An Introduction to Learning

Lecture 8

Todd M. Gureckis, Lila Davachi
Department of Psychology
New York University
{todd.gureckis,lila.davachi}@nyu.edu
Agenda for Today

- Introduction to Cognitive or ‘Conscious’ Memory
  - History
  - Lessons from H.M.: there are different forms of memory!
  - Distinguishing cognitive memory?
  - Forgetting
  - Encoding
  - Context and Retrieval
1 History
Traditions diverge...

Pavlov

Learning

Ebbinghaus

Memory

Experimental Methods!
Traditions diverge...

Pavlov

Learning

Ebbinghaus

Memory

Experimental Methods!
Ebbinghaus’s Forgetting Curve

- Ebbinghaus (1885) studied memory for nonsense syllables
- Forgetting increases as time progresses...
2 Lessons from H.M.
What could he do?

Recall

Pattern-Analysis Skill

Knowing ‘how’ vs Knowing ‘what’
What could he do?

Visual Recall

A  

B  

C  Control
Cognitive memory (place learning) vs. procedural memory (response learning)
Memory Systems

Long-term memory

Declarative memory
- Episodic memory
- Semantic memory

Nondeclarative memory
- Procedural memory: skills, habits
- Priming
- Simple classical conditioning
- Habituation, sensitization
- Reflex pathways

Brain regions:
- Medial temporal lobe, diencephalon
- Basal ganglia
- Neocortex
- Amygdala, cerebellum
Memory Systems (updated)

- **Rapid encoding of flexible associations**
  - Hippocampus
  - Neocortex

- **Slow encoding of rigid associations**
  - Basal ganglia
  - Cerebellum
  - Neocortex

- **Rapid encoding of single or unitized items**
  - Parahippocampal gyrus
  - Neocortex

- Episodic memory
  - Procedural memory
    - Classical conditioning
    - Semantic memory
  - Familiarity
  - Priming

Henke, 2010
3 Distinguishing ‘cognitive memory’
• Conscious access to representations (semantic and/or episodic)

• Flexibility/inferential
Flexibility
(A) Constant start position training

(B) Novel probe tests

(C) Swim paths on probe tests
   Normal rats
   Rats with hippocampal damage

(D) Mean latency (seconds)
   - Normal
   - Hippocampal damage

Trial
Flexibility

Retrieval or Encoding?
Conscious access to representations

Flexibility/inferential
4 Forgetting
Forgetting

What are the reasons why we fail to remember things?

Trace decay?

Today we will focus on interference: Learning about one thing can mess up your memory for other things.

We will show that some kinds of forgetting that appear to be attributable to trace decay may, in fact, be caused by interference.
Interference

Two mechanisms for interference:

New memories might overwrite or damage old memories

Even if new & old memories co-exist peacefully, you might still end up getting competition at retrieval
Interference and competition

- **Competition Assumption** (Anderson, Bjork, & Bjork, 1994):
  - A cue activates all of its associates to some degree
  - The activated associates compete for access to awareness
  - **Competitors** are any associates other than the target memory

- Interference:
  - Increases with the number of competitors a target has
  - **cue-overload principle**:
    - The tendency for recall to decrease with the number of to-be-remembered items paired with the same cue

Adapted from Anderson and Neely (1996).
Proactive Interference

- Studied in the lab using the AB-AC interference paradigm
- Forgetting increases as a function of the number of previously studied lists
- Especially if the learning experiences are similar (e.g. share a cue)
Retroactive Interference

- Baddeley and Hitch (1977):
  - **Task:**
    - Rugby players were asked to recall the names of teams they played earlier in the season
  - **Control:**
    - Some players missed certain games, leaving a measure of forgetting due to decay as opposed to interference from intervening games
  - **Results:**
    - Time wasn’t a good predictor of forgetting
    - Forgetting increased with the number of intervening games; new learning can interfere with old memories

**Conclusion:** forgetting is caused by interference, not trace decay
How to avoid interference?

Shereshevskii: A case study in distinctive encoding

- Subject of “The Mind of a Mnemonist” by A. Luria

- Newspaper reporter with apparently unlimited capacity for memorization

- Experienced synesthesia: e.g., words evoke visual impressions, and sometimes sensations of taste and touch

- Synesthesia led to hyper-distinctive encoding

- Reduces interference

- Incredible memory for details but problems categorizing & generalizing
Mnemonics

Mnemonic techniques = techniques for improving recall

Almost all mnemonic systems work by fostering distinctive encoding

Peg-word mnemonic:

one is a bun
two is a shoe
three is a tree
four is a door
five is a hive
six is sticks
seven is heaven
eight is a gate
nine is wine
ten is a hen
Mnemonics

Peg-Word Mnemonic

one is a bun
two is a shoe
three is a tree
four is a door
five is a hive
six is sticks
seven is heaven
eight is a gate
nine is wine
ten is a hen
Mnemonics

Peg-Word Mnemonic

one is a bun                      donkey
two is a shoe                     cup
three is a tree                   elephant
four is a door                    stocking
five is a hive                    fire engine
six is sticks                     ring
seven is heaven                   caterpillar
eight is a gate                   saucepan
nine is wine                      rabbit
ten is a hen                       top hat
Mnemonics

Peg-Word Mnemonic

one is a bun
two is a shoe
three is a tree
four is a door
five is a hive
six is sticks
seven is heaven
eight is a gate
nine is wine
ten is a hen
Effective mnemonics require two things:

- Subjects need to form distinctive memory traces -- distinctiveness prevents the memory traces from getting confused with (or overwriting) one another

- Subjects need a way of finding these distinctive traces

- Peg-word gives us a way

  One is a bun
  Two is a shoe...
Availability and Accessibility

- Availability: Is the memory retrievable in principle?

