Grounded Feedback in a Fraction Addition Tutor

Eliane Stampfer Wiese, *Carnegie Mellon* stampfer@cs.cmu.edu
Kenneth R. Koedinger, *Carnegie Mellon* koedinger@cs.cmu.edu

Objective: To determine when providing a second representation as feedback during problem solving enhances learning in mathematics.

Perspective: 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? By "grounding" we mean connecting to a more concrete or familiar foundation on which to make sense or build meaning. 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 form that is easier to reason with (Mathan & Koedinger, 2005; Nathan, 1998). Grounded feedback is grounded "in the world" and in the student's prior knowledge (Stampfer & Koedinger, 2012).

We hypothesize that grounded feedback allows students to apply their prior conceptual knowledge to the more-familiar feedback representation and then decide if their 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 shaded and one with 1/4 shaded should alert the learner to her error: she expects 1/10 to have more shaded, but sees that it has less. The more-familiar rectangle representation serves to disambiguate the meaning of the less-familiar symbolic representation (Ainsworth, 1999). Importantly, in grounded feedback students do not directly manipulate the familiar 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 a familiar representation to facilitate sense making and self-evaluation, transfer is encouraged by having students act directly on the to-be-learned representation. However, few experiments have compared grounded feedback to other types of support. Some found robust improvements in learning compared to problem-solving with right/wrong feedback (Mathan & Koedinger, 2005; Nathan, 1998), while others found no differences (Yerushalmy, 1991), or benefits for worked examples over grounded feedback when using three representations (Rau, Aleven, Rummel, & Rohrbach, 2012). More research is needed to identify the contexts and design strategies that make grounded feedback useful.

The current study is a follow-up to a previous experiment comparing a grounded feedback fraction addition tutor to a symbols-only control with right/wrong feedback (Stampfer & Koedinger, 2012). That study found greater pretest to posttest gains for the control condition, though the grounded condition caught up at the delayed posttest. One explanation for these results was that grounded feedback students often seemed unable to correctly interpret and integrate both representations (Stampfer & Koedinger, 2012), perhaps due to gaps in prior knowledge or still-developing meta-cognitive skills. 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.

Methods and Materials: We experimentally compared grounded feedback to immediate, step-level right/wrong feedback in intelligent tutors 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

