Contemporary Cognitive Load Theory Research: The Good, the Bad and the Ugly

Paul A. Kirschner[1]*, Paul Ayres[2], Paul Chandler[3]

[1] Centre for Learning Sciences and Technologies, Open University of the Netherlands

[2] School of Education, University of New South Wales, Australia

[3] Faculty of Education, Wollongong University, Australia

* Correspondence can be sent to Paul A. Kirschner, Centre for Learning Sciences and Technologies (CELSTEC), Open University of the Netherlands, Valkenburgerweg 177, 6419AT Heerlen, The Netherlands. E-mail: paul.kirschner@ou.nl

Abstract

This paper reviews the 16 contributions of the special issue entitled *Current Research in Cognitive Load Theory*. Each paper is briefly summarized and some critical comments made. The overall collection is then discussed in terms of the positive contributions they make to the field of learning and instruction, and cognitive load theory in particular (the *good*), as well as problematical issues such as unresolved explanations and conflicting results (the *bad*) and the special case of measuring cognitive load (the *ugly*).

Contemporary Cognitive Load Theory Research: The Good, the Bad and the Ugly

Since its development in the 1980's, cognitive load theory (CLT) has identified many effects such as worked examples, temporal and spatial split-attention, redundancy, modality and expertise reversal (for summaries see Sweller, 1999; Van Merriënboer & Ayres, 2005; Van Merriënboer & Sweller, 2005) that influence instructional design. The papers in this special issue show the broadening of the research base and indicate the variety of research groups studying CLT. This article discusses 16 studies which are based on the 3rd Internal Conference on Cognitive Load Theory in Heerlen, the Netherlands. They represent a cross-section of contemporary research into CLT. We have utilized the well-known colloquialism the *good, the bad and the ugly*, in the title, as it provides a useful framework to sums up our observations on the papers. There are very positive aspects to the studies particularly the focus on well conceptualized random controlled studies which measure learning outcomes. It is an unfortunate reality that the field of cognition and learning continues to lack from evidenced based theory driven research and the current studies provides an ongoing valuable contribution to the field. There also some issues that we consider more problematic and require attention to ensure the future quality of cognitive load research. We have grouped the papers together into three sections, which we have called: a) learning in complex environments, b) learner control and choice; and c) animated and multimedia instruction. We start our discussion by reviewing each individual paper, and conclude by identifying the different issues that could be considered the good, the bad and the ugly of the current papers

Individual Paper Summaries and Discussions

Learning in Complex Environments

As society and work environments become more interconnected and complex, it is increasingly relevant that cognition and learning research is carried out in environments that ‘mirror’ the complexity of the real world. One of the major strengths of cognitive load research over the years is that it has increasingly focused on complex learning environments and real life school and industry settings (Chandler & Sweller, 1991, 1992, 1996; Ward & Sweller, 1990). The first set of contributions to this special issue all deal in one way or another with learning and instruction in complex environments.

The contribution by Schwonke, Renkl, Salden, and Aleven (this issue) investigated the effects of different ratios of worked solution steps (high assistance) and to-be-solved problem steps (low assistance) on cognitive skill acquisition in geometry using a cognitive tutor. Based on the *Assistance Dilemma* (see Koedinger & Aleven, 2007), they assumed an inverted U-shaped relation between the number of worked-steps and learning outcomes with neither the highest proportion of guidance (i.e., many worked steps with little problem solving practice) nor the lowest proportion would be optimal for learning. The authors assume in their research that they can directly assess extraneous load by asking learners how difficult it was to study or solve the problems. We will return to this in the discussion further on in this article. In agreement with P. A. Kirschner, Sweller and Clark (2006), the authors found that problem solving induced more extraneous cognitive load than example based learning–irrespective of the ratio of worked steps and to-be-solved steps. Together with a pattern of substantial negative correlations between extraneous load and learning outcomes, this corroborated the notion that the worked-example effect is based upon reduction of extraneous load.

