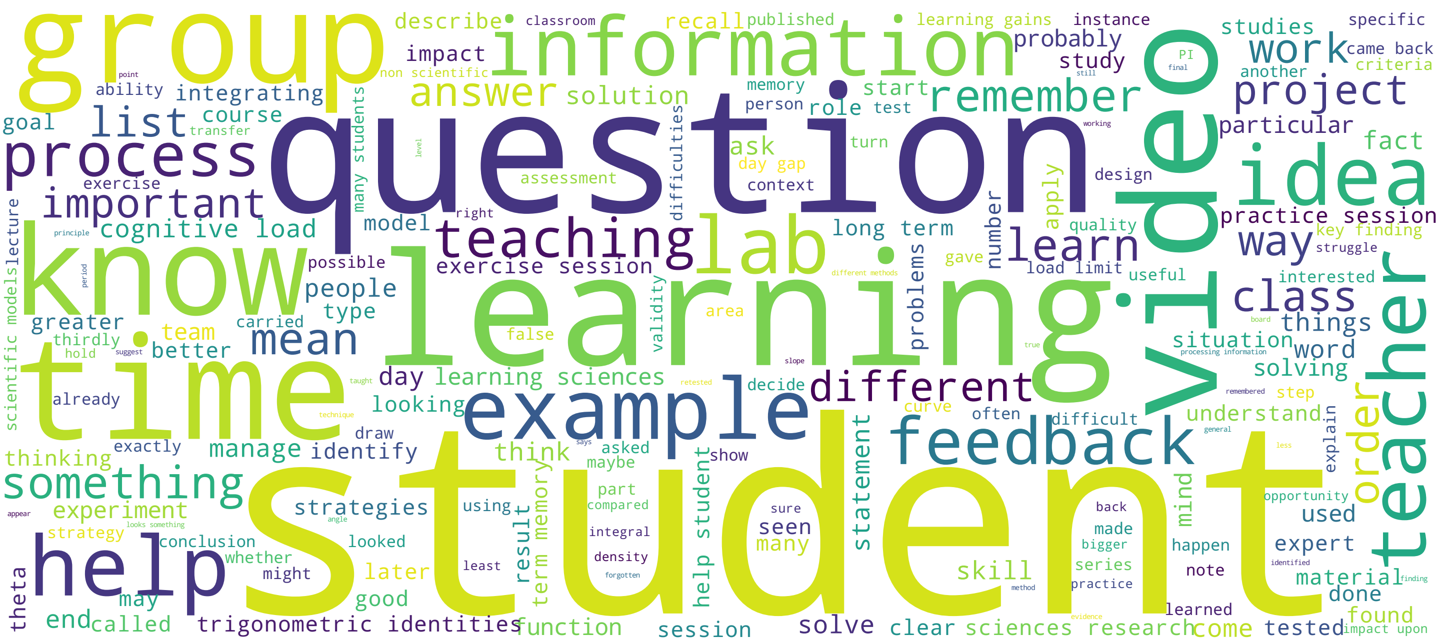
Evidence about Learning (Part II)

Foundations for Teaching in Higher Education

Roland Tormey
Ingrid Le Duc
Cécile Hardebolle

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE



Search MOOC

Video

EPFL

Goals of this video

You should be able to describe what research evidence says about the role of the following things in learning:

- **Cognitive load**
- **Spaced repetitions**
- **Metacognition**

In the last video, we introduced three key findings from the learning sciences research. First, it’s students actively processed information. Secondly, that they need to engage with their naïve or non-scientific prior models which they implicitly bring with them to scientific education and thirdly, that feedback plays a crucial role in terms their learning. In this video, we look at three further important findings from the learnings in sciences research. And so by the end of this video you should be able to describe what the learning sciences research says about the idea of cognitive load, what it says about spaced repetitions and thirdly, you should be able to describe what is meant by the term “metacognition” and what is its impact on student learning.

Notes

Summary

0m 04s

Let’s start with an exercise…

I will read a series of statements.

For each one:

(a) decide if it is true or false

(b) Remember the last word in the statement.

After the last statement write down the list of last words.

In order to understand better the first of these ideas, let's carry out a small experiment. So I'm going to read to you a series of statements. For each statement, you should decide, is the statement true or false and then you should try to remember the last word from the statement. So I'll give you a series of statements and you try to remember those list of last words whilst at the same time deciding if each statement is true or false. Here goes. Dogs have ears. Cats read philosophy. Sausages grow on trees. Milk comes from cows. Wood floats on water. Vampires crave tea. Oranges are fruit. Wine has wings. Butter is made from iron. Now when you've done that, stop the video here and then write down the list of words that you need to remember.

Notes

Summary

0m 52s

Let’s start with an exercise…

The list:

- Ears
- Philosophy
- Trees
- Cows
- Water
- Tea
- Fruit
- Wings
- Iron

Okay. So if you managed to do that you'll have a list of words which looks something like this list here. It's not really important whether or not you manage to remember all the words. What's important to note is that it probably became more difficult to try to manage the two different tasks keeping in mind a list of words while at the same time deciding if each statement was true or false. And that's interesting. Right. Because actually it's not a very long list of words. These are not very difficult statements. How is it that it can be hard to do this when it's possible for people to understand you know Newtonian physics, thermodynamics or chemical bonding. And so this is linked to this idea of 'cognitive load'.

