The acquisition of procedural skills:

An analysis of the worked example effect using animated demonstrations

R. David Lewis and Ann E. Barron Ed.D.

Abstract

Many educators suggest active rather than passive instruction. However, Sweller and Cooper found that learners who passively studied worked examples were significantly more efficient than learners who actively solved problems (Sweller & Cooper, 1985; Cooper & Sweller, 1987). This study tests this “worked example effect” (and the “variability effect”) using animated demonstrations. It considers the performance of four treatment groups, learners that: (a) study animated demonstrations, (b) study demonstrations and practice similar procedures, (c) study demonstrations and practice different procedures, (d) only practice procedures. Performance time and accuracy are compared, one week after initial instruction. Results were consistent with the “worked example effect.” It also considers instructional condition, learning, and performance efficiency. In each case, significant differences were not found.

Proposal Summary

Objectives or purposes

Sweller and Cooper found that learners who studied worked examples during early schema acquisition significantly out-performed their peers, who learned the same procedures through active problem solving (Sweller & Cooper, 1985, Cooper & Sweller, 1987). This has since been described as the “worked example effect” (Sweller, 1993; Sweller, van Merriënboer, & Paas, 1998). Lewis (2005) claims that animated demonstrations act as animated worked-examples. The primary objective of this study is to test this hypothesis and to determine if animated demonstrations are subject to Sweller’s worked example effect.

In addition, Tuovinen and Sweller (1999) found that learners who studied worked examples performed significantly better than those learning through discovery-based practice. The current study replicates their instructional conditions, but contrasts the cognitive load and learner performance of those using animated demonstrations with those using discovery-based problem solving.

Finally, Paas and van Merriënboer (1994) describe another cognitive load learning effect, the variability effect. This effect suggests that the variability of worked examples encountered during instruction, is important. Specifically this effect suggests learners are better able to abstract a problem schema if they are provided with varied examples of that problem type.

Perspective or theoretical framework

Cognitive load theory explains why some methods of instruction are more difficult (Sweller, 1988; Sweller and Chandler, 1994). It suggests that since working memory is limited, learners may be bombarded by information during early schema acquisition, and, if not properly managed, this cognitive overload may impair learning and subsequently deteriorate performance (Sweller, 1988). Because working memory resources are limited, novices easily become distracted by irrelevant aspects of a problem, and often make errors during learning through problem solving (Sweller, 1998). So rather than trying to solve problems initially, Sweller and Cooper (1985) suggest learners should use that time more effectively to first learn the underlying problem schema, perhaps by studying worked examples, and then later practice to automate these newly learned skills.

Sweller and his colleagues define worked examples this way: “A worked example is a step-by-step demonstration of how to perform a task or how to solve a problem” (Clark, Nguyen, & Sweller, 2006, p. 190). To date, little to no cognitive load research has been conducted using animation, other than the work of Mayer and his colleagues (Mayer, 2001). However, Mayer’s instructional materials are more aptly described as “animated explanation,” rather than animated demonstration.

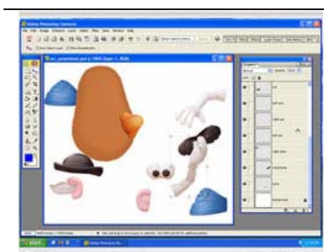
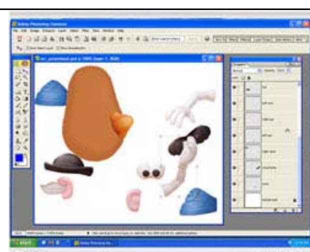
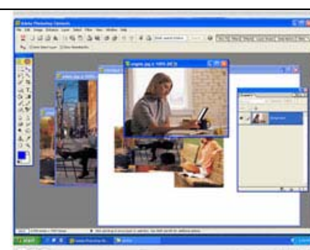
Mayer describes his work as asking questions about scientific explanations: “By ‘explanation’ we mean a description of a causal system containing parts that interact in a coherent way, such as a description of how a pump works or how the human respiratory system works.” (Mayer & Sims, 1994, p.389). Clark and Mayer (2003) even describe these as “two different e-learning goals” that teach learners to “inform and perform” (p.17).

