Part 1: Teaching with Working Memory in mind
Cognitive Load Theory
What is it, and why should I care?

Dylan Wiliam
@dylanwiliam

I've come to the conclusion Sweller's Cognitive Load Theory is the single most important thing for teachers to know bit.ly/2kouLOq

6:16 PM - Jan 26, 2017

21 Retweets 418 Likes 650 Favorites
Working Memory and Long Term Memory

- Sensory Memory
- Working Memory
- Long Term Memory

Flow:
- Incoming Information
- Sensory Memory
- Working Memory
- Long Term Memory

Process:
- Encoding
- Retrieval
- Rehearsal

Issues:
- Forgetting
Knowledge, long term memory and schema

- More on this later!
- **Infinite** in capacity
- Retrieval can be a challenge: need **cues**
- Schema = elements of information, organised
- As expertise develops, *schema* become increasingly complex
Working memory

- Receives information from environment
- Temporarily holds information for processing
- Limited capacity for new information (approx 4 items)
- Retrieves suitable schema from Long Term Memory
Cognitive Overload

• If working memory is overloaded:
  
  • Content can be misunderstood
  
  • Content will not be effectively encoded into Long-term Memory
  
  • Learning will be slowed down

• Cognitive Load is too great (cognitive overload)
“For to every one who has will more be given, and he will have abundance; but from him who has not, even what he has will be taken away.”

—Matthew 25:29
CLT: Applications

Worked Example Effect

• Unguided problem solving places a heavy burden on Working Memory.

• Worked examples (a problem that has already been solved, with each step explained) reduce this burden.

• Students are more likely to remember the problem solving strategy later.
Example: calculation questions

• Important tips:
  • Check with Maths department
  • Consistent approach throughout school
  • Model it
  • Insist that pupils show working out
  • Avoid “triangle method” if possible
Example process

1. Identify **what you know** and **what you need to find out**

2. **Select** equation

3. **Write out** equation

4. Check **prefixes/units**

5. **Substitute** into equation

6. **Simplify/rearrange** as necessary

7. **Calculate** answer

8. **Check** answer (is it reasonable?)

9. Check **units**, significant figures/decimal places.
1. Identify what you know and what you need to find out
2. Select equation
3. Write out equation
4. Check prefixes/units
5. Substitute into equation
6. Simplify/rearrange as necessary
7. Calculate answer
8. Check answer (is it reasonable?)
9. Check units, significant figures/decimal places.

(ii) To heat the water in the tank from 50°C to 58°C the immersion heater transfers 4032 kJ of energy to the water.

Calculate the mass of water in the tank.

Specific heat capacity of water = 4200 J/kg°C

...............................................................................................................
...............................................................................................................
...............................................................................................................
...............................................................................................................

Mass = .............................................................. kg
CLT: Applications

Expertise Reversal Effect

• The exception to the Worked Example Effect.

• Use of worked examples becomes less effective as learners’ expertise increases, eventually becoming redundant or even counterproductive.

• Solution: give two worked examples, then reduce scaffolding (partly solved examples, fill in the gaps etc.).
CLT: Applications

Split attention effect

- Multi-tasking not possible. Attention jumps from one task to another.

- High cognitive load - both tasks/sources of information need to be in Working Memory at same time and need to be mentally integrated.

- Can demonstrate with a simple timed task...
CLT: Applications

Split Attention Effect

• Instructions:

• Pair up with a partner. One person to do the task, the other to time/record results.

• Task A

• Time how long it takes to write the alphabet (A to Z) and then the numbers 1 to 26. Note down the time.
CLT: Applications

Split Attention Effect

• Task B

• Time how long it takes to write the alphabet (A to Z) and then the numbers 1 to 26, but alternating between the two, e.g. A1 B2 C3…

• Solution: try not to split sources of information, e.g. position text related to diagram ON the diagram (not positioned separately).
Split attention effect: example

Taken from the Ben Rogers’ blog: https://wp.me/p4Sf4J-C7
Split attention effect: example

Taken from the Ben Rogers’ blog: https://wp.me/p4Sf4J-C7

Mass of plank = 32kg.

