Hammill Institute on Disabilities

Learning Disability Quarterly
2014, Vol. 37(2) 71–86

DOI: 10.1177/0731948713507262
ldq.sagepub.com
SAGE

Using Evidence-Based Multimedia to Improve Vocabulary Performance of Adolescents With LD: A UDL Approach

Michael J. Kennedy, PhD[1], Cathy Newman Thomas, PhD[2], J. Patrick Meyer, PhD[1], Kat D. Alves, MEd[1], and John Wills Lloyd, PhD[1]

Abstract

Universal Design for Learning (UDL) is a framework that is commonly used for guiding the construction and delivery of instruction intended to support all students. In this study, we used a related model to guide creation of a multimedia-based instructional tool called content acquisition podcasts (CAPs). CAPs delivered vocabulary instruction during two concurrent social studies units to 32 SWD and 109 students without disabilities. We created CAPs using a combination of evidence-based practices for vocabulary instruction, UDL, and Mayer's instructional design principles. High school students with and without learning disabilities completed weekly curriculum-based measurement (CBM) probes (vocabulary matching) over an 8-week period along with two corresponding posttests. Students were nested within sections of world history and randomly assigned to alternating treatments (CAPs and business as usual) that were administered sequentially to each group. Results revealed that students with and without disabilities made significant growth on CBMs and scored significantly higher on the posttests when taught using CAPs.

Keywords

Universal Design for Learning, multimedia instruction, instructional design, vocabulary instruction, adolescent students with LD

Universal Design for Learning (UDL) is included in important public policies that affect students with and without disabilities. UDL is endorsed within Individuals With Disabilities Education Act Amendments (2004), which instructs teachers, "to the extent possible, [to] use universal design principles in developing and administering any assessments" (Section 300.160), and through technology (Section 300.704) to maximize its use to provide access to the general curriculum. The blueprint for the reauthorization of the Elementary and Secondary Education Act (U.S. Department of Education, 2010) specifically mentions UDL with respect to assessment; curricula, and instructional supports; technology, including technology for science, math, engineering, and technology; and for students with disabilities (SWD) and English Language Learners. The Higher Education Opportunity Act of 2008 defines UDL and provides guidelines for its use in teacher preparation, and the National Educational Technology Plan (Atkins et al., 2010) includes a section on UDL, describing its principles and including information about UDL as a means to provide access to the curriculum for diverse learners.

Clearly, UDL has had a substantial influence on policy. However, policies offer no guarantee that educator practice will be impacted in ways that change instruction for the better (Klingner, 2004). With that in mind, the purpose of this article is to consider the role of UDL in special education in light of the present study's findings of an intervention influenced by UDL, and other instructional design principles and evidence-based practices (EBPs). In addition, we present and describe a new framework for creating effective multimedia for use when teaching SWD that may help translate theory and policy into practice.

Questioning the Use and Evidence Base for UDL in Special Education

Three Interrelated Areas of Concern

Despite UDL's influence on policy, an important, but oft-unasked question is, "To what extent has UDL similarly affected the practice of general and special educators working with students with specific learning disabilities (LD)?" There are at least three ways to address this question. First,

[1]University of Virginia, Charlottesville, USA
[2]University of Missouri, Columbia, USA

Corresponding Author:
Michael J. Kennedy, Curry School of Education, University of Virginia, P.O. Box 400273, Charlottesville, VA 22904, USA.
Email: MKennedy@Virginia.edu

it is important to consider the goals and purposes of special education. Can instruction that is universally designed be appropriate for students with special needs? Relatedly, we must consider the evidence-base for UDL: In which settings and instructional conditions is UDL effective for students with LD? Finally, researchers, practitioners, and other stakeholders need to resolve the extent to which UDL can explicitly guide teacher practice in a way that is observable and measurable within various content areas.

Paradox of UDL. To our first point and question, UDL is somewhat paradoxical within the field of special education. To illustrate, as the Center for Applied Special Technology (CAST; 2011)—arguably the foremost champion of UDL—explains in its mission statement, UDL offers a framework to design instruction that meets the needs of all students. However, the purpose of special education is to provide individualized instruction to help mitigate the impact of a specific disability given a task or set of demands (Kauffman & Hallahan, 2005), and implies that practices will be grounded in evidence (Zigmond, 2006). Because the principles of UDL are intentionally broad so that they can be used in any subject area and with any student, how then should researchers and practitioners in our field reconcile this inherent flexibility with the charge to provide individualized and evidence-based instruction, which, by comparison, is prescribed and rigid? In this regard, we argue that UDL should not be abandoned, but rather subjected to a level of scrutiny expected for interventions used when teaching students with LD.

Empirical evidence for students with LD? Relatedly, a second reason for caution in adopting the UDL framework for use in special education (and in the context of this article, specifically for students with LD) is the limited empirical data that demonstrate UDL's effectiveness on the academic performance of SWD. Although the theory of learning that undergirds UDL is appropriate for many students with LD (Rose & Meyer, 2002), this is not sufficient on its own to categorize UDL as an EBP. To illustrate, when we searched for studies that met Gersten et al.'s (2005) quality indicators for group experimental and quasi-experimental research in special education and that used UDL as the framework to design and deliver vocabulary instruction to adolescents with LD (the topic of this study), we found none. The search did yield three studies that used technology to deliver vocabulary instruction (e.g., Kennedy, Deshler, & Lloyd, 2013; Vaughn et al., 2009; Xin & Rieth, 2001), however, none used UDL as a theoretical framework tied explicitly to students with LD. An article by Proctor et al. (2011) used UDL as its framework, and met the standard for inclusion (use of quasi-experimental design to measure vocabulary performance), but focused on English Language Learners. In sum, there is a significant need for additional support for UDL via publication of peer-reviewed research.

Operational definition and measurement issues. Finally, UDL provides additional challenges in that there is widespread agreement regarding the need for instruction that "has something for everyone" (Basham & Marino, 2013), but, in the same breath spurs disagreement and confusion among policymakers, researchers, and practitioners who do not have reliable mechanisms to identify and measure universally designed instruction (Ralabate et al., 2012). To illustrate, Ralabate et al. (2012) examined UDL implementation at the state and local levels and found that although it enjoys broad recognition, and is included in many state technology plans, there is confusion about what UDL is, how it relates to other initiatives, and how to implement it. Thus, UDL is difficult to describe in that there are no reliable, valid, or widely available measures for universal screening, progress monitoring, implementation fidelity, or outcome assessments (Edyburn, 2010). Although researchers are developing and investigating measures to evaluate UDL as a framework rather than as a collection of individual practices (Basham & Gardner, 2010), the ill-defined nature of UDL may cause difficulties for special educators charged with providing specially designed instruction to students with LD (King-Sears, 2009).

Summary. This discussion is not intended to discredit UDL. In fact, the opposite is true—There is room and merit in marrying existing scientifically valid practice with innovation supported under the UDL framework. However, explicit instructional design and practice, grounded in strong theory and reliable and valid measurement, is needed to help ensure that specially designed instruction and UDL's relationship is extended and bolstered on the basis of evidence for students with LD (Kennedy & Deshler, 2010). As long as these standards are met, teachers can, and should consider UDL as a guiding framework that may interface with existing and relevant instructional approaches (Kennedy, Deshler, et al., 2013; Kennedy & Wexler, 2013).

To address these open questions and existing limitations of UDL for teaching students with LD, we conducted a research study using a multimedia-based intervention that reflects the principles of UDL, but is more specific given a new instructional design framework (the Multimedia Design Framework [MDF]) we propose is intended to shape and improve vocabulary performance of adolescents with and without LD. The intervention is called Content Acquisition Podcasts (CAPs; Kennedy, Deshler, et al., 2013), and in this study is used to provide vocabulary instruction to students with and without LD in a world history course.

A Framework for Integrating UDL Design Principles With EBPs

What Are CAPs?

CAPs are multimedia-based instructional modules with looks and sounds structured in a way consistent with (a) UDL's principles of multiple means of representation and engagement (CAST, 2011), and (b) Mayer's (2008, 2009) cognitive theory of multimedia learning (CTML) and validated instructional design principles. However, quality instructional design principles alone do not constitute an acceptable EBP for students with LD (Kennedy & Deshler, 2010). To achieve that standard, CAPs are (c) infused with content-specific instructional practices, such as explicit instruction for teaching vocabulary terms/concepts, and are subjected to empirical tests of their capacity to augment student performance on dependent measures of interest (Kennedy, Deshler, et al., 2013). Therefore, CAPs represent an innovation in multimedia design that profits from a logical, and opportunistic combination of previously unaffiliated theories, empirical findings, and approaches to instruction for students with LD.

