it results in a magnitude-congruent answer).

How might fraction magnitude knowledge and fraction arithmetic knowledge be related? The dynamic view proposes the two are independent and become progressively more so over time (Anderson, 1983). In contrast, the simultaneous activation view argues that arithmetic computation errors are the product of a lack of relevant concepts being simultaneously activated to reduce implausible solutions (in this case, magnitude knowledge; Hiebert, 1987). To test these two views, Byrnes and Wasik (1991) ran two studies to establish temporal precedence between fraction magnitude knowledge and fraction arithmetic knowledge and empirically intervene to teach fraction arithmetic knowledge with or without magnitude. Results showed that teaching with magnitude did not improve learning above a purely procedural approach. Students made arithmetic errors despite having an abundance of magnitude knowledge, suggesting they had the conceptual knowledge to reject the arithmetic errors, but were not bringing that knowledge to bear, as the simultaneous view would suggest. However, while the magnitude instruction included a *demonstration* of using fraction bars and coordinating between the two representations, students did not actually *practice* this skill themselves. Coordinating the fraction bars and fraction symbols is not trivial for students (Stampfer & Koedinger, 2013), and this coordination may be a pre-requisite skill for the activation of conceptual knowledge in a procedural context. To that end, the grounded feedback condition includes pre-instruction on interpreting the fraction bars, and the grounded feedback tutor showed students the magnitudes of the converted fractions and sums that they were proposing. The dynamic magnitude representations were intended to help students bring their existing magnitude knowledge to bear while practicing the procedure. Our experiment supports this notion, as the grounded feedback students matched the control students on fraction addition gains while outperforming them on more conceptual questions.

A previous experiment comparing grounded feedback to a symbols-only control (the correctness tutor) found similar pretest to posttest gains for both conditions, though the grounded condition had greater pre-test to delayed gains (Wiese, 2015). One explanation for why the grounded condition did not outperform the control was that grounded students often seemed unable to correctly interpret and integrate both representations (Ainsworth, 1999; Wiese, 2015). The current experiment investigates if pre-instruction on the feedback representation and a longer intervention time can lead to greater learning gains relative to a control.

Grounded Feedback for Fraction Addition

Figure 1 shows a screenshot from the grounded feedback tutor, constructed with CTAT (Aleven, McLaren, Sewall, & Koedinger, 2006). Students input numbers at the bottom of the interface, while fraction bars reflect the converted and sum fractions in a more concrete form. The fraction bars aim to ground the symbolic fractions by making their magnitude more salient. In addition, the grounding relies on students' prior knowledge of equivalence: equivalent fractions have the same magnitude, so equivalent fraction bars have the same amount colored in. Grounded feedback allows students to see the consequences of their errors and thus may promote students' evaluation of their own work (e.g., a student may guess that 8/24 + 9/24 = 17/48, but the fraction bars show 17/48 is too small). While the grounded feedback tutor offers on-demand text hints, it does not provide step-level right/wrong feedback, and does not prevent students from erasing correct inputs. A previous experiment compared this tutor to a correctness tutor, which did not include fraction bars but did provide immediate step-level feedback (correct inputs were colored green and incorrect inputs were colored red) (Wiese, 2015). In the correctness tutor, students were not permitted to change correct inputs. With both tutors, students were required to solve the current problem correctly before moving on.

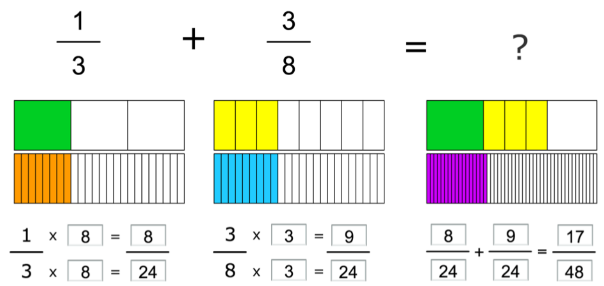

Figure 1: Grounded feedback tutor. Top row of fractions and yellow and green fraction bars are given, second row of bars dynamically shows students' inputs as they are typed in boxes at the bottom. Not shown: the window for on-demand hints, and the "done" button.

