

Using Cognitive Load Theory to Structure MOOCs and Other Forms of Computer-Based Learning

Ouhao Chen¹, Geoff Woolcott¹, John Sweller²

1. School of Education, Southern Cross University, Lismore, Australia, 2480
2. School of Education, University of New South Wales, Australia, 2052

Correspondence concerning this article should be addressed to John Sweller, School of Education, University of New South Wales. E-mail: j.sweller@unsw.edu.au; Phone (w) +61 2 9385 1984.

Abstract

A massive, open, online course (MOOC) is a form of computer-based learning that offers open access, internet-based education for unlimited numbers of participants. However, the general quality and utility of MOOCs has been criticized. Most MOOCs have been structured with minimal consideration of relevant aspects of human cognitive architecture and instructional design principles. This paper suggests cognitive load theory, with its roots embedded in our knowledge of human cognitive architecture and evolutionary educational psychology, is ideally placed to provide instructional design principles for all forms of computer based learning, including MOOCs. The paper outlines the theory and indicates instructional design principles that could be used to structure online learning and to provide an appropriate base for instructional design when using computer-based learning.

Keywords: Cognitive Load Theory, MOOCs, Instructional Design, Human Cognitive Architecture, Evolutionary Educational Psychology.

Introduction

The increasing availability of MOOCs since their inception in 2008 provides open access, online courses to anyone who wishes to enroll outside institutional walls, without necessarily obtaining a formal qualification (Hew, 2015; McAuley, Stewart, Siemens, & Cormier, 2015). This availability and an obvious demand for MOOCs has led to research investigating factors that may improve their instructional quality (Deimann & Vogt, 2015; Moe, 2015; Siemens, 2013). Most MOOCs, however, are not organized according to current knowledge of human cognition, despite arguments of their origins in the learning sciences (see discussion in Clarà & Barberà, 2014). In this paper, we suggest that cognitive load theory (Sweller, 2012; Sweller, Ayres, & Kalyuga, 2011), which is based on human cognitive architecture, can be used to build more instructionally effective online learning materials, including MOOCs. We will begin by outlining some of the theoretical and practical issues that have been raised with respect to online learning, followed by a summary of cognitive load theory and some of the instructional effects generated by the theory that we hypothesise are relevant to computer-based learning such as MOOCs.

Some issues that have been raised concerning the effectiveness of computer-based learning and MOOCs

Since their inception, MOOCs have been constructed with little basis in the cognitive processes required for effective and efficient online learning. Despite this omission, most of the criticisms of MOOCs relate to general factors such as motivation, participation and study time, rather than providing a deeper analysis of cognitive processes (Champaign et al., 2014; El-Hmoudova, 2014; Hew, 2015; Zheng, Rosson, Shih, & Carroll, 2015). We will briefly discuss some of those general factors.

Some general factors associated with constructing effective MOOCs

Champaign et al. (2014) investigated correlations between the time spent on specific online courses and students' performance and skill. A strong negative correlation was found between students' skill levels and time spent studying online courses, indicating that students with higher expertise in a specific domain required less time spent on learning relevant materials. There was a strong positive correlation between time spent on using online learning courses and conceptual learning showing that the more time students spent on using online learning resources, the better conceptual learning would be. The authors expressed concern, however, that the wide demographic range and skill sets made it difficult to isolate and correlate factors associated with individual learning.

Hew (2015) investigated design factors related to student engagement in online courses by examining student reflection data from participants in three student MOOCs on programming languages, literature, and arts & design. He found that there were five factors, in order of importance, that were influential in student engagement in these courses: (1) the courses were problem-oriented with clear explanations; (2) lecturers were willing to interact with students and their passion for teaching was evident; (3) students could share knowledge and ideas among peers; (4) students were involved in learning and thinking about what they were learning; and, (5) there were a variety of online resources and activities that engaged students according to their diverse learning preferences. Although this study was not an investigation of cognition or instructional design, Hew does point to building knowledge in small increments as an essential feature of online learning.

El-Hmoudova (2013) explored the issue of large course numbers but high drop-out rates by considering various learning style preferences indicated by students. Many students had medium levels of preference for Visual, Active and Sequential learning styles. Students who indicated a strong preference for a particular learning style, also required that the online learning environment fitted that style. Zheng et al. (2015) also emphasized drop-out rates by

conducting interviews to investigate learners' motivation and learning patterns that might influence drop-out rates when using MOOCs. Results indicated that preparing students to get ready to transit from high-school to college motivated students' access to MOOCs. Another factor motivating students to participate in MOOCs was social interaction—the initial motivation in joining MOOCs, at least for some students, was to connect with other people. Some learning patterns were also identified as important in completing MOOCs, such as taking them as regular school classes or as a member of a study group with students working together online to reach learning goals. Of additional interest with regard to motivation is the view of Wang and Baker (2015) that students who completed MOOCs were more interested in the course content itself, whereas students who failed to complete the course regarded MOOCs as merely a learning experience.