Kalyuga and Hanham (this issue) researched whether learning flexible problem-solving skills could be enhanced by explicitly instructing learners in generalized forms of schematic knowledge structures applicable to a greater variety of problems. They reasoned that such structures direct learner attention toward essential characteristics of novel problems and be associated with top-down types of transfer (Bassok & Holyoak, 1993). To this end they propose a general schematic framework for describing different man-made products in technical domains based upon the assumption that technical objects can be characterized by their functions / purpose, the processes utilized in their operation, and their internal structure; the *functions-processes-structures (FPS) framework*. Kalyuga and Hanham compared direct instruction in the FPS schematic framework with a conventional format without explicit schema-based instruction. The schema-based instruction used two alternative procedures, a gradual hierarchical multilevel introduction of the FPS schema and a single-level introduction. The differences between the three formats were not significant on a retention test but were significant for transfer, where the hierarchically structured condition performed marginally better (i.e., not significant) than both other groups. They speculate that the results indicated that there was a “possible effect of enhancing learner abilities in handling relatively new (transfer) problems” when explicit instruction via the FPS schema was used.

Berthold, Röder, Knörzer, Kessler, and Renkl (this issue) studied whether and to what extent *explanation prompts* as an instructional support feature induce active processing while at the same time looking at whether, as Sweller (2006) states, they take cognitive load to the upper limit of working memory capacities when learning with complex learning materials. Their contribution to this issue continues Berthold, Eysink, and Renkl’s (2009) work on using explanation prompts to focus the processing of concepts to the field of procedural learning. To this end, they studied the positive and negative effects of conceptually-oriented explanation prompts in a complex e-learning module on learning outcomes (i.e., conceptual and procedural knowledge), learning processes (i.e., prompts responses and annotations), and cognitive load (i.e., intrinsic and extraneous load). They found that explanation prompts had a positive effect on both the quantity and quality of explanations during learning, but that the prompts also prevented learners from considering necessary arithmetical operations during learning, thereby, hindering acquisition of procedural knowledge. They call this the *double-edged* effect of using prompts.

While modern education strongly advocates working in groups and collaborative learning environments, very little research actually moves beyond fuzzy “feel good” explanations as to how and why group learning can be beneficial. The work of F. Kirschner, Paas and Kirschner (this issue) distinguishes itself from much of the field by examining the cognitive underpinnings of collaborative learning. Their contribution to this issue continues their earlier work (F. Kirschner, Paas & Kirschner, 2009) which examined the interaction between complexity and group learning. In a nutshell, F. Kirschner et al.’s (this issue) research found that the efficiency of group learning is tempered by information complexity (referred to as element interactivity within cognitive load theory). Quite simply, when learning material is low in complexity, individual learning is superior. However, when learning complex information collaborative learning is more effective than individual learning. They confirm the “group” effect with mathematics based material. Both performance and cognitive load measures confirm their hypotheses. The result is significant in that it indicates that the “group” effect may have a degree of robustness and generality across disciplines. Interestingly, the authors propose that self-efficacy expectations (or more precisely collaborative collective efficacy) may be a meditating factor involved with the group effect. They suggest that when groups work together on complex instructions they have a heightened level of confidence in their ability, as they are aware they can spread working memory load amongst other members of the group. This interpretation of the results while interesting is at this stage speculative and requires further research including incorporating measures of collective efficacy (Hanham & McCormick, 2009). However, it is apparent from this paper and previous work that group learning is proving to be a promising vein of research for cognitive load theory. In addition, the research has the potential to build research bridges between purely cognitive and motivational understandings of group learning.

Learner Control and Choice

Giving learners control of their learning environments is quite a seductive prospective. The move from instructor based to student controlled learning environments certainly has much inherent appeal. Though the last decade has seen a strong global push in this direction at all levels of education, the empirical realities of this push are a little different. While some learners benefit from learner control others struggle greatly (Kopcha & Sullivan, 2007; Katz & Assor, 2007). As Corbalan, Kester and Van Merriënboer (this issue) so concisely note, "[T]oo much control causes cognitive overload and even experts might experience difficulties in selecting, sequencing and pacing huge amounts of information". The discussion of the following papers will examine studies that focus on the role of learner control and provided much needed evidenced based guidelines for their use in education.

Within a CLT framework, research has consistently shown that while high knowledge learners and advanced learners may benefit from student control formats, less experienced learners flounder due to the heavy cognitive load requirements of unstructured learner-controlled environments. Extensive research within the expertise reversal paradigm effect (see Kalyuga, Chandler & Sweller, 1998; Kayluga & Sweller, 2005; Kayluga, 2007) has established that the level of expertise is the key factor mediating the use of control in learning environments. The higher the expertise the more effective learner control is.

