Abstract Title Page

Not included in page count.

Title: Towards Automated Support for Small-Group Instruction: Using Data from an ITS to Automatically Group Students

Authors and Affiliations:

Maria Mendiburo, Laura Williams, James Segedy, and Ted Hasselbring, *Institute for Software Integrated Systems (ISIS) and Department of Special Education, Vanderbilt University, Nashville, TN, USA*

Background / Context

The National Council of Teachers of Mathematics (2006) and the National Mathematics Advisory Panel (2008) recommend that teachers place particular emphasis on the teaching of fractions in upper elementary school. In this paper we explore a form of blended learning in which upper elementary school students learn fractions by rotating between computer-based instruction and teacher-led instruction during the same class period. The intervention discussed in this paper took place in three fifth-grade* mathematics classrooms at a public school in Middle Tennessee. It was one of a series of research studies completed as part of a three-year, education technology development project called HALF (*H*elping *A*t-Risk Students *L*earn *F*ractions). The goal of the HALF project, which began in June of 2010, is to design and pilot test a complete, blended learning intervention program that teaches basic fractions concepts.

Small group instruction is an essential component of the HALF instructional model, which is similar in many ways to the instructional model for the READ 180 intervention program (Mayer, Alexander, De Vivo, Aguhob, Davidson, 2013). READ 180 is among the most successful reading intervention programs on the market today and has been in use in America's schools for more than 15 years. HALF, like READ 180, uses a rotational model that has been found to be extremely successful in elementary, middle, and high school classes (Slavin, Cheung, Groff, & Lake, 2008; Lang, Torgesen, Petscher, Vogel, Chanter, Lefsky, 2009). Students begin each day of the HALF intervention program with a brief whole class discussion. After the discussion, students begin one of three rotations: computer-based adaptive instruction using an intelligent tutoring system (ITS) for fractions learning, game-based fluency practice, or small group instruction led by a teacher. Small group instruction allows the teacher to deliver either remediation or enrichment lessons; remediation lessons help students overcome difficulties or misconceptions encountered during their interactions with the ITS, and enrichment lessons provide students with opportunities to extend their understanding of fractions content through teacher-facilitated, higher-level discussions with their peers that go beyond the computer-based lessons they've already mastered.

Purpose / Objective / Research Question / Focus of Study

In this paper, we explore the use of learning analytics as a method for easing the cognitive demands on teachers implementing the HALF instructional model. Learning analytics has been defined as "the measurement, collection, analysis and reporting of data about learners and their contexts for the purposes of understanding and optimizing learning and the environments in which it occurs" (Vatrapu, Teplovs, & Fujita, 2011). When used in a classroom setting in conjunction with teachers, learning analytics systems leverage both the expertise of the teacher and the capabilities of the technology to synergistically optimize the learning resources available for students (Segedy, Sulcer, & Biswas, 2010). By automatically analyzing data generated as students' work through the fractions ITS, the HALF software should be able to provide the teacher with immediately actionable suggestions for effectively grouping (i) students who need remediation about related concepts, and (ii) students who have already mastered related concepts and are ready for an additional challenge. As an initial step in developing this technology, we report the results of a study that explored how to automate the instructional decisions made by a researcher with content expertise and in-depth knowledge of the HALF intervention program using data generated by the fractions ITS.

* Fifth-grade students in the US are typically 10-11 years old.

Setting and Population

The sample of students was comprised of three intact mathematics classes taught by a single teacher at a traditional public middle school in the state of Tennessee. All 75 participating students were in fifth-grade at the time of the intervention. Of the students, 49% were male and 70% were Caucasian. Additionally, 30% were from economically disadvantaged backgrounds and 14% had identified learning disabilities. One student in the sample was classified as an English language learner. Separate from this intervention, the students took the state standardized test, the Tennessee Comprehensive Assessment Program (TCAP). On the mathematics portion of the TCAP, 8% of students were ranked as Below Basic, 28% as Basic, 33% as Proficient, and 31% as Advanced.

Intervention / Program / Practice

Two weeks prior to the start of the intervention, the students in all three classes completed a diagnostic fractions assessment. The assessment included 30 multiple-choice questions drawn from the test bank of 243 items that were developed for the HALF intervention program. Based on the students' results, we developed two sequences of lessons, one for students at an early stage of fractions development (Sequence 1) and a more advanced stage (Sequence 2). Sequence 1 included more lessons about the most basic fractions concepts than Sequence 2. Previous research from the HALF project helped us determine that some concepts in fractions are harder for students to learn than others in predictable ways. We used this information to help organize the lessons within each sequence into an appropriate difficulty progression for that group of students.

The intelligent tutoring system used during the intervention included two main components: (i) instructional videos and (ii) guided practice with an embedded scaffolding framework. When students first began learning each lesson using the ITS, they viewed a short (1-2 minute) instructional video that included an explanation of the new fractions content and a demonstration of an example problem. Students were given the option to view additional examples, if desired, and then were moved to the guided practice mode. This guided practice consisted of multiple stages within each lesson, and each stage included a different set of embedded scaffolds, with earlier stages incorporating more scaffolds than later stages. Students needed to complete the most difficult stage (i.e. with the least scaffolds) in order to complete a lesson.

