Cognitivist Theories I: Cognitive Load Theory (A Hitchhiker's Guide to Learning Theory)

Patrick Parrish, WMO

Background

Cognitive Load Theory is based on a very simple premise. That premise is that we have a limited ability to process information at any given time, that we can keep only so many things on our minds at once. We can be at times more or at times less focused and attentive, but our capacity doesn't vary that much.

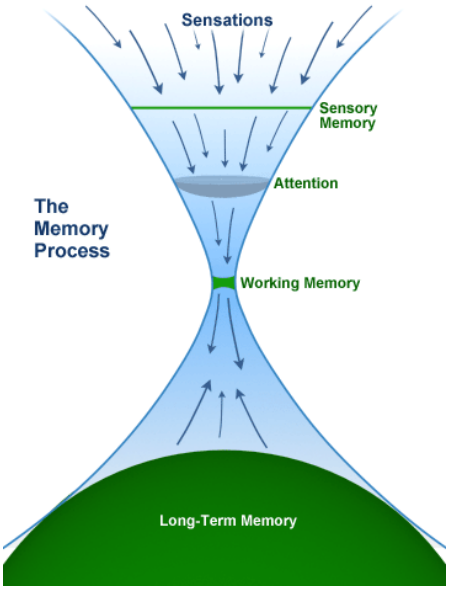
Cognitive science research tells us pretty clearly that our minds have what is called limited "working memory," which is the part of our minds where we hold things in our consciousness as we think about them. If we are comparing two things, for example a satellite image and an NWP analysis, we can only focus on a few parameters at a time as we do the comparison. When we ask a learner to focus on too many things at once, we may be increasing cognitive load beyond their capacity and reducing their ability to learn.

Some everyday examples are trying to remember a long telephone number, or the names of the people you meet at a large party. Trying to keep these things in your mind is a bit like juggling. Research has repeatedly shown that the number of things people can keep on their mind is typically about 7, plus or minus 2 (5 to 9). When we group or chunk things, this effectively increases the number of things quickly available to working memory, because the chunks can be thought of as one thing, but the limit is still there.

For weather forecasters, applying conceptual models, overlaying products, or combining satellite channels are strategies to reduce the cognitive load required for analyzing data products. These strategies allow them to see patterns and indications of features that would be otherwise require long examination of individual data points, and likely be missed due to high cognitive load.

The research that led to this theory began in the 1950's, but it has always been clear that we can perceive and think about only a few things at once.

In this diagram, working memory is represented by the bottleneck in the middle, the thinnest point. The top and bottom of the diagram are drawn to indicate that those features are so large they can't be represented. There is an almost infinite world of potential sensations and ideas that we could attend to if we had the capacity, and our long-term memories are also enormous (even though we have a hard time retrieving memories at times). But to perform any serious thinking, like we have to do when we are learning and performing even basic cognitive tasks, we have to ignore all



but a few things available from the outside world or from our long-term memories.

There are two other important elements in the diagram. The first is Sensory Memory, which is where the images we see and sounds we hear are temporarily stored. We aren't conscious of all of these, so sensory memory is larger than our working memory. And then there is Attention, which acts like a lens to determine which things we will make available to and keep in our working memories. An overloaded sensory memory that vies for our limited attention, such as too many conversations and or too much information on an instructional visual, frustrates our ability to focus attention and have a well-functioning working memory.

I suspect there should be another lens between working memory and long-term memory, but I don't think there is a single term for that—there are many things that limit what we can access from long term memory. The one that is most important for learning purposes is how much practice we have had using what we have stored in long term memory. Cognitive and neuroscience research suggest that practice both strengthens and increases the number of links to information stored in long term memory, making it easier to retrieve when needed.

Applications

This is an information processing theory, and it doesn't necessarily discuss the kinds of thinking people need to do with the mental objects they work with. It is a learning theory in that it is often used to support decision making for information and media design, in interface design, and in managing the complexity of learning activities. It reminds us not to overload students with information, or to provide it a little at a time.

A related theory is dual-coding theory, which suggests that we process visual and verbal information with different parts of our brain. Research demonstrates that we can attend to more things if some are visual and some verbal, rather than if they were all in a single mode.

One key way cognitive load theory is used is to call for simplifying presentations of information by reducing how much content they contain. If we want learners to attend to it all, not miss things, or not get frustrated by being overloaded, try to keep individual items to less than ten on any slide, graphic, or section of a text. Dual coding theory additionally suggests that if we present some information orally, in narration or live speech, and some visually, we might be able to get away with more. We violate the recommendations of cognitive load theory when we provide a complex graphic or animation on a web page followed by a long segment of text, because this requires learners to keep the graphic in working memory at the same time they are reading about it. It's better to use narration while the graphic is presented, or to repeat the graphic during a long explanation. In the same way, it's a good idea to design graphical interfaces so that they provide limited options, or group options in clearly organized ways. Richard Mayer is currently well known for his research on cognitive load theory for multimedia designs.

Another way cognitive load theory is used is to suggest that learning tasks need to be kept as simple as possible until learners develop the required expertise. When learning tasks require the use of large

amounts of information, they can be overwhelming for novices. Some strategies for teaching complex cognitive tasks include providing frequent supporting hints, prompts, and feedback (known as "scaffolding"), providing just-in-time support on the job, sequencing tasks from simple to complex, and teaching parts of tasks before requiring learners to practice whole tasks.

Limitations

Cognitive load theory is seductive in its simplicity and apparent irrefutability. But because Cognitive Load Theory deals almost exclusively with information processing, it is rather limited in its description of learning. For example, it says little about how knowledge is structured in memory, and doesn't discuss how increased engagement and increased attention facilitates information processing. It doesn't tell us about the nature of complex cognitive tasks, such as problem solving, other than that they **are** complex. And it doesn't address how communities of professionals work together to solve problems.

Using only cognitive load theory to explain learning is in a way like explaining a shirt by describing its buttons and button holes. Sure they are necessary, but they don't begin to describe what protects you from the cold wind. I think it's time we put out our thumbs for another ride.