Using a cognitive load framework, Mihalca, Salden, Corbalan, Paas and Miclea (this issue) examined the role of learner control on performance and instructional efficiency using a genetics training program. In their study comparing three types of instruction (i.e., non-adaptive program control, adaptive program control, and learner control), they predicted that adaptive control would be more effective than both other groups as it better met the needs of learners than program control and was less load bearing than learner controlled environments. While there is some evidence that adaptive control was effective in terms of instructional efficiency the results did not generalize to test-performance measures (near or far transfer). While the study showed considerable promise for embedding adaptive program control into technology based instruction, there is a clear need to tease out the testing issues and replicate this work.

Schwamborn, Thillmann, Opfermann, and Leutner (this issue) investigated the potential benefits of utilizing learner generated and instructor provided illustrations with textual science based instructions. Utilizing CLT, they predicted that both learner and instructor generated computer illustrations would be advantageous for learning. Results indicated that instructor generated illustrations aided comprehension and understanding of science materials and led to less cognitive load and lower perceived task difficulty. Learner generated illustrations did not seem to be as effective. The findings were consistent with CLT and demonstrated that for inexperienced students, generating their own illustrations may impose excessive cognitive load and reduce learning. As the authors point out, this work needs to be extended with more experienced students where the results may turn out differently. Their work again demonstrates the instructional danger of introducing learner control too early in the learning process. This work encountered some of the load measurement issues that were common to most papers in this issue that will be addressed elsewhere in the wider context of the special issue.

Corbalan, Kester and Van Merriënboer (this issue) provide a clear assessment of the research on learner control to date. The authors constructed a well-designed study to investigate the role of learner versus program control on surface versus structural features of genetics-based materials. Results indicated that students benefitted more from learner control when they could exercise that control over selecting tasks with different surface elements. The results extended to test performance, mental effort and instructional efficiency. The “added value” of learner control did not seem to extend to structural learner control conditions. This study provides yet more evidence-based guidelines on the circumstances under which learner control can be effective.

Wetzels, Kester and Van Merriënboer (this issue) examined the broader issue of activating prior knowledge in instructional settings. While CLT has focused on the role of prior knowledge in constructing learning environments, less work has been done on ensuring activation and utilization of learners’ prior knowledge. The authors examined the role of mobilization and perspective-taking in activating prior knowledge. Both techniques can be regarded as “schema activators” in that they help build bridges with what we know and allow us to meaningfully link it to what we may be currently learning. Using biology instructions, the authors propose that mobilization would be more effective learning devices for learners with low prior knowledge while perspective-taking would be more effective for learners with high prior knowledge. As predicted, they found that the usefulness of schema activation was tempered by the prior knowledge of learners, with low knowledge learners benefiting more from mobilization and high prior knowledge learners from perspective taking. However, the expected findings only generalized to the learning phase and did not hold for transfer tasks. The authors suggest the need for future research into other possible schema activation techniques such as problem based discussions and self-explanations.

Zhang, Ayres and Chan (this issue) continued the innovative group research of F. Kirschner, Paas and Kirschner (2009a, b) and F. Kirschner, Paas and Kirschner (this issue) by further examining the role of group work from a cognitive load perspective using web design materials. As with the F. Kirschner et al. (this issue) research, the authors assert that collaborative learning will be more effective than individual learning when learning complex information. They utilized webpage design as the instructional domain because of its high level of complexity. In addition to a group versus individual comparison, the authors included two task conditions namely an open-ended task (group exercised their choice of web based design) and a closed task (where the group had no choice other than to design a personal web page). This study extends the F. Kirschner et al. (2009a, b; this issue) work by conducting the study under normal classroom and homework conditions. Results from the study confirmed the “group” effect but only for the open-ended task. As predicted, students working on open-ended tasks outperformed students working on predetermined tasks by the teacher. The findings of this work are significant as it not only indicates that the “group” effect may have a degree of robustness and generality not only across disciplines but may also be applicable to “real” educational settings.

Studies into Animation and Multimedia Instruction

With the widespread availability and use of computers in teaching and learning has come a greater use of animation and multimedia. This increase may have positive effects but if not used judiciously and/or properly may lead to sub-optimal environments. For example, research has shown that under many conditions animations are not more effective than pages in a book (see Tversky, Morrison, & Betrancourt, 2002), and multimedia designs can be negatively impacted by split-attention (see Ayres & Sweller, 2005) or redundancy (see Chandler & Sweller, 1996). The articles in this section use CLT to continue our understanding of the factors influencing instructional animation and multimedia designs.

