Cognitive Load Theory in action

In this article **Dr Ollie Lovell** provides a tantalising sample of some of the ideas that he has included in his recently published book, *Cognitive load theory in action*. The concepts are relevant to teachers of students of all learning abilities and all ages – very practical principles of learning that can help teachers understand how to avoid leaving students floundering and puzzled as to what to do next.

Cognitive Load Theory (CLT) is a series of instructional recommendations built upon knowledge of how humans learn. The recommendations provide the basis for teaching strategies that apply to all ages and all disciplines.

I first encountered CLT through a now famous tweet by the renowned educationalist Dylan William in early 2017 when he wrote, *I have come to the conclusion that Cognitive Load Theory is the single most important thing for teachers to know*. And that tweet sent me down a rabbit hole of researching CLT out of which I still haven't emerged (Lovell, 2020).

There are two key reasons why I am particularly passionate about Cognitive Load Theory. The first is that it is based upon fundamental immutable principles around how people learn. The second is the solid research basis on which it is built, that stretches back four decades. To me, this is a powerful combination of foundational theory and experimental validation, that can give teachers confidence that it holds real promise for the classroom.

I am a classroom teacher, and the head of the mathematics department in a secondary school. Over the past year my explorations into CLT have been focussed around trying to convert its findings into practical and actionable advice for teachers. This work has been conducted in conjunction with the originator of CLT, John Sweller, which has proved a very fruitful partnership. To my mind, this is the kind of collaboration that education would benefit from more of, a classroom-based teacher translating academic research into practical terms for a teacher audience, and doing so under the oversight of academics, who are checking the interpretations for consistency with the underlying theories and research at every turn!

In this short article I will share some of the key ideas that have emerged from this project and give examples of some practical classroom strategies that emerge from the key CLT principles.

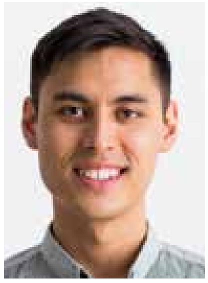
I have tried to make the key foundational principles of CLT accessible and understandable through what I have called 'the ABCDE of CLT'. The acronym stands for:

- **A**rchitecture of human memory,
- **B**iologically primary vs. secondary knowledge,
- **C**ategories of cognitive load,
- **D**omain general vs. domain specific skills, and
- **E**lement interactivity.

Due to space limitations, I will restrict the discussion to only three of these five principles for the purposes of this article – covering only the 'ACE' of CLT, before providing some examples of the practical implications of the theory.

The A of CLT: Architecture of Human Memory

Why is it that students often forget to do simple things, like start a sentence with a capital, or end it with a full stop? This question, and many others, can be easily answered once we have a solid understanding of the architecture of human memory.



We can characterise the resources that all humans think with in a simplified three-part model.

First there is the environment. This an unlimited external store of information. This is where books, notes, presentations like this one, and google exist.

We then have our 'working memory'. Working memory is the site of our consciousness – it's where all of the thinking happens. Working memory is an internal store of information, but it is a *limited* internal store of information.

Finally, we have our long-term memory. Once we know something, this is where that information is stored. And, for all intents and purposes, long-term memory is an unlimited internal store of information. We do not have any evidence to suggest that it can be filled within a human lifetime.

As you can see in Figure 1, attention brings information from the environment into our working memory for processing. Rehearsal within working memory, and making connections to what is already known, is how learning happens. Once information is within long-term memory it can be retrieved, and it is also frequently forgotten.

But the key feature of the model shown in Figure 1 is that whilst the environment and long-term memory are unlimited in their capacity, working memory is not. What this means is that working memory is the bottleneck of thinking. Whenever we teach our students, working memory is one of the crucial constraints that we deal with.