A hallmark of CAPs is that they are intentionally constrained to include only the key content for a topic to not overwhelm the viewer's limited cognitive load levels (Kennedy & Wexler, 2013). Thus, CAPs range in length from 1 to 3 min. In addition, CAPs use vivid images and occasional on-screen text to help the viewer create meaningful representations of content to help facilitate learning (Austin, 2009). The goal of CAPs is to develop background knowledge and vocabulary, accommodating the capacity of working memory to improve learning potential. Although we use CAPs in this study to provide evidence-based vocabulary instruction in a world history course, the idea is that instructors could use this tool to design, package, and deliver instruction for any subject. A sample CAP can be viewed at https://vimeo.com/49191997.

The framework of CAPs is grounded in strong applied theory, but this alone is insufficient to qualify as specially designed instruction. To meet that standard, CAPs are filled with carefully selected evidence-based instructional practices that are appropriate for the content being delivered. To illustrate, six specific instructional practices, grounded in the empirical literature on vocabulary instruction (e.g., Bryant, Goodwin, Bryant, & Higgins, 2003; Ebbers & Denton, 2008; Jitendra, Edwards, Sacks, & Jacobson, 2004), constitute a menu of practices embedded into the instructional routine used within CAPs. These include (a) promoting word consciousness (e.g., pronunciation, spelling, syllables, prefix, suffix, root words; Reed, 2008), (b) providing direct instruction of word meanings (Archer & Hughes, 2011), (c) providing guided practice and scaffolding (Dexter et al., 2011), (d) providing instruction that promotes awareness of closely related terms (Graves, 2006), (e) using the keyword mnemonic strategy (Mastropieri, Scruggs, & Levin, 1987), and (f) providing a statement of purpose/rationale for why the student needs to learn a given term or concept (Deshler & Shumaker, 2006). These six elements of effective vocabulary instruction are represented within a checklist called the Vocabulary Planning Framework (see Kennedy, Lloyd, Cole, & Ely, 2012). When combined with UDL and Mayer's instructional design principles (described below), the use of these EBPs for vocabulary instruction may form the base for a multimedia-based practice that can be tested for use with students. Figure 1 provides three screen shots of CAPs, demonstrating how, for example, they promote word consciousness, use the keyword mnemonic strategy, and provide students rationale to learn the term.

To provide researchers and practitioners with explicit guidance for creating, selecting, and evaluating multimedia (such as CAPs) for use with students with LD, we offer a conceptual framework called the MDF (see Figure 2). In the following discussion, we present the components of the MDF, and describe how it informed the research reported in this article, and why it may be a viable guide for future research and practice.

The MDF

MDF Phase 1: The purpose of instruction. Given wide and unchecked use of multimedia in education, it is self-evident that all multimedia are not created equal. In other words, gratuitous use of multimedia offers instructors no guarantees of success (as measured by student learning), and may actually be counterproductive to that end (Clark, 1994; Kennedy et al., 2012). To illustrate, many students with LD struggle with cognitive processing speed, and focusing on key, rather than extraneous details (Swanson, 2001). Thus, as researchers, practitioners, or other stakeholders continue work in this domain, it is important to remember the end goal should always be measurable student performance on dependent variables of interest, and not rapid creation and adoption of appealing and trendy multimedia (Clark, 2009; Mayer, 2011). This is one compelling reason why the MDF is necessary and used in this research. Figure 3 contains two columns: Column 1 lists the steps of the MDF, column 2 notes how the steps are reflected in the creation of the CAPs developed for this study.

MDF Phase 2: Preplanning questions. Although presented as such in Figures 2 and 3, the MDF is not intended to be a linear design process. Instead, we aim for researchers and practitioners to answer the questions within Phase 2, while simultaneously considering the design framework for UDL detailed by CAST (2011) and Mayer's design principles (2008). The rationale for this nonlinear approach is reflected in our desire to promote forethought with respect to how teachers think about the roles and purposes of multimedia in instruction. To that end, note in Figure 3 how, from a practical perspective, CAPs fill a need for students (e.g., weak vocabulary skills), and address common problems of practice faced by teachers in many secondary-level courses (e.g., vocabulary instruction is not a priority, limited time given competing demands). However, the multimedia is also designed to provide evidence-based instruction for the population of interest while simultaneously addressing practical needs.

This screen shot shows the use of the keyword mnemonic strategy in CAPs. Although the KMS is typically a student-generated strategy (see Mastropieri et al., 1987), given the two-dimensionality of CAPs, the keyword and remembering sentence are provided to the student. Depending on the user's familiarity with the KMS, the CAPs could contain direct instruction regarding its use. See https://vimeo.com/37765820 for additional details.

This screen shot shows how the CAPs creator can cue the user's attention to the important status of the term/concept being taught. The CAPs in this study had the explicit statement: "The term ______ is an important concept in our world history course. It is important because ______."

This screen shot shows how CAPs can use simple text and shapes to explicitly teach students important components of words, including morphological components, or semantic features. CAPs in this study had the following explicit instruction: "You can figure out the meaning of an unknown term if you know what its root word means. The word *alliance* has the root word *ally. Ally* means *friend.* Therefore, the term *alliance* has something to do with being friends." See https://vimeo.com/37764041 for more details.

Figure 1. Screen shots from various evidence-based components of CAPs.
Note. CAP = content acquisition podcast.

MDF Phase 3: Consider principles of UDL. UDL's first principle, providing multiple means of representation, is a logical fit for CAPs. CAPs offer students a new method for receiving high-quality instruction in a format they may prefer to traditional methods, and benefit from academically (Kennedy, Deshler, et al., 2013). CAPs are also a logical example of the UDL principle of multiple means of engagement. In a recent study, students who learned using CAPs reported being motivated to learn compared with other methods for learning vocabulary (Kennedy, Deshler, et al., 2013). Finally, although not previously tested or reported in this article, CAPs may offer students the means to express knowledge. We expand on the potential role for CAPs in offering multiple means of action/expression in the discussion section.

MDF Phase 4: Consider Mayer's principles. The CTML is an applied, learner-oriented theory intended to guide instructional designers in their quest to create effective multimedia. Mayer's (2009) CTML is an outgrowth of cognitive load theory (CLT; Chandler & Sweller, 1991) and Paivio's (1986) dual processing principle (DPP). Specifically, this theory and 12 accompanying instructional design principles provide a framework for designing instruction that meets explicit standards for maximizing a user's essential cognitive processes, while limiting extraneous cognitive load (DeLeeuw & Mayer, 2008). See Table 1 for the 12 principles and effect sizes for each principle based on research spanning the past 20+ years. These principles offer instructional designers a specific and measurable framework for translating UDL's broad principles into practice. Although not specific to special education, Mayer's work in the field of multimedia learning and cognition is well-known and influential in the field, and offers a level of specificity in instructional design not currently available via UDL or other design principles.

MDF Phase 5: Evaluating outcomes. Multimedia instruction should be monitored to examine whether the intended outcomes actually occur. In the case of this study, at Phase 1, classroom performance and pretesting confirmed the need

Phase 1: Purpose of Instruction

1.1 Create multimedia utilizing evidence-based practices and validated instructional design principles that helps support measurable learning gains for students with LD

Phase 2: Pre-Planning Questions

2.1 What are the individual learning needs of the student(s) with disabilities?
2.2 Given the demands of the instructional task(s), what evidence-based practice(s) is indicated?
2.3 Why are you considering multimedia as an instructional tool/approach?
2.4 Given (a) this student's individual learning needs, and (b) the demands of the task, how would the selected evidence-based practice interface with the planned tool/approach?
2.5 What does the tool/approach need to do and look like to accomplish your goal?
2.6 Thinking ahead, what are possible pitfalls of combining this evidence-based practice with a multimedia-based delivery option?

