Article Critique:

The Effects of Diagrams and Time-Compressed Instruction on Learning and Learners' Perceptions of Cognitive Load

Colin M. Gray

Department of Instructional Systems Technology

School of Education

Indiana University Bloomington

December 13, 2010

Abstract

The article *The Effects of Diagrams and Time-Compressed Instruction on Learning and Learners' Perceptions of Cognitive Load* by Raymond S. Pastore is critiqued using the framework provided by the American Educational Research Association (AERA). An extended summary of the original article is provided, followed by description through eight independent rating categories. Each rating category is supported by a minimum of two criteria, which support and define the rating assessment. Article appropriateness and applicability to the field is discussed.

Article Critique: The Effects of Diagrams and Time-Compressed Instruction on Learning and Learners' Perceptions of Cognitive Load

Summary

Overview

In this study, Pastore (2009) examines the effect of time-compressed audio on learning, and related effects of this compression on cognitive load. Through research by Mayer (2005) and Foulke and Sticht (1967), Pastore postulates that moderately compressed instruction will result in similar learning outcomes as compared to uncompressed instruction, while more extreme compression will result in decreased learning due to increased cognitive load. This study examines participants in a learning scenario, assigned to one of six experimental treatments randomly, using a quasi-experimental research model.

Theoretical Framework

This study utilizes three primary theoretical frameworks within the scope of multimedia instruction: cognitive load theory, multiple external representations, and time-compressed instruction. As a composite, these three strands of research serve as a foundation for the research questions and methods employed in this study. Cognitive load theory is the primary predictive framework used in this study, based on early work by Chandler and Sweller (1991) and DeLeeuw and Mayer (2008). This theory refers to the interworking of short-term memory, widely understood to have the capacity of seven information units, plus or minus two units (Miller, 1956). This work, in tandem with the information-processing model by Brunken *et al.* (2003) suggests that while an individual's working memory is limited, their long-term memory is unlimited. In this research study, the compression of time would be considered an extraneous

form of cognitive load, in contrast to the intrinsic and germane load, which are affected by element interactivity and schema development, accordingly.

The research on multiple external representations of instruction affects the information processing of the learning event, but using a different lens to explain the learning phenomena. In this study, the Dual Coding Theory (DCT) by Paivio and Csapo (1969), Paivio (1979), and Clark and Paivio (1991) elucidates the functioning of parallel auditory and visual channels of processing, concluding that each channel processes information independently yet in an interconnected way. The research of Mayer and Sims (1994) extends this theory, to form the Cognitive Theory of Multimedia Learning (CTML), which assumes that dual channel-based learning will result in superior instruction, as compared to a single channel-based instructional event (Pastore, 2009).

Finally, the research on time-compressed instruction reveals a certain tolerance by learners to understand and comprehend compressed audio in the context of a learning event. Research by Orr (1968) indicates that learners can understand and process speech within certain compression parameters, followed by the research of Heiman *et al.* (1986), which reveals that comprehension began to degrade at a compression rate of 50%.

Research Questions

Based on the extant research on time-compressed instruction, DCT, and cognitive load theory, Pastore (2009) studies what effect time-compression and use of visual devices has on student achievement and perceptions of cognitive load. Because of the independency of channels in DCT, Pastore (2009) notes that "presenting verbal representations as audio rather than text in a multimedia environment better supports learning" (p.489). Subsequent to this conclusion, Pastore expects that achievement will be negatively correlated to audio compression, assuming

that participants in the 50% compression experimental group will experience the lowest academic achievement, based on the research of Heiman *et al.* (1986). Based on the research of Mayer and Moreno (2003), learners assigned to low or moderate compression experimental groups should experience lower perception of cognitive load than their counterparts in the 50% compression group.

Method

This study utilized a convenience sample of participants in a university environment. The sample was composed of 216 undergraduate education majors, including a range of academic classifications from freshman to senior. The majority of the participants were 18-22 years old, and the sample was comprised of 64% females and 36% males.

The participants were randomly assigned to one of six experimental groups, including: 0% compression with visuals, 0% compression with no visuals, 25% compression with visuals, 25% compression with no visuals, 50% compression with visuals, and 50% compression with no visuals. Instructional material used in this research study was developed by Dwyer (1965), including a script regarding the physiology of the human heart. This script was later adapted by Dwyer and Lamberski (1983) to include 19 line drawings with color-coded regions to coincide with existing instructional concepts. This instruction was adapted for a multimedia learning environment using Adobe Flash CS3, resulting in 21 static pages of content. Review behavior, consisting of the learner's opportunity to replay the content or return to a previous slide through the use of a back button, was allowed in the user environment, and results were captured for further data analysis. The audio was compressed to the 25% and 50% level using the Audacity sound editor, resulting in audio of a shorter tempo with unaltered pitch.

