Improving teacher candidate knowledge using content acquisition podcasts

Michael J. Kennedy a,*, Cathy Newman Thomas b, Sara Aronin c, Jennifer R. Newton d, John Wills Lloyd e

a University of Virginia, Curry School of Education, Department of Curriculum, Instruction, and Special Education, Bavaro Hall Room 327, PO Box 400273, Charlottesville, VA 22904, United States
b University of Missouri, United States
c West Virginia University, United States
d James Madison University, United States

doi:10.1016/j.compedu.2013.08.010

Abstract

There are many open questions with respect to theory and empirical support for methods used in college teaching, especially when technology is incorporated into instruction. In this article, we report the results of a study of a multimedia-based instructional tool called Content Acquisition Podcasts (CAPs) that provides university instructors with a tool that is grounded in applied theory and has advanced through several iterations of developmental and experimental testing as suggested by Clark (2009). CAPs are a form of enhanced podcasts (still images synchronized with audio) that incorporate Mayer’s cognitive theory of multimedia learning (2009), and accompanying instructional design principles (2008) to ensure the looks and sounds of instruction do not overwhelm the limitations of users’ cognitive processes. This article reports data from one of the first five experimental tests of CAPs in which undergraduate teacher candidates received instruction related to an introductory course in special education. In this study, teacher candidates from two universities were randomly assigned either to watch a CAP or read a textbook chapter containing the same content for two topics: characteristics of students with learning disabilities or high-functioning autism. We employed a pretest-posttest-maintenance design to evaluate participant performance on dependent measures of knowledge. Results indicate that when participants learned with CAPs, they had significantly higher scores on content-knowledge tests at both posttest and maintenance assessments than when they studied via the usual text-based materials.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Pause for a moment, and think back to the prevailing methods used by professors to teach (especially large enrollment) courses you took as an undergraduate and graduate student—and then, if appropriate, reflect on your “go to” approach for teaching today. Now, please answer the following three questions based on your reflection. Question one: Did you envision a professor standing at the front of the classroom for extended periods of time delivering a lecture based on text-heavy PowerPoint slides (or overhead transparencies, depending on your age) and asking the occasional question? Question two: To what extent do the methods used by your professors constitute a mirror image of how you teach today? And finally, depending on your answers to questions one and two: Can you specify a theoretical framework or evidence base for the instructional methods that were, or are being used? If the spirit of your answers is: Yes-Yes-No, you are not alone. These questions are not intended to be rhetorical, or imply criticism of our professors, colleagues, and selves. Instead, our goal is to maintain consciousness of an issue that, for many, has become an unconscious aspect of daily practice—the atheoretical nature of most college teaching (Barr & Tagg, 1995; D’Avanzo, 2003; Handelsman et al., 2004); and, specific to the purpose of this article, the relatively unchecked expansion of technology into this domain (Clark, 1994, 2009). This is a critical dialog, as unproven applications of technology are infused into already questionable instructional practices on a massive scale (Clark & Estes, 2008; Clark & Mayer, 2011; Hew & Cheung, 2013) and show no signs of abating (Atkins et al., 2010). This is problematic, particularly in content areas such as teacher preparation.

* Corresponding author. Tel: +1 434 924 0827; fax: +1 434 924 7461.
E-mail addresses: MKennedy@Virginia.edu (M.J. Kennedy), ThomasCat@Missouri.edu (C.N. Thomas), Sara.Aronin@mail.wvu.edu (S. Aronin), NewtonJr@jmu.edu (J.R. Newton), JohnL@Virginia.edu (J.W. Lloyd).
The purpose of this article is to extend a program of research regarding a multimedia-based instructional tool called Content Acquisition Podcasts (CAPs). CAPs are an instructional approach that addresses the need to use theory when designing and delivering instruction to students, including undergraduates and graduates. CAPs are multimedia-based instructional vignettes created using Mayer’s cognitive theory of multimedia learning (Mayer, 2009) and accompanying instructional design principles (2008; see Fig. 1). The CTML is based on cognitive load theory (Chandler & Sweller, 1991) and the dual processing principle (Paivio, 1986). This applied theory provides guidance to instructional designers seeking to create multimedia-based materials that are a good match for how people learn. A sample CAP can be seen at https://vimeo.com/40105175. Although CAPs are undergirded by a well-known, and oft-tested theory, this tool is not useful without the inclusion of content that lends itself to this delivery format. Therefore, we use Clark’s (2009) framework for guiding research and development processes of instructional designers.

1.1. Need for theory and empirical validation in higher education instruction

The identification of theory alone is insufficient for selecting or developing multimedia-based instructional applications that are to be used in unique learning environments (Clark & Estes, 2008; Clark & Mayer, 2011; Mayer, 2011). Clark (2009) holds that an instructor needs to be able to articulate clearly his or her instructional method and theoretical framework and to define how various elements of his or her pedagogy fits within the overall learning goals of the course and curriculum. Although evaluating one’s teaching given these critical elements is easier said than done, Clark presents a four-part instructional design framework to support instructional designers and instructors in creating and evaluating multimedia-based pedagogies, practices, and tools. We selected this framework to guide our program of research in higher education pedagogy given its iterative research and development cycle, and corresponding emphasis on creating multimedia that is both powerful and relevant for the end user. In the case of the present study, the end user is undergraduate teacher candidates enrolled in a required introductory course on special education.

Clark’s (2009) framework calls for programs of research that (a) use descriptive research methods to identify specific theory (or theories) that may be helpful in addressing specific problems of practice; (b) undertake a program of predictive research, which includes the traditional cycle of research and development for validating a new theory-driven intervention; (c) analyze the generic features of the technology being developed and tested, which includes careful examination and corresponding description of how and why the active