Phase 3: Consider Principles of UDL Framework:
(http://www.udlcenter.org/aboutudl/udlguidelines)

As you think about the development of this tool/approach in the context of this student's individual needs and the demands of the content area, you should answer:
3.1 Does tool/approach provide multiple means of representation?
3.2 Does tool/approach provide multiple means of engagement?
3.3 Does tool/approach provide multiple means of action and expression?
For each: Use the CAST (2011) 2.0 design principles to explain and defend your planned combination of the evidence-based practice and multimedia-based tool/approach

Phase 4: Consider Mayer's Principles (see Figure 4):

To finish the design process, it is important to carefully consider and plan literally every word, image, and sound the learner is exposed to. The reason for this is to carefully manage the learner's cognitive load to the extent possible. Mayer's design principles provide explicit guidance that can be used to that end. Resources available to support the use of these principles are available in Figure 4, and at: http://people.virginia.edu/~mjk3p/docs/CAP_Production_Steps_MK.pdf.

Phase 5: Evaluate Outcomes

The evaluation process should be iterative with data informing design and implementation.
5.1 Does the multimedia design reflect principles of UDL? Mayer's principles? Are there evidence-based practices embedded in the multimedia delivery?
5.2 Does data exist to support need for the intervention and match of intervention to need?
5.2 Was the intervention implemented with fidelity?
5.3 Did the technology function without interruption/failure?
5.4 How was individual student progress monitored toward IEP goals and grade level/content standards?
5.5 Were the intended outcomes achieved? If not, why not?

Figure 2. Multimedia Design Framework.

for vocabulary intervention. In addition, it is imperative that all instruction achieves a level of implementation fidelity that can be validated prior to use. Throughout Phases 2, 3, and 4 of this study, opportunity exists to check instructional decisions and design features against prescribed principles and EBPs. Finally, designing outcome measures that are reliable and valid is core to the mission of improving the utility of multimedia when teaching students with LD. When making decisions regarding the continued use of multimedia, instructors need to be aided by solid data.

Previous study. In a precursor to the current study (Kennedy, Deshler, et al., 2013), the first experimental test of CAPs' potential to improve vocabulary performance of adolescents with LD was conducted in a high school world history course. Results showed that students who learned using this tool significantly outperformed classmates who learned using non-validated instructional materials. In addition, students who learned using CAPs reported significantly higher levels of satisfaction than classmates in the comparison condition. The present study is an attempt to replicate those findings, and improve the reliability and validity of measurement.

Purpose. The key research question for this study is whether and to what extent adolescents with and without LD improve vocabulary performance in social studies

Multimedia Design Framework (MDF) Steps	How This Study's CAPs Were Produced Using the MDF
Phase 1: Purpose of Instruction	Support adolescents with LD and improve vocabulary performance in world history
Phase 2: Pre-Planning Questions **2.1** What are the individual learning needs of the students with disabilities? **2.2** Given the demands of the instructional task(s), what evidence-based practice(s) is indicated? **2.3** Why are you considering multimedia as an instructional tool/approach? **2.4** Given (a) this student's individual learning needs, and (b) the demands of the task, how would the evidence-based practice you selected interface with the planned tool/approach? **2.5** Ideally, what does the tool/approach need to do and look like? **2.6** What are possible pitfalls of combining this evidence-based practice with a multimedia-based delivery option?	**2.1** Weak overall literacy skills, difficulty learning content-specific words via reading, or during lecture **2.2** Explicit instruction, strategy instruction (keyword mnemonic strategy), semantic feature analysis, morphological instruction, multiple exposures to words **2.3** General Ed teacher does not provide evidence-based vocab instruction during class, multimedia can provide supplement and also be re-watched by students at various times/places to gain multiple exposures **2.4** Vocabulary instruction is provided orally, and using visualizations-both lend to use of multimedia **2.5** Short instructional videos that help provide students with explicit definitions and visual cues for one essential term/concept at a time to help build meaningful cognitive representations **2.6** Language in the videos might be too complex for some viewers, selected visuals might not be clear/easily understood, time needed to create a large number of CAPs
Phase 3: Consider Principles of UDL **1.1** Does tool/approach provide multiple means of representation? **1.2** Does tool/approach provide multiple means of engagement? **1.3** Does tool/approach provide multiple means of action and expression?	**3.1** Yes, CAPs provide an alternative mode of presenting instruction to students using visuals, simplified explanations, and a format for learning they may be familiar and comfortable with **3.2** Yes, CAPs are motivating in that they (hopefully) help increase achievement, and offer students a flexible tool for taking charge of their learning **3.3** Not as constructed and used in this study, but easily could be if students created their own CAPs
Phase 4: Consider Mayer's Principles	**4.1** Mayer's principles are used when writing scripts, selecting images, and essentially, when making all production decisions. Final products are juried for content by general ed content expert, and for adherence to this model using the rubric
Phase 5: Evaluate Outcomes **5.1** Does the multimedia design reflect principles of UDL? Mayer's principles? Are evidence-based practices embedded in the multimedia delivery? **5.2** What is your fidelity check for these principles? **5.3** Is there data to support need for intervention and match to student need? **5.4** Did the technology function without interruption/failure? **5.5** How was individual student progress monitored toward IEP goals and grade level/content standards? **5.6** Were the intended outcomes achieved? If not, why not?	**5.1** Yes, CAPs reflect UDL's principles of engagement and representation, and Mayer's 12 instructional design principles. Yes, CAPs include several evidence-based practices for teaching vocabulary **5.2** Based on a rubric of Mayer's principles and expert review, CAPs adhere to these principles **5.3** Yes, student grades in world history are low, also students with LD scored below peers during first semester **5.4** No problems reported **5.5** Students' reading-related goals (e.g., comprehension, oral reading fluency, vocabulary) are measured using CBMs. Unit tests are used to measure performance compared to general education standards **5.6** Yes, students made gains in vocabulary and overall world history performance when using CAPs as an instructional supplement

Figure 3. Sample use of MDF to create CAPs in present study.
Note. MDF = Multimedia Design Framework; CAP = content acquisition podcast.

Table 1. Mayer's Design Principles as Aligned With the Triarchic Model of Cognitive Load.

Triarchic model of cognitive load (DeLeeuw & Mayer, 2008)	Research-based instructional design principles (Mayer, 2009)	Brief description of Mayer's (2008, 2009) instructional design principles
Limit Extraneous Processing	Coherence Principle (*ES* = 0.97)	Instructional materials are enhanced when irrelevant or extraneous information is excluded
	Signaling Principle (*ES* = 0.52)	Learning is enhanced when explicit cues are provided that signal the beginning of major headings or elements of the material being covered
	Redundancy Principle (*ES* = 0.72)	Inclusion of extensive text (transcription) on screen along with spoken words and pictures hinders learning. Carefully selected words or short phrases, however, augment retention (Mayer & Johnson, 2008)
	Spatial Contiguity Principle (*ES* = 1.12)	On screen text and pictures should be presented in close proximity to one another to limit eye shifting during instructional presentations
	Temporal Contiguity Principle (*ES* = 1.31)	Pictures and text should shown on screen should correspond to the audio presentation
Manage Essential Processing	Modality Principle (*ES* = 1.02)	People learn better from spoken words and pictures than they do from pictures and text alone
	Segmenting Principle (*ES* = .98)	People learn better when multimedia presentations are divided into short bursts as opposed to longer modules
	Pretraining Principle (*ES* = 0.85)	People learn better when there is an advance organizer that highlights and reviews key content prior to instruction
Foster Generative Processing	Multimedia Principle (*ES* = 1.39)	People learn better from pictures and spoken words than from words alone
	Personalization, Voice, and Image Principles (*ES* = 1.11)	Narration presented in a conversational style result in better engagement and learning than more formal audio presentations. Images should be nonabstract, and clearly represent the content being presented.

following instruction using CAPs, as compared with when CAPs are not used? Results are discussed in light of how the UDL framework can intersect with Mayer's instructional design principles and evidence-based instructional methods to support the learning of students with LD.

Method

Participants

The University Human Subjects Committee, the participating school district's research review board, the principal of the school, parents of all students, and the students gave permission to conduct research. The school district was located in an urban community of 146,867 residents. At the time of the study, the selected high school had a student enrollment of 1,159, 83% of whom received free or reduced-price lunch. Researchers recruited a world history teacher (Mr. Awesome) responsible for teaching five total sections of 10th-grade world history to participate in the study. Mr. Awesome was selected given the recommendation of the principal as a strong educator of SWD included within his courses. A total of 141 urban high school students participated in the study. Two groups of students participated: (a) SWD (n = 32) and (b) students without disabilities (n = 109).