Notes

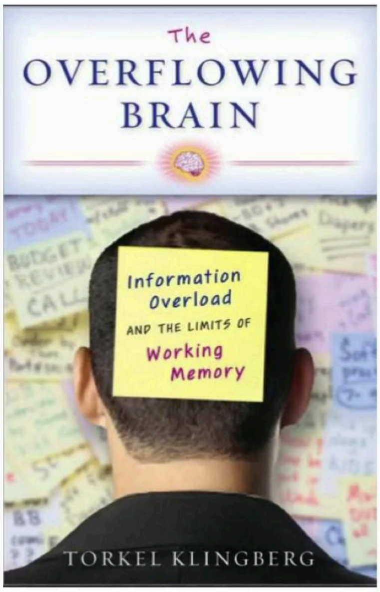
Summary

2m 06s

Limits to our processing capacity

“The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information” (1956)

- George A. Miller, Harvard University



So the experiment you've just done is described in this book “The Overflowing Brain”, but the idea which underpins it comes originally from a paper which was published by George Miller from Harvard University in 1956. It's one of the most famous papers in the history of psychology and it's called “The magical number seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information.” Miller was interested in how many discrete pieces of information a person could hold in their memory at any given time. His answer was 7 plus or minus 2. In fact, more recent research suggests that people can't really hold 7 pieces of information especially if they need to process them.

Notes

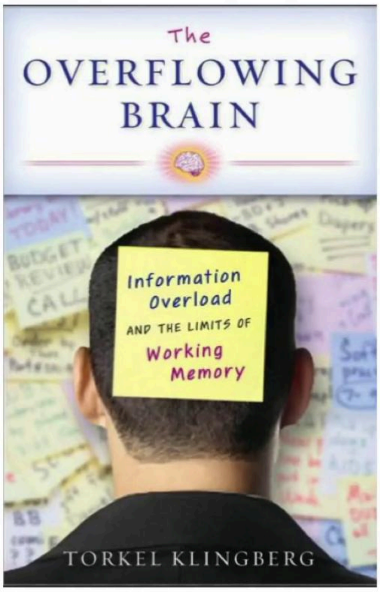
Summary

2m 58s

Limits to our processing capacity

"The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information" (1956)

- George A. Miller, Harvard University
- Limits to processing…
  - **New information** or
  - **Actively recalled information**
- Experts have **connected** information
- Novices have **unconnected** information



So how is it that we can hold all of this information about thermodynamics or classical mechanics in our mind and use it but at the same time we have this cognitive load limit. It's because the cognitive load limit applies to new information or information that we are actively needing to recall. Information that we have stored in our memory but we have to go looking for it in our mind. Our cognitive load limits are also affected by how we store information in our memory. Novices tend to store information as discrete individual pieces or chunks. For example, if there are three key equations describing movement when gravity is the only acceleration, novices will store them separately and recall them one by one. Experts will know how these three equations relate to each other and how they can each be derived from a single equation. They will typically recall all three equations as a single piece of information because they have already knitted them together. Novices will end up treating the equations as three discrete pieces of information leaving limited space for thinking about other things. Experts will treat the three as a single piece of information leaving them much more free head space for thinking.