Prior Research on the Grounded Feedback Tutor

Prior work found that students learned from the grounded feedback tutor, but also indicated that students found the feedback unclear. Participants in a think-aloud study used the fraction bar feedback to identify and fix mistakes (Stampfer, Long, Aleven, & Koedinger, 2011) and a classroom study with 5th graders found learning benefits (Wiese, 2015). However, those students did not use the fraction bars effectively - they often clicked the "done" button when the fraction bars did not line up (Wiese, 2015). A follow-up study assessed how well 5th graders could evaluate fraction addition equations when fraction bars were provided as scaffolds. Equations were presented in four formats: three included fraction bars, and one was a numbers-only control (Fig. 2) (Stampfer & Koedinger, 2013). Students saw one correct and one incorrect equation in each format, and were asked to indicate if the equation was true or false. Incorrect sums were obtained by adding the numerators and denominators independently. The average of students' scores with the numbers-only format was 21%, far below their performance with the fraction bars. Still, performance with the fraction bars was low: 79% with the pictures-only format, 64% with pictures and numbers, and 46% with half pictures and numbers (Stampfer & Koedinger, 2013). These scores indicate that, while the fraction bars improve performance, they are not enough for students to reliably determine when an equation is correct or not, explaining students' confusion with the tutor.

Pre-Instruction on the Fraction Bar Representation

To help students interpret the fraction bar representations, the current grounded feedback tutor includes up-front instruction on the fraction bars. The instruction consists of multiple-choice problems, beginning with questions on fraction equivalence (expected to be within students' prior knowledge; Stampfer & Koedinger, 2013) and gradually fading in the addition operations and fraction symbols. This progression is based on concreteness fading (Fyfe, McNeil, Son, & Goldstone, 2014). Students were given immediate

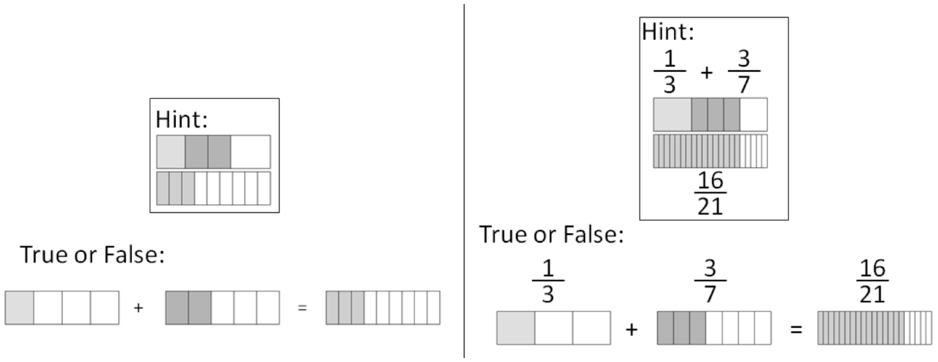

Figure 2: Sample addition questions in the three formats that included fraction bars. From left to right: pictures only, pictures and numbers, half pictures and numbers. The numbers-only format showed the symbolic equation and answer options without any fraction bars.

correctness feedback and on-demand hints. Sample problems are shown in Figs. 3-5.

Experiment: Grounded vs. Correctness

This experiment compared learning with the grounded and correctness feedback tutors, using a pretest-intervention-posttest design. Both tutors included the same brief instruction on using the tutor software and on fraction addition. The grounded feedback tutor included the pre-instruction on fraction bars.

Materials, Participants, and Procedures

The 29-question pre- and posttests included 12 symbolic fraction addition items and 9 evaluation items that proposed a fraction addition equation and asked if the sum was correct, too big, or too small (3 each of pictures only, numbers only, and both pictures and numbers). Answers were scored 1 if correct and 0 otherwise. Two matched tests (same problem types, different numbers) were counterbalanced, question order was determined randomly, and half of the tests were given in reversed question order.

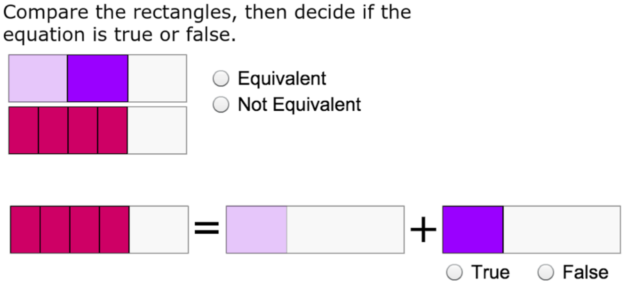
Compare the rectangles, then decide if the equation is true or false.