SWD. All SWD in this study had an Individualized Education Plan (IEP) on file with the school and received special education services. Each student received daily special education services embedded within their core academic content classes taught by a general education teacher (e.g., social studies, science, mathematics, and language arts), and also had a study skills course taught by a special educator. These 32 students were in the 10th grade at the time of the study, and had a mean age of 16.9 years. In terms of demographic characteristics, 76% are male and 24% are female; 63.3% are African American, 26.7% are Hispanic, and 10% are Caucasian. Permission to collect individual socioeconomic status could not be obtained from the school district's human subjects review board. However, given that approximately half of the 10th graders in the school are enrolled in one of the five sections of world history participating in this project, and 83% of students received free or reduced-price lunch, we assumed that more than three quarters of students in this study received free or reduced-price lunch.

Approximately 84% (n = 27) of the SWD were individuals with specific LD. All but one of these students had LD in an area related to reading, which, according to the National Joint Committee on Learning Disabilities' (NJCLD; 1991) definition, means these students have significant difficulty acquiring and using skills and knowledge related to reading. The rationale for adopting this definition is, in content-area classrooms, most content demands are tied to reading processes and the accompanying cognitive processing tasks (Shanahan & Shanahan, 2008). Of the remaining five students, three received special education services under the behavior disorders category, and two qualified in the category of intellectual disability.

At the time of the study, the district used the *Wechsler Individual Achievement Test* (WIAT-II) as part of their special education identification and reevaluation processes. Because SWD in this study were administered the WIAT-II at different times given their respective reevaluation cycles, the aggregate data reported here should be interpreted with some caution. Given its relevance to the goals of this study, we report aggregate results from the reading subtests (word reading, reading comprehension, pseudoword decoding). Based on age-based scores, the average standard scores for students with LD (n = 27) were 81.3 (SD = 3.6) for word reading, 77.6 (SD = 4.6) for reading comprehension, and 74.3 (SD = 6.5) for pseudoword. The corresponding average percentile ranks fall approximately at the 10th, 6th, and 4th percentiles, respectively. The average percentile ranks in word reading, reading comprehension, and pseudoword decoding for the 3 students with EBD were 18th, 5th, and 10th, respectively, and the two students with intellectual disabilities each scored at or below the 1st percentile for each subtest. The mean *Wechsler Intelligence Scale for Children–IV* (WISC-IV) IQ for the 27 students with LD was 91.3 (SD = 8.6). The mean WISC-IV IQ for the 3 students with EBD was 98.1 (SD = 13.9), and it was 76 (SD = 2.8) for the students with intellectual disabilities.

General education students. Of the 109 general education student participants in this study, African American students represented the largest ethnic group (72%), Caucasian students were the next largest group at 13%, and Hispanic students constituted 10%. The mean age of participants was 16.3 years, 46% were female, 54% were male, and 92% were in the 10th grade. No standardized achievement data were available for these students. We used students' first-semester grade point average (GPA) in world history to sort them into one of three levels of achievement status: (a) High Achiever—85% or above, (b) Typical Achiever—84% to 70%, and (c) Low Achiever—69% and below.

Procedures

Research activities for this study occurred across two concurrent units of world history: Renaissance and Revolutions (RR) and Exploration and Expansion (EE). Mr. Awesome, the participating teacher, had an average of 28.2 students spread across five sections of world history that met for 50 min, 5 times a week. He also taught an average of 6.2 SWD per section who were included alongside their peers without disabilities. It was not possible to assign students randomly to experimental or control conditions; therefore, randomization occurred at the section level. Researchers randomized the order in which the students within the five sections received instruction using CAPs, or standard practice for the entire unit. To illustrate, students in section 1 were taught using standard practice for the RR Unit and then were taught using CAPs during the EE Unit.

CAP production and validation. The first author and a research assistant produced a total of 81 CAPs for use in this study. Two reviewers each of whom had approximately 40 hr of experience in producing and evaluating CAPs for a previous study used a production rubric to score CAPs independently for the present study (e.g., Kennedy, Deshler, et al., 2013). Reviewers completed one rubric per CAP (n = 81) to gauge adherence to Mayer's instructional design principles. Two separate reviewers with expertise in adolescent literacy and vocabulary instruction used the CAP Vocabulary Instruction eWorksheet (VIeW; see Kennedy et al., 2012) to determine the CAPs' utilization of EBPs for vocabulary instruction. Reviewers used understanding of these principles and their judgment to complete this task. Feedback from the reviewers informed revisions prior to use in the study. Interscorer reliability across the reviewers was 92%. We revised the CAPs based on the feedback and then invited reviewers to review them a second time. CAPs were not used in the study until all concerns were satisfied. Finally, Mr. Awesome had the last word regarding the content of each CAP to ensure they were accurate in content and appropriate for his students. The teacher's requested changes were made to problematic CAPs prior to use. The most common feedback was to ensure pictures selected and displayed in the CAP (which included some maps) accurately represented content being discussed.

Research procedures. Students completed a pretest for each respective unit on the school day following completion of the previous unit's exam. The pretest was untimed, and occurred during class. Students were told to do their best, but the score would not count for or against their grade. For daily instruction, Mr. Awesome was provided with preproduced CAPs stored on the school's intranet, organized into folders by topic. For the appropriate sections of his world history course, researchers instructed Mr. Awesome to use his overhead LCD projector and classroom speakers to play the CAP when he introduced each new term. Prior to the study, the first author met with Mr. Awesome to demonstrate use of CAPs within typical instruction. Mr. Awesome frequently used PowerPoint slides to help organize his lectures and provide students with written notes to be copied into their notebooks, so he would insert a slide for the new term that cued the need to play the CAP. Each CAP was shown during class twice, once when first introduced, and again during a review prior to the exam. Mr. Awesome tracked this information using a log emailed to the researchers.

The SWD watched each CAP two additional times during their study skills period, again following the pattern of watching the video on or near the day of first presentation, and again prior to the exam. The special education teacher who worked in Mr. Awesome's classroom tracked this information using a log that she emailed to the researchers. This ability for students to review CAPs multiple times based on individual needs, and without the need for face-to-face instruction by a teacher or peer demonstrates one of the key affordances of well-designed, student-centered multimedia platforms such as CAPs.

For students in the non-CAP sections, Mr. Awesome provided text-based definitions of the vocabulary terms and concepts using his overhead projector once when first introducing the term, and again during the review. Students were also required to copy the definition into their notebooks using the textbook glossary or another resource. The special education teacher reviewed these terms with students during study skills periods prior to the exam. The researchers did not monitor this instruction; instead, teachers kept logs of which terms were explicitly taught and/or reviewed to ensure all were covered during the scope of the unit. The researcher monitored the logs and provided Mr. Awesome prompting questions (e.g., Have you taught imperialism yet in Period 4?) when discrepancies occurred.

Students completed weekly curriculum-based measurement (CBM) probes on the Friday of each week (Monday was used for makeups). Probes were completed during class, and were scored by the first author and a research assistant. Because probes were simple matching items, there were no unusual discrepancies regarding scoring. Further detail regarding the probes is noted in a subsequent section. Mr. Awesome told students that the weekly probe would not count as a quiz grade, but rather went toward their participation score. He also informed them that each probe would likely contain terms that he had not yet taught, and so to not become discouraged if they could not answer all of the questions. Finally, students took the posttest for each of the two units during class as usual. These grades did count toward their course grade. Further detail regarding the posttest is also provided in a later section.

Measures

Pretest and posttest. Mr. Awesome, the participating teacher in this study, was required to use two common assessments created by a committee of world history teachers and content specialists at his district office as the unit test for two concurrent units. All world history teachers in the district used these tests; district officials used student data to monitor teacher's progress, and to make revisions to the tests and the curriculum. Researchers received permission for Mr. Awesome to use these assessments as a pretest. Respectively, the assessments on the *renaissance* and *age of exploration* contained 40 and 42 items, and were scored out of 100 points. The items on both instruments contained multiple-choice, short-answer, and essay questions. The multiple-choice items had three distracters and one correct answer. The short-answer questions required the student to write 1 to 3 sentences using basic and applied knowledge. Finally, essay questions required students to write 5 to 10 sentences using basic and applied knowledge to construct coherent responses.