Notes

Summary

3m 40s

An example

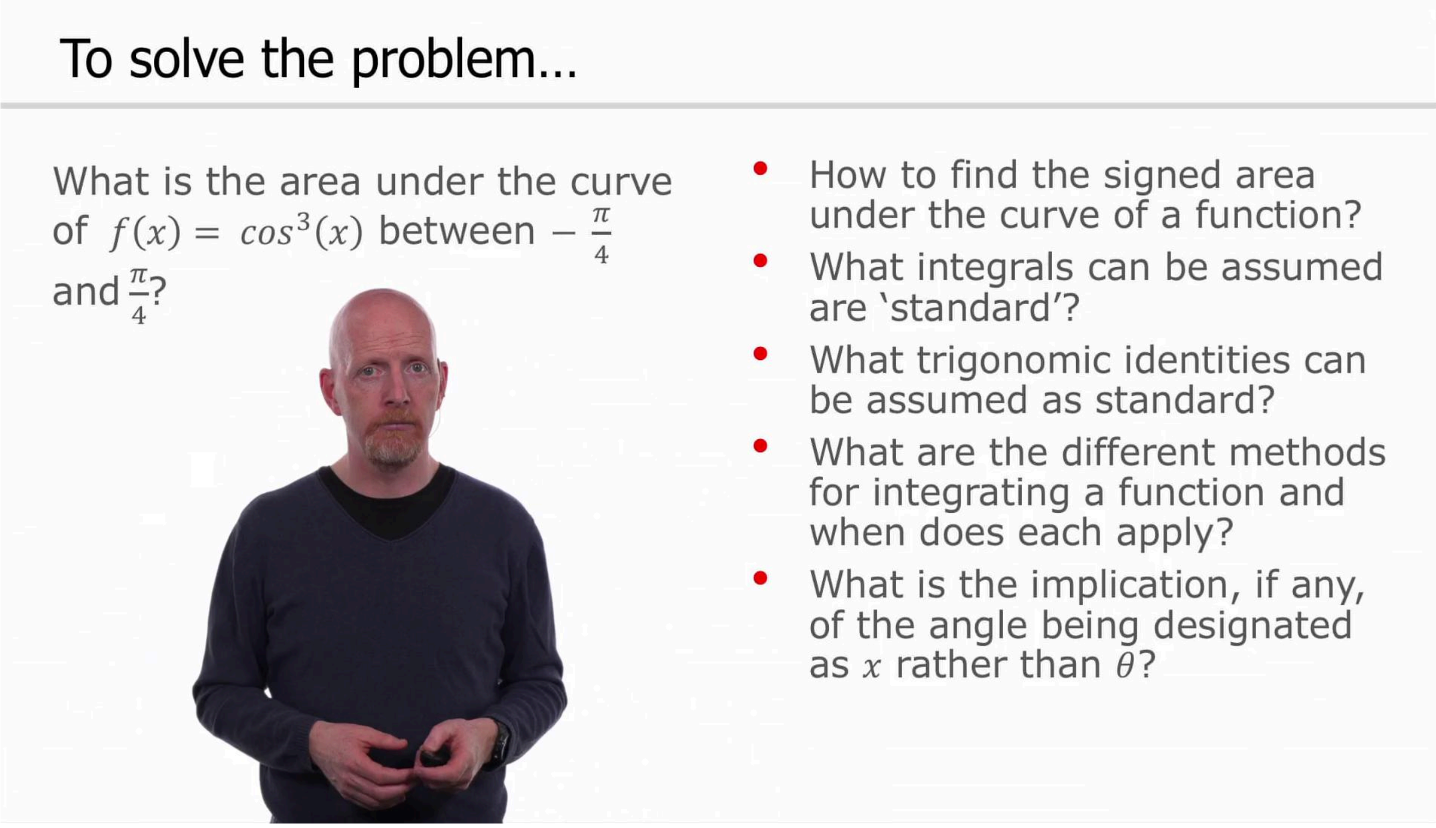
- A calculus class on integrating functions using trigonomic identities.
- The teacher has integrated $f(\theta) = cos^2(\theta)$ in class.
- One of the exercises is "What is the area under the curve of $f(x) = cos^3(x)$ between $-\frac{\pi}{4}$ and $\frac{\pi}{4}$? "

This has clear implications for teaching. Let's think of an example. Imagine a calculus class and in the calculus class the teacher is teaching students about difficult to integrate functions and integrating them using trigonometric identities. In the course of the class the teacher does an example on the board for the students to see. The teacher integrates F of theta is equal to cos squared of theta in the class and then the students go into the exercise session. And in the exercise session the students are faced with a problem which on the surface of it looks quite similar to the previous one. Here the problem is that they need to identify the area under a curve of F of X is equal to cos cubed of X between minus PI over 4 and PI over 4. The student will be able to use the same general approach, that is, they'll be able to use trigonometric identities in order to solve this problem. But the same trigonometric identities probably isn't the best fit in this case. They'll need to use different trigonometric identity and they also probably need to use a different approach to integration in order to do it, a different integration technique.

Notes

Summary

5m 01s



So to solve this problem what information does the student need to draw into their mind. First of all, the student needs to think about how to find the area, the signed area under a curve of a function. This is something they know. They've dealt with previously but it's something they have to actively recall. It's not something that's simply ready in the forefront of their brain. Another thing that they have to actively recall or at least look up is what integrals can they assume are standard what integrals do they have to calculate and which ones can they simply provide. They'll have a similar issue with trigonometric identities. It may be that some of the trigonometric identities involved, they won't have looked at for weeks, months or maybe even years and yet now they'll need to be able to bring it to mind and use it straight away. Then they have the problem of integrating a function. There are different methods of integrating a function and they need to know when these different methods are going to apply. They need to be able to look at the function and recognize what's the most appropriate tool for integrating the particular function they have.

Notes

Summary

6m 13s

To solve the problem…

What is the area under the curve of $f(x) = cos^3(x)$ between $-\frac{\pi}{4}$ and $\frac{\pi}{4}$?

- How to find the signed area under the curve of a function?
- What integrals can be assumed are 'standard'?
- What trigonometric identities can be assumed as standard?
- What are the different methods for integrating a function and when does each apply?
- What is the implication, if any, of the angle being designated as x rather than θ?

Finally, in class the teacher integrated a function which involved theta and that makes sense to the students because theta is often used to describe an angle but in this case the variable X is used. What does that mean? Is it still an angle? Is it a number? What happens when you replace theta with X, if anything? For the student in order to solve this relatively straightforward problem which is really simply reapplication of a technique they've seen in class to this exercise, they need to manage multiple sources of information. They need to manage a large quantity of different pieces of information all at the same time. And here they reach the limits of cognitive load.

