WhaleWatch: An Intelligent Multimedia Math Tutor

David M. Hart, Ivon Arroyo, Joseph E. Beck and Beverly P. Woolf
Center for Computer-Based Instructional Technology
Department of Computer Science
University of Massachusetts, Amherst
United States
{dhart,ivon,beck,bev}@cs.umass.edu

Carole R. Beal
Department of Psychology
University of Massachusetts, Amherst
United States
cbeal@psych.umass.edu

Abstract: Mathematics training is essential for participation in science and engineering careers, yet many students, especially girls, dislike and avoid math, and are therefore unprepared for university science majors and graduate programs. The goal of this project is to increase students' interest in math and their confidence in their ability to learn math through an intelligent, model-based multimedia tutor, WhaleWatch. Based on a dynamically updated student model, WhaleWatch selects problems of appropriate difficulty and provides help and instruction as needed. The results of several evaluation studies indicate that WhaleWatch had a positive impact on students' math self concept and beliefs in the value of learning mathematics. These results also suggest beneficial tutor modifications that will be incorporated WhaleWatch's successor, AnimalWatch.

Increasing Girls' Self-Confidence in Math

It is well documented that women are under-represented in math and science career fields, a problem that is getting worse instead of better in areas such as Computer Science. This problem can be attributed in part to math avoidance among girls, who frequently drop out of mathematics courses when they become optional at the secondary level, and as a result perform less well on math achievement tests (Beal, 1994; Beller & Gafni, 1996). Because these same math courses are prerequisites for most math and science careers, many girls find themselves shut out of these fields early on. Even the most mathematically gifted girls have less math training and are less interested in science careers than their male peers (Benbow, 1992). One phenomena often cited as a turning point for girls in their math education is a drop in their self-confidence that occurs between 5th grade (11 years old) and 8th grade (14). Before fifth grade girls participate fully and perform as well as boys; after this period they lose interest.

A collaborative of researchers from Psychology, Computer Science and Education have joined in a project to create an intelligent tutoring system called WhaleWatch whose goal is to raise girls' self-confidence in their math skills during this critical period in their development. Why an intelligent computer tutor? Feedback from teachers, whether conscious or unconscious, has been shown to reinforce boys' self-confidence while lowering girls' (Beal, 1994; Boggiano & Barrett, 1991). The WhaleWatch tutor is designed to be encouraging of both girls and boys in its response.

An inherent danger exists, however, in addressing this problem with computers – computer programs have traditionally been designed by men, and typically use competitive and excitement-oriented paradigms to engage the user. Studies have shown that girls are not as attracted to these paradigms, preferring instead activities whose content is less game-like and which feature more characterization and role playing. They also like to work together, interacting cooperatively rather than competitively. WhaleWatch has been designed with these features in mind, using the study of endangered species such as Right Whales as a context for problem-solving

[1], and providing the student the opportunity to select a mentoring character as a guide through the activity. While it does not yet provide a good method for students to collaborate while working with the tutor, we are exploring ways to incorporate this.

WhaleWatch is being developed with the collaboration of two local school systems, the Frontier Regional Schools in suburban/rural Western Massachusetts, and the Springfield Public Schools, and urban school setting also in Western Massachusetts. These two systems are an interesting contrast in demographics and in the kinds of computer resources available for students, which provides a wide range of conditions for testing WhaleWatch. We have tested a WhaleWatch prototype in four different trials in these systems between 1996 and the present. These formative tests have yielded encouraging data on the potential for boosting girls' self-confidence in math, while at the same time giving us critical feedback on the effectiveness of the tutor for both girls and boys. This feedback will be incorporated into the design of an expanded tutor under development now called AnimalWatch, which will build on WhaleWatch by adding more species and a broader range of math topics.

