Journal of Educational Psychology
2012, Vol. 104, No. 1, 250–263

0022-0663/11/$12.00 DOI: 10.1037/a0026147

Verbal Redundancy in Multimedia Learning Environments: A Meta-Analysis

Olusola O. Adesope
Washington State University, Pullman

John C. Nesbit
Simon Fraser University, Burnaby

Verbal redundancy arises from the concurrent presentation of text and verbatim speech. To inform theories of multimedia learning that guide the design of educational materials, a meta-analysis was conducted to investigate the effects of spoken-only, written-only, and spoken–written presentations on learning retention and transfer. After an extensive search for experimental studies meeting specified inclusion criteria, data from 57 independent studies were extracted. Most of the research participants were postsecondary students. Overall, this meta-analysis revealed that outcomes comparing spoken–written and written-only presentations did not differ, but students who learned from spoken–written presentations outperformed those who learned from spoken-only presentations. This effect was dependent on learners' prior knowledge, pacing of presentation, and inclusion of animation or diagrams. Specifically, the advantages of spoken–written presentations over spoken-only presentations were found for low prior knowledge learners, system-paced learning materials, and picture-free materials. In comparison with verbatim, spoken–written presentations, presentations displaying key terms extracted from spoken narrations were associated with better learning outcomes and accounted for much of the advantage of spoken–written over spoken-only presentations. These findings have significant implications for the design of multimedia materials.

Keywords: multimedia, review, narration, redundant, audio

Advances in information technologies have resulted in greater availability of multimedia materials intended to improve teaching and learning (Mayer, 2009; Mayer & Moreno, 2002; Schnotz & Rasch, 2005; Zhu & Grabowski, 2006). Teachers and instructional designers can now deploy interactive multimedia that present instructional materials in various combinations of on-screen texts, images, video, audio, and animation. A frequently appearing configuration in this multimedia landscape is verbally redundant, simultaneously presented, spoken narration and written text (Kalyuga, Chandler, & Sweller, 2004), which we refer to as *spoken–written* presentation. Read-along books accompanied by recorded audio, lectures in which instructors read from projected text, and television programs with closed captions are examples of spoken–written presentation.

The term *redundancy* has been frequently used by researchers to describe the presentation of the same information through different presentational modes and sense modalities (Sweller, 2005b). Sweller (2005b) observed that the effects of redundancy have been "discovered, forgotten, and rediscovered many times over many decades" (p. 159). Examining past research on redundancy, Moreno and Mayer (2002b) concluded that "response times and memory are facilitated when redundant signal information is presented simultaneously in two sensory channels rather than in one" (p. 156). For example, Lewandowski and Kobus (1993) found that, with two separate samples, participants who were presented with the same words concurrently in auditory and visual modes recalled more words than did participants who were presented with words in a single mode. Miller (1937, as cited in Sweller, 2005b) showed that presentation of redundant materials facilitated better reading and text comprehension than did presentation of nonredundant materials.

Although Moreno and Mayer (2002b) claimed that "the converging findings in the theoretical and applied literatures is that bisensory redundant stimulus presentations are more efficiently processed by the human cognitive system" (p. 157), some analyses of recall from verbally redundant texts have produced contradictory results. For example, Koroghlanian and Sullivan (2000) found that concurrent spoken–written presentations did not yield learning benefits when compared with written-only presentations. Kalyuga et al. (2004) found in two experiments that participants who learned from nonconcurrent spoken–-written presentations recalled more information than did those who learned from concurrent spoken–written presentations. In a third experiment, they found that concurrent spoken–written presentation resulted in poorer recall than did spoken-only presentation. In research with learners of English as a second language, Atkinson, Clark, Harrison, Koenig, and Ramirez (2007) found that redundant presentations were advantageous under some conditions and detrimental under other conditions. It appears that the mixed results in verbal redundancy research and the difficulty of drawing conclusions about its effects on learning may be attributed to varying learner

This article was published Online First November 7, 2011.

Olusola O. Adesope, Educational Psychology Program, Department of Educational Leadership and Counseling Psychology, College of Education, Washington State University, Pullman; John C. Nesbit, Faculty of Education, Simon Fraser University, Burnaby, British Columbia, Canada.

Support for this research was provided by a grant from the Social Sciences and Humanities Research Council of Canada (410-2007-1157) to John C. Nesbit.

Correspondence concerning this article should be addressed to Olusola O. Adesope, Department of Educational Leadership and Counseling Psychology, Washington State University, Pullman, WA 99164-2114. E-mail: olusola.adesope@wsu.edu

characteristics, research methods, outcome variables, and features of the redundant presentations. Researchers have acknowledged the contrasting findings in the literature on verbal redundancy and remarked on the need for resolution (Moreno & Mayer, 2002b). The principal goal of this review is to reconcile the outcomes in verbal redundancy research through a meta-analysis that examines the different conditions under which verbally redundant presentations affect learning. There is presently no meta-analysis on the instructional effects of verbal redundancy and no summative analysis of potential moderating variables. Consequently, researchers, teachers, and instructional designers lack a comprehensive account of the different conditions under which spoken–written materials facilitate or inhibit learning and have little empirical guidance toward a theory of how learning processes are affected by verbal redundancy.

Cognitive Theories of Verbal Redundancy

There is a substantial body of research suggesting that learning outcomes are affected by the availability of cognitive resources for schema acquisition (Kalyuga et al., 2004; Sweller, 2005b). The effects of verbally redundant presentations on learning may depend on whether they enhance or interfere with schema acquisition: Learning is enhanced if coordinating or linking spoken and written representations promotes schema acquisition but is inhibited if coordinating and linking the representations is a resource-consuming cognitive activity whose direct effects are transient. In this section, we review prominent cognitive theories underlying verbal redundancy research.

Cognitive Load Theory (CLT)

Cognitive load is a theoretical construct that refers to the demands placed on the limited capacity of working memory as learners process instructional information (Ayres, 2006; Bobis, Sweller, & Cooper, 1993; Kalyuga, 2011; Kirschner, 2002; Kirschner, Ayres, & Chandler, 2011; Mayer & Moreno, 2003; Paas, Renkl, & Sweller, 2003, 2004; Park, Moreno, Seufert, & Brünken, 2011; Plass, Moreno, & Brünken, 2010; Sweller, 1999; Sweller, van Merriënboer & Paas, 1998; van Merriënboer & Sweller, 2005). Sweller (2005b) noted that “redundant material interferes with rather than facilitates learning” (p. 159). Generally, cognitive load researchers have used the term *redundancy effect* to refer to situations in which learning is hindered by redundant presentations (Chandler & Sweller, 1991; Kalyuga, Chandler, & Sweller, 1998; Rey & Buchwald, 2011; Sweller, 2005a, 2005b). In CLT, redundancy effects are often explained in terms of split attention and are thought to occur when multimedia instructional materials require learners to split their attention among disparate sources of information and mentally integrate these multiple sources of information before meaningful learning can occur (Florax, & Ploetzner, 2010). For example, learning is inhibited when on-screen text is sequentially added to an already intelligible diagram because the text redundantly describes with words what is already represented with a diagram thus requiring learners to expend additional cognitive resources to integrate the sequential information (Chandler & Sweller, 1991; Moreno & Mayer, 2002b). Similarly, researchers have found that learning is inhibited when learners are presented with concurrent animation, narration, and on-screen text (Mayer & Moreno, 1998; Moreno & Mayer, 1999).

CLT also predicts that concurrent spoken–written presentations will inhibit learning by inducing extraneous processing (Mayer, 2005b). Kalyuga et al. (2004) claimed that mental integration of information from two or more redundant sources needlessly “consumes cognitive resources that become unavailable for learners to process essential information” (p. 568).

