Learning and Individual Differences 110 (2024) 102423

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Cognitive load theory and individual differences

John Sweller

School of Education, University of New South Wales, Sydney, NSW 2052, Australia

ARTICLE INFO

Keywords:
Cognitive load theory
Element interactivity
Individual differences through learning

ABSTRACT

Cognitive load theory is an instructional theory based on the following assumptions. (1) Information can be divided into biologically primary information that we have evolved to process and acquire unconsciously, and secondary information that requires conscious effort. (2) Secondary information is either acquired during problem solving or from other people. It is processed by a limited capacity, limited duration working memory before being stored for subsequent use in long-term memory. Once stored, information may be transferred back to working memory to govern action relevant to the extant environment. Working memory has no known capacity or duration limits when dealing with information transferred back from long-term memory. (3) The major consequence that flows from instruction is the accumulation of information in long-term memory. (4) That differential knowledge results in individual differences in expertise and provides a major source, possibly the only source, of modifiable cognitive differences between learners. (5) Learners with differential knowledge held in long-term memory require different instructional procedures leading to the expertise reversal effects of cognitive load theory.

Educational relevance: Cognitive load theory is an instructional theory based on our knowledge of evolutionary psychology leading to human cognitive architecture. That architecture specifies individual differences due to either biological or environmental factors. Information held in long-term memory provides the major source of environmentally mediated individual differences. The theory uses randomised, controlled trials to determine instructional effectiveness.

Cognitive load theory (Sweller et al., 2011; Sweller et al., 2019) uses our knowledge of human cognitive architecture to design instructional procedures. Common instructional procedures are analysed from the perspective of that knowledge. If a procedure appears not to conform well with what is known of human cognitive architecture compared to possible alternatives, a randomised, controlled trial can be run to compare the instructional effectiveness of the original procedure to the suggested alternative or alternatives. When a newly designed instructional procedure proves to be superior to a conventional procedure, a new cognitive load theory effect is established.

The generation of new cognitive load theory effects has resulted in the theory being under constant development for over 40 years (Sweller, In press). The purpose of this development was to provide a vehicle for instructional design rather than to consider issues associated with individual differences. Nevertheless, in generating instructional effects, factors associated with individual differences have been clarified by the theory. The intent of this paper is to indicate and discuss those individual differences within a cognitive load theory framework.

There are many cognitive load theory effects, but all effects have a single goal – optimising working memory load to facilitate the accumulation of knowledge held in long-term memory. That accumulation of knowledge leads to individual differences and is the major factor leading to individual differences that is considered by cognitive load theory. While there are many cognitive factors leading to individual differences, such as working memory capacity, visuospatial processing, or gender (e.g., Bevilacqua, 2017; Castro-Alonso et al., 2019), knowledge held in long-term memory is arguably the only one that can be systematically and readily varied. For this reason, cognitive load theory is largely concerned with procedures that can facilitate knowledge accumulation.

Individual differences in knowledge held in long-term memory can considerably alter how people process information and those differences have instructional implications that provide the basic rationale for cognitive load theory. The remainder of this paper, beginning with human cognitive architecture, explains how and why.

1. Human cognitive architecture

There are two general aspects of human cognitive architecture that

E-mail address: j.sweller@unsw.edu.au.

https://doi.org/10.1016/j.lindif.2024.102423
Received 20 December 2023; Received in revised form 8 February 2024; Accepted 17 February 2024
Available online 21 February 2024

are relevant to instructional design: (a) Categories of knowledge based on evolutionary psychology and (b) the organisation of cognitive structures and functions when dealing with the category of knowledge relevant to instruction.

1.1. Using evolutionary psychology to categorise knowledge

There are two categories of knowledge that are relevant to instructional design, biologically primary and secondary knowledge (Geary, 2002, 2005, 2008, 2012; Geary & Berch, 2016; Sweller, 2022). Biologically primary knowledge consists of information that we have evolved to acquire easily and automatically. It is acquired unconsciously without instruction from others. All that is required is membership of a functioning society. Learning to listen to and speak a native language provides one obvious example but there are many others such as learning to solve problems by means-ends analysis (Newell & Simon, 1972) which involves noting one's current problem state, noting the goal state, extracting differences between the two states, and finding a problem-solving operator that can reduce those differences. We all use this strategy without conscious awareness and without it being explicitly taught. We automatically acquire the strategy during normal activities in our environment.