WhaleWatch: The Tutor

WhaleWatch uses intelligent tutoring techniques to help students learn fractions, decimals and percentages at a 5th-6th grade level. In contrast to common drill-and-practice systems, intelligent tutoring systems modify themselves to conform to the students' learning styles (Anderson, et al., 1995). WhaleWatch begins students with whole-number arithmetic problems in order to build a profile of what the student already knows and how fast the student can advance. Once the student demonstrates mastery of whole numbers, the tutor presents simple fractional problems that require increasingly challenging application of the cognitive subtasks involved in solving the problems (e.g. adding fractions with like denominators, adding fractions with different denominators, etc.).

Interaction between the student and WhaleWatch is built around word problems that are contextualized using the endangered species theme. These include questions about Right Whales' feeding habits, migration patterns and population, as we see in (Fig. 1). Graphics of whales are used to help set the scene for students. AnimalWatch will expand on this contextualization, adding more narrative, using various other endangered species, and increasing the connection between the problems and the students' ability to manipulate the objects onscreen in order to solve the problems presented.


_____
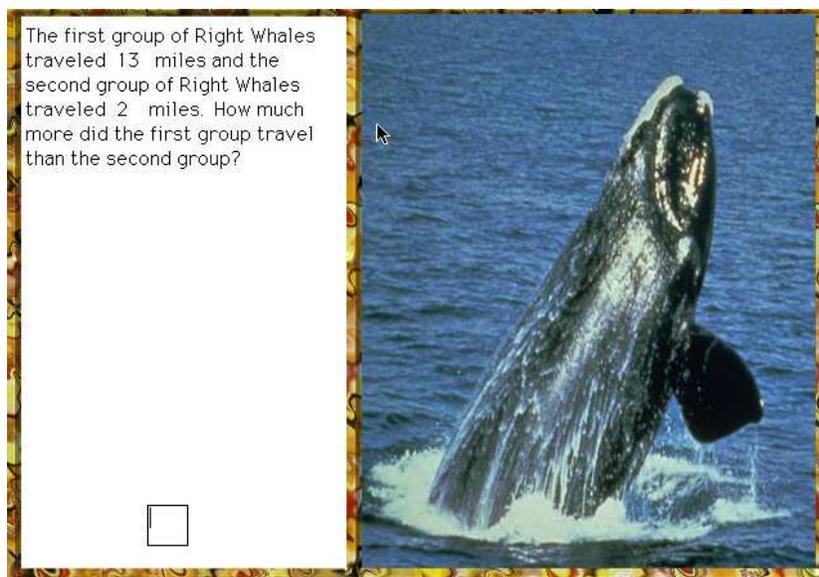
[1] Of all the sciences, women are best represented in biology and especially environmental science.

Figure 1: WhaleWatch Tutoring Interface. The students are asked to solve word problems about Right Whales, an endangered species.