Amadieu, Mariné and Laimay (this issue) investigated the influence of cuing on animated instructions. A cuing strategy that zoomed in on each critical step of a dynamic system and hid irrelevant information was compared with an identical non-cued animation. Consistent with previous CLT research (see Sweller & Chandler, 1994), results indicated no significant effects for low element interactivity materials, other than improvements of test scores over learning repetitions. For high element interactivity materials there was a significant interaction in that the cuing group improved scores over repetitions, but the non-cued group did not. For a problem solving task which was only completed at the end of the final repetition, the cued group performed significantly better. The cognitive load measures provided similar patterns. No significant differences were found on the mental effort measure; however, a combined difficulty scale did reveal results. There was an interaction that showed that the cued group found the materials less difficult over the repetitions. Hence support can be found that the cued group experienced less cognitive load over the period of trials. As the authors argued, searching and extracting relevant elements within an animation could be considered an additional task and therefore can be considered as extraneous cognitive load and helpful to learning. The results generally support this position.

De Koning, Tabbers, Rikers, and Paas (this issue) also investigated cuing within instructional animations. In this study, the density of elements presented during the animation was varied, giving a 2 (cued vs. non-cued) x 2 (high vs. low speed) design. In the high-speed condition more elements had to be processed than in the low-speed condition. Comprehension and transfer tests were conducted and a single mental effort subjective measure was collected after each phase (i.e., instruction, comprehension and transfer). Results indicated no significant group differences or interactions on the test scores. For mental effort, one significant effect was found after instruction, in that less mental effort was reported when the animation was presented at high speed compared with low speed. Overall it can be concluded that varying the speed of the animation had no effect on learning outcomes. Although less mental effort was required in studying the higher-speed animation, performance data did not corroborate cognitive load measures. Interestingly, no cuing effect was found and as argued by the authors, may suggest that the cuing effect is not very robust.

Spanjers, Wouters, Van Gog and Van Merriënboer (this issue) investigated the effectiveness of segmentation in the context of the expertise reversal effect (see Kalyuga, Ayres, Chandler, & Sweller, 2003). Results indicated no interaction for learning outcomes, although learners in the segmented condition scored higher than those in the continuous condition. Significant interactions were found on the mental effort and efficiency measures. In both cases, the interactions were caused by low prior knowledge students benefiting from segmentation, whereas, for students with high prior knowledge segmenting made no difference. The finding that the usefulness of segmentation is in line with other research showing that instructional animations cause extraneous cognitive load and need some kind of compensatory mechanism that reduces this load (see Ayres & Paas, 2007a, b). However, this need for a buffer to the negative influence of animated design disappears for learners with expertise in the domain, as hypothesized it is likely that they are more able to cope with the transient information. Whereas test scores did not find an expertise reversal effect, measures of mental effect and efficiency did. Students with low prior knowledge invested less mental effort and performed more efficiently in the segmented format, indicating reduced cognitive load. Not obtaining the required interaction on test scores weakens the results to some degree, although learning more efficiency is an advantage (Van Gog & Paas, 2008).

Schmidt-Weigand and Scheiter (this issue) investigated the influence of high versus low spatial wording of text and the availability of an accompanying animation on learning. Results indicated no textual main effect for retention, transfer and visual learning tasks, but a significant advantage for groups who had the additional animations. Also, there was a significant interaction on retention and visual tests, where the animated design was most effective in conjunction with low spatial text. It should be noted that the means for the transfer scores were higher than the retention scores, suggesting that they may not have been a true test of transfer as claimed. For the cognitive load measures, the low spatial text groups reported higher overall cognitive load than the high spatial text groups. Moreover, cognitive load in the animated format was lower than in the non-animated format (all measures except germane load), but no significant interactions were identified. The results are clear in that additional animations led to superior learning with reduced cognitive load. Adding animations are particularly helpful for text that is considered less spatial in nature. However, the authors struggle to pinpoint the underlying reasons for the effect.

Kühl, Scheiter, Gerjets and Edelmann (this issue) investigated the effectiveness of animated instructions and the modality effect. In a 2 (animation vs. static) x 2 (spoken vs. written text) design students were required to learn about fish motion. On test scores, animations were shown to be superior to statics on the transfer measure only (no difference for factual knowledge or pictorial recall), and spoken text was superior to the written text for factual knowledge and pictorial recall (no difference on transfer). An extraneous load measure yielded a marginal effect for animation only (higher for statics) and a germane load measure found only a modality effect (higher for static). These results were partially in line with what was expected, positive effects were found for both animation and a combination of spoken text and visualizations (modality effect). However, for each main effect only certain tests were decisive. The authors go to some lengths to explain why some tasks were more suitable than others; however, these without further evidence are speculative. Incorporating seductive details (interesting but irrelevant information) into learning environments has been common-place in education at all levels for decades, assumedly to motivate learners.