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Limit Extraneous Processing</strong></td>
<td>Coherence Principle &lt;i&gt;ES = .97, 14 Studies&lt;/i&gt;</td>
<td>Each CAP only contains information relevant to the history term/concept being presented</td>
</tr>
<tr>
<td></td>
<td>Signaling Principle &lt;i&gt;ES = .32, 6 Studies&lt;/i&gt;</td>
<td>Each CAP contains recurring explicit cues to signal the beginning of a new section</td>
</tr>
<tr>
<td></td>
<td>Redundancy Principle &lt;i&gt;ES = .72, 5 Studies&lt;/i&gt;</td>
<td>Each CAP only contains carefully selected key text</td>
</tr>
<tr>
<td></td>
<td>Spatial Contiguity Principle, &lt;i&gt;ES = .112, 5 Studies&lt;/i&gt;</td>
<td>The on-screen text and pictures in each CAP are presented in close proximity to one another</td>
</tr>
<tr>
<td></td>
<td>Temporal Contiguity Principle &lt;i&gt;ES = .31, 8 Studies&lt;/i&gt;</td>
<td>Pictures and text within each CAP correspond to the audio presentation</td>
</tr>
</tbody>
</table>

| Manage Essential Processing                              | Modality Principle <i>ES = .102, 17 Studies</i>                         | CAPs are multimedia; therefore this principle is addressed |
|                                                           | Segmenting Principle <i>ES = .98, 3 Studies</i>                         | Each CAP is broken either into segments as noted by instructions to pause the video, or main ideas are separated into separate CAPs |
|                                                           | Pretraining Principle <i>ES = .85, 5 Studies</i>                        | Each CAP begins with an explicit statement of purpose and an advance organizer for the term |

| Foster Generative Processing                             | Multimedia Principle <i>ES = .139, 11 Studies</i>                     | The CAPs are multimedia; therefore this principle is addressed |
|                                                           | Personalization, Voice, and Image Principles <i>ES = .111, 11 Studies</i> | The narration in each CAP is presented in a conversational style and by a human voice. The speaker’s image is not on the screen. Images are non-abstract and easily recognizable by viewers |

*Note: Effect sizes are summaries of empirical research conducted by Mayer and his colleagues and are reported in Mayer (2008).
elements of the theory and intervention function as they do with respect to user performance; and finally (d) investigate the new intervention in situated settings to evaluate the extent to which the theory and active elements within the tool can support instruction and learning in unique settings. This framework provides scaffolding for the development and experimental testing of the program of research associated with the intervention described in this article; content acquisition podcasts (CAPs).

1.2. Conducting descriptive research

A key first step in Clark's (2009) framework is to identify an existing problem of practice and then recognize the need for a new technology or instructional approach as opposed to the common situation where researchers, technology developers, or instructors develop new multimedia-based tools and retrofit them into instruction. The setting for the problem of practice addressed in this study is a course required for licensure in almost all teacher preparation programs, Teaching Students with Disabilities in the General Education Classroom. The specific and observable problem addressed by this program of research is the phenomena of limited face-to-face instructional time to provide teacher candidates with all of the content knowledge and pedagogical practices needed to address the learning needs of students with exceptionalities in the general education classroom in a traditional 3-credit hour class (Reschly, Holdheide, Behrstock, & Weber, 2009). Success in dealing with this problem in the short term entails measuring and confirming student knowledge with respect to static content knowledge (e.g., characteristics of students with specific disabilities, knowledge of special education law), and for the long term, evaluating performance (e.g., implementation of evidence-based practices) during practicum and early career teaching experiences. Creating reliable and valid measures for evaluating students’ static knowledge in this course is not difficult, assuming the researcher uses best practices for item construction and psychometric testing (e.g., American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Haladyna, 2004).

1.2.1. Use of podcasts to deliver instruction

Several researchers have brainstormed strategies for conserving face-to-face instructional time by way of reducing the duration of lectures, and freeing up in-class time to include hands-on learning experiences, such as case studies. One approach researchers and instructors are experimenting with is “flipped classroom” experiences that provide content outside of class time and are often asynchronous (Bull, Ferster, & Kjellstrom, 2012). Despite its popularity and recent coverage in the media, there is limited empirical support for this approach. Another approach for delivering core content asynchronously is podcasting (Evans, 2008). Many professors in higher education record their lectures (audio podcast), or sync their PowerPoint slides to the recording (enhanced podcast, using apps such as Tegrity) and upload it to a course management site, or an online repository, such as iTunesU, for students to access on demand (Fernandez, Simo, & Sallan, 2009; Griffin, Mitchell, & Thompson, 2009; Lin, Zimmer, & Lee, 2013; Lonn & Teasley, 2009; O’Bannon, Lubke, Beard, & Britt, 2011). Although this process is widely employed in higher education, important questions remain regarding the effectiveness of the practice as an instructional tool (Heilesen, 2010; Hew & Cheung, 2013; Kay, 2012).

Research confirming user satisfaction and support (creator and consumer) for podcasts in higher education coursework is abundant (see Evans, 2008; Fernandez et al., 2009; Lin et al., 2013; O’Bannon et al., 2011; Wu et al., 2012). However, given criticisms of atheoretical technology development and use (e.g., Clark, 2009; Mayer, 2011), simply adopting a new instructional tool because it available and accessible and easy to use, create, and consume is not a compelling rationale to implement that tool. Thus, it is critical to examine the evidence-base for podcasting in higher education.

Heilesen (2010) found only 10 studies in the professional literature between the years of 2002–2009 that incorporated an empirical test of podcasts’ efficacy in learning within higher education coursework. Of those 10, only three used an experimental design with a control condition. Although the three higher quality studies found positive results for podcast use, this is too few to recommend wide adoption of this practice, although, it seems to have been widely adopted irrespective of the slim evidence base (e.g., Chester, Buntine, Hammond, & Atkinson, 2011). In addition, most studies in the podcasting literature fail to specify a theoretical or conceptual framework for explaining why or how podcasts have the potential to impact student learning.

Researchers are evaluating various multimedia-related theories to examine the extent to which any of them may provide an appropriate framework for rethinking use of podcasts in teacher preparation coursework to address this limitation within the current literature base (Kennedy, Driver, Pullen, Ely, & Cole, 2013). In the present program of research, we hypothesize that Mayer’s CTML (2009) provides a framework for creating multimedia-based enhanced podcasts that allow easy development and delivery of content, moreover, in a manner that results in improved learning outcomes.

1.2.2. Previous descriptive research

Kennedy, Newton, Haines, Walther-Thomas, and Kellems (2012) used a design experiment method (Kelly, 2004) to evaluate the extent to which teacher education students (n = 11) would respond favorably to the use of CAPs to learn essential course content during a semester. The researchers created a total of 13 CAPs and instructed students to watch them in addition to typical expectations for class preparation (e.g., read the textbook). Qualitative feedback from the students was positive regarding the practical nature of CAPs as an addition to the course textbook. Students also noted their preference for spending face-to-face instructional time working on case studies and other activities rather than monotonous lectures. In addition, students provided ideas for improving CAPs. Their feedback included calls for CAPs to include specific pause points so they could take notes, and also, in general, slow down the speed of the videos. Newer CAPs contain embedded pauses and a more deliberate pace, which addresses user feedback, and better adheres to Mayer’s (2008) signaling, segmenting and pre-training principles. Following reflection and revision using design methods, the next logical step in this program of research, and in Clark’s (2009) framework was to test CAPs experimentally.