Notes

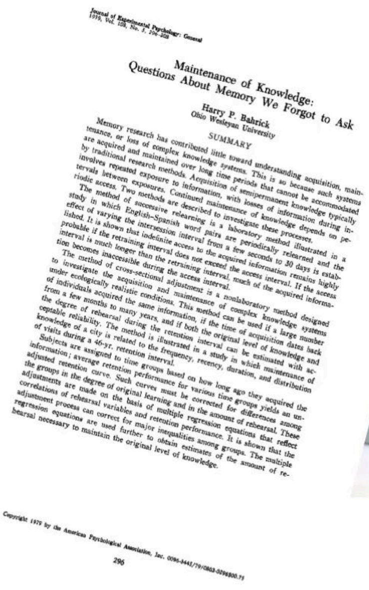
Summary

7m 24s

“Maintenance of Knowledge: Questions About Memory We Forgot to Ask” (1979)

- Harry Barhick, Ohio Wesleyan University

Journal of Experimental Psychology



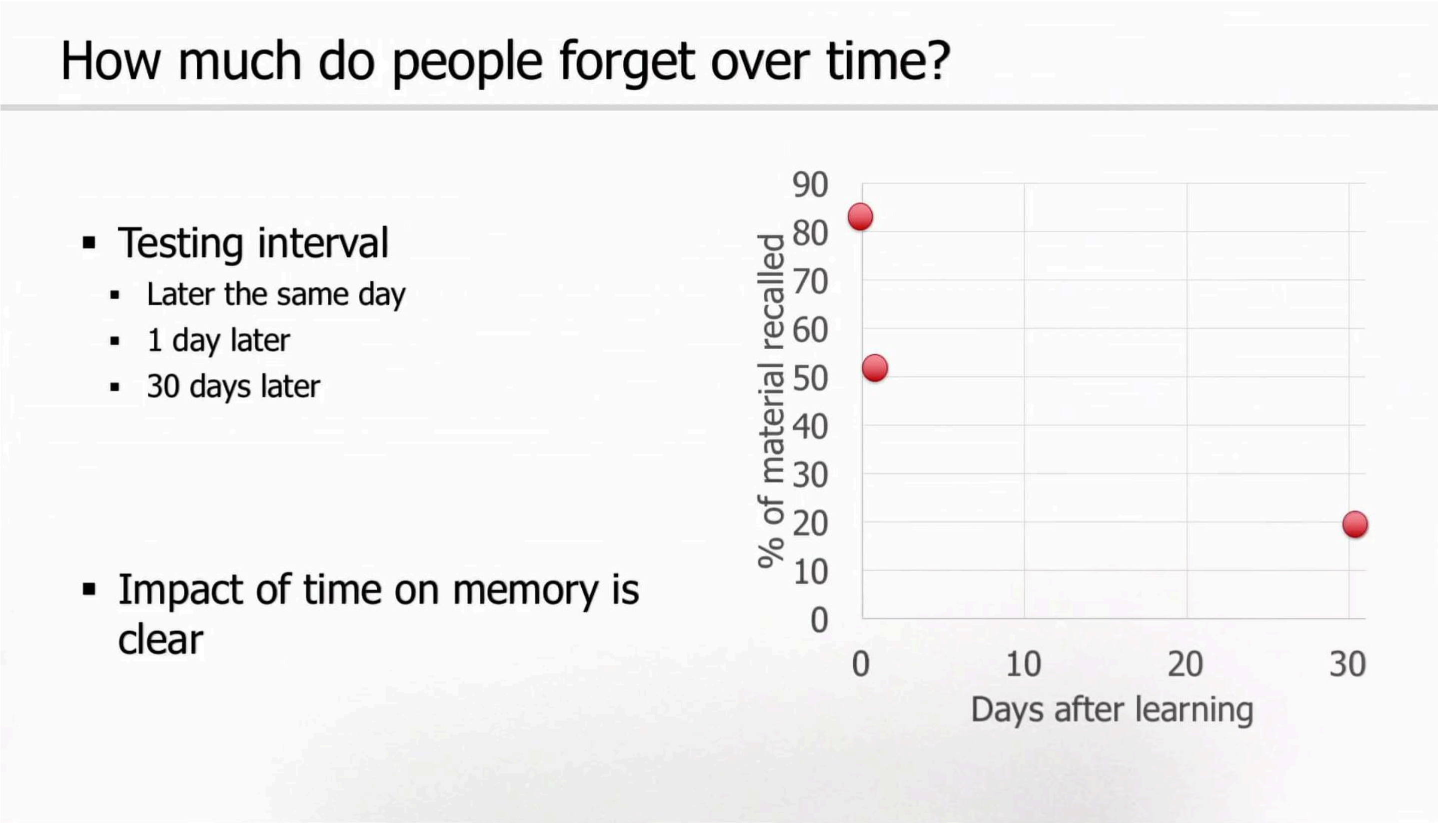
© Alain Herzog

So the first of the key findings we're addressing in this video is about the role of cognitive load in learning. The second is about this idea of spaced repetition when we repeatedly engage with an idea over time and to do this we look at a study which was published in 1979 by Harry Barhick, called “Maintenance of Knowledge: Questions About Memory We Forgot to Ask.” In his study, Barhick asked people to remember things and then having asked them to remember of list the things he came back to them later to test how much they remember.

