Journal of Applied Research in Memory and Cognition 1 (2012) 251–253

Contents lists available at SciVerse ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Commentary

Promising principles: Translating the science of learning to educational practice

David B. Daniel*

Department of Psychology, James Madison University, 91 E. Grace St, MSC 7704, Harrisonburg, VA 22807, United States

ARTICLE INFO

Article history:
Received 9 October 2012
Accepted 10 October 2012

The poor connections between practitioners and researchers have many resultant costs. Perhaps the most often cited is the difficulty of getting research into practice: well-meaning researchers generate ideas, findings, practices, and programs that are not optimally communicated to practitioners, and often not optimally designed for use in practice.
Catherine Snow– Harvard Graduate School of Education

At a time when education is under scrutiny and budgets are tight, Roediger and Pyc (2012) argue that it makes good sense to rely upon robust scientific findings as a source of potential high-impact/low-cost suggestions rather than investing in unproven but flashy alternatives (e.g., Daniel & Willingham, 2012). Policy-makers and educators should insist upon evidence of effectiveness before investing time and resources to "improve" educational practice. While cognitive psychologists are familiar with promising findings from the Science of Learning for improving classroom practice, educators are not. Roediger and Pyc (2012) have done a nice job restating the information presented by many others over the years, highlighting the more promising findings from the Science of Learning and their potential applications (see also Bransford, Brown, & Cocking, 2000; Dunlosky, Rawson, Marsh, Nathan & Willingham, in press; Halpern & Hakel, 2002, 2003; Matlin, 2007; Pashler et al., 2007; Worrell et al., 2009, etc.).

Unfortunately, Roediger and Pyc (2012) have made the same error made by researchers in this tradition for many years: they are recommending ubiquitous use of lab-based findings before these strategies have undergone the contextual vetting and design that would allow practitioners and policy-makers to confidently leverage them as reliable pedagogical tools in typical classroom contexts. So, while I agree that the reviewed findings may have potential for positive impact in educational contexts, the recommendations for classrooms are premature. I believe that the conclusion that ought to be drawn from their review is that we need a targeted investment in translational research and the development of productive methods of practice with the goal of understanding how, when and under what constraints to apply these strategies in educational contexts.

Consider the following translational steps for moving from the lab to the classroom:

(1) Exploration of promising findings in the lab – yields hypothesis for practice
(2) Careful experimentation in select classroom contexts – yields promising principle
(3) Development and design of classroom/teacher-friendly methods integrating promising principle into everyday practice – yields promising practice (and encourages practitioner fidelity to practice)
(4) Coordinated experimentation in more representative and complex settings (clinical trials, etc.) – yields best practice
(5) Dissemination and continued refinement

The strategies reviewed by Roediger and Pyc (2012) have compellingly accomplished the first step, yielding the potential to impact classroom learning: a hypothesis for practice. The literature has also begun to demonstrate growing success with selected materials in a small number of classrooms, allowing us to begin defining the parameters of strategies that may hold promise for classroom use. Roediger and Pyc (2012) jump to recommending the reviewed strategies for broad classroom use (step 5) before accomplishing the essential tasks of developing these promising principles into usable methods (step 3) with demonstrable impact in a variety of subjects in typical classrooms (step 4).

1. Promising principles

Research-derived strategies are considered promising when they have demonstrated success in the lab and have also been tested in controlled classroom settings that allow us to demonstrate that the improvements were, indeed, due to these treatments (Worrell et al., 2009). As described above, a promising principle, however, has several steps of development before we can

* Tel.: +1 540 568 7359; fax: +1 540 568 4916.
E-mail address: danieldb@jmu.edu

http://dx.doi.org/10.1016/j.jarmac.2012.10.004

confidently recommend it for ubiquitous classroom practice. The reviewed findings still need to be translated into successful as well as usable pedagogical strategies and then vetted in representative classroom contexts before recommendations for their use can be made with confidence.

It is also important to insure that the laboratory practices align with the goals of the educators. Recommendations for practice need to take into account artifacts of experimental design that may not align with the demands and/or goals of other contexts. If the goals or methods are not aligned with those deployed by educators, a promising strategy could be resisted by the educator for very good reasons, few of which are on the researcher's radar. Most studies of the testing effect, for example, tend to use quiz questions identical to the questions on the exam. One would be hard-pressed to find an educator who would give such a practice high value. Several studies have demonstrated the potential for the testing-effect to increase learning for related questions and far transfer (see Carpenter, 2012 for a review), but these paradigms have not been fully developed for broad educational application. While studies in this area are beginning to move into representative classrooms, many have not done so in an authentic fashion.