1.3. Conducting predictive research

The second phase of Clark’s (2009) framework is too often the completion of the typical research and development cycle—conducting predictive research. A pilot study using an experimental design conducted by Kennedy, Hart, and Kellems (2011) explored the use of Mayer’s
CTML (2009) to design CAPs to deliver core course content to students. Undergraduate teacher candidates were randomly assigned to either an audio-only or a CAP condition where they received instruction related to traumatic brain injury (TBI) and the No Child Left Behind (NCLB, 2002) law. Results showed students in the CAP group significantly outperformed the audio-only group for both topics at posttest; effect sizes were medium to large (Cohen’s d = .64, .82, respectively). Although limited to two topics, this result provided rationale to conduct further studies of CAPs' utility to provide instruction for other course content and compare its effects with other types of learning.

In a follow-up study, Kennedy and Thomas (2012) randomly assigned preservice teachers from two large universities to either a CAP or text-based group to learn about School-Wide Positive Behavioral Interventions and Supports (SW-PBIS). On a researcher-developed test of knowledge, participants in the CAP group significantly outperformed peers who read a portion of a book chapter paired with a graphic organizer. Furthermore, effect sizes (Cohen’s d) were large (.98 and .97 at posttest and maintenance, respectively) in favor of the CAP group’s progress across the three time points. In a third experimental study, Kennedy, Ely, et al. (2012) randomly assigned 168 teacher candidates to (a) watch a CAP before reading a textbook chapter (Pre-CAP), (b) watch a CAP following reading (Re-CAP), or (c) only read. A pretest-posttest design was used to measure learning of two topics—characteristics of students with learning disabilities (LD) and students with high functioning autism (HFA). Students who watched CAPs in addition to reading significantly outperformed the Text-GO group, but there were not significant differences between the two CAP groups. Effect sizes for the Pre-CAP group were large for LD (d = 1.24) and medium for HFA (d = .63), and for the Re-CAP group, effect sizes were large for LD (d = .94) and for HFA (d = .94). In summary, instructional materials (CAPs) created with guidance from Mayer’s CTML (2009) and instructional design principles (2008) resulted in higher content knowledge for teacher candidates.

1.4. Determining generic technology

The third phase of Clark’s (2009) framework calls for careful examination of the active ingredients within new applications of technology. In practical terms, Clark calls on researchers to answer the question: How do you know which components of the new instructional tool ‘caused’ the effect you have observed and reported? To address this question, Mayer and his colleagues have published results of experiments that examine how various design principles impact student learning (see Mayer, 2009, 2011) and provide strong guidance for anyone seeking an applied method for creating multimedia learning tools. Therefore, as the active ingredients in CAPs are examined across these studies, Kennedy et al. (2011), Kennedy, Newton, et al. (2012), and Kennedy, Ely, et al. (2012) considered how each of Mayer’s instructional design principles has guided the looks and sounds of instruction. Because of the comprehensiveness of Mayer’s research, and the detailed empirical record supporting this framework, we consider Mayer’s existing research as a proxy for address this third phase of Clark’s framework. However, future research will aim to answer Clark’s question in a more precise fashion.

1.5. Testing situated technology

Finally, Clark (2009) recognizes technologies that have an observed effect in one content area or instructional purpose may not automatically transfer to other areas or domains. Therefore, it is important to conduct experimental tests of technology in various content areas, and for different instructional purposes. One strength of the present program of research is that Mayer’s CTML (2009) and instructional design principles (2008) are content-neutral. For example, Kennedy, Driver, et al. (2013) tested an application of Mayer’s CTML and instructional design principles to create a CAP that delivered instruction for phonological awareness (PA) knowledge. Undergraduates (n = 149) in a teacher preparation course were randomly assigned to either watch a CAP on PA, or read a user-friendly article containing the same content. A pretest–posttest–maintenance design was used; students who watched the CAP significantly outperformed the text-only group at posttest and maintenance (d = .86; .97). Kennedy and colleagues (Kennedy, Deshler, et al., 2013) also used CAPs to teach vocabulary terms to high school students with and without disabilities. Students with and without disabilities who learned using CAPs with embedded evidence-based practices for vocabulary learning outperformed others who learned with audio-only podcasts. These findings replicate the previously mentioned studies in teacher education, and provide evidence that this instructional model can be applied for varied instructional purposes and across different instructional settings. In the present study, CAPs that deliver instruction for teaching content related to the characteristics of students with LD and HFA are experimentally tested to provide additional empirical evidence for this learning tool.

1.6. Research questions

The present study sought to answer the following questions:

1. Do preservice general education teachers exposed to CAPs covering content about students with disabilities learn as well or better than peers exposed to the same content through chapter readings paired with graphic organizers (Text-GO)?
2. Is learning durable when measured by maintenance probes administered in controlled and naturalistic conditions?

2. Method

2.1. Participants and setting

The participants for this study were 164 students from two large Midwestern universities enrolled in coursework required to prepare preservice general education teachers for inclusive classrooms. These schools of education and universities are similar to other American flagship universities with respect to the structure of their teacher preparation programs and the overall characteristics of students. Seventy-eight students enrolled in three sections of the same course participated from University 1, and the remaining 86 participants were enrolled in one section of the course from University 2. Eighty-one percent of participants were female, and all participants were either in their third or fourth year of study within their respective teacher preparation programs. The courses at the two universities covered comparable content and met similar objectives for their respective state requirements for licensure of future teachers. Admission at both universities
required completion of at least two years of pre-education coursework with a GPA of 2.75 or higher. Researchers obtained human subjects permission from both universities prior to data collection. Participants consented to have their data included in published analyses and reports.