Notes

Summary

8m 12s



And so what you can see is that, for example, when he tested them later the same day, they remembered most of the things. When he tested him the next day they had forgotten almost half of the things and when he came back to them thirty days later, a different group they remembered only about 20% and this constitutes a curve which looks something like this. Which is actually called the Ebbinghaus curve or the Ebbinghaus forgetting curve. It's a well-established phenomenon in learning sciences research. People forget a lot in the first day or two. After two or three days they probably only remember of 20% and then that will be maintained over a period of time. What we can see is that there's a clear impact of time on memory. If a student has learned something in a lecture, by the next day they will probably have forgotten about half of it, by the day after that they'll probably have forgotten about 80% of it and they'll maintain that that low level of recall over an extended period of time unless we do something else.

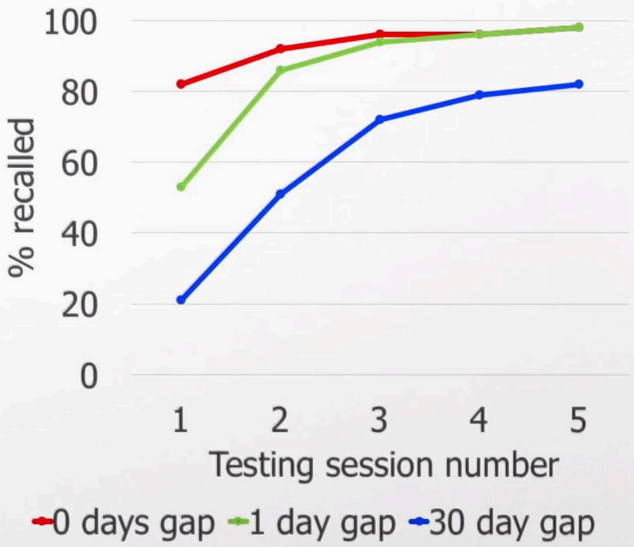
Notes

Summary

8m 51s

How much does repetition help learning?

- Retesting interval
  - **Later the same day**
  - **1 day later**
  - **30 days later**
- Bigger learning gains from longer spaces between sessions