Dual Processing Theory of Multimedia Learning (DPTML)

Verbal redundancy has been explained using the dual processing theory of multimedia learning (Mayer, 2005a; Moreno & Mayer, 2002b). The DPTML makes the following three fundamental assumptions: (a) working memory has separate auditory and visual systems for processing verbal and nonverbal information, and these systems are partially independent (Baddeley, 1992); (b) each of these two working memory systems is limited in processing capacity (Chandler & Sweller, 1992; Mayer, 2009; Moreno & Mayer, 2002b; van Merriënboer & Sweller, 2005); and (c) learning from both verbal and visual materials occurs when relevant information in each mode is selected, organized, and integrated with information in the other mode (Clark & Paivio, 1991; Paivio, 1986).

DPTML predicts that when learners concurrently read and hear a message, they represent the written information in visual working memory and the corresponding spoken information in verbal working memory (Moreno & Mayer, 2002b). Since verbal and visual working memories are functionally distinct and draw from different cognitive resources, learners can efficiently hold verbal and visual representations in working memory while building referential connections between them (Mayer & Moreno, 1998). Information that is misperceived in one modality is available through the other modality. Thus, DPTML predicts greater learning, retention, and transfer from spoken–written presentations than written-only and spoken-only presentations (Mayer, Heiser, & Lonn, 2001).

Individual Differences

The effects of verbal redundancy may vary with individual differences among learners. Considering the evidence that low prior knowledge learners show stronger multimedia effects than do high prior knowledge learners (Mayer & Gallini, 1990; Mayer, Steinhoff, Bower, & Mars, 1995), we predicted that low prior knowledge students would learn more from verbally redundant materials than would those with high prior knowledge. Similarly, while redundancy might support nonfluent readers by helping them decode unfamiliar words, we predicted that fluent readers would receive no learning benefit from redundant materials.

There is evidence that students who have weak reading or language skills benefit from spoken–written presentations. For example, in a study of students of English as a second language (ESL) learning about the human circulatory system in a multimedia environment, Atkinson et al. (2007) found that verbally redundant on-screen text and audio narration raised retention and conceptual understanding of the multimedia material in comparison with presentations limited to unimodal audio narration or on-screen text. However, other research has suggested that such findings may not apply to all ESL learners. For example, Diao and Sweller (2007) found that learning was hindered in older ESL students with many years of experience studying the English language when the students were presented with concurrent written and spoken words.

In other contexts, spoken–written presentations may serve as accessibility options for those who are hard of hearing or blind. Such materials may significantly remove or minimize barriers to accessing information for learning (Adesope & Nesbit, 2005; Wald, 2008).

Signaling Text

How verbal redundancy is implemented could determine the effects of learning from spoken–written presentations. There is evidence that learning is enhanced when spoken narration is accompanied by brief, written phrases that replicate only key terms and phrases in the narration. Mayer and Johnson (2008) found that participants who studied a diagram accompanied by a spoken narration recalled more when key phrases from the narration consisting of only two or three words appeared at an appropriate location in the diagram as the phrases were being spoken.

Purpose of the Meta-Analysis

Research on verbal redundancy has produced mixed results. Some studies have found positive effects of verbal redundancy (McNeill, 2004; Mousavi, Low, & Sweller, 1995; Ritzhaupt, Barron, & Kealy, 2011; Ritzhaupt, Gomes, & Barron, 2008), and others have reported that learning with verbally redundant spoken–written presentations did not improve performance on achievement-related tasks (Jamet & LeBohec, 2007; Kalyuga et al., 2004; Mayer et al., 2001). Mixed results have been found within some studies (Atkinson et al., 2007; Craig, Driscoll, & Gholson, 2004, Craig, Gholson, & Driscoll, 2002; Moreno & Mayer, 2002b), with positive effects occurring in some conditions and negative effects in others. These contradictory outcomes are not surprising, considering that verbal redundancy research has been conducted with a wide variety of multimedia features, learning conditions, comparison treatments, and outcome measures. Hence, the goal of this investigation is to conduct a meta-analysis to reconcile inconsistencies in verbal redundancy research and estimate the specific effects of learning with spoken–written presentations under different conditions. Specifically, this meta-analysis sought answers to the following research questions:

1. What are the effects of learning from concurrent spoken–written presentations in comparison with spoken-only and written-only presentations?

2. How do verbal redundancy effect sizes vary with different educational levels, participants' reading fluency, and prior domain knowledge?

3. How are verbal redundancy effects moderated by pacing of presentation, degree of correspondence between the audio and the text, size of text segments, and the presence or absence of images in the learning materials?

4. How are verbal redundancy effects moderated by contextual features of studies, outcome constructs, and test formats?

5. How are verbal redundancy effects influenced by methodological features of the research?

Method

Selection Criteria

For this meta-analysis, a study was deemed eligible for inclusion if it

1. investigated the effect of verbal redundancy within a multimedia learning environment by comparing the effects of written-only or spoken-only presentations with those of spoken–written presentations;

2. reported measurable cognitive outcomes including recall, transfer, or response times;

3. reported sufficient data to allow for effect size calculations;

4. randomly assigned participants to groups, or reported pretest or other prior data to control for preexisting differences between the groups; and

5. was publicly available either online or in library archives.

Studies that reported insufficient information for effect size calculations were excluded. For multiple reports of the same study (e.g., dissertation, conference proceedings, and journal article), the version published as a journal article was coded but other versions were used to make the coding features more extensive and accurate.

Location and Selection of Studies

We used the query *redundan** OR *bisensory stimul** as keywords to conduct a comprehensive and systematic search on the following electronic databases: Digital Dissertations, ERIC, PsycARTICLES, PsycINFO, and Web of Science. The reference sections of a number of articles that investigated the effect of verbal redundancy on learning were also searched (e.g., Kalyuga et al., 2004; Mayer & Johnson, 2008; Moreno & Mayer, 2002b). The search procedure returned a total of 1,857 articles.

There were two filtering phases to determine whether articles returned by the search should be included. In the first filtering phase, the abstract or online text of articles found in the search were read to determine eligibility for further examination. Full-text copies were obtained for the 159 articles that passed the first filtering phase. In the second filtering phase, we reviewed the 159 full-text copies by applying the selection criteria stated earlier. A total of 33 articles met all inclusion criteria.

Data from the 33 articles were extracted and entered into an extensive coding form. The coding variables were organized into the following eight major categories: (a) study identification, (b) research questions, (c) sample information, (d) treatment and control conditions, (e) research design, (f) recruitment and consent, (g) dependent variables, and (j) results. In cases where a coded variable was not explicitly reported, we made reasonable inferences and noted that the coding was based on inference. Although this practice is somewhat subjective and less ideal than directly coding information reported in the articles, documentation of the inference provided a degree of control and allowed us to rate the methodological quality of the studies and evaluate the impact it had on the effect sizes. Rating methodological quality can help in explaining the variability among studies included in a meta-analysis (Bernard, Abrami, Lou, & Borokhovski, 2004). The mean interrater agreement between two researchers on all coded variables was about 98%.

Extraction and Calculation of Effect Sizes

Because some articles reported multiple studies, the 33 articles produced 57 studies. For each study included in this meta-analysis, we obtained Cohen's *d* effect size, a standardized estimate of the difference in achievement scores between students who learned with spoken–written presentations compared with those who learned with either spoken-only or written-only presentations. Specifically, Cohen's *d* was calculated as the difference between the experimental (spoken–written presentations) and control (spoken-only or written-only presentations) mean scores divided by the pooled standard deviation of the two groups. Because differential sample sizes across studies may bias the effect size obtained by Cohen's *d*, Hedges and Olkin (1985) proposed the use of Hedges's *g* to reduce the bias. Hedges's *g* (Hedges, 1981; Hedges & Olkin, 1985, p. 81) was computed and reported throughout this meta-analysis as an unbiased estimate of the standardized mean difference effect size. In a few cases where basic descriptive statistics were not provided, effect sizes were estimated from other statistics provided in the studies using conversion formulas (Cooper & Hedges, 1994; Glass, McGaw, & Smith, 1981). Positive effect sizes indicate benefits of spoken–written verbal presentations over spoken-only or written-only presentations.