Put simply, this is why students forget to do simple tasks like end their sentences with a full stop. For beginning writers, the act of writing is very cognitively demanding. It utilises the vast

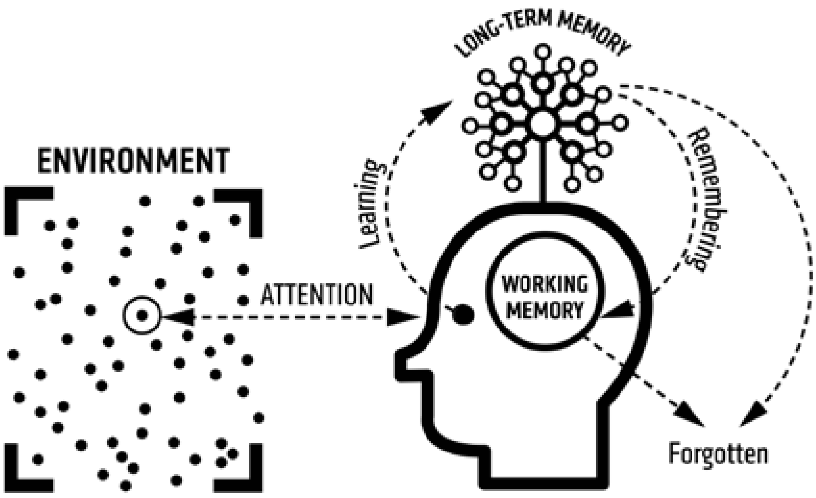


Figure 1. Model of the architecture of human memory

majority of working memory capacity, and this means that there's very little left to remember to do simple tasks.

Acknowledging this fact, Cognitive Load Theory is designed around this fundamental idea: *Working memory is the bottleneck of thinking*.

The C of CLT: Categories of intrinsic and extraneous load

The capacity of working memory as a bottleneck can be understood further by considering the categories of intrinsic and extraneous load, as characterised in Figure 2.

When students juggle information in their working memory, we say that they are encountering 'cognitive load'. But not all cognitive load is created equal. Some of it is desirable, and is required for the student to learn the key ideas. We call this cognitive load that is directly associated with the learning task, 'intrinsic load'.

On the other hand, there are things that teachers do, or environmental factors, that distract students from the core information at hand. This also generates cognitive load, but we call this type of load 'extraneous load', because it is extraneous to the learning task.

Cognitive Load Theory is built around the key idea that *instruction is enhanced by reducing extraneous load and optimising intrinsic load*. Said another way, when we reduce the distractions that students face (reduce extraneous load), we free up working memory capacity, which can then be directed to productive learning (optimise intrinsic load).

The vast majority of the CLT effects and recommendations (including the three introduced later on in this article) are focussed around reducing extraneous load, and this originates in large part from the manner and structure of instructional presentation. A small number of the effects relate to the optimisation of intrinsic load as well.

The E of CLT: Element Interactivity

If we are to reduce extraneous load and optimise intrinsic load, how are we to actually do it? A good first step is to understand the source of cognitive load, which is *element interactivity*.

Put simply, all cognitive load comes from the number of elements that we are asking students to consider at any one time, and the number of interactions between those elements. If we ask students to do a simple task, like learn the location of the country of Poland on a map of Europe, there are not many interacting elements in the task, so it will be low in cognitive load.

If, however, we ask students to find the coordinates (latitude and longitude) of Poland on that same map, we suddenly introduce a whole lot more elements that they must contend with. One of the key insights of CLT is that there are many things that teachers do that unwittingly increase the number of elements that students must contend with, but that are superfluous to learning. These unnecessary elements represent extraneous load, and they are what CLT's instructional recommendations assist us in reducing.

The other key point in relation to element interactivity is that not all 'elements' are created equal. Crucially, new information is very burdensome on working memory, and even a small number of elements of information can overload students. In contrast, familiar information, information that has already been learnt by students and stored in long-term memory, can be drawn upon by working memory and utilised with very little working memory strain. This key insight, that information stored in long-term memory eases the burden on working memory, illustrates the importance of students truly learning information, and undercuts the common, but incorrect, narrative that students no longer have to learn 'facts', because everything is available on google.