Calculate the vertical supporting force from the rock.
Split attention effect: example
Taken from Rebecca Torrance Jenkins, *Using educational neuroscience and psychology to teach science Part 1*. SSR December 2017

![Heart Diagram](image)

**Key**
- **A** vena cava
- **B** aorta
- **C** pulmonary artery
- **D** pulmonary veins
- **E** right atrium
- **F** left atrium
- **G** right ventricle
- **H** left ventricle

**Figure 2** The spatial contiguity principle: (a) reducing extraneous load by integrating labels with visualisation; (b) extraneous load is increased when labels are not integrated with visualisation
What did you see?
Cognitive Load: Applications

Modality Effect

• Visuals are often more effective. Words can over complicate. Can we make better use of visuals?

• Can decrease cognitive load by using both auditory and visual channels simultaneously.

• Try not to use text to explain a diagram. Explain verbally instead.

• Avoid visual “noise”. Keep diagrams and images simple.
Taken from David Paterson’s blog post, Cognitive load and practical work research – an update: [https://wp.me/p8KXzn-bG](https://wp.me/p8KXzn-bG)
Taken from Adam Boxer’s blog post, Cognitive Load in Chemistry Practicals: http://bit.ly/2qdpEI6
Investigating the extension of a spring
Aim:
To investigate the extension and work done when applying forces to a spring.

A.
Zero on ruler is level with bottom of spring.

B.
Add 100g mass. Wait for it to stop moving.

C.
Measure new position of bottom of spring. This is the extension.

D.
Repeat step B and step C five more times.

Results

<table>
<thead>
<tr>
<th>Mass (g)</th>
<th>Force (N)</th>
<th>Extension (cm)</th>
<th>Extension (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>0.98</td>
<td>0.98</td>
<td>0.001</td>
</tr>
<tr>
<td>200</td>
<td>1.96</td>
<td>1.96</td>
<td>0.002</td>
</tr>
<tr>
<td>300</td>
<td>2.94</td>
<td>2.94</td>
<td>0.003</td>
</tr>
<tr>
<td>400</td>
<td>3.92</td>
<td>3.92</td>
<td>0.004</td>
</tr>
<tr>
<td>500</td>
<td>4.90</td>
<td>4.90</td>
<td>0.005</td>
</tr>
<tr>
<td>600</td>
<td>5.88</td>
<td>5.88</td>
<td>0.006</td>
</tr>
</tbody>
</table>
Cognitive Overload: Causes

The Redundancy Effect

• The redundancy effect occurs when information that includes redundant material results in less learning than the same information minus the redundant material.

• Adding extra or repeated information that isn’t necessary for understanding can be counterproductive.

• Key point: don’t read out slides/notes whilst pupils are reading them. “Clogs up” auditory channel of working memory (phonological loop).
Cognitive Load: Applications

Goal-free Effect

• Typically, novices will adopt a means-ends approach to problem solving (work backwards from the goal to the problem, then forwards again).

• This strategy = heavy cognitive load

• Pupils often get the answer, but forget the strategy.
Goal-free effect: example

Taken from the Ben Rogers’ blog: https://wp.me/p4Sf4J-B8

1 (b) A uniform plank of wood of mass 32 kg and length 4.0 m is used by a boy to help him cross a ditch. In the ditch is a rock, which is used to support the plank horizontally 0.80 m from one end, as shown in Figure 1. The other end of the plank is supported by the bank.

Figure 1

Calculate the vertical supporting force from the rock when the plank is placed in position as shown in Figure 1.

[2 marks]
Goal-free effect: example

Taken from the Ben Rogers’ blog: https://wp.me/p4Sf4J-B8

What do we know, and what can we find out about the situation above?

A uniform plank of wood of mass 32 kg and length 4.0 m is used by a boy to help him cross a ditch. In the ditch is a rock, which is used to support the plank horizontally 0.80 m from one end, as shown in Figure 1. The other end of the plank is supported by the bank.
Summary: CLT Effects

• Be aware of the potential to overload working memory. **Try to limit new information to 3 or 4 items at a time.**

• Reduce extraneous load by removing (where possible) tasks/distractions that do not contribute to learning.