2. From principle to practice

It is necessary to provide teachers guidelines and specific techniques for applying promising findings, rather than simply communicating the findings and their potential promise (Bransford et al., 2000). How do we integrate these principles into a system for teachers to efficiently achieve educator-desired learning goals? These strategies have not yet been designed as educational interventions, optimized for the classroom or aligned with curricular goals. We are left little idea how to deploy this set of recommendations in typical classrooms or how they may interact with each other to enhance, or subvert, learning. While Roediger and Pyc (2012) review relevant classroom studies that offer promise and then provide examples of what potential applications may look like, they are not able to provide adequate guidance, structure or boundaries to encourage efficient, productive and responsible educational use given the present state of this literature.

For example, several of the reviewed techniques are useful for fact learning, which is important. But how do we encourage the more flexible and conceptual learning that characterizes the goals of many educators? Since explanatory questioning has demonstrated success with far transfer, would it be useful to supplement testing with such questioning? Or, would they interact to subvert learning? Should we interleave within academic subjects or would it be best to interleave across subjects? What are the boundary conditions associated with each strategy? Strategies that demonstrate effects under constrained and supported conditions in the lab require design that supports optimal processing in representative contexts (Daniel & Poole, 2009).

3. Increasing complexity, decreasing control

As researchers, we need to acknowledge that the classroom is complex. The lab exists to remove complexity in order to allow for more intense and targeted investigation. With each step from the lab toward the classroom, new variables and, importantly, new interactions become relevant. We need to be open to the potential that recommended pedagogical strategies may interact with content, other strategies, or competing goals to produce a novel, and potentially undesirable, outcome.

Without an appreciation for the complexity of the context in which one hopes to apply promising principles, efforts toward the classroom can yield no improvement or, as can often be the case, lead to practices that subvert learning. In fact, many of the pedagogical features in college-level textbooks are based upon findings from the lab. Yet, the use of these strategies can yield negative correlations with learning as typically used outside of the lab (Gurung, 2003, 2004). Consider signaling devices in textbooks such as headings and bold-words: while they enable readers to prioritize and organize the content in structured settings (Lorch, 1989), they also encourage readers to skip or skim the unsignalled material (Nevid & Lampmann, 2003) in typical classroom use, resulting in poorer exam performance (Gurung, 2003, 2004). If we are truly interested in pursuing educational implications, it is important to understand that what works in a controlled context may have a very different effect when applied without the support inherent to high-quality curricular design. Once achieved, positive results under laboratory conditions must also be translated to useable techniques for educators to insure fidelity of the treatment. The field has to engage in more targeted adaptation before recommending classroom applications.

4. A call to action

I fully appreciate the enthusiasm associated with robust findings and hope to encourage learning scientists to consider productive educational applications. However, this enthusiasm must be tempered with an appreciation for the strengths, challenges and affordances associated with more and less complex contexts. The jump from the lab to the classroom is challenging and involves much more than simply explaining findings and main effects. Scientists ask different questions in a different way than do educators, work at different levels of analysis, and often rely upon different outcome measures. Partnering with educators to understand the demands of their context, systematically addressing levels of complexity and attending to curricular design are essential if we are to successfully address the challenges of education with the science of learning. Roediger and Pyc (2012) have made a provocative argument that may be best realized with collaborative translational research to responsibly guide the integration of promising scientific findings to effective educational practice.

References

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How People Learn*. Washington, DC: National Academy Press.

Bucuvalas, A. (2003). *HGSE News. Personal interview with Catherine Snow*. Retrieved http://www.gse.harvard.edu/news/features/snow06012003.html

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, *21*(279), 279–283.

Daniel, D. B., & Poole, D. A. (2009). The ecology of pedagogy: How collaborative research can prevent us from harming students. *Perspectives on Psychological Science*, *4*(1), 91–96.

Daniel, D. B., & Willingham, D. T. (2012). Electronic textbooks: Why the rush? *Science*, *333*(6076), 1569–1571.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, in press.

Gurung, R. A. R. (2003). Pedagogical aids and student performance. *Teaching of Psychology*, *30*, 92–95.

Gurung, R. A. R. (2004). Pedagogical aids: Learning enhancers or dangerous detours? *Teaching of Psychology*, *31*, 164–166.

Halpern, D., & Hakel, M. (Eds.). (2002). *New directions in teaching and learning: Using the principles of cognitive psychology as a pedagogy for higher education*. San Francisco: Jossey-Bass.