Research Design, Data Collection, and Analysis

The goal of this study was to create an algorithm that could assign students to small groups on a day-to-day basis relative to the student's daily progress through the intelligent tutoring system. During the week-long intervention, the teacher-researcher manually assigned the students in each class to one of three small groups and wrote lesson plans for each group. This was a very time-consuming process that took two to three hours per day to complete. It also relied heavily on the teacher-researcher's experience and personal knowledge of pedagogy and content. Since we intend for teachers that are not part of our research team to be able to implement the HALF intervention program in their classes, both of these issues speak to the need for an algorithm that uses decision rules that replicate the teacher-researcher's expertise.

The criteria used to make these grouping decisions fell into two categories, *primary* and *secondary* criteria. Primary criteria for each student included: (i) overall performance using the

ITS (total percentage of questions answered correctly across all stages and all lessons attempted that day), performance on the highest stage of the last lesson attempted, and the total number of fractions problems answered while using the ITS. Point values were assigned to each criteria to form each student's primary remediation score (PRS). We determined that any student with a PRS greater than 1 should be automatically assigned to a group receiving a remediation lesson during the next day's class. Any student with a PRS score equal to 0 was assigned to an enrichment group. Students with a PRS equal to 1 were assigned to a teacher check-in and were subsequently assigned either enrichment or remediation by the teacher, depending on the result of that check-in. Students absent on the prior day were automatically assigned to whichever group they would have been in previously, as well as receiving a check-in from the teacher. A list of secondary criteria was used to refine the assignment decision-making process and to prioritize students' need for remediation. These criteria were assigned point values and used to modify the PRS to create a final remediation score (FRS). Adding these secondary criteria to the algorithm allowed us to reduce the probability of unnecessarily assigning a student to remediation who was otherwise succeeding in the learning environment and increase the probability that all students in the class would experience remediation and enrichment lessons. The full list of primary and secondary criteria and the point values for creating the PRS and FRS are given in Figure 1

The group assignment algorithm also took into account the minimum and maximum group sizes for successful remediation and enrichment lessons. The HALF instructional model includes three distinct components that we determined students should experience during each day of the intervention, so we typically prescribe a 3-group rotational model for typical-sized classes. The algorithm was also designed to assign a variable number of remediation or enrichment groups depending on the needs of the students. Based on the intervention, we determined that remediation groups should be no larger than 1/3 of the total class size and no smaller than 1/6 of the total class size. Enrichment groups could function at larger than 1/3 of the class size if necessary, but classes with more than one enrichment group should have an approximately equal number of students in each group. The algorithm was developed in such a way that if the maximum size of one remediation group was reached and there were not enough students to reach the minimum size necessary to create a second remediation group, it assigned students with the highest FRS to remediation. Students with lower FRS scores were assigned to an enrichment group, but they were also assigned to receive a teacher check-in. If there was more than one remediation group in a class, the composition of students in each group was determined by the content of the students' current lessons.

Results

Once we completed the process of creating the algorithm, we used the ITS data generated by the system on the third day of the intervention to test how well the algorithm replicated the teacher-researcher's decision rules for assigning students to groups for the fourth day of the intervention. We evaluated the results in three ways (Table 1). First, the teacher-researcher reanalysed the data from the third day of the intervention and wrote a new set of group assignments and lesson prescriptions. The changes between these two groups show how the teacher-researcher's thinking about what factors were actually important when deciding which students would benefit most from remediation evolved during this process. As the group assignments were an evolving process, we expected to see a number of discrepancies. The comparisons in Table 1 show that the two main discrepancies between these categories were that not enough students were originally referred for remediation and that more students would have benefitted from a teacher check-in. The assignment of additional students to remediation was due to the discovery that there are more criteria than originally thought that could indicate a student is struggling. The additional teacher check-ins arose from a number of factors, such as the recognition that several students did not fall clearly into a remediation or enrichment category and might benefit from a brief conversation with the teacher to see if a problem could be quickly identified and ameliorated.

The second result of interest is a comparison of the small groups generated by the automated algorithm to the revised groups created by the content expert. This comparison shows that the overall performance of the algorithm was excellent as compared to the expert-generated groups. There were no clear patterns of over- or under-assignment by the algorithm to a particular type of group or to a teacher check-in. Evaluation of the automatically-generated groups by the content expert showed that the groups' compositions were sufficiently accurate to consider the rule set a successful initial attempt at modeling the thought process and logic behind small group assignment. This is related to the third result from this study, which was how students mis-assigned by the algorithm could be more accurately assigned in the future. The overall result of the algorithm's performance was that fewer than 10% of the students in the sample were potentially assigned to a less-appropriate small group, as determined by the content expert's evaluation of the algorithm-generated small groups. In these cases, it is not necessarily clear what the best assignment for that student on that day would be. As we continue the algorithm development process, it is our goal to use additional field data to refine the assignment rules as to eliminate any potential for incorrect assignment.