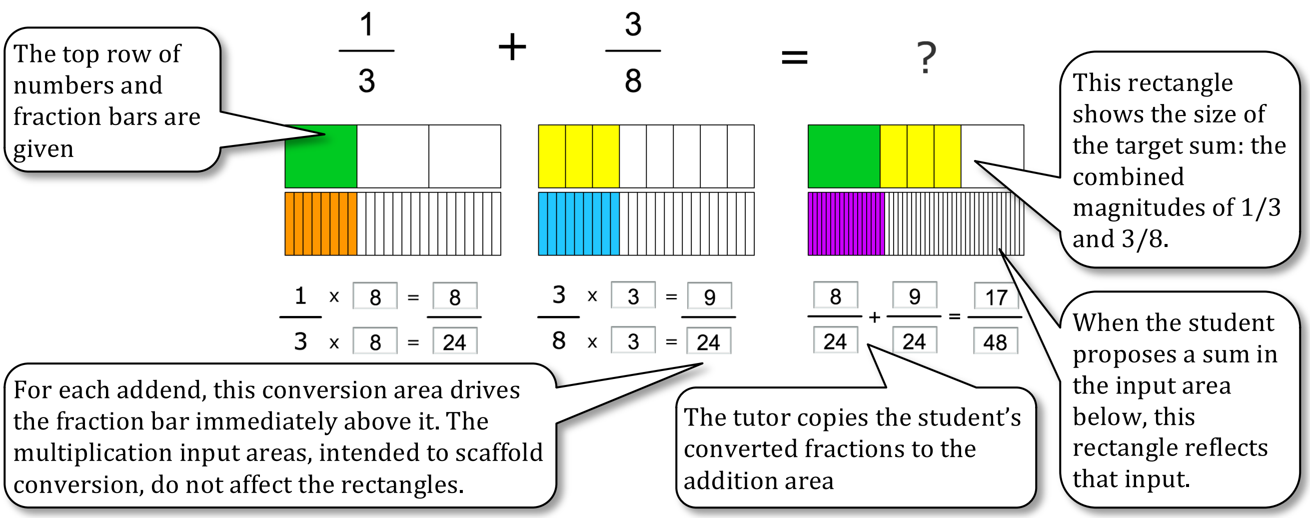

Figure 1. Grounded feedback tutor. This student has converted correctly but added incorrectly. The tutor does not provide step-level right/wrong feedback. The control tutor has no rectangles, and inputs are immediately colored green if correct and red otherwise. Both tutors include a message window for text hints (not shown).

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.

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). For these problems, students were given immediate right/wrong feedback and on-demand hints. Sample problems are shown in Figures 2-4.

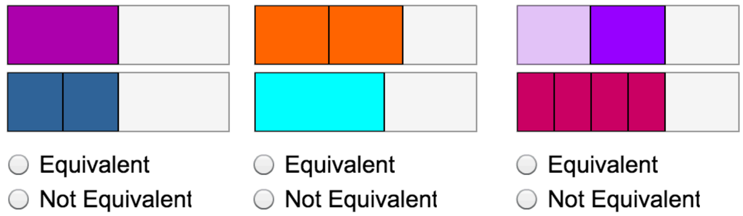

Figure 2. Question 1 of the up-front fraction bar instruction. 72% of students solved the problem, without hints, on their first try.

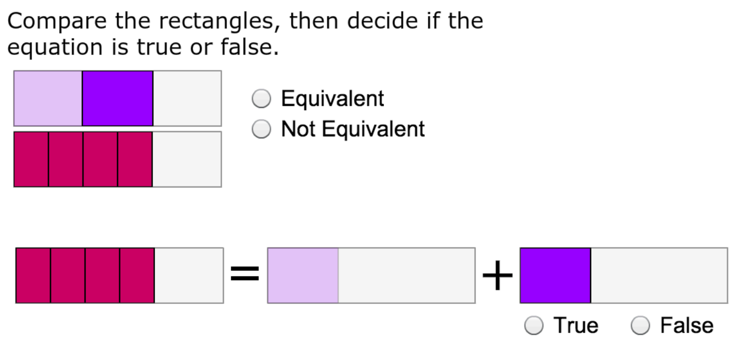

Figure 3. Question 5 of the up-front fraction bar instruction. 53% of students solved the problem, without hints, on their first try.

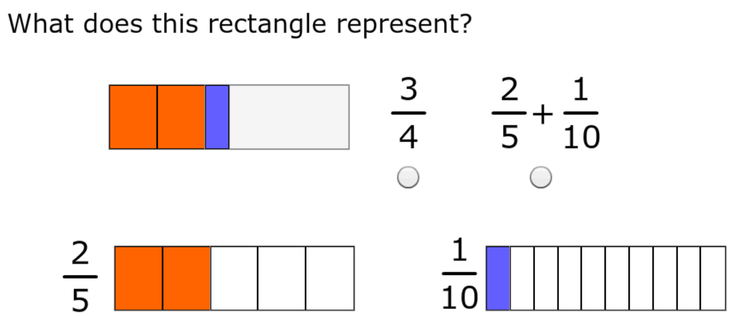

Figure 4. Question 10 of the up-front fraction bar instruction. 81% of students solved the problem, without hints, on their first try.

The control tutor gave the same fraction addition problems as the grounded feedback tutor, but did not show any fraction rectangles. Instead, the control tutor provided immediate right/wrong feedback on each step, and did not allow students to change correct inputs. The control tutor did not include the up-front fraction bar instruction. With both tutors, students were required to solve the current problem correctly before moving on to the next one. All components of the study were done online, and student actions were logged to DataShop (Koedinger et al., 2010). To analyze learning, we examined students' pre- and post-test scores.