When a student has trouble solving a problem, WhaleWatch initiates a tutoring interaction such as the one shown in (Fig. 2) that provides tailored hints and guidance that helps the student work through the problem. Similar problems involving the same subskills are given until the student can successfully work the problems. The example shown in (Fig. 2) is a tutoring aid to help a student understand Equivalent Fractions. To solve the fractions problem shown here, $1/4 + 1/2$, the fractions must first be converted to equivalent forms (i.e. fractions of the same denominator). The window shown overlaying the picture of the whale is an interactive calculator for finding equivalent fractions. As the student clicks on the Up and Down buttons, the fraction is multiplied or divided. At the same time, the Cuisenaire Bars shown under the fractions are divided into parts that coincide with the fraction above. For example, the bar under the 1/4 fraction in the top right is divided into four parts, one of which is colored to show the fraction four given by the numerator one. By alternately multiplying and dividing the upper and lower fractions, the student can use this calculator to find their equivalent forms. Once they are found, 1/4 and 2/4, the student can add them and enter the answer in the boxes at the lower left in the main problem window. A correct answer elicits positive auditory feedback, while an incorrect answer elicits encouragement to try again. The next problem the student is given will be easier or harder, based on the student's ease in solving this one.
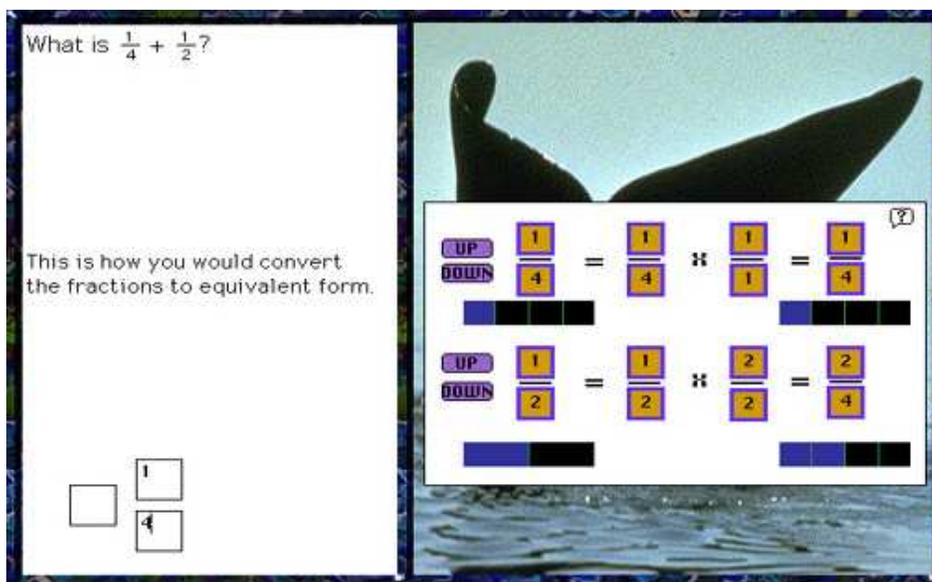
Figure 2: WhaleWatch Hinting. When the student makes a mistake solving a word problem, WhaleWatch offers hints in the right window pane that are designed to help the student understand her mistake and master the subskill in question.

For WhaleWatch to be successful in its tutoring, it must maintain an accurate assessment of the student's strengths and weaknesses in this task domain. Online self-assessment surveys conducted as students work with WhaleWatch have shown that the tutor generates a more accurate assessment of each student's abilities than the students themselves (Beck, et al., 1997a). WhaleWatch uses Artificial Intelligence techniques for problem generation, hint selection and student modeling (Beck, et al., 1997b; Beal, et al., 1998). Multimedia is used judiciously to engage the student by animating key concepts and providing interactive manipulables based on those used by classroom teachers.

WhaleWatch Testing and Results

Classroom trials of the WhaleWatch prototype for purposes of formative evaluation have been conducted on four occasions, three times in Deerfield, and once in Springfield. Typically these trials involve two or three classes for as many as 60 students who use the tutor for up to five hours (in 60- or 90-minute sittings over several days). Standardized tests of self-confidence (Eccles, et al., 1993) are given by paper and pencil as pretests before students start the trials, and then afterward the same tests are given again. In future trials we will introduce standardized testing for math proficiency as well. We briefly summarize results of these trials here.

Evaluation Study 1

The first evaluation study was designed to determine if WhaleWatch could in fact have a positive effect on girls' self-confidence. Results of this study, conducted with three classes in the spring of 1997, showed that WhaleWatch did increase girls' self-confidence, raising it to a level on par with the boys'. The boys' confidence level, which started out higher, remained the same (Beal, et al., 1998; Beck, et al., 1997a; Beck, et al., 1999).

Evaluation Study 2

While the first evaluation study gave promising results, it did not clearly show that the change in girls' confidence was due to the help and guidance of the tutor itself. To test this hypothesis, we designed an experiment in which half of the students from each class using the "tutor" would get all the help and guidance the system had provided in Evaluation Study 1 (the treatment group), and half would get none of this help (the control). This second group would see word problems about whales and could enter solutions, but would get no feedback other than the answer was correct or incorrect. In effect, these students were doing drill-and-practice while the treatment group would receive the full benefits of intelligent tutoring.