Halpern, D. F., & Hakel, M. D. (2003). Applying the Science of Learning to the University and Beyond: Teaching for Long-term Retention and Transfer. Change, July/August 2–13.

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, *1*, 209–234.

Matlin, M. (2007). How cognitive psychology can enhance your students' learning. In S. A. Meyers & J. R. Stowell (Eds.), Essays from E-xcellence in Teaching (Chapter 9), Vol. 7. Retrieved http://teachpsych.org/resources/ebooks/eit2007/eit2007.php

Nevid, J. S., & Lampmann, J. L. (2003). Effects on content acquisition of signaling key concepts in text material. *Teaching of Psychology*, *30*, 227–230.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., et al. (2007). *Organizing instruction and study to improve student learning*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ncer.ed.gov

Roediger, & Pyc. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, *1*, 242–248.

Worrell, F. C., Casad, B. J., Daniel, D. B., McDaniel, M., Messer, W. S., Miller, H. L., Jr., et al. (2009). Promising principles for translating psychological science into teaching and learning. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (expanded ed., pp. 129–144). Washington, DC: American Psychological Association.

Teaching of Psychology, 37: 199–203, 2010

ISSN: 0098-6283 print / 1532-8023 online
DOI: 10.1080/00986283.2010.488542

Routledge
Taylor & Francis Group

TECHNOLOGY AND TEACHING

They Hear, But Do Not Listen: Retention for Podcasted Material in a Classroom Context

David B. Daniel
James Madison University

William Douglas Woody
University of Northern Colorado

This study examined the retention of students who listened to podcasts of a primary source to the retention of students who read the source as text. We also assessed students' preferences and study habits. Quiz scores revealed that the podcast group performed more poorly than did students who read the text. Although students initially preferred podcasts, their preferences changed immediately after the quiz. Podcasts might be a useful tool to supplement or enrich course-related material, but they are not as effective as text for delivering primary content.

The popularity of portable MP3 players as tools for students to conveniently listen to or view course lectures and content has been increasing in higher education. Audio-podcasting, akin to creating files to allow a student to listen to a lecture or reading on a tape-recorder, has become very popular, with several publishers offering content in this format. Although a number of papers report provocative uses and student enthusiasm for podcasted material (e.g., Campbell, 2005; Evans, 2008; Rosell-Aguilar, 2007) scant evidence exists with regard to actual, rather than perceived, learning impact.

Similar to many other instances of technology, podcasts have been lauded for the student enthusiasm they generate and for the convenience; students can listen to them anywhere and anytime. In an often-cited report, Duke University distributed iPods to first-year students and evaluated student use and satisfaction via surveys and focus groups (Duke University, Office of Information Technology, 2005). Findings included significant student enthusiasm for the concept and perceptions of positive learning impacts. Scholars have reasoned that students' enjoyment will correlate with positive learning outcomes, as students might engage in an activity more often and more thoroughly if they enjoy it. These findings are in line with large student surveys regarding perceived benefits of technology. In a survey of students at member institutions, Educause, a group that advocates and documents technological innovation in higher education, found that students rated convenience, not learning, as the number one benefit of instructional technology (Kravik, Caruso, & Morgan, 2004). In fact, students rated convenience and time savings more than five times higher than learning benefits. It is possible that student enthusiasm for podcasts might be based on the convenience, enjoyment, or perceived learning rather than actual learning impact.

Despite students' beliefs that podcasts are effective learning tools, how well should instructors expect students to learn material presented only in an audio format? During the 1980s, cognitive psychologists studied participants' recall after reading or listening to text. Across several studies, reading text led to better recall than listening to text (Dixon, Simon, Nowak, & Hultsch, 1982; Green, 1981; Hildyard & Olson, 1982). Scholars have reported contrary findings (e.g.,

Sannomiya, 1982, 1984), but other researchers have raised methodology questions regarding these equivocal results (see Rickheit, Strohner, Müsseler, & Nattkemper, 1987). This body of basic research raises applied questions about the use of podcasts to present primary course content, but other advantages, such as convenience, accessibility, and enjoyment, might offset these concerns.

Previous research argued that students are not very good judges of their own learning (see Dunning, Johnson, Ehrlinger, & Kruger, 2003, for a review). In a study of students' perceived learning, use of pedagogical aids, and actual performance, Gurung and Daniel (2005) summarized reports of negligible to negative correlations between student use of such aids and student exam scores, despite positive student perceptions of the learning impact of these tools (see also Gurung, 2003, 2004). Clearly, initial student preference and self-report of learning are not the best indicators of student learning.