Data regarding the technical adequacy of this instrument was not made available to the research team. However, based on a calculation of test–retest reliability using each participant's pretest and posttest scores, the approximate bivariate correlation for the two time points of the R&R instrument is .69, and the EE instrument is .65. Students in each group were comparable on the pretests given before each unit. For the RR Pretest, students scored an average of 14.62 and 14.66 (out of 100) in the two groups, respectively. For the EE Pretest, average scores were 14.21 and 14.03 for each group. Independent-samples t tests for the RR ($t_{139} = -0.07, p = .94$) and EE unit pretests ($t_{139} = 0.35, p = .73$) revealed no statistically significant differences between groups.

Curriculum-based measures. Researchers created CBM probes for in this study using methods described by Espin, Busch, Shin, and Kruschwitz (2001). Mr. Awesome, the teacher participating in this study, identified a total of 81 items that included terms and definitions, and historical figures and descriptions he planned to teach during the two concurrent units. Researchers randomly selected (with replacement) 20 terms/historical figures and 22 definitions/descriptions from the bank of terms and historical figures to form eight probes. The two additional definitions/descriptions on each probe served as distracters. Definitions and descriptions were taken directly from the course textbook. Probes were set up so that terms and historical figures were presented in alphabetical order on the left-hand column, and definitions and descriptions were placed in random order in the right-hand column. Coefficient alpha for all eight CBM scores was 0.90 when using scores from the entire group.

Design

Because the students in this study were nested within sections of Mr. Awesome's world history course, we could not use a traditional between-groups design. Therefore, we randomly counter balanced each of the five sections so that each section was taught using CAPs for one unit, and not during the other. This design permitted us to analyze performance using between and within-subjects groupings. Pretest, CBM, and posttest data were collected for each student across both units. For the purpose of comparing the between-groups performance, we collapsed average CBM and posttest scores across the five sections in the alternating conditions into two groups (CAP and Business as Usual—BAU). In addition, we examined performance on the CBMs, which is nested within students, using a multilevel growth model. This type of model allows estimation

Table 2. Descriptive Statistics for All Students on the Eight Weekly CBM Probes and Two Unit Posttests.

Period	*n*	Probes 1 and 5	Probes 2 and 6	Probes 3 and 7	Probes 4 and 8	RR and EE Posttest
RR Unit (Probes 1–4)—SWD						
3 (CAP)	5	6.0 (1.6)	7.0 (1.6)	8.4 (1.3)	8.8 (.84)	78.2(2.3)
5 (CAP)	10	5.5 (1.1)	7.1 (1.4)	7.8 (1.3)	9.8 (1.8)	78.4 (3.6)
1 (No CAP)	6	5.3 (1.9)	6.2 (1.2)	7.0 (1.1)	6.5 (.55)	67.0 (5.2)
2 (No CAP)	7	3.6 (1.5)	4.9 (1.8)	5.4 (2.0)	6.6 (1.6)	72.1 (5.7)
4 (No CAP)	4	3.5 (1.3)	4.3 (1.0)	5.3 (1.3)	5.5 (1.7)	71.3 (4.6)
M CAP	15	5.7 (1.2)	7.1 (1.4)	8.0 (1.3)	9.5 (1.6)	78.3 (3.1)
M No CAP 17	4.2(1.7)	5.2 (1.6)	5.9 (1.7)	6.3(1.4)	70.1 (5.5)	
RR Unit (Probes 1–4)—Gen Ed						
3 (CAP)	24	5.9 (1.7)	7.8 (2.2)	8.5 (2.2)	9.8 (2.0)	83.4 (6.1)
5 (CAP)	20	6.8 (1.7)	9.0 (2.0)	9.8 (2.2)	11.0 (2.6)	85.4 (6.3)
1 (No CAP)	21	5.1 (2.0)	6.2 (2.8)	7.2 (3.1)	8.2 (2.5)	80.5 (7.0)
2 (No CAP)	21	5.2 (1.9)	6.3 (2.2)	7.3 (2.7)	8.5 (2.5)	82.1 (7.3)
4 (No CAP)	23	4.4 (1.7)	6.1 (2.6)	6.9 (2.8)	7.6 (2.8)	78.2 (7.3)
M CAP	44	6.3 (1.7)	8.3 (2.2)	9.0 (2.3)	10.3 (2.4)	84.3 (6.2)
M No CAP 65	4.9 (1.9)	6.2 (2.5)	7.1 (2.8)	8.1 (2.6)	80.2 (7.3)	
EE Unit (Probes 5–8)—SWD						
1 (CAP)	6	12.0 (1.4)	12.7 (1.9)	14.0 (2.4)	15.0 (1.4)	81.5 (4.2)
2 (CAP)	7	10.7 (2.5)	11.9 (2.7)	14.6 (2.2)	15.4 (1.5)	79.3 (4.8)
4 (CAP)	4	10.5 (1.3)	12.8 (1.9)	14.5 (1.3)	15.8 (1.3)	80.8 (2.6)
3 (No CAP)	5	7.8 (1.3)	10.2 (2.3)	9.0 (1.6)	11.6 (2.4)	73.2 (2.8)
5 (No CAP)	10	9.5 (2.6)	11.1 (3.1)	11.6 (2.7)	12.5 (3.4)	74.7 (6.3)
Mean CAP	17	11.1 (1.9)	12.4 (2.1)	14.4 (2.0)	15.4 (1.4)	80.4 (4.0)
Mean No CAP 15	8.9 (2.3)	10.8 (2.8)	10.7 (2.7)	12.2 (3.1)	74.2 (5.3)	
EE Unit (Probes 5–8)—Gen Ed						
1 (CAP)	21	14.7 (2.2)	16.5 (2.6)	16.8 (2.5)	17.8 (2.1)	86.2 (6.0)
2 (CAP)	21	13.9 (3.3)	15.8 (2.9)	17.2 (2.4)	18.1 (2.1)	88.0 (5.4)
4 (CAP)	23	15.6 (3.3)	17.1 (2.8)	18.1 (2.0)	18.6 (1.5)	84.7 (6.0)
3 (No CAP)	24	11.7 (2.2)	13.3 (2.1)	14.3 (2.5)	15.2 (2.7)	80.2 (7.0)
5 (No CAP)	20	13.5(2.9)	14.4 (3.3)	15.1 (2.6)	16.5 (2.5)	80.0 (6.9)
M CAP	65	14.7 (3.0)	16.5 (2.8)	17.4 (2.3)	18.2 (1.9)	86.2 (5.9)
M No CAP	44	12.5 (2.7)	13.8 (2.7)	14.6 (2.6)	15.8 (2.7)	80.1 (6.9)

Note. CBM probes are out of 20 points; Posttest is out of 100 points. RR = Renaissance and Revolutions; EE = Exploration and Expansion; CAP = content acquisition podcast; SWD = students with disabilities; CBM = curriculum-based measurement.

of the variance in initial conditions and linear growth while also estimating the average trend.

Results

Between-Groups Analyses

SWD. To evaluate group differences during the RR and EE units for SWD, we conducted a series of one-way ANOVAs. Table 2 contains descriptive data for participants on the two posttests and eight CBM probes. For the first set of ANOVAs, SWD from Periods 3 and 5 taught using CAPs in the RR unit are compared with their peers from Periods 1, 2, and 4 taught using a BAU approach. Thus, students from Periods 3 and 5 were combined to form the CAP group, and their peers from Periods 1, 2, and 4 are the BAU group. Comparisons are made using a mean CBM score derived from probes 1 to 4 for each group, and the mean group score on the RR posttest.

SWD taught using CAPs (n = 15) had significantly higher average CBM scores on Probes 1 to 4 (M = 7.6, SD = 1.1) than students in the BAU condition (n = 17, M = 5.4, SD = 1.3) $F(1, 31) = 20.0$, $p < .001$; $d = 1.83$. Similarly, SWD in the CAP condition (M = 78.3, SD = 3.1) significantly outscored SWD in the BAU condition on the posttest (M = 70.1, SD = 5.5), $F(1, 515) = 28.8$, $p < .001$, $d = 1.84$. Table 2 contains more information related to the ANOVAs.

