New results from Using a Web-based Interactive Landform Simulation Model (WILSIM) in a General Education Physical Geography Course

Wei Luo Department of Geography, Northern Illinois University, DeKalb, IL 60115
Michael Konen Department of Geography, Northern Illinois University, DeKalb, IL 60115

ABSTRACT

Luo et al. (2005) reported that careful design of the questions is very important to accurate assessment of the effectiveness of dynamic simulation and visualization on student learning. We followed the lessons learned in that study and improved some of the poorly designed pre- and post-test questions used for assessing the effectiveness of a Web-based Interactive Landform Simulation Model (WILSIM). We tested the new questions in a large section of the same general education physical geography course in the springs of 2006 and 2007 before and after students used WILSIM to simulate 3 different scenarios of landform evolution. We found that the post-test scores were significantly higher than pre-test scores, whereas previous study showed no statistically significant differences.

INTRODUCTION

Visualization and simulation have been increasingly used in education to enhance learning (Gordin and Pea, 1995) because they supposedly exploit the power of the human visual system for discerning patterns and observing processes that are otherwise difficult to conceptualize. They also offer an opportunity to engage students in authentic inquiry as part of their learning process and thus enhance their understanding (Edelson and Gordin, 1998; Edelson et al., 1999). This is particularly relevant to the study of landform evolution. Present day landforms are the result of a variety of complex interacting physical processes operating over geological time scales. Because of the short human life span, long-term landform evolution cannot be observed directly. Furthermore, the interaction of differing physical processes involved are difficult to infer from the limited temporal observations of present day forms. Computer based visualization and simulation could potentially help students learn the concepts and processes involved in landscape evolution.

However, studies on the effectiveness of visualization and animation in improving learning have produced mixed results (e.g., Jensen et al., 2002; Tversky et al., 2002; Anglin et al., 2004). Luo et al. (2005) conducted a study on use of a Web-based Interactive Landform Simulation Model (WILSIM) (Luo et al., 2004) in a general education physical geography course to determine whether the interactive simulation exercise improved students' scores on tests (consisting of multiple choice questions) administered before and after the exercise, which consists of simulating landform evolution under 3 different scenarios: spatially uniform erodibility, spatially different erodibility and spatially different tectonic uplift rate (Luo et al., 2005). They found that although the mean of post-test scores was slightly higher than that of the pre-test, the difference was not statistically significant. They also found that students tended to do better in those questions that are more directly related to the simulation scenarios and poorer in those questions that are more related to terminology. They suggested that the insignificant difference between post- and pre-test scores could be partially attributed to the poor design of some questions.

Multiple choice questions (MCQs) consist of two components: (1) the stem, which can be a question or an incomplete statement, and (2) several alternative options that answer the question or complete the incomplete statement, of which one is the correct answer and the rest are distracters. MCQ has become a popular format for exams, especially for assessing students in large enrollment classes, because MCQs can be easily graded, statistically analyzed, and the result can be both accurate and objective (Race et al., 2005, Brown, 1997). However, much care needs to be taken regarding the design and validation of the questions in order to take all the advantages of MCQs and avoid their disadvantages, such as hard to test high-level skills and guessing by students (Race et al., 2005; Fuhrman, 1996). Well constructed MCQs are characterized by clarity and brevity in the stem and un-ambiguity in choices. Previous studies (e.g., Gronlund,1988; Fuhrman, 1996; Brown, 1997; Race et al., 2005) have suggested many tips for developing good MCQs, for example, (1) The stem should be clear, simple and as concrete as possible; (2) Avoid using negative form in the stem; (3) The distracters should be plausible to the uninformed and all choices should have similar sentence structure and similar length; (4) Avoid "All of the above" and use "None of the above" with care in choices.

Here we report on new results using WILSIM again in large sections of the same introductory physical geography course in springs of 2006 and 2007 after addressing lessons learned in the previous study, including improving the design of MCQs.