Data Analysis

Standard guidelines for conducting a meta-analysis were followed in all data analyses (Adesope, Lavin, Thompson, & Ungerleider, 2010; Bernard et al., 2009; Cooper & Hedges, 1994; Egger, Smith, & Phillips, 1997; Hedges & Olkin, 1985; Lipsey & Wilson, 2001; Nesbit & Adesope, 2006). Data were analyzed using Comprehensive Meta-Analysis 2.2.048 (Borenstein, Hedges, Higgins, & Rothstein, 2008) and SPSS Version 18 for Windows.

The weighted mean effect sizes were aggregated to form an overall weighted mean estimate of the effect of learning with spoken–written presentations (i.e., $g+$). This approach allowed more weight to be assigned to studies with larger sample sizes. The significance of each weighted mean effect size was determined by its 95% confidence interval. When the lower limit of a confidence interval was greater than zero, the mean effect size was interpreted as indicating a statistically detectable result favoring the use of verbally redundant presentations. The width of a confidence interval indicates the degree of variability in the results across the studies in one category.

A critical step in any meta-analysis is determining if the observed effect sizes that are combined into a mean all estimate the same population effect size. This property of the sample, called homogeneity of variance, was tested by the *Q* statistic generated by the Comprehensive Meta-Analysis software. When the observed effect sizes have homogeneity of variance, *Q* approximates a χ^2 distribution with $k - 1$ degrees of freedom, where *k* is the number of effect sizes or studies for a particular subset. If *Q* exceeded the critical value of the χ^2 distribution (i.e., $p < .05$), the mean effect size was reported to be significantly heterogeneous, meaning that the effect sizes in a category had more variability than would be expected from sampling error and indicating that the effect sizes did not estimate a common population mean (Lipsey & Wilson, 2001). The I^2 statistic was reported to more comprehensively interpret the result of the homogeneity test (Higgins & Thompson, 2002; Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006). When *Q* is less than or equal to the degree of freedom concomitant with a subset, I^2 is assigned a value of zero. Negative values of I^2 are assigned a value of zero such that I^2 lies between 0% and 100%. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity. Researchers have suggested that percentages of around 25% ($I^2 = .25$), 50% ($I^2 = .50$), and 75% ($I^2 = .75$) should be interpreted to mean low, medium, and high heterogeneity, respectively (Higgins & Thompson, 2002).

Comprehensive Meta-Analysis software was used to evaluate the effect of removing potential outliers from the distribution of effect sizes and to examine whether the removal of such potential outliers would significantly influence the homogeneity results (Hedges & Olkin, 1985). The forest plot of the 57 standardized mean difference effect sizes was assessed, and seven potential outlying studies were removed one at a time. In each case, the recalculated results did not improve the fit of the remaining effect sizes to a simple model of homogeneity. Consequently, we decided to include all 57 studies in further analyses.

Results

A total of 33 research reports yielding 57 independent effect sizes were analyzed. Again, throughout this meta-analysis, positive effect sizes indicated an advantage for spoken–written presentations. Table 1 shows a summary of the variables coded for each study, including the study identification, domain of study, grade level of participants, total number of participants involved in each study, comparison materials (i.e., the instructional materials provided to the control or comparison group), pacing of presentation, level of redundancy between the text and audio, whether images were included, and the associated unbiased effect size (Hedges's *g*).

Effects of Learning With Spoken–Written Versus Written-Only or Spoken-Only Presentations

Tables of results include the number of participants (*N*) in each category, the number of findings (*k*), the weighted mean effect size ($g+$) and its standard error (*SE*), the 95% lower and upper confidence intervals, the results of a test of homogeneity (*Q*) with its concomitant degrees of freedom (*df*), and the percentage of variability that could be attributed to true heterogeneity or between-

Table 1
Descriptive Information and Effect Sizes for Coded Studies

Study	Domain	Grade range	*N*	Comparison material	Pacing of presentation	Level of redundancy	Images included?	Effect size (*g*)
Atkinson et al. (2007)	Biology	7	48	Audio	LP	High	Yes	1.16*
Atkinson et al. (2007)	Biology	7	45	Text	LP	High	Yes	−0.25
Barron (1994)	Computer lit.	PS	101	Audio	LP	Mixed	Yes	−0.12
Barron & Kysilka (1993)	Computer lit.	PS	60	Text	LP	Mixed	—	0.16
Barton & Dwyer (1987)	Physiology	PS	18	Text	SP	High	Yes	−0.84
Barton & Dwyer (1987)	Physiology	PS	34	Text	SP	High	Yes	0.24
Chu (2006)	Computer lit.	PS	112	Audio	SP	High	Yes	−0.15
Chu (2006)	Computer lit.	PS	112	Audio	SP	High	Yes	0.24
Craig et al. (2004), Exp. 1	Computer lit.	PS	30	Audio	LP	High	Yes	0.54
Craig et al. (2004), Exp. 1	Computer lit.	PS	30	Audio	LP	High	No	0.29
Craig et al. (2004), Exp. 1	Computer lit.	PS	30	Audio	SP	High	Yes	−0.29
Craig et al. (2004), Exp. 1	Computer lit.	PS	30	Audio	SP	High	No	0.16
Craig et al. (2002), Exp. 2	Meteorology	PS	48	Audio	LP	High	Yes	−0.58*
Craig et al. (2002), Exp. 2	Meteorology	PS	47	Text	LP	High	Yes	0.18
Diao et al. (2007)	Reading	PS	159	Audio	SP	High	No	0.35*
Diao & Sweller (2007)	Reading	PS	57	Text	SP	High	No	−0.50
Jamet & LeBohec (2007)	Memory model	PS	90	Audio	SP	High	Yes	−0.97*
Kalyuga et al. (1999), Exp. 1	Mechanical	PS	23	Text	LP	High	Yes	−0.22
Kalyuga et al. (1999), Exp. 1	Mechanical	PS	23	Audio	LP	High	Yes	−1.29*
Kalyuga et al. (2000), Exp. 1	Mechanical	PS	44	Text	LP	High	Yes	−0.41
Kalyuga et al. (2004), Exp. 3	Mechanical	PS	21	Audio	SP	High	No	−1.09*
Koroghlanian & Klein (2004)	Biology	8–12	103	Text	LP	Low	Yes	−0.15
Koroghlanian & Sullivan (2000)	Computer lit.	PS	134	Text	LP	High	Yes	−0.10
Lai (2000)	Computer lit.	PS	126	Text	SP	High	Yes	−0.05
Lai (2000)	Computer lit.	PS	127	Text	SP	High	Yes	0.56*
Leahy et al. (2003)	Meteorology	4–7	30	Text	SP	High	Yes	−0.94*
LeBohec & Jamet (2008)	Accounting	PS	123	Audio	SP	Mixed	Yes	−0.14
Linebarger (2001)	Reading lit.	K–3	76	Audio	SP	High	Yes	0.32
Markham (1999)	Social studies	PS	59	Audio	SP	High	Yes	0.71*
Markham (1999)	Biology	PS	59	Audio	SP	High	Yes	1.03*
Mayer et al. (2001), Exp. 1	Meteorology	PS	41	Text	SP	Low	Yes	−0.69*
Mayer et al. (2001), Exp. 1	Meteorology	PS	37	Text	SP	Low	Yes	−0.43
Mayer et al. (2001), Exp. 2	Meteorology	PS	109	Text	SP	Mixed	Yes	−0.89*
Mayer & Johnson (2008), Exp. 1	Meteorology	PS	90	Audio	SP	Low	Yes	0.25
Mayer & Johnson (2008), Exp. 2	Mechanical	PS	62	Audio	SP	Low	Yes	0.30
McNeill (2004)	History	PS	37	Audio	LP	High	No	0.45
McNeill (2004)	History	PS	37	Text	LP	High	No	0.44
Montali & Lewandowski (1996)	Reading	8–9	12	Text	LP	High	No	0.34
Montali & Lewandowski (1996)	Reading	8–9	12	Audio	LP	High	No	0.74
Montali & Lewandowski (1996)	Reading	8–9	12	Text	LP	High	No	0.72
Montali & Lewandowski (1996)	Reading	8–9	12	Audio	LP	High	No	0.90
Moreno & Mayer (2002a), Exp. 2	Botany	PS	28	Text	LP	High	Yes	0.54
Moreno & Mayer (2002a), Exp. 2	Botany	PS	28	Audio	LP	High	Yes	−0.15
Moreno & Mayer (2002b), Exp. 1	Meteorology	PS	38	Audio	SP	High	No	0.80*
Moreno & Mayer (2002b), Exp. 1	Meteorology	PS	36	Audio	SP	High	Yes	0.69*
Moreno & Mayer (2002b), Exp. 2	Meteorology	PS	34	Audio	SP	High	Yes	−0.85*
Moreno & Mayer (2002b), Exp. 2	Meteorology	PS	35	Audio	SP	High	Yes	0.35
Moreno & Mayer (2002b), Exp. 3	Meteorology	PS	35	Audio	SP	High	No	1.43*
Moreno & Mayer (2002b), Exp. 3	Meteorology	PS	36	Audio	SP	High	No	1.27*
Neuman& Koskinen (1992)	Science	7–8	69	Captioned text	SP	High	No	0.26
Olson & Wise (1992)	Reading lit.	2–6	90	Text	LP	Low	No	0.75*
Rehaag & Szabo (1995)	Mathematics	10	82	Text	LP	High	—	−0.01
Reitsma (1988)	Reading	1–2	37	Guided reading	SP	High	No	0.25
Ritzhaupt et al. (2008)	Computer lit.	PS	183	Audio	SP	Low	—	1.85*
Rockwell& Singleton (2007)	History	PS	132	Text	LP	High	No	−0.19
Severin (1967)	Nature	7	62	Audio	SP	High	No	0.41
Severin (1967)	Nature	7	67	Text	SP	High	No	0.07