2.2. Procedures

Researchers pooled consenting participants into one large group, and then randomly assigned them to one of two experimental conditions. Students in Group 1 received instruction using CAPs (CAP Group), but students in Group 2 accessed content using chapters from relevant introductory textbooks, and also had access to a graphic organizer and outline of core content (Text-GO Group). The purpose of the GO was to provide a scaffold for students who would access content by reading, and the categories highlighted in this outline were identical to those emphasized in the CAP. The groups participated in two experiments. Experiment one investigated participant learning of the characteristics of students with learning disabilities (LD), and the second investigated learning of the characteristics of students with high functioning Autism (HFA). All participants completed a pretest, posttest, and maintenance probe of knowledge related to characteristics of students with LD and HFA. Researchers conducted all research activities during class time. Pretest data for LD and HFA were collected early in the semester and before content related to LD or HFA was introduced.

2.2.1. Text + graphic organizer (Text-GO) condition

The materials used for the Text-GO condition came from published textbooks. Researchers selected chapters written by respected experts in the fields of LD (Bryant, Smith, & Bryant, 2008) and HFA (Stichter, Conroy, & Kaufman, 2008). Graphic organizers and outlines of the chapter sections were created using Inspiration® software following chapter headings and subheadings to support the content reading. Our research team decided to provide the comparison group with the graphic organizer because in previous experiments the CAP group significantly outperformed the text-only group. Therefore, we sought to determine if strengthening the comparison condition by providing the text-only group with an additional scaffold for comprehension would affect the outcome.

For the Text-GO condition, the instructor or a graduate assistant remained in the classroom with the appropriate students. Each participant received a folder and envelope; the folder contained the reading and graphic organizer, and the envelope contained the posttest. Instructors directed participants when to begin reading. As each participant finished, they were asked to raise their hand before proceeding to the posttest in order to ensure no one used the reading when completing the test.

2.2.2. CAP condition

The researchers created the CAPs to be tested in this experiment, one on the topic of LD and another on the topic of HFA. Both CAPs were carefully aligned with Mayer's CTML (2009) and instructional design principles (2008), which were used as a checklist to ensure fidelity to the design principles (see Fig. 1). Members of the research team independently verified CAPs' respective adherence to Mayer’s principles, and then came together for discussion, and to plan revisions. Revisions were made prior to implementation. Each CAP mirrors content from the selected readings used in the Text-GO condition. Content match was assured between the CAPs and readings through a sentence-by-sentence analysis of the CAPs to ensure no information was presented in the videos but not in the readings, and vice versa.

Researchers used a 15" Macbook Pro laptop, the 2011 version of Microsoft’s presentation software, PowerPoint®, and images from www.ThinkStockPhotos.com and Google Images in the public domain to create CAPs. Each CAP took approximately 60 min to create and upload. Completed files were posted on Vimeo, a video-sharing site, for user access. The CAP on LD can be viewed at https://vimeo.com/72439473, and the CAP on HFA may be accessed at https://vimeo.com/72518420. Detailed production steps to support instructors in creating their own CAPs have been described in previous articles, including Kennedy et al. (2011) and Kennedy and Thomas (2012). Additionally, CAPs teaching how to make CAPs may be viewed at https://vimeo.com/24179998 (part 1) and https://vimeo.com/24182724 (part 2).

Students in the CAP group completed instruction in computer labs housed at each university, while researchers adhered to an implementation fidelity protocol for the experimental condition. Participants wore headphones, and accessed the CAP through a password-protected Vimeo site at individual computer terminals. The supervising instructor gave a signal for participants to start the CAP simultaneously. As in the Text-GO condition, when participants completed watching the podcast, they signaled the proctor, received an envelope, and began the posttest. Researchers circulated the room to ensure students did not navigate away from the CAP or engage in any activities that would detract from the fidelity of the experiment. Research activities were untimed for both conditions.

2.3. Dependent measures

2.3.1. Knowledge test

The knowledge test consisted of 30 multiple-choice items, including 15 questions each on topic of LD and HFA. This test was designed to assess content knowledge of the facts, vocabulary, and concepts (Haladyna, 2004) of these disability related topics, and to measure the effects of instruction (AERA, American Psychological Association, & National Council on Measurement in Education, 1999). Researchers used the same instruments at pretest, posttest, and maintenance. Although the order of the questions was scrambled at each time point to reduce the impact of order effects, use of the same measure was needed to answer our research questions regarding differences in learning following instruction with CAPs. Because participants in both conditions received the test repeatedly, we are able to eliminate repeated testing as a reason for one or the other group having had higher results. Participants did not receive feedback of any kind on test performance until after the experiment was complete, and test scores were not included in students' grades. To help control for recency effects, we scheduled a delay of two weeks between the pretest and posttest, followed by an additional three weeks between the posttest and maintenance probe at University 1 and six weeks at University 2.

Test development procedures included setting the test purpose and mapping the scope and sequence for this test (AERA et al., 1999). Two nationally and internationally recognized experts from the fields of LD and HFA respectively reviewed the appropriate dependent measure and provided critical feedback. The researchers then revised the dependent measures, thereby establishing content validity (AERA et al.). The LD and HFA instruments have also been employed in another study of the impact of CAPs on student learning (Kennedy, Ely, et al., 2012).
Fig. 2 contains sample items from the LD and HFA instruments. The full instruments are available from the first author of this study upon request.

Although the instrument contains knowledge-based items, our intent was to measure one construct each related to the characteristics of students with of LD and HFA. Therefore, for each data set (pretest, posttest, and maintenance for the LD and HFA measure), researchers ran multidimensional item response theory models with one and two factor solutions (Bock, Gibbons, & Muraki, 1988). The Bayesian Information Criterion (BIC) fit statistic (Chalmers, 2012) indicates that the one factor solution was best for the LD experiment (pretest = 4753.1 for one factor, 4858.5 for two; posttest = 5019.2 for one factor, 5167.6 for two; maintenance = 5255.7 for one factor and 5389.9 for two) and the HFA experiment (pretest = 4812.3 for one factor and 4978.6 for two; posttest = 5134.5 for one factor, 5208.8 for two; maintenance = 5332.5 for one factor and 5401.3 for two). Chronbach’s alpha levels for the three time points are all within the acceptable range for social science research (pretest $\alpha = .79$; posttest $\alpha = .80$; maintenance $\alpha = .83$).

2.4. Design

This quasi-experimental study utilizes both a $2 \times 2$ and $2 \times 3$ between group design, with performance alternately measured at two (pretest and posttest) and three time points (pretest, posttest, and maintenance). Students were randomly assigned to either Group 1 (CAP condition) or Group 2 (Text-GO condition). The independent variable was the type of instruction (CAP or Text-GO), and the dependent variables are performance on the posttest and maintenance probe for LD and HFA, respectively. Split-plot, repeated-measures ANOVA was used to analyze data, and one-way ANOVAs are used to evaluate directionality of significant findings.