Participants, Procedures, and Data: 194 students from 9 classes at a local public school participated in the experiment (60 4th graders and 134 5th graders). 31 students were removed from the sample because they were absent during the pre- or post-test, or they spent less than 45 minutes on their assigned tutor, leaving 163 students (78 grounded, 85 control).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 (including the fraction bar instruction for the grounded feedback condition), and then took a 15-minute posttest the next day. Pre- and post-tests were administered on a computer and students could not return to previously answered questions.

The 29-question pre- and post-tests included 12 symbolic fraction addition items, 9 evaluation items, and 8 other items. Evaluation items 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; see figures 5-6). Answers were scored 1 if correct and 0 otherwise. Two matched tests were counterbalanced, question order was determined randomly, and half of the tests were given in reversed question order.

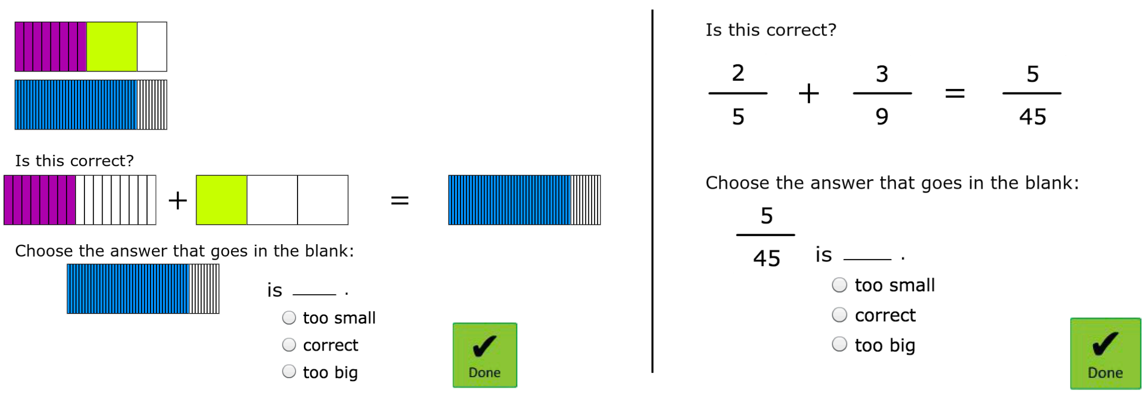

Figure 5. Sample evaluation items with pictures only (left) and numbers only (right).

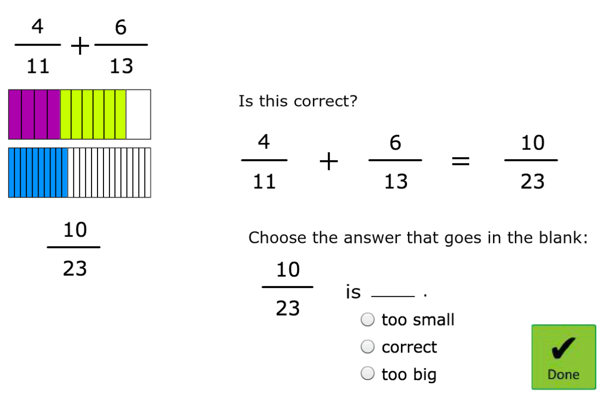

Figure 6. Sample evaluation item with pictures and numbers.

Results: Both conditions improved from pre- to post-test. Figures 7 and 8 show the average scores for the overall pre- and post-tests and for the addition, evaluation, and other items, by condition. Paired samples t-tests show all within-condition differences from pre- to posttest are significant ($p < .01$) – evidence of learning from both tutors.