Figure 3: Question 1. 53% of students solved the problem, without hints, on their first try.

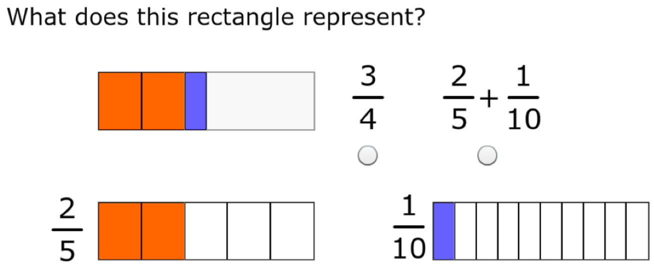
What does this rectangle represent?

Figure 4: Question 10. 81% of students solved the problem, without hints, on their first try.

194 students from 9 classes at a local public school participated in the experiment (60 4th graders and 134 5th graders). The school tracked students by achievement, and teachers identified their classes as high (3), average (5), or low (1). 31 students were removed from the sample because they were absent during the pre- or posttest, or they spent less than 45 minutes on their assigned tutor, leaving 163 students (78 grounded, 85 correctness). The experiment took place at the school during class time over four consecutive days. All random assignment was within-class. Students were given a 15-minute pretest, worked with a randomly assigned tutor for up to 80 minutes, and then took a 15-minute posttest the next day. The tests were administered on a computer and students could not return to previously answered questions.

Results

Table 1: Average scores (and standard deviations) for overall tests and subtests.

Condition	Test	Total	Addition	Evaluation	Other
Correctness	Pre	.43 (.20)	.32 (.27)	.42 (.26)	.60 (.24)
	Post	.59 (.22)	.49 (.30)	.63 (.26)	.69 (.18)
Grounded	Pre	.42 (.19)	.35 (.26)	.42 (.23)	.57 (.23)
	Post	.63 (.22)	.55 (.32)	.69 (.23)	.71 (.22)

Did the grounded condition learn more than the correctness condition? Overall, yes. Table 1 shows the average scores for the overall pre- and posttests and for the

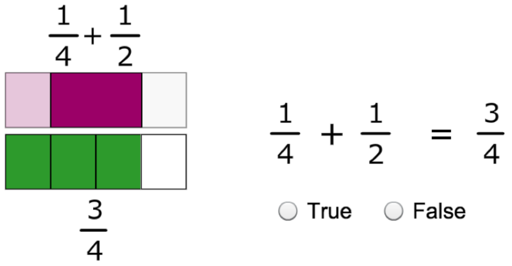

Figure 5: Question 14. 68% of students solved the problem, without hints, on their first try.

three subtests, by condition. To test that pretest differences were not significant, an ANOVA was run on pretest score, with pretest order, pretest form, class tracking level, and condition as fixed factors, and class as a random factor. The first model included all main effects and two-way interactions. After removing non-significant interactions and main effects, the final model included a marginal effect for order ($p = .07$), a marginal order by pretest form interaction ($p = .08$) and a significant class by pretest form interaction ($p = .04$). Condition was not significant ($p = .7$). Paired samples t-tests show all within-condition differences from pre- to posttest are significant ($p < .01$). To test if condition had a significant effect on learning, we re-ran the final model, this time on posttest score, with pretest score as a covariate. The first model included all two-way interactions with pretest score. After removing non-significant interactions and main effects, the final model included class and total pretest score as significant main effects (both $p < .01$) and condition as a marginal main effect ($p = .065$), in favor of grounded feedback. The same tests were repeated on the addition and evaluation subtests – condition was not significant in either case.

How did transfer from the grounded tutor to a symbols-only assessment compare to transfer from the symbols-only tutor to a dual-representation assessment? To determine if there were condition differences for scores on the numbers only and pictures and numbers evaluation items, a MANOVA was run on the posttest scores for each scaffold type, with corresponding pretest scores as covariates and class and condition as fixed factors. The condition by class interaction was not significant in the multivariate test so the model was re-run without it. Multivariate tests showed pretest scores and class were significant ($p < .04$), as was condition ($p = .047$), in favor of grounded feedback. Condition was significant on the posttest score for the pictures and numbers scaffold ($p = .015$, again in favor of grounding), but not for the numbers only scaffold. Figure 6 shows the estimated marginal means for the two scaffold types, by condition.