While most individuals acquire biologically primary information automatically without tuition or effort, the speed and extent to which such information is acquired is likely to lead to individual differences. For example, some children will learn to listen to and speak their native language much more rapidly and to a greater extent than other children. Some of the early conceptions of intelligence as an immutable factor for any individual are possibly based on what now is referred to as biologically primary knowledge. Nevertheless, it needs to be noted that since biologically primary knowledge is dependent on the environment in which a person develops, especially the social environment, the acquisition of biologically primary knowledge is unlikely to be immutable. A child who develops in a richer social context is likely to acquire superior native language skills than a child who develops in an impoverished social context, despite both children acquiring their skills automatically and without explicit instruction.

Biologically secondary knowledge consists of information needed by members of a society that those members have not specifically evolved to acquire but nevertheless are able to acquire. Unlike primary knowledge, it usually requires conscious effort and needs to be explicitly taught. Virtually everything taught in education and training institutions consists of biologically secondary knowledge. Such institutions were established precisely because without them, biologically secondary knowledge is unlikely to be acquired. Reading and writing provide an obvious historical example. Written language was invented thousands of years ago but until the advent of mass education relatively recently, the vast majority of most populations, despite having automatically acquired the biologically primary skills needed to listen and fluently talk in their native language, did not learn to read and write. Most people do not learn to read and write without adequate tuition. Because we have not specifically evolved to acquire it, biologically secondary knowledge acquired during education belongs to a different category of information than biologically primary knowledge that we have specifically evolved to acquire. That difference is critical to how we process information belonging to the two different categories and critical to instructional procedures and individual differences.

1.2. Cognitive architecture associated with biologically secondary knowledge

While we have not evolved to acquire and use specific examples of biologically secondary information, we can describe a clear, general system that humans have evolved to acquire, process, store, and use this important information. That system recapitulates the one used by evolution by natural selection (Sweller, 2022; Sweller & Sweller, 2006). It is a natural information processing system.

1.2.1. Acquiring biologically secondary information

There are two ways humans can acquire biologically secondary information: during problem solving or from other people. Initially, all biologically secondary information must be discovered or constructed by someone, commonly using a problem-solving strategy such as means-ends analysis (Newell & Simon, 1972) as described above. Ultimately, all problem-solving strategies dealing with novel information must rely on a random-generate-and-test process at some point. In the absence of appropriate knowledge held either by the problem-solver or by someone else to whom the problem-solver has access, the problem-solver has no options available but to randomly generate a move and determine whether it is effective in providing useful information. If problem solvers are in a position where they have no knowledge concerning the advisability of determining which of multiple possible moves should be made at a particular point, they have no choice but to randomly choose one move and consider the consequences of that move. As will be discussed below, obtaining information in this manner can be complex, difficult, and slow, but is unavoidable. Our general skill in obtaining information under these circumstances is biologically primary but can include many biologically secondary sub-skills such as knowing how to read or use a library or computer.

Rather than using problem solving, whenever available, we tend to obtain information from others because it is a far more efficient way of acquiring biologically secondary information. Arguably, humans are the dominant mammalian species on earth at least in part because of our unique ability to rapidly share very large amounts of information between us. It may be difficult for us to acquire biologically secondary information by problem solving, but once acquired, it can be rapidly and relatively easily shared with others. The immensely complex skill set that enables us to efficiently present and obtain information from others is biologically primary and so does not need to be taught. Educational institutions were developed based on the assumption that students already have or will automatically obtain that skill set.