As reported by Park, Moreno, Seufert and Brunken (this issue), the field seemingly is split over their value. Considerable research has shown that seductive details have negative effects on learning while other research indicates that their inclusion has no effect. Park et al. (this issue) devised a clever cognitive load intervention to shed light into the division in the field. Specifically, by manipulating the modality effect (Ginns, 2005; Moreno & Mayer, 1999); Tindall-Ford, Chandler & Sweller, 1997), the authors examine the inclusion of seductive details with low and high cognitive load instructional formats. Using biology materials, they assert on the basis of their study that the inclusion of seductive details may be of use in low load learning environments when learners have sufficient additional processing capacity to process and make use of seductive details. Results from performance measures indicate some limited support for this. The authors assert that seductive details may indeed have an integral place in learning as motivator. As discussed with the work of F. Kirschner et al. (this issue), research that brings together arousal and motivation factors within cognitive processes will likely lead cognitive load research in the future. While the study reported by Park et al. (this issue) has potential there needs to be far more research on seductive details within a cognitive load framework before any firm conclusions at all can be made to their potential value. In addition, the authors failed to find any support in terms of cognitive load measures which leads them to question the usefulness of the additivity hypothesis of cognitive load theory. However, the results may just as likely be a measurement issue as the scale used in their study is yet another non-validated variation of the mental load scale (Paas, 1992) which is increasingly being identified in the field as being highly problematic (see Van Gog & Pass, 2008). Indeed, measuring mental load has become the single most problematic issue in cognitive load theory and will be discussed at greater length in the summation paper.

Lee and Kalyuga (this issue) investigated the redundancy and expertise reversal effects within a multimedia environment. Using a common practice in China of teaching the native language, pinyin was used to create a redundancy format. Because Chinese characters are difficult to learn, they are often supported by the use of pinyin, which is a phonic transcription system based on English. In this experiment, pinyin was positioned above Chinese sentences that the participants were required to learn, and therefore as both forms of language say the same thing there is the potential for redundancy. Full pinyin, partial pinyin or no-pinyin was used to create three different treatments. In a 3 x 2 design two groups of learners (high or low prior knowledge) were differentiated by prior knowledge tests. Results indicated no significant differences on sentence comprehension or recognition of characters, but a significant interaction for pronunciation. For low prior knowledge students, no group differences were found but for students with higher prior knowledge, the partial pinyin strategy was superior to the no pinyin format. No redundancy effects were found. These results are interesting for a number of reasons. Firstly, if there was a true effect here, then the widespread use of pinyin characters may have some educational limitations. Secondly, the expertise reversal effect found here is in itself unusual, in that the main effect of the treatment was found for higher ability students. In most reported cases of expertise reversal, the effect disappears for high ability students (for a summary, see Kalyuga et al., 2003). To explain these results, the authors argue that it may be due to high intrinsic cognitive load experienced by low knowledge learners in particular, or that the redundant components are sufficiently short to not interfere too much with working memory processing. However, further experiments are needed to pinpoint the underlying answer.

Conclusions

It all sounds so simple. Cognitive load – the amount of *mental effort* that a learner expends - is based upon human cognitive architecture which consists of a severely limited working memory with partly independent processing units for visual/spatial and auditory/verbal information, which interacts with a comparatively unlimited long-term memory (Sweller, 1988). According to Paas, Tuovinen, Tabbers, and Van Gerven (2003), mental effort is "the aspect of cognitive load that refers to the cognitive capacity that is actually allocated to accommodate the demands imposed by the task; thus, it can be considered to reflect the actual cognitive load" (p. 64). Cognitive load is caused by, or maybe we should say dependent upon, the number of novel elements in learning materials that need to be kept in working (i.e., short-term) memory and the degree of interaction between those novel elements. The problem or task has a certain amount of cognitive load that is *intrinsic* to the task itself and which is affected by the expertise of the learner. In addition to the load intrinsic to the task, the way the learning task is presented and thus the way one learns and/or carries out the task also brings along a certain amount of cognitive load with it. If that load is facilitative of and/or functional for learning, then the load is considered to be *germane* for learning; if that load does not promote or advance learning, then the load is considered to be *extraneous* for learning. The goal of research on cognitive load is not necessarily minimizing cognitive load during learning, but optimizing it for learning. This means making sure that (1) instructional design keeps extraneous load to a minimum, (2) any load incurred by an instructional design is germane in nature, and (3) the correlation between total cognitive load and learning is optimized. Piece of cake, right?