In conjunction with this evaluation study we administered for the first time an online Piagetian pretest designed to identify each student's level of mathematical cognitive development (Arroyo, et al., 1999a; Arroyo, et al., 1999b). Results of this test allowed us to correlate student performance and changes in self-confidence with cognitive ability.

Results of this evaluation study, detailed in (Beck, et al., 1999) held several surprises. Reassuringly, girls benefited from the help provided by the tutor, though this was not seen as clearly as in Evaluation Study 1 [2]. Boys, on the other hand, had their confidence levels increased in the control condition (no tutor help), while their confidence level *decreased* in the treatment condition (tutor help). We hypothesize that boys felt constrained by the amount of time taken by hinting (hints could be elaborate), and subconsciously preferred not to be slowed down. This would argue for changing the hinting mechanism to be less elaborate and allow user termination of each hint during delivery. AnimalWatch is being designed to support this method of hinting.

Another surprising finding is that the tutor was most helpful to a group of students in the mid range of cognitive abilities, as determined by the Piagetian pretest (Arroyo, et al., 1999b). These students' performance using the tutor was increased significantly, while students at the upper and lower cognitive ranges were not significantly helped. This suggests that we broaden the scope of the hinting capabilities as we create AnimalWatch, producing hints that are more concrete and that use interactive manipulables for the low cognitive ability students, while also providing more abstract and symbolic hinting for students at the higher end of the cognitive scale (Arroyo, et al., 1999b). An example of the former, a more concrete hint, is shown in (Fig. 3).

This AnimalWatch hint employs Cuisenaire Rods, a commonly used classroom manipulative. In the simple addition-of-wholes example shown, the student is assisted through the steps in adding 25 and 79. Unlike

---

[2] It appears that the assignment of girls to control and treatment groups, while random, produced biased samples. Specifically, the girls in the treatment group had confidence levels and pretest scores that were significantly higher than those in the control group. In fact, their confidence levels before using the tutor were higher than the confidence levels of girls from the previous year *after* they had used the tutor. This seems to have produced a ceiling effect on confidence levels in the girls' treatment group, preventing us from seeing an increase corresponding to that seen in Evaluation Study 1.

WhaleWatch, which used a static set of values for its hints, AnimalWatch hints will be passed the values of the actual problem the student is trying to solve. Cuisenaire Rods representing each value can be manipulated by the student. Unit boxes are dragged into the lower right box, where the sum box (shown here as zero) accumulates a value for the number of Units dragged in. Similarly, rods representing Tens are dragged into lower center box, where their sum is incremented with each bar that is added. Next the student is encouraged to move groups of ten unit boxes from the Units box to the Tens box, and then to move groups of Tens to the Hundreds box (lower left). This form of concrete, manipulative-driven hint is designed to be particularly useful to students at the lower end of the cognitive ability scale. AnimalWatch will include hint-selection capabilities that will factor in differences in student ability (among other attributes) and select hints that are more targeted than those in WhaleWatch.
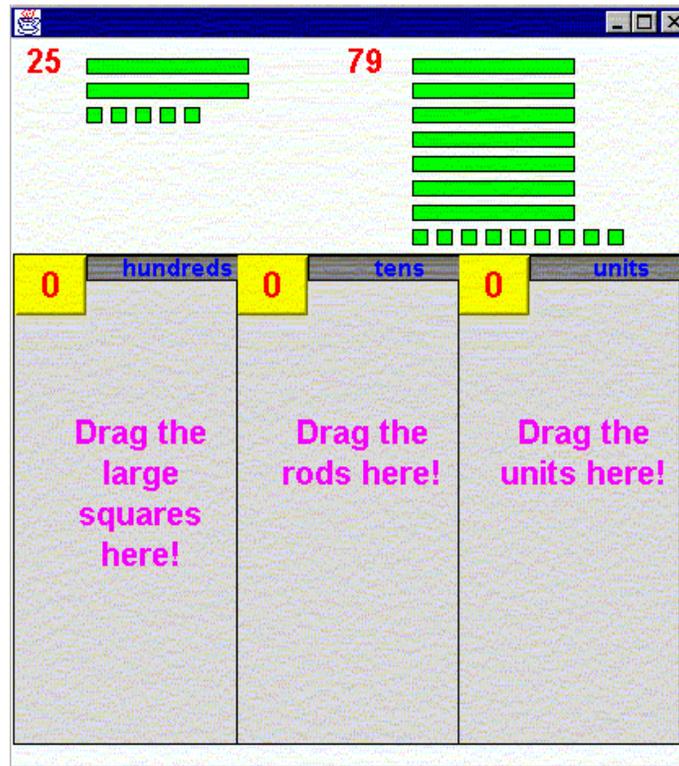
Figure 3: AnimalWatch hinting will be more interactive, allowing students to manipulate objects onscreen as they explore the concepts underlying operations on fractions, decimals and percentages.