Some factors relating specifically to the learning quality of MOOCs

There are researchers, however, who have commented on the lack of investigation into upgrading the current low learning quality of MOOCs, including concerns expressed about quality of content, design, and instructional delivery (Bali, 2014; Fischer, 2014). Ossiannilsson, Altinay, and Altinay (2015) have shown that learning design and learning environment affect learners' experiences and, along with the openness, personalization, choice-based learning, and a collaborative learning context, are factors that might be used to refer to aspects of the quality of MOOCs. Such factors may account for Nkuyubwatsi's (2013) evaluation of the learning environment and design of MOOCs compared with face-to-face learning, from the students' perspective. Nkuyubwatsi suggested that MOOCs might be superior to face-to-face courses in terms of openness, availability, diversity, and delivery, but face-to-face teaching might offer more interactivity among peers. Fischer (2014) has argued that since a large number of MOOCs are designed to follow traditional teaching formats and,

as well, are built on business models, they may not be optimal, as far as learning quality goes, for online learning.

What is clear from most critical studies is the lack of investigation in the pedagogy aspects of MOOCs (Bali, 2014; Deimann & Vogt, 2015; Moe, 2015). Equally, it is apparent that neither MOOCs that are truly interactive and provide collaborative learning spaces, nor those that are designed to deliver content (see Siemens, 2013) appear to take into account the cognitive demands placed on online learners.

Based on this review of MOOCs and other computer-based learning research studies, the future design and evaluation of computer-based learning environments may need to be grounded in human cognition in order to design effective online instruction. In the next section, cognitive load theory, which is based on human cognitive architecture, will be introduced in order to comment on cognitive aspects of the structure of MOOCs and other forms of computer-based learning.

Cognitive load theory

Human cognitive architecture

Knowledge can be divided into biologically primary and biologically secondary categories (Geary, 2007, 2008, 2012; Geary & Berch, 2016). Biologically primary knowledge is obtained naturally and effortlessly, as we have evolved to acquire this category of knowledge, therefore, we do not need instruction to acquire it. Most generic cognitive skills such as general problem solving strategies are biologically primary (Tricot & Sweller, 2014).

With the development of human culture, a different category of knowledge is required, and the nature of that new knowledge, biologically secondary knowledge, is distinct from biologically primary knowledge. Unlike the acquisition of biologically primary knowledge, secondary knowledge requires conscious effort and explicit instruction (Kirschner, Sweller, &

Clark, 2006; Sweller, Kirschner, & Clark, 2007) with much of the knowledge being domain-specific rather than generic-cognitive in nature. As biologically secondary knowledge requires explicit instruction, we need to consider how we should best impart the relevant information. Cognitive load theory focuses on this issue and it is this issue that is critical to the success of all forms of computer-based instruction.

The theory is based on human cognitive architecture. That architecture is concerned with the organization of the structures, functions and processes that allow us to learn, think and solve problems associated with the biologically secondary knowledge that is central to instructional design (Sweller et al., 2011). The architecture can be described as a natural information processing system analogous to the system used by biological evolution. It can be described by five basic principles (Sweller & Sweller, 2006).

Information store principle. In order to function, human cognition requires very large amounts of domain-specific information stored in long-term memory (De Groot, Gobet, & Jongman, 1996). There are no known limits to the capacity and duration of information stored in long-term memory. Expertise in any educationally relevant area is dependent on the previous acquisition and storage of large amounts of biologically secondary, domain-specific knowledge.

The borrowing and reorganizing principle. The large amounts of information required for cognitive functioning are obtained primarily by borrowing information from other people. We read what others write, listen to what they say, and imitate what they do (Bandura, 1986). The ability to obtain information from others occurs in few other species and is a biologically primary skill. That borrowed information is usually restructured and reorganised when combined with previous information held in long-term memory before being stored.

Randomness as genesis principle. When information cannot be borrowed from others, it can be generated during problem solving using a random generate and test procedure. Randomly generated information needs to be tested for effectiveness with effective information stored in long-term memory and ineffective information abandoned.

Narrow limits of change principle. Random generation and testing of moves occurs in working memory but cannot function effectively if many elements must be tested. This lack of effectiveness is because working memory has limited capacity (Cowan, 2001; Miller, 1956) and limited duration (Peterson & Peterson, 1959) when dealing with novel information from the environment. In addition, the limited amount of information that can be processed in working memory avoids large and rapid changes of long-term memory that could destroy its functionality.

Environmental organizing and linking principle. This last principle reveals how stored information can be retrieved from long-term memory and be used to determine action. Activated by external signals from the environment, relevant information stored in long-term memory can be transferred to working memory to determine appropriate actions. Unlike when dealing with novel information from the environment, working memory can handle unlimited amounts of well-structured information transferred from long-term memory (Ericsson & Kintsch, 1995).

The cognitive architecture described in this section provides the base for cognitive load theory. We obtain novel information from other people using the borrowing and reorganizing principle or, if such information is unavailable, attempt to generate information using the randomness as genesis principle. That novel information must be processed in a limited capacity, limited duration working memory according to the narrow limits of change principle before being stored in long-term memory via the information store principle. Once stored, the

now processed, familiar information is triggered for use by environmental cues according to the environmental organizing and linking principle.