Unfortunately this is not the case. In many recent presentations and journal articles and even in the articles within this special issue we see many good things evolving, but also a number of problems associated with *optimizing cognitive load*. Designing instructional experiments is not simple, because of the many complex factors that interact.

By observing the techniques of controlled randomized experiments hypotheses can be directly tested; however, there are at times confounding variables like prior-knowledge, element interactivity and redundancy, which are difficult to control, making it difficult to generate significant effects, or isolate the underlying factors. It is quite a struggle controlling variables and explaining the underlying factors in terms of the different types of cognitive load present. This is often made more confusing because of unsatisfactory attempts to coincide cognitive load measures with performance effects. Some articles in this collection have experienced such difficulties as the following conclusions indicate.

Good Aspects of the Special Issue

A very positive aspect of the papers in this collection is that the fundamentals of experimental psychology have been closely adhered to in that the papers are well conceived and theoretically driven, and most importantly nearly all are randomly controlled studies. Unlike much educational research, actual learning outcomes are measured instead of just perceptions or opinions of learning. Indeed, the abandonment of carefully designed randomized controlled studies in much educational research over recent years demonstrates just how difficult designing this important research can be.

Furthermore, our knowledge of the field has been advanced in a number of directions, including some highly innovate work, as the following concluding summary illustrates.

Much research in the field of educational psychology has restricted itself to simple learning environments. The studies dealing with *complex learning environments* have successfully made the move to real learning in real complex settings. Schwonke et al. (this issue) has advanced out knowledge as to how support and guidance (Van Merriënboer & Kirschner, 2007) can successfully be faded to adapt to the learning needs and expertise levels of learners. Kalyuga and Hanham (this issue) showed that transfer can be significantly enhanced by helping learners develop generalized schemata in which the new, to be acquired, knowledge can be flexibly assimilated and later used.

A second very important contribution is the expansion of cognitive load theory. CLT, and the instructional interventions which it has spawned, has often been seen as all encompassing and universal. Berthold et al. (this issue) as well as F. Kirschner et al. (this issue) have shown that is not necessarily the case. Berthold et al. showed that the use of an often employed intervention – explanation prompts – while positively affecting some learning might lead to so much extraneous load that other learning is impeded; the *double-edged sword*. F. Kirschner et al. illustrated that it is possible to attribute at least some of the presumed cognitive effects usually ascribed to CLT may also involve motivational and affective factors.

The studies in *learner control and choice* section unearthed some positive findings about learner control and choice. Corbalan et al. (this issue) found that students could benefit from learner control when they had control over task selection. Similarly Zhang et al. (this issue) found that task selection, a preference for open-ended task rather than system selected tasks, led to increased learning in a collaborative setting. Like most research into CLT effects, levels of prior knowledge were found to be an important moderating variable. Mihaca et al. (this issue) found evidence that adaptive program training could be effective by catering to students' individual prior knowledge. Wetzels et al. (this issue) found that low-knowledge learners could benefit most from mobilization techniques while high-knowledge learners benefitted most from perspective taking in schema activation strategies. However, Schwamborn et al. (this issue) found that asking inexperienced learners to generate their own illustrations in learning about a scientific concept was ineffective compared with instructor provided illustrations.

The studies on *animation and multimedia learning* have added to our knowledge of this field. Amadieu et al. (this issue) collected evidence suggesting that cuing can be effective as it is more likely to lead to stronger growth over learning cycles than non-cued animations, and enhances problem solving, particularly for materials high in element interactivity. However, de Koning et al. (this issue) found that cuing may lack robustness and dependent upon on particular conditions. Schmidt-Weigand and Scheiter (this issue) showed that animations could be successfully used in conjunction with textual information; however best results were achieved if the text information was low in spatial wording. There has been much research comparing animations with static representations, with mixed results (see Tverskey et al., 2002). Kühl et al. (this issue) found that animations could be more effective than statics when learning about fish movement. They also found a multimedia effect in that spoken text was superior to written text. Some new knowledge was also found concerning the segmentation of animation. Previous research (e.g., Mayer & Chandler, 2001), has found segmenting to be effective, but the findings of Spanjers et al. (this issue) suggest that perhaps only students with low levels of prior knowledge benefit most from this strategy. Lee and Kalyuga (this issue) also found an expertise reversal effect (see Kalyuga et al., 2003), in that high knowledge Chinese learners benefited from the additional use of pinyin information, even though it was considered redundant. This result is unusual in that in most cases of the expertise reversal effect, differences tend to be found in low knowledge learners only. Finally, Park et al. (this issue) widened the research into the multimedia effect by demonstrating that seductive details can be used effectively in conjunction with spoken text (narrative).

