Comments on Sweller, van Merriënboer & Paas' 2019 Cognitive Load update.

Anne Watson March 2019

(see https://link.springer.com/article/10.1007/s10648-019-09465-5)

CL: cognitive load

WM: working memory

LTM: long term memory

BPK: biologically primary knowledge

BSK: biologically secondary knowledge

The leading academics of CL theory summarise where it has got to and where it might go next. Following my paper that critiques the ways in which CL theory is applied in educational advice coming from the DfE, I am now commenting on this summary. The starting place is repeated from my earlier posting.

1. I have no argument with the authors of CL theory and its development to describe how the mind learns given a particular model of memory and its possible limitations when performing particular kinds of task.

2. My argument is with the direct application of this theory to the classroom learning of mathematics. Much of the application ignores
    a. that mathematics contains its own symbols, chunked concepts, language, notations and diagrams that enable people to manage cognitive load. This means that complex tasks and multi-stage reasoning can be stored in paper, screen and artefacts rather than solely in the mind.
    b. that 12 years of schooling gradually transforms learners from novices to experts in particular conceptual strands and also in appropriate behaviour for doing mathematics
    c. that 'new' mathematical ideas are rarely totally new but more often are developments, adaptations and extensions of familiar ideas, drawing on structures in long-term memory (schema) and supportive artefacts

3. The tasks and tests used to generate and validate elements of CL theory need to be examined before assuming their findings are relevant for continuous mathematical learning over 12 years of school. Too often the tasks are not part of a coherent subject discipline, or are extremely unlike tasks undertaken in school (e.g. using dual-task theory that arises from setting unconnected tasks to justify arguments to reduce CL). Interventions are inauthentic and do not take normal pedagogical features into account (e.g. post-task discussion of key points; use of tools when multi-step calculation would slow down problem-solving; developing confidence in working in unfamiliar and complex situations). Too often tasks and tests are about getting answers by performing skills. Generally, the results are due to randomised control

trials and not to continuous subject education over time. Delayed post-tests might be carried out with students one or two years on who are not necessarily following related subject pathways.

I had two laugh-out-loud moment when reading this paper. The first is to do with Figure 1, which is a diagram of 4C/ID (a four-component instructional design model). The figure introduces 8 kinds of symbol that have to be understood and coordinated, along with four wordy text boxes, and then, because it is hard to decipher, eighteen lines of text to explain what the diagram means. This breaks many of the 'rules' of CL theory applied to giving new information.

My second laugh-out-loud moment came when reading that after sustained cognitive effort there is some depletion of working memory and you have to go off and do something else to let it recover. I read this after having to stop reading and going off to make a cup of tea!!! There are other moments like this second one, such as recognising that emotion pays a part, imagery may be helpful, using multiple modes is worth thinking about, a spiral curriculum might help learning, and critical use of variation is a useful. I am delighted that CL theory is catching up with what it is really like to do mathematics and what good teachers already know from their professional experience. So research subjects in intervention studies and lab experiments do the same kinds of things with their minds as students in school classrooms – amazing!

Now to pick my way through the paper and point out what seems worth saying and what does not. Warning: I have not read all the studies that contribute to this paper, but have looked at enough to know that my third point above still applies.

Firstly, CL theory is about processing **new** information and constructing knowledge in long-term memory. It is not about reprocessing familiar ideas in new ways, or restructuring and extending existing schema, and so far does not seem to say anything much about the formation of chunks of remembered experience – structures and their associated processes and meanings – in the LTM. Expert chess players needing a repertoire of positions and moves is mentioned as a seminal study, but the fact that they also play full games while this is being accumulated, and those games may or may not present opportunities to use the most recently-learnt structure, is not part of that research. However the chess study is widely used to support 'learn the moves then do the same maths' approaches to teaching. CL theory does not appear to test adaptations and extensions of familiar 'information'.

Ages ago the notion of **'germane' cognitive load** was introduced to accompany intrinsic load (roughly the number of elements in the task) and extrinsic load (the working environment) to notice that some load is germane for the processes of learning, so minimising overall load might not be helpful for learning. Now this label has been abandoned because it seems that 'germane' load can be described as the work of shifting attention from extraneous load to the load that is intrinsic to the current situation we might call that 'analysing the problem' – once that work is done germane load no longer adds to total load.

