

Wright State University

CORE Scholar

Biological Sciences Faculty Publications

Biological Sciences

2013

Game-Based Curricula in Biology Classes: Multi-Level Assessment of Science Learning

Troy D. Sadler

William L. Romine

Wright State University - Main Campus, william.romine@wright.edu

Parker E. Stuart

Follow this and additional works at: <https://corescholar.libraries.wright.edu/biology>



Part of the [Biology Commons](#), [Medical Sciences Commons](#), and the [Systems Biology Commons](#)

Repository Citation

Sadler, T. D., Romine, W. L., & Stuart, P. E. (2013). Game-Based Curricula in Biology Classes: Multi-Level Assessment of Science Learning. .

<https://corescholar.libraries.wright.edu/biology/297>

This Conference Proceeding is brought to you for free and open access by the Biological Sciences at CORE Scholar. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.



From the

AERA Online Paper Repository

<http://www.aera.net/repository>

Paper Title Game-Based Curricula in Biology Classes:
Multilevel Assessment of Science Learning

Author(s) Troy D. Sadler, University of Missouri - Columbia;
William Romine, University of Missouri; Parker Stuart, University
of Missouri

Session Title Division C Section 1d Poster Session: Science
Learning and Teaching

Session Type Poster Presentation

Presentation Date 4/28/2013

Presentation Location San Francisco, California

Descriptors Assessment, Science Education, Technology

Methodology Quantitative

Unit Division C - Learning and Instruction

Each presenter retains copyright on the full-text paper. Repository users should follow legal and ethical practices in their use of repository material; permission to reuse material must be sought from the presenter, who owns copyright. Users should be aware of the [AERA Code of Ethics](#).

Citation of a paper in the repository should take the following form:
[Authors.] ([Year, Date of Presentation]). [Paper Title.] Paper presented at
the [Year] annual meeting of the American Educational Research
Association. Retrieved [Retrieval Date], from the AERA Online Paper
Repository.

Game-based Curricula in Biology Classes: Multi-level Assessment of Science Learning

Purpose of the Study

Video games played on computers, game consoles, tablets and smart phones have become a very popular medium in our society. A growing body of scholarship suggests that video games can support substantial learning (for reviews see Gee, 2007; Honey & Hilton, 2011; Squire, 2011). Research has documented ways in which games engage students deeply in thinking, learning, and collaboration (Gee, 2005). Evidence suggests that games can support development of critical thinking (Squire, 2006), creativity (Annetta, Cheng, & Holmes, 2010), and mastery and application of target concepts (Squire, 2011).

A review of the literature specific to gaming in science education provides evidence of games supporting student learning in science, but the majority of studies have investigated the effects of games on relatively small samples (Barab, et al., 2007; Hickey, et al., 2009; Nelson, 2007; Neulight, et al., 2007). (Ketelhut and colleagues' (2010) study of approximately 2,000 students is a notable exception.) Most of these studies have also explored the effects of gaming at the middle school level. Only Annetta et al. (2010) and Rosenbaum et al. (2007) examine high school students, and in both cases, the work does not focus primarily on student learning of content. Additionally, in those studies when student learning of biological content knowledge was the focus, the content tended not to be core biological knowledge (e.g., structure and function, genetics and evolution).

Based on the literature, we know video games can support student learning of science. We sought to build on these findings and to explore the effects of a game-based curriculum with students of varying academic levels. (We use the phrase "academic levels" to refer to the various tracks that are commonly used in high school science education

including general level courses, honors courses and advanced level courses such as Advanced Placement and International Baccalaureate.)

The study was guided by the following three research questions:

- 1) Can a game-based, biotechnology curriculum support student learning of biological principles?
- 2) What are the effects of a game based curriculum relative to varying academic levels?
- 3) How can multi-level assessment be used to illuminate patterns of learning associated with innovative curricula?