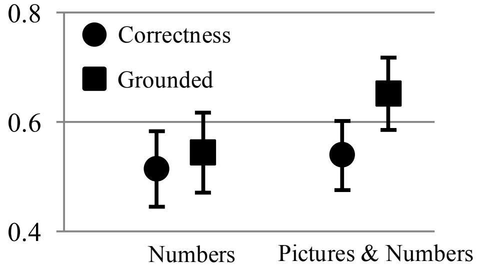

Figure 6: Estimated marginal means for posttest evaluation items that included numbers, with 95% confidence intervals (y-axis is from .4 to .8).

Did Students Learn from the Fraction Bar Pre-Instruction? The fraction bar instruction aimed to help students interpret the grounded feedback. One measure of success is how often students pressed the "done" button when the proposed sum differed from the correct some by more than .1: on average .16 times per problem (.34 on average for the first 20 problems, compared to .99 for the 20 problems in the previous study; Wiese, 2015). Another measure of learning comes from a two-question pre- and posttest bracketing the pre-instruction. Similar to the question shown in Fig. 5, the test questions proposed a fraction addition equation with the fractions represented both symbolically and as fraction bars. Students indicated if the proposed sum was correct, too big, or too small. These pre- and posttests included one true equation and one false equation, where the sum was obtained by adding the numerators and denominators independently. Both before and after instruction, the average score was 63% correct. Errors were categorized as whole number error, other error, or skipped. A whole number error indicates incorrect transfer from whole number addition: answering 'correct' to a sum obtained by adding the numerators and denominators of the addends, and answering 'too big' to the correct sum. Answers that were not correct or whole number errors were coded as other. Table 2 shows the proportion of each error at the fraction bar pre- and posttest (this table includes the 95 students who completed this section, not just the 78 grounded students included in the other analyses).

Table 2: Proportion of correct answers and error types for the fraction bar pre- and posttest

	Correct	Whole Number Error	Other Error	Skipped
Pre	63%	30%	6%	1%
Post	63%	23%	13%	1%

Table 3: Pearson correlations between error types on evaluation items and performance on free-response fraction addition items. $*p < .03$

Response on Fraction Addition Items	Whole Number Error	Other Error	Correct
Percent Correct	-.42*	.12	.30*
Rate of Whole-Number Error	.31*	-.11	-.21*

After the fraction bar instruction, students had fewer whole number errors. To determine if one type of error indicates better understanding, we examined correlations between each type of error and proficiency at fraction addition problems. The study pretest included two evaluation questions that were isomorphic to those used in the fraction bar pre- and posttest, and 12 free-response symbolic fraction addition problems. For this analysis we include students who saw both of the evaluation questions, and calculated scores and error rates on the addition items based on the questions that students saw (i.e., disregarding questions that students ran out of time for). Table 3 shows the correlations between occurrence of each error type and (1) score on the fraction addition items and (2) rates of student-generated whole-number errors on the addition items. These results show that correct responses on the evaluation items are correlated positively with correct responses on the fraction addition items and correlated *negatively* with whole-number errors on the fraction addition items; the reverse is true for whole-number errors on the evaluation items.

Case Studies: Using Grounded Feedback

Is grounded feedback easier to work with than correctness feedback? On average, students in the grounded condition solved fewer fraction addition problems (38 vs. 74 for correctness), took longer per problem (65 seconds per problem vs. 40), and requested more hints per problem (1.4 vs. 0.4), indicating that grounded feedback was more difficult.

How did students make use of the grounded feedback? Log data suggests two pathways: responding to the grounded feedback directly to diagnose and correct errors, and using grounded feedback to decide when to ask for a hint. Figures 7-8 illustrate the first strategy for a student converting 3/8 to 24ths. The student is adding 1/3 and 3/8, and got a hint for the denominator of the first fraction that said to multiply 3 by 8. The student correctly chose to multiply 8 by 3 to get the denominator for the second fraction, but then decided to multiply the numerator by 6. Figure 7 shows the student's interface at this point. The grounded feedback shows that 18/24 is bigger than 3/8. Next, the student tries 10 as a numerator (still to big), and then 9 (Fig. 8). After the grounded feedback shows that 9/24 equals 3/8, the student updates the multiplication area to show 3 x 3 = 9. In this case, the student does not seem able to find the equivalent fraction using symbols alone: the student does not begin by multiplying the numerator and denominator by 3. Instead, the student appears to use the grounded feedback to inform a guess-and-check strategy, identifying the direction of the error and correctly deciding when that part of the problem is complete (after converting the second fraction, the student moves on to the sum).