1.2.2. Processing biologically secondary information: working memory

Once acquired through the senses, biologically secondary information must be processed using a limited capacity (Cowan, 2001; Miller, 1956) limited duration (Peterson & Peterson, 1959) working memory. With respect to capacity, we can remember no more than about 7 elements of novel information in working memory and process no more than about 3–4 elements where processing means organising, combining, classifying, or handling the information in some fashion. These limits apply only to novel information. Familiar information has different characteristics that will be discussed below. With respect to duration, we can hold novel information in working memory for only a few seconds. After about 20 s, most information is lost. We can retain information in working memory for longer periods by repetition.

It also might be noted that working memory is the source of human consciousness. We only are aware of information being processed in working memory. We are unaware of the far greater amount of information held in long-term memory, discussed next.

1.2.3. Storing biologically secondary information: long-term memory

Once processed, secondary information can be stored for later use in long-term memory. While working memory when dealing with novel information is limited in capacity and duration, there are no known limits of the long-term memory store. Because at any given moment we only are aware of information held in working memory, we can have difficulty appreciating the enormous size of that memory store. To provide an example, Simon and Gilmartin (1973) estimated that chess grand masters have memorised over one hundred thousand board configurations, indicating the size of long-term memory. Anyone who has skill in any complex domain derives that skill from the extensive information that they have stored in long-term memory. That memory store is the source of expertise and as discussed in more detail below, possibly the only source of malleable and alterable cognitive differences between individuals. Those differences between individuals in the content of their long-term memory are generated by differences in what they have learned.

1.2.4. Using biologically secondary information

The ultimate purpose of this human cognitive architecture is to permit us to function appropriately in the various contexts with which we are faced. Information stored in long-term memory can be transferred back to working memory to govern activity. Environmental signals are used to trigger appropriate long-term memory elements to be used to determine that activity. As an example, anyone reading this text is faced with a large number of complex squiggles. Of course, having spent many years learning to read English, we immediately use the signals provided by the text to connect those squiggles to the large amount of information held in long-term memory associated with reading. The text signals to us which information should be transferred to working memory to generate suitable activity, in this case, interpreting the squiggles in English language terms.

While, as indicated above, working memory is very limited in duration and capacity when dealing with novel information from the environment, the characteristics of working memory are dramatically altered when dealing with familiar information transferred from long-term memory. Learning complex activities such as reading, mathematics, playing chess etc., occurs within the confines of our limited working memory when dealing with novel information from the environment. In contrast, when the same information is transferred to working memory from long-term memory rather than from the environment, it assumes the characteristics of long-term memory. There are no known limits to the amount of information that working memory can process when that information comes from long-term memory. In this manner, knowledge transforms us leading to individual differences.

2. Human cognitive architecture and individual differences

Evidence for the potential size of long-term memory and its consequences for individual differences was provided by De Groot (1965) who found that chess masters, shown a chess board configuration taken from a real game for a few seconds, could replicate it with very high accuracy compared to much lower accuracy for week-end players. Chase and Simon (1973) also obtained that result but in addition found the difference disappeared using random board configurations indicating that the difference between experts and novices was due to information stored in long-term memory and not due to differences in working memory.

A good chess player has memorised a large number of configurations along with the most appropriate move for each configuration. Similar results have been obtained in a variety of fields (Chiesi et al., 1979; Egan & Schwartz, 1979; Jeffries et al., 1981; Sweller & Cooper, 1985). Expertise in any field derives from a huge amount of information relevant to that field that is stored in long-term memory. When faced with a problem, an expert knows exactly what to do because he or she has encountered that problem, probably repeatedly, and can use a memorised solution. In contrast, a novice has to use a problem-solving strategy such as means-ends analysis. Instead of using a potentially large amount of information stored in long-term memory, because of the absence of that information, a novice has to use his or her limited capacity, limited duration working memory. It is this difference in stored information held in long-term memory, associated with the characteristics of human cognitive architecture, that lead to the differences between novices and experts.