Theoretical Framework

Design of the project and the game that is explored within the project was informed by situated perspectives on teaching and learning (Cobb & Bowers, 1999; Lave & Wenger, 1991). Given this orientation, the design team prioritized opportunities for learners to engage in a rich environment that featured the materials, tools, and language derived from the target environment (i.e., a biotechnology laboratory). The project team also employed *multi-level assessment* (Ruiz-Primo et al., 2002; Hickey & Pellegrino, 2005) as a model for how we explored student learning in the context of the intervention explores. In this model, assessment opportunities are considered in terms of distance from the curriculum of interest. The distance of an assessment is determined by the extent to which that assessment is associated with the curriculum and learning experiences associated with that curriculum. For example, a teacher asking her students to explain ideas they are confronting in the midst of a learning experience represents assessment that is positioned very close to the curriculum. A quiz which challenges students to formalize their ideas will

still be closely related to the curriculum but not as close as the questions asked within the midst of the experience. A unit test, which challenges students to apply ideas from a curriculum to new contexts, will be even farther removed as compared to the quiz; however, the unit test is much closer to the curriculum than an end-of-course exam. Multi-level assessment offers a framework for considering how data derived from assessments at various distances can inform questions about the effectiveness of a particular intervention.

Methods

As a part of this study, we developed and implemented a computer-based game, *Mission Biotech* (MBt) as a tool for teaching biology. In MBt, players assume the role of a biotechnologist challenged to diagnose the pathogenic agent causing an emergent epidemic. Throughout the diagnosis process, non-player characters as well as clues and tools within the environment scaffold player use of biotechnology processes in order to find the virus behind the virtual outbreak. By the end of the game, players extract DNA, build a database of viral symptoms and characteristics, perform reverse transcription, conduct real-time polymerase chain reactions (PCR), and analyze and interpret data sets to draw conclusions and make decisions. They encounter and use virtual laboratory equipment and materials such as micropipettors, thermocyclers, centrifuges, water baths, lab notebooks, reagents, etc.

We collaborated with ten high school biology teachers who had participated in a professional development sequence and expressed interest in partnering with our team. The teachers implemented a standardized MBt instructional sequence which featured significant time devoted to game play as well as other classroom activities designed to support student understandings of content and processes featured within the game

environment. The ten teachers implemented MBt in 31 different sections of biology courses. We grouped the biology courses in three academic levels: general biology courses (Gen), honors biology courses (Hon), and advanced biology courses (Adv). A total of 642 students from the 31 different biology classes participated in the study; although, not all of these students completed every assessment.

Data Sources

The project team used the multi-level assessment framework described above to design assessments and collect data. Data were collected through proximal and distal assessments. For the proximal assessment, the project team designed a twenty item multiple-choice test directly aligned with the MBt curriculum. The proximal instrument was analogous to a unit test that might be administered by teachers following a curricular unit. The instrument was reviewed by an expert panel and pilot tested. Ultimately, the instrument was reduced to 19 items. The items demonstrated satisfactory collective reliability on both pre- ($\alpha = 0.715$) and post-tests ($\alpha = 0.826$).

The distal assessment was an 18 item multiple-choice exam aligned not with the MBt curriculum but rather with the eight content standards upon which MBt was based. The exam was pilot tested with 128 students not involved in the current study. The items demonstrated satisfactory collective reliability on both pre- ($\alpha = 0.836$) and post-tests ($\alpha = 0.853$). For analyses featured in this study, we used pre- and post-test data from the proximal and distal assessments.

We used the proximal MBt unit test and the distal standards-aligned exam to measure the effect of the MBt intervention on students' understanding of biological content knowledge using a quasi-experimental, pre-post repeated measures multiple analysis of

variance (MANOVA) design. The between subjects factor of academic level (Gen, Hon, and Adv) was tested in addition to the effect of the intervention.

Results

Multivariate tests reveal a significant difference between mean vectors across time ($\Lambda = 0.671$, $F = 127.742$, $p << 0.01$, $\eta^2 = 0.329$) and between ability levels ($\Lambda = 0.635$, $F = 66.63$, $p << 0.01$, $\eta^2 = 0.203$). We also find the difference between vectors of mean gain to be significantly different between ability levels ($\Lambda = 0.962$, $F = 5.074$, $p << 0.01$, $\eta^2 = 0.019$). Univariate tests on proximal and distal knowledge (see table 1) show that all differences are statistically significant at $\alpha = 0.01$ with the exception of the Time-Level interaction for proximal assessment ($F = 1.41$, $p = 0.24$, $\eta^2 = 0.01$).

