How to schedule multiple graphical representations? A classroom experiment with an intelligent tutoring system for fractions

Martina A. Rau, Carnegie Mellon University, Human-Computer Interaction Institute, 5000 Forbes Ave, Pittsburgh, PA 15213, USA, marau@cs.cmu.edu,
Nikol Rummel, Ruhr-Universität Bochum, Institute of Educational Research, Universitätsstraße 150, 44801 Bochum, Germany, nikol.rummel@rub.de
Vincent Aleven, Laura Pacilio, Zelha Tunc-Pekkan, Carnegie Mellon University, Human-Computer Interaction Institute, 5000 Forbes Ave, Pittsburgh, PA 15213, USA, aleven@cs.cmu.edu, lpacilio@andrew.cmu.edu, zelha@cs.cmu.edu

Abstract: Providing learners with multiple representations of the learning content has been shown to enhance learning outcomes. When designing problem sequences with multiple representations, designers of intelligent tutoring systems must decide how to schedule the representations. Prior research on contextual interference has demonstrated that interleaving different types of *learning tasks* can foster a deep understanding of the underlying concepts. Do the same advantages apply to interleaving *representations?* In a classroom experiment, we compared four conditions that varied the practice schedules of multiple graphical representations between interleaving and blocking. The multiple-representation conditions were compared to three single-representation control conditions. During their regular classroom instruction, 290 4th and 5th-grade students worked for five hours with versions of an intelligent tutoring system for fractions. On several dependent measures, interleaving multiple graphical representations led to better learning results than blocking multiple graphical representations. Findings from a think-aloud study give insight into the underlying cognitive processes.

Introduction

Graphical representations of learning contents are often used for instruction (Ainsworth, 2006). When used in learning technology, graphical representations can be especially useful since they allow for interactions that are physically impossible or very difficult to realize, for instance by dragging and dropping symbolic statements into a chart that automatically updates to display the information (Moyer, Bolyard, & Spikell, 2002). However, learning with multiple graphical representations is challenging. An important prerequisite for benefiting from *multiple* representations is the acquisition of *representational fluency*: students need to conceptually understand each of the representations, and they need to be able to use them to solve problems (Ainsworth, 2006). Furthermore, students need to develop *representational flexibility*: they need to understand the differences and similarities between the representations, they need to learn to relate the different representations to one another, and to use the different representations interchangeably to solve problems (Ainsworth, 2006; de Jong et al., 1998).

Fractions are one of many areas in mathematics where multiple graphical representations are used extensively (National Mathematics Advisory Panel, 2008). There are different conceptual interpretations of fractions, such as the measurement concept and the part-whole interpretation (Charalambous & Pitta-Pantazi, 2007). Each conceptual interpretation can be illustrated using a different graphical representation, such as number lines for measurement, or area models (e.g., circles or rectangles) to support part-whole interpretations. Multiple graphical representations may help students understand different conceptual aspects of fractions and thus gain a robust understanding of fractions. In a prior study, we found experimental evidence that students working with multiple graphical representation of fractions (e.g., circles, rectangles, and number lines) outperform students who work with a single graphical representation (a number line), although only when prompted to explain how the graphical representations (e.g., half a circle) relate to the symbolic representation (e.g., 1/2) (Rau, Aleven, & Rummel, 2009). These results demonstrate that understanding individual graphical representations (i.e., by relating graphically displayed information to the concepts of numerator and denominator) is essential in order for learners to benefit from multiple graphical representations.

When designing instruction that uses multiple graphical representations, curriculum designers must decide how to temporally sequence the different graphical representations. How frequently should the curriculum alternate between graphical representations? Practice schedules are likely to impact how students understand each graphical representation and, consequently, how well they learn the underlying mathematical concepts. In particular, it may matter whether the different representations are practiced in a “blocked” manner (e.g., A – A – B – B) or are interleaved with practice of other representations (e.g., A – B – A – B). Research on contextual interference has investigated scheduling effects of different task types. Results show that *interleaving task types* leads to better learning results than blocked practice (Battig, 1972; de Croock, van Merrienboer, &

Paas, 1998). A common interpretation of this finding is that interleaved practice encourages deep processing (de Croock et al., 1998). Since students cannot hold all relevant knowledge in working memory, they must reactivate task-specific knowledge as it comes up again in the task sequence. Another interpretation is that when frequently switching between task types, students will be more likely to abstract a common principle from the tasks than when switching infrequently between them (de Croock et al., 1998). They can do so, for instance, by comparing different task types to one another.