2.1. Instructional implications of cognitive load theory: some cognitive load theory effects

Cognitive load theory has used this cognitive architecture to generate a series of instructional effects. A cognitive load theory effect is demonstrated when instructional information structured according to cognitive load theory principles results in superior learning compared to a more traditional instructional procedure. All cognitive load theory effects are based on multiple, replicated, randomised, controlled trials. All are intended to facilitate the development of biologically secondary information leading to expertise in domain-specific areas (Tricot & Sweller, 2014).

One effect, the element interactivity effect (Chen et al., 2023; Sweller, 1994, 2010), is central to cognitive load theory and the cognitive load theory effects as well as being central to how the theory deals with individual differences. Some information consists of multiple, interacting elements and because they interact, they must be processed simultaneously. Element interactivity has consequences for the functioning of the human cognitive system through its effects on both working memory and long-term memory.

2.2. The effects of element interactivity on working memory and long-term memory: intrinsic cognitive load

Processing multiple elements simultaneously imposes a heavy working memory load. That working memory load can lead to either *intrinsic* or *extraneous* cognitive load. Intrinsic cognitive load is caused by the intrinsic nature of the subject matter the learner is dealing with. For example, assume a student is learning the translation of individual words from one natural language to another. He or she can learn that the translation of *the cat* in English is *le chat* in French and similarly *the dog* translates to *le chien* in French. Both translations can be learned independently of each other. Element interactivity is low. Much biologically secondary information such as learning the symbols of the Chemistry Periodic Table is similarly low. Low element interactivity information may be very difficult to learn if there are many elements, but it does not impose a high working memory load. Learning that *Fe* is the symbol for *iron* independently of learning that *Cu* is the symbol for copper is relatively easy with each imposing a low working memory load. Learning the vocabulary of a foreign language can take many years despite the fact that each vocabulary item imposes a light cognitive load. The difficulty stems from the enormous number of items that must be learned, not because they are high in element interactivity that imposes a heavy working memory load.

In contrast, other information that must be processed and learned is high in element interactivity and is difficult for that reason. Learning French grammar or learning to balance chemical equations requires the consideration of multiple elements simultaneously. Much of mathematics is high in element interactivity. For example, learning to solve a problem such as $(a + b)/c = d$, *solve for a,* is relatively high in element interactivity because no move can be made without considering every element of the equation. As a consequence of high element interactivity, learning can be difficult, not because there are many elements but rather, because the limited number of elements interact. Having to simultaneously consider multiple elements, including how they interact, imposes a high intrinsic cognitive load. It is called *intrinsic* because it is an intrinsic, unavoidable aspect of the information that must be processed.

Based on the above cognitive architecture, when high element interactivity information is processed in working memory and sufficiently practiced, it will eventually be stored for subsequent use in long-term memory, but it will not be stored in long-term memory in the same form as it was processed in working memory. With sufficient practice, the multiple elements that constitute high element interactivity information can be "chunked" (Miller, 1956) into a single element. In this way, knowledge held in long-term memory influences element interactivity by transforming multiple elements into a single element.

When that single element is transferred back to working memory to generate activity relevant to the environment, it no longer imposes a heavy intrinsic cognitive load. That single element can be processed in working memory much more readily than its constituent multiple elements which explains the importance of information held in long-term memory. As indicated previously, it explains why chess masters recognise an entire board configuration and its relevant moves rather than just many individual chess pieces, or why anyone competent in elementary algebra will immediately recognise the above algebra problem and know immediately that the first move should be to multiply out the denominator on the left-hand side of the equation followed by subtraction of the addend. Critically, it is that information held in long-term memory that leads to a major source of individual differences.

An expert can perform at a vastly higher level than a novice because the expert is not faced with information that is high in element interactivity. In effect, experts deal with information that for them, is low in element interactivity and so has a low intrinsic cognitive load. The same information is very high in element interactivity for novices and so has a high intrinsic cognitive load. It follows from these processes that there are two ways in which intrinsic cognitive load can be altered. It was indicated above that some information imposes a low intrinsic cognitive load because the elements do not interact. Accordingly, intrinsic cognitive load can be altered by the obvious procedure of altering what needs to be learned. Of course, altering what needs to be learned is frequently not possible. Much biologically secondary information is essential in many circumstances which is why it is mandated in educational curricula. The alternative way of reducing intrinsic cognitive load is to change the structure of the information held in long-term memory. Learning achieves this aim and is a major source of individual differences.