Drawing on memory of past learning is less taxing on working memory than having to juggle new ideas – indeed no one yet knows if there are limitations in WM where long-term memory is supplying much of the material. This means, in mathematics, that the long term build up of concepts over time does not need to be restricted to the results about WM and CL to anything like the extent that is needed for new and unconnected ideas. Elements in a task interact and, for any task, this may mean different things for different people depending on their LTM. These observations mean that CL is not a feature of the task alone, but is a feature of the task and its place in the curriculum and the long term memories of the individuals doing the task. I think teachers already know this, and also know that students may need to recall useful schema when meeting an extension of an earlier concept. Indeed many use this observation to prepare students for forthcoming work by revisiting older ideas so they are more easily available from LTM. The desirable learning might then be the interaction of several previously known elements which previously have been unconnected, but learning the interaction is very difficult if the elements are not already known. This is what doing maths is like, so it is indeed remarkable that CL academics have caught up and verified with research that this is the case.

Variation gets a look-in: while variation allows similarities and differences to be discerned, this increases intrinsic CL because there is more to look at and process, but it also increases transfer of learning to different situations. In mathematics, we would say that this is because identifying sameness and difference can lead to a sense of underlying structure, and this is what makes transfer easier, i.e. recognising familiar structure rather than a familiar example. Learning about mathematical structure rather than skills and factual knowledge would always be my aim when thinking about teaching, so this is fine with me (see also Krutetskii's seminal study of the knowledge that gifted Russian mathematics students have). However, in this paper the effect is summarised by saying 'vary everything' as a contrast to presenting examples with surface similarity, but varying everything creates high CL if the object of learning is a general structure and teachers are aware that, as in science, controlling variation draw attention more effectively to aspects of a structure/concept.

Another useful result is that methods of teaching novices are not appropriate for teaching people as their expertise grows.

This, and other CL research, indicates that the simple small step idea currently going around is not justified by a need to reduce CL for much of school mathematics – CL is already reduced because of the build up of concepts over time, and the fact that elements of knowledge are not isolated, and that students are becoming experts.

A very useful distinction is offered between biologically primary knowledge and biologically secondary knowledge (as Dewey recognised in the late 19th century). BPK refers to those functions that have evolved and are part of natural human activity and, in this body of work, that includes transferring previous information to new situations, working on unfamiliar problems, planning, regulating thought processes, or ability to construct knowledge, and (they don't mention this in their list) making distinctions and generalising. They say these cannot be taught. BSK are those things that have to be in a formal curriculum and require conscious effort to be learnt. They claim that these require explicit instruction but Dave Hewitt has drawn attention to two kinds of BSK: that which has to be told (e.g. names for things; how to ‘read’ algebraic statements; etc.) and that which can be deduced by reasoning (e.g. how to solve equations once you know the symbolism and the meaning of ‘equals’). I am not sure where logical reasoning fits in the CL literature and of course no one would say that students have to reason everything out for themselves, however it does seem to me that applying BPK to mathematical objects can do much of the work for understanding mathematical ideas. They dig up the example I refer to in my earlier post about knowing what to do with a particular algebraic equation, assuming that ‘multiplying both sides by a’ has to be explicitly taught. No it doesn’t, but you might decide to explicitly teach it either by showing how this retains the equality, or by asking students what they can do that retains the equality and gets them nearer the required result, or by recognising the rational structure of the expressions in the equation and knowing transformations of these.

Because some problem solution moves are generic and come free with the way our brains work they can be useful in learning BSK, but ‘pointing out to students that a generic cognitive skill should be used on a particular class of specific problems can be instructionally effective’. I find it hard to believe that teachers of mathematics need decades of cognitive science to learn that. It also turns out that it is a biological given that we might randomly try out methods, chuck the ineffective ones away and keep the effective ones without any explicit instruction. Who knew?

After a review of this kind of research finding they offer a 4-component model for instruction: task, information, procedural information and practice of parts of a task. They say that learning tasks are ‘preferably based on real life tasks’. This loses me except to say that it challenges whether the model is appropriate for mathematics, for which we sometimes, but not always, might find a real-life task that can generate some abstract learning but more often than not this is not appropriate for abstract mathematical concepts. It also does not appear to relate to what is said elsewhere – real-life contexts usually come with heavy loads of extrinsic information so this is all a bit puzzling.

The model is equivocal about inserting practice while doing complex tasks, saying it is not helpful during complex learning but prior automation of recurrent skills does free up cognitive resources. Instead they advocate giving just-in-time information in complex tasks if it is needed. It was hard to make sense of this (I need to go away and make another cup of tea to replenish my WM) but I have scrawled ‘Mike Ollerton might like this’ in the margin.

I see very little in this paper that supports step-by-step with low CL approaches to mathematics pedagogy throughout the subject and throughout school.