Against the background of research on blocking versus interleaving different task types, one could hypothesize that *interleaving* practice with different *graphical representations* (i.e. switching frequently between them) may lead to deeper processing of conceptual fractions knowledge and may encourage students to abstract a robust conceptual understanding from the multiplicity of graphical representations. These processes may help students acquire representational flexibility. On the other hand, our own previous work (Rau et al., 2009) leads to the hypothesis that developing representational fluency (i.e., coming to understand each single representation) may be a prerequisite for developing representational flexibility. Blocked practice with the representations may be successful in promoting representational fluency because it provides students with the opportunity to become fluent with one graphical representation before starting to work with a second graphical representation. This could be achieved by switching infrequently or with moderate frequency between the representations. If representation flexibility builds on representational fluency, as just argued, then students may benefit most from a condition that gradually moves from blocking to an increasingly interleaved schedule of multiple graphical representations.

To test the hypotheses mentioned above, we compared four conditions that varied the practice schedules of multiple graphical representations in a classroom experiment. Students in all conditions worked on the same problems, but multiple graphical representations were presented either in a blocked, moderately interleaved, fully interleaved, or increasingly interleaved manner. We also employed three single graphical representation control conditions in which students worked either with only a circle, a rectangle, or a number line. These control conditions allow us to replicate the results from our first study (which, as mentioned showed advantages for learning with multiple graphical representations over learning with a number line; see Rau, et al., 2009) and to extend this finding to working with only area models. In order to gain insights into the cognitive processes underlying the most successful practice schedule, we additionally conducted a small think-aloud study. We report the results of the think-aloud study after discussing the results from the classroom experiment.

We investigated the effects of interleaving multiple graphical representations in the context of a proven intelligent tutoring system technology, namely Cognitive Tutors (Koedinger & Corbett, 2006). Cognitive Tutors have a proven track record in improving students' mathematics achievement (Koedinger & Corbett, 2006). We developed several versions of an example-tracing tutor for fractions learning, using Cognitive Tutor Authoring Tools (Aleven et al., 2009). This type of tutor behaves like a Cognitive Tutor but relies on examples of correct and incorrect solution paths rather than on a cognitive model. The design of the fractions tutor was informed by results from our previous studies (Rau, Aleven, & Rummel, 2009; 2010), and by small-scale user studies.

Experimental Study

We investigated the effects of blocking versus interleaving multiple graphical representations in a classroom experiment. We expected that interleaving multiple graphical representations of fractions would enhance the acquisition of representational flexibility whereas blocking and moderately interleaving representations would enhance the acquisition of representational fluency. We expected that increasingly interleaving representations would be most successful if representational flexibility builds on representational fluency.

Methods

The tutors used in the study included three interactive graphical representations of fractions: circles, rectangles, and number lines (Figure 1). Our fractions tutor curriculum covered six task types: identifying fractions from multiple graphical representations, making multiple graphical representations of symbolic fractions, reconstructing the unit from unit fractions, reconstructing the unit from proper fractions, identifying improper fractions from multiple graphical representations, and making multiple graphical representations of improper fractions. The tutoring system takes a conceptually-focused approach in introducing fractions. A common theme throughout the fractions tutor was the unit of the fraction (i.e., what the fraction is taken of). The concept of the unit is being introduced in the first task types, and revisited during the later task types as students learn about improper fractions. Figure 2 shows an example of a problem in which students make circle representations for two given symbolic fractions and are then prompted to reflect on the relative size of the two fractions.

Students solved each problem by interacting both with symbols and with interactive graphical representations. Students manipulated the graphical representations in various ways: by clicking on fraction pieces to highlight them, by dragging and dropping fraction pieces, and through buttons to change the partitioning of the graphical representations. The tutor interfaces updated interactively after each step to show the next step to work on, and to emphasize parts of the graphical representations that were conceptually relevant for the subtask at hand through color-highlighting.

Students received error feedback and hints on all steps. Error feedback messages were designed to make students reconsider their answer using the multiple graphical representations, or by reminding them of a previously introduced principle. Hint messages provided conceptually oriented help in relation to the graphical representation. Each problem included conceptually oriented prompts to help students relate the multiple graphical representations to the symbolic notation of fractions. We found these prompts to be effective in an earlier experimental study (Rau et al., 2009).