Element interactivity and types of cognitive load

The above architecture explains how information is both acquired and subsequently used. As such, it can be used to guide instructional design, including computer-based instructional design (e.g., MOOCs and computer-assisted learning). We will discuss why some information imposes a heavy working memory load next, using the concept of element interactivity.

Element interactivity. Element interactivity provides an index of the complexity of tasks and also is the basic concept used to determine categories of cognitive load. Interactive elements are elements that must be processed in working memory simultaneously as they are logically connected (Sweller, 2010). For example, if a student is learning how to manipulate an algebraic equation, element interactivity and working memory load are high because any change in one part of an equation has consequences for the entire equation. The elements of the equation interact and it is the interaction of those elements that must be learned. In contrast, if a student is learning some of the vocabulary of a second language, the translation of many individual words can be learned independently because they are not logically connected. For this material, element interactivity and cognitive load are low because the translation of each word can be learned without reference to or consideration of any of the other words.

Two categories of cognitive load. Two categories of cognitive load, intrinsic and extraneous load, are discussed in a recent version of cognitive load theory (Kalyuga, 2011). Intrinsic load is imposed by the nature of learning materials or the levels of element interactivity determined by the learning materials. The examples above of learning to manipulate algebraic equations or learning the translation of vocabulary items of a second language provide instances of variation in intrinsic cognitive load.

Intrinsic cognitive load can be changed not only by varying the nature of the information to be learned, but also by varying the expertise of learners. For example, if an expert in elementary algebra, rather than a novice student, is faced with an algebraic equation, element interactivity and intrinsic cognitive load may be low rather than high because, using the environmental organizing and linking principle, an expert can use information in long-term memory to treat the entire equation as a single element.

Extraneous load, unlike intrinsic load, is influenced by how learning materials are presented. This type of load should be reduced or eliminated as it interferes with learning. Element interactivity also determines extraneous load. If element interactivity is altered by instructional procedures without changing what needs to be learned, then extraneous cognitive load has been altered. For example, when studying a worked example, fewer interactive elements must be processed simultaneously in working memory compared to having novice learners attempt to solve the same problem because problem-solving generates higher levels of element interactivity (Sweller, 2010). When studying a worked example, only the moves needed for solution are processed. When solving a novel problem, additional moves not required for solution and leading to dead-ends are likely to be processed as well resulting in an increase in element interactivity.

Some cognitive load theory guidelines associated with constructing MOOCs

Cognitive load theory has not been applied to constructing MOOCs but there is every reason to suppose that the instructional design recommendations of cognitive load theory are directly applicable to MOOCs and other computer-based learning procedures. We can hypothesise that at least some of the problems indicated in the literature that have been associated with MOOCs are due to inadequate instructional designs that ignore what we know of human cognition and of evolutionary educational psychology. In this section, we will consider instructional guidelines based on cognitive load theory.

Cognitive load effects under a cognitive load theory framework are demonstrated when randomized, controlled trials indicate superior learning using an instructional condition generated by the theory compared to a more conventional instructional condition. We hypothesise that these effects can be used to indicate how to effectively structure computer-based learning platforms such as MOOCs to facilitate students' learning. They will be discussed in this section. Each effect discussed below can be used directly in computer-based learning.

The goal-free effect

Goal-free problems are designed to reduce extraneous load imposed by using means-ends analysis (Sweller, Van Merriënboer, & Paas, 1998). Assume a student is required to solve a geometry problem by finding a value for Angle x . If the student is a novice, means-ends analysis is a common way to solve the problem. This method requires students to focus on the goal of the problem and find a sequence of problem solving operators that can reduce differences between this goal and the givens of the problem, a process requiring many interacting elements that impose a heavy working memory load.

However, goal-free problems may be open questions, such as asking students to find the value of as many angles as possible. For this problem, the student only needs to look at the givens of the problem and find any new angle that can be calculated from those givens. Because only one angle needs to be considered at a time rather than a sequence of related angles leading to a goal, element interactivity is reduced and learners are in a better position to learn to recognize problem states and the moves that lead from those states. Ayres (1993) showed that using such goal-free problems is superior to using traditional goal-specific problems.

The worked example effect

Worked examples that provide full problem-solving guidance showing learners how to solve a problem are superior to learner problem solving that provides no guidance, indicating the worked example effect (Renkl, 2014). From the theoretical perspective outlined above, studying worked examples means borrowing well-structured knowledge from others via the borrowing and reorganizing principle, while problem solving in the absence of domain-specific, problem solving knowledge, requires learners to randomly generate solutions and test them for effectiveness via the randomness as genesis principle. Processing a worked example requires learners to deal with fewer interactive elements than randomly generating and testing potential moves, which would result in an increase in extraneous cognitive load.

Many research studies have investigated and supported the worked example effect in a variety of different domains (Cooper & Sweller, 1987; Kyun, Kalyuga, & Sweller, 2013; Paas, 1992; Rourke & Sweller, 2009; Sweller & Cooper, 1985). Podcast have been employed to investigate if worked examples can enhance students' mathematics performance (Kay & Edwards, 2012; Kay & Kletskin, 2012). Students responded positively to these example-based podcasts and improved their mathematics performance. These experiments, however, did not compare example-based design with other forms of instructions.