Note. Dashes indicate that data were not reported. LP = learner-paced; SP = system-paced; lit. = literacy; PS = postsecondary; Exp. = experiment.
* $p < .05$.

studies variability (I^2). Table 2 shows the overall analysis of the weighted means of all statistically independent effect sizes as well as the results for comparison treatments.

Overall, Table 2 shows a small but statistically detectable effect of learning with spoken–written presentations. The overall sample was heterogeneous, $Q(56) = 304.93$, $p < .001$, $I^2 = .82$. The total variability that could be attributed to true heterogeneity or between-studies variability was 82%, indicating that 82% of the variance was between-studies variance (i.e., could be explained by study-level covariates) and 18% of the variance was within-study variance based on sampling error. Generally, this means that there was more variability among the individual effect sizes than would be expected for samples from a single population. Significant heterogeneity justifies examination of study features that may have acted as moderating variables responsible for the variability.

Table 2 shows that there is variability in the magnitude of effect sizes across comparison treatments, ranging from −0.04 to 0.29. Prior to conducting moderator analyses, we examined the effect of learning with spoken–written presentations in comparison with written-only or spoken-only presentations. Table 2 shows that students did not learn more from spoken–written than written-only presentations. However, spoken–written presentations were found to enhance learning when compared with spoken-only presentations.

The between-levels difference was statistically significant, indicating variation among the comparison treatments, $Q_B(1) = 22.77$, $p < .001$. Post hoc analysis revealed that using spoken-only comparison treatments was associated with effect sizes that were significantly greater than those associated with written-only comparison treatments. Considering the variability between studies that had written-only comparison and those with spoken-only comparison, we decided against collapsing the two comparison treatments together. Furthermore, because the overall mean effect size for studies that used written-only comparison was negligible ($g = -0.04$) and moderator analyses mostly produced small negligible effects, a decision was made not to report the results of the moderator analyses for written-only comparison in this article. The remainder of the Results section reports moderator analyses for spoken-only comparisons.

Analysis of Moderator Effects for Spoken-Only Comparisons

To determine under what conditions verbal redundancy may enhance or inhibit learning, results of moderator analyses were organized by presentational features, participant characteristics, contextual features, outcome constructs, test formats, and methodological features. Tables 3 to 7 show the results categorized according to study features or moderators that accounted for the variability in the overall findings for spoken-only comparisons.

Presentation features.

Pacing of presentation. Table 3 shows the weighted mean effect sizes for presentational features. There is some evidence that the type of pacing of presentation (self-paced or system-paced) may affect the way learners comprehend and learn with multimedia materials (Kalyuga et al., 2004; Nesbit & Adesope, 2011; Tversky, Morrison, & Betrancourt, 2002). A study was coded as system-paced if participants had no control over the pacing of the presentation or multimedia material. It was coded as learner-paced if participants were able to exercise a degree of control over the presentation of the material (e.g., pause, play, previous, next). Most of the studies were system-paced. System-paced studies with spoken–written presentations were associated with a statistically detectable effect size ($g = 0.34$). The learner-paced studies were not ($g = 0.12$). One plausible explanation for this result is that learners in system-paced conditions who misperceive information from one source may not have time to reread or relisten to the missed information. If the system-paced presentation uses both modalities, then information in one modality can be recovered from the other modality. On the other hand, when learners have full control over the pace of the materials, they may not benefit from spoken–written presentations because they are provided with affordances to navigate, pause, rewind, or slow down the pace of the written or spoken presentation for effective apprehension.

Degree of correspondence between audio and text. In Table 3, degree of correspondence was coded high when there was exact correspondence between the written text and spoken narration; that is, when the written text reproduced the entire narration word-for-word. Degree of correspondence was coded low when abridgements or key terms derived from spoken segments were used as written text.

Most of the studies used high degree of correspondence. Studies with low and high degrees of correspondence between spoken and written texts were associated with statistically detectable effect sizes ($g = 0.99$ and $g = 0.21$, respectively). Since the between-levels difference was significant, a post hoc analysis was conducted. The result showed that presentations with a low degree of correspondence were associated with a significantly higher weighted mean effect size than those with a high degree of corre-

Table 2
Weighted Mean Effect Sizes for Comparison Treatments

			Effect size			Test of heterogeneity			
Variable	*N*	*k*	*g*+	*SE*	95% CI	*Q*	*df*	*p*	I^2 (%)
All	3,452	57	0.15*	0.03	[0.08, 0.22]	304.93	56	< .001	81.64
Comparison treatment									
Text (on-screen) only	1,480	23	−0.04	0.05	[−0.14, 0.06]				
Speech (audio) only	1,972	34	0.29*	0.05	[0.20, 0.39]				
Between levels (Q_B)						22.77	1	< .001	

* $p < .05$.

Table 3
Weighted Mean Effect Sizes for Presentational Features

Variable	N	k	Effect size $g+$	Effect size SE	95% CI	Test of heterogeneity Q	df	p
Pacing of presentation								
Learner-paced	369	10	0.12	0.11	[−0.09, 0.32]			
System-paced	1,603	24	0.34*	0.05	[0.24, 0.44]			
Between levels (Q_B)						3.54	1	.06
Degree of correspondence between audio and text								
Low	335	3	0.99*	0.12	[0.76, 1.23]			
High	1,413	29	0.21*	0.05	[0.11, 0.32]			
Mixed low and high	224	2	−0.13	0.14	[−0.40, 0.14]			
Between levels (Q_B)						46.01	2	< .001
Materials segmented while being presented?								
No	655	7	0.53*	0.08	[0.37, 0.69]			
Yes	1,015	22	0.15*	0.06	[0.03, 0.28]			
Not reported	302	5	0.30*	0.12	[0.07, 0.53]			
Between levels (Q_B)						12.72	2	< .001
Images included in material?								
No	541	12	0.45*	0.09	[0.28, 0.63]			
Yes	1,248	21	0.06	0.06	[−0.05, 0.17]			
Not reported	183	1	1.85*	0.18	[1.51, 2.20]			
Between levels (Q_B)						97.96	2	< .001
Animation included in material?								
No	1,216	21	0.24*	0.06	[0.12, 0.35]			
Yes	573	12	0.06	0.08	[−0.10, 0.23]			
Not reported	183	1	1.85*	0.18	[1.51, 2.20]			
Between levels (Q_B)						86.85	2	< .001

* $p < .05$.

spondence. Our analysis suggests that students learned more when written texts were abridgements or key terms derived from spoken narrations (low degree of correspondence) rather than verbatim reproductions (high degree of correspondence).