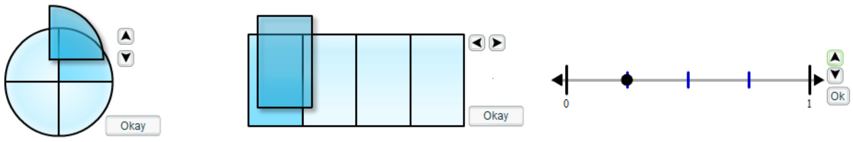

Figure 1. Interactive representations used in fractions tutor: circle, rectangle, and number line.

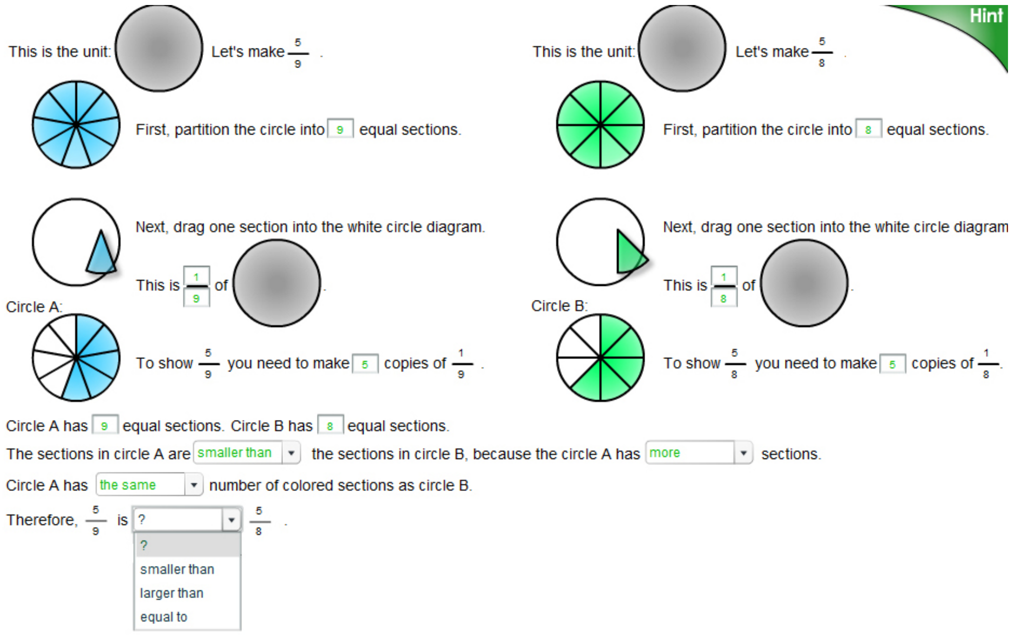

Figure 2. Making a circle given a symbolic fraction, combined with prompts to compare the two fractions. Reflection prompts are implemented with drop-down menus shown in the bottom half of each problem.

A total of 587 4th- and 5th-grade students from six different schools (31 classes) participated in the study during their regular mathematics instruction. We excluded students who missed at least one test day, and who completed less than 67% of all tutor problems (to ensure that students in the multiple graphical representations conditions encountered all three graphical representations). This results in a total of N = 290 (n = 63 in blocked, n = 53 in moderate, n = 52 in fully interleaved, n = 62 in increased, n = 21 in single-circle, n = 20 in single-rectangle, n = 19 in single-number-line).

Prior to working on the fractions tutor, students completed a pretest. The pretest took about 30 minutes. On the following day, all students started working with the fractions tutor. Students accessed the tutoring system from the computer lab at their schools and worked on the tutor for fractions for about five hours as part of their regular math instruction for five to six consecutive school days (depending on the length of the respective school's class periods). All students worked on the fractions tutor at their own pace, but the time students spent with the system was held constant across classrooms and across experimental conditions. On the day following the tutoring sessions, students took the immediate posttest which took about 30 minutes to complete. Seven days after the posttest, students completed an equivalent delayed posttest.