The use of worked examples will only be effective, however, if they are structured in a manner that will reduce extraneous cognitive load. Providing students with ineffectively structured worked examples to study, or providing examples to learners who do not need examples to study, can have negative rather than positive effects. These issues are discussed in several of the sections below.

The problem completion effect

Completion problems can act as a substitute for traditional worked examples which provide full solutions. Completion problems provide explicitly some solution steps but leave some key steps for learners to complete. For example,

$$2x + 10 = 14$$

$$2x = 14 - 10$$

$$x = ?$$

Van Merriënboer (1990) conducted the first extensive research on the effectiveness of completion tasks within the framework of cognitive load theory using a computer programming course. Students were randomly assigned to two groups: a traditional strategy group, in which students were required to design and code new computer programs, and a problem completion group, in which students were required to modify and extend existing computer programs. The problem completion group was superior to the traditional strategy group according to the post-test results. Van Merriënboer and De Crook (1992) undertook a similar study using computer programming. Their results on post-tests also indicated that problem completion was better than traditional problem solving.

Paas (1992) compared three conditions - worked example, problem completion and conventional problem solving - to investigate the problem completion effect. Results showed that worked example and problem completion groups were superior to a conventional problem solving group. However, the superiority of problem completion tasks may be limited to far transfer post-test tasks (Van Merriënboer, Schuurman, De Crook & Paas, 2002).

The split-attention effect

Split attention occurs when students are required to split their attention between multiple sources of information that have been separated either spatially or temporally and so need to be mentally integrated if they are individually unintelligible and unlearnable in isolation (Sweller et al., 2011). If the same information is physically integrated, thus obviating the need to mentally integrate it, extraneous cognitive load is reduced and learning is enhanced due to the reduction in the number of interacting elements.

This effect has been documented in multimedia learning. For example, Kalyuga, Chandler, and Sweller (1999) tested this effect on computer-based instructional material consisting of diagrams and texts. They found that physically integrated information enhanced learning compared to split-attention information. Al-Shehri and Gitsaki (2010) compared a split-attention design with an integrated design for learners' online reading performance. Results supported the view that online integrated materials enhanced students' learning. Similarly, Liu, Lin, Tsai and Paas (2012) investigated the split-attention effect on mobile learning in the domain of physics. The results also revealed that when text and picture were embedded in a mobile device, students performed better.

The redundancy effect

Within the framework of cognitive load theory, any information which is not required for learning is regarded as redundant information (Sweller et al., 2011). Examples are the provision of identical information in multiple forms such as listening to and reading the same text or the provision of unnecessary decorative information (seductive details) such as cartoons associated with text. Such information is likely to be processed by working memory and therefore, increases extraneous working memory load. Redundant information can look much like split-source information that also deals with multiple sources of information, but the redundancy effect emphasizes that the multiple sources of information can be understood without reference to each other. For example, in multimedia learning, redundancy may occur when the same information is presented as a diagram and as text, or as text presented in both written and spoken forms, as indicated above.

When redundancy occurs, rather than physically integrating the multiple sources of information, one source should be eliminated. For example, Chandler and Sweller (1991) presented one group of biology learners with a diagram of the flow of blood through the

heart, lungs and body along with a set of written statements describing the same flow of blood. Another group was presented the same diagram without the statements. The results indicated enhanced learning by the group without the written statements. Mayer, Heiser, and Lonn (2001) found that the results of university learners who studied narrations with animations simultaneously were superior to those of learners who were presented the same animations with narrations as well as on-screen text which repeated the content of the narrations.

The modality effect

Just as the use of diagrams and text depends on the relations between diagrams and text resulting in the split-attention or redundancy effects, the use of visual and auditory information also depends on the logical relations between the two sources of information. As indicated above, repeating spoken information with the same written information leads to redundancy and should be avoided. Nevertheless, under certain closely defined conditions, the use of both visual and auditory information can be advantageous. Those conditions are similar to the ones leading to the split-attention effect. If a diagram, which is always visual, cannot be understood without text, thus providing the conditions under which the split-attention effect is obtainable, rather than physically integrating the diagram and written text, it may be preferable to present the text in spoken rather than written form.

Working memory has two sub-components, the visuospatial sketch pad and the phonological loop (Baddeley, 1992). Tindall-Ford, Chandler, and Sweller (1997) pointed out that using both visual and audio channels is superior to only using the visual channel when learning. The explanation may be that using both channels increases the usable capacity of working memory (Penney, 1989). Many research studies have investigated the modality effect (Kalyuga, Chandler, & Sweller, 2000; Mayer & Moreno, 1998; Tindall-Ford et al.,

1997). Tindall-Ford et al. (1997) compared text presented in audio format and diagrams in visual format to text and diagrams both presented in a visual format. The results indicated that the mixed format performed better than the single format on a subsequent test. Similarly, Mayer and Moreno (1998) investigated the modality effect in multiple experiments explaining the process of lightning formation or a car's braking system. One group received the narrated text describing major steps, the second group was presented the same words on screen. Results confirmed the superiority of the mixed format.