Students taught using CAPs in the RR unit (Periods 3 and 5) rotated into the BAU condition for the EE unit, and their counterparts who started in the BAU condition

(Periods 1, 2, and 4) rotated into the CAP condition. For the EE unit, SWD taught using CAPs (n = 17) had significantly higher average CBM scores on Probes 5 to 8 (M = 13.3, SD = 1.6) than students in the BAU condition (n = 15, M = 10.7, SD = 2.5), $F(1, 61) = 18.1$, $p < .001$, d = 1.24. Similarly, SWD in the CAP condition (M = 80.4, SD = 4.0) significantly outscored SWD in the BAU condition (M = 74.2, SD = 5.3) on the RR posttest, $F(1, 347) = 21.0$, $p < .001$, d = 1.32.

General education students. We used the same procedures and analyses to examine group differences for the general education students as reported for the SWD. General education students taught using CAPs for the RR unit (n = 44) had significantly higher average CBM scores on Probes 1 to 4 (M = 8.5, SD = 2.0) than students in the BAU condition (n = 65, M = 6.6, SD = 2.3), $F(1, 97) = 20.1$, $p < .001$; d = 84. Similarly, students in the CAP condition (M = 84.3, SD = 6.2) significantly outscored students in the BAU condition (M = 80.2, SD = 7.3) on the posttest, $F(1, 440) = 9.3$, p = .003, d = .61. Table 3 contains more information related to the ANOVAs.

Finally, for the EE unit, students taught using CAPs (n = 65) had significantly higher average CBM scores on Probes 5 to 8 (M = 16.7, SD = 2.3) than students in the BAU condition (n = 44, M = 14.2, SD = 2.5), $F(1, 167) = 28.8$, $p < .001$, d = 1.04. Similarly, SWD in the CAP condition (M = 86.2, SD = 5.9) significantly outscored students in the BAU condition (M = 80.0, SD = 6.9) on the posttest, $F(1, 989) = 25.0$, $p < .001$, d = .95.

Table 3. One-Way ANOVA Between-Groups Results.

Group	*n*	*M (SD)*	*df*	*F*	*p*	*d*
RR Unit—CBM Average—SWD						
CAP	15	7.6 (1.1)	1	20	.000	1.83
No CAP	17	5.4 (1.3)				
RR Unit—Posttest—SWD						
CAP	15	78.3 (5.5)	1	28.8	.000	1.84
No CAP	17	70.1 (3.1)				
EE Unit—CBM Average—SWD						
CAP	17	13.3 (1.6)	1	18.1	.001	1.24
No CAP	15	10.7 (2.5)				
EE Unit—Posttest—SWD						
CAP	17	80.4 (4.0)	1	21	.001	1.32
No CAP	15	74.2 (5.3)				
RR Unit—CBM Average—General Ed						
CAP	44	8.4 (2.0)	1	20.1	.000	0.84
No CAP	65	6.6 (2.3)				
RR Unit—Posttest—General Ed						
CAP	44	84.3 (6.2)	1	9.3	.003	0.61
No CAP	65	80.2 (7.3)				
EE Unit—CBM Average—General Ed						
CAP	65	16.7 (2.3)	1	28.8	.000	1.04
No CAP	44	14.2 (2.5)				
EE Unit—Posttest—General Ed						
CAP	65	86.2 (5.9)	1	25	.000	0.95
No CAP	44	80.1 (6.9)				

Note. CBM Average Scores are out of 20 Points, Posttests are out of 100 Points. RR = Renaissance and Revolutions; EE = Exploration and Expansion; CAP = content acquisition podcast; SWD = students with disabilities; CBM = curriculum-based measurement.

Within-Groups Growth Analyses

As noted, the posttest for the RR and EE units contained items assessing student knowledge of content that extended beyond instruction provided in the CAPs. Although academic vocabulary performance is a reliable predictor of overall performance in content areas (Espin & Campbell, 2012), in this study CBM performance is the stronger measure of CAPs' impact on student learning. Therefore, we conducted a second set of growth analyses using CBM data to examine the impact of the CAP intervention on student learning across the experiment.

To account for the variability in student CBM score trends, we analyzed data with a multilevel growth model such that CBM scores were nested within student. This type of model allows us to estimate the variance in initial conditions and linear growth while also estimating the average trend. We built our model incrementally starting with the fully unconditional model. The intraclass correlation for the unconditional model was 0.08, which indicates that 8% of the variance in CBM scores is attributable to student differences. Across the 8 weeks of instruction, CBM scores tended to become more variable, which implies that students were learning at different rates.

Our second model included effects for initial status and linear growth. Reliability was 0.81 for both of these Level 1 coefficients. In addition, adding these coefficients, reduced within student variance by 87%. However, a significant portion of variance in initial status and linear growth remained. To help explain the variability in initial status and linear growth and account for outstanding variance at Level 1, we constructed our final model with three Level 2 covariates and an additional Level 1 covariate. At Level 2, we added a variable for RR pretest scores, and a variable for EE pretest scores. These variables were grand mean centered and included to account for prior achievement. The third Level 2 covariate was a variable that indicated whether a student was in General Education or Special Education. This variable allowed us to evaluate the difference in growth between these two student populations.

Although our intervention was aimed at the student level (Level 2), the study design did not permit us to include intervention as a Level 2 variable. Recall that half of the students received instructing using CAPs during the first 4 weeks while the other half received BAU instruction. In the second 4 weeks, the situation was reversed. That is, intervention changed during the course of the study. Therefore, we included this variable at Level 1 as a nonrandom

Table 4. Growth Model Results.

Fixed effect	Coefficient	*SE*	*t*	*df*	*p* value
Intercept, π_0					
Intercept, β_{00}	3.83	0.22	17.72	137	<.001
Disability, β_{01}	0.19	0.46	0.41	137	.682
RR Pretest, β_{02}	−0.236	0.09	−2.5	137	.014
EE Pretest, β_{03}	0.041	0.1	0.4	137	.689
Week, π_1					
Intercept, β_{10}	1.719	0.04	44.58	137	<.001
Disability, β_{11}	−0.287	0.09	−3.25	137	.002
RR Pretest, β_{12}	0.004	0.02	−0.23	137	.82
EE Pretest, β_{13}	0.009	0.02	0.47	137	.637
Intervention, π_2					
Intercept, β_{20}	2.665	0.13	20.61	1,119	<.001
Random effect	**Variance**	***df***	**x^2**	***p* value**	
Intercept, r_0	1.88	137	674.72	<.001	
Week, r_1	0.32	137	414.71	<.001	
Level 1, e	1.47	137			

time-varying covariate. More concisely, our Level 1 model was, $Y_{ti} = \pi_{0i} + \pi_{1i}(\text{Week}) + \pi_{2i}(\text{Interventation}) + e_{ti}$, where Y_{ti} is a student's CBM score at time t, π_{0i} represents a student's initial status, π_{1i} represents the mean linear growth rate, and π_{2i} indicates the effect of the intervention. Our Level 2 model (i.e., student level) is as follows,

$$\pi_{0i} = \beta_{00} + \beta_{01}(\text{SPE}) + \beta_{02}(\text{RR Pretest}) + \beta_{03}(\text{EE Pretest}) + r_{0i},$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(\text{SPE}) + \beta_{12}(\text{RR Pretest}) + \beta_{13}(\text{EE Pretest}) + r_{1i},$$

$$\pi_{2i} = \beta_{20}(\text{CAP Intervention}).$$

Results from the full model indicated that grow parameters were stable and related in expected ways. Reliability estimates for initial status and mean linear growth in the full model were .80 and .67, respectively. The correlation between initial status and linear growth was −.09, which indicates that students with a lower initial status tended to learn at a faster rate.

Table 4 shows the estimates of fixed and random effect from the full model. Those estimates indicate that among general education students, initial CBM scores are 3.83 points and increase at a rate of 1.72 points per week on average. Initial CBM scores for special education students were 0.19 points higher than initial CBM scores for general education students, but their average linear growth is 0.29 points less per week. That said, the effect of the intervention more than compensates for these different growth rates. Student CBM scores are 2.67 points higher during the time when they are exposed to the intervention. This effect is illustrated by the gap between the two lines in Figure 4. Indeed, including intervention in the model reduces individual difference in CBM scores by 23%.