Conclusions

In this paper, we presented a design-based research approach for applying learning analytics to the process of supporting teachers in implementing the HALF instructional model in a classroom environment. Our results show that by explicating and automating a pedagogical and content expert's decision-making processes in relation to grouping students, systems can be designed to support future implementations of HALF without having to carefully analyze each student's learning trajectory in the ITS. This has significant implications in terms of the potential for using ITSs to interpret and synthesize the large amounts of data they collect as students learn with them. In particular, these systems can, with a fair amount of accuracy, provide teachers with *immediately actionable suggestions* for how to design small group instruction. These suggestions can ultimately allow teachers to more efficiently allocate their class time by taking advantage of the data provided by the ITS without dedicating hours of additional time to interpreting each individual student's progress.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Lang, L. H., Torgesen, J. K., Petscher, Y., Vogel, W., Chanter, C., Lefsky, E. (2009). Exploring the relative effectiveness of reading interventions for high school students. Journal of Research on Educational Effectiveness, 2(2), 149-175. doi:10.1080/19345740802641535

Mayer, M., Alexander, F., De Vivo, K., Aguhob, M., and Davidson, J. (2013). *READ 180: A heritage of research.* New York, NY: Scholastic Inc. Retrieved from http://read180.scholastic.com/reading-intervention-program/research/view/350

National Council of Teachers of Mathematics (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence.* Reston, VA: National Council of Teachers of Mathematics.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: US Department of Education. Retrieved from www2.ed.gov/about/bdscomm/list/mathpanel/ report/final-report.pdf

Segedy, J.R., Sulcer, B., & Biswas, G. (2010). Are ILEs ready for the classroom? Bringing teachers into the feedback loop. In V. Aleven, J. Kay, J. Mostow (Eds.), *Lecture Notes in Computer Science, 6095*, 405-407.

Vatrapu, R., Teplovs, C., Fujita, N., & Bull, S. (2011). Towards visual analytics for teachers' dynamic diagnostic pedagogical decision-making. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 93-98). New York, NY: ACM.

Appendix B. Tables and Figures

Not included in page count.

	Points	Criteria	Rationale
Primary Remediation Score	+1	Score on highest lesson stage less than 75% when at least 3 questions answered	Gives weight to students who were not completely successful with the most difficult content encountered that day. Question number modifier ensures that an adequate number of questions are being used to make the decision.
	+1	Score on all questions less than 70%	Gives weight to students who struggled with some concepts. This is separate from the previous criteria because it was possible for a student to have difficulty with the hardest stage but not overall, or vice versa.
	+1	Score less than 60% on a stage that included additional instructional supports	The two easiest stages (of a possible 5) in each lesson included additional support for learning. A student who struggled even with this additional instruction received extra weight.
	+1	Fewer than 6 questions attempted during the session	Depending on the lesson and other instructional content seen by the student that day, it was possible for a student to see a relatively small number of questions. Fewer than 6 indicates that the student was taking too long per question, was not using time appropriately, or was having trouble with the workspace.
	+1	Student previously referred for remediation but was never assigned to a group	Due to group size and lesson constraints, some students assigned to remediation could not participate in a remediation group that day. Extra weight was given these students.
	+1	Student still on first lesson of the content progression	Gives weight to a student who has not progressed along with classmates or who has been absent and might need help catching up.
	Remove all points	Student a full lesson ahead of majority of classmates	Removes priority from students who have been succeeding on their own. If they are still struggling the next day, they can be remediated at that time.
Final Remediation Score	Variable; number of days since last remediation minus 1	Number of days since student participated in a remediation group	Gives priority to students who have not recently participated in remediation. This is secondary as to not penalize students who have not needed a remediation group and continue to do well.
	Variable; number of days working on the same lesson minus 1	Number of days working on the most recent lesson	Gives weight to student s who have been working on the same lesson for more than one day. This is secondary as to not penalize students who are working on a longer lesson that might take more than one day to complete.
	+1	Any given stage attempted more than twice	Gives weight to students who are repeatedly moving between stages with more and less instructional supports but not moving forward.
	-1	More than one lesson attempted that day	Reduces priority on students who have already completed at least one lesson that day.

Table 1
Results

	Original	Hand-revised	Algorithm
Class 1 - 20 students			
Number of remediation groups	1	2	2
Number of enrichment groups	2	1	1
Initial group assignments			
# assigned remediation	7	10	11
# assigned enrichment	6	3	5
# assigned teacher check-in	7	7	4
Final group assignments			
# assigned remediation	7	11	13
# assigned enrichment	13	9	7
Class 2 - 28 students			
Number of remediation groups	0	0	0
Number of enrichment groups	3	3	3
Initial group assignments			
# assigned remediation	0	0	0
# assigned enrichment	27	25	23
# assigned teacher check-in	1	3	5
Final group assignments			
# assigned remediation	0	0	0
# assigned enrichment	28	28	28
Class 3 - 23 students			
Number of remediation groups	1	1	1
Number of enrichment groups	2	2	2
Initial group assignments			
# assigned remediation	6	6	5
# assigned enrichment	17	9	9
# assigned teacher check-in	0	8	9
Final group assignments			
# assigned remediation	6	9	8
# assigned enrichment	17	14	15