The transient information effect

The modality effect indicates the potential advantages of presenting information in auditory form. However, there are limits to the use of the auditory modality when presenting text. Spoken text, unlike written text, is transient. If written text is lengthy and high in element interactivity, learners can readily choose to read difficult parts repeatedly. They can stop to think about sections of written text and return to it as needed. Such activity can be difficult or impossible with spoken text because spoken text is transient. What is heard now, disappears to be replaced by new information. If it is high in element interactivity, the new information may need to be processed in working memory along with the previous information but that previous information has disappeared. Based on cognitive load theory, high element interactivity information should not be presented in spoken form. Students may not be able to process audio information before it disappears and it may be physically difficult or impossible to retrieve such transient information, leading to the transient information effect.

The transient information effect indicates that instructional procedures that present high element interactivity information in a transient form such as spoken text, can result in learning decrements compared to the equivalent written text. The transient information effect

has implications for the modality effect. Leahy and Sweller (2011) obtained a reverse modality effect with written information superior to spoken information. A conventional modality effect with superior spoken information was obtained when the high element interactivity text was divided into smaller segments.

Spoken information is not the only source of transient information. Many videos and animations are equally transient. Wong, Leahy, Marcus, and Sweller (2012) investigated the transient information effect using both spoken and animation-based instruction. They obtained similar results to Leahy and Sweller (2011) testing for the modality effect. For animations, results indicated that lengthy, animation-based instruction was less effective than static graphics, while, for segmented information, animation-based information was superior.

The transient information effect is particularly relevant to computer-based learning systems. Modern instructional technology allows us to use spoken words, videos and animations much more readily than in the past. The ability to use those techniques does not in itself justify this use. Human cognition predates instructional technology and provides the ultimate determinant of the effectiveness of any particular procedure.

The imagination effect

When studying a procedure or concept, imaging the procedure or concept may be superior to traditional studying, indicating the imagination effect (Leahy & Sweller, 2004). Some research studies have shown positive results when applying an imagination strategy for procedural and concept learning (Leahy & Sweller, 2004, 2005, 2008). Leahy and Sweller (2004) compared an imagination strategy with traditional study. Results indicated that using an imagination strategy would be better than studying instructions. Leahy and Sweller (2005, 2008) extended their results of 2004 and indicated that the imagination effect might be more likely to be obtained with high element interactivity materials.

Based on multimedia learning, Leutner, Leopold and Sumfleth (2009) compared drawing pictures with imagining pictures of molecular structures in chemistry. Results indicated that imagining pictures decreased levels of cognitive load and improved students' understanding, whereas, drawing pictures increased levels of cognitive load and decreased understanding.

The isolated elements effect

The isolated elements effect indicates that initially presenting a set of isolated elements of information rather than whole complexes of interactive elements in instructional materials may reduce excessive intrinsic load. A disadvantage of this strategy is that the information is in the form of isolated elements and so students are unlikely to learn the relations between the isolated elements initially. However, these isolated elements could help learners form partial schemas at the outset and then form the whole schema at the following phase after receiving instructions concerning the relations among those isolated elements (Pollock, Chandler, & Sweller, 2002).

Blayney, Kalyuga, and Sweller (2010) compared two instructional formats: one used an isolated-interactive elements format; and another used a fully interactive elements format. This experiment was in the domain of accountancy with university students as participants. Novices benefited more from an isolated-interactive elements format, compared to more knowledgeable learners who demonstrated superior results after studying with a fully interactive elements format. In this experiment, when more knowledgeable learners were presented with an isolated-interactive elements format, they needed to integrate those isolated simple elements with their knowledge base, which required extra working memory resources, interfering with learning. The results also indicate that the isolated-element strategy interacts

with learners' expertise, which is relevant to the expertise reversal effect discussed in the next section.

The expertise reversal effect

All instructional effects, including all cognitive load effects, have limits beyond which they no longer apply. The expertise reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003) provides a limit on most cognitive load effects due to an interaction between the characteristics of learners and tasks. Assume a new, cognitive load theory-based instructional procedure that is superior to a traditional procedure for novice learners. With increases in expertise, the relative effectiveness of the new procedure may decrease. Eventually, there may be no advantage of the new procedure and even a reversal of the relative effectiveness of the two procedures.

The redundancy effect provides the most common explanation of the expertise reversal effect. Information that is necessary for novices may become redundant and deleterious for more expert learners. For example, novices may require worked examples to assist them in understanding the solutions to problems. With increasing expertise, those worked examples may become decreasingly effective and eventually redundant, resulting in the expertise reversal effect (Kalyuga, Chandle, Tuovinen, & Sweller, 2001). As another example, Kalyuga et al. (2000) obtained an expertise reversal effect when testing learning from visual diagrams with or without audio text. During this experiment, students had training sessions to increase their expertise. Initially, the diagrams plus audio text condition was superior but as the students became more expert, the visual only condition became superior to the visual plus audio text group.

Computer-based learning is ideally suited to dealing with the expertise reversal effect. According to the effect, the nature of instruction should change with increasing levels of

expertise. By assessing learners on-line and altering their instruction depending on their levels of knowledge, it is possible to provide instruction that is more appropriate than a one-size-fits-all regime (Kalyuga & Sweller, 2004, 2005).