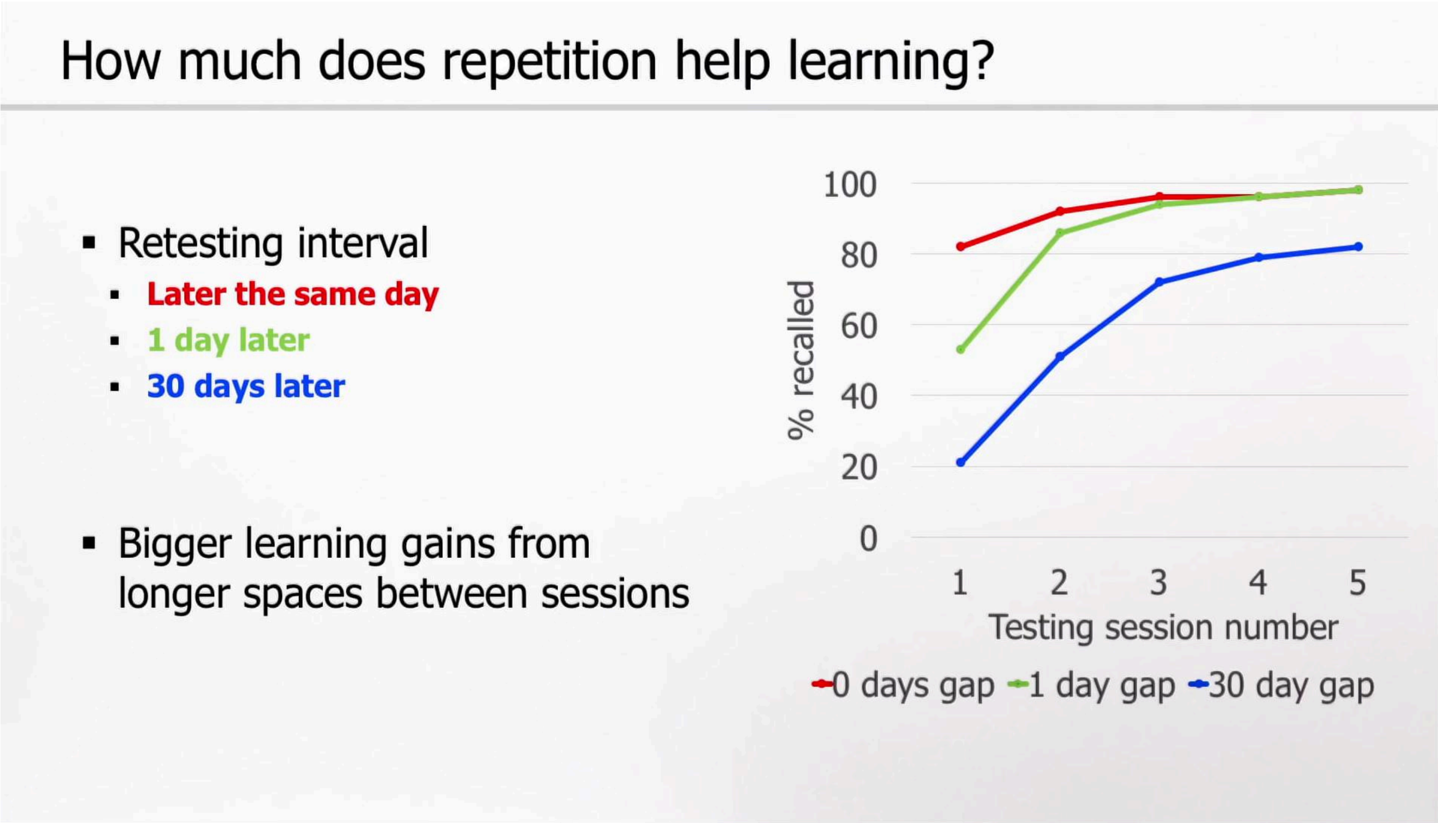


And so what Barrack did was he did something else. He came back to each group and having tested them he then gave him an opportunity to relearn the material, to get back up to full competence in the material and then he retested them. For one group, he retest them later the same day and so this group there, this session here, which is their third session, they have learned the material, they've been tested and relearned it, and then coming back third time, they show a slight increase in terms of recall in the third session. This group here there's a one-day gap between each session. So, this group here, this is their first day after initial learning. This is their second day after initial learning. You can see that the slope here is much greater for this group and it's also greater for this group who are being retested or relearning 30 days later. So what you can see is that the slope is bigger in situations where the gap is greater. The bigger the learning gains come from longer spaces between sessions. And in fact, Barrack returned again and again and again over time to see what happened when people were given opportunities to relearn material over a period of time where there was a spaced interval between.

Notes

Summary

9m 57s



And what you find is again that the bigger learning gains tend to come from situations where there are longer spaces and in particular, the biggest learning gains seem to come where there was a 30-day gap between each study session as compared to where there was a one-day gap or indeed where all of the study happened on the same day.

Notes

Summary

11m 21s

How does spacing practice impact on long term learning

- Test taken 30 days after final practice session
- Group with longer intervals between practice sessions showed better long term memory

Barhick was interested if whether or not spaced practice, leaving intervals between each practice session, would have an impact on long-term memory. And so he gave the participants a test at the end of the experiment. In fact, the test happened 30 days after their last practice session.

Notes

Summary

11m 46s

How does spacing practice impact on long term learning

Space between practice sessions	Test scores
0 days	68
1 day	86
30 days	95

- Test taken 30 days after final practice session
- Group with longer intervals between practice sessions showed better long term memory

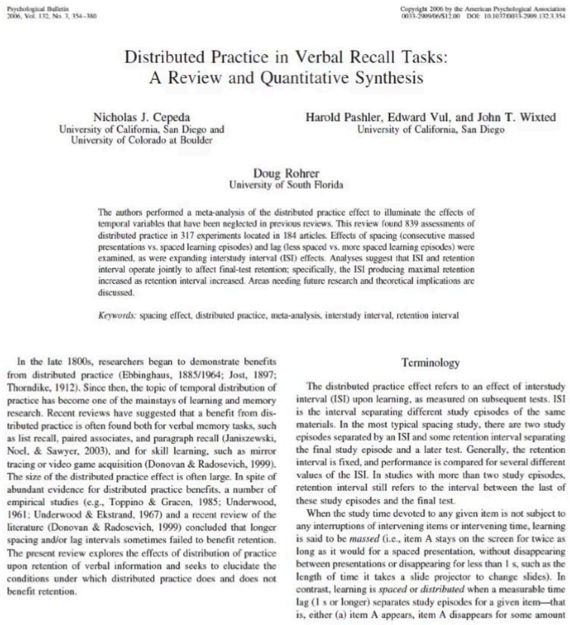
What he found was that the group who had the longest interval between each practice session were the group that had the best test score. So in fact, it does appear that spacing practice has an impact on long-term memory.

Notes

Summary

12m 04s

A consistent research finding



Psychological Bulletin
2006, Vol. 132, No. 3, 354–380

Copyright 2006 by the American Psychological Association
0033-2909/06/$12.00 DOI: 10.1037/0033-2909.132.3.354

Distributed Practice in Verbal Recall Tasks: A Review and Quantitative Synthesis

Nicholas J. Cepeda
University of California, San Diego and
University of Colorado at Boulder

Harold Pashler, Edward Vul, and John T. Wixted
University of California, San Diego

Doug Rohrer
University of South Florida

The authors performed a meta-analysis of the distributed practice effect to illuminate the effects of temporal variables that have been neglected in previous reviews. This review found 839 assessments of distributed practice in 317 experiments located in 184 articles. Effects of spacing (consecutive massed presentations vs. spaced learning episodes) and lag (less spaced vs. more spaced learning episodes) were examined, as were expanding interstudy interval (ISI) effects. Analyses suggest that ISI and retention interval operate jointly to affect final-test retention; specifically, the ISI producing maximal retention increased as retention interval increased. Areas needing future research and theoretical implications are discussed.