Segmentation. The size of segments is the amount of text presented at one time. The size of segments varied considerably in the studies we analyzed. While some studies typically presented one sentence at a time, others presented several sentences at a time. Nonsegmented presentations displayed all the text at one time. Researchers have claimed that the size of text segments may influence the outcome of learning with spoken–written presentations (Kalyuga et al., 2004).

Table 3 shows that although both segmented and nonsegmented studies were associated with statistically detectable effect sizes, the nonsegmented presentations were associated with a larger effect size ($g = 0.53$) than were the segmented ones ($g = 0.15$). This result suggests that verbal redundancy is more beneficial when texts are not segmented. The between-levels difference was statistically significant, $Q_B(2) = 12.72$, $p < .001$. Post hoc analysis revealed that nonsegmented presentations were associated with a higher weighted mean effect size and were significantly different from segmented presentations. It may be that learners more easily miss portions of nonsegmented texts, especially when they cannot control the pace at which such texts are presented, and spoken–written presentations help learners recover misperceived portions of the texts.

Images and animation. Some studies included in this meta-analysis presented images or animation. We analyzed the effect of verbal redundancy moderated by inclusion or exclusion of images or animation in the learning materials. Table 3 shows that spoken–written presentations without images or animation were associated with statistically detectable effects ($g = 0.45$ and $g = 0.24$, respectively). When images or animation was included in the presentations, spoken–written presentations did not facilitate learning in comparison with spoken-only presentations ($g = 0.06$ and $g = 0.06$, respectively). Between-levels differences were statistically significant. Post hoc analyses revealed that achievement outcomes in studies that did not include images in spoken–written presentations were different from those that included images.

Participant characteristics.

Educational level, reading fluency, and prior domain knowledge. Table 4 presents analyses of participants' educational level, reading fluency, and prior domain knowledge as moderator variables. Most of the participants in this meta-analysis were postsecondary students. Spoken–written verbal presentations were associated with statistically detectable effect sizes for intermediate (Grades 4–7, $g = 0.53$), secondary (high school, $g = 0.82$), and postsecondary ($g = 0.26$) students. The between-levels difference for educational level was not statistically significant. This suggests that the effects of learning with spoken–written presentations are positive regardless of whether students are in intermediate, secondary, or postsecondary levels of education.

Although the studies we examined were not aimed at increasing the reading ability of participants, it is possible that reading fluency may moderate the achievement effects on spoken–written presentations. Statistically detectable effect sizes were found for participants who (a) were able to read fluently in their first language (L1, $g = 0.24$); (b) were early or nonfluent readers in their first language ($g = 0.82$); and (c) were second-

Table 4
Weighted Mean Effect Sizes for Participant Characteristics

Variable	N	k	Effect size $g+$	Effect size SE	95% CI	Test of heterogeneity Q	Test of heterogeneity df	Test of heterogeneity p
Educational level								
Primary (K to Grade 3)	76	1	0.32	0.23	[−0.13, 0.76]			
Intermediate (Grades 4–7)	179	3	0.53*	0.15	[0.23, 0.83]			
Secondary (Grades 8–12)	24	2	0.82*	0.39	[0.04, 1.59]			
Postsecondary	1,693	28	0.26*	0.05	[0.16, 0.36]			
Between levels (Q_B)						4.70	3	.20
Reading fluency								
Fluent L1	1,671	29	0.24*	0.05	[0.14, 0.34]			
Early L1 or nonfluent L1	24	2	0.82*	0.39	[0.04, 1.59]			
L2	277	3	0.58*	0.13	[0.33, 0.82]			
Between levels (Q_B)						7.73	2	.02
Prior domain knowledge								
Low	1,499	25	0.29*	0.05	[0.18, 0.39]			
High	157	2	−0.05	0.17	[−0.37, 0.28]			
Mixed	24	2	0.82*	0.39	[0.04, 1.59]			
Not reported	292	5	0.46*	0.12	[0.23, 0.69]			
Between levels (Q_B)						8.12	3	.05

Note. L1 = first language; L2 = second language.
* $p < .05$.

language (L2, $g = 0.58$) readers. The results for reading fluency suggest that although spoken–written presentations are beneficial for readers at different fluency levels, they are especially beneficial for second-language learners and early readers who are not fluent in their first language.

Most of the studies reported that their participants had low prior domain knowledge. The studies with low prior knowledge learners were associated with a statistically detectable effect size ($g = 0.29$), and studies with high prior knowledge learners were associated with a nonsignificant effect size ($g = -0.05$). The between-levels difference was marginally nonsignificant.

Contextual factors. Table 5 presents an analysis of contextual moderator variables, including the setting where the research was conducted (laboratory or classroom), as well as duration of the entire study and treatment. When learning activities were reported as part of a formal course of study or were conducted entirely in a classroom under the supervision of an instructor, the setting for such studies was coded as *classroom*. When learning activities were not reported as part of a formal course of study or were conducted in a laboratory, the setting was coded as *laboratory*. We observed a large variation in treatment duration for the verbal redundancy studies we analyzed. While many studies had treatment durations of less than 15 min, others had an hour or more of treatment. We used approximate median splits on duration of treatment (split at 1 hr) and duration of study (split at 1 week) to create two categories for each of these variables. Treatment duration was fully reported in all but four studies, and study duration was reported in all but three studies.

Table 5
Weighted Mean Effect Sizes for Contextual Features

Variable	N	k	Effect size $g+$	Effect size SE	95% CI	Test of heterogeneity Q	Test of heterogeneity df	Test of heterogeneity p
Setting								
Classroom	675	11	0.20*	0.08	[0.04, 0.35]			
Laboratory	1,297	23	0.36*	0.06	[0.23, 0.46]			
Between levels (Q_B)						2.26	1	.13
Treatment duration								
Less than 1 hour	1,266	24	0.13*	0.06	[0.02, 0.25]			
Greater than or equal to 1 hour	312	6	0.27*	0.12	[0.04, 0.50]			
Not reported	394	4	0.91*	0.11	[0.70, 1.13]			
Between levels (Q_B)						39.35	2	< .001
Study duration								
Less than or equal to 1 week	1,251	26	0.15*	0.06	[0.03, 0.26]			
Greater than 1 week	453	5	0.30*	0.10	[0.11, 0.49]			
Not reported	268	3	1.13*	0.14	[0.86, 1.40]			
Between levels (Q_B)						43.17	2	< .001

* $p < .05$.

Both classroom and laboratory studies were associated with statistically detectable effect sizes. The classroom studies produced a small effect size ($g = 0.20$), while the laboratory studies produced a near medium effect size ($g = 0.36$). However, the between-levels difference was not statistically significant.

A pattern showing higher mean effect sizes for longer treatment and study durations was observed. Treatments that were equal to or greater than 1 hr in length were associated with a higher statistically detectable effect size ($g = 0.27$) than were shorter duration treatments ($g = 0.13$). Similarly, studies conducted for over a week were associated with a higher mean effect size ($g = 0.30$) than were studies conducted for a week or less ($g = 0.15$). Post hoc analyses showed that studies that did not report the treatment and study durations were different from those that reported these study features. However, no differences were statistically detected between the long and short durations on either variable.