Methods

This study contrasts learner performance given four instructional treatments (see Figure 1). All conditions watch a brief overview giving learners some context. This overview is a short web-based presentation (~ 2 minutes) that provides learners with a brief introduction to graphic design, but is not animated. It only presents learners with graphics depicting graphic designers and briefly describes Adobe Photoshop Elements 2.0 (with screen shots). Once this overview concludes a JavaScript randomly assigns learners into four separate instructional conditions.

Figure 1. Instructional materials include an overview and the four conditions.

Overview

Condition 1 Demo	Condition 2 Demo + Practice	Condition 3 Demo 2+ Practice	Condition 4 practice
			

The four instructional conditions (Figure 1) are: (1) an animated demonstration (demo); (2) an animated demonstration with the Mr. Potatohead problem (see Figure 2) (demo+ practice); (3) a different collage-based animated demonstration, with the Mr. Potatohead problem as practice (demo2+practice); and (4) a discovery based practice condition with the Mr. Potatohead documents serving as practice (practice only). All instructional materials were developed with Techsmith Camtasia Studio® 4.0 (Techsmith, 2006) and are designed to teach a novice how to use Adobe Photoshop layers. Specifically, learners learn how to select, move, rotate, and hide layers within an Adobe Photoshop Elements document (Adobe Systems, 2002). Both animated demonstrations are about 10 minutes long, and demonstrate the same procedures within different contexts.

Figure 2. Week one: learners reassemble the Mr. Potatohead problem.

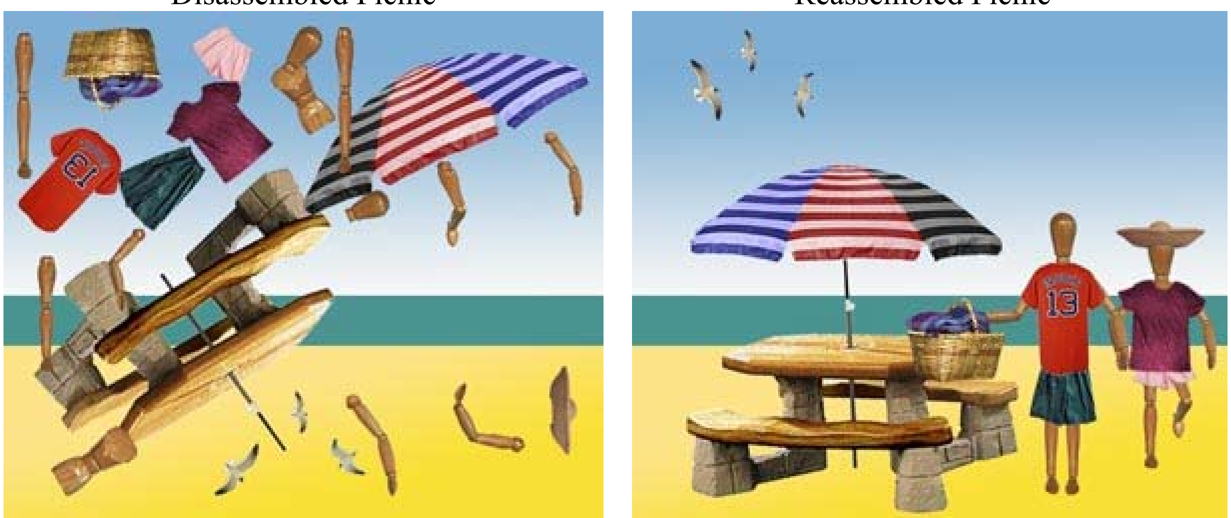
Disassembled Mr. Potatohead | Reassembled Mr. Potatohead

After viewing the animated demonstration, learners in condition 1 (the demonstration only group) were asked to fill out the "week one mental effort survey" and then to return the following week. Learners in conditions 2, 3, and 4 were allowed to practice their skills and asked to reconstruct the Mr. Potatohead document (Figure 2). Finally learners in group 4 (the practice only group) received no additional instruction following the overview, but were asked to assemble the Mr. Potatohead document.