Keywords: spacing effect, distributed practice, meta-analysis, interstudy interval, retention interval

In the late 1800s, researchers began to demonstrate benefits from distributed practice (Ebbinghaus, 1885/1964; Jost, 1897; Thorndike, 1912). Since then, the topic of temporal distribution of practice has become one of the mainstays of learning and memory research. Recent reviews have suggested that a benefit from distributed practice is often found both for verbal memory tasks, such as list recall, paired associates, and paragraph recall (Janiszewski, Noel, & Sawyer, 2003), and for skill learning, such as mirror tracing or video game acquisition (Donovan & Radosevich, 1999). The size of the distributed practice effect is often large. In spite of abundant evidence for distributed practice benefits, a number of empirical studies (e.g., Toppino & Gracen, 1985; Underwood, 1961; Underwood & Ekstrand, 1967) and a recent review of the literature (Donovan & Radosevich, 1999) concluded that longer spacing and/or lag intervals sometimes failed to benefit retention. The present review explores the effects of distribution of practice upon retention of verbal information and seeks to elucidate the conditions under which distributed practice does and does not benefit retention.

Terminology

The distributed practice effect refers to an effect of interstudy interval (ISI) upon learning, as measured on subsequent tests. ISI is the interval separating different study episodes of the same materials. In the most typical spacing study, there are two study episodes separated by an ISI and some retention interval separating the final study episode and a later test. Generally, the retention interval is fixed, and performance is compared for several different values of the ISI. In studies with more than two study episodes, retention interval still refers to the interval between the last of these study episodes and the final test.

When the study time devoted to any given item is not subject to any interruptions of intervening items or intervening time, learning is said to be *massed* (i.e., item A stays on the screen for twice as long as it would for a spaced presentation, without disappearing between presentations or disappearing for less than 1 s, such as the length of time it takes a slide projector to change slides). In contrast, learning is *spaced* or *distributed* when a measurable time lag (1 s or longer) separates study episodes for a given item—that is, either (a) item A appears, item A disappears for some amount

Now this is the result of a single experiment but in fact this is an experiment which has been repeated again and again and again since it was first carried out by Barhick.

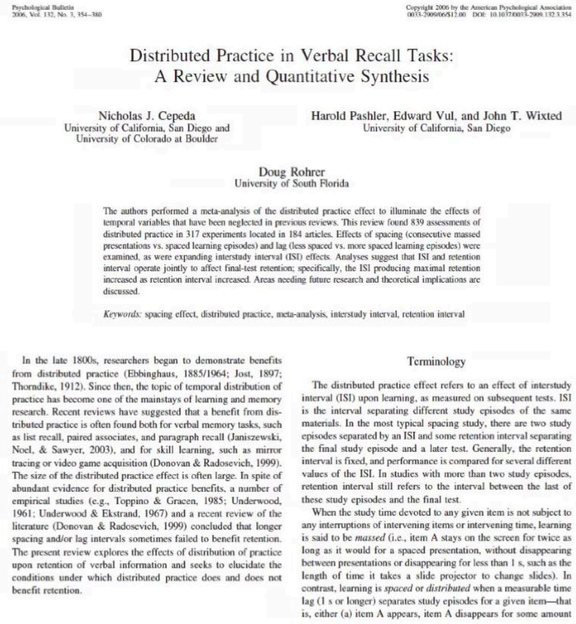
Notes

Summary

12m 20s

A consistent research finding

- "separating learning episodes by a period of at least 1 day… is extremely useful for maximizing long-term retention…
- Every study …with a retention interval longer than 1 month demonstrated a benefit from distribution of learning across weeks or months"
- 317 separate experiments


Psychological Bulletin
2006, Vol. 132, No. 3, 354–380
Copyright 2006 by the American Psychological Association
0033-2909/06/$12.00 DOI: 10.1037/0033-2909.132.3.354

Distributed Practice in Verbal Recall Tasks: A Review and Quantitative Synthesis

Nicholas J. Cepeda
University of California, San Diego and University of Colorado at Boulder

Harold Pashler, Edward Vul, and John T. Wixted
University of California, San Diego

Doug Rohrer
University of South Florida

The authors performed a meta-analysis of the distributed practice effect to illuminate the effects of temporal variables that have been neglected in previous reviews. This review found 839 assessments of distributed practice in 317 experiments located in 184 articles. Effects of spacing (consecutive massed presentations vs. spaced learning episodes) and lag (less spaced vs. more spaced learning episodes) were examined, as were expanding interstudy interval (ISI) effects. Analyses suggest that ISI and retention interval operate jointly to affect final-test retention; specifically, the ISI producing maximal retention increased as retention interval increased. Areas needing future research and theoretical implications are discussed.