Outcome constructs and test formats. Table 6 presents the results of different outcome variables and test formats. The outcome variables were divided into three categories: retention, transfer, and mixed retention and transfer. The test formats were divided into two categories: objective formats (e.g., multiple choice items) and mixed formats (e.g., multiple choice, free recall, and short answer). Most of the studies used mixed retention and transfer outcomes. Retention outcomes and multiple choice objective-item test formats were associated with statistically detectable effect sizes. Between-levels differences were statistically significant for outcome variables and test formats. Post hoc analyses revealed that the retention-only outcome was associated with a significantly different and larger effect size ($g = 0.53$) than was the transfer-only outcome ($g = 0.01$) and mixed retention and transfer outcome ($g = 0.11$). Similarly, the multiple choice objective test format was associated with a significantly different and larger effect size ($g = 0.63$) than were mixed test formats ($g = 0.11$).

Methodological research features. Table 7 presents the effect size variations related to methodological quality of the research. This includes our rating of treatment fidelity of the studies examined; the reliability of measures that were used in the primary studies; and whether research was published in a journal, presented at a conference, or written as a dissertation. Most of the studies were published in journals and coded high in treatment fidelity. Mean effect sizes were statistically detectable irrespective of whether the treatment fidelity was coded as medium or high.

When the reliability of measuring instruments was reported, spoken–written presentations were associated with a moderate, statistically detectable mean effect size ($g = 0.63$). On the other hand, there was no statistically detectable effect of spoken–written presentations when reliability was not reported ($g = 0.06$). The between-levels difference was statistically significant. Post hoc analysis revealed that studies that reported the reliability of measures were significantly different from and produced a much larger effect size than those that did not report the reliability of the measures. Studies published in journals were associated with a statistically detectable mean effect size ($g = 0.30$), while nonpublished dissertations were not ($g = 0.10$). This result may be attributed to two issues: (a) the higher methodological quality of published research or (b) a possible publication bias favoring published studies that report statistically significant effect sizes. We now examine the extent to which publication bias constitutes a threat to the results of this meta-analysis.

Examining Publication Bias

Researchers have claimed that there is an inherent bias against publishing nonstatistically significant results in peer-reviewed journals (Orwin, 1983; Rosenthal, 1979) and that studies with nonstatistically significant findings are often not published or are reported in the less accessible gray literature (Lipsey & Wilson, 2001). This "file drawer" problem poses a threat to the validity of results by potentially skewing findings of a meta-analysis toward a more positive mean effect size. Researchers have developed several approaches for estimating publication bias in meta-analyses and have proposed that publication bias (or lack of it) is established through consistent results among the different methods (Ferguson, 2007; Rothstein, Sutton, & Borenstein, 2005).

In the current work, three approaches were computed with Comprehensive Meta-Analysis software to examine the potential existence of publication bias. First, the funnel plot (which reveals the estimates of the unbiased effect size compared with the standard error) shows a symmetrical distribution around the weighted mean effect. Symmetric funnel plots suggest the absence of publication bias (Duval & Tweedie, 2000; Song, Khan, Dinnes, &

Table 6
Weighted Mean Effect Sizes for Outcome Constructs and Test Formats

			Effect size			Test of heterogeneity		
Variable	*N*	*k*	*g*+	*SE*	95% CI	*Q*	*df*	*p*
Outcome variable								
Retention	940	14	0.53*	0.07	[0.39, 0.66]			
Transfer	160	2	0.01	0.16	[−0.31, 0.33]			
Mixed retention and transfer	872	18	0.11	0.07	[−0.03, 0.24]			
Between levels (Q_B)						21.84	2	< .001
Test format								
M/C objective	750	11	0.63*	0.08	[0.48, 0.79]			
Mixed item types	1,222	23	0.11	0.06	[−0.01, 0.22]			
Between levels (Q_B)						29.52	1	< .001

Note. M/C = multiple choice.
* $p < .05$.

Table 7
Weighted Mean Effect Sizes for Methodological Features

Variable	N	k	Effect size $g+$	SE	95% CI	Test of heterogeneity Q	df	p
Treatment fidelity								
Medium	48	1	1.16*	0.31	[0.55, 1.76]			
High	1,924	33	0.27*	0.05	[0.18, 0.37]			
Between levels (Q_B)						8.03	1	< .001
Reliability								
Reported	731	10	0.63*	0.08	[0.48, 0.78]			
Not reported	1,241	24	0.10	0.06	[−0.01, 0.22]			
Between levels (Q_B)						29.36	1	< .001
Document type		30	0.30*	0.05	[0.20, 0.40]			
Journal publication	1,663							
Conference paper	48	1	1.16*	0.31	[0.55, 1.76]			
Dissertation	261	3	0.10	0.12	[−0.14, 0.34]			
Between levels (Q_B)						10.32	2	.01

* $p < .05$.

Sutton, 2002). Second, Egger's linear regression test (Egger, Smith, Schneider, & Minder, 1997) was used to more fully investigate the results of the funnel plot through an examination of the unbiased effect sizes and standard errors. Results of this test further corroborated the result of the funnel plot, clearly showing the absence of publication bias ($p = .51$). Third, a "Classic Fail-safe N" test was performed to determine the number of null effect studies needed to raise the p value associated with the average effect above an arbitrary alpha level (set at $\alpha = .05$). This test revealed that 236 additional qualified studies would be required to nullify the overall effect size found in this meta-analysis.

The consistent results of these three tests suggest the absence of a level of publication bias that would pose a significant threat to the validity of the overall result of this meta-analysis.

Discussion

Although, overall, a statistically detectable learning benefit for verbal redundancy was found, the small size of this effect and its large heterogeneity render it theoretically and educationally inconsequential. A meaningful interpretation of verbal redundancy effects can only be made by considering moderating conditions including pacing of presentation, degree of correspondence between audio and text, reading fluency of learners, and inclusion of images or animation. Specifically, this meta-analysis provided answers to the following research questions.

What Are the Effects of Learning From Concurrent Spoken–Written Presentations in Comparison With Spoken-Only and Written-Only Presentations?

The meta-analysis revealed that in comparison with written-only presentations, verbally redundant, spoken–written presentations did not facilitate learning. On the other hand, verbally redundant spoken–written presentations were associated with enhanced learning in comparison with spoken-only presentations. This result suggests that adding text to an audio narration is beneficial but that adding audio narration to text is not. Perhaps learners are more likely to misperceive speech elements in a fixed-pace presentation than misread text, and supplementing speech presentations with text may help them to recover from failed speech perception.

How Do Verbal Redundancy Effect Sizes Vary With Different Educational Levels, Participants' Reading Fluency, and Prior Domain Knowledge?

The meta-analysis revealed that, compared with spoken-only presentations, spoken–written presentations were associated with enhanced learning for students in intermediate (Grades 4–7), secondary (high school), and postsecondary levels. The effects of verbal redundancy were moderated by individual differences related to the reading fluency of learners. Although spoken–written presentations tended to benefit learners at all levels of reading fluency, we found moderate and large effect sizes for L2 ($g = 0.58$) and nonfluent L1 ($g = 0.82$) learners but only a small effect size for fluent L1 learners ($g = 0.24$). It may be that L2 and, somewhat paradoxically, less fluent L1 readers are able to use written content to recover from failed comprehension of spoken content. The meta-analysis also revealed that studies conducted with low prior knowledge learners found advantages for spoken–written presentations ($g = 0.29$), while studies with high prior knowledge learners did not ($g = -0.05$).

How Are Verbal Redundancy Effects Moderated by Pacing of Presentation, Degree of Correspondence Between Audio and Text, Segment Size, and the Presence or Absence of Images in the Learning Materials?

