

Testing the Multimedia Principle in the Real World: A Comparison of Video vs. Text Feedback in Authentic Middle School Math Assignments

Korinn S. Ostrow
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609
(508) 373 - 2676
ksostrow@wpi.edu

Neil T. Heffernan
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609
(508) 831 - 5569
nth@wpi.edu

ABSTRACT

As technology evolves and design options for web-based homework support systems expand, researchers are left with questions regarding best practices. These platforms often provide students correctness feedback meant to guide learning and offer dynamic tutoring to help students solve difficult problems. Feedback typically consists of bland text and worked examples, but as hypermedia gains prevalence, researchers are turning their focus to the appropriate use of such elements in e-learning environments. The following study assesses the effects of feedback medium within a randomized controlled trial conducted using ASSISTments, an adaptive math tutor. Results suggest that video feedback enhances learning outcomes and is well perceived by student users. These findings are of particular interest to the Learning Sciences, with intent to optimize e-Learning design.

Keywords

E-Learning, Cognitive Theory, Multimedia Principles, Feedback, Adaptive Tutoring, ASSISTments, Randomized Controlled Trial.

1. INTRODUCTION

A leader in the field of e-Learning, Richard Mayer has defined various multimedia principles for the optimal design of technology supported learning environments such as web-based homework support systems [3]. Rooted in cognitive theory, these principles call for the design of learning environments that are driven by an active learning process and that take the restraints of cognitive load and working memory into consideration [3][7]. Still, researchers seeking to enhance student engagement, motivation, and persistence, they are left questioning how to optimize the learning environment without overloading learners.

Mayer also posits that learners utilize separate information processing channels to internalize information; under the *redundancy principle*, material offered through one channel (i.e., a narrated passage) should not be simultaneously presented through another (i.e., text accompanying the narration) [3]. When such circumstances occur, the learner's attention is split across redundant content, depressing intake from both channels and

hampering learning. Further, the *modality effect* suggests that learning gains are greater for narrated content than for content presented as text [7]. Based on these principles, the use of video, when presented without redundant textual explanation should appeal to both auditory and visual processing channels without risking overload.

Video is not novel to education, and it is growing increasingly popular due to the concept of the "flipped classroom," which often parallels the use of web-based homework support systems. While the quality of evidence for the flipped classroom has not yet proven impressive [4], the trend speaks to the growing accessibility of technological resources in education. Self-recorded video lectures and feedback offer teachers the opportunity to be deeply involved in student learning while simultaneously enhancing ownership of the technology [5].

Contrary research has suggested that video is not universally successful in promoting learning gains. In his early work on the effect of educational movies, Pane [9] noted mixed results as a function of content, offering evidence that the use of video may improve the speed of immediate recall, yet potentially harm long-term learning. Negative effects of video may include prolonged time-on-task that potentially leads to boredom or frustration, the inability to appropriately convey abstract content material, and the likelihood of technological difficulties that prevent students from adequately accessing materials.

In the present study, the ASSISTments platform is used to compare the delivery methods of feedback messages within a mathematics e-Learning environment. Prior research has found that dynamic graphics are more effective than static graphics in mathematics realms [7], and thus, we hypothesize that video will have a positive effect on learning gains within this system. Since its inception, ASSISTments has delivered significant results surrounding the use of textual feedback in the form of scaffolding and hints [10][11][12]; the present study serves as a preliminary exploration into replacing textual feedback with video.

Thus, we pose the following research questions:

1. Are learning outcomes enhanced when scaffold feedback is delivered using video rather than text?
2. Can we determine if students disproportionately internalize feedback based on the medium, given next question performance and response time?
3. Based on self-report measures, do students respond positively to the addition of video to their assignment?

2. METHODS

2.1 Participants

A set of six questions requiring students to use the Pythagorean theorem was assigned to 139 8th grade students using

ASSISTments. This student population was comprised of four classes of differing skill levels that spanned four suburban middle schools in Massachusetts and Ohio. All students were familiar with ASSISTments and used the system on a regular basis as part of classwork and homework assignments.