A number of instruments were used in this study, including: an initial demographic questionnaire, prior knowledge pretest, knowledge posttests, a posttest questionnaire, and a drawing test. The prior knowledge pretest consisted of general factual recall within the field of physiology, and included 36 multiple-choice questions. This instrument was also used to ensure that prior knowledge of participants was equivalent across all available experimental groups. The posttest questionnaire included a single Likert question assessing the perceived cognitive load of the student prior to the posttest evaluation instrument. The knowledge posttest instrument consisted of four internal tests: drawing, identification, terminology, and comprehension. These scales were found by Dwyer (1978) to have acceptable Kuder-Richardson Formula 20 (KR-20) reliability scores of .83, .81, .83, and .70, respectively.

The research study was conducted over two regular class sessions, with the first session containing a demographic survey and prior knowledge test. At the second class, each participant was randomly assigned to one of the six possible experimental conditions. Following the conclusion of the learning intervention, each participant, regardless of experimental group, was given a final questionnaire and the drawing test. Following the completion of these instruments, the participants were given a knowledge posttest.

Results

Data from this research study were analyzed using SPSS statistical software, using a 2×3 ANOVA to test the pretest instrument, followed by a 2×3 MANOVA to evaluate the four variables present in the posttest. Finally, cognitive load and review of instruction variables were evaluated using a 2×3 ANOVA.

The evaluation of the prior knowledge pretest using a 2×3 ANOVA revealed no significant differences in visual or compression groups. Descriptive statistics were performed on

this data, including mean, standard deviation, and frequency across all six experimental conditions. According to Pastore (2009), "there were not any significant violations of the assumptions of independent observations, normal distribution, or homogeneity of variance" in this data set (p. 495).

The knowledge posttests were evaluated using a 2×3 MANOVA, to determine the effect of the visual and compression independent variables on the learning scales present in the knowledge posttest. These learning scales include: drawing, identification, terminology, and comprehension. All scales were evaluated using the KR-20 reliability test, and were found to be adequately reliable, including scores of .91, .81, .76, and .70, respectively. Prior to the assessment of MANOVA on the six treatment groups, each learning variable was assessed by Pastore (2009) for "normality, homogeneity of variance-covariance matrices, linearity, and multicollinearity and singularity," finding no major violation of assumptions (p.495). Descriptive statistics were collected on all treatment groups and learning variables. Using Levene's test, it was found that the assumption of equal variances was not met by the data from the drawing ($p < .001$), terminology ($p < .001$), and comprehension ($p = .033$) tests, although the identification test ($p = .392$) did meet this assumption. Based on these results, the alpha level was changed to .025 rather than .05, as suggested by Tabachnick and Fidell (2007). The results of the 2×3 MANOVA indicate a significant difference between experimental groups. Significance was found on the independent variables visual and compression, although the visual by compression interaction was not statistically significant.

Univariate analysis was conducted to examine the effect of each independent variable on the knowledge subtests. The comprehension subtest resulted in a significant interaction effect of visuals and compression, while no other subtest has a significant measured interaction effect.

There was a significant effect of compression on the drawing ($F(2, 210) = 13.444$), identification ($F(2, 210) = 11.273$), and terminology ($F(2, 210) = 5.042$) subtests in the knowledge posttest. Tukey's Honestly Significance Difference test (HSD) post hoc analysis was used to determine what elements led to the difference in effect. Tukey's HSD revealed the 0% and 25% compression groups for both the drawing and identification subtests were not significantly different when comparing group means, but participants in both of these groups were significantly higher than those in the 50% compression group. Within the terminology subtest, statistically significant differences were found between the 0% and 50% compression groups, but not between the 25% compression and any other group.

A 2×3 ANOVA was conducted to examine the role of cognitive load in the research study. Descriptive statistics were performed, and no major violations of ANOVA assumptions were present. A significant effect for the visual condition ($F(1, 2010) = 31.376$) was found, and a lower effect was found for the effect of compression on cognitive load ($F(2, 210) = 11.04$). No significant effect was found for the compression and visual interaction. Using Tukey's HSD, participants in the 0% and 25% compression groups perceived less cognitive load than did participants in the 50% compression group.

Finally, a 2×3 ANOVA was conducted to evaluate the effect of the independent variables on review within the learning intervention. No significant effect or interaction was found between review behaviors and visuals or compression.

