Investigating the Influence of Transitory Information and Motivation during Instructional Animations

Paul Ayres and Amina Youssef, University of New South Wales, Sydney, Australia
Email: p.ayres@unsw.edu.au, a.youssef@unsw.edu.au

Abstract: This study investigated the cognitive load theory prediction that the inconsistent findings concerning the effectiveness of instructional animations are exacerbated by their transitory nature. Three groups were compared who received different but equivalent forms of instruction in learning a topic in economics. One group received an animation presentation with integrated text and diagrams, a second group received a static diagram presentation with integrated text and diagrams, and the third group received a static diagram presentation with non-integrated text and diagrams in a classical split-attention design. Results indicated that the animated design was superior to the static integrated design only on test questions that closely resembled the presented information. No other significant group differences were identified. Furthermore a battery of self-rating measures of cognitive load and motivational items indicated that test performance was predicted by motivation and concentration levels, rather than by mental effort or task difficulty.

This study had two main aims. The first aim was to advance the research on instructional animations, and secondly, to explore the use of multiple self-rating measures of cognitive load. The paper begins by briefly describing some of the main principles of cognitive load theory as it used as the theoretical framework for the study.

In the last decade cognitive load theory has developed into a well-known theory of learning and instruction. From its early development in the 1980s, where it focused on explaining why problem solving was an ineffective method of learning, it has more recently become a theory, which considers cognition as a natural information processing system similar to evolution by natural selection (Sweller, 2003, 2004). At the centre of CLT is the interaction between long-term memory (LTM) and working memory (WM). Whereas WM is very limited in both capacity and duration (see Miller, 1956; Peterson & Peterson, 1959), the capacity of LTM is extremely large, and stored information can last a lifetime. Furthermore it is argued by CLT theorists that information is stored in LTM in the form of schemas (Chi, Glaser & Rees, 1982). By recalling schematic knowledge from LTM into WM, multiple bits of connected information can be treated as single elements, thus relatively increasing the capacity of WM. Hence humans are able to learn complex material and solve difficult problems in spite of the limitations of WM. However, when learning novel information, where schemas are not available, the limited capacity of WM can be a severe impediment to learning. Much of the research into CLT has investigated the various conditions under which WM resources become overstretched and how these conditions can be alleviated by instructional designs that take into account the human cognitive architecture (van Merriënboer & Sweller, 2005).

CLT identifies three categories of cognitive load: intrinsic, extraneous and germane (see Sweller, van Merriënboer & Paas, 1998). *Intrinsic* load is the load caused by the complexity (element interactivity, see Sweller & Chandler, 1991, 1994) of the materials to be learnt, *extraneous* load is the load caused by the instructional procedures, and *germane* load is the load directly invested in schema acquisition. Whereas intrinsic load is considered fixed, dependent upon only the prior knowledge of the learner (se Kalyuga, Ayres, Chandler & Sweller, 2003), extraneous load is under the control of the instructional designer. An effective design lowers extraneous load and induces germane load. Over two decades CLT research has identified a number of strategies, such as worked examples and the modality effect, to optimize learning environments (see Sweller, 1999; van Merriënboer & Ayres, 2005). In the last few years, researchers have extended the research on CLT to outline a theory to explain the conditions under which instructional animations may be designed more effectively.

The results of studies that have used dynamic representations (animations) as a learning tool are somewhat inconclusive. A review by Tversky, Morrison and Betrancourt (2002) found that animation was no more effective than static representations, although it was postulated that animation might be best used in environments that "convey real time changes and reorientations in time and space". (p. 257). Mayer, Hegarty, Mayer and Campbell (2005) found static diagrams to be more effective than animations when learning about mechanical systems. In contrast, a meta-analysis by Höffler and Leutner (2007) found evidence that a number of studies showed animation to be more effective than static pictures, particularly when the animation was highly realistic and/or procedural motor knowledge was involved. As well as research that has compared static diagrams with animations, a number of studies have looked to improve the animated designs themselves. For example, researchers have shown that animations can be improved by more directed extraction of information (Lowe, 1999, 2003, 2004) user interactivity (Hasler, Kersten & Sweller, 2007), segmenting the presentation into smaller chunks (Mayer & Chandler, 2001; Moreno, 2007), and by signaling information (De Koning, Tabbers, Rikers & Paas, 2007; Lusk & Atkinson, 2007; Moreno, 2007). Whereas a number of these compensatory tactics have been successfully employed and a number of explanations tendered, the research field has lacked a comprehensive theory to explain why animations are sometime effective and other times not. However, CLT has provided some further insights. CLT theorists argue that animations can be highly transitory in nature (see Ayres, Kalyuga, Marcus & Sweller, 2005; Ayres & Paas, 2007a, 2007b). Animations that involve information disappearing from the screen as the animation progresses, forces learners to process current information while trying to remember previous information, leading to heavy demands on working memory load. In such situations, animation creates an extraneous cognitive load. This argument gives a plausible explanation as to why stopping an animation (user-interactivity) or segmenting it into smaller parts is beneficial- both strategies lower working memory load by negating the effects of transitory information. The study reported here continues the research into the transitory effect. We argue that another way to overcome transitory information it to ensure that no information disappears from the screen. To test this prediction the following study was completed.