Analysis of the Time-Level interactions from the two assessments reveals an interesting result. On the proximal assessment, students from all three academic levels show similar patterns of content gains (see figure 1). However, there is a significant interaction effect on the distal assessments. The General level group shows gains of greater magnitude than either the Honors or Advanced level groups (see figure 2). This interpretation is further supported by an analysis of the effect sizes for the pre/post gains for each academic level (see table 3). Post-hoc, paired t-tests suggest that post-test scores for the proximal and distal assessments were significantly higher than the pre-test scores. Effect sizes of the changes observed on the proximal test for all three academic levels were relatively high ranging from 0.75 to 1.03. On the distal assessment, only the General level demonstrated changes with a high effect size ($d=0.82$); changes observed in the other levels were statistically significant, but the effect sizes were considerably more modest ($d=0.29$) (Cohen, 1988).

Scholarly Significance

In this study, we used multiple assessments positioned at different distances from the curriculum to generate “a better picture of the extent of the effect” (Ruiz-Primo et al., 2002, p. 371) of an innovative, game-based curriculum. All three of the groups studied (students from General, Honors and Advanced biology classes) demonstrated statistically and practically significant gains on the post-test, distal assessment as compared to pre-test scores. We interpret this to mean that students learned at least some of the biological content featured in the game-based curriculum.

While the distal exam is certainly not as far removed from the curriculum as a state or national standardized test, it serves as a proxy of sorts considering the limited temporal nature of the intervention. The fact that all groups represented in this study demonstrated statistically significant gains on the distal exam provides strong evidence of the potential of game-based curricula to support student learning that can translate on the kinds of metrics that are most valued in the current system of schooling (Orpwood, 2001). A two to three week game-based unit, in and of itself, is not going to drastically improve student scores on comprehensive, standardized tests, but our research suggests that incorporating a game-based unit in science courses can support the kind of learning that is expected and tested on these kinds of examinations.

Researchers using multi-level assessments have consistently documented that distal assessments are less sensitive to learning gains as compared to more proximal assessments (Barab, et al., 2007; Cross, Taasobshirazi, Hendricks, & Hickey, 2008; Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). Our results support this conclusion. Effect sizes associated with the univariate ANOVAs indicate that changes documented on the proximal

assessment ($\eta^2 = 0.26$) were twice as large as changes on the distal assessment ($\eta^2 = 0.13$). Because we were able to disaggregate the data based on academic level, an additional pattern became observable. The effect sizes, calculated using Cohen's d , for students in the Honors and Advanced classes dropped in moving from analysis of the proximal to the distal assessments (Hon: 1.03 to .29; Adv: .75 to .29). This result is consistent with other multi-level assessment studies of innovative curricula. However, the effect size for students in the General classes on the distal assessment ($d = .82$) was only slightly lower than the effect size observed on the proximal assessment ($d = .91$).

Had we not adopted a multi-level assessment approach, the conclusions drawn from this study would have been more limited. If research questions 1 and 2 were addressed with data only from the proximal exam, then we would have concluded that the game-based curriculum was equally successful in supporting learning for all students. If the distal assessment had provided the only data under consideration, then we would have concluded that the game-based curriculum was fairly effective for the General group but not particularly effective for the more advanced students. By considering both data sources, we are able to see that the game supported learning in all three groups, but that it was more helpful in supporting the lower ability group in their performance on a proxy for a standardized examination than their peers who demonstrated stronger performance on the distal exam prior to the learning experience. This more nuanced interpretation of the available data is quite useful in terms of making sense of the effects of an innovative curriculum. This research serves as a case for the use of multi-level assessment in the study and evaluation of game-based and other innovative curricula.