A delayed test was conducted one week after initial instruction. It was during this delayed test that all learners were asked to reassemble a different scene – the picnic problem (see Figure 3).

Figure 3. Week two: learners reassemble the picnic problem.

Disassembled Picnic | Reassembled Picnic



Performance data were only collected during week two. TechSmith Morae® 1.01 served as the primary tool for data collection (Techsmith, 2004). This usability software recorded the learner's onscreen computer interaction and produced a coded movie file of each learner's performance. Analysis of these movie files produced the two main dependent variables (performance time & accuracy). Performance time was measured in seconds using Techsmith Morae. Accuracy was measured using a rubric, based upon the number of problem solving operators required to solve the problem, for a total of 48 possible points.

This study implemented the instructional condition efficiency procedure initially developed by Paas and van Merriënboer (1993), along with a more recent procedure that was developed by Paas, Tuovinen, Tabbers, and Van Gerven (2003). Mental effort ratings were gathered during week one (training) and week two following performance (see Tables 1 and 2).

Table 1.

Relative condition efficiency (adapted from Paas, 2007) uses mental effort ratings from the performance phase.

	Acquisition phase	Performance phase
Performance score		X
Mental effort rating		X

Table 2.

Learning efficiency (adapted from Paas, 2007) uses mental effort ratings from the acquisition phase.

	Acquisition phase	Performance phase
Performance score		X
Mental effort rating	X	

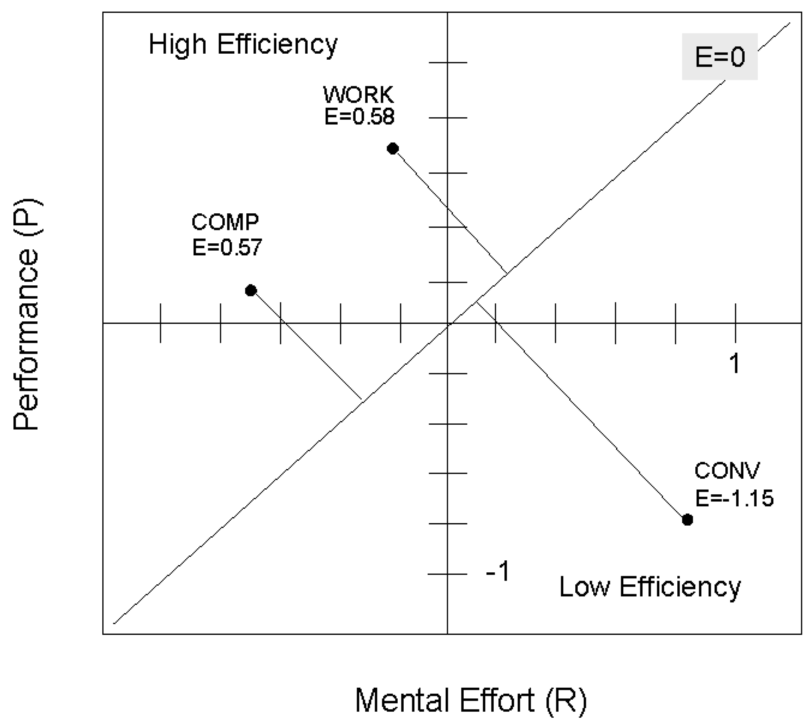
During week one, both a pre-treatment and post-treatment survey were used to gather basic demographic data and mental effort ratings during training. In addition, post-performance surveys were conducted to allow learners to rate their mental effort during performance. These ratings were combined with performance scores, given the formula in Figure 4, and plotted on a graph as in Figure 5, according to the procedure outlined by Paas & van Merriënboer (1993).

Figure 4. Instructional condition efficiency includes both performance and mental effort.

$$E = \frac{Z_{performance} - Z_{mentaleffort}}{\sqrt{2}}$$

When considering these graphs, one should note that the instructional conditions located above the diagonal line have greater relative efficiency scores because they have a higher group performance with lower invested mental effort (Paas & van Merriënboer, 1993). Conversely low instructional efficiency (below the line) is the result of low task performance and high effort (Tuovinen & Paas, 2004).