Figure 3 illustrates the practice schedules of task types and graphical representations for the four multiple graphical representations conditions. In all conditions, students worked through the same sequence of task types and fraction problems, and switched task types after every 9 of a total of 108 problems. Each task type was revisited three times. This procedure corresponds to the most successful level of interleaving task types in our prior experiment (Rau et al., 2010). We randomly assigned students to one of seven conditions. In the *blocked* condition, students switched graphical representations after 36 problems. In the *moderate* condition, students switched representations after every six problems. In the *fully interleaved* condition, students switched representations after each problem. In the *increased* condition, the length of the blocks was gradually reduced from twelve problems at the beginning to a single problem at the end. To account for possible effects of the order of graphical representations, we randomized the order in which students encountered the graphical representations. Finally, students in the three single graphical representation conditions worked on all tutor problems with only the circle, the rectangle, or the number line, respectively.

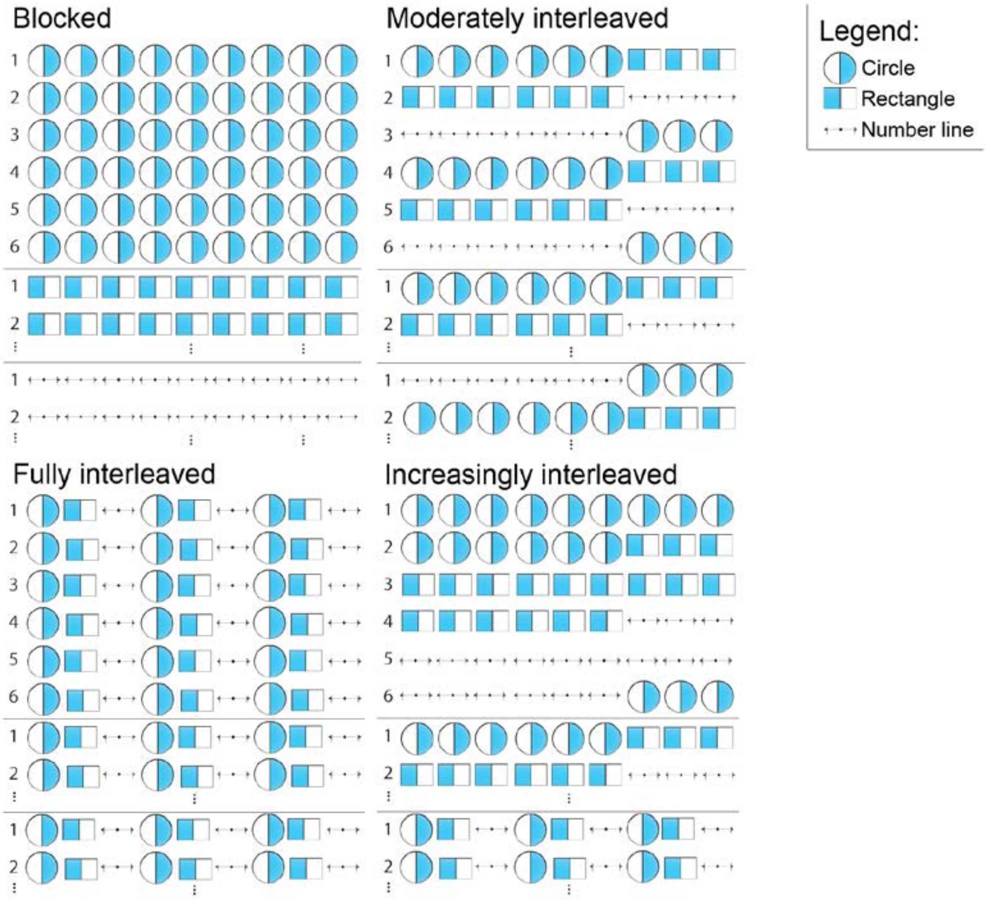

Figure 3: Practice schedule for multiple graphical representations conditions for all six task types. Each task type was revisited three times. Task types are indicated by numbers 1-6, representations are indicated by the different shapes.

We assessed students' knowledge of fractions at three test times. Three equivalent test forms were created, and we randomized the order in which they were administered. The tests included four knowledge types: fluency with area models (i.e., circles and rectangles), fluency with the number line, conceptual transfer and procedural transfer. The fluency items included identifying fractions given a graphical representation, making a graphical representation given a symbolic fraction, and recreating the unit given a graphical representation of both unit fractions and proper fractions. Conceptual transfer items included proportional reasoning questions with and without graphical representations. Procedural transfer items included comparison questions with and without graphical representations. The theoretical structure of the test (i.e., the four knowledge types just mentioned) resulted from a factor analysis performed on the pretest data. Test items including the number line seemed to be more challenging for students than area models.