2.2 Design

The Pythagorean theorem problem set was derived from pre-existing ASSISTments certified material, based on Common Core State Standards and chosen in an attempt to match 8th grade fall curriculum. The structure of the problem set relied on three questions with text feedback (A, B, C) and three isomorphic questions with video feedback (A*, B*, C*). Each question and its morph were of similar difficulty and were therefore considered interchangeable (i.e., A and A*). The questions are available at [8] for further comparison.

The fixed question patterns depicted in Table 1 were rooted in the intention to allow all students an equivalent opportunity to experience both feedback styles. Thus, the four groups were designed to house fixed question patterns from which we could assess the impact of video versus text at various points throughout the problem set. Random assignment was attained by allowing ASSISTments to allocate students into one of the four groups at the start of the assignment. As depicted in Table 1, students assigned to Group 1 received video feedback if they answered question 1 incorrectly, text feedback if they answered question 2 incorrectly, and so on.

Table 1. Group design

Linear Order	1	2	3	4	5	6
Group 1	A*	B	C*	A	B*	C
Group 2	A	B*	C	A*	B	C*
Group 3	A*	B*	C	A	B	C*
Group 4	A	B	C*	A*	B*	C

*Depicts question morph with video feedback

Video content was designed to mirror textual feedback in an attempt to provide identical assistance through both mediums. Each video simply featured the lead researcher reading a feedback message while referring to the question content on a whiteboard. Figure 1 depicts question C* with video feedback, while Figure 2 depicts the question morph (C) with text feedback. All video material can be accessed at [8].

Both types of feedback were set to load incrementally with incorrect responses or if the student requested to break the question down into steps. Videos were set to play automatically, allowing students to gain information with equal efficiency regardless of feedback medium, and perhaps making it harder or

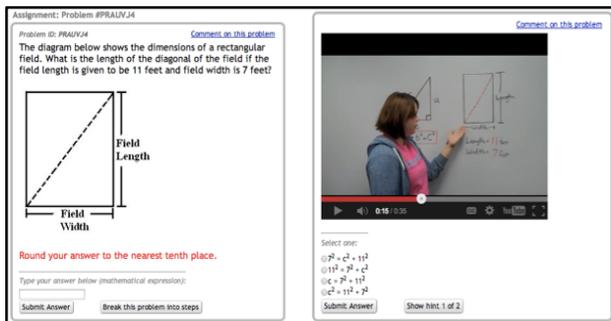


Figure 1. Video feedback for question C*

more inconvenient to “game the system,” or click through the scaffold steps in rapid succession.

For each group, four post-test survey questions asked students to judge, using a simple three-measure Likert scale (i.e., not at all, somewhat, a lot), if they felt video feedback was helpful, if it was enjoyable, if they would prefer similar videos in future assignments, and what effect video feedback had on their focus. For the entire student experience, see [8].

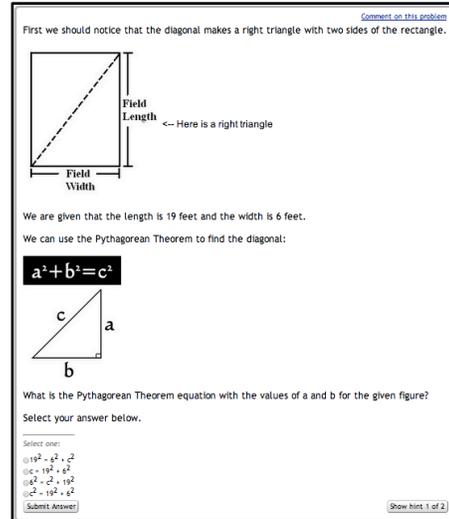


Figure 2. Text feedback for question C

2.3 Procedure

The problem set was assigned to students in the manner consistent with their teacher’s usual use of ASSISTments (i.e., as either classwork or homework). Students were free to work at their own pace and were not required to complete the assignment in one sitting. Log data was compiled for each student’s performance, including elements such as first action, correctness, response time, attempts, and hints requested. Delegating random assignment to the tutor produced results that were less than optimal, as significantly fewer students were assigned to Group 2 and Group 4. However, assessment of the code controlling ASSISTments’ random assignment function concluded that this anomaly was not influenced by any student attribute or system characteristic.