Students like the idea of podcasts, but do they learn primary content as well from listening to it as they do from reading it? This issue becomes more important as higher education begins to explore the possibility of audio text supplements to deliver a course's primary content and vocabulary. This study investigated student preference and performance on podcasted versus text-based primary content to begin to provide instructors as well as publishers with guidelines and challenges for audio-podcast use. Additionally, the students who heard podcasts participated in a focus group in which they provided feedback about their learning experiences with the podcasts.

Method

Participants

Participants were 48 students (12 men, 36 women) in a developmental psychology course at a medium-sized regional university who participated as part of a course requirement. We treated all participants in accordance with American Psychological Association (APA) ethical guidelines (APA, 2002).

Materials and Procedure

We randomly assigned students to either read the 3,330-word article or listen to a 21 min, 42 sec podcast of "Mindful of Symbols" by DeLoache (2005) in preparation for a quiz. After 2 days of time to read or listen to the article, all students used a 9-point Likert scale (1 = *not at all*, 9 = *extremely*) to complete prequiz measures of their perceived knowledge and understanding of the material for the quiz, the difficulty of the material, how much they learned, and how much they enjoyed the reading or podcast. Participants also reported the amount of time they spent studying, their activities concurrent with studying (e.g., walking while listening to the podcast), the location of studying, and competing activities they performed (e.g., talking on the phone). Finally, they used the 9-point Likert scale to report the degree to which they would prefer a podcast over reading to learn important material.

All participants then completed a 10-question multiple-choice quiz about the article (e.g., "According to the author, what is the first type of symbolism that infants and young children master?"). After completion of the quiz, participants answered the last Likert scale question again.

In addition to the quantitative data collection previously described, the 23 students who heard the podcast participated in a focus group discussion immediately following the quiz and provided feedback regarding their perceptions of positive and negative aspects of podcasts as primary learning tools.

Results and Discussion

Experimental Data

Data included quiz scores, responses on the prequiz measure, and responses to the postquiz question. The 25 participants who read the article scored higher on the quiz ($M = 8.16$, $SD = 1.11$) than the 23 participants who heard the podcast ($M = 5.91$, $SD = 1.56$), $t(46) = 5.78$, $p < .001$, $d = 1.70$. Despite claims from many, these results reflect the basic research from the 1980s (e.g., Dixon et al., 1982; Green, 1981; Hildyard & Olson, 1982) and suggest that podcasts do not deliver primary content as well as textbooks. Students remember primary content better when they read instead of listen to it.

To explore students' perceptions of their learning, we used a MANOVA with condition as a between-participants independent variable and students' self-reports of their knowledge, their comprehension, the difficulty of the material, and the amount they learned from the text or podcast as dependent variables. There was a multivariate main effect for condition, Wilks's

Table 1. Means, Standard Deviations, and Univariate Statistics for Participants' Perceptions of Learning as a Function of Condition

Measure	Text		Podcast		Univariate Results
	M	*SD*	*M*	*SD*	
Knew (remembered) material	6.56	0.96	5.52	1.73	$F(1, 46) = 6.76$, $p < .05$, partial $\eta^2 = .13$
Understood (comprehended) material	7.08	0.91	6.04	1.94	$F(1, 46) = 5.76$, $p < .05$, partial $\eta^2 = .11$
Difficulty level of material	4.00	1.29	5.04	1.40	$F(1, 46) = 7.23$, $p < .05$, partial $\eta^2 = .14$
Learn from text/podcast	6.40	0.96	5.78	1.51	$F(1, 46) = 2.92$, $p < .10$, partial $\eta^2 = .06$

Lambda, $F(4, 43) = 3.40$, $p < .05$, $\eta_p^2 = .24$. As shown in Table 1, univariate results revealed that students who read the article reported that they knew more, understood more, had less difficulty, and, marginally, learned more than did students who heard the podcast. Despite these differences, students did not report spending different amounts of time reading the text (M = 25 min, 2.40 sec, *SD* = 13 min, 15.70 sec) and listening to the podcast (M = 24 min, 20.87 sec, *SD* = 7 min, 4.88 sec), $t(46) = .22$, $p = .83$, $d = .06$, and students similarly enjoyed reading the text (M = 5.60, *SD* = 1.47) and listening to the podcast (M = 5.48, *SD* = 2.04). The difference in enjoyment was not significant, $t(46) = .24$, $p = .81$, $d = .07$. Despite claims of greater flexibility for podcasts and student preference for them when asked, students did not spend different amounts of time interacting with podcasts and text, and they did not rate them as differentially enjoyable when asked after actually interacting with the material in both media. Even with these similarities in time and enjoyment, the quiz performance difference remained.