Analysis

As mentioned, we analyzed the data of $N = 290$ students. There was no significant difference between conditions with respect to the number of students excluded ($\chi^2 < 1$). There were no significant differences between conditions at pretest for any dependent measure, *p*s > .10. There was no significant effect for order of multiple graphical representations for any dependent measure, $F(5, 285) = 1.56$, *p*s > .10.

We used a hierarchical linear model (HLM, see Raudenbush & Bryk, 2002) with four nested levels to analyze the data. We modeled performance on the tests for each student (level 1), differences between students nested within classes (level 2), differences between classes nested within schools (level 3), and differences between schools (level 4). More specifically, the following HLM model was fitted to the data:

$score_{ij}$ = $test_j$ + $condition_i$ + $test_j$*$condition_i$ + $preScore_i$ + $preScore_i$*$condition_i$ + $numProblems_i$ + $student(class)_i$+ $class(school)_i$ + $school_i$, **1**)

with the dependent variable $score_{ij}$ being $student_i$'s score on the dependent measures at $test_j$ (i.e., immediate or delayed posttest). In order to analyze whether students with different levels of prior knowledge benefit differently from our conditions, we included students' pretest scores as a covariate ($preScore_i$), and modeled the interaction of pretest score with condition ($preScore_i$*$condition_i$).

Since the HLM described in (1) uses students' pretest scores as a covariate, it does not allow us to analyze whether students in the various conditions improved from pretest to immediate and delayed posttest. To analyze learning gains, we included pretest score in the dependent variable, yielding:

$score_{ij}$ = $test_j$ + $condition_i$ + $test_j$*$condition_i$ + $numProblems_i$ + $student(class)_i$+ $class(school)_i$ + $school_i$, **2**)

with the dependent variable $score_{ij}$ being $student_i$'s score on the dependent measures at $test_j$ (i.e., pretest, immediate posttest, or delayed posttest).

We used planned contrasts and post-hoc comparisons to clarify results from the HLM analysis. All reported *p*-values were adjusted using the Bonferroni correction for multiple comparisons.

Table 1: Improvement of test scores at immediate posttest (post) over pretest (pre) and delayed posttest (delayed) over pretest by knowledge types and conditions.

Condition	**Effect**	**Fluency with area models**	**Fluency with the number line**	**Conceptual transfer**	**Procedural transfer**
blocked	post > pre delayed > pre	n.s. $p < .05$, $d = .52$	n.s. $p < .01$, $d = .39$	$p < .05$, $d = .42$ $p < .05$, $d = .39$	n.s. n.s.
moderately interleaved	post > pre delayed > pre	n.s. n.s.	n.s. $p < .01$, $d = .50$	$p < .05$, $d = .29$ $p < .05$, $d = .30$	n.s. $p < .05$, $d = .45$
fully interleaved	post > pre delayed > pre	$p < .05$, $d = .45$ $p < .05$, $d = .38$	$p < .01$, $d = .51$ $p < .01$, $d = .75$	$p < .01$, $d = .34$ $p < .01$, $d = .60$	n.s. n.s.
increasingly interleaved	post > pre delayed > pre	$p < .05$, $d = .38$ $p < .05$, $d = .55$	$p < .01$, $d = .43$ $p < .01$, $d = .46$	n.s. n.s.	n.s. n.s.
single-circle	post > pre delayed > pre	n.s. n.s.	n.s. n.s.	n.s. n.s.	n.s. n.s.
single-rectangle	post > pre delayed > pre	n.s. n.s.	n.s. n.s.	n.s. n.s.	n.s. n.s.
single-number-line	post > pre delayed > pre	n.s. n.s.	n.s. n.s.	n.s. n.s.	n.s. n.s.

Table 2: Differences between conditions at immediate posttest (post) and delayed posttest (delayed) by type of knowledge. "n.s." indicates non-significant results. "-" indicates that no post-hoc comparisons were computed.

Effect	**Test**	**Fluency with area models**	**Fluency with number line**	**Conceptual transfer**	**Procedural transfer**
fully interleaved > blocked, moderately interleaved, increasingly interleaved	post delayed	- -	n.s. n.s.	n.s. $p < .05$, $d = .33$	- -
increasingly interleaved > blocked, moderately interleaved, fully interleaved	post delayed	$p < .10$, $d = .30$ $p < .10$, $d = .30$	- -	- -	- -
moderately interleaved > blocked, fully interleaved, increasingly interleaved	post delayed	- -	- -	- -	n.s. n.s.