Figure 5. This is an example instructional condition efficiency graph (adapted from Paas & van Merriënboer, 1993).



Finally, this study included a metric that is based completely on objective measures performance time (PT) and performance (P) (see Figure 6). This construct does not include the more subjective, mental effort ratings. This metric, should be described as performance efficiency and treated in a manner similar to instructional condition efficiency, described in Paas & van Merriënboer (1993).

Figure 6. This formula derives a new metric called performance efficiency.

$$\text{Performance Efficiency} = \frac{PT - P}{\sqrt{2}}$$

Data source

Fifty six pre-service teachers viewed the instructional materials as a part of an introductory instructional technology course at a large southeastern university. Seven students were lost due to attrition or technical issues. Three participants answered yes to a post performance survey question that asked if they had used Photoshop during the past week; therefore, these performances were removed from the dataset, leaving a total of n=46 participants.

Results and Discussion

An ANOVA was conducted to compare the mean performance time and accuracy scores for each group of learners. The results for performance time were significant $F(3, 45) = 3.04$ $p<0.05$ MSE= 25.72. A post hoc analysis, a Tukey test (α=0.05) revealed significant differences between groups 1 and 2 (the demo only and demo +practice group) showing that the demonstration only group took significantly longer than the demo+practice group (mean performance times were 1202 and 864 seconds, respectively). This is consistent with early animated demonstrations studies that describe an “animation deficit” (Palmiter, 1991; Palmiter & Elkerton, 1993; Lipps, Trafton & Gray, 1998).

An ANOVA was also used to compare accuracy group mean scores. These results were somewhat surprising, as there were no significant differences between the group means; $F(3, 45) = 0.03$ $p=0.99$ MSE=7.37. Thus even though it took the demonstration only learners longer to complete the problem, they were able to compete with their peers and actually performed better than any other group (demo $\bar{x} = 42.67$, demo+practice $\bar{x}=42.06$, demo2+practice $\bar{x}=42.3$, practice $\bar{x}=41.73$). This was the case, even though they had not practiced the procedures one week earlier. These results are consistent with the worked example effect, but the differences were not significant. It should be stated that most cognitive load researchers recommend practice after viewing a worked example (an animated demonstration in this case) for this is necessary to automate skills and thus improve performance times.

Finally the instructional condition, learning, and performance efficiency metrics were implemented. The procedures described by Paas and van Merriënboer (1993) and Paas et al, (2003) called for an ANOVA to be run, to test for significant differences between groups.

In each of the efficiency measures, the group means cluster near the point of origin (see Figures 7, 8, & 9). A series of ANOVAs were run as a part this procedure but no significant differences were found. The instructional condition efficiency ANOVA revealed an $F(3, 45) = 1.00$ $p=0.39$; Learning efficiency, $F(3, 45) = 0.55$ $p=0.64$; and Performance efficiency, $F(3, 45) = 1.50$ $p=0.22$.

Educational or scientific importance of the study

Perhaps the most important reason to study animated demonstration is that this form of instruction is generalizable to all computer based applications. Certainly instructional designers should develop these presentations to make efficient use of both visual and verbal modalities, involving dual-coding theory (Paivio, 1978) and allowing for multimedia learning (Mayer, 2001). Computer-based procedures have become very important in our knowledge worker society, thus this efficient form of instruction is very valuable and should be studied in great detail.

Figure 7. Instructional condition efficiency results given the instructional conditions.