Discussion

Clark (1994, 2009) and Mayer (2009, 2011), who are among the foremost leaders in the field of multimedia design and learning, often criticize the use of multimedia in instruction. These scholars, respectively, note the important difference between multimedia that looks appealing by way of cutting-edge graphics and other design features that are attractive to users, and multimedia that meets specific benchmarks for supporting cognition and has undergone experimental testing (Clark, 2009; Mayer, 2009, 2011). Indeed, much of the published literature in the field of multimedia learning reports user's satisfaction with greater frequency than empirical tests of learning (Heileson, 2010; Kennedy, Hart, & Kellems, 2011; Kennedy, Driver, Pullen, Ely, & Cole, 2013).

The present study offers an example of multimedia-based instruction that goes beyond user preferences. To do so, we held ourselves to a higher standard regarding multimedia that is grounded in appropriate theory, but also capitalizes on design principles that support user cognition and learning. The MDF, which reflects an agglomeration and synthesis of the principles of UDL (CAST, 2011), Mayer's design principles (2008), and EBPs for vocabulary instruction, guided our processes during development and research of the CAPs tested in this study.

When researchers combine various theories and practices, they add complexity with respect to measurement and

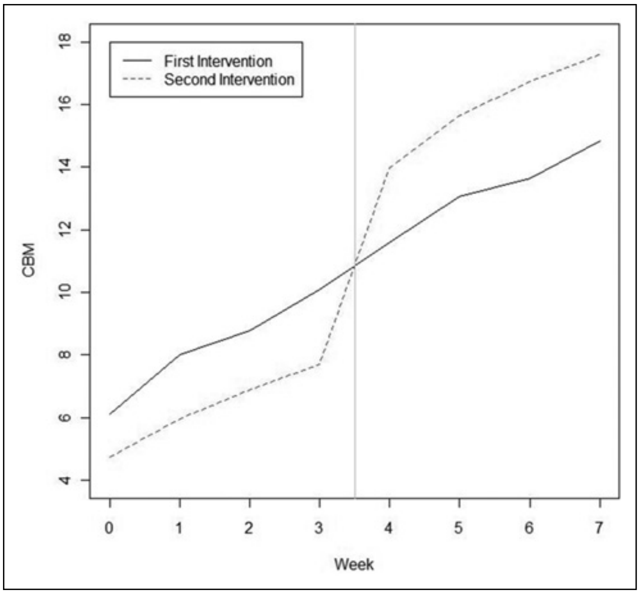

Figure 4. Trend lines for mean CBM scores among students receiving instruction using CAPs during the RR or EE units.
Note. CAP = content acquisition podcast; RR = Renaissance and Revolutions; EE = Exploration and Expansion.

interpretation. As we reflect on the present study, we must ask if the observed gains in performance are attributable to UDL, Mayer's principles, the embedded evidence-based vocabulary practices, or all of the above. To disentangle the shared variance using statistics requires a series of experiments isolating instruction designed and delivered using UDL, Mayer's principles, and the various vocabulary practices. This work constitutes a program of research being undertaken by members of this research team. With respect to understanding the present results, we offer the conceptual argument reflected within the MDF.

This argument draws its significance from the opening deconstruction of UDL's role in special education. Because the principles of UDL are intentionally broad, and do not inherently lend themselves to checks of fidelity or empirical measurement, there is a need to bolster precision in both of these domains. Mayer's (2008, 2009) framework provides a more measurable set of principles to shape the looks and sounds of multimedia that still reflect the opportunity to provide students with multiple means of representation, engagement, and action and expression. Conceptually, the Phases of the MDF provide guidance that integrates these different-but-complementary approaches to instructional design. Finally, the UDL framework and Mayer's respective principles of instructional design have limited utility unless they can actually be used to design and deliver instruction for specific groups of learners. This is an aspect of using multimedia to design and provide instruction that is overlooked by researchers, practitioners, and other stakeholders with great regularity (Clark, 1994, 2009; Mayer, 2009, 2011). As special educators, our role is to provide evidence-based, individualized instruction to SWD. Most educational multimedia, especially videos found on You Tube, educational Apps, and software simply do not meet this standard (Kennedy, Deshler, et al., 2013), and should not be looked to as a primary source of instruction for our students.

Results of this study show that, when instructed with CAPs built on valid instructional design principles and evidence-based instructional methods, SWD learned vocabulary terms and concepts at a faster rate and in a more powerful way than when taught using a BAU approach. This study's findings are similar to those from Kennedy, Deshler, and colleagues' (2013) first test of CAPs for promoting vocabulary performance. Effect sizes should be interpreted with caution given limited sample sizes (especially of SWD) and lack of standardized measurement; however, the large effect sizes provide justification to continue this line of research. In addition, although students without disabilities taught using CAPs still outperformed SWD who used CAPs on CBMs and the posttests, based on mean CBM and posttest scores, the SWD taught using CAPs closed gaps on students without disabilities taught using the BAU approach. This result is preliminary, and subject to question given the limited scope of this study, but the finding is encouraging.

Limitations

This study has several important limitations. First, although we used a quasi-experimental design, only 141 students participated. In addition, only 32 SWD participated. Although these are not small numbers in social science research, the students were enrolled in one high school, thus representing a homogeneous group. The small number of SWD —and specifically, LD—does not permit appropriately powered statistical examinations of differences at the various time points. Although mean scores are clearly in favor of students at times when they learned using CAPs, the observed results may ultimately be misleading, given the substantial changes in mean scores and overall interpretability that could be introduced by even one or a few more additional student scores. That said, a close examination of the mean differences show students in the CAP condition only scored, on average, approximately three questions better than students in the comparison condition. Clearly, that difference is enough to result in statistical significance; but may or may not represent educational significance. However, multiplied over the many units that comprise a semester, a year, or a sequence of courses, the difference becomes more substantial. Future research should aim to boost the number of participants and carefully examine and interpret the educational significance of results.

Second, the first author and a research assistant created all of the CAPs used in the study. Although the CAPs were created using a production rubric based on Mayer's CTML and a checklist for effective elements of vocabulary instruction, and were reviewed by experienced colleagues, important questions remain about the ability of other teachers or researchers to create CAPs. This is an important question to be answered by future research. Relatedly, researchers did not explicitly observe the vocabulary instruction that took place during the comparison conditions. Although teachers were asked to note when they taught or reviewed each term using a log, this process was not standardized or controlled experimentally. Thus, errors with respect to whether specific terms were taught when teachers said they were are possible.

Third, the researchers created the measures used in the study. Standardized measures of vocabulary knowledge for specific content areas (e.g., world history) do not exist, and other standardized measures were not appropriate for use in the study given the research questions. Furthermore, given the limited scope of this experiment with respect to terms that were taught as well as the duration of the study (approximately 8 weeks), growth on a standardized measure would likely be unachievable. An important question to be addressed by future research is the extent to which CAPs used across a semester or entire year may impact achievement.

Implications for Practice

CAPs offer a means of instruction that differs from the type of instruction frequently found in secondary-level content-area courses (Kennedy & Ihle, 2012), and is a mechanism to help motivate learners as well (Kennedy, Driver, et al., 2013). Future research could explore ways students could be taught to create their own CAP, to support the valued outcome related to becoming goal-directed learners. Therefore, researchers in the field of special education should continue to use EBPs to support the individualized needs of SWD. Coupling these practices with innovative methods for packaging and delivering instruction, such as UDL, is also recommended. Researchers should organize and conduct programs of research that utilize experimental and quasi-experimental designs to carefully manipulate independent variables, and measure their impact on reliable and valid dependent measures.

Our goal is that the MDF may provide researchers and practitioners with a starting point for integrating broad and specific instructional design principles and practices with individual student learning needs. At minimum, even for those who are not creating new multimedia from the ground up, considering the questions posed throughout the five phases of the MDF can steer decisions away from convenient, but potentially ineffective instructional options.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Archer, A., L., & Hughes, C. A. (2011). *Explicit instruction—Effective and efficient teaching*. New York, NY: Guilford Press.

Atkins, D., Bennett, J., Brown, J. S., Chopra, A., Dede, C., Fishman, B., & Williams, B. (2010). *Transforming American education: Learning powered by technology national educational technology plan 2010*. Washington, DC: U.S. Department of Education, Office of Educational Technology.