• Tackle intrinsic load by **teaching smaller chunks of knowledge** and then building up to form a bigger picture.

• Make good use of **worked examples.** Use “faded” examples (where there are gaps etc) to avoid the Expertise Reversal effect.
Summary: CLT Effects

• Use/create resources and activities that reduce the **Split Attention effect**.

• Decrease cognitive load by using both auditory and visual channels simultaneously.

• **Avoid redundant material/info.** Avoid reading out slides whilst pupils are reading them.

• Use “goal free” tasks when pupils are still developing expertise.
Further reading etc

• Working Memory (including interviews with Alan Baddeley): https://www.simplypsychology.org/working%20memory.html

• Cognitive Load Theory
  - NSW Department of Education - Cognitive load theory: Research that teachers really need to understand
  - Blake Harvard (The Effortful Educator) blog post: Cognitive Load Theory and Applications in the Classroom
  - Ben Rogers’ blog posts: https://readingforlearning.org
  - David Paterson’s blog posts: https://dave2004b.wordpress.com
  - Adam Boxer’s blog posts (https://achemicalorthodoxy.wordpress.com) including this post on CLT and practicals.
Part 2: Making the most of long-term memory
The Challenge

- 2 year (often 3 year) course with terminal exams.
- More challenging (especially for low attaining pupils).
- Retention of knowledge and understanding from early Y9 to end of Y11.
- Speed of rollout = little time to plan curriculum.
 Working Memory & Long Term Memory

Incoming Information → Sensory Memory → Working Memory

Retrieval → Encoding

Rehearsal

Forgetting → Forgetting → Forgetting
Working Memory

• Receives information from environment via sensory memory

• Temporarily holds information for processing

• Limited capacity for new information (approx 4 items)

• Retrieves suitable schema from Long Term Memory
Knowledge, long term memory and schema

- Infinite in capacity
- Retrieval can be a challenge: need **cues**
- *Schema* = elements of information, organised
- As expertise develops, *schema* become increasingly complex
Knowledge, long term memory and schema

More knowledge = more connections = greater understanding = more retrieval cues
Knowledge, long term memory and schema

- More elaborate schema = more pathways and cues to aid retrieval
- *Schema* = **only single space** in working memory
- Extensive practice leads to **automation**. Schema accessed automatically **without burden on working memory** (e.g. driving)
Forgetting

Why do we forget?

● Interference theory (memories interfere and disrupt each other)
  ○ Proactive: old memories disrupt new memories
  ○ Retroactive: new memories disrupt old memories

● Lack of consolidation (changes to neural networks)

● Retrieval failure theory (retrieval cues)
There are two categories of memory strength:

1. **Retrieval Strength**: how easily something is recalled NOW

2. **Storage Strength**: how deeply embedded/learned something is.
Desirable Difficulties

1. Studying information causes **both** Retrieval and Storage Strength to increase.

2. After studying, the **higher** the Storage Strength, the **slower** the loss of Retrieval Strength (forgetting is slower)

3. When restudying, the increase in Storage Strength is inversely related to current Retrieval Strength: **the harder it is to bring to mind, the greater the improvement in memory.**

Desirable Difficulties

Difficult (but accomplishable) tasks may slow down learning initially, but have greater long term benefits.

Learning vs Performance

Problems...

1. Learning and performance are different

2. Conditions that maximise gain in retrieval strength differ from conditions that maximise storage strength

3. If learners interpret current retrieval strength as storage strength, they can develop a preference for poorer learning conditions.
Examples of desirable difficulties

- Desirable difficulties are conditions that help trigger encoding and retrieval processes that support learning, understanding and remembering.

- Note: if learner does not have background knowledge or skills to respond to this successfully then they become undesirable difficulties.

- Well tested examples:
  - Spaced practice
  - Retrieval practice
  - Interleaving
  - Varying conditions of practice (e.g. location)
The Spacing Effect

- Spacing/distributing learning >> cramming.

- First discovered by Ebbinghaus in 1885! Well established and thoroughly tested.