Results

To investigate whether students learned from the fractions tutor, we analyzed learning gains using the simpler HLM described in formula (2). The main effect of test time was significant for fluency with the number line, $F(2, 867) = 20.09$, $p < .01$, partial $\eta^2 = .03$, for fluency with area models, $F(2, 867) = 17.54$, $p < .01$, $\eta^2 = .02$, conceptual transfer, $F(2, 867) = 38.78$, $p < .01$, partial $\eta^2 = .03$, and marginally significant for procedural transfer, $F(2, 867) = 2.84$, $p < .10$, partial $\eta^2 = .01$. The interaction between test time and condition was significant for fluency with area models $F(12, 862) = 2.06$, $p < .05$, partial $\eta^2 = .01$. These results show that students (regardless of condition) improved on fluency with the number line, area models, procedural and conceptual transfer. On fluency with area models, students' learning gains also depended on the condition.

To further clarify these results, we computed post-hoc comparisons that compared students' scores at the immediate posttest and the delayed posttest compared to the pretest, respectively. Table 1 provides a summary of these post-hoc comparisons. Generally, we found significant learning gains at the delayed posttest for most of the multiple graphical representations conditions on fluency with area models, fluency with the number line, and conceptual transfer. On procedural transfer, only the moderate condition showed significant learning gains at the delayed posttest. Finally, we found no significant learning gains for the single graphical representation conditions except for the single-circle condition at the delayed posttest on conceptual transfer.

To analyze the effect of practice schedules of multiple graphical representations, we computed the HLM presented in formula (1) for only the multiple graphical representations conditions. There was no significant main effect of condition on any posttest scale, indicating that there was no global effect of practice schedules of multiple graphical representations across immediate and delayed posttests. An interaction between test time and condition was marginally significant for fluency with area models, $F(3, 867) = 2.57$, $p < .10$, $\eta^2 = .01$, indicating that the effect of practice schedules depends on test time. The interaction between pretest score and condition was marginally significant for conceptual transfer, $F(3, 219) = 2.52$, $p < .10$, $\eta^2 = .02$, demonstrating that students with different pretest scores benefit from different practice schedules.

To clarify the interaction between test time and condition, we used post-hoc comparisons separately for the immediate and the delayed posttest. To limit the number of comparisons, we only compared the most successful multiple graphical representations condition against the other three multiple graphical representations conditions, as summarized in Table 2. We found some support for a benefit of interleaving multiple graphical representations: the fully interleaved condition significantly outperformed the blocked, the moderately interleaved, and the increasingly interleaved conditions on conceptual transfer at the delayed posttest. Furthermore, we found a marginally significant advantage for the increasingly interleaved condition over the blocked, moderately interleaved, and fully interleaved conditions on fluency with the number line at the immediate and the delayed posttests.

To clarify the interaction between pretest score and condition on proportional reasoning items, we computed post-hoc comparisons for students with extremely low or high pretest scores. For students with a pretest score of 15%, 20%, and 25%, we found a significant advantage for the interleaved over the blocked condition (*p*s < .05). We found no differences for high prior knowledge students.

To analyze the difference between multiple graphical representations conditions and the single graphical representation conditions, we computed the HLM described in formula (1). We used planned contrasts to compare the multiple graphical representations conditions to the single graphical representation conditions. We found a significant advantage for the multiple graphical representations conditions over the single graphical representation conditions for number line test items at delayed posttest ($p < .05$, d =.29).

Think-aloud study

In order to gain further insight into the cognitive processes underlying the benefits of the interleaved practice schedules, we conducted a small-scale think-aloud study with six students who worked on the fully interleaved version of the tutoring system. As argued above, one hypothesized mechanism is that interleaved practice leads students to abstract across multiple graphical representations, for instance by comparing them to one another. Alternatively, it has been hypothesized that interleaved practice leads to the reactivation of representation-specific knowledge. Specifically, we were interested in what kinds of spontaneous comparisons students are making between the graphical representations, at what points in the curriculum they make comparisons, and whether students who fail to make spontaneous comparisons can be prompted to do so.