NEW RESULTS

We tested WILSIM in the large section of the same general education course (Survey of Physical Geography) course in the springs of 2006 and 2007. The majority of the students in the course are non-geoscience majors. We followed the same general procedure adopted in Luo et al. (2005), i.e., participating students were asked to (1) take the pre-test, (2) use WILSIM to simulate landform evolution under 3 different scenarios, (3) answer questions after each scenario, and (4) take the post-test and relevant questionnaires. The pre- and post-tests, which are exactly the same and each consists of 9 questions, were designed to test students' general understanding of landform evolution processes and associated basic terminology. The questions after each scenario were geared specifically to that scenario and there are 4 questions for scenario 1, 2 for scenario 2, and 3 for scenario 3. The questionnaires were designed to gauge students' perception of, and response to the

Before (used in 2004)	After (used in 2006 and 2007)
Where in the landscape do you think deposition (accumulation of eroded and transported sediments) would be most unlikely to occur?	Where in the landscape do you think deposition (accumulation of eroded and transported sediments) would most likely occur?
A. In places where the slope is very steep	A. In places where the amount of sediments exceed the carrying capability of the stream
B. In places where the slope is very gentle	B. In places where there is a change in slope from steep to gentle
C. In places where there is a change in slope from steep to gentle	C. In places where the stream enters an open water body such as lake or ocean and slows down
D. In places where the stream enters an open water body such as lake or ocean	*D. All of the above*

Table 1. Comparison of a question and answer choices before and after change. *Italic* font indicates correct answer.

Before Change		After Change			
2004 Pre-test	2004 Post-test	2006 Pre-test	2006 Post-test	2007 Pre-test	2007 Post-test
47.46%	43.64%	68.70%	70.78%	70.83%	80.45%

Table 2. Comparison of percentages of correct answers to the question shown in Table1 before and after change.

landform simulation model. Based on the lessons learned in Luo et al. (2005) and students' qualitative feedback, we made the following changes:

(1) revised the poorly designed questions (see table 1 for an example);
(2) put the instructions in a separate PDF file and suggested that students print them out to follow along;
(3) briefly introduced WILSIM by demonstrating and explaining how it works in class before students start to work on the simulation scenarios on their own.

We modified the questions in the pre- and post-tests, following the tips of MCQ design as suggested in the literature as much as possible. For example, we changed the question that contains negative form in the stem to positive form and also made the choices consistent in grammatical structure and similar in length (see Table 1). Even though we still have "all of the above" as one of the choices, the result is an amazing > 20% increase in the percentage of students who answered this question correctly in both pre- and post-tests in both after years (2006 and 2007) as compared with the before year (2004) (see Table 2). The before question (contains negative "unlikely" in 2004) apparently confused a lot of students and even caused the percentage of correct answers to decrease from pre-test to post-test. After the change, the percentage of correct answers in post-test is higher than that in the pre-test in both 2006 and 2007. Changes to other poor questions also resulted in some improvement but not as dramatic as this example.

To determine if the differences between the means of total scores of pre-tests and post-tests were significant after the three changes described above, we conducted a t-test for related (or dependent) samples (e.g., Heiman, 1992). The paired sample sizes were 210 and 135 for 2006 and 2007, respectively. The null hypothesis is that the means of pre- and post-test scores are the same: $\mathbf{H_0}$: $\mu_{pre} = \mu_{post}$ and the alternative hypothesis is that the post-test mean will be higher than the pre-test mean: $\mathbf{H_1}$: $\mu_{pre} < \mu_{post}$. The results are shown in Table 3. In both semesters, the post-test means were higher than those of the pre-test and the differences were highly statistically significant ($p = 0.00215$ and 0.00252 for 2006 and 2007, respectively). We believe that changing the design of poor questions and demonstrating and explaining how WILSIM works in class made the difference.

The percentage of correct answers in the pre- and post-tests show similar pattern as compared that with 2004, i.e., students continue to do better in those pre- and post-test questions that are more directly related to the simulation scenarios (e.g., geologic time scale and erodibility) and not as well in those questions that are more related to terminology (e.g., tectonics) (Luo et al., 2005). The results from the questions after each specific scenario show no significant difference as compared with those in 2004.