Bad Aspects of the Special Issue

Whereas many positives can be found in the studies conducted, there are also some less positive aspects. A dominant theme to emerge throughout the contributions is that some effects were found that could not be explained by the authors. Often these were unexpected, or involved conflicting information. Many measures of performance were often collected in these studies, for example recall, factual information, near and far transfer, which on the surface represent good research. However, this was often a double-edged sword as effects were often found on some measures, but not others, and these differences could often not be satisfactory explained and/or the explanations given were highly speculative. Furthermore, different measures of cognitive load (discussed in more detail in the next section) were also collected, that either were not helpful or presented direct conflicts. Whereas authors sometimes provided good and/or plausible reasons for unexpected or conflicting results, and suggested further research, many unresolved issues were left hanging. This situation may bring into doubt the worthiness of many single study experiments. When serious questions need to be answered the tried and trusted way is to complete further experiments and experimentally resolve such issues. It may be advisable to conduct follow up experiments when obvious discrepancies occur and to draw conclusions not matched by the actual findings. Carefully designed follow up studies have been the cornerstone of cognitive load research for two decades and it may be time for future research to return to this strong base.

Finally, we note a tendency of some of the researchers to grasp at what could be called the “marginally significant” or “trend” straws. Failing to find significant results, non-significant results are presented and sometimes even exposed to post-hoc analyses which actually may not be carried out on non-significant main effects.

Ugly Aspects of the Special Issue

Since the early development of CLT there has been a need for measures of cognitive load. Claiming specific effects are caused by increases in cognitive load fits the theory but without independent measures, is a circular argument. Whereas classical dual-task methodologies for monitoring working memory differences have been applied (see Chandler & Sweller, 1996; Sweller, 1988), they may be somewhat limited for studies that involve complex instructional designs. For many years the subjective measure of cognitive load, a mental effort scale, developed by Paas (1992), and a difficulty scale (see Cerpa, Chandler & Sweller, 1996) have been used successfully, without too much controversy. However, with the conceptualization of three different types of load into intrinsic, extraneous and germane (Sweller et al., 1998), there has been a greater need for individual measures to support theoretical arguments based on different types of load (see Ayres, 2006). To some extent this has become the holy grail of CLT research. We seriously doubt whether this is possible. Researchers have certainly tried to develop new scales, based on more precise wording of difficulty or mental effort linked to an aspect of the learning process, but with little success. As discussed by Van Gog and Paas (2008), there have been considerable discrepancies in the wording of cognitive load measures, when they are collected, and how efficiency is used. The studies in this special issue reflect this somewhat chaotic state, and from our perspective we consider it the *ugly* side of CLT research.

It is clear from simple examination of these studies that there are number of issues associated with cognitive load measures. Often performance test results do not correlate with the subjective measures. There may be one significant effect (either performance or load) but not both, and occasionally they conflict each other or the underpinning theoretical argument. Inevitably, attempts to measure more than one type, fails as either they are highly correlated or inconsistent. Sometimes researchers do not report individual measures and simply group them together under the umbrella of overall cognitive load. In other words, measuring cognitive load has become highly problematical.

To sum up, the papers in this special issue have some excellent aspects to them. They are theory driven and generally use randomized designs in controlled settings. They have extended our knowledge of the field and added some exciting new directions to cognitive load theory, particularly collaboration, user-control and learning in complex environments. There are however some problematical issues; namely an over-reliance on single experiment studies often leaving unexplained and conflicting results, and ineffective uses of cognitive load measures.

References

Amadieu, F., Mariné, C., & Laimay, C. (this issue). The attention-guiding effect and cognitive load in the comprehension of animations. *Computers in Human Behavior, xx*, xxx-xxx

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction, 16*, 389-400.

Ayres, P. & Sweller, J. (2005). The split-attention principle in multimedia learning. In R.E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 135-146). New York: Cambridge University Press.

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, 6*, 811-820.

Bassok, M., & Holyoak, K. J. (1993). Pragmatic knowledge and conceptual structure: Determinants of transfer between quantitative domains. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 68-98). Norwood, NJ: Ablex.