Main hypotheses of the study

The study consisted of three treatment groups and participants were required to learn some economics content, a domain rarely used in CLT research. One group (*Animated*) was required to learn from an animated presentation. To overcome the transitory effect the computer-based presentation was designed so that no information disappeared from the screen. To eliminate other potential sources of extraneous cognitive load, namely the split-attention effect (see Ayes & Sweller, 2005; Chandler & Sweller, 1992), all text and diagrams were integrated together. A second group (*Static integrated*) received an equivalent presentation to the Animated group except that a set of static diagrams replaced the animation. Again all text and diagrams were integrated to avoid a split-attention scenario. The third group (Static non-integrated) was identical to the Static integrated group, except that the text and diagrams were deliberately kept separate in order to induce a split-attention effect. By inducing such an effect it was expected that the least learning would occur and therefore this group served as a control condition. The following two hypotheses were tested:

- *Hypothesis 1:* Learners will benefit more from an animated instructional format than a static format
- *Hypothesis 2:* Learners will benefit more from an integrated format than a non-integrated format.

Cognitive load measures

In much of cognitive load theory research a global self-rating measure of cognitive load has been used based on the original Paas (1992) 9-point scale. This scale was worded as follows for the lowest mental effort rating (1): "In solving or studying the preceding problem I invested very, very low mental effort", and was administered after students had taken the test questions. However, Van Gog and Paas (2008) have pointed out that many researchers have used variations of the original scale and changed the wording of the scale considerably. Often using *difficulty* instead of mental effort. Further, most researchers administer the scale following instruction rather than the test phase. Accordingly, Van Gog and Paas have argued that these variations may pose a threat to the validity of the instrument. In more recent times researchers have become interested in measuring intrinsic, extraneous and germane load individually. For example, Ayres (2006) has measured changes in intrinsic load, while others have investigated methods to measure all three loads (see Cierniak, Scheiter, & Gerjets, 2007; Scheiter, Gerjets, & Opfermann, 2007). An interesting finding in the Cierniak et al. study was that a *concentration* measure was found to mediate the split attention effect. Furthermore, it has been argued by Moreno (2007) that motivational factors mediate learning and influence cognitive engagement. Whereas research into motivation is a huge field in its own right, it is rarely used in CLT. As a result of these new directions and findings, a number of subjective cognitive load and motivational measures were introduced into this study. The broad aim of the subjective measures, which was *exploratory* in nature, was to collect information on:

- Potential measures of intrinsic, extraneous and germane load
- Concentration and motivation measures
- Predictors of performance

Method

Participants

The participants used in the study were 45 (18 male & 25 female, mean age of 20.4 years) undergraduate students from an Australian university enrolled in first year Education subjects. No student had formally studied an economics course at high school or university. Each student was randomly assigned to one of the three groups (each with N=15).

Materials

Background knowledge materials

The economics topic to be learnt in this study was *Demand*. In order for students to gain some basic knowledge of the topic before embarking on the main learning phase some background knowledge materials were developed. A two-page (A4) handout provided basic information in the form of definitions and examples on the Demand schedule, the Demand curve, and shifts of the Demand curve. It also explained how as the price of a *Good* increases the quantity demanded of that Good also decreases and vice versa. A graph illustrated the relationship between price and demand, and showed that a typical Demand curve is sloping downwards from left to right. A paper-based knowledge test (pre-test) was constructed, comprising of multiple choice and short-answer questions, to determine the students' understanding of the economics concepts taught in the background knowledge.