Students reported other activities they did while they were reading and the locations for reading. The choices of activities included walking, sitting, working out, driving, doing chores, and other. Although students did not differ across all categories of reported activities for reading or listening locations as a function of condition, $\chi^2(4, N = 48) = 6.38$, *ns*, 88% of participants who read the text reported sitting, and 60.9% participants who listened to the podcast reported sitting, $\chi^2(1, N = 48) = 4.70$, $p < .05$, $\varphi = .31$. Students listening to podcasts were less likely to sit and study than were students who read the material. There were no significant differences in the locations in which students read or listened to the material. Students who listened to podcasts were not more likely to take advantage of the potential for flexibility for study locations and activities provided by podcasts.

Table 2 shows the number and percentage of participants who reported doing each of the other competing activities (e.g., talking on the phone, doing other computer activities) while they were reading or listening to the article. Students who read the article reported engaging in a mean total of 3.36 (*SD* = 2.78) competing activities, and students who listened to the podcast reported engaging in a mean total of 2.48 (*SD* = 2.21) competing activities; this difference was not significant, $t(46) = 1.20$, $p = .23$, $d = .35$. As shown in Table 2, students in the podcast condition were more likely than students in the text condition to report doing other computer activities, $\chi^2(1, N = 48) = 4.17$, $p < .05$, $\varphi = .30$. Although students in the text condition reported a greater total of noncomputer activities such as talking on the phone, watching television, or having people present (M = 3.12, *SD* = 2.59) than

Table 2. Number and Percentage of Participants Who Self-Reported Competing Activities as a Function of Condition

Activity	Text		Podcast	
	n	%	*n*	%
Television	5	20.0	1	4.3
Music	9	36.0	3	13
Roommates/friends present	13	52.0	8	34.8
Unknown people present	6	24.0	3	13
Both friends and unknown people present	6	24.0	2	8.7
Respond to instant messaging/e-mail via the Internet	8	32.0	5	21.7
Facebook/MySpace	6	24.0	6	26.1
Other computer activities	5	20.0	11	47.8*
Text message	10	40.0	5	21.7
Answer phone	7	28.0	5	21.7
Talk on phone	5	20.0	3	13.0
Other	4	16.0	8	34.8
Total	84		61	

Note. "Other" responses included "get ready," "distracted," "interrupted," "other reading," "sat with significant other," "cleaned room," and "showered."

*$p < .05$, $\varphi = .30$.

students in the podcast condition reported ($M = 2.04$, $SD = 1.89$), this difference was not significant, $t(46) = 1.63$, $p = .11$, $d = .47$. Students who listen to podcasts on computers do more computer activities, perhaps due to the distractions inherent in the wide variety of easily accessible computer activities. Any technological device that plays podcasts (e.g., computers, iPods, or cellular phones) might have other features that could be more interesting and distracting than the class material in the podcast.

Both before and after the quiz, participants reported the degree to which they would prefer a podcast over reading to learn important material. To evaluate responses, we used a repeated measures ANOVA with condition as a between-participants independent variable and the timing of the quiz (pre and post) as a within-participants variable. There was a main effect for condition; students preferred text ($M = 6.02$, $SD = 1.25$) over podcasts ($M = 4.41$, $SD = 1.88$), $F(1, 46) = 12.31$, $p < .01$, $\eta_p^2 = .21$. There was also a main effect for time; scores decreased between the pretest ($M = 5.73$, $SD = 2.09$) and the posttest ($M = 4.77$, $SD = 1.87$), Wilks's Lambda, $F(1, 46) = 19.98$, $p < .001$, $\eta_p^2 = .30$. More importantly, however, a significant interaction existed between condition and time, Wilks's Lambda, $F(1, 46) = 18.41$, $p < .001$, $\eta_p^2 = .29$. We performed simple contrasts to investigate the interaction. The difference between the pretest scores ($M = 6.04$, $SD = 1.46$) and posttest scores ($M = 6.00$, $SD = 1.38$) for the text group was not significant, $t(46) = .15$, $p = .88$, $d = .04$. The difference between pretest ($M = 5.39$, $SD = 2.61$) and posttest ($M = 3.43$, $SD = 1.34$) for the podcast group was significant, $t(46) = 5.38$, $p < .001$, $d = 1.55$. Although students who listened to podcasts preferred podcasts before the quiz, after the quiz the preference for podcasts decreased. Although students did not immediately learn about their performance on the quiz (i.e., their grades), merely taking the test alerted them to the limits in their comprehension after listening, and the significant change in the preferences for the podcast group reflected this realization.