- Cramming may generate high retrieval strength (information is fresh) but will not lead to significant gains in learning (storage strength).

- Spacing learning may be more challenging, but leads to much greater gains.

- How can we make use of this in the classroom?
Fig. 10. Proportion of items answered correctly on an initial test administered in each of six practice sessions (prior to actual practice) and on the final test 30 days after the final practice session as a function of lag between sessions (0 days, 1 day, or 30 days) in Bahrick (1979).
Spaced repetition
Interleaving

- A form of spacing.
- Mix up topics and tasks.
- Provides desirable difficulties, but also encourages pupils to notice connections between topics (transfer) and identify which problem solving solutions to use.
- Many studies show it to be effective in both cognitive thinking and motor skills (e.g. sports)
Fig. 13. Percentage of correct responses on sets of problems completed in practice sessions and on a delayed criterion test in Rohrer and Taylor (2007). Error bars represent standard errors.
Generation Effects (the testing effect/retrieval practice)

- Retrieval is a powerful “memory modifier” (Bjork 1975)

- Any time a learner looks up an answer or is told the answer to something that they could (drawing on current cues and past knowledge) generate instead, a powerful learning opportunity is missed.

- Tests can be used as learning events (rather than just for assessment purposes)
Figure 1. Study design comparing the effects of study versus retrieval practice. Derived from Roediger and Karpicke, 2006.
Figure 2. Effects of repeated studying versus repeated retrieval practice. Derived from Roediger and Karpicke, 2006.
Retrieval Practice

- Not just tests/exams! Quizzing/self-quizzing.
- This benefit occurs even in the absence of feedback and in the absence of an opportunity to re-study the information.
- Can also benefit from interleaving (mix up questions from different topics)
- Additional benefit: provides pupils with useful feedback as to what has/hasn’t been learnt.
- How can we make effective use of this in the classroom?
Retrieval Roulette

• See Adam Boxer’s post and collated resources:
  
  • http://bit.ly/RetrievalRoulette

• Use at start of lessons before teaching new content.

• Can be sent home for parents/pupils to use.
Homework

- Homework CAN be highly effective, but needs to be:
  - Short and manageable
  - Task orientated, focussed tasks (rote learning, practice/rehearsal)
  - Closely monitored by teachers
- Homework can provide an excellent opportunity for retrieval practice
The Spiral Curriculum

● Learning is not linear in nature.

● A spiral curriculum has benefits to learning because:
  ○ Material is revisited
  ○ Complexity is increased with each cycle (simple schema become more complex)
  ○ Complexity of material can increase with student age

● How should we organise the curriculum?
Threshold Concepts

Threshold concepts represent

“a transformed way of understanding...something without which the learner finds it difficult to progress within the curriculum”

(Meyer and Land, 2003)

“The easiest way to envisage a threshold concept is as a “gateway” or “portal” that leads the learner to a previously undiscovered, and perhaps inaccessible, way of thinking.”

(Stokes, King and Libarkin, 2007)
Threshold Concepts

Threshold concepts are:

- **Transformative**: they change the way students think

- **Irreversible**: once learned they are unlikely to be un-learned or forgotten

- **Integrative**: they build connections between previously separate concepts

- **Troublesome**: they may require students to deal with knowledge that is conceptually difficult, counter-intuitive or “alien” in some way
Threshold Concepts

Discuss:

- What are the threshold concepts in GCSE Science?

- Pick a threshold concept:
  - What do students need to understand to be able to access the threshold concept?
  - What can students now access once they’ve understood the threshold concept?
Further reading

- The Learning Scientists blog: http://www.learningscientists.org/

- Rosenshine (2012). *Principles of instruction. Research based strategies that all teachers should know.*


Next Steps

- Twitter: #CogSciSci


- CogSciSci mailing list

- Books:
  - *What if everything you knew about education was wrong?* - David Didau
  - *What does this look like in the classroom?* - Carl Hendrick and Robin Macpherson
  - *Applying Cognitive Science to Education* - Frederick Reif
  - Plus for Physics teaching: *The Big Ideas in Physics and How to Teach Them* - Ben Rogers

info@bobpritchard.education
@rjpritchard