References

- Annetta, L. A., Cheng, M.-T., & Holmes, S. (2010). Assessing twenty-first century skills through a teacher created video game for high school biology teachers. *Research in Science and Technological Education, 28*(2), 101-114. doi: 10.1080/02635141003748358.
- Barab, S. A., Sadler, T. D., Heiselt, C., Hickey, D., & Zuiker, S. (2007). Relating narrative, inquiry, and inscriptions: Supporting consequential play. *Journal of Science Education and Technology, 16*(1), 59-82. doi: 10.1007/s10956-006-9033-3.
- Cross, D., Taasobshirazi, G., Hendricks, S., & Hickey, D. T. (2008). Argumentation: A strategy for improving achievement and revealing scientific identities. *International journal of science education, 30*(6), 837-861. doi: 10.1080/09500690701411567.
- Cobb, P., & Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher, 28*(2), 4-15. doi: 10.3102/0013189X028002004.
- Gee, J. P. (2005). *Why video games are good for your soul*. Pleasure and learning. Common Ground.
- Gee, J.P. (2007). *What video games have to teach us about learning and literacy*. Palgrave Macmillan: New York. doi: 10.1145/950566.950595.
- Hickey, D. T., Ingram-Goble, A. A., & Jameson, E. M. (2009). Designing assessments and assessing designs in virtual educational environments. *Journal of Science Education and Technology, 18*(2), 187-208. doi: 10.1007/s10956-008-9143-1.
- Hickey, D. T., & Pellegrino, J. W. (2005). Theory, level, and function: Three dimensions for understanding transfer and student assessment. In J. P. Mestre (Ed.), *Transfer of*

learning from a modern multidisciplinary perspective (pp. 251-293). Greenwich, CT: Information Age Publishers.

Honey, M. A. & Hilton, M. H. (2011) *Learning Science: Computer Games, Simulations, and Education*. National Research Council. Washington, DC: The National Academies Press.

Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology*, 16(1), 99-111. doi: 10.1007/s10956-006-9038-y.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Nelson, B. C. (2007). Exploring the use of individualized, reflective guidance in an educational multi-user virtual environment. *Journal of Science Education and Technology*, 16(1), 83-97. doi: 10.1007/s10956-006-9039-x.

Neulight, N., Kafai, Y. B., Kao, L., Foley, B., & Galas, C. (2007). Children's participation in a virtual epidemic in the science classroom: Making connections to natural infectious diseases. *Journal of Science Education and Technology*, 16(1), 47-58.

Orpwood, G. (2001). The role of assessment in science curriculum reform. *Assessment in Education*, 8, 135-151. doi: 10.1007/s10956-006-9029-z.

Rosenbaum, E., Klopfer, E., & Perry, J. (2007). On location learning: Authentic applied science with networked augmented realities. *Journal of Science Education and Technology*, 16(1), 31-45. doi: 10.1007/s10956-006-9036-0.

- Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching*, 39(5), 369-393. doi: 10.1002/tea.10027.
- Squire, K. (2006). From content to context: Videogames as designed experience. *Educational researcher*, 35(8), 19-29. doi: 10.3102/0013189X035008019.
- Squire, K. (2011). *Video Games and Learning: Teaching and Participatory Culture in the Digital Age*: Teachers College Press.

Table 1. Proximal and distal test results.

<i>Proximal</i>				
	Pre-test	Post-test		Cohen's
Level	Mean(SD)	Mean(SD)	T ^a	D ^b
Gen (n = 70)	6.17(3.15)	10.13(5.26)	5.77	0.91
Hon (n = 248)	7.95(4.01)	12.83(5.34)	12.77	1.03
Adv (n = 219)	11.14(5.03)	15.28(5.95)	8.83	0.75
<i>Distal</i>				
	Pre-test	Post-test		Cohen's
Level	Mean(SD)	Mean(SD)	T ^a	D ^b
Gen (n = 71)	5.96(2.83)	9.18(4.82)	5.37	0.82
Hon (n = 233)	10.30(4.66)	11.61(4.55)	4.22	0.29
Adv (n = 220)	14.22(3.88)	15.25(3.08)	4.50	0.29

^ap << 0.01 for all tests

^bEffect size: D>0.80 high; 0.50<D<0.79 moderate; 0.20<D<0.49 low (Cohen, 1988)

Figure 1. Results from the proximal test.