Learning materials

A different presentation was constructed for each group, but all depicting the same concepts, diagrams and text. Each computer-based presentation consisted of two parts. The first part explained how the demand curve moves given an increase in demand, while the second part demonstrated a decrease in demand. The two static group presentations were identical apart from the positioning of the text on the slide. For an increase in demand two power point slides were presented consecutively. The first slide (static image) showed the position of the original demand curve (line), and the second slide showed the final position of the curve following an increase in demand. Arrows were used to indicate the motion of the curve on the second slide (see Figure 1). On each slide, text was used to explain why the curve was moving in such a fashion. However for the Static integrated group, the text was integrated into the diagram, whereas for the Static non-integrated group, all text was positioned at the bottom of the slide. The first slide was visible for 95 seconds to be replaced by the second slide also visible for 95 seconds. For the second part of the presentation the format of the materials was identical, the only difference being that the two slides showed a decrease in demand. Altogether the whole presentation, including some further qualifying instructions took 420 seconds. For the Animated group, all the material was identical to the Static integrated group, except in place of the two slides showing the curve before and after a change in demand, one single animated display was presented, showing the demand curve move according to the change in conditions. It was predicted that this dynamic representation of the curve moving in real time would reinforce the underlying principles of the economic concept of demand. Furthermore the text was synchronized to appear at the right time. For each part (increase and decrease of demand) the display lasted 190 seconds. In this display no information was lost (disappeared from the screen) at anytime, which prevented any transitory effects. The whole display lasted 420 seconds, and each component lasted the exact time as the static versions.

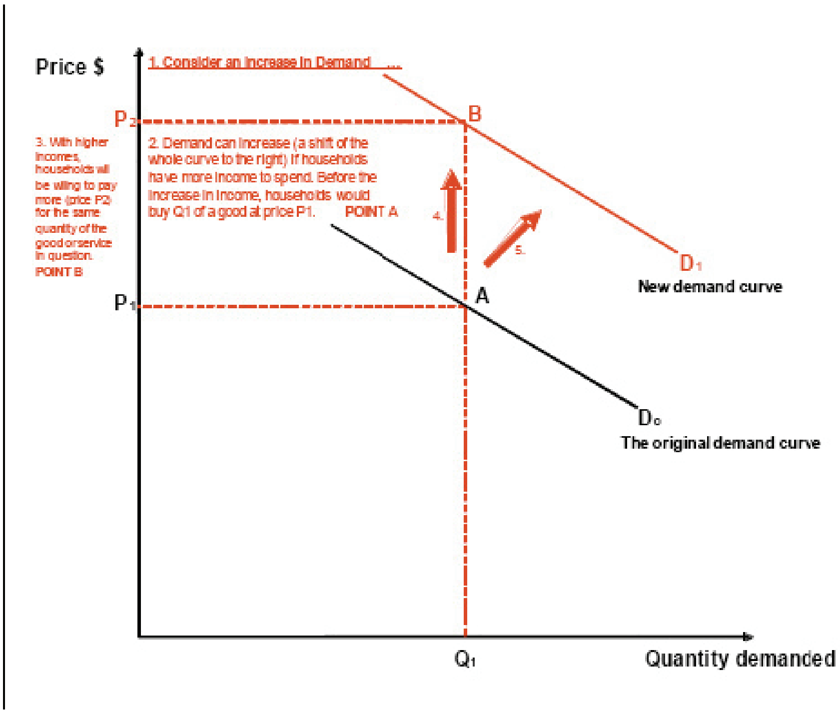

Figure 1. Second slide presented for the Static integrated group.

Testing materials. Two assess the effectiveness of the different presentations two paper-based tests were constructed. One (post-test) tested basic understanding of the material and consisted of three types of different knowledge: background knowledge (4 questions), economic theory (10 questions), and Economic diagrams (2 questions). The second (transfer) tested whether the knowledge learnt could be transferred (9 questions).