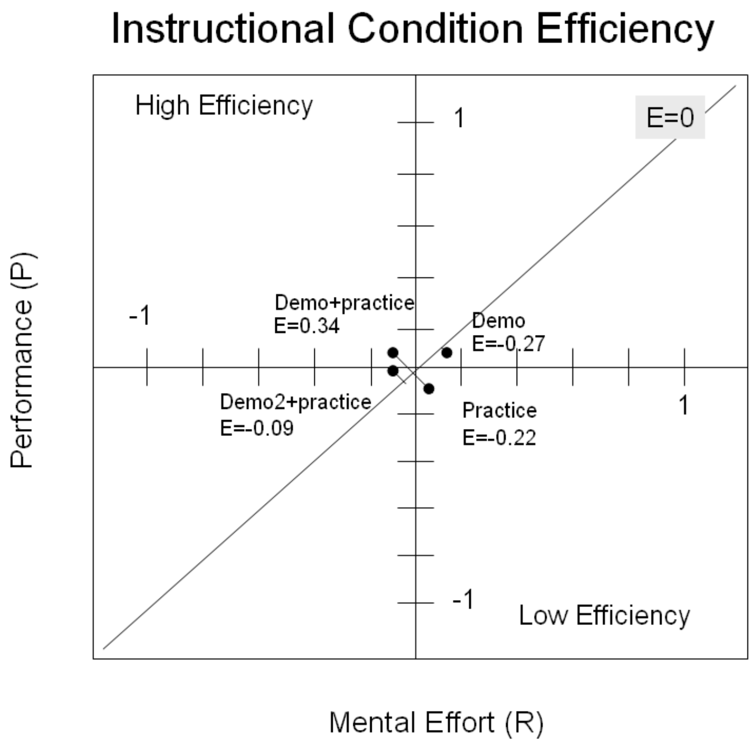


Figure 8. Learning efficiency results given the instructional conditions.

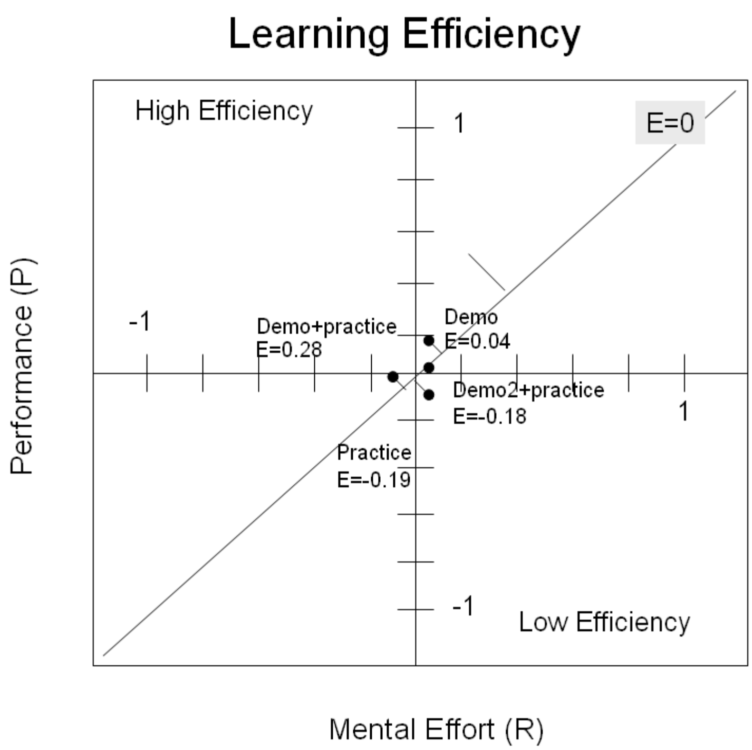
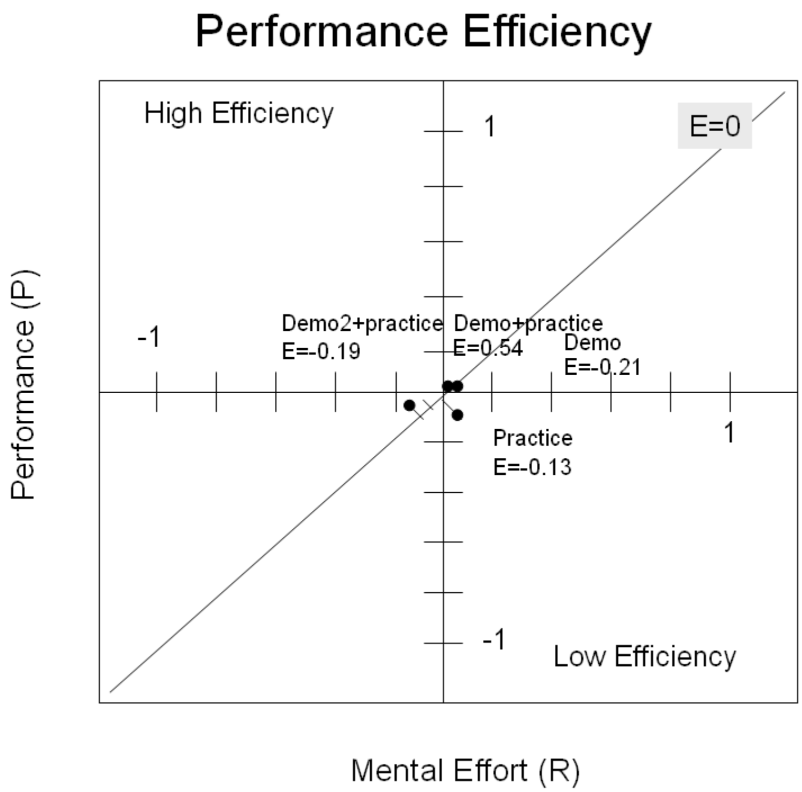


