Engagement, Understanding, and Achievement: Iterative Assessment and Refinement of Khipu Master

Ellen Jameson, Dr. Daniel T. Hickey, Indiana University Bloomington,
ejameson@indiana.edu, dthickey@indiana.edu

Abstract: We designed the Khipu Master game to support several 5th grade Common Core mathematics standards related to the base-10 number system. Taking the design from its starting point, we are using design-based research methods to refine both the design and the assessment instruments through a series of iterations. We have finished conducting a pilot study on basic elements of the game and a feasibility study on the first iteration, and are in the middle of collecting data for a comparison study on the second iteration. So far, gains on the pre/post assessment appear to have increased modestly after the first iteration.

Introduction

In the Khipu Master game, 5th-grade students arrive in pre-colonial Machu Picchu, Peru and progress toward their goals by learning to use a historical base-10 positional number system as an alternative representation of the modern standard base-10 number system taught in grades 1-5 (Brezine, 2011). Khipu were an Inca system of record-keeping which made use of a number system directly analogous to our natural numbers, storing numbers as knots in strings. Khipu were even used to depict multiple relationships across sets of data using string grouping, color, and other traits (Ascher & Ascher, 1981).

This highly contextualized alternative representation of the concepts underlying the targeted math standards allowed us to structure our assessment of student learning in the game into three levels of increasingly formal learning outcomes (Hickey & Jameson, 2012). *Engagement* in learning is documented by informally interpreting the patterns of student gameplay, and then the written artifacts generated in the game. *Understanding* is assessed using "curriculum-oriented" performance assessment items aligned to game concepts in the primary and secondary targeted standards. Gains in *achievement* are measured with "standards-oriented" test items aligned to the primary targeted standards. In Khipu Master, engagement is measured in a mid-game written response to a reflective prompt, while understanding and achievement are assessed in pre-and post-tests.

Theories of situativity and apprenticeship have been applied to many educational games (Barab et al, 2007), and are central to our design choices in Khipu Master. In this study we are focused on whether we can use evidence from assessments before, during, and after gameplay to adjust design and assessment elements based on these theories in a way that improves outcomes in subsequent iterations of the game. Our design intent is for the interactive khipu to be an affordance for students to test their base-10 assumptions in a new context (Greeno, 1991). The *khipukamayoq* takes players on as apprentices, and in that role scaffolds the use of khipu, gives automated feedback depending on student actions with the khipu and, with the help of the teacher, targeted feedback on student written reflections within the game, eventually welcoming players as khipu masters themselves (Lave and Wenger, 1991).

Methods

For each implementation, before and after the implementation, students complete a mathematics assessment before and after playing the game. The assessment included ten "curriculum-oriented" items that assessed their knowledge of based ten in the context of the khipu problems. These items are "proximal" in that they are a relatively direct assessment of the skills the students were expected to learn in the game in the context they were learned.

The assessment also included ten "standards-oriented" items that assessed knowledge of base ten in conventional test contexts, drawn randomly from released test that were aligned to two targeted standards but unrelated to any particular curricular context. These items were "distal" in that they required students to transfer their new knowledge in a context that was different than the context in which they learned it. These items predict whether the knowledge students took away from the game is likely to impact achievement on high-stakes test items that are aligned to the targeted standard.

The assessments were completed online, and two different versions of each assessment were developed by making minor changes in the item values. Counterbalanced forms of each were administered and the assessment was

completed online the period before beginning gameplay and the period after gameplay was completed.

We conducted a pilot implementation in a small number of classrooms, and used the result to update the design of the game, narrowing its focus to a few closely related standards, and expanded the scaffolding of conceptual elements of those standards. We tested this iteration in a larger feasibility study. After the feasibility implementation (of the first iteration), we adjusted the amount, style, and timing of feedback from the master *khipukamayoq* character.

Learning Results from the Feasibility Study So Far (first iteration)

Scores on the curriculum-oriented assessment increased from 3.06 to 4.81. This gain of 1.73 points was statistically significant [F (1, 15) = 21.0, p <.001] and very unlikely to have occurred by chance. Given a pooled standard deviation of 1.62, this gain of 1.73 points represented a gain of 1.06 SD, which just exceeded the threshold of 1 SD that is generally needed to consistently yield a statistically significant "echo" on a corresponding standards-oriented achievement test. As shown in Figure 1, the gains were very similar across the five teachers. This suggests that it is the designed and enacted game, rather than the way the game was taken up by the various teachers, which was responsible for the learning outcomes.

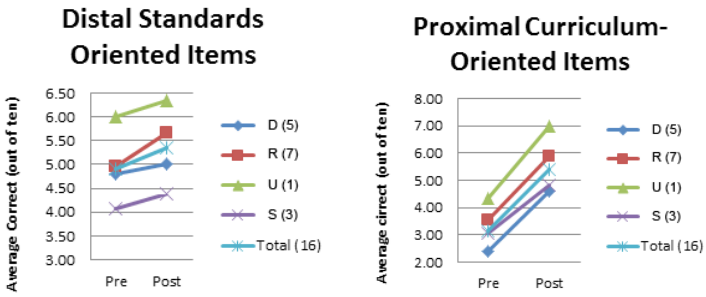

Figure 1 (a: Proximal Curriculum-Oriented Items, b: Distal Standards-Oriented Items)

Scores on the standard-oriented assessment increased from 4.06 to 4.37. This gain of .29 points did not reach statistical significance [F (1,15) = 0.8, p = .38. Given a pooled standard deviation of 1.9, this gain of .29 points represented a gain of 0.15 SD. A gain of this magnitude may or may not reach statistical significance with a larger sample size. Examination of scores by teacher shows gain in all classrooms, with somewhat larger gains in Mr. R and somewhat smaller gains in Ms. D. This may well represent chance differences.

References

Ascher, M, & Ascher, R. (1997). *Mathematics of the Incas: Code of the Quipu*. New York: Dover Publications.

Barab, S. A., Zuiker, S., Warren, S., Hickey, D., Ingram-Goble, A., Kwon, E-J., Kouper, I., & Herring, S. C. (2007). Situationally Embodied Curriculum: Relating Formalisms and Contexts. *Science Education*, 91(5), 750-782.

Brezine, C. (2011). "Khipu Problem Description." Personal communication, October 2011.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education*, 22, 170-218.

Hickey, D.T. & Jameson, E. (2012). Designing for participation in immersive educational videogames. In D. Ifenthaler, D. Eseryel, X. Ge (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives.* New York: Springer.