Although participants at both universities completed the pretest, intervention (CAP or Text-GO), and posttest prior to any formal instruction from the course instructor (to maintain experimental control), implementation of the maintenance probe was purposefully varied between the two universities. Researchers intentionally introduced a naturalistic instructional method to measure performance at maintenance at University 2. To illustrate, students in both conditions (CAP and Text-Go) at University 1 completed the posttest and maintenance probes prior to the introduction of content in class, but students in both conditions at University 2 completed the maintenance probe for LD and HFA after the content was discussed in class and through assigned readings. At University 2, this additional instruction included class lectures on the content in addition to assigned readings. This instruction contained the same information as contained in the CAP/Text-Go conditions, along with new information. Therefore, results are reported for the LD and HFA experiments, respectively, in three parts. First, results from the posttest for the full group of students at both universities are jointly reported. Second, maintenance scores for students at University 1 are reported. Finally, maintenance scores for students at University 2 are reported with the provision that the results include effects of additional instruction beyond the CAP or text-based intervention.

**Sample Questions for LD and HFA Instruments**

**Sample LD Questions**

Learning disabilities may be best described as

a. A disorder characterized by subaverage intelligence  
b. A disorder characterized by an inability to learn  
c. A disorder resulting from traumatic brain injury  
d. A social skills disorder

Strategy instruction is an effective method for students with LD because

a. Strategy instruction helps students process and organize information  
b. Strategy instruction lowers the difficulty level of the instruction  
c. Strategy instruction removes the need for high level critical thinking  
d. Strategy instruction helps students develop peer relationships

**Sample HFA Questions**

When Brian talks, he often includes catch phrases he has heard on television shows and in advertising jingles. This is an example of___________.

a. Delayed echolalia  
b. Social stories  
c. Pragmatic language  
d. Advanced vocabulary

To support students with High Functioning Autism in general education classrooms, teachers should:

a. Prepare students for changes in routines  
b. Lower academic expectations  
c. Ignore social problems as they are unavoidable  
d. Reduce the quantity of work expected
3. Results

Results are organized and reported by experiment (LD and HFA), and also (a) all participants at posttest, (b) participants at University 1 at maintenance, and (c) participants at University 2 at maintenance. This separation reflects the different procedure used at the two universities for the maintenance portion of the experiments.

3.1. LD experiment

3.1.1. Posttest data for all participants

Although we randomly assigned participants from both universities (n = 164) to groups, we conducted a one-way ANOVA to rule out any group differences at pretest. The results of the ANOVA show there were no significant differences between the groups on the LD pretest, \( F(1, 162) = 1.36, p > .05 \). To analyze data, we conducted a 2 x 2 split-plot, fixed-factor repeated measures ANOVA. The between-subjects factor was group assignment (group), and the within-subject factor was time (pretest, posttest). The results showed there was not a significant difference between the CAP group \( (M_{\text{CAP}} = 12.2) \) and Text-GO group \( (M_{\text{Text-GO}} = 12.4) \) at pretest, \( F(1, 162) = 1.36, p > .05 \). However, the mean for the CAP group \( (M_{\text{CAP}} = 12.4) \) was significantly higher than for the Text-GO group \( (M_{\text{Text-GO}} = 10.2) \) at posttest, \( F(1, 156) = 43.10, p < .001, d = 1.09 \). According to Cohen (1988), this is a large effect, and is similar to previous studies using CAPs to deliver instruction to teacher candidates in this domain.

3.1.2. Maintenance data for students at university 1

Students at University 1 (n = 86) completed the LD maintenance probe three weeks after the day of the experiment with no additional instruction about students with LD. Therefore, experimental control was sustained, and these results reflect a measure of the intervention’s impact on student learning without contamination from other factors.

To analyze these data, we conducted a 2 x 2 split-plot, fixed-factor repeated-measures ANOVA. The between-subjects factor was group assignment (group), and the within-subjects factor was time (pretest, posttest, and maintenance). The results showed there was not a significant difference between the CAP group \( (M_{\text{CAP}} = 6.1) \) and Text-GO group \( (M_{\text{Text-GO}} = 6.3) \) at pretest, \( F(1, 162) = 1.36, p > .05 \). However, the CAP group \( (M_{\text{CAP}} = 6.1) \) did not significantly outperform the Text-GO group \( (M_{\text{Text-GO}} = 6.3) \) at posttest, \( F(1, 162) = 1.36, p > .05 \). However, the CAP group \( (M_{\text{CAP}} = 12.4) \) did significantly outperform the Text-GO group \( (M_{\text{Text-GO}} = 9.9) \) at posttest, \( F(1, 162) = 24.33, p < .001, d = 1.09 \), and maintenance \( F(1, 162) = 11.56, p < .001, d = .81 \). In addition, Cohen’s d indicates a large, and medium-to-large effect size at both time points, again consistent with previous experimental work in this domain.

Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>p</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – A5</td>
<td>44</td>
<td>12.7</td>
<td>1.89</td>
<td>.000</td>
<td>1.09</td>
</tr>
<tr>
<td>Text-Only Group – A5</td>
<td>77</td>
<td>10.2</td>
<td>2.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U1</td>
<td>40</td>
<td>12.2</td>
<td>1.98</td>
<td>.000</td>
<td>1.09</td>
</tr>
<tr>
<td>Text-Only Group – U1</td>
<td>41</td>
<td>9.9</td>
<td>2.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U2</td>
<td>51</td>
<td>12.5</td>
<td>1.80</td>
<td>.000</td>
<td>1.00</td>
</tr>
<tr>
<td>Text-Only Group – U2</td>
<td>36</td>
<td>10.8</td>
<td>2.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – A5</td>
<td>84</td>
<td>6.6</td>
<td>2.05</td>
<td>.864</td>
<td>.00</td>
</tr>
<tr>
<td>Text-Only Group – A5</td>
<td>80</td>
<td>6.6</td>
<td>2.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U1</td>
<td>43</td>
<td>6.1</td>
<td>2.07</td>
<td>.574</td>
<td>.09</td>
</tr>
<tr>
<td>Text-Only Group – U1</td>
<td>43</td>
<td>6.3</td>
<td>2.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U2</td>
<td>41</td>
<td>7.2</td>
<td>1.90</td>
<td>.441</td>
<td>.20</td>
</tr>
<tr>
<td>Text-Only Group – U2</td>
<td>37</td>
<td>6.8</td>
<td>2.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U1</td>
<td>37</td>
<td>10.7</td>
<td>2.40</td>
<td>.001</td>
<td>.81</td>
</tr>
<tr>
<td>Text-Only Group – U1</td>
<td>39</td>
<td>8.9</td>
<td>2.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U2</td>
<td>41</td>
<td>10.2</td>
<td>2.50</td>
<td>.599</td>
<td>.13</td>
</tr>
<tr>
<td>Text-Only Group – U2</td>
<td>36</td>
<td>9.9</td>
<td>2.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. A5 – All Students; U1 – University 1; U2 – University 2. Scores are out of 15 points.
3.1.3. Maintenance data for students at university 2