Self-rating measures. To measure cognitive load and other motivational-type factors fourteen separate subjective items were constructed (see Table 1). Eight items were grouped and presented before the test questions (Pre) and six after the test (Post). Each group was presented on a single sheet of A4 paper and students were asked to rate each item according to a 9-point Likert scale. In order to potentially measure the individual cognitive loads, items (Pre-1, Pre-2, Pre-3) were constructed that focused on aspects of the presentation and therefore might be considered extraneous, the content (Intrinsic: Pre-5, Pre-6, Post-3) and the effort made in learning, including motivation and concentration, (Germane: Pre-4, Pre-7, Pre-8, Post-1, Post-2, Post-6). In addition, two items were added that evaluated the usefulness of the presentations (Post-4 and Post-5). Measures of both understanding and learning were used (Pre-2 and Pre-3). A variation of Paas' original measure (Paas, 1992), which grouped both 'solving and studying' together, was separated into two measures: one focusing on mental effort during the presentation (Post-1) and the other on solving the test problem (Post-2). These were presented after the test even though they referred to the instructional phase consistent with the original instrument (see Van Gog & Paas, 2008). A concentration item was also included and repeated in the pre and post-testing phases (Pre-8 and Post-6).

Table 1: Cognitive load and motivational measures.

Item No.	Item Description	Mean (SD)
Pre-1	How difficult was it for you to follow the demonstration? (Pre-test)	3.4 (1.79)
Pre-2	How difficult was it for you to learn from the demonstration? (Pre-test)	3.9 (1.92)
Pre-3	How difficult was it for you to concentrate during the demonstration? (Pre-test)	5.5 (1.63)
Pre-4	How much were you motivated by the demonstration? (Pre-test)	4.0 (1.44)
Pre-5	How difficult was it for you to understand the economics concepts? (Pre-test)	3.9 (1.54)
Pre-6	How difficult was it for you to learn the economics concepts? (Pre-test)	4.2 (1.55)
Pre-7	How much effort did you put into learning the economics material? (Pre-test)	5.5 (1.36)
Pre-8	How high was your concentration level when trying to learn the economics material? (Pre-test)	5.1 (1.27)
Post-1	In studying the demonstration how much mental effort did you invest? (Post-test)	5.6 (1.35)
Post-2	In solving the test questions how much mental effort did you invest? (Post-test)	6.2 (1.19)
Post-3	How difficult was it for you to answer the test questions? (Post-test)	5.6 (1.39)
Post-4	How helpful was the demonstration in learning the economics material? (Post test)	5.9 (1.46)
Post-5	How helpful was the demonstration in understanding the economics material? (Post test)	6.0 (1.50)
Post-6	How much did you concentrate when trying to learn the economics material? (Post-test)	5.6 (1.11)

Procedure

Participants were tested in groups of two, each with a separate computer, and were required to complete seven phases in the study. In the first phase participants were given five minutes to read and understand the background information handout. They were given ten minutes to complete the Pre-test. In the third phase, participants viewed one of the three presentations for seven minutes in total. In the fourth phase, participants were then required to complete the first batch of 8 cognitive load measures (Pre), which took approximately three minutes. In the fifth phase, participants were complete the post-test in 12 minutes. Following on from the Post-test the Transfer test was administered (Phase six) and five minutes was allocated to this task. The final component required participants to complete the remaining set of 6 cognitive load measures, which took approximately three minutes. The total study time was approximately forty-five to fifty minutes.

Results and discussion

Learning tests

For the pre-test, a 1-way ANOVA revealed no significant group difference ($F(2, 42)= 1.13$, $p > 0.05$). Because no group differences were found on the pre-test it can be assumed that prior knowledge differences did not influence the overall results of this study. For the post-test, group means and standard deviations were calculated for each of the three different types of questions (see Table 2). A MANOVA using the three types of questions as variables produced a significant result under the Roy's largest root test: $F(3, 41) = 3.56$, $p < 0.05$. The univariate tests were non-significant for background knowledge and economic theory question subgroups, but there was a significant group difference for the economic diagram questions ($F(2, 42) = 3.482$, MSe= 2.156, $p < 0.05$). A Tukey Post-Hoc test found that the animation group performed better that the static integrated group ($p < 0.05$), but all other comparisons were non-significant. Means and SDs for the transfer test questions were Animated (M =2.58, SD =1.69), Static integrated (2.54, 1.07), and Static non-integrated (2.62, 1.12). A 1-way ANOVA revealed no significant group difference on the transfer test questions ($F(2, 42) = 0.01$, $p > 0.05$).