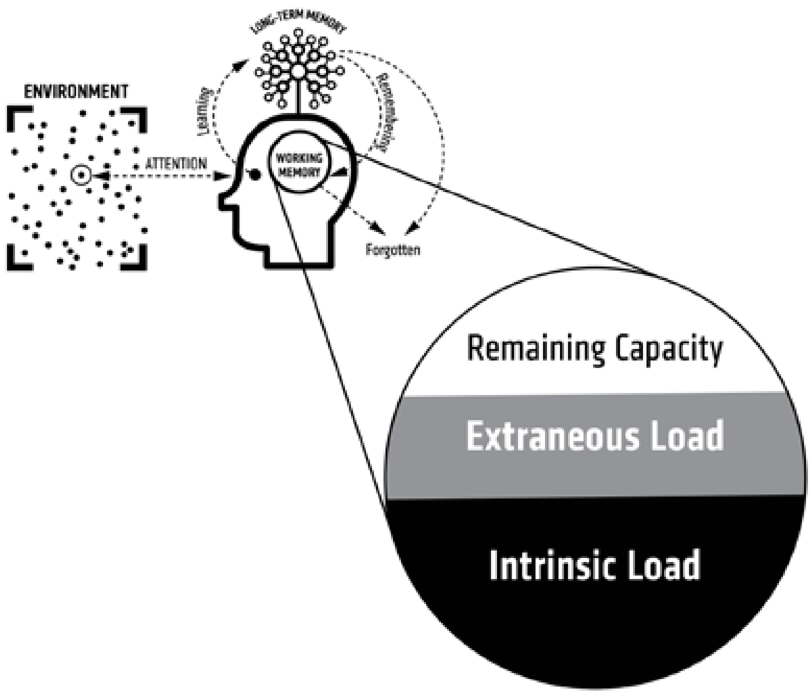


Figure 2. Categories of cognitive load

Instructional recommendations: Practical implications of CLT

Now that we're aware of the 'ACE' of Cognitive Load Theory, we can start to consider some of its instructional recommendations that will allow us to reduce the extraneous load that our students encounter, and improve instruction.

The split-attention effect

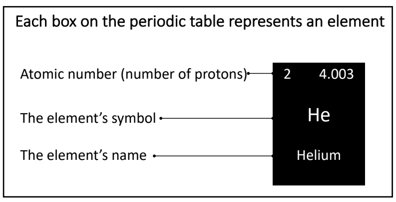
Some time ago I was in a year 7 or 8 science class and the teacher was teaching their students about the periodic table. She had a slide on the board that looked something like this:

> **Learning to read the periodic table**
>
> ❑ Each box on the periodic table represents an element
>
> ❑ The atomic number gives the number of protons in the element
>
> ❑ The one or two letters in the middle of the box gives the element's symbol
>
> ❑ The word at the bottom of each box gives the element's name

Each student also had their own periodic table in hand, and the teacher was guiding them to look at the table in their hands, and relate it to the key information on the slide.

Now, the teacher's instructions were clear, and this task seems quite simple, but as I looked around the room, quite a few of the students seemed puzzled. They were looking up at the slide, then down at the periodic table, and trying to make sense of the two together. What was happening here was that the format of instruction was actually causing the students to split their attention between two places, and mentally integrate the written instructions on the board, with the periodic table in their hands. There were a set of interacting elements that were separated, and the students had to mentally integrate them. This mental integration was causing extraneous cognitive load.

After the class, I spoke to the teacher about the lesson, talked about the idea of split-attention, and we came up with the following, altered slide:



This format reduced the need for students to try to work out what-goes-with-what, and eased the burden on working memory.

Similar alterations have been demonstrated to be effective in randomised controlled trials. For example, Bobis, Sweller and Cooper (1993) found that students were better able to complete a folding task when split attention was reduced by integrating the instructions onto the paper itself. In Figure 3 below, the left-hand side of the page shows a split-attention format, with the printed instructions to be used by the students presented separately from the circle of paper they were meant to be folding. The right-hand side shows an integrated format, with the instructions actually printed on the to-be-folded circle.

Split-attention can also show up in the way that we write or speak. Consider the following two explanations of human cognitive architecture.

Split attention format:

> *There are three key resources we all draw upon in order to think: the environment, working memory, and long-term memory. The first is an unlimited external store of information. The second is a limited internal store of information. And the third is a practically unlimited internal store of information*

Integrated format:

> *There are three key resources we all draw upon in order to think: the environment, working memory, and long-term memory. The environment is an unlimited external store of information. Working memory is a limited internal store of information. And long-term memory is a practically unlimited internal store of information.*

As you can hopefully see, in the first example, readers must hold in working memory what is meant by the 'first', 'second', and 'third' as they read on. In the second example, the three components are repeated at the start of each sentence, reducing the need for readers to keep in mind what is meant by 'first', 'second', and 'third', and therefore easing the burden on working memory.