Students at University 2 completed the LD maintenance probe following instruction from the professor and through other course materials. Therefore, these results reflect effects from the intervention conditions and instruction from the professor and other course materials.

To analyze data, researchers conducted a 2 x 3 split-plot, fixed-factor repeated measures ANOVA. The between subjects factor was group assignment (group), and the three levels of the within subject factor (time) were performance on the pretest, posttest, and maintenance probe. Levene's test for equality of variances was not significant for the pretest, posttest and maintenance probe. However, Mauchly's test of sphericity was significant, Mauchly's W = .881, p = .01, thus, Greenhouse-Geisser corrections are used to evaluate and interpret results.

There was a significant main effect for group F(1, 74) = 9.64, p = .003, $\omega^2 = .02$; time F(1.78, 74) = 92.35, p < .000, $\omega^2 = .44$, and the interaction between group and time F(1.79, 74) = 3.34, p = .044, $\omega^2 = .01$. Therefore, 47% of the variance in this model is explained by the sum of effects from the predictor variable (CAP vs. Text-GO). To examine the findings of the significant interactions further, we examined group differences with a simple effects test using a Bonferroni correction at both time points (.05/3 = .0167). Results show the CAP group ($M_{pre} = 7.2$) did not differ significantly from the Text-GO group ($M_{pre} = 6.8$) at pretest F(1, 74) = .599, p = .441. However, the CAP group ($M_{post} = 12.5$; $M_{maint} = 10.2$) significantly outperformed the Text-GO group ($M = 10.6$; $M_{maint} = 9.9$) at posttest $F(1, 74) = 18.17$, p < .000, $d = 1.00$, but not at maintenance $F(1, 74) = .278$, p = .599, $d = .13$. The effect at posttest is large, and consistent with other findings. Fig. 3 illustrates the results for Universities 1 and 2 across the three assessments on the tests of students’ knowledge about LD.

3.2. HFA experiment

3.2.1. HFA experiment pretest–posttest data for all participants

First, we conducted a one-way ANOVA to rule out any group differences at pretest. The results of the ANOVA showed there were no significant differences between the groups on the HFA pretest, $F(1, 150) = 1.21$. To analyze data, we conducted a 2 x 2 split-plot, fixed-factor repeated-measures ANOVA. The between-subjects factor was group assignment (group) and the two levels of the within-subjects factor (time) were performance on the pretest and posttest. Levene’s test for equality of variances was not significant for the pretest or posttest. Table 2 provides the raw score means, standard deviations, $p$ values, and effect sizes (Cohen’s $d$) for the HFA experiment.

There was a significant main effect for group F(1, 150) = 20.58, p < .000, $\omega^2 = .03$; time F(1, 150) = 718.13, p < .000, $\omega^2 = .60$, and the interaction between group and time F(1, 150) = 24.68, p < .000, $\omega^2 = .02$. Therefore, 64% of the variance in this model is explained by the sum of effects from the predictor variable (CAP vs. Text-GO) in our experiment. To examine the finding of the significant interactions further, we examined group differences with a simple effects test using a Bonferroni correction at both time points (.05/2 = .025). Results show the CAP group ($M_{pre} = 8.5$) did not score significantly differently than the Text-GO group ($M_{pre} = 8.3$) at pretest, $F(1, 150) = 1; p = .501$. However, the CAP group ($M_{post} = 13.9$) did significantly outscore the Text-GO students ($M_{post} = 12.0$) at posttest $F(1, 150) = 57.61$, p > .000, $d = 1.21$. Cohen's test indicates a large effect size.

Fig. 3. Results for test of learning disabilities knowledge.
3.2.2. HFA experiment pretest–posttest–maintenance data for students at university 1

Students at University 1 (n = 86) completed the HFA maintenance probe three weeks after the day of the experiment with no additional instruction about children with HFA. Therefore, experimental control was maintained, and these results reflect a measure of the intervention’s impact on student learning without contamination from other factors.

To analyze data, we conducted a 2 \times 3 split-plot, fixed-factor repeated-measures ANOVA. The between subjects factor was group assignment (group), and the three levels of the within-subjects factor (time) were performance on the pretest, posttest and maintenance probe. Levene’s test for equality of variances was not significant for the pretest, posttest and maintenance probe. In addition, Mauchly’s test of sphericity was not significant, Mauchly’s W = .99, p = .991.

There was a significant main effect for group F(1, 71) = 31.30, p < .000, ω² = .18; time F(1, 71) = 120.58, p < .000, ω² = .22, and the interaction between group and time F(1, 71) = 20.33, p < .000, ω² = .04. Therefore, 44% of the variance in this model is explained by the sum of effects from the predictor variable (CAP vs. Text-GO). To examine the finding of the significant interactions further, we examined group differences with a simple effects test using a Bonferroni correction at both time points (.05/3 = .0167). Results show the CAP group (M_{Pre} = 8.0) did not significantly differ from the Text-GO group (M_{Pre} = 7.7) at pretest F(1, 71) = 1.6, p = .249. However, the CAP group (M_{Post} = 13.8; M_{Maint} = 11.3) significantly outperformed the Text-GO group (M = 11.3; M_{Maint} = 9.1) at posttest F(1, 71) = 40.29, p < .000, d = 1.43, and maintenance F(1, 71) = 33.11, p < .000, d = 1.33. Cohen’s test indicates a large effect size at both time points.