The above results partially support the first hypothesis. Evidence emerged in the post-test that the economic diagrams questions were better answered by the animation group than the static integrated group. It is notable that the economic diagrams questions closely imitated the presentation materials, and therefore were perhaps an example of the *congruence principle* (see Tversky et al., 2002). Tversky et al. have argued that one condition in which animations may be effective is when they closely resemble the material to be learnt. But, overall the results in favour of the animated group are weak. No transfer effects were found, nor were there any differences on test items that examined background knowledge or economic theory. Surprisingly, the second hypothesis was not supported. The split-attention group did not perform any worse than the animated or static-integrated groups. This latter finding could imply that the study materials to be learnt could be low in element interactivity as the split-attention format was expected to have a negative result. Previous research has found that many of the cognitive load theory effects only occur when element interactivity is high (Sweller and Chandler, 1994).

Table 2: Mean group scores on the post-test

Group		Background knowledge	Economic theory	Economic diagrams	Combined scores
Animated	Mean	3.40	5.43	1.67	10.50
	SD	0.91	2.61	0.62	3.34
Static integrated	Mean	3.47	4.43	0.93	8.83
	SD	1.06	2.05	0.88	2.50
Static non-integrated	Mean	3.53	4.57	1.47	9.57
	SD	0.74	2.03	0.83	2.37

Cognitive load measures

Factors.

The overall means and standard deviations for the cognitive load measures are shown in Table 1. To investigate the potential that the fourteen items form specific constructs (factors) an exploratory factor analysis was conducted using the Principal Axis Factoring method with Varimax Rotation and Kaiser Normalisation on the whole sample. In order to test reliability, a Cronbach alpha coefficient was calculated for each identified factor (Eigenvalues > 1). As a result four factors were identified.

The first factor (Initial Eigenvalue = 4.5, variance explained = 31.8%, Cronbach alpha = 0.84) consisted of the five items Pre-6 (How difficult was it for you to learn the economics concepts?), Pre-5 (How difficult was it for you to understand the economics concepts?), Pre-1 (How difficult was it for you to follow the demonstration?), Pre-2 (How difficult was it for you to learn from the demonstration?) and Post-3 (How difficult was it for you to answer the test questions?). The grouping of these items are consistent in that they refer to difficulty in learning the content, following the presentation and answering the test questions. It has nominally been named *Difficulty in learning* and appears to be a construct that includes both intrinsic and extraneous cognitive load.

The second factor (Initial Eigenvalue = 3.3, variance explained = 23.7%, Cronbach alpha = 0.84) consisted of five items Post-6 (How much did you concentrate when trying to learn the economics material?), Pre-8 (How high was your concentration level when trying to learn the economics material?), Pre-7 (How much effort did you put into learning the economics material?), Post-2 (In solving the test questions how much mental effort did you invest?) and Post-1 (In studying the demonstration how much mental effort did you invest?). This collection of items is concerned with effort in learning and concentrating, and from this perspective may be considered a mental effort construct. It has been nominally named *Effort and concentration in learning*. It cannot be considered Germane load in the classical cognitive load theory sense because it is not known if this mental effort was invested totally in schema acquisition.

The third factor (Initial Eigenvalue = 1.6, variance explained = 11.1%, Cronbach alpha = 0.94) consisted of two items Post-4 (How helpful was the demonstration in learning the economics material?) and

Post-5 (How helpful was the demonstration in understanding the economics material?) These two items in this factor are highly correlated (0.89, $p < 0.01$) and are an appraisal of the presentations viewed, and was named *Demonstration helpfulness*. The final factor (Initial Eigenvalue = 1.2, variance explained = 8.2%, Cronbach alpha = 0.63) consisted of the two remaining items Pre-4 (How much were you motivated by the demonstration?) and Pre-3 (How difficult was it for you to concentrate during the demonstration?). This was the weakest factor identified and was tentatively labeled *Motivation*. However, of interest here is that the concentration item focusing on the presentation is linked to the motivation item, but not with the other concentration items that focus on learning the content (Pre-8, Post-6).