The split-attention effect can be remembered with the rhyme, *'Information that must be combined should be placed together in space and time!'*

The transient information effect

Another way that students' working memories can be put under unnecessary strain is through transient information. Transient information just means, 'information that disappears'. Whenever we introduce new information to students, then expect them to hold it in working memory while they do something with it, we're exposing them to transient information.

The most common occurrence of this that is seen in classrooms is due to the use of slideshows. Because of the

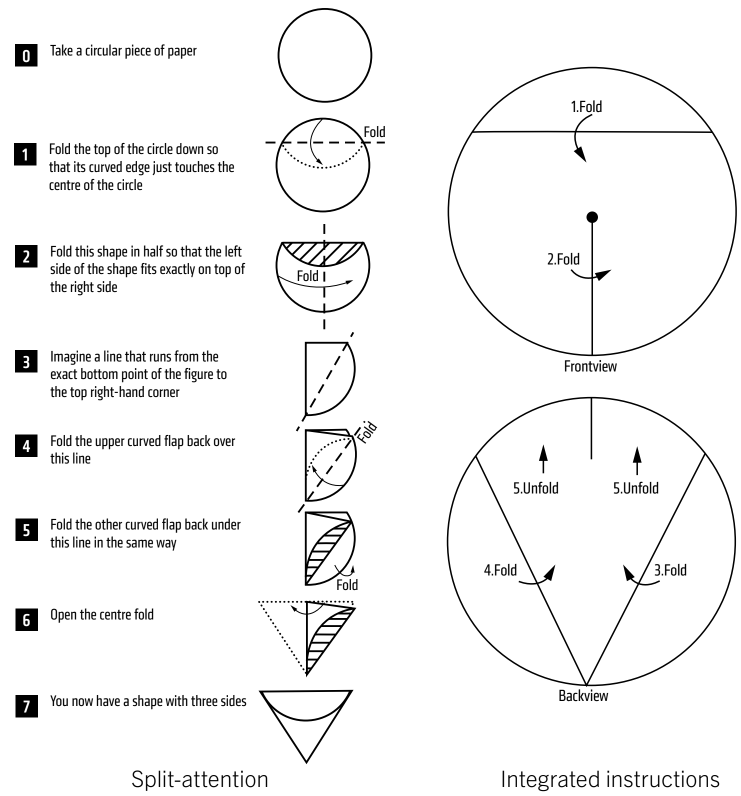


Figure 3: Split-attention versus integrated instructions in a paper-folding task (Bobis et al.1993, p. 16).

nature of the PowerPoint presentation, when the instructor changes slides, often a lot of the information disappears.

For example, imagine that a teacher is teaching their students about punctuation, and provides two slides of instructions, as follows:

> Punctuation: The full stop
>
> Full stop
> Rule: Use full stop at the end of a sentence.
> Example: I rode my bike to the shop.
>
> 2

> Punctuation: The comma
>
> Comma
> Rule: Use an example to separate items in a list.
> Example: At the shop I bought some oat milk, bread, broccoli and vegan sausages.
>
> Rule: Commas sit at natural pause points within a sentence.
> Example: I like cats, especially those with big furry tails.
>
> 3

A following slide then requires students to use this information in an activity:

> Punctuation: You try
>
> **You try**: Complete the following by placing a comma or a full stop at the locations marked (once you're done, change the letter at the start of each new sentence to a capital)
>
> *Dogs are mammals that have been bred to live with humans _ not in the wild _ they have been bred by humans for a long time _ and were the first animals to ever live with humans _ there are many types of dogs _ such as beagle _ retriever _ Jack Russell _ the dingo is also a dog _ but many dingoes have become wild animals again _*
>
> 4

In such an instance, the transience of the information, the fact that it's no longer visible to students when it comes time to complete the activity, means that the students must hold that new information (what full stops and commas are, and how to use them), in working memory whilst completing the task. This is likely to cause a high load on working memory for students unfamiliar with these punctuation marks, and could be aided by simply reducing the transience of information, which could be achieved by simply adding a small 'remember' box:

> Punctuation: You try
>
> **Remember**
> **Full Stop**: use a full stop at the end of a sentence.
> **Comma**: Use a comma to separate items in a list.
> **Comma**: Commas sit at natural pause points within a sentence.
>
> **You try**: Complete the following by placing a comma or a full stop at the locations marked (once you're done, change the letter at the start of each new sentence to a capital)
> *Dogs are mammals that have been bred to live with humans _ not in the wild _ they have been bred by humans for a long time _ and were the first animals to ever live with humans _ there are many types of dogs _ such as beagle _ retriever _ Jack Russell _ the dingo is also a dog _ but many dingoes have become wild animals again _*
>
> 2

Worked examples

One of the most widely known recommendations arising from Cognitive Load Theory is that, for novice learners, worked examples can be an effective instructional strategy. However, it wasn't until I really dove into the research that I realised that what I, and many other teachers bring to mind when we hear the term, 'worked examples', is not at all what the Cognitive Load Theory research is referring to.

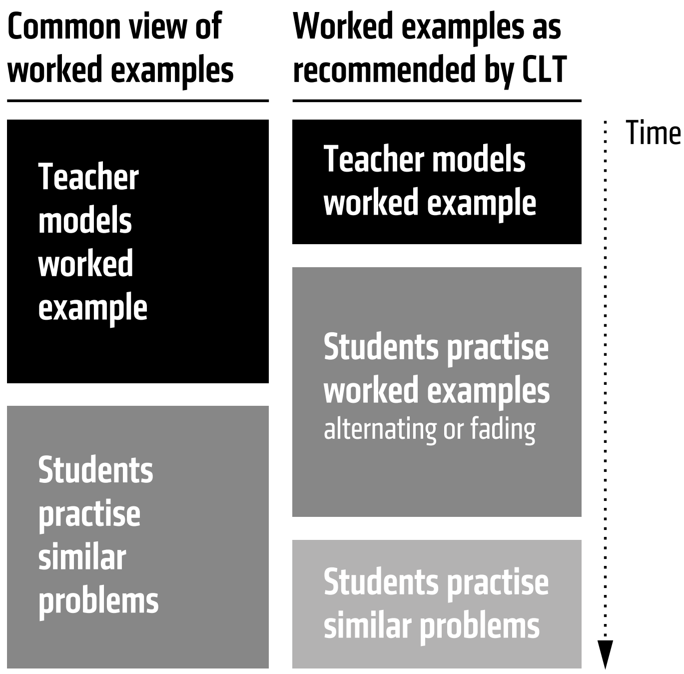


Figure 4. Common view versus CLT view of 'worked examples'

Most teachers (including myself, prior to exploring the CLT research), see worked examples as something like the following: The teacher models two questions on the board, asks the students if they have any questions (perhaps does a check for understanding), then sets the students free on some independent practice. However, this is not at all what CLT means by 'worked examples'.

In the CLT literature, worked

Worked examples	**Similar problem 1.**
We are practising using the phrase **'…In contrast, …'** 'In contrast' is used to show that the ideas presented before and after the 'in contrast' are opposites, or almost opposites.	Write a sentence using '… *In contrast*, …' for each of the topics below. *Don't forget this comma!*
Food example: *My favourite food is baked beans. In contrast, my brother hates them!*	Food example:
Film example: *Harry thinks that The Matrix is an excellent film. In contrast, I think it's boring.*	Film example:
Sport example: *I am very bad at soccer. In contrast, Faduma is amazing!*	Sport example:
Music example: *My Dad loves classical music. In contrast, my mum is a big fan of heavy metal.*	Music example:
Example from school subjects (harder): *In English we indicate that we are asking a question by raising the pitch of our voice at the end of a sentence. In contrast, in Mandarin a question is indicated by saying 'ma' at the end of a sentence.*	Example from school subjects (harder):

Figure 5. Worked examples in a lesson on sentence structure

examples are a form of instruction that sits between teacher modelling, and student independent practice, and extends the period of scaffolded learning. Figure 4 captures the distinction.