The guidance fading effect

Based on the expertise reversal effect, a learning environment should be tailored with changes in learners' expertise. The guidance fading strategy (Renkl, 2014) provides a way to appropriately tailor a learning environment. For example, we can change full worked examples into completion tasks (see the problem completion effect), followed by problem solving tasks (Van Merriënboer, Kirschner & Kester, 2003; Van Merriënboer, 1990). This fading strategy is based on the assumption that learners will still have sufficient working memory resources to process increasing problem-solving demands with an increase in expertise.

Renkl, Atkinson, Maier, and Staley (2002) compared backward and forward fading with traditional worked example and problem solving pairs. For both classroom and lab-based experiments, the fading strategy was superior to traditional worked example, problem solving pairs.

The pace of fading is another issue which should be mentioned when using a fading strategy. Reisslein, Sullivan, and Reisslein (2007) compared three speeds of fading: immediate, fast, and slow fading. For the immediate fading group, learners solved problems directly after instruction that used worked examples, while for the fast fading group, learners received a full worked example and then steps were omitted gradually (an omitted step for the second task, then two omitted steps for the next task and so on) based on a backward fading procedure, and for the slow fading group, every two worked examples, a step was omitted. The results of this experiment indicated an interaction between speed of fading and levels of

learner's expertise. More knowledgeable learners benefited more from immediate and fast fading conditions, compared to less experienced learners who learned more from the slow fading condition. Therefore, well-guided and slow-paced instructional procedures that reduce extraneous load are more suitable for novices.

The element interactivity effect

The expertise reversal effect is usually treated as an independent, cognitive load effect but recently, it has been suggested that it is a specific example of a more general effect, the element interactivity effect (Chen, Kalyuga, & Sweller, in press). The element interactivity effect has been known for many years (Sweller, 1994, 2010; Sweller & Chandler, 1994). Like the expertise reversal effect, it is a second order effect that provides a limit to a first order effect. In the case of the expertise reversal effect, that limit is provided by levels of expertise. In the case of the element interactivity effect, the limit is provided by levels of element interactivity.

The element interactivity effect occurs when cognitive load effects that can be obtained using information that is high in intrinsic cognitive load disappear using information that is low in intrinsic cognitive load (Sweller & Chandler, 1994; Tindall-Ford et al., 1997). Indeed, not only can an effect disappear using low element interactivity material, it can reverse. Chen, Kalyuga, and Sweller (2015, 2016) obtained a conventional worked example effect using high element interactivity information but a reverse effect for low element interactivity information, providing a strong example of the element interactivity effect. This reverse of the worked example effect is called the generation effect and has been known for longer than the worked example effect. It occurs when learners provided an answer to a memory task perform worse on a subsequent task than learners asked to generate an answer themselves.

As indicated previously, there are two ways in which element interactivity can be altered. The first is to alter the nature of the task. For example, a task can be altered to reduce the number of interacting elements, as occurs when the isolated elements effect is obtained. A second way is to alter the levels of expertise of learners. It is this second process that leads to the relations between the expertise reversal and element interactivity effects.

Consider a person faced with solving the problem: $(a + b)/c = d$, solve for a . For someone just beginning to learn how to manipulate algebraic equations, this problem is likely to be a relatively high element interactivity, difficult problem. There are many elements that must be simultaneously processed in working memory in order to successfully solve the problem. In contrast, anyone with algebraic competency is likely to be able to easily process all of the elements in working memory. Many people can readily solve such problems mentally, without effort. Knowledge held in long-term memory can be retrieved and used to solve the problem using the environmental organizing and linking principle, as discussed above. For a person who has solved many such problems, the entire problem and its solution constitutes a single element. Element interactivity is low, rather than high, for an expert in this area.

Providing an expert with, for example, a worked example, would increase rather than decrease working memory load due to redundancy, explaining the expertise reversal effect. In that sense, the expertise reversal effect is merely a specific example of the element interactivity effect that is caused by changes in element interactivity with changes in expertise.

There are instructional consequences of this relation. In one experiment, Chen et al. (2015) taught novices how to calculate areas of geometric shapes. Two types of materials were investigated. One included regular geometry formulae that were low in element

interactivity, while another was designed to teach students how to calculate the areas of some geometric shapes by applying those geometry formulae. Calculating the areas was high in element interactivity for these students. Results indicated that high guidance via worked examples should be provided when studying materials high in element interactivity, indicating a worked example effect, whereas, for materials low in element interactivity, students performed better after generating answers by themselves rather than being shown the answer, indicating a generation effect.

However, with an increase in learners' expertise in a second experiment, the generation effect was found for both sets of materials as the increase in learners' expertise decreased the levels of element interactivity. High guidance for more knowledgeable learners became redundant, even for more complex information that imposed high levels of extraneous load. The results confirmed the element interactivity effect by indicating that the worked example effect is obtainable only for materials high in element interactivity. Simultaneously, the results indicated an expertise reversal effect with the worked example effect obtained for novices but a reverse worked example effect for more expert learners.