Predictors of performance

A total correlation matrix of all test scores and the cognitive load measures indicated that performance on the post-test questions was significantly correlated (0.43, $p < 0.01$) with only one item (Pre-4: How much were you motivated by the demonstration?) and hence was the only significant predictor using a linear regression analysis ($r^2 = 0.19$). In addition, it was also found that Pre-4 was significantly correlated with Pre-3 (-0.46, $p < 0.01$) and Pre-8 (0.29, $p = 0.05$), both of which are measures of concentration; and Post-5 (0.39, $p < 0.01$) and Post 6 (0.30, $p < 0.05$) both of which are 'usefulness' measures. Two of these items Pre-3 ($r^2 = 0.10$; How difficult was it for you to concentrate during the demonstration?) and Post-4 ($r^2 = 0.15$; How helpful was the demonstration in learning the economics material?) were significant predictors of the motivation item. Because the test was conducted after the presentation, items connected to the presentation are legitimate predictors of test performance. In summary it can be seen that motivation during the presentation predicted test performance, and in turn concentration and perceived helpfulness of the presentation predicted motivation, although they did not predict performance directly.

It was also notable that not one of the subjective measures correlated with the transfer scores, although Post-6 (How much did you concentrate when trying to learn the economics material?) just failed to reach significance ($r = -0.30$, $p = 0.06$). This is an interesting result, as it means that the less concentration led to an increase in test performance.

In this study there were no significance group differences on overall performance measures. However, on the subgroup of test items called economic diagrams a significant result was found. Consequently a correlation matrix was calculated for this test measure with the subjective cognitive load measures. Significant correlations were found with difficulty in learning from the presentation (Pre-2: $r = 0.41$, $p < 0.01$), difficulty in understanding the content (Pre-5: $r = 0.33$, $p < 0.05$), difficulty in learning the content (Pre-6: $r = 0.34$, $p < 0.05$), and concentration levels when trying to learn (Pre-8: $r = - 0.30$, $p < 0.05$; Post-6: $r = -0.35$, $p < 0.05$). The significant predictors of performance on these particular types of questions were Pre-5 ($r^2 = 0.12$; difficulty in understanding the content) and Pre-8 ($r^2 = 0.09$; concentrating when trying to learn). However, the particular direction of the correlations meant that those who found the demonstration more difficult and performed better in the test results, as did those who concentrated less. These results are somewhat counter-intuitive and may be attributed to overall motivation of the students in conjunction with the design of the study materials.

Conclusions

The two main aims of this study were to test the effectiveness of an animated design, free of transitory information, in learning some economics content, and to explore the use of multiple measures of cognitive load including motivational items. Regarding the first aim there was weak support that the animation presentation was effective. It was hypothesized, based on the argument of Ayres & Paas (2007a, 2007b) that by eliminating transitory information, extraneous load would be reduced, and consequently animations would become a more effective learning tool. The evidence does not confirm this hypothesis, nor does it refute it. The most notable results were found with test questions that most resembled the presentations, the animation group was more effective than a static integrated group, and secondly there was no disadvantage to a group that received a presentation in a split attention format. The latter result can be plausibly explained by element interactivity, as it argued by cognitive load theorists that most cognitive load effects are only found with high element interactivity materials. In contrast, the former result may be explained by the congruence principle due to the similarity of the test question subset and the presentation (Tversky et al., 2002).

For the second aim the use of multiple cognitive load measures produced a number of interesting findings. The finding that the motivation item (How much were you motivated by the demonstration) was the only predictor of the overall post-test results is significant as it suggests that motivation is a factor, which should be utilized more frequently in CLT research. A point argued by Moreno (2007) and acknowledged by Paas, Tuovinen, Van Gervan & Darabi (2005). The exploratory factor analysis indicated four factors, two of which each explained over 20% of the variance. The first factor, named *Difficulty in learning,* seemed to contain both intrinsic and extraneous cognitive load measures. If this is a real construct, then attempts to measure intrinsic and extraneous loads separately may be complicated by a closer relationship between them than most cognitive load theorists assume. The second factor, named *Effort and concentration in learning*, is notable because concentration levels are correlated with mental effort. Concentration items were also found to be predictors of the motivation item and test performance on the economic diagram questions. The significant role played by concentration is also consistent with the research by Cierniak et al. (2007) who found concentration to be a mediator of the split-attention effect, and again links to the importance of motivation. Finally the finding that two distinct factors emerged which correspond to content difficulty and mental effort in learning, suggest that they are different constructs, therefore supporting the concern of van Gog and Paas (2008) that researchers are using different measures in CLT research.