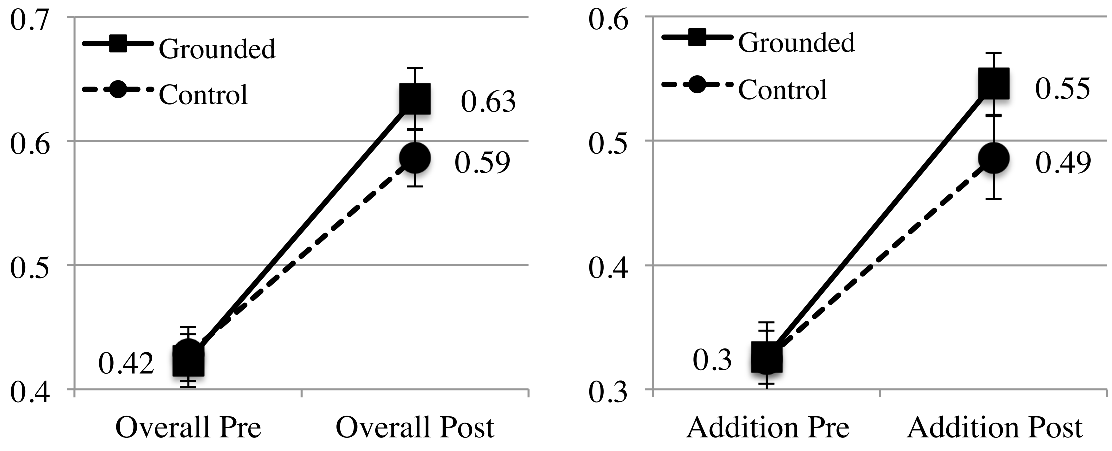

Figure 7. Average scores for the overall pre- and post-test (left) and for the addition items (right), by condition, with standard error bars (note: y-axis scales are different).

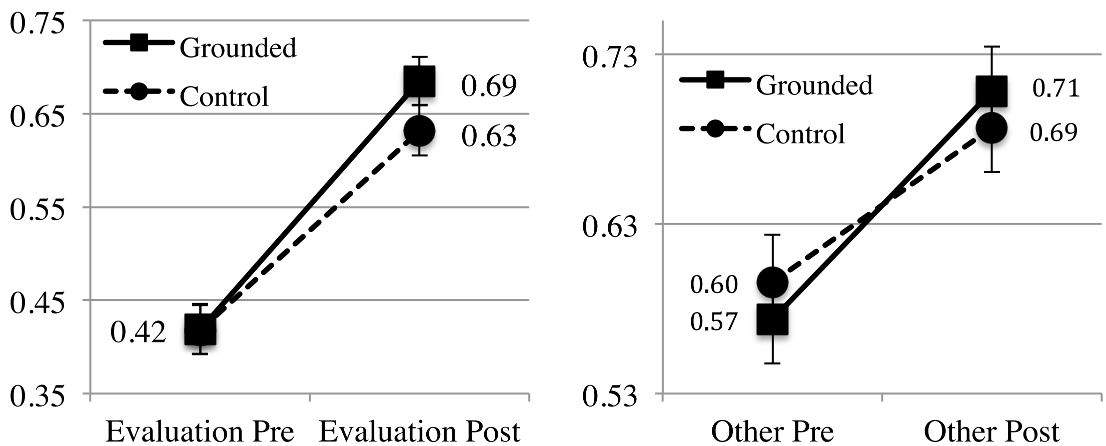

Figure 8. Average scores for the evaluation items at pre- and post-test (left) and for the other items (right), by condition, with standard error bars (note: y-axis scales are different).

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. There were no significant condition differences at pre-test ($F(1, 142) = .123$, $p = .7$).

To test if condition had a significant effect on learning, we re-ran the model used for pre-test score, this time on post-test 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 (class: $F(8,152) = 4.19$, $p < .01$; total pretest score: $F(1, 152) = 81.6$, $p < .01$) and condition as a marginal main effect ($F(1, 152) = 3.46$, $p = .065$), in favor of grounded feedback. The same tests were repeated on the addition and evaluation items, but condition was not significant.

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 post-test scores for each evaluation type, with corresponding pretest scores as covariates and class and condition as fixed factors. Multivariate tests showed pretest scores and class were significant (pictures and numbers pre-test: $F(2, 150) = 3.5$, $p = .031$; numbers only pre-test: $F(2, 150) = 12.1$, $p < .001$; class: $F(16, 303) = 2.29$, $p = .003$), as was condition ($F(2, 150) = 3.11$, $p = .047$), in favor of grounded feedback. Condition was significant on the post-test score for the pictures and numbers items ($F(1, 151) = 6.07$, $p = .015$, again in favor of grounding), but not for the numbers only items ($F(1,151) = .356$, $p = .552$). Figure 9 shows the estimated marginal means for the two evaluation types, by condition.