It follows that the expertise reversal effect and the element interactivity effect are intertwined (Chen et al., in press). Increasing the levels of learners' expertise reduces the levels of element interactivity, requiring reduced levels of guidance.

Application of cognitive load effects to MOOCs and other forms of computer-assisted learning

While many of the above cognitive load effects have been tested using various forms of computer-assisted learning, as far as we are aware, they have not been incorporated into MOOCs. Accordingly, in the absence of explicit evidence using randomised, controlled trials, our suggestion that MOOCs could be improved by the use of cognitive load theory and its

associated effects has the status of a hypothesis, indeed, a series of hypotheses, at least one for each effect, that require testing.

There are good grounds for assuming that the cognitive load effects apply equally to MOOCs as to any other form of presentation. Humans have evolved to process biologically primary, generic-cognitive information and biologically secondary, domain-specific information without reference to whether that information is presented digitally or as hard copy. For example, there is no evidence that text or diagrams presented on a screen differ from the same text or diagrams presented as hard copy. Our cognitive system is likely to process both in the same way with working memory and long-term memory functioning identically in both cases.

Of course, some effects are more likely when presented on a screen than when presented as hard copy. For example, the negative effects on working memory of transient information effects are far more likely when information is presented digitally than as hard copy simply because it is easier to present transient information digitally. Nevertheless, our cognitive system reacts in the same way to transient information irrespective of the presentation vehicle.

Conclusions

Cognitive load theory provides an integrated instructional design theory leading from our understanding of human cognitive architecture to the development of instructional design principles that we hypothesise are directly applicable to constructing computer-based learning modules such as MOOCs. The theory has provided guidelines for the construction of effective learning systems, and even though these guidelines have not been applied to MOOC development, they are directly applicable in online learning environments. The lack of

research studies on applying those cognitive load principles to MOOCs points to a fertile area of future research.

Cognitive load theory, along with its instructional recommendations, can be used to explain at least some of the failures associated with the use of MOOCs, especially their failure to convince learners to complete a course. Additionally, the emphasis of cognitive load theory on human cognitive architecture to generate hypotheses and randomized controlled experiments to test out those hypotheses provides us with recommendations that, because they have been carefully tested, are likely to be effective in MOOC environments. Computer-based learning modules based on cognitive load theory principles should be very different to most systems currently available. We expect such cognitive load theory-based systems to be superior to systems that do not use the recommendations of the theory.

References

- Al-Shehri, S., & Gitsaki, C. (2010). Online reading: A preliminary study of the impact of integrated and split-attention formats on L2 students' cognitive load. *ReCALL*, 22, 356-375.
- Ayres, P. (1993). Why goal-free problems can facilitate learning. *Contemporary Educational Psychology*, 18, 376-381.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556-559.
- Bali, M. (2014). MOOC pedagogy: gleaning good practice from existing MOOCs. *Journal of Online Learning and Teaching*, 10, 44-56.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*: Englewoods Cliffs, New Jersey: Prentice-Hall, Inc.
- Blayney, P., Kalyuga, S., & Sweller, J. (2010). Interactions between the isolated–interactive elements effect and levels of learner expertise: Experimental evidence from an accountancy class. *Instructional Science*, 38, 277-287.
- Champaign, J., Colvin, K. F., Liu, A., Fredericks, C., Seaton, D., & Pritchard, D. E. (2014). *Correlating skill and improvement in 2 MOOCs with a student's time on tasks*. Paper

- presented at the Proceedings of the first ACM conference on Learning@ scale conference (pp. 11-20). ACM
- Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.
- Chen, O., Kalyuga, S. & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology, 107*, 689-704.
- Chen, O., Kalyuga, S., & Sweller, J. (2016). Relations between the worked example and generation effects on immediate and delayed tests. *Learning and Instruction, 45*, 20-30.
- Chen, O., Kalyuga, S., & Sweller, J. (in press). The expertise reversal effect is a variant of the more general element interactivity effect. *Educational Psychology Review*. doi:10.1007/s10648-016-9359-1
- Clarà, M., & Barberà, E. (2014). Three problems with the connectivist conception of learning. *Journal of Computer Assisted Learning, 30*, 197-206.
- Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*, 347-362.
- Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences, 24*, 154-176.
- De Groot, A. D., Gobet, F., & Jongman, R. W. (1996). *Perception and memory in chess: Studies in the heuristics of the professional eye*. Assen The Netherlands: Van Gorcum & Co.
- Deimann, M., & Vogt, S. (2015). Editorial. *The International Review of Research in Open and Distributed Learning, 16*.
- El-Hmoudova, D. (2014). MOOCs motivation and communication in the cyber learning environment. *Procedia-Social and Behavioral Sciences, 131*, 29-34.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211-245.
- Fischer, G. (2014). Beyond hype and underestimation: identifying research challenges for the future of MOOCs. *Distance Education, 35*, 149-158.