From WhaleWatch to AnimalWatch

Over the next three years WhaleWatch will be expanded to include many of the new features described above, and will also have more content added (increasing the average student contact time to 20 hours) and will see the introduction of other endangered species for added context, necessitating the name change to Animal-Watch. Larger scale trials with more contact time will be conducted in Deerfield and Springfield during this time.

References

Anderson, J.R., Corbett, A.T., Koedinger, K.R., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, 4 (2), 167-207.

Arroyo, I., Beck, J.E., Schultz, K., & Woolf, B.P. (1999a). A Useful Prediction Variable for Student Models: Cognitive Development Level. Submitted to the *International Conference on User Modeling*, 1999.

Arroyo, I., Beck, J.E., Schultz, K., & Woolf, B.P. (1999b). Piagetian Psychology in Intelligent Tutoring Systems. Submitted to *Artificial Intelligence in Education* , 1999.

Beal, C.R., Beck, J.E., Woolf, B.W., & Rea-Ramirez, M. (1998). WhaleWatch: An Intelligent Model-Based Mathematics Tutoring System. In *Proceedings of the International Federation for Information Processing World Computer Congress*, Information Technology & Knowledge Systems, 1998. Vienna, Austria and Budapest, Hungary.

Beal, C.R. (1994). *Boys and Girls: The Development of Gender Roles*. New York, NY: McGraw Hill.

Beck, J.E, Stern, M., & Woolf, B.P. (1997a). Cooperative Student Models. *In Proceedings of the Eighth World Conference on Artificial Intelligence in Education*, 1997. 127-134.

Beck, J.E, Stern, M., & Woolf, B.P. (1997b). Using the Student Model to Control Problem Difficulty. *In Proceedings of the Seventh International Conference on User Modeling*, 1997. 277-288.

Beck, J.E., Arroyo, I., Woolf, B.P., & Beal, C.R. (1999). Affecting Self-Confidence with an ITS. Submitted to *Artificial Intelligence in Education*, 1999.

Beller, M,, & Gafni, N. (1996). The 1991 International Assessment of Educational Progress in Mathematics and Sciences: The Gender Differences Perspective. *Journal of Educational Psychology*, 88, 365-377.

Benbow, C.P. (1992). Academic Achievement in Mathematics and Science Between Ages 13 and 23: Are There Differences Among Students in the Top One Percent of Mathematical Ability? *Journal of Educational Psychology*, 84, 51-61.

Boggiano, A.K., & Barrett, M. (1991). Strategies to Motivate Helpless and Mastery-Oriented Children: The Effect of Gender-Based Expectancies. *Sex Roles*, 25, 487-510.

Eccles, J.S., Wigfield, A., Harold, R.D., & Blumenfield, P. (1993). Age and Gender Differences in Children's Self and Task Perceptions During Elementary School. *Child Development*, 64, 830-847.

Schofield, J.W., Eurich-Fulcer, R., & Britt, C.L. (1994). Teachers, Computer Tutors, and Teaching: The Artificially Intelligent Tutor as an Agent for Classroom Change. *American Educational Research Journal*, 31, 579-607.

Acknowledgments