Keywords: spacing effect, distributed practice, meta-analysis, interstudy interval, retention interval

In the late 1800s, researchers began to demonstrate benefits from distributed practice (Ebbinghaus, 1885/1964; Jost, 1897; Thorndike, 1912). Since then, the topic of temporal distribution of practice has become one of the mainstays of learning and memory research. Recent reviews have suggested that a benefit from distributed practice is often found both for verbal memory tasks, such as list recall, paired associates, and paragraph recall (Janiszewski, Noel, & Sawyer, 2003), and for skill learning, such as mirror tracing or video game acquisition (Donovan & Radosevich, 1999). The size of the distributed practice effect is often large. In spite of abundant evidence for distributed practice benefits, a number of empirical studies (e.g., Toppino & Gracen, 1985; Underwood, 1961; Underwood & Ekstrand, 1967) and a recent review of the literature (Donovan & Radosevich, 1999) concluded that longer spacing and/or lag intervals sometimes failed to benefit retention. The present review explores the effects of distribution of practice upon retention of verbal information and seeks to elucidate the conditions under which distributed practice does and does not benefit retention.

Terminology

The distributed practice effect refers to an effect of interstudy interval (ISI) upon learning, as measured on subsequent tests. ISI is the interval separating different study episodes of the same materials. In the most typical spacing study, there are two study episodes separated by an ISI and some retention interval separating the final study episode and a later test. Generally, the retention interval is fixed, and performance is compared for several different values of the ISI. In studies with more than two study episodes, retention interval still refers to the interval between the last of these study episodes and the final test.

When the study time devoted to any given item is not subject to any interruptions of intervening items or intervening time, learning is said to be *massed* (i.e., item A stays on the screen for twice as long as it would for a spaced presentation, without disappearing between presentations or disappearing for less than 1 s, such as the length of time it takes a slide projector to change slides). In contrast, learning is *spaced* or *distributed* when a measurable time lag (1 s or longer) separates study episodes for a given item—that is, either (a) item A appears, item A disappears for some amount

One review of the evidence around this phenomenon which was carried out in 2006, found 184 different articles which involved 317 separate experiments. And what they found is that spacing out practice sessions does have an impact on long-term memory. In particular, separating learning episodes by at least one day turns out to be extremely useful for long-term attention. And at a very basic what this means is, if you want to learn something it's good to come back to it the day after you first engaged with it or a few days after. Longer term intervals between practice sessions also appear to be particularly beneficial. Every study that had a retention interval longer than one month, demonstrated a benefit of this distribution of learning across a long-term memory period.

Notes

Summary

12m 30s

Metacognition: Putting it all together

- Many students use
  - Inefficient note-taking strategies
  - Less effective study strategies
  - Ineffective problem solving strategies
- Need to manage their own learning
- Metacognition: “Planning, monitoring, debugging and evaluating your own learning”

So here I come to the final of the key findings from the research literature on learning sciences and this looks at the process of how students manage their own learning. And so when we look at studies of what students actually do when they're learning, it turns out that students are often using inefficient strategies. In lectures for example, we know that students would be well served taking handwritten notes but many students do not do so. In study we know that rereading notes and underlining is amongst the least effective strategies that are available but we also know it's one of those which is most commonly used by students. And when it comes to actually solving exercise problems, we know that many students look at the problem and decide very quickly on a solution and then continue with that solution irrespective of its success when actually the student will be much better served using a more structured and careful problem-solving process. What does this mean? It means that in order to be successful, students need to manage their own learning. And this is what's referred to as “metacognition.” Metacognition is a process of planning, monitoring, debugging and reviewing a student's learning.

Notes

Summary

13m 30s

Conclusion

- Six key findings from learning science research:
  - **Actively process** information
  - Deal with **implicit misconceptions**
  - Use **feedback**
  - Manage **cognitive load**
  - **Spaced repetitions** are valuable
  - **Metacognition** helps learning
- How to apply these principles in practice?

Some students will do this more or less naturally but other students may need help. The good news about metacognition is that it can be taught. If teachers encouraged students to plan effectively and to pay attention to what they're doing while they're learning, well students can actually learn plan and debug their own learning. So in conclusion over these two videos we've identified six different key ideas from the learning sciences research which can impact upon the quality of teaching and the quality of learning. First, students need to be actively processing information. If you as a teacher are doing more of the information processing than the students then that's possibly a problem and one that we need to look at. Secondly students don't just learn scientific models but actually they already have implicit naïve, non-scientific models of the physical world and we need to question these models before they can actively use the new models they're getting. Otherwise they'll learn these new models, they'll be able to solve mathematical problems with them, but they won't apply them to real life situations. Thirdly feedback is important. Students need to get feedback and students need help to learn from that feedback.

Notes

Summary

14m 55s

Conclusion

- Six key findings from learning science research:
  - **Actively process** information
  - Deal with **implicit misconceptions**
  - Use **feedback**
  - Manage **cognitive load**
  - **Spaced repetitions** are valuable
  - **Metacognition** helps learning
- How to apply these principles in practice?

Fourthly cognitive load is important. Students will struggle to remember things that an expert will find easy to recall. Students will struggle to deal with multiple pieces of information when, for the expert, it will look like it's straightforward. Spacing of practice is also important. Having all of the practice on a particular thing in a single day is likely to be less effective than spreading it over time. And finally students need help to plan, monitor, debug and evaluate their learning. If they can learn to do that, this can have a particularly positive impact on their learning over time. Okay. These are principles. These are findings from the research. But what does that mean in actual practice? How do these things impact upon what teachers should be doing with students in classrooms, in labs. That's exactly what we're going to look at over the next series of videos.

Notes

Summary

16m 16s