Methods

Six 5th-grade students participated in the think-aloud study. The think-aloud study was conducted in our laboratory and included three sessions. During the first session, students took the same pretest that was used in our experimental study. The pretest took about 30 minutes to complete. During the second session, students worked for one hour on a subset of problems taken from the interleaved version of the tutoring system while being prompted to think aloud, following the procedure described in Ericsson and Simon (1984). In the third session, students worked with similar tutor problems for one hour while being prompted to relate the different graphical representations to one another. We varied the type of prompts based on a within-subjects design: the prompt questions were either implicit (i.e., without directly prompting comparisons between the representations; e.g. "How is this problem the same as the last two you did?" or "How is this problem different from the last one you did?"), or explicit (i.e., directly referring to aspects that the different representations share; e.g., "What is the unit in the circle / rectangle / number line?" or "How are the rectangle and the circle and the number line the same / different?"). All students received two implicit prompts and four explicit prompts, in a fixed sequence.

Students' utterances were recorded and transcribed. We combined top-down and bottom-up approaches in developing a coding scheme: the experimenters identified types of comparisons that students might make prior to the think-aloud study, and then refined the coding scheme after viewing the transcripts from the think-aloud study. Comparisons between graphical representations were coded as surface comparisons if they either referred to the color of the representation, the shape of the representation, or the action performed on the representation (e.g., dragging and dropping). Comparisons were coded as conceptual if they referred to the corresponding features of the representations (i.e., numerator, denominator, unit), or the magnitude represented.

Results and Discussion

The results from the pretest indicate that all students had a good understanding of fractions. During the spontaneous comparison phase of the think-aloud study, we found only five instances of comparisons. These five comparisons were uttered by five of the six students. All five comparisons were surface comparisons.

When prompted to compare the different representations, students generated 138 instances of comparisons overall. Table 3 summarizes the average number of comparisons coded as surface and conceptual comparisons per implicit and explicit prompt. Given the small number of students, a statistical test on the types of comparisons in response to implicit and explicit prompts is not warranted. Table 3 shows that students generated substantially more surface than conceptual comparisons per prompt. We can also see that the implicit prompts yielded most of the surface comparisons, but almost none of the conceptual comparisons. In contrast, explicit prompts seem to have yielded more of the conceptual comparisons and fewer of the surface comparisons, compared to the implicit prompts.

Table 3: Average number of surface and conceptual comparisons per implicit and explicit prompts.

	Implicit prompts	**Explicit prompts**	
Surface	4.17	2.33	*2.94*
Conceptual	0.58	1.63	*1.28*
	2.38	*1.98*	

Conclusions

Taken together, the results from the experiment demonstrate significant learning gains for students who worked with a tutoring system that supports learning with multiple graphical representations of fractions, but not for those students who worked with only a single graphical representation. The gains persist until one week after the study when we administered the delayed posttest. Learning gains were found for students in all multiple graphical representations conditions on all posttest scales except procedural transfer. The fact that students' performance on procedural transfer does not improve may be due to the fact that comparing fractions was not the focus of the tutor, and that the comparison tasks in the procedural transfer items were too difficult for students. We did not find evidence for learning in the single representation groups. The lack of learning gains in the single representation control conditions demonstrates that the learning gains in the other conditions are not due to practice effects with the test format. Furthermore, the finding that the multiple graphical representations conditions outperform the single graphical representation conditions on number line items at the delayed posttest suggests an advantage of learning with multiple graphical representations over learning with a single graphical representation. Taken together with the lack of learning gains in the single representation conditions, these results replicate our earlier finding that multiple graphical representations lead to better learning of fractions than a single graphical representation (Rau et al., 2009), and extend this finding from working with a number line only to working with circles or rectangles only.

We argued that interleaving multiple graphical representations may help students acquire representational flexibility, whereas blocking graphical representations will help students acquire fluency with each representation. Our results confirm that the practice schedule for learning with multiple graphical representations matters: interleaving multiple graphical representations enhances the acquisition of representational flexibility as assessed by the conceptual transfer scale of the test – in particular for students with low prior knowledge. On the other hand, we did not find differences between experimental conditions for representational fluency (i.e., on area model and number line items) – with the exception of a marginally significant advantage of the increasingly interleaved condition on fluency with area models. Our results therefore suggest that students in all multiple representations conditions were able to acquire representational fluency from the tutoring system; and that – contrary to our hypothesis – blocking graphical representations does not contribute to the development of representational fluency. We hypothesized that interleaving graphical representations will promote the acquisition of representational flexibility. The significant advantage of the fully interleaved condition on conceptual transfer supports this hypothesis. Reasoning that representational flexibility may build on representational fluency, we had hypothesized that increasingly interleaving representations is the best option. Our findings do not support this hypothesis. Rather, the finding that especially low prior knowledge students benefit from fully interleaved representations may suggest that developing representational flexibility is particularly important for novice learners and that, perhaps, representational flexibility is a prerequisite for the development of representational fluency. Although more research is needed to investigate whether our findings generalize to other domains, based on our findings we can carefully conclude that designers of intelligent tutoring systems should employ an interleaved practice schedule of multiple graphical representations in order to enhance robust conceptual understanding of fractions.