Figure 9. Performance efficiency results given the instructional conditions.



References

Adobe Systems. (2002). *Adobe Photoshop Elements 2.0* [Computer program] Mountain View, CA

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Anderson, J. R., Albert, M. V., Fincham, J.M. (2005). Tracing problem solving in real time: HMRI analysis of the subject-paced tower of Hanoi. *Journal of Cognitive Neuroscience, 17* 1261-1274.

Clark R.C. & Mayer R.E. (2003). e-learning and the science of instruction. San Francisco, CA: Pfeiffer

Clark, R.C., Nguyen, F., and Sweller, J. (2006). *Efficiency in learning: evidence-based guidelines to manage cognitive load.* San Francisco, CA: Pfeiffer.

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*. 79(4), 347-362.

Lipps, A. Trafton, J. and Gray, W. (1998). Animation as documentation: A Replication with Reinterpretation. In the *Proceedings for the 45th annual conference of the Society for Technical Communication Arlington*, VA. retrieved June 12, 2003 from http://www.stc.org/proceedings/ConfProceed/1998/PDFs/00006.PDF

Lewis, D. (2005). Demobank: a method of presenting just-in-time online learning. *Paper presented at the Annual meeting for the Association for Educational Communications and Technology*, Orlando, FL

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 84, 389–401.

Mayer, R. (2001). *Multimedia learning*. Cambridge: Cambridge University Press

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429–434.

Paas, F. (2007). *A multidimensional approach to the mental efficiency of instructional conditions*. Retrieved June 5, 2007 from http://www.ou.nl/Docs/Expertise/OTEC/Projecten/onderzoeksvoorstellen%20PDF/Paasproject34%5B1%5D.pdf

Paas, F., Tuovinen, J. E., Tabbers, H. K., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38* (1), 63–71.

Paas, F. & van Merriënboer, J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, 35(4) 737-743

Paas, F. & van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86, 122–133.

Paas, F., Tuovinen, J. E., Tabbers, H. K., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38* (1), 63–71.

Paivio, A. (1978). The relationship between verbal and perceptual codes. In E. C. Carterette and M. P. Friedman (Eds.) *Handbook of perception: Vol. IX Perceptual processing.* New York: Academic press.

Palmiter, S. (1991). *Animated demonstrations for learning procedural, computer-based tasks*. Doctoral Dissertation. University of Michigan, Ann Arbor, (Proquest No. 9124076)

Palmiter, S. & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction, 8*(3), 193-216.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(1), 257-285.

Sweller, J. (1993). Some cognitive processes and their consequences for the organisation and presentation of information. *Australian Journal of Psychology*, *45*(1) 1-8.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12 (3) 185-233.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2(1), 59-89.

Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.

Techsmith (2006). *TechSmith Camtasia Studio 4.0* [Computer program] Okemos, MI

Techsmith (2004). *TechSmith Morae 1.0.1* [Computer program] Okemos, MI

Tuovinen, J.E. & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*. 91(2) 334-341