Limitations

Pastore (2009) mentioned several limitations to this study, including: lack of user control, limited generalizability due to research study population, and the role of the designer in the creation of compressed materials. Based on the results of the research study, participants may

want the ability to set the pace of instruction in reaction to the type and quality of content being presented. Further, the research study participants were relatively homogenous in background and age, and additional research in other cultures and learner populations may provide different data. Finally, Pastore (2009) mentions "it will be up to the designer to decide where time compression is appropriate based on the conclusions outlined in this paper" (p.502).

Critique

Choice of Problem/Topic (Rating = 4)

This research study focuses on the intersection of multimedia learning, DCT, cognitive load theory, and time-compression of audio. While each of these elements has an independently comprehensive collection of research to justify its conclusions, the combination of frameworks could resolve issues not yet anticipated in the research literature.

From a pragmatic standpoint, Pastore notes the need for instruction to be completed in less time, as other business or education needs dictate. While this goal is somewhat consumerist, the goal of completing more instruction in less time does fit the dictum of instructional design—making instruction more efficient, effective, or engaging (Merrill, 2008). Further, it is possible that this goal of making learning more efficient from a time standpoint may also increase engagement, through the use of multimedia.

This topic is also important from a research perspective, in that this research study synthesizes four topic areas that have no precedent in the research literature. While much is known about each of these frameworks in isolation, no sound predictions might be made in terms of effect without comprehensive data. Also, from the literature cited in Pastore's study, it appears that time-compression in instruction has not been actively studied in recent years, with the

exception of a study by He and Gupta (2001). Additional research within the domain of modern audio editing capabilities may yield helpful results due the availability and affordance of this technology.

Theoretical Framework (Rating = 5)

This study references the contributions of four intersecting frameworks—Cognitive Load Theory, Dual Coding Theory (DCT), Cognitive Theory of Multimedia Learning (CTML), and Time-Compressed Instruction—all of which were adequately described both in isolation, and in reference to the other frameworks grounding the study. Each framework referenced and informed specific aspects of the research design and research questions, serving as a helpful guide to making informed predictions, or hypotheses, regarding the results of the study.

Cognitive load theory grounded the discussion of information processing, and the introduction of additional environmental constraints within the learning intervention were discussed within the framework of extraneous load. While extraneous load was decreased in certain experimental groups due to the use of visual imagery, the addition of time-compressed narration added extraneous load to other experimental group permutations.

The use of multiple external representations of learning content has been studied through a number of lenses, and is discussed in this research study in the context of DCT by Paivio and the CTML by Mayer. Through a synthesis of these theories, Paivio (1979) predicts that a learner can process auditory and visual stimuli independently, while Mayer and Sims (1994) predict that learning outcomes in a dual channel presentation of content will be higher than in a single channel presentation.

Time-compressed instruction is foundational to this study, in regard to research design and assignment of experimental groups. While existing research in the implementation of time-

compressed instruction may serve as a cursory guide of its effect on an instructional environment, correlating this effect with the modern conception of multimedia learning serves to guide future research in each area. The predictions within this research study indicate a compilation of all four frameworks, and Pastore (2009) concludes that "cognitive load is decreased when participants are presented with a multimedia presentation rather than a single media representation" (p.489).

Methods (Rating = 4)

Although the sample was well constructed within the research context, the use of a convenience sample to construct the sample pool inherently limits the generalizability of results. More positively, the use of random assignment within the six available experimental conditions, consistent with a standard quasi-experimental model, in association with a moderately sized sample, increases the statistical power of this research study.

While sampling in the general population is often more difficult, a broader cross-section of participants increases the generalizability of study results and decreases the chance of attributes surrounding the convenience sample, known or unknown, from conflicting with the independent variable assigned in the research design. In this study, the accessible population may not adequately represent the target population for the study. As Fraenkel and Wallen (2009) note: "the more narrowly researchers define the population, the more they save on time, effort, and (probably) money, but the more they limit generalizability" (Location 2812-2828).

The use of random assignment allows for greater generalizability of the research study results within highly-similar populations, and the sample size allows for greater statistical power than might be found in a smaller study. The issue of generalizability is important to consider, as the representativeness of the accessible population as compared to the target population affects

the applicability of study results and conclusions to a population beyond the initial study. According to Fraenkel and Wallen (2009), "generalization is made more plausible if data are presented to show that the sample is representative of the intended population on at least some relevant variables" (Location 3098-3117). While Pastore notes the nature of the sample as convenience, and possible purposive (due to the targeted nature of education students), the definition of the study's target and accessible population is not explicitly defined, resulting in limited applicability beyond this particular educational context.