- Geary, D. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Psychological Perspectives on Contemporary Educational Issues* (pp. 1-99). Greenwich, CT: Information Age Publishing.
- Geary, D. (2008). An evolutionarily informed education science. *Educational Psychologist*, *43*, 179-195.
- Geary, D. (2012). Evolutionary Educational Psychology. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook* (Vol. 1, pp. 597-621). Washington, D.C.: American Psychological Association.
- Geary, D., & Berch, D. (2016). Evolution and children's cognitive and academic development. In D. Geary & D. Berch (Eds.), *Evolutionary perspectives on child development and education* (pp. 217-249). Switzerland: Springer.
- Hew, K. F. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCs. *British Journal of Educational Technology*, *47*, 320-341.
- Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review*, *23*, 1-19.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*, 23-31.
- Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology*, *93*, 579-588.
- Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, *92*, 126-136.
- Kalyuga, S., & Sweller, J. (2004). Measuring Knowledge to Optimize Cognitive Load Factors During Instruction. *Journal of Educational Psychology*, *96*, 558-568.
- 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.
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351-371.

- Kay, R. H., & Edwards, J. (2012). Examining the Use of Worked Example Video Podcasts in Middle School Mathematics Classrooms: A Formative Analysis. *Canadian Journal of Learning and Technology*, 38, 1-20.
- Kay, R., & Kletskin, I. (2012). Evaluating the use of problem-based video podcasts to teach mathematics in higher education. *Computers & Education*, 59, 619-627.
- Kirschner, P., 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*, 41, 75-86.
- Kyun, S., Kalyuga, S., & Sweller, J. (2013). The effect of worked examples when learning to write essays in English literature. *The Journal of Experimental Education*, 81, 385-408.
- Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, 25, 943-951.
- Leahy, W., & Sweller, J. (2004). Cognitive load and the imagination effect. *Applied Cognitive Psychology*, 18, 857-875.
- Leahy, W., & Sweller, J. (2005). Interactions among the imagination, expertise reversal, and element interactivity effects. *Journal of Experimental Psychology: Applied*, 11, 266-276.
- Leahy, W., & Sweller, J. (2008). The imagination effect increases with an increased intrinsic cognitive load. *Applied Cognitive Psychology*, 22, 273-283.
- Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior*, 25, 284-289.
- Liu, T. C., Lin, Y. C., Tsai, M. J., & Paas, F. (2012). Split-attention and redundancy effects on mobile learning in physical environments. *Computers & Education*, 58, 172-180.
- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93, 187-198.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90, 312-320.

- McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2015). The MOOC model for digital practice (2010). URL: http://www.elearnspace.org/Articles/MOOC_Final.pdf (Stand: 01.06. 2012).
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 343-352.
- Moe, R. (2015). *MOOCs as a Canary: A Critical Look at the Rise of EdTech*. Paper presented at the E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (pp. 1037-1042).
- Nkuyubwatsi, B. (2013). *Evaluation of Massive Open Online Courses (MOOCs) from the learner's perspective*. Paper presented at the European Conference on e-Learning (p. 340).
- Ossiannilsson, E., Altinay, F., & Altinay, Z. (2015). Analysis of MOOCs practices from the perspective of learner experiences and quality culture. *Educational Media International*, 52, 1-12.
- 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.
- Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, 17, 398-422.
- Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193-198.
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12, 61-86.
- Reisslein, J., Sullivan, H., & Reisslein, M. (2007). Learner achievement and attitudes under different paces of transitioning to independent problem solving. *Journal of Engineering Education*, 96, 45-56.
- Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, 38, 1-37.
- Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *The Journal of Experimental Education*, 70, 293-315.

- Rourke, A., & Sweller, J. (2009). The worked-example effect using ill-defined problems: Learning to recognise designers' styles. *Learning and Instruction, 19*, 185-199.
- Siemens, G. (2013). Massive open online courses: Innovation in education. *Open Educational Resources: Innovation, Research and Practice, 5*.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*, 123-138.
- Sweller, J. (2012). Human cognitive architecture: Why some instructional procedures work and others do not. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook (Vol. 1)*. Washington, DC: American Psychological Association.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. NY: Springer.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185-233.
- Sweller, J., & Cooper, G. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*, 59-89.
- Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology, 4*, 434-458.
- Sweller, J., Van Merriënboer, J. J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.
- Sweller, J., Kirschner, P., & Clark, R. E. (2007). Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist, 42*, 115-121.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3*, 257-287.
- Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review, 26*, 265-283.
- Van Merriënboer, J. J. (1990). Strategies for programming instruction in high school: Program completion vs. program generation. *Journal of Educational Computing Research, 6*, 265-285.

- Van Merriënboer, J. J., & De Croock, M. B. (1992). Strategies for computer-based programming instruction: Program completion vs. program generation. *Journal of Educational Computing Research*, 8, 365-394.
- Van Merriënboer, J. J., Kirschner, P., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, 38, 5-13.
- Van Merriënboer, J. J., Schuurman, J., De Croock, M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction*, 12, 11-37.
- Wang, Y., & Baker, R. (2015). Content or platform: Why do students complete MOOCs? *Journal of Online Learning and Teaching*, 11, 17-30.
- Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, 22, 449-457.
- Zheng, S., Rosson, M. B., Shih, P. C., & Carroll, J. M. (2015). *Understanding student motivation, behaviors and perceptions in MOOCs*. Paper presented at the Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 1882-1895). ACM.