Table 2 explains initial group assignment and the process for excluding students from analysis. A total of 139 students were originally assigned (OA) the problem set. Six students failed to log enough progress to initiate a group assignment and were therefore excluded. Of the remaining 133 students, 13 students did not complete the problem set (I), and 31 students tested out (TO) (these students answered each question correctly and failed to receive feedback of either style). A disproportionate number of students tested out of Group 3, likely as a function of random assignment and small sample size.

Table 2. Students excluded from analysis

	OA	I	TO	G	Remaining
Group 1	35	4	7	0	24
Group 2	29	3	6	4	16
Group 3	43	4	11	2	26
Group 4	26	2	7	4	13
Total	133*	13	31	10	79

*Six students failed to initiate a condition.

In prior research, “gaming the system” within ASSISTments has been defined as consistent answer seeking behavior displayed in rapid succession (i.e., clicking through all hints or scaffolds for completion) [1]. As such, “gamers” were operationalized as any student who clicked through question A (or A*) and its four scaffolds, regardless of feedback medium, at a rate faster than five seconds per response. By this loose definition, a total of 10 students qualified as “gamers” (G) and were removed prior to analysis as shown in Table 2.

Our primary analysis assessed student performance on the second question as a function of the feedback medium they experienced after incorrectly answering the first question. For question 1, Groups 1 and 3 were presented video feedback (A*), while Groups 2 and 4 were presented text feedback (A). We were therefore able to collapse these groups when analyzing second question performance. Based on Table 2, the removal of gamers significantly differs when the Groups are collapsed: for Groups 1 and 3, only 7.1% of students are removed from the remaining sample, while Groups 2 and 4 lose 43.5% of the remaining students. Considering our operational definition of gamers, and noting that Groups 2 and 4 received text feedback upon incorrectly answering question 1, the discrepancy found here suggests that video feedback may deter gaming. To better understand this bias, the proceeding analysis is carried out both with and without gamers for comparison.

3. RESULTS

3.1 Second Question Analysis

After considering the aforementioned exclusion methods, 79 students were remaining for analysis (89 when gamers were included). To address our initial research question, we assessed second question performance in students who had received feedback on question 1, as summarized in Table 5. Learning outcomes were enhanced for students who received video feedback ($M = 0.77$, $SD = 0.43$) rather than text feedback ($M = 0.63$, $SD = 0.50$), approaching significance at $p = 0.143$, with an effect size¹ of 0.32, 95% CI [-0.28, 0.91]. When gamers were included to analyze the effect of the selection bias, the improvement for students who had received video ($M = 0.76$, $SD = 0.44$) versus text ($M = 0.52$, $SD = 0.51$) became statistically significant, $p < .05$, with an effect size of 0.50, 95% CI [-0.03, 1.03].

Table 5. Summary of second question analysis

	Video	N	Text	N	ES ¹	95% CI
Performance						
w/o Gamers	0.77 (0.43)	35	0.63 (0.50)	16	0.32	[-0.28, 0.91]
w/ Gamers	0.76 (0.44)	37	0.52 (0.51)	23	0.50*	[-0.03, 1.03]
Resp. Time						
w/o Gamers	134.86 (118.76)	35	421.77 (1122.27)	16	-0.45	[-1.05, 0.15]
w/ Gamers	129.72 (117.46)	37	307.33 (943.50)	23	-0.30	[-0.82, 0.23]

Note. Time is depicted in seconds as: mean (standard deviation). * $p < .05$.

Further analysis of second question performance suggested that response time was faster for students who had received video ($M = 134.86$, $SD = 118.76$) rather than text ($M = 421.77$, $SD = 1122.27$), approaching significance at $p = 0.068$, with an effect size of -0.45, 95% CI [-1.05, 0.15]. When gamers were included

¹ Effect sizes are reported throughout using Hedges correction [2].

for comparison, students who had incorrectly answered the first question and received video feedback performed faster ($M = 129.72$, $SD = 117.46$) than those receiving a text scaffold ($M = 307.33$, $SD = 943.50$), but results were not significant and the effect size dropped to -0.30, 95% CI [-0.82, 0.23]. As gaming was defined as rapidly clicking through questions and feedback, it is not surprising that time measures would drop in this manner. While these results portray consistent trends approaching significance, they should be taken with caution, as the number of students who received text feedback was disproportionately smaller than the number of students who received video feedback.