Student Focus Group Outcomes

In a focus group following the quiz, we asked the 23 students in the podcast group how the podcasts could be made more valuable. Several issues emerged that might be helpful for future investigation as well as podcast development. Students reached a near unanimous consensus on five points: (a) The lack of signaling devices (e.g., bold words, italics) in the podcasts made it difficult to prioritize the reading and focus on the important points; (2) podcasts lack visuals such as charts and graphs that reinforce the reading; (3) the students were much less likely to review sections of the podcast than they would have been when reading it; (4) the more the voiceover in the podcast sounded like a professional reader (e.g., not casual and conversational), the less enjoyable the podcast; and (5) it was easier to listen to the podcast on the computer than to go through the trouble of downloading it to an MP3 player. Although the last suggestion might explain the results that the podcast and text groups did not differ much in where they interacted with the material, the suggestion that the podcasts might be more effective if the learners were also supplied visual support (e,g., signaled text and supporting graphs) was not tested in this study.

Conclusions

The results argue for caution when relying on audio podcasts to deliver primary course content. Students in the podcast group performed relatively poorly on the quiz and reported that they knew less, understood less, experienced more difficulty with the material, and, marginally, learned less than did students in the text condition. Despite the popular claims that podcasts allow for more flexibility of use, efficiency, and enjoyment (e.g., Campbell, 2005; Duke University, Office of Information Technology, 2005) and despite a lower likelihood of sitting to study, the students in this sample did not differ in where they interacted with the material, how long they studied, or how much they enjoyed the content.

As expected, students who listened to podcasts initially preferred podcasts as learning tools in this study. This finding joins an ever growing list of student preferences for pedagogy and techniques that do not positively affect their actual, as opposed to perceived, learning (e.g., Gurung & Daniel, 2005; Wesp & Miele, 2008). Student perception of learning is seldom a reliable basis for performance-based measures (e.g., Dunning et al., 2003). Interestingly, directly after taking the quiz, even without formal feedback regarding their performance, students in the podcast group realized that the podcasts were not effective tools for their learning and performance. It is possible that students are more likely to gauge the effectiveness of very poor strategies if they have experience using them and reflecting on them in an evaluated context.

The findings reported here suggest that audio podcasts are not effective learning tools for the mastery of primary course content, such as vocabulary and core concepts. The use of audio podcasts remains untested for delivering secondary content that reinforces, extends, and contextualizes the primary concepts of a course or concept. Indeed, enriching primary content in this manner might be the ideal use for audio podcasts.

References

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57,* 1060–1074.

Campbell, G. (2005). There's something in the air: Podcasting in education. *Educause Review, 40,* 32–47.

DeLoache, J. S. (2006). Mindful of symbols. *Scientific American Mind, 17*(1), 71–75.

Dixon, R. A., Simon, E. W., Nowak, C. A., & Hultsch, D. F. (1982). Text recall in adulthood as a function of level of information, input modality, and delay interval. *Journal of Gerontology, 37,* 358–364.

Duke University, Office of Information Technology. (2005). *Duke University iPod first year experience* (Final Evaluation Rep.). Retrieved November 9, 2008, from http://cit.duke.edu/pdf/reports/ipod_initiative_04_05.pdf

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12,* 83–87.

Evans, C. (2008). The effectiveness of m-learning in the form of podcast revision lectures in higher education. *Computers & Education, 50,* 491–498.

Green, R. (1981). Remembering ideas from text: The effect of modality of presentation. *British Journal of Educational Psychology, 51,* 83–89.

Gurung, R. A. R. (2003). Pedagogical aids and student performance. *Teaching of Psychology, 30,* 92–95.

Gurung, R. A. R. (2004). Pedagogical aids: Learning enhancers or dangerous detours? *Teaching of Psychology, 31,* 164–166.

Gurung, R. A. R., & Daniel, D. B. (2005). Do text-based pedagogical features enhance student learning? In D. S. Dunn & S. L. Chew (Eds.), *Best practices for teaching introduction to psychology* (pp. 41–55). Mahwah, NJ: Lawrence Erlbaum Associates.

Hildyard, A., & Olson, D. R. (1982). On the comprehension and memory of oral v. written discourse. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 19–33). Norwood, NJ: Ablex.