3.2.3. HFA experiment pretest–posttest–maintenance data for students at university 2

Students at University 2 completed the HFA maintenance probe following instruction from the professor and other course materials. Therefore, these results reflect effects from the intervention conditions plus instruction from the professor and other course materials.

To analyze data, we conducted a 2 \times 3 split-plot, fixed-factor repeated-measures ANOVA. The between subjects factor was group assignment (group), and the three levels of the within subject factor (time) were performance on the pretest, posttest, and maintenance probe. Levene’s test for equality of variances was not significant for the pretest, posttest and maintenance probe. Mauchly’s test of sphericity was significant, Mauchly’s W = .754, p < .000, thus, Greenhouse-Geisser corrections are used to evaluate results.

There was a significant main effect for group F(1, 71) = 3.76, p = .056, ω² = .00; time F(1, 71) = 153, p < .000, ω² = .59, and the interaction between group and time F(1.6, 71) = 3.33, p = .050, ω² = .01. Therefore, 60% of the variance in this model is explained by the sum of effects from the predictor variable (CAP vs. Text-GO). To examine the finding of the significant interactions further, we examined group differences with a simple effects test using a Bonferroni correction at both time points (.05/3 = .0167). Results show the CAP group (M_{Pre} = 9.0) did not score significantly differently than the Text-GO group (M_{Pre} = 8.9) at pretest F(1.06) = 1, p = .908. However, the CAP group (M_{Post} = 14.1; M_{Maint} = 12.8) did score significantly higher than the Text-GO group (M_{Post} = 12.3; M_{Maint} = 12.3) at posttest F(1, 71) = 23.96, p > .000, d = 1.21, but not at maintenance F(1, 71) = .41, p = .522, d = .00. There is a large effect favoring the CAP group at posttest. Fig. 4 illustrates the results for both University 1 and 2 across the three assessments on the test of knowledge about HFA.

4. Discussion

In recent years, the use of multimedia to deliver instruction in higher education, including preservice teacher education, has soared (Atkins et al., 2010; Kukulska-Hulme, 2012). Despite this increase in usage, an identifiable evidence base for best practice has struggled to keep pace (Bull, Knezek, Roblyer, Schrum, & Thompson, 2005; Hew & Cheung, 2013; Kirschner, Sweller, & Clark, 2006; Lawless & Pellegrino, 2007; Mayer, 2004; Schmid et al., 2009). This gap in understanding of the impact of technology presents an opportunity for researchers concerned about the quality of instruction occurring in higher education to critique this enterprise (e.g., Clark, 2009; Heilesen, 2010; Hew & Cheung, 2013). In part, educational technology adoption has accelerated in response to public perception of technology as a potential

### Table 2

Mean scores, standard deviations, p values, and effect sizes (Cohen’s d) for high functioning autism experiment.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>p</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pretest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – A5</td>
<td>84</td>
<td>8.5</td>
<td>1.82</td>
<td>.501</td>
<td>.09</td>
</tr>
<tr>
<td>Text-Only Group – A5</td>
<td>80</td>
<td>8.3</td>
<td>2.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U1</td>
<td>43</td>
<td>8.0</td>
<td>1.68</td>
<td>.429</td>
<td>.17</td>
</tr>
<tr>
<td>Text-Only Group – U1</td>
<td>43</td>
<td>7.7</td>
<td>1.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U2</td>
<td>41</td>
<td>9.0</td>
<td>1.85</td>
<td>.908</td>
<td>.05</td>
</tr>
<tr>
<td>Text-Only Group – U2</td>
<td>37</td>
<td>8.9</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Posttest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – A5</td>
<td>78</td>
<td>13.9</td>
<td>.99</td>
<td>.000</td>
<td>1.21</td>
</tr>
<tr>
<td>Text-Only Group – A5</td>
<td>74</td>
<td>12.0</td>
<td>1.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U1</td>
<td>39</td>
<td>13.8</td>
<td>1.03</td>
<td>.000</td>
<td>1.43</td>
</tr>
<tr>
<td>Text-Only Group – U1</td>
<td>39</td>
<td>11.3</td>
<td>2.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U2</td>
<td>39</td>
<td>14.1</td>
<td>.94</td>
<td>.000</td>
<td>1.21</td>
</tr>
<tr>
<td>Text-Only Group – U2</td>
<td>35</td>
<td>12.8</td>
<td>1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maintenance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U1</td>
<td>37</td>
<td>11.3</td>
<td>1.16</td>
<td>.000</td>
<td>1.33</td>
</tr>
<tr>
<td>Text-Only Group – U1</td>
<td>36</td>
<td>9.1</td>
<td>2.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP Group – U2</td>
<td>41</td>
<td>12.3</td>
<td>1.35</td>
<td>.981</td>
<td>.00</td>
</tr>
<tr>
<td>Text-Only Group – U2</td>
<td>36</td>
<td>12.3</td>
<td>1.64</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: A5 = All Students; U1 = University 1; U2 = University 2. Scores are out of 15 points.
panacea for the ills of education (Bushaw & McNee, 2009), along with policy proposing its use to promote educational reforms (Atkins et al., 2010).

Technology has largely enjoyed a privileged status in higher education teaching practice because of its ease of use paired with students’ perceptions of enjoyment and effectiveness (Kay, 2012; Lin et al., 2013; Lonn & Teasley, 2009; O’Bannon et al., 2011). To illustrate, it can be tempting to adopt the latest presentation format or widget as an exciting way to “spice up” teaching and perceived student engagement; however, as university instructors it is incumbent upon us to reflect not on the surface-level properties of new programs or tools, but instead make targeted inquiries regarding theoretical and empirical underpinnings. Furthermore, if new tools are adopted, they should be accompanied by attempts to experimentally test the impact on dependent variables of interest, specifically, student learning (Clark, 2009; Heilesen, 2010).

In this study, CAPs supported teacher candidates in learning content about the characteristics of students with LD and HFA. Results confirm that students who learn using CAPs answer more retention questions at posttest than students who read the same content and had access to a graphic organizer and outline of the key content. Students at University 1 also answered more questions correctly at maintenance than the comparison students. This finding replicates prior work in this field. The students at University 2, who received typical in-class instruction after the posttest, but before the maintenance probe, did not replicate this result. One rationale for this finding is the purpose and utility of CAPs. To illustrate, CAPs are not intended as a replacement for in-class instruction. Instead, Author considers CAPs a supplemental tool that students can use prior to reading as an advance organizer (Pre-CAP), or after reading or lecture (Re-CAP) in preparation for a quiz or other assignment. In this case, additional instruction provided by the instructor helped students at University 2 “catch up” to the CAP learners. This finding is promising because the results at University 1 demonstrate that if in-class time constraints are significant, CAPs are a better option than simply reading text alone, and the learning proved to be durable. Given the results at University 2, future research should compare CAPs directly to typical in-class instruction, such as lecture and activities, preferably with a large enough sample to address teacher effects and curricular differences.