3.2 Response Time Within Feedback

To address our second research question, we examined students’ overall experience within each type of feedback. Students saw a total of 186 scaffold levels of video feedback, and 171 scaffold levels of text feedback while completing their assignment. On average, response time during video feedback ($M = 202.51$, $SD = 337.99$) was longer than response time during text feedback ($M = 35.18$, $SD = 28.74$) approaching significance at $p = 0.085$, with an effect size of 0.68, 95% CI [0.47, 0.90]. When gamers were included for comparison, students saw a total of 241 levels of video feedback, and 231 levels of text feedback. Average response times dropped within both feedback styles, yet response time during video feedback remained longer ($M = 169.28$, $SD = 268.44$) than response time during text feedback ($M = 28.38$, $SD = 21.67$), approaching significance at $p = 0.076$, with an effect size of 0.73, 95% CI [0.54, 0.92].

These results suggest that there was no significant difference in the overall number of feedback levels experienced by students as a function of feedback medium. On average, students spent 3 minutes and 23 seconds within video feedback and only 35 seconds within text feedback. When gamers were considered, less time on average was spent within each feedback style, with students spending 2 minutes 49 seconds within video feedback and only 28 seconds within text feedback. Thus, students consistently spent more time within video feedback, suggesting that they actually took the time to watch the videos and internalize the content whereas they seemed to gloss over text feedback.

3.3 Survey Response Analysis

Of all students available for analysis, 53 answered the four post-test survey questions. Student responses are proportioned in Table 8. Taken together, we consider the survey results to suggest that video feedback is well perceived by students. Essentially, 83% of students reported that they would at least somewhat prefer ASSISTments to use video more often. Coupled with the student performance findings discussed above, we feel that video may be a beneficial tool for ASSISTments and that further exploration regarding the long-term effect on learning is required.

Table 8. Student responses to post-test survey questions

	Not at all	Somewhat	A lot
How helpful did you find the videos as you completed your assignment?	14%	43%	43%
How much did you enjoy the videos?	17%	57%	26%
Would you like it if more of your ASSISTments assignments used videos?	17%	43%	40%
Did you feel more focused on your assignment when the question had videos?	30%	38%	32%

4. CONTRIBUTIONS AND DISCUSSION

Although Mayer's work has been a predominant influence on the field of multimedia infused learning, much of his research has assessed college undergraduates in psychology labs. Thus, his results suggest a massive and seemingly unrealistic effect when compared to most real-world educational interventions. According to recent research detailing average effect sizes in educational settings, Lipsey, et al. [6] note that at the middle school level, researcher developed studies with specialized topics tend to show strength with effect sizes of approximately 0.43. The present study is on par with this trend, with effect sizes for second question analysis ranging from 0.32 to 0.51. We argue that these results provide a contribution to the Learning Sciences and help establish a basis for future research.

Based on our findings, we feel that video feedback may be a significantly beneficial tool for e-Learning. Immediate learning gains, represented by second question performance after receiving feedback on question 1, were significantly greater in students who experienced video feedback. Our results suggest that the use of video forces the learner to slow down and internalize the concept that is being taught, as depicted by consistent trends for response times within the feedback experience. Although text feedback consistently provides a faster alternative for skilled readers, perhaps adaptively slowing the pace more closely mimics the actions of a human tutor.

It should be noted that video feedback appears to have deterred gaming behavior. This may have been due in part to novelty, but was likely a function of the automatic nature of video playback. When a student tried to game through a question, each scaffold level would present another video until they were all playing simultaneously. A slightly more qualitative inspection of gaming behavior within this problem set suggested that at least three of the students labeled as gamers corrected their behavior after being exposed to video feedback. Future research is required to determine if video feedback provides a beneficial intervention for this population in general.