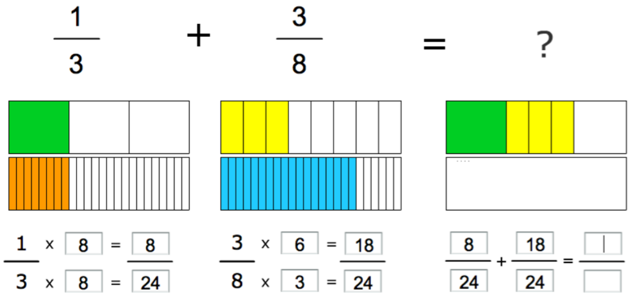

Figure 7: The grounded feedback tutor. The student is converting 3/8 to 24ths

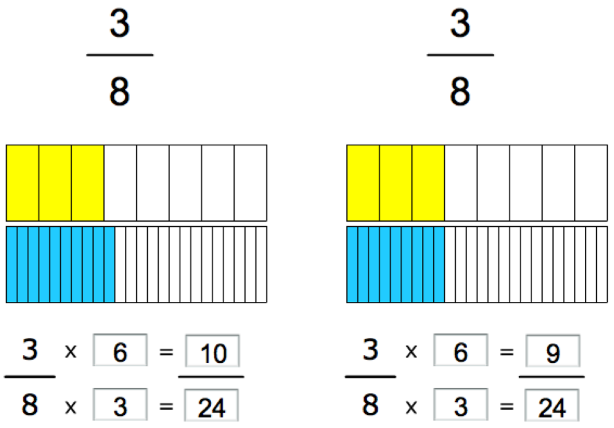

Figure 8: Grounded feedback for each guess-and-check conversion attempt

In other cases, the feedback may facilitate learning from hints. In one example, a student adding 4/9 and 1/9 entered 5/18 for the sum (the whole-number error). The student seems to interpret the feedback as showing an error, but appears unsure of how to fix it. Instead of pressing the done butting or guessing, the student asks for hints until the answer is provided. On the next problem, the student converts the addends incorrectly, and then uses the whole-number strategy on the converted fractions, again asking for a hint only after entering the incorrect sum (perhaps the student pays more attention to the addition section of the interface than the converting sections, or the student might not realize that the converted fractions should be equivalent to the addends). This student does not attempt the whole-number strategy on any subsequent problems. Here, the grounded feedback appears to have shaken this student's confidence in that incorrect strategy, perhaps facilitating acceptance of the correct strategy offered in the hints.

Discussion

Correctness feedback is easier to work with than grounded feedback, indicated by students solving many more correctness problems, spending less time per problem, and requesting fewer hints on each problem. How does the additional difficulty of grounded feedback affect learning? The marginal significance in favor of grounded feedback on overall learning and the non-significant difference on the addition subtest indicates that grounded feedback is no worse than correctness. The differences in learning on the evaluation items with pictures and numbers also suggest that the additional difficulties in grounded feedback are desirable. Those items include the same representations present in the grounded tutor. The numbers-only evaluation items only included the symbolic representation present in the correctness tutor. Therefore, the pictures and numbers items can be considered target items for the grounded students while the numbers only items are transfer, and visa versa for the correctness students. With this view, the grounded feedback students were better than the correctness students at transferring their knowledge to the less-familiar format: Grounded students scored just as well on the numbers only problems as the correctness students, while outperforming them on the pictures and numbers items. At the very least, the similar performance of both conditions on the fraction addition items and numbers only evaluation items shows that including the fraction bars during learning did not impede students' performance with numbers on the posttest.

Did students learn from the fraction bar tutorial? Scores on the evaluation items bracketing the pre-instruction did not change. However, students decreased their rates of whole number errors, switching to other errors instead. Whole number errors are negatively correlated with solving symbolic fraction addition problems correctly and are positively correlated with adding both numerators and denominators independently on such problems, while other errors are not correlated with either behavior. Therefore, whole number errors appear to be more harmful than other errors, and a decrease in whole number errors suggests that students benefitted from the tutorial.