Berthold, K., Röder, H., Knörzer, D., Kessler, W., & Renkl, A. (this issue). The double-edged effects of explanation prompts. *Computers in Human Behavior*, *xx*, xxx-xxx

Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research, 15*, 345-367.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology, 10*, 151-170.

Corbalan, G., Kester, L., & Van Merriënboer, J. J G. (this issue). Learner-controlled selection of tasks with different surface and structural features: Effects on transfer and efficiency. *Computers in Human Behavior, xx*, xxx-xxx

Craik, F. I. M., & Lockhart, P. (1072). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

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*, 731-746.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (this issue). Attention cueing in an instructional animation: The role of presentation speed. *Computers in Human Behavior, xx*, xxx-xxx

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313-331.

Hanham, J., & McCormick, J. (2009). Group work in schools with close friends and acquaintances: Linking self-processes with group processes. *Learning and Instruction, 19*, 214-227.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23-31.

Kalyuga, S., Chandler, P. & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40*, 1-17.

Kalyuga, S., & Hanham, J. (this issue). Instructing in generalized knowledge structures to develop flexible problem solving skills. *Computers in Human Behavior, xx*, xxx-xxx

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development, 53*, 83-93.

Katz, I. & Assor, A. (2007). When Choice motivates and when it does not. E*ducational Psychology Review*, *19*, 429-442.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009a). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review, 21*, 31-42.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009b). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior, 25*, 306-314.

Kirschner, F., Paas, F., Kirschner, P. A. (this issue). Superiority of collaborative learning with complex tasks: A research note on an alternative affective explanation. *Computers in Human Behavior,xx, xx-xx.*

Kirschner, P. A., Sweller , J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *46*, 75-86.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review*, *19*, 239-264.

Kopcha, T. J., & Sullivan, H. (2007). Self-presentation bias in surveys of teachers' educational technology practices. *Educational Technology Research and Development*, *55*, 627–646.

Kühl, T., Scheiter, K., Gerjets, P., & Edelmann, J. (this issue). The influence of text modality on learning with static and dynamic visualizations. *Computers in Human Behavior, xx*, xxx-xxx

Lee, C. H., & Kalyuga, S. (this issue). Effectiveness of on-screen pinyin in learning chinese: An expertise reversal for multimedia redundancy effect. *Computers in Human Behavior, xx*, xxx-xxx

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.

Mihalca, L., Salden, R. J. C. M, Corbalan, G., Paas, F., & Miclea, M. (this issue). Effectiveness of cognitive-load based adaptive instruction in genetics education. *Computers in Human Behavior, xx*, xxx-xxx

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. J*ournal of Educational Psychology, 91*, 358-368.

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. E., Tabbers, H., & van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-71.

Park, B., Moreno, R., Seufert, T., & Brünken, R. (this issue). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior, xx*, xxx-xxx

Schmidt-Weigand, F., & Scheiter, K. (this issue). The role of spatial descriptions in learning from multimedia. *Computers in Human Behavior, xx*, xxx-xxx

Schwamborn, A., Thillmann, H., Opfermann, M. & Leutner, D. (this issue). Cognitive load and instructionally supported learning with provided and learner-generated visualizations. *Computers in Human Behavior, xx*, xxx-xxx

Schwonke, R., Renkl, A., Salden, R. J. C. M, & Aleven, V. (this issue). Effects of different ratios of worked solution steps and problem solving opportunities on cognitive load and learning outcomes. *Computers in Human Behavior, xx*, xxx-xxx

Spanjers, I. A. E., Wouters, P., Van Gog, T., & Van Merriënboer, J. J. G. (this issue). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior, xx*, xxx-xxx

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

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

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction, 16*, 165–169.

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

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

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3,* 257-287.

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 it's Design Implications for E-Learning. *Educational Technology, Research and Development*, *53*, 5-13.

Van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning*. Mahwah, NJ: Lawrence Erlbaum.

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.

Wetzels, S. A. J., Kester, L., &. Van Merriënboer, J. J. G. (this issue). Adapting prior knowledge activation: Mobilisation, perspective taking, and learners' prior knowledge. *Computers in Human Behavior, xx*, xxx-xxx

Zhang, L., Ayres, P., & Chan, K. K. (this issue). Examining different types of collaborative learning in a complex computer-based environment: A cognitive load approach. *Computers in Human Behavior*, *xx*, xxx-xxx