Kravik, R. B., Caruso, J. B., & Morgan, G. (2004). *ECAR study of students and information technology, 2004: Convenience, connection, and control roadmap.* Retrieved November 9, 2008, from http://www.educause.edu/ECAR/ECARStudyofStudentsandInformat/155030

Rickheit, G., Strohner, H., Müsseler, J., & Nattkemper, D. (1987). Recalling oral and written discourse. *Journal of Educational Psychology, 79,* 438–444.

Rosell-Aguilar, F. (2007). Top of the pods—In search of a podcasting "pedagogy" for language learning. *Computer Assisted Language Learning, 20,* 471–492.

Sannomiya, M. (1982). The effect of presentation modality on text memory as a function of difficulty level. *Japanese Journal of Psychonomic Science, 1,* 85–90.

Sannomiya, M. (1984). The modality effect on text processing as a function of presentation rate. *Japanese Journal of Psychonomic Science, 3,* 17–20.

Wesp, R., & Miele, J. (2008). Student opinion of the quality of teaching activities poorly predicts pedagogical effectiveness. *Teaching of Psychology, 35,* 360–362.

Notes

1. We would like to thank Worth Publishers for providing the podcasts and article used in this study.

2. Send correspondence to William Douglas Woody, School of Psychological Sciences, University of Northern Colorado, Greeley, CO 80639; e-mail: william.woody@unco.edu

Using Web-Based Quizzing to Improve Exam Performance: Lessons Learned

David B. Daniel
University of Maine at Farmington

John Broida
University of Southern Maine

This study examined the utility of Web-based quizzing. We assigned 3 classes to a no-quiz, in-class quiz, or Web-based quiz condition. Midsemester results demonstrated a positive effect for in-class quizzing but not Web-based quizzing. After several adjustments in quiz presentation and duration, the Web-based group increased exam performance to a level equivalent to the in-class quiz group for the second half of the semester. These results illustrate that online quizzing can be as effective as in-class quizzing, but only under specific conditions.

A number of studies have provided evidence that routine quizzing increases student performance on exams (e.g., Connor-Greene, 2000; Grover, Becker, & Davis, 1989; Taraban, Maki, & Rynearson, 1999). However, quizzing often consumes valuable class time and requires that someone grade and record student performance. The increased effort required from the instructor diminishes the practical utility of quizzing as class size increases. Thus, quizzes are not often used in courses with large enrollment where the positive impact in exam performance may be most appreciated.

Alternatively, Web-based quizzing outside of class may be a useful tool for incorporating the benefits of quizzing without sacrificing valuable class time (Brothen & Wambach, 2001). An additional benefit of this system is that the quizzes can be automatically scored and recorded for the instructor, reducing the burden placed on the instructor as class size increases.

Many faculty, however, express concerns that computer-based teaching tools may not be as effective as traditional teaching strategies (Brewster, 1996). We examined the utility of Web-based quizzing and its effects on multiple choice exam scores in a moderately sized psychology course.

Method

Participants

One hundred and twenty-five students enrolled in three sections of Child and Adolescent Development at a public liberal arts university in New England participated. We assigned each of the three sections to a no-quiz ($n = 44$), an in-class quiz ($n = 42$), or a Web-based quiz ($n = 39$) condition.

Procedure

The in-class quiz group received 16 weekly chapter-based quizzes in the first 15 min of class. Students in the Web-based quiz group received the same quizzes available for self-administration on the Web 24 hr preceding class, also allowing 15 min for completion. The outcome measures were four exams. For the first half of the semester, we presented identical quizzes to both quiz groups, and all three groups completed the same exams. In addition, the same lecturer taught all classes using the same notes, with the lectures lasting approximately the same amount of time in each class (60 min twice weekly).

Results 1

As seen in Table 1, although in-class quizzing seemed to demonstrate a positive impact on exam scores at midsemester, a cursory view of exam performance did not yield an obvious impact of Web-based quizzing on exam scores when compared to the no-quiz group. We computed the mean of Exams 1 and 2 for each group for the analysis. A one-way ANOVA revealed significant effects for quiz condition $F(2, 122) = 46.69$, $p < .01$, $\eta^2 = .43$. Bonferroni post hoc tests indicated significant differences between the in-class quiz group and both the Web-based quiz group, $t(79) = 7.46$, $p < .001$, and the no-quiz group, $t(84) = 9.38$, $p < .001$. There were no significant differences between Web-based quizzing and the no-quiz groups, $t(81) = .64$, $p > .05$. Consequently, we decided to pursue possible explanations to account for the lack of impact for Web-based quizzing on exam scores.