2.3. *The effects of element interactivity on working memory and long-term memory: extraneous cognitive load*

Levels of element interactivity are not only caused by the intrinsic nature of the information. They also are affected by the way the information is presented and received, leading to extraneous cognitive load. Cognitive load theory is primarily concerned with extraneous cognitive load caused by instructional design. There are many cognitive load theory effects which are summarised elsewhere (Sweller et al., 2011; Sweller et al., 2019). Based on the above cognitive architecture, all are intended to facilitate the transfer of information from working memory to long-term memory. This transfer, as indicated above, can result in large differences in individual performance on specific tasks. Element interactivity is central to all extraneous cognitive load effects. The worked example effect, which arguably is the most important extraneous cognitive load theory effect, will be used to explain how extraneous cognitive load effects are generated by variations in element interactivity.

2.3.1. *The worked example effect*

This effect occurs when learners presented with worked examples of problem solutions to study learn more than learners presented the same problems to solve (Chen et al., 2015, 2016a, 2016b; Cooper & Sweller, 1987; Kyun et al., 2013; Paas & van Merrienboer, 1994; Renkl, 2014a, 2014b; Sweller & Cooper, 1985). The theoretical explanation of the effect provided by cognitive load theory is that studying a worked example reduces element interactivity compared to solving a problem (Sweller, 2010). A learner attempting to solve a problem using means-ends analysis must consider for each move they make, their current problem state, the goal state, multiple differences between the two states, and an indefinite number of possible moves to reduce those differences. In contrast, a good worked example clearly indicates each problem state and the move required to reach the next state which is also depicted. A worked example considerably reduces the number of interacting elements that require consideration. Based on the worked example effect, expertise can be acquired much more easily and rapidly by studying worked examples rather than solving problems. A similar reduction in element interactivity explains the advantage of other extraneous cognitive load effects (Sweller, 2010), including the redundancy effect, discussed next.

2.3.2. *The redundancy effect*

The redundancy effect occurs when learners are presented with redundant information that is not required for learning (Chandler & Sweller, 1991). Experimentally, the effect is demonstrated when learners presented with redundant information learn less than learners presented the same information with redundant information excluded. Redundant information may be the same information presented in more than one modality or form such as identical written and spoken text, diagrams and text providing the same information, or completely unrelated information added to essential information such as cartoons added solely for their humour without any useful, informational content. Adding redundant information increases the number of elements with which learners must deal resulting in an unnecessary increase in working memory load.

The common technique of providing the same information in spoken and written form provides an example of the redundancy effect. A PowerPoint presentation in which large slabs of orally presented information also are presented in written form on a screen does not facilitate learning by an audience. Instead, by increasing extraneous cognitive load, it can interfere with learning (Kalyuga et al., 2004).

2.3.3. *The expertise reversal effect*

Providing students with information that they do not need because they have already acquired the information provides another example of the redundancy effect. To provide an example, while worked examples facilitate learning compared to solving the equivalent problems, as expertise increases, studying worked examples may no longer be effective. With increased expertise, the effect first decreases in size, then disappears, and may eventually reverse with problem solving superior to studying worked examples (Kalyuga et al., 2001). The cause is the decrease in element interactivity associated with increases in expertise. Worked examples are highly effective when learners are dealing with novel, high element interactivity information because, as indicated above, a worked example reduces element interactivity. As element interactivity decreases with increasing expertise, studying a worked example becomes redundant because at some point, all that is needed is practice at problem solving rather than instruction in how to solve a particular class of problems. The same instructional procedure, studying worked examples, which is beneficial for novices because it reduces element interactivity, becomes redundant when levels of expertise have increased to the point where all that is needed is practice at problem solving. The result is the expertise reversal effect (Kalyuga et al., 2003) with no advantage for worked examples or even an advantage of problem solving over studying worked examples.