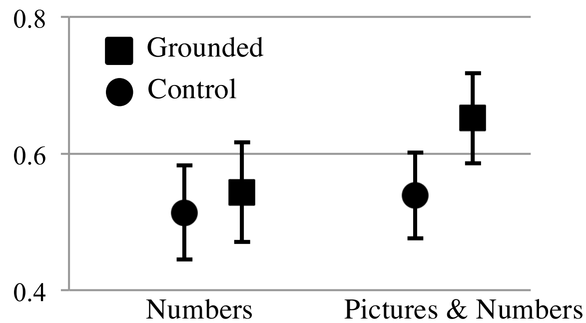

Figure 9. Estimated marginal means for the numbers only and pictures and numbers posttest evaluation items, with 95% confidence intervals.

Significance: This study provides evidence that grounded feedback may be more effective for learning than step-level right/wrong feedback in the target representation. The results from the symbolic fraction addition items indicate that students who learned with concrete and symbolic representations (the grounded feedback condition) could demonstrate transfer to symbols-only questions. Since total time with the tutors was controlled, these results show that giving students time to practice with the feedback representation did not detract from their learning of the target material.

The results from the evaluation items further suggest that grounded feedback promotes transfer better than a symbols-only control. The numbers-only evaluation items only included the symbolic representation present in the control tutor, while the pictures and numbers evaluation items included both representations from the grounded 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 control students. With this view, the grounded feedback students were better than the control students at transferring their knowledge to the less-familiar format: Grounded students scored just as well on the numbers only problems as the control students, while outperforming them on the pictures and numbers items.

By comparing grounded feedback to a robust control, this study adds to the very limited literature on grounded feedback. It provides a demonstration of effectiveness of the approach, not only in producing learning, but also in possibly producing more learning than a tight, high-bar control condition. While previous studies with adults found benefits for grounded feedback (Mathan & Koedinger, 2005; Nathan, 1998), studies with children did not (Rau et al., 2012; Stampfer & Koedinger, 2012; Yerushalmy, 1991). The current study provided up-front instruction on interpreting the grounded feedback, a component not present in previous work on the fraction addition tutor (Stampfer & Koedinger, 2012). This indicates that children may benefit more from grounded feedback when they are supported in interpreting the feedback and coordinating the symbolic and feedback representations – meta-cognitive tasks which may come easier to adults. While the study is certainly not conclusive, these results suggest that grounded feedback holds promise for middle-school students when they are given instruction in interpreting it. Further research is warranted, especially on the mechanisms and developmental trajectories for student learning with grounded feedback.

Acknowledgements: This work is supported in part by Carnegie Mellon University's Program in Interdisciplinary Education Research (PIER) funded by grant number R305B090023 from the US 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. Retrieved from http://www.springerlink.com/index/k037578x2q825156.pdf

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.

Koedinger, K. R., Baker, R. S. J., Cunningham, K., Skogsholm, A., Leber, B., & Stamper, J. (2010). A Data Repository for the EDM community : The PSLC DataShop. In C. Romero, S. Ventura, M. Pechenizkiy, & R. S. J. d. Baker (Eds.), *Handbook of Educational Data Mining* (pp. 1–21). Boca Raton, FL: CRC Press.

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.

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

Rau, M. A., Aleven, V., Rummel, N., & Rohrbach, S. (2012). Sense Making Alone Doesn't Do It: Fluency Matters Too! ITS Support for Robust Learning with Multiple Representations. In S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems* (pp. 174–184). Springer Berlin Heidelberg. doi:10.1007/978-3-642-30950-2

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

Stampfer, E., & Koedinger, K. R. (2012). Tradeoffs between Immediate and Future Learning. *Paper presented at the European Association for Learning and Instruction*. Bari, Italy.

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.

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

Yerushalmy, M. (1991). Effects of Computerized Feedback on Performing and Debugging Algebraic Transformations. *Journal of Educational Computing Research*, *7*(3), 309–330.