What might be the mechanisms leading to the advantage of interleaving multiple graphical representations over blocking multiple graphical representations? Are students actively abstracting across graphical representations by explicitly comparing them to one another? Or are they, as argued above, reactivating knowledge that is specific to the graphical representations? Findings from our think-aloud study with the fully interleaved version of the tutoring system suggest that students did not spontaneously relate the different graphical representations to one another. The benefit from interleaving multiple graphical representations does not seem to stem from conscious abstraction across the different representations, but from more subliminal processes, such as repeatedly reactivating knowledge about the specific representations. On the other hand, the finding that students generate a good number of conceptual comparisons between the graphical representations when explicitly prompted to do so, suggests that students might benefit from explicit support in relating the different representations to one another.

Although these considerations are based on the findings from only a small-scale think-aloud study, we believe that they provide interesting insights into the processes that may underlie the benefits of interleaving multiple graphical representations. Future work should investigate whether indeed the advantage of interleaving representations results from repeated reactivation of representational knowledge. Furthermore, future work should explore the benefit of explicitly supporting students in conceptually relating multiple graphical representations and in making sense of their differences and similarities in how they depict fractions.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183-198.

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2009). Example-tracing tutors: A new paradigm for intelligent tutoring systems. *International Journal of Artificial Intelligence in Education, 19*(2), 105-154.

Battig, W. F. (1972). Intratask interference as a source of facilitation in transfer and retention. In R. F. Thompson & J. F. Vos (Eds.), *Topics in learning and performance* (pp. 131-159). New York: Academic Press.

Charalambous, C. Y., & Pitta-Pantazi, D. (2007). Drawing on a Theoretical Model to Study Students' Understandings of Fractions. Educational Studies in Mathematics, 64(3), 293-316.

de Croock, M. B. M., Van Merrienboer, J. J. G., & Paas, F. (1998). High versus low contextual interference in simulation-based training of troubleshooting skills: Effects on transfer performance and invested mental effort. *Computers in Human Behavior, 14*(2), 249-267.

de Jong, T., Ainsworth, S. E., Dobson, M., Van der Meij, J., Levonen, J., Reimann, P., et al. (1998). Acquiring knowledge in science and mathematics: The use of multiple representations in technology-based learning environments. In M. W. Van Someren, W. Reimers, H. P. A. Boshuizen & T. de Jong (Eds.), Learning with Multiple Representations. Oxford.

Ericsson, & Simon. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.

Koedinger, K., & Corbett, A. (2006). Cognitive Tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 61-78): Cambridge University Press.

Moyer, P., Bolyard, J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching children mathematics, 8*, 372-377.

National Mathematics Advisory Panel (2008). *Foundations for Success: Report of the National Mathematics Advisory Board Panel*: U.S. Government Printing Office.

Rau, M. A., Aleven, V., & Rummel, N. (2009). Intelligent tutoring systems with multiple representations and self-explanation prompts support learning of fractions. In Dimibrova, V., Mizoguchi, R., du Boulay, B. (Eds.). *14th International Conference on Artificial Intelligence in Education,* pp. 441-448. IOS Press, Amsterdam.

Rau, M. A., Aleven, V., & Rummel, N. (2010). Blocked versus Interleaved Practice with Multiple Representations in an Intelligent Tutoring System for Fractions In Aleven, V., Kay, J., Mostow, J. (Eds.) *10th International Conference of Intelligent Tutoring Systems,* pp. 412-422, Springer, Heidelberg.

Raudenbush, S.W., Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods.* Sage Publications, Newbury Park.

Acknowledgments

This work was supported by the National Science Foundation, REESE-21851-1-1121307. We thank Ken Koedinger, Mitchell Nathan, Kathy Cramer, Peg Smith, Jay Raspat, Michael Ringenberg, the Datashop and CTAT teams, Brian Junker, Howard Seltman, Cassandra Studer, the students, teachers, and school principals.