A good example demonstrating the efficacy of this type of approach comes from Ward and Sweller (1990). In this study, one group of students were given standard homework that consisted of ten practice problems. Another group were given 'alternating worked examples', which means that they were given five worked problems, and five very similar problems for them to do themselves. In this context, the students who did the alternating worked examples, even though they did only half as many questions, did significantly better. The key idea here is that students often spend much longer in the novice stage than we might expect, and persisting with more structured support for longer, in the form of worked examples, can often be beneficial.

Worked examples are often considered primarily the purview of 'algorithmic' subjects like mathematics. However, they can also be effectively used in many other arenas. Figure 5 provides an example of what worked examples could look like in supporting students with their writing.

The worked example demonstrated above, and the homework example, are both forms of 'alternating' worked examples. The student reads or sees a worked problem presented, then attempts a similar one herself, then again encounters another worked example, then does yet another herself, and so on.

Another form that has been shown to be very effective is the *faded* worked example. Faded worked examples start with a fully worked example, then omit a line of working in each subsequent problem, requiring students to do more and more of the problem type on their own. This approach is captured in Figure 6.

This approach could be used with the 'in contrast' example above, with a maths problem, or even when teaching primary students to 'count on' or 'skip count'. Here, for example, is how a faded approach could be used to teach skip counting, as lead by the teacher:

1. Students count out loud together to 36.
2. Count to 36 emphasising (loudly, it's fun!) every third number.
3. Count to 36 whispering the numbers in between every third number.
4. Count to 36 emphasising every third number and whispering all other numbers.
5. Count to 36 only saying every third number but tapping along on desks for other numbers.
6. Do the same, but this time without overtly tapping.
7. Do the same, challenge students not to bob their heads or do anything that can externally be seen as counting.
8. Skip count from 0 to 36 by threes.

Conclusions

I have only been able to present a few of the key ideas of Cognitive Load Theory within this short article. I have discussed the 'ACE' of CLT, and provided practical examples relating to the split-attention effect, the transient information effect, and alternating and faded worked examples. But we've really only been able to scratch the surface. There are many other valuable effects – all with intriguing names - that it's important for teachers to know about, such as the redundancy effect, the modality effect, the expertise-reversal effect, and the goal-free effect. There are also other CLT recommendations, such as those targeted at optimising intrinsic load. These strategies include pre-teaching, segmentation, and carefully considering and designing sequencing and combination of concepts. I hope that this article has whetted your appetite, and encouraged you to explore this fruitful area of research, and the wide array of classroom recommendations, further.

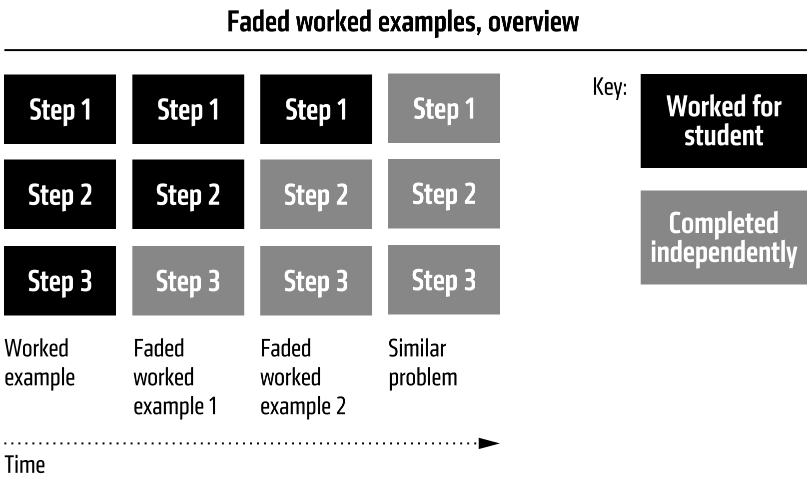


Figure 6. Overview of faded worked examples

References

Bobis, J., Sweller, J. & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction, 3(1)*, 1-21.

Lovell, O. (2020). *Cognitive Load Theory in Action*. John Catt Educational. https://www.ollielovell.com/book/

Ward, M. & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, 7(1), 1-39.

Dr Ollie Lovell is a head of senior mathematics and a learning specialist based at a secondary college in Melbourne. He writes books on education, presents in-services, and runs a very informative mailing list - https://www.ollielovell.com. Ollie would also be delighted if readers found out what Social Venture Partners do at https://svpmelbourne.org.au