Regardless of the cause, let us assume for a moment that these effects are valid and reliable, and that video feedback significantly enhances student performance. With the growing popularity of web-based homework support systems and the ubiquitous nature of video servers such as YouTube and SchoolTube, teachers and instructional designers may be overlooking a valuable tool. The videos used in this study were of low production quality, shot in a single take, and featured a non-professional actress reading from a script. Teachers with years of expertise in providing feedback could arguably record a short video on their smartphone or tablet that would outperform the content used in this study. The use of video within e-Learning environments has the potential to streamline the process of repetitive one-to-one tutoring and boost the teacher's efficiency in the classroom. While pedagogical agents have become a popular tool for feedback delivery within e-Learning environments, the same messages may carry significantly more power when delivered by the student's teacher. A multitude of brief interactions offering personalized and appropriately timed feedback, guidance, and motivation, could become an important step toward truly adaptive tutoring.

Future implementations of this study should utilize a more powerful pre/post-test design with additional far transfer items and the use of open-ended survey response options to gauge student feedback. We also suggest that future endeavors compare a purely video condition to a control containing only textual

feedback, perhaps using an AB design with multiple content topics to maintain fair treatment. Future work should also attempt to pinpoint critical elements driving the effects of video, such as motivation, novelty, personalization, and engagement.

5. ACKNOWLEDGEMENTS

We acknowledge funding from NSF (#1316736, 1252297, 1109483, 1031398, 0742503), ONR's "STEM Grand Challenges," and IES (#R305A120125, R305C100024). Thanks to S.D.O. & L.P.B.O.

6. REFERENCES

- [1] Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A. & Koedinger, K. 2008. Why students engage in "Gaming the System" behavior in interactive learning environments. *Journal of Interactive Learning Research*. 19, 2, 185-224.
- [2] CEM. 2013. Effect Size Calculator. Centre for Evaluation & Monitoring, Durham University. Accessed 11/8/2013 at <https://tinyurl.com/nt4snvo>.
- [3] Clark, R.C. & Mayer, R. E. 2003. *e-Learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: Pfeiffer.
- [4] Goodwin, B. & Miller, K. 2013. Research says/evidence on flipped classrooms is still coming in. *Educational Leadership: Technology-Rich Learning*. 70, 6, 78-80.
- [5] Kelly, K., Heffernan, N., D'Mello, S., Namias, J., & Strain, A. 2013. Adding teacher-created motivational video to an ITS. In Proceedings of the Florida Artificial Intelligence Research Society Conference (FLAIRS 2013). 503-508.
- [6] Lipsey, M.W., Puzio, K., Yun, C., Hebert, M. A., Steinkamp-Fry, K., Cole, M.W., Roberts, M., et al. 2012. Translating the statistical representation of the effects of education interventions into more readily interpretable forms. *National Center for Special Education Research*. Washington, DC.
- [7] Mayer, R.E. (Ed). 2005. *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.
- [8] Ostrow, Korinn. 2013. Pythagorean Theorem Math Study. Accessed 11/8/2013.
 - a. Student Experience by Group, Randomized Controlled Trial, Data: <https://tinyurl.com/lr529jp>
 - b. Video Scaffolds: <https://tinyurl.com/mcc4w8z>
- [9] Pane, J.F. 1994. Assessment of the ACSE science learning environment and the impact of movies and simulations. Carnegie Mellon University, *School of Computer Science Technical Report CMU-CS-94-162*, Pittsburgh, PA.
- [10] Razzaq, L. & Heffernan, N.T. 2006. Scaffolding vs. hints in the ASSISTments system. In Ikeda, Ashley & Chan (Eds). *Proceedings of the Eighth International Conference on Intelligent Tutoring Systems*. Springer-Verlag: Berlin. 635-644.
- [11] Singh, R., Saleem, M., Pradhan, P., Heffernan, C., Heffernan, N., Razzaq, L. Dailey, M. O'Connor, C. & Mulchay, C. 2011. Feedback during web-based homework: the role of hints. In Biswas et al. (Eds). *Proceedings of the Artificial Intelligence in Education Conference*. 328-336.
- [12] Wang, Y. & Heffernan, N. 2011. The "assistance" model: leveraging how many hints and attempts a student needs. In Proceedings of the *Florida Artificial Intelligence Research Society Conference (FLAIRS 2011)*.