Most of the studies in this meta-analysis used system-paced learning materials. The system-paced studies produced a statistically detectable effect size favoring spoken–written presentation in comparison with spoken-only presentations. Collectively, the learner-paced studies did not find statistically detectable benefits associated with spoken–written presentations. Tversky et al.'s (2002) apprehension principle may be useful in explaining these findings. According to these authors, multimedia learning resources "must be slow and clear enough for observers to perceive movements, changes, and their timing" (p. 258). They suggested that this can be accomplished by allowing learners to stop, start, pace, and review multimedia learning materials. When learners can exercise control over the pace at which multimedia materials are presented, they may not need redundant speech and text to cognitively process the materials effectively because they can review either a written-only or spoken-only presentation until they have fully understood it.

An important finding of this meta-analytic review is that presentations that use written key terms from spoken narrations are associated with greater learning than are fully redundant spoken–written presentations. One reason why less correspondence between audio and text (partially redundant presentations) might be beneficial for learning is that such presentations may focus students' attention on key terms and more efficiently cue retrieval and formation of appropriate schemas. From the perspective of multimedia learning theory, partially redundant presentations may signal important concepts (Mayer & Johnson, 2008). Signaling is used "to guide the search for specific information and to simplify decisions the reader may have to make about which information is relevant" (Mautone & Mayer, 2001, p. 378). From the perspective of the multiple external representation model advanced by Ainsworth (1999), visually presenting key terms during an audio narration may constrain or guide the interpretation students place on the narrated content.

The meta-analysis revealed that verbal redundancy was not beneficial when the materials included concurrent pictorial components such as diagrams and animation. Studies of the split-attention effect, such as those found in verbal redundancy studies included in this meta-analysis, have tended to use concurrently presented diagrams or animation that are intelligible without reference to text or speech elements. We speculate that verbal redundancy did not help in studies that presented diagrams or animation because the need to use text to correct a failed perception of the audio narration was obviated by the presence of those pictorial components. Learners who failed to perceive speech could always refer to the pictorial component to understand what was said, and the addition of redundant text offered no benefit.

How Do Contextual Features, Outcome Constructs, and Test Formats Affect Learning From Spoken–Written Presentations?

The meta-analysis revealed that studies conducted in the classroom and those conducted in the laboratory produced statistically detectable effect sizes. Furthermore, the analysis revealed a pattern showing higher mean effect sizes for longer treatment and study duration. The use of retention outcomes and multiple choice objective test formats were associated with statistically detectable effect sizes.

How Are Verbal Redundancy Effects Influenced by Methodological Features of the Research?

Treatment fidelity was coded either medium or high. Across both medium and high treatment fidelity, the meta-analysis revealed that spoken–written presentations were associated with statistically detectable effects. Further results showed that spoken–written presentations were associated with a moderate, statistically detectable mean effect size when reliability of measures used was reported. Studies published in journals produced a statistically detectable mean effect size, while dissertation studies did not. This result could be attributed to the higher methodological quality of published research.

Limitations and Future Research Directions

While evidence was found for factors that moderated the overall effect of learning with spoken–written presentations in comparison with spoken-only presentations, there is a need for further research to refine the understanding of these moderating effects. Specifically, more studies are needed on the effects of partial redundancy and its uses as a method for signaling key concepts. We also note that the results of this meta-analysis do not directly apply to using multimedia resources for teaching reading, as we included only one study that investigated the effects of verbal redundancy in primary or elementary grades.

Verbally redundant, spoken–written learning resources may be especially useful in the presentation of accessible materials for disabled learners who have minimal to moderate visual, auditory, or cognitive impairments that restrict their ability to process verbal information. Although none of the research included in this meta-analysis examined the use of spoken–written presentations for disabled learners, it seems that an examination of verbal redundancy in relation to the use of assistive technology devices (e.g., screen readers that offer text-to-speech functionality for disabled learners) would be illuminating.

Most of the studies included in this meta-analysis were conducted with undergraduate university students. There is a limited understanding of the learning effectiveness of redundant speech and text for students in K–12. More high quality studies on the effects of verbal redundancy with this group of learners are required. We also note the brevity of the instructional treatments, as many of the studies relied on treatment durations of under an hour. Investigations with longer instructional treatments as well as with students of varied verbal abilities would enhance the generalizability of verbal redundancy research.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80,* 207–245. doi:10.3102/0034654310368803

Adesope, O. O., & Nesbit, J. C. (2005). Toward accessible learning resources. In G. Richards (Ed.), *Proceedings of World Conference on e-Learning in Corporate, Government, Healthcare, and Higher Education 2005* (pp. 1802–1807). Chesapeake, VA: Association for the Advancement of Computing in Education.

Ainsworth, S. E. (1999). The functions of multiple representations. *Computers & Education, 33,* 131–152. doi:10.1016/S0360-1315(99)00029-9

*Atkinson, R. K., Clark, D., Harrison, C., Koenig, A., & Ramirez, F. (2007). *Multimedia learning environments: Do the redundancy and modality principles apply to English language learners?* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20,* 287–298. doi:10.1002/acp.1245

Baddeley, A. (1992, January 31). Working memory. *Science, 255,* 556–559. doi:10.1126/science.1736359

*Barron, A. E. (1994). Audio instruction in multimedia education: Is textual redundancy important? *Journal of Educational Multimedia and Hypermedia, 3,* 295–306.

*Barron, A. E., & Kysilka, M. L. (1993). The effectiveness of digital audio in online-based training. *Journal of Research on Computing in Education, 25,* 277–289.

*Barton, E. A., & Dwyer, F. M. (1987). The effect of audio redundancy on students' ability to profit from printed-verbal/visualized instruction. *International Journal of Instructional Media, 14,* 93–98.

Bernard, R. M., Abrami, P. C., Lou, Y., & Borokhovski, E. (2004). A methodological morass: How can we improve the quality of quantitative research in distance education? *Distance Education, 25,* 175–198. doi:10.1080/0158791042000262094

Bernard, R. M., Abrami, P. C., Wade, A., Borokhovski, E., Tamim, R., Surkes, M. A., & Bethel, E. C. (2009). A meta-analysis of three interaction treatments in distance education. *Review of Educational Research, 79,* 1243–1289. doi:10.3102/0034654309333844

Bobis, J., Sweller, J., & Cooper, J. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction, 3,* 1–21. doi:10.1016/S0959-4752(09)80002-9

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. R. (2008). Comprehensive Meta Analysis (Version 2. 2.048). Englewood, NJ: Biostat.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293–332. doi:10.1207/s1532690xci0804_2

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62,* 233–246. doi:10.1111/j.2044-8279.1992.tb01017.x

*Chu, S. (2006). *Investigating the effectiveness of redundant text and animation in multimedia learning environments* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (AAT 3210347)

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3,* 149–210. doi:10.1007/BF01320076

Cooper, H. M., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis.* New York, NY: Russell Sage Foundation.

*Craig, S. D., Driscoll, D., & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents. *Journal of Educational Multimedia and Hypermedia, 13,* 163–183.

*Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94,* 428–434. doi:10.1037/0022-0663.94.2.428

*Diao, Y., Chandler, P., & Sweller, J. (2007). The effect of written text on learning to comprehend spoken English as a foreign language. *American Journal of Psychology, 120,* 237–261.

*Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction, 17,* 78–88. doi:10.1016/j.learninstruc.2006.11.007

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel plot based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455–463. doi:10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., & Phillips, A. N. (1997). Meta-analysis: Principles and procedures. *British Medical Journal, 315*(7121), 1533–1537.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple graphical test. *British Medical Journal, 315*(7109), 629–634.