Further discussion is warranted to situate the results of the current study within the existing research on CAPs and podcasting in higher education. As noted, this study is the sixth to examine the results of teacher candidates’ learning using CAPs. In the first four studies (Kennedy et al., 2011; Kennedy, Newton, et al., 2012; Kennedy & Thomas, 2012; Kennedy, Ely, et al., 2012) the CAPs had important differences by way of instructional design. However, the early CAPs did not feature built-in pauses with comprehension or discussion questions. This omission could have caused CAPs to unintentionally violate Mayer’s coherence, signaling, and segmenting principles due to the uninterrupted delivery of content.

Another improvement in the CAPs used in this study is an upgrade in the quality and timing of images per Mayer’s personalization, and spatial and temporal contiguity principles. Although the CAPs in this line of research always use royalty free images, Author’s videos now feature professional-quality images from the service www.ThinkStockImages.com. ThinkStock offers a larger number of high quality and royalty free images than can be found randomly using Google or other searches. Because adherence to Mayer’s model requires use of clear images that explicitly represent the content being provided, the use of the higher quality images is essential for comprehension and memory. To illustrate this update, first watch an older CAP on the characteristics of students with LD at https://vimeo.com/14444476, and then see its upgrade at https://vimeo.com/72439473. In summary, the CAPs created for this and more recent studies (Kennedy, Deshler,
et al., 2013; Kennedy, Driver, et al., 2013) have stronger adherence to Mayer’s principles and are, in simple terms, the beneficiaries of time and experience as Author’s instructional team continues to use this tool in research and practice—Clark’s iterative design process uses ongoing data to improve the products.

In terms of contribution to the existing podcasting literature in higher education, a key strength of this program of research on CAPs is that there are multiple experimental replications across different topics in the field of teacher preparation. A close examination of Heilesen’s (2010) review shows that the few studies in this domain that report experimental data come from independent research teams that each published one qualifying study. While it is possible that these respective researchers are just getting started and are currently conducting follow up research, ‘one and done’ publications are limited in what they contribute to the literature base, and therefore, to our understanding of podcasts’ capacity to spur and augment student learning. Programs of research that answer different research questions by continuously testing theories, which concurrently offer new insights into instructional design and learning, are needed. Toward this end, the primary Author and research team are focused on developing and sustaining research on CAPs, and extending use of CAPs into other domains (i.e., teaching vocabulary terms to adolescents with and without disabilities; Kennedy, Deshler, et al., 2013; Kennedy, Thomas, Meyer, Alves, & Lloyd, 2013).

4.1. Limitations

There are several important limitations in this study. First, the sample was drawn from a convenience sample of preservice teacher candidates and other undergraduates from authors’ respective universities. In addition, only 164 students participated. Although in social science research 164 is not typically considered small, future research should seek to conduct experiments with much larger samples and across universities, programs, regions, and for a variety of content. Doing so would permit examination of interactions among subject and instructional variables that would permit researchers to probe theoretically important issues about why and how multimedia-based instructional practices benefit students’ learning. Second, in these experiments, research activities occurred in a laboratory setting and are relatively limited in scope respective to the small number of topics (two) and duration of the intervention. It is possible that the comparison condition was less motivating, given required reading in class, than typical conditions for required readings that occur in a setting of the learner’s preference. Future experiments might investigate conditions in which users had more control of CAPs, and manipulate whether and how students use them in an organic fashion, and if there is a corresponding increase in learning depending upon usage frequency and patterns. In addition, the questions on the pretest might signal viewer’s attention to key content needed for later assessments. Furthermore, although the dependent measures are reliable given standards for preliminary social science research, multiple measures, including open-ended and qualitative evaluations of knowledge, understanding, and ability to apply knowledge to address real world problems are important, and should be addressed in future studies of CAPs.

Although substantial efforts were taken to maintain experimental control for University 1 between the posttest and maintenance probe, it is possible that an overachieving student or students could have pursued the content on their own. However, the quizzes used for the experiment were ungraded and participation was voluntary, so there was no impetus for students to seek additional knowledge; still, this unknown variable raises an important question for interpreting results. Finally, one research team has completed all of the published work on CAPs. Despite the fact that we argued that this program of research on CAPs is a strength in the foregoing discussion, it is also a limitation in that CAPs have not yet been adopted into the research lines of others; when other adopt and adapt CAPs procedures and replicate the findings presented here and elsewhere, the generality of these results will expand (Gersten, Baker, & Lloyd, 2000).

5. Conclusion

Given the criticism levied at the beginning of this article regarding the numbers of professors who incorporate instructional methods and tools into their teaching without considering the theoretical foundations of learning and instruction, this is a significant point of concern for the use of technology in higher education. While some empirical evidence now exists to justify the use of podcasting (Heilesen, 2010), researchers and university instructors should question the atheoretical nature of generic audio-only or enhanced podcasts.

This study reports the sixth experimental replication of CAPs’ utility to deliver knowledge-level instruction to teacher candidates. Effect sizes across these studies are medium to large in size, and reflect several different content areas (characteristics of students with TBI, LD and HFA; PBIS; NCLB, and phonological awareness instruction). Although this is by no means an exhaustive list of topics important for teacher educators, this preliminary work provides rationale to continue testing of the instructional design phase that Clark (2009) terms situated learning. An important unanswered question for future research conducted by this and other research teams is to explore the impact of different types of CAPs on student learning. For example, experimental exploration of the impact on learning given various design options such as (a) use of images with real people compared to cartoons, (b) various lengths of CAPs, (c) location and use of simple recall compared to inferential questions throughout CAPs, (d) the inclusion of video in addition to or in place of still images, and (e) interactive or ‘clickable’ questions embedded within CAPs that require the user click the correct answer to continue. Conducting further experiments with the intent of exploring the effect of these various design features on learning will help refine the CAP process within Clark’s (2009) framework.

References