A Midsemester Correction

On the second exam, students in both quiz groups anonymously described common methods to "cheat" on the quizzes. Students in the Web-based group reported a number of strategies for cheating; most prominently reported were printing and sharing of quizzes, looking up answers in the book during the quiz, using an online glossary opened in a window adjacent to the quiz, and working in groups.

Table 1. Overall Results: Group by Mean Exam Score

	Exams 1 and 2[a]		Exams 3 and 4[a]	
Group	*M*	*SD*	*M*	*SD*
No quiz	48.95	6.18	46.45	6.57
Web-based quiz	49.75	4.89	60.90	4.78
In-class quiz	59.45	5.37	61.60	6.71

[a]Out of 75.

As a result of this information, we changed the Web-based quiz parameters. In particular, we added an additional pool of 100 questions from the supplied testbank to the publisher-provided 10 question quizzes, enabling each student to receive a different random selection of questions on each quiz. Each student's quiz, therefore, consisted of different questions. In addition, we removed the glossary from the Web site and reduced the time allowed for each 10-item quiz from 15 to 7 min. With these changes, the students resumed the quizzing schedule outlined in Method 1.

Results 2

We computed the mean of Exams 3 and 4 for each group and performed a one-way ANOVA on the means. Results of this analysis revealed significant effects for quiz condition, $F(2, 122) = 81.70$, $p < .01$, $\eta^2 = .57$. Bonferroni post hoc tests indicated significant differences between the no-quiz group and both the Web-based quiz group, $t(81) = 11.47$, $p < .001$, and the in-class quiz group, $t(84) = 10.45$, $p < .001$. Unlike the results in the first half of the semester, there were no longer significant differences between Web-based quizzing and the in-class quiz groups, $t(79) = .54$, $p > .05$. In sum, the Web-based quiz group increased exam performance to a level equivalent to the in-class quiz group and significantly higher than that of the no-quiz group (see Table 1).

Discussion

Although quizzes are generally effective, Web-based quizzes do not always positively affect exam performance as compared to in-class quizzes (see also Brothen & Wambach, 2001). Students in the Web-based quiz group used strategies to optimize their quiz performance without mastering the text. Efforts to discourage these efforts by randomly assigning questions from a larger test bank and decreasing the amount of time allowed for the quiz were effective.

Unfortunately, much of the Web content currently available from publishers in various formats is of the type that does not discourage such activities by students: Every student receives the same questions in the identical order with no time limits imposed. Furthermore, not all Web-based platforms allow for the corrections evaluated in this study (see Brooks, 2001). We argue that these, and possibly other, adaptations of publisher-provided content are necessary to fully obtain benefits similar to in-class quizzing with Web-based products.

Note that the present results do not indicate that online quizzing has any benefits for the student beyond in-class quizzing. Web-based quizzes with immediate feedback did not improve performance more than in-class quizzing. However, the advantage for the instructor, in addition to possibly increasing student exam scores, is the availability of the class time required to administer in-class quizzes as well as the time needed to grade and record the quizzes.

Our study specifically targeted a larger survey-type course where multiple-choice and short answer exams are used as one form of evaluation. Although our data support the use of quizzing to increase success on objective tests, further study will be needed to evaluate the potential effects of Web-based interactions on other outcome measures (i.e., essay exams and papers).

References

Brewster, J. (1996). Teaching abnormal psychology in a multimedia classroom. *Teaching of Psychology, 23,* 249–252.

Brooks, S. (2001). *Comparing course offering software: WebCT vs. Blackboard.* Retrieved October 21, 2003, from www.edb.utexas.edu/multimedia/PDFfolder/WebCTvsBlackboard.pdf

Brothen, T., & Wambach, C. (2001). Effective student use of computerized quizzes. *Teaching of Psychology, 28,* 292–294.

Connor-Greene, P. A. (2000). Assessing and promoting student learning: Blurring the line between teaching and testing. *Teaching of Psychology, 27,* 84–88.

Grover, C. A., Becker, A. H., & Davis, S. F. (1989). Chapters and units: Frequent versus infrequent testing revisited. *Teaching of Psychology, 16,* 192–194.

Taraban, R., Maki, W. S., & Rynearson, K. (1999). Measuring study time distributions: Implications for designing computer-based courses. *Behavior Research Methods, Instruments & Computers, 31,* 263–269.

Note

Please send correspondence to David B. Daniel, Department of Psychology, University of Maine at Farmington, 234 Main Street, Farmington, ME 04938, e-mail: dbdaniel@maine.edu.