Austin, K. A. (2009). Multimedia learning: Cognitive individual differences and display design techniques predict transfer learning with multimedia learning modules. *Computers & Education*, *53*, 1339–1354. Retrieved from http://dx.doi.org/10.1016/j.compedu.2009.06.017

Basham, J. D., & Gardner, J. E. (2010). Measuring Universal Design for Learning. *Special Education Technology Practice*, *12*, 15–19.

Basham, J. D., & Marino, M. T. (2013). Understanding STEM education and supporting students through universal design for learning. *Teaching Exceptional Children*, *45*, 8–15.

Bryant, D. P., Goodwin, M., Bryant, B. R., & Higgins, K. (2003). Vocabulary instruction for students with learning disabilities: A review of the research. *Learning Disability Quarterly*, *26*, 117–128. Retrieved from http://dx.doi.org/10.2307/1593594

Center for Applied Special Technology. (2011). *Universal design for learning guidelines version 2.0*. Wakefield, MA: Author.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293–332.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, *42*, 21–29.

Clark, R. E. (2009). Translating research into new instructional technologies for higher education: The active ingredient process. *Journal of Computing in Higher Education*, *21*, 4–18. Retrieved from http://dx.doi.org/10.1007/s12528-009-9013-8

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*, 223–234. Retrieved from http://dx.doi.org/10.1037/0022-0663.100.1.223

Deshler, D. D., & Shumaker, J. B. (2006). *High school students with disabilities: Strategies for accessing the curriculum*. New York, NY: Corwin Press.

Dexter, D. D., Park, Y. J., & Hughes, C. A. (2011). A meta-analytic review of graphic organizers and science instruction for adolescents with learning disabilities: Implications for the intermediate and secondary science classroom. *Learning Disabilities Research & Practice*, *26*, 204–213.

Ebbers, S. M., & Denton, C. A. (2008). A root awakening: Vocabulary instruction for older students with reading difficulties. *Learning Disabilities Research & Practice*, *23*, 90–102. Retrieved from http://dx.doi.org/10.1111/j.1540-5826.2008.00267.x

Edyburn, D. L. (2010). Would you recognize universal design for learning if you saw it? Ten propositions for new directions for the second decade of UDL. *Learning Disability Quarterly*, *33*, 33–41.

Espin, C. A., Busch, T. W., Shin, J., & Kruschwitz, R. (2001). Curriculum-based measurement in the content areas: Validity of vocabulary-matching as an indicator of performance in social studies. *Learning Disabilities Research & Practice*, *16*, 142–151. Retrieved from http://dx.doi.org/10.1111/0938-8982.00015

Espin, C. A., & Campbell, H. M. (2012). They're getting older… but are they getting better? The influence of curriculum-based measurement on programming for secondary-school students with learning disabilities. In C. A. Espin, K. L. McMaster, S. Rose, & M. M. Wayman (Eds.), *A measure of success—The influence of curriculum-based measurement on education* (pp. 149–164). Minneapolis: University of Minnesota Press.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*, 149–164.

Graves, M. F. (2006). *The vocabulary book: Learning & instruction*. New York, NY: Teacher's College Press.

Heileson, S. B. (2010). What is the academic efficacy of podcasting? *Computers & Education*, *55*, 1063–1068. doi:10.1016/j.compedu.2010.05.002

Individuals With Disabilities Education Act Amendments of 2004, Public Law 108-446. (2004). Retrieved from http://www.gpo.gov/fdsys/pkg/PLAW-108publ446/html/PLAW-108publ446.htm

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children*, *70*, 299–322.

Kauffman, J. M., & Hallahan, D. P. (2005). *Special education: What it is and why we need it*. Boston, MA: Allyn & Bacon.

Kennedy, M. J., & Deshler, D. D. (2010). Literacy instruction, technology, and students with learning disabilities: Research we have, research we need. *Learning Disability Quarterly*, *33*, 289–298.

Kennedy, M. J., Deshler, D. D., & Lloyd, J. W. (2013). Effects of multimedia vocabulary instruction on adolescents with learning disabilities. *Journal of Learning Disabilities*. doi:10.1177/0022219413487406

Kennedy, M. J., Driver, M. K., Pullen, P. C., Ely, E., & Cole, M. T. (2013). Improving teacher candidates' knowledge of phonological awareness: A multimedia approach. *Computers & Education*, *64*, 42–51. doi:10.1016/j.compedu.2013.0.1.010

Kennedy, M. J., Hart, J. E., & Kellems, R. O. (2011). Using enhanced podcasts to augment limited instructional time in teacher preparation. *Teacher Education and Special Education*, *34*, 87–105. Retrieved from http://dx.doi.org/10.1177/0888406410376203

Kennedy, M. J., & Ihle, F. M. (2012). The old man and the sea: Navigating the gulf between special educators and the content area classroom. *Learning Disabilities Research & Practice*, *27*, 44–54. Retrieved from http://dx.doi.org/10.1111/j.1540-5826.2011.00349.x

Kennedy, M. J., Lloyd, J. W., Cole, M., & Ely, E. (2012). Specially designed vocabulary instruction in the content areas: What does high quality instruction look like? *Teaching Exceptional Children*, *45*, 7. Retrieved from http://tecplus.org/articles/article/1/0

Kennedy, M. J., & Wexler, J. (2013). Helping students succeed within secondary-level STEM content. Using the "T" in STEM to improve literacy skills. *Teaching Exceptional Children*, *45*, 26–33.

King-Sears, M. E. (2009). Universal design for learning: Technology and pedagogy. *Learning Disability Quarterly*, *32*, 199–201.

Klingner, J. K. (2004). The science of professional development. *Journal of Learning Disabilities*, *37*, 248–255. doi:10.1177/00222194040370031001

Mastropieri, M. A., Scruggs, T. E., & Levin, J. R. (1987). Learning-disabled students' memory for expository prose: Mnemonic versus non-mnemonic pictures. *American Educational Research Journal*, *24*, 505–519.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, *63*, 760–769. Retrieved from http://dx.doi/10.1037/0003066X.63.8.760

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press. Retrieved from http://dx.doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2011). *Applying the science of learning*. New York, NY: Pearson.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*, 380–386. doi:10.1037/0022-0663.100.2.380

National Joint Committee on Learning Disabilities. (1991). Learning disabilities: Issues on definition. *Asha*, *33*(Suppl. 5), 18–20.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.

Proctor, C. P., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing*, *24*, 517–544. doi:10.1007/s11145-009-9218-2

Ralabate, P., Dodd, E., Vue, G., Karger, J., Smith, F., Carlisle, A., & Eidelman, H. (2012). *Understanding the impact of the race to the top and ARRA funding on the promotion of universal design for learning*. Wakefield, MA: National Center on UDL.

Reed, D. K. (2008). A synthesis of morphology interventions and effects on reading outcomes for students in grades k-12. *Learning Disabilities Research & Practice*, *23*, 36–49. Retrieved from http://dx.doi.org/10.1111/j.1540-5826.2007.00261.x

Rose, D. H., & Meyer, A. (2002). *Teaching every student in the digital age: Universal design for learning*. Alexandria, VA: ASCD.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy [Special issue: Adolescent Literacy]. *Harvard Educational Review*, *78*, 40–59.

Swanson, H. L. (2001). Research on interventions for adolescents with learning disabilities: A meta-analysis of outcomes related to higher-order processing. *The Elementary School Journal*, *101*, 331–348. Retrieved from http://dx.doi.org/10.1086/499671

U.S. Department of Education, Office of Educational Technology. (2010). *Transforming American education: Learning powered by technology*. Washington, DC: Author.

Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., & Francis, D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade English language learners: Findings from two experimental studies. *Journal of Research on Educational Effectiveness*, *2*, 297–324. doi:10.1080/19345740903167018

Xin, J. F., & Rieth, H. (2001). Video-assisted vocabulary instruction for elementary school students with learning disabilities. *Information Technology in Childhood Education Annual*, *1*, 87–103.

Zigmond, N. (2006). Reading and writing in co-taught secondary school social studies classrooms: A reality check. *Reading and Writing Quarterly*, *22*, 249–268. Retrieved from http://dx.doi.org/10.1080/10573560500455711