All cognitive load theory effects are due to variations in element interactivity. The expertise reversal effect is assumed to occur because element interactivity is reduced as expertise increases. Based on the cognitive architecture that underlies cognitive load theory, information stored in long-term memory reduces element interactivity. In turn, element interactivity explains and relates the worked example effect, the redundancy effect and the expertise reversal effect. For novice learners in a particular area, studying worked examples reduces extraneous cognitive load by reducing element interactivity and so facilitating learning. With increasing expertise, studying worked examples becomes redundant and so the advantage of studying worked examples disappears due to the redundancy effect. Practice at solving problems becomes more important than studying worked examples, leading to the expertise reversal effect. Each of these effects facilitates the acquisition of information held in long-term memory by reducing the load on working memory. That increase in knowledge leads to expertise and a major source of individual differences.

3. Conclusions

The cognitive architecture used by cognitive load theory, along with the cognitive load theory effects, provide information on how learning creates individual differences. We all know that knowledge is transformational. While there are many factors that lead to individual differences, the acquisition of knowledge is a major, possibly the only cognitive factor that is readily malleable and alterable. Differences between individuals in their knowledge have profound consequences for their ability to navigate their way in specific areas. A person with knowledge of a particular historical epoch or the science associated with energy production is in a vastly better position to think critically about the relevant subject areas than a person who does not have that requisite knowledge held in long-term memory. Cognitive load theory, through the cognitive load theory effects, is concerned with techniques that facilitate the acquisition of that knowledge.

CRediT authorship contribution statement

John Sweller: Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

No competing interest.

References

Bevilacqua, A. (2017). Commentary: Should gender differences be included in the evolutionary upgrade to cognitive load theory? *Educational Psychology Review, 29*, 189–194.

Castro-Alonso, J. C., Wong, M., Adesope, O. O., Ayres, P., & Paas, F. (2019). Gender imbalance in instructional dynamic versus static visualizations: A meta-analysis. *Educational Psychology Review, 31*, 361–387.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.

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. (2016a). 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. (2016b). When instructional guidance is needed. *Educational and Developmental Psychologist, 33*, 149–162.

Chen, O., Paas, F., & Sweller, J. (2023). A cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review, 35*.

Chiesi, H., Spilich, G., & Voss, J. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behaviour, 18*, 257–273.

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). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–114.

De Groot, A. (1965). *Thought and choice in chess*. Mouton (Original work published 1946).

Egan, D. E., & Schwartz, B. J. (1979). Chunking in recall of symbolic drawings. *Memory & Cognition, 7*, 149–158.

Geary, D. (2002). Principles of evolutionary educational psychology. *Learning and Individual Differences, 12*, 317–345.

Geary, D. (2005). *The origin of mind: Evolution of brain, cognition, and general intelligence*. American Psychological Association.

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.)*, Vol. 1. APA Educational Psychology handbook* (pp. 597–621). 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). Springer.

Jeffries, R., Turner, A., Polson, P., & Atwood, M. (1981). Processes involved in designing software. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 255–283). Erlbaum.

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

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46*, 567–581.

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.

Kyun, S., Kalyuga, S., & Sweller, J. (2013). The effect of worked examples when learning to write essays in English literature. *Journal of Experimental Education, 81*, 385–408.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice Hall.

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

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

Renkl, A. (2014a). The worked-out examples principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 391–412). Cambridge University Press.

Renkl, A. (2014b). Toward an instructionally oriented theory of example-based learning. *Cognitive Science, 38*, 1–37.

Simon, H., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology, 5*, 29–46.

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. (2022). The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. *Educational Psychology Review, 34*, 2229–2241.

Sweller, J. (2024). The development of cognitive load theory: Replication crises and the incorporation of other theories can lead to theory expansion. *Educational Psychology Review* (in press.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

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

Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology, 4*, 434–458.

Sweller, J., van Merriënboer, J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292.

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review, 26*, 265–283.