Ferguson, C. J. (2007). Evidence for publication bias in video game violence effects literature: A meta-analytic review. *Aggression and Violent Behavior, 12,* 470–482. doi:10.1016/j.avb.2007.01.001

Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labeling, and spatial proximity. *Learning and Instruction, 20,* 216–224. doi:10.1016/j.learninstruc.2009.02.021

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128. doi:10.2307/1164588

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558. doi:10.1002/sim.1186

Huedo-Medina, T. B., Sánchez-Meca, F., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: I^2 or Q statistic? *Psychological Methods, 11,* 193–206. doi:10.1037/1082-989X.11.2.193

*Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology, 32,* 588–598. doi:10.1016/j.cedpsych.2006.07.001

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review, 23,* 1–19. doi:10.1007/s10648-010-9150-7

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1–17. doi:10.1518/001872098779480587

*Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13,* 351–371. doi:10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6

*Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92,* 126–136. doi:10.1037/0022-0663.92.1.126

*Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46,* 567–581. doi:10.1518/hfes.46.3.567.50405

Kirschner, P. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction, 12,* 1–10. doi:10.1016/S0959-4752(01)00014-7

Kirschner, P., Ayres, P., & Chandler, P. (2011). Contemporary cognitive load theory research: The good, the bad and the ugly. *Computers in Human Behavior, 27,* 99–105. doi:10.1016/j.chb.2010.06.025

*Koroghlanian, C. M., & Klein, J. (2004). The effect of audio and animation in multimedia instruction. *Journal of Educational Multimedia and Hypermedia, 13,* 23–46.

*Koroghlanian, C. M., & Sullivan, H. J. (2000). Audio and text density in computer-based instruction. *Journal of Educational Computing Research, 22,* 217–230. doi:10.2190/5QVQ-3XXQ-X1JC-T2F2

*Lai, S. (2000). Increasing associative learning of abstract concepts through audiovisual redundancy. *Journal of Educational Computing Research, 23,* 275–289. doi:10.2190/XKLM-3A96–2LAV-CB3L

*Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentation should and should not be a component of multimedia instruction. *Applied Cognitive Psychology, 17,* 401–418. doi:10.1002/acp.877

*Le Bohec, O., & Jamet, E. (2008). Levels of verbal redundancy, note-taking and multimedia learning. In J.-F. Rouet, R. Lowe, & W. Schnotz (Eds.), *Understanding multimedia documents* (pp. 79–101). New York, NY: Springer.

Lewandowski, L. J., & Kobus, D. A. (1993). The effects of redundancy in bimodal word processing. *Human Performance, 6,* 229–239. doi:10.1207/s15327043hup0603_3

*Linebarger, D. L. (2001). Learning to read using television: The effects of captions and narration. *Journal of Educational Psychology, 93,* 288–298. doi:10.1037/0022-0663.93.2.288

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Applied Social Research Methods Series Vol. 49). Thousand Oaks, CA: Sage.

*Markham, P. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals, 32,* 321–328. doi:10.1111/j.1944-9720.1999.tb01344.x

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93,* 377–389. doi:10.1037/0022-0663.93.2.377

Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York, NY: Cambridge University Press.

Mayer, R. E. (2005b). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 183–200). New York, NY: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82,* 715–726. doi:10.1037/0022-0663.82.4.715

*Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93,* 187–198. doi:10.1037/0022-0663.93.1.187

*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

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90,* 312–320. doi:10.1037/0022-0663.90.2.312

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction, 12,* 107–119. doi:10.1016/S0959-4752(01)00018-4

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43–52. doi:10.1207/S15326985EP3801_6

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43,* 31–43. doi:10.1007/BF02300480

*McNeill, A. L. (2004). *The effects of training, modality, and redundancy on the development of a historical inquiry strategy in a multimedia learning environment* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (AAT 3142239)

*Montali, J., & Lewandowski, L. (1996). Bimodal reading: Benefits of a talking computer for average and less skilled readers. *Journal of Learning Disabilities, 29,* 271–279. doi:10.1177/002221949602900305

Moreno, R., & Mayer, R. E. (1999). Visual presentations in multimedia learning: Conditions that overload visual working memory. In D. P. Huijsmans & A. W. M. Smeulders (Eds.), *Lecture notes in computer science: Visual information and information systems* (pp. 793–800). Berlin, Germany: Springer-Verlag.

*Moreno, R., & Mayer, R. E. (2002a). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94,* 598–610. doi:10.1037/0022-0663.94.3.598

*Moreno, R., & Mayer, R. E. (2002b). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94,* 156–163. doi:10.1037/0022-0663.94.1.156

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation models. *Journal of Educational Psychology, 87,* 319–334. doi:10.1037/0022-0663.87.2.319

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76,* 413–448. doi:10.3102/00346543076003413

Nesbit, J. C., & Adesope, O. O. (2011). Learning from animated concept maps with concurrent audio narration. *Journal of Experimental Education, 79,* 209–230. doi:10.1080/00220970903292918

*Neuman, S., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning in context for language minority students. *Reading Research Quarterly, 27,* 94–106. doi:10.2307/747835

*Olson, R. K., & Wise, B. (1992). Reading on the computer with orthographic and speech feedback: An overview of the Colorado Remediation Project. *Reading and Writing, 4,* 107–144. doi:10.1007/BF01027488

Orwin, R. G. (1983). A Fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8,* 157–159. doi:10.2307/1164923

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38,* 1–4. doi:10.1207/S15326985EP3801_1

Paas, F. Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32,* 1–8. doi:10.1023/B:TRUC.0000021806.17516.d0

Paivio, A. (1986). *Mental representations.* New York, NY: Oxford University Press.

Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior, 27,* 5–10. doi:10.1016/j.chb.2010.05.006

Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory.* New York, NY: Cambridge University Press.

*Rehaag, D., & Szabo, M. (1995). *An experiment on effects of redundant audio in computer based instruction on achievement, attitude and learning time in 10th grade mathematics.* Paper presented at the annual meeting of the Association for Educational Communications Technology, Anaheim, CA.

*Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-while-listening, and independent reading with computer-based speech feedback. *Reading Research Quarterly, 23,* 219–235. doi:10.2307/747803

Rey, G. D., & Buchwald, F. (2011). The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied, 17,* 33–48. doi:10.1037/a0022243

Ritzhaupt, A. D., Barron, A. E., & Kealy, W. A. (2011). Conjoint processing of time-compressed narration in multimedia instruction: The effects on recall, but not recognition. *Journal of Educational Computing Research, 44,* 203–217. doi:10.2190/EC.44.2.d

*Ritzhaupt, A. D., Gomes, N. D., & Barron, A. E. (2008). The effects of time-compressed audio and verbal redundancy on learner performance and satisfaction. *Computers in Human Behavior, 24,* 2434–2445. doi:10.1016/j.chb.2008.02.017

*Rockwell, S. C., & Singleton, L. A. (2007). The effect of the modality of presentation of streaming multimedia on information acquisition. *Media Psychology, 9,* 179–191. doi:10.1080/15213260709336808

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641. doi:10.1037/0033-2909.86.3.638

Rothstein, H., Sutton, A., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments.* Hoboken, NJ: Wiley. doi:10.1002/0470870168

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development, 53,* 47–58. doi:10.1007/BF02504797

*Severin, W. J. (1967). Another look at cue summation. *AV Communication Review, 15*, 233–245.

Song, F., Khan, K. S., Dinnes, J., & Sutton, A. J. (2002). Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *International Journal of Epidemiology, 31,* 88–95. doi:10.1093/ije/31.1.88

Sweller, J. (1999). *Instructional design in technical areas.* Camberwell, Victoria, Australia: ACER Press.

Sweller, J. (2005a). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York, NY: Cambridge University Press.

Sweller, J. (2005b). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–167). New York, NY: Cambridge University Press.

Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296. doi:10.1023/A:1022193728205

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Does it facilitate? *International Journal of Human-Computer Studies, 57,* 247–262. doi:10.1006/ijhc.2002.1017

van Merriënboer, J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17,* 147–177. doi:10.1007/s10648-005-3951-0

Wald, M. (2008). Learning through multimedia: Speech recognition enhancing accessibility and interaction. *Journal of Educational Multimedia and Hypermedia, 17,* 215–233.

Zhu, L., & Grabowski, B. (2006). Web-based animation or static graphics: Is the extra cost of animation worth it? *Journal of Educational Multimedia and Hypermedia, 15,* 329–347.

Received April 29, 2010
Revision received October 4, 2011
Accepted October 4, 2011 ■