Data Source(s) (Rating = 4)

A wide range of instruments are used to collect data for this research study, ranging from review data collection within the learning environment to pre- and post-test measures, to a drawing test and attitudinal data collection. Through these data sources, a wide variety of data is represented, but the data collection may not be sufficiently focused to allow for direct application to the research questions originally proposed by this research study.

Although many instruments are utilized for data collection in this study, many of the conclusions are based on data derived from the knowledge posttest. Pastore (2009) notes the reliability of the subtests contained within the knowledge posttest, and the reported KR-20 reliability measures exceed the averages from Dwyer's (1978) original averages for these scales. As noted by Fraenkel and Wallen (2009), "a reliable instrument is one that gives consistent results," and this instrument has been proven over multiple research studies to produce reliably consistent results (Location 3300-3319).

The validity of the knowledge posttest is potentially more difficult to assess. While Pastore notes the reliability of this instrument and its development by Dwyer, no indication in terms of face or external validity is provided. While the data is reliable, as indicated by the KR-

20 results, the targeting of these questions to "measure what it is supposed to measure" is important to assure that inferences can be made from the resulting posttest data (Fraenkel and Wallen, 2009, Location 3300-3319). Further, Trochim and Donnelly (2008) note that threats to validity can be minimized through logical argumentation or measurement and observation.

Conclusions/Interpretations (Rating = 4)

In general, the hypotheses suggested by Pastore at the conclusion of the literature review were validated by the results of the data collection and subsequent analysis. In particular, compression was found to have little impact on learning outcomes at the 0% and 25% levels, while 50% compression significantly curtailed learning achievement. Another area of concern may be implementation of these findings in a designed learning environment, and potentially negative affects on the learner if instructional designers do not utilize reasonable compression guidelines.

While compression was found to be limited in its effect on learning at the 0% and 25% levels, the disparity in data between the 25% and 50% levels causes the actual level at which extraneous load overwhelms the learner to be difficult to ascertain. While the 25% level provides a helpful practical limit for compression, additional granularity of data may allow for additional exploration of the outer limits for extraneous load as it applies to multimedia learning.

Additionally, Pastore suggests that a compression level of 25% would be allowable in a designed learning environment, using corporate implementation as a potential exemplar. While similar results on the effect of auditory compression on learning might be found within the corporate context, the lack of random sampling and limited target population nullifies this recommendation.

Quality of Writing/Organization (Rating = 4)

The quality of writing and organization of the paper was good, overall. Some elements were referenced inconsistently, however, resulting in potential confusion. For instance, the final questionnaire was referred to alternately as: "Likert-style question," "Likert-style questionnaire," and by the header "Cognitive load" (Pastore, 2009, p.494,494).

Contribution to Field (Rating = 4)

While the author cites this field of research as a gap in the current literature, the value of this research in the larger multimedia or computer-based instruction space is not well understood. While time-compression of audio resources may in fact increase the efficiency of instruction with little effect on learner achievement, this study represents only a small segment of the field of multimedia instruction.

Similarly, while the instructional materials and instruments have a research history in the field, the age of these materials may inhibit their application to contemporary multimedia learning research due to the staticity of the image approach. As referenced in the literature review, Pastore recognizes that proper application of multimedia to a learning experience can have the potential to increase the efficiency and quality of learning, but the same level of value has not been established for instructional compression on a purely linear time-based scale.

Overall Recommendation (Rating = Accept with Revisions)

This study provides valuable data that could be helpful in overcoming a gap in the current literature. While the study must include more guarded language in terms of audience generalization and application, I would accept the study for publication if the author would make appropriate revisions. In particular, the use of parallel language when addressing specific

instruments would aid in readability. Also, specific statements addressing the participant selection process, including the use of intact groups or convenience samples, would assist in clarifying the role of the target and accessible populations for this study. Finally, any conclusions or interpretations of the data must address only the target and accessible populations defined by the scope of the sample.

References

Fraenkel, J. R., & Wallen, N. E. (2009). *How to Design and Evaluate Research in Education* (7th ed.) [Amazon Kindle version]. New York: McGraw-Hill.

Merrill, M. D. (2008). What makes e3 (effective,efficient,engaging) instruction? [Manuscript] Retrieved from http://mdavidmerrill.com/Papers/Finding_e3_instruction_EdTech%5BFinal%5D.pdf

Pastore, R. S. (2010). The effects of diagrams and time-compressed instruction on learning and learners' perceptions of cognitive load. *Educational Technology Research & Development, 58*, 485-505.

Trochim, W. M. K., & Donnelly, J. P. (2008). *The Research Methods Knowledge Base* (3rd ed.). Mason, Ohio: Atomic Dog.