The student self assessment shows that about 70% of the students strongly agree or agree that WILSIM help them understand concepts and processes involved in landform evolution and the specific effects of different factors such as erodibility contrast and tectonic uplift. This is consistent with the findings of Luo et al. (2005). The majority of the comments are very positive. Because of the changes made to the current assessment instrument, students' complaints regarding switching between different windows and not fully understanding the conceptual framework of the model have largely disappeared. The negative comments are mainly related to computer glitches and asking for higher resolution images.

CONCLUSIONS

After addressing the lessons learned in previous study and incorporating students' comments into the new design of the experiment (e.g., improving the poorly designed questions and demonstrating WILSIM in class before students run the scenarios), the post-test scores significantly increased over pre-test scores. This study demonstrated that the negative form in the stem of a MCQ was particularly confusing to students and changing it to positive form greatly improves the percentage of correct answers. The new results suggest that visualization and simulation do help improve

Enrollment			n	Mean	Standard Deviation	t	p
2006	380	Pre-test Post-test	210	6.952 7.276	1.903 1.892	-2.887	0.00215
2007	234	Pre-test Post-test	135	6.515 6.903	2.040 1.835	-2.853	0.00252

Table 3. Dependent t-test statistics between pre- and post-tests in 2006 and 2007.

student learning. The poor design of some questions in previous study might have masked out the positive effect of visualization and simulation on enhancing students' learning, as suggested in Luo et al. (2005). Such positive effect are also demonstrated, in both current and previous studies, in the students' self perception about visual learning and the fact that the majority of them can correctly answer questions specifically related to landform simulation scenarios. Students' comments about WILSIM are also much more positive than the previous study. The implication of this study is that some of the mixed results reported in literature about the effects of visualization and simulation on learning may be the result of poor design of the assessment instruments and not necessarily of the learning exercises.

ACKNOWLEDGEMENT

The WILSIM project is supported by the National Science Foundation under Grant No. 0127424. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We also would like to thank John A. Knox and Alan C. Czarnetzki for helpful reviews.

REFERENCES

Anglin, G. J., H. Vaez, K. L. Cunningham, 2004, The Role of Static and Animated Graphics, in Jonassen, D.H., editor, Handbook of research for educational communications and technology (2nd edition). Bloomington, IN, AECT, p. 865-916.

Brown, G., with Bull, J., and Pendlebury, M., 1997, Assessing Student Learning in Higher Education. London, Routledge, 317p.

Edelson, D. C., and Gordin, D., 1998, Visualization for learners: A framework for adapting scientists' tools, Computers and Geosciences, v. 24, p. 607-616.

Edelson, D. C., M. Brown, D. N. Gordin, D. A. Griffin, 1999, Making visualization accessible to students, GSA Today, February 1999, p. 8-10.

Fuhrman, M., 1996, Developing Good Multiple-Choice Tests and Test Questions, Journal of Geoscience Education, v. 44, p. 379-384.

Gordin, D. N., and Pea, R. D., 1995, Prospects for scientific visualization as an educational technology, Journal of the Learning Sciences, v. 4, p. 249-279.

Gronlund, N.E., 1988, How to construct Achievement Tests, Englewood Cliffs, Prentice Hall, 160p.

Heiman, G. W., 1992, Basic statistics for the behavioral sciences, Boston, Houghton Mifflin, 567 p.

Jensen, D., B. Self, D. Rhymer, J. Wood, and M. Bowe, 2002, A rocky journey toward effective assessment of visualization modules for learning enhancement in Engineering Mechanics, Educational Technology and Society, v. 5, p. 150-162.

Luo, W., K.L. Duffin, E. Peronja, J.A. Stravers, and G.M. Henry, 2004, A Web-based Interactive Landform Simulation Model (WILSIM), Computers and Geosciences, v. 30, p. 215-220.

Luo, W., Stravers, J., and K. Duffin, 2005, Lessons Learned from Using a Web-based Interactive Landform Simulation Model (WILSIM) in a General Education Physical Geography Course, Journal of Geoscience Education, v. 53, p. 489-493.

Race, P. Brown S. and Smith B., 2005, 500 tips on assessment, 2nd ed., RoutledgeFalmer, London and New York, 171 p.

Tversky, B., J. Morrison, M. Betrancourt, 2002, Animation: Can It Facilitate?, International Journal of Human Computer Studies, v. 57, p. 247-262.