These results indicate that a longer intervention time (80 vs. 40 minutes) and the inclusion of fraction bar pre-instruction addressed the shortcomings of the grounded condition in the previous study (Wiese, 2015). Still, the case studies point to further areas for improvement. Even with the grounded feedback, students do not always seem to recognize when their work is incorrect (e.g., a student may recognize when a proposed sum is incorrect but may not recognize when a converted fraction is incorrect). Including correctness feedback with the grounding may help: Instead of relying on the grounding alone to evaluate the action and diagnose the error, the correctness feedback will evaluate the error, freeing cognitive resources to focus on the diagnosis.

Conclusions

This study shows an advantage for grounded feedback, and certainly no disadvantage, compared to a strong control condition. Students in this study seemed to be better able to interpret the grounded feedback than students in the previous study (Wiese, 2015), although the measures used (rates of incorrectly pressing the "done" button and performance on evaluation items) may be overly coarse. Control students did purely procedural practice with the fraction addition items, and improved on all test sections from pre-test to post-test. Even though students in the grounded condition had their mental resources split between the procedure and the magnitude concepts, they improved just as much on symbolic fraction addition, and outperformed the control on the conceptual evaluation items with symbols and magnitude. The dynamic view would suggest that the grounded condition's improvement on the conceptual items should come at a cost to the procedural ones. That grounded students improved as much as the control on the procedural items offers some support to the simultaneous activation theory.

Acknowledgements

This work is supported in part by the Pittsburgh Science of Learning Center through NSF award SBE-0836012, and by Carnegie Mellon University's Program in Interdisciplinary Education Research (PIER) funded by Grant R305B090023 from the U.S. Department of Education, and by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C100024 to WestEd. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2-3), 131–152. doi:10.1016/S0360-1315(99)00029-9

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2006). The cognitive tutor authoring tools (CTAT): Preliminary evaluation of efficiency gains. In M. Ikea, K. Ashkely, & T.-W. Chan (Eds.), *Intelligent Tutoring Systems* (pp. 61 – 70). Springer-Verlag Berlin Heidelberg.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Byrnes, J. P., & Wasik, B. A. (1991). Role of Conceptual Knowledge in Mathematical Procedural Learning. *Developmental Psychology*, *27*(5), 777–786.

Fyfe, E. R., McNeil, N., Son, J., & Goldstone, R. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, *26*(1), 9–25.

Hiebert, J. (1987). *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, New Jersey: Erlbaum.

Mathan, S., & Koedinger, K. R. (2005). Fostering the Intelligent Novice: Learning From Errors With Metacognitive Tutoring. *Educational Psychologist*, *40*(4), 257–265. doi:10.1207/s15326985ep4004_7

Nathan, M. J. (1998). Knowledge and Situational Feedback in a Learning Environment for Algebra Story Problem Solving. *Interactive Learning Environments*, *5*(1), 135–159.

Ni, Y., & Zhou, Y.-D. (2005). Teaching and Learning Fraction and Rational Numbers: The Origins and Implications of Whole Number Bias. *Educational Psychologist*, *40*(1), 27–52.

Ohlsson, S. (1996). Learning from Performance Errors. *Psychological Review*, *103*(2), 241–262.

Sarama, J., & Clements, D. H. (2009). "Concrete" Computer Manipulatives in Mathematics Education. *Child Development Perspectives*, *3*(3), 145–150. doi:10.1111/j.1750-8606.2009.00095.x

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296. doi:10.1016/j.cogpsych.2011.03.001

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.

Stampfer, E., Long, Y., Aleven, V., & Koedinger, K. R. (2011). Eliciting Intelligent Novice Behaviors with Grounded Feedback in a Fraction Addition Tutor. In *Proceedings of the 15th annual conference on Artificial Intelligence in Education*.

Uttal, D. H., Amaya, M., Maita, M. del R., Hand, L. L., Cohen, C. A., O'Doherty, K., & Deloache, J. S. (2013). It works both ways: Transfer difficulties between manipulatives and written subtraction solutions. *Child Development Research*, *2013*. doi:10.1155/2013/216367

Wiese, E. S. (2015). *Toward sense making with grounded feedback*. Carnegie Mellon University. Retrieved from http://reports-archive.adm.cs.cmu.edu/anon/hcii/CMU-HCII-15-104.pdf