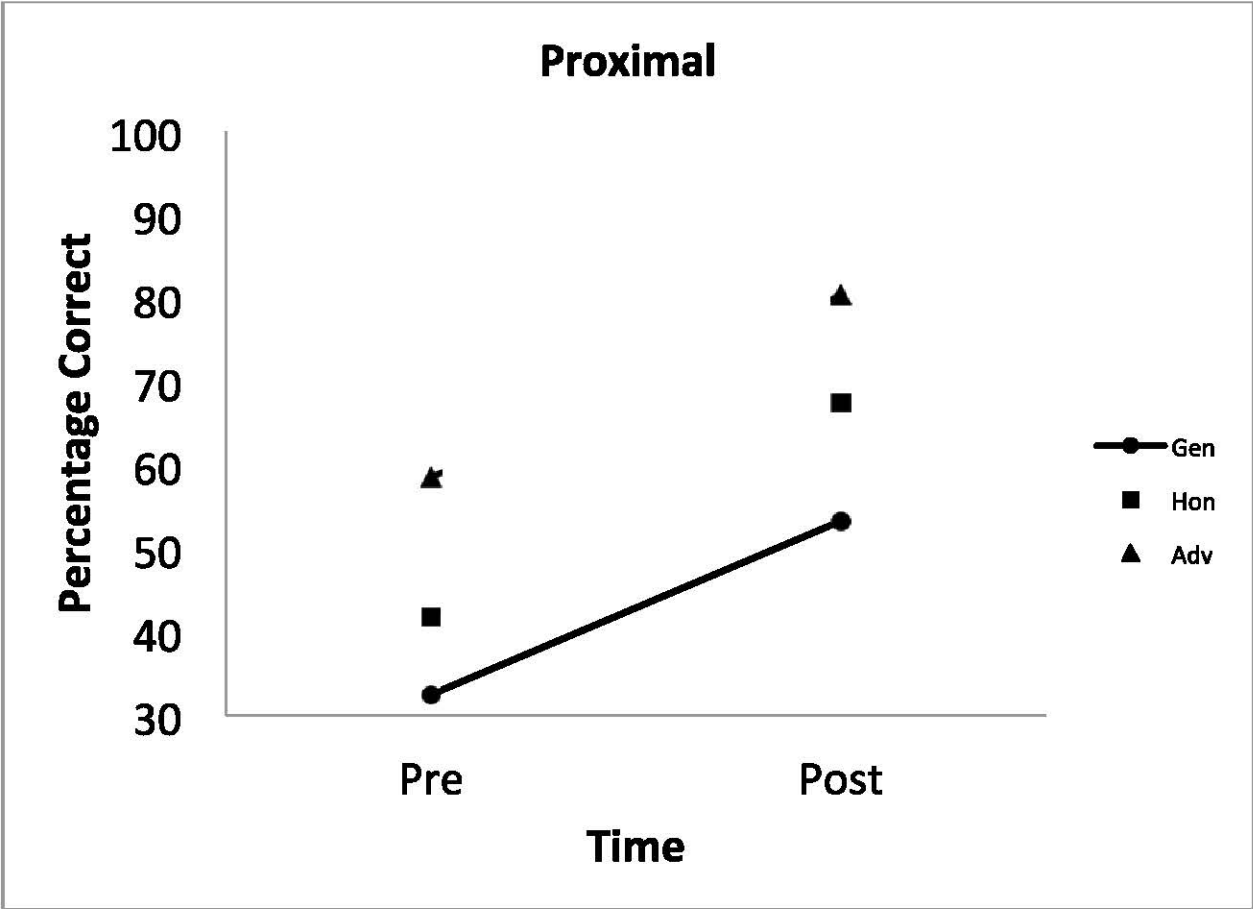


Figure 2. Results from the distal test.