The study had some limitations. Firstly the materials may have been so low in element interactivity that the likelihood of testing the overall hypotheses was reduced. However, future research could incorporate more complex material. If an animated group can be found to be more effective than a static group on material less similar to the presentation then the transitory argument is strengthened. Also a direct comparison with a animated group where information disappeared from the computer screen could have tested the transient hypothesis more directly. The sample size was too small to provide great confidence in the factor analysis, which must be considered entirely exploratory. Again future research could include a much larger sample to investigate the identified factors further. The motivational scale used was somewhat unsophisticated, and future experiments could include more traditional measures of motivation found in the motivational literature. Nevertheless the study did show that an animated design is effective in an economics domain when it depicts graphical information. The study also suggests that motivational issues are highly important when designing instructional materials. Finally, the results have also identified some potential new directions for CLT researchers in measuring cognitive load. That is, a battery of items may provide some robust factors that may enable more sensitive measures of cognitive load to be identified.

References

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, *10,* 1-12.

Ayres, P., Kalyuga, S., Marcus, N., and Sweller, J. (September, 2005). *The conditions under which instructional animation may be effective.* Paper presented at an International Workshop and Mini-conference, Open University of the Netherlands: Heerlen, The Netherlands.

Ayres, P. & Paas, F. (2007a). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology, 21*, 695-700.

Ayres, P. & Paas, F. (2007b). Can the Cognitive Load Approach Make Instructional Animations More Effective? *Applied Cognitive Psychology, 21*, 811-820.

Ayres, P., & Sweller, P. (2005). The Split-Attention Principle in Multimedia Learning. In R.E. Mayer (ED.), *The Cambridge Handbook of Multimedia Learning* (pp. 135-146). Cambridge: Cambridge University Press.

Cierniak, G., Scheiter, K., & Gerjets, P. (2007). Subjective and objective load measures: Subjective load ratings vs. Dual-task methodology. Paper presented at the Cognitive Load Theory Conference, UNSW, Sydney, 24-26 March.

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.

Chi, M., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7-75) Hillsdale, NJ: Erlbaum.

De Koning, B. B., Tabbers, H.K., Rikers, R.M.J.P., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology, 21,*7731-746

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21*, 713-729.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. Learning and Instruction, doi: 10.1016/j.learninstruc.2007.09.013

Lowe, R. K. (1999). Extracting infromation from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225-244.

Lowe, R. K. (2003). Animation and Learning: selective processing of information in dynamic graphics. *Learning and Instruction, 13,* 157-176.

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14,* 257-274.

Lusk, M. M., & Atkinson, R.K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology, 21,*7747-764.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93,* 390-397.

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When Static Media Promote Active Learning: Annotated Illustrations Versus Narrated Animations in Multimedia Instruction. *Journal of Experimental Psychology: Applied 11,* 256-265.

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review, 63,* 81-97.

Moreno, R. (2007). Optimizing learning from animations by minimizing cognitive load: Cognitive and affective consequences of signaling and segmentation methods. *Applied Cognitive Psychology, 21,*765-781.

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., Tuovinen, J., Van Merriënboer, J. J. G., & Darabi, A. (2005). A motivational perspective on the relation between mental effort and performance: optimizing learner involvement in instruction. *Educational Technology Research and Development, 53*(3), 25-34.

Peterson, L., & Peterson, M. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193-198.

Scheiter, K., Gerjets, P., & Opfermann, M. (2007). Online measurement of different types of cognitive load with an adjusted version of NASA-TLX. Paper presented at the Cognitive Load Theory Conference, UNSW, Sydney, 24-26 March.

Sweller, J. (1999). *Instructional Design.* Camberwell, Victoria, ACER Press.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction, 4,* 295-312.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, *Vol. 43* (pp.215-266). San Diego: Academic Press.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science, 32*, 9-31.

Sweller, J., & Chandler, P. (1994). Why Some Material Is Difficult to Learn. *Cognition & Instruction, 12,* 185.

Sweller, J., & Chandler, P. (1991). Evidence for Cognitive Load Theory. *Cognition & Instruction, 8,* 351.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review, 10,* 251-296.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57,* 247-262.

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, *43*, 16-26.

Van Merriënboer, J. J. G., & Ayres, P. (2005). Research on cognitive load theory and its design implications for E-learning. *Educational Technology, Research and Development, 53,* 5-13.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review, 17,* 147-177.

Acknowledgements

The research was funded by a grant from the UNSW Faculty of Arts and Social Sciences awarded to the first author.