- Accessibility: Assuming that the memory is retrievable in principle, can it be retrieved?

For the peg-word mnemonic:

- Distinctive encoding ensures that the memory is available

- Structured retrieval probes ensure that the memory can be accessed
Method of Loci

Come up with a familiar route, comprised of multiple places (loci)

For example: The route from this classroom to my office

Start in the classroom
Walk to the elevator
down the hall
Into my lab
Into my office

Stash mental images of to-be-remembered stimuli at different places on the route

Example: Remember grocery list of strawberries, milk, bread, chicken, grapes
Methods of memory masters

Memory champions can memorize random digit sequences that are hundreds of digits long. How?

One sample strategy:

Take the digit sequence & break it into 6 digit chunks

For each 6 digit chunk:

The first 2 digits correspond to a **person**
The second 2 digits correspond to an **action**
The third 2 digits correspond to an **object**

e.g., 235637
person #23 (e.g. Todd)
action #56 (e.g. jumping)
object #37 (e.g. turtle)
Methods of memory masters

- Next step: Method of loci

- Stash the images in different loci along the route (e.g. put image of Todd jumping on a turtle in the office..)

- To recall the digit sequence, restart the route
- Look in each location
- Read out the image that’s stored there
5 Encoding

- Attention
- Levels of Processing
- Organization
- Spacing
- Prior Knowledge
5 Encoding

- Attention
- Levels of Processing
- Organization
- Spacing
- Prior Knowledge
### Levels of Processing

- **Levels-of-Processing Hypothesis** (Craik & Lockhart, 1972):
  - Information can be processed on a variety of levels, from the most basic (visual form), to phonology, to the deepest level (contextual meaning).
  
  The depth of processing helps determine the durability in LTM.

<table>
<thead>
<tr>
<th>Level of Processing</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Visual Form</td>
<td>“DOG” comprises the letters D, O, and G</td>
</tr>
<tr>
<td>2) Phonology</td>
<td>Rhymes with FOG</td>
</tr>
<tr>
<td>3) Semantics (Meaning)</td>
<td>A four-legged pet that often chases cats and chews on bones</td>
</tr>
</tbody>
</table>
Levels of Processing
Craik and Tulving (1975)

**Task:**
- Participants viewed words and were asked to make three different types of judgments:
  - Visual processing (e.g. “Is LOG in upper case?”)
  - Phonological (e.g. “Does DOG rhyme with LOG?”)
  - Semantic (e.g. “Does DOG fit in the sentence: ‘The ___ chased the cat’?”)
- Finally, participants were asked to recognize the words they had seen before in a surprise test

**Results:**
- Words that were more deeply processed were more easily recognized
  - Particularly for questions with a “YES” response
Levels of Processing
Craik and Tulving (1975)

Task:
- Participants viewed words and were asked to make three different types of judgments:
  - Visual processing (e.g. “Is LOG in upper case?”)
  - Phonological (e.g. “Does DOG rhyme with LOG?”)
  - Semantic (e.g. “Does DOG fit in the sentence: ‘The ___ chased the cat’?”)
- Finally, participants were asked to recognize the words they had seen before in a surprise test

Results:
- Words that were more deeply processed were more easily recognized
- Particularly for questions with a “YES” response
The crucial role of retrieval cues

- Depth of processing / elaborative encoding is only part of the story

- A major determinant of memory success is how well the retrieval cues at test match the processing that subjects did at study...

- The better the match, the better subjects’ recall will be
Transfer-Appropriate Processing

The Transfer-Appropriate Processing Principle:
- Recall is better if the processing requirements of the test match the processing conditions at encoding

Morris, Bransford, and Franks (1977) tested the principle:
- Task:
  - Participants made either a phonological or semantic judgment about each item on a word list
  - The learning was incidental: participants didn’t know they would have to later recall the words
- The final test was either:
  - A standard recognition test for the learned words
  - A rhyming recognition test for learned words
Transfer-Appropriate Processing
Morris, Bransford, and Franks (1977)

Results:
- Standard recognition test: Deeper processing led to better performance
- Rhyming recognition test: Shallower rhyme-based encoding task led to better performance because it matched the testing situation

Conclusion:
- It only makes sense to talk about a learning method’s efficacy in the context of the type of final test
Interference: Brown-Peterson Paradigm

- simple paradigm for looking at interference effects
- present triplets of stimuli
- have people count backwards for some interval
- recall the most recently presented triplet
Interference: Brown-Peterson Paradigm

- 3 words will appear (e.g., apple orange banana)

- then you will see a 4-digit number (e.g., 2764)

- when the 4 digit number appears, start counting down by 7s from that number until I say stop
Brown-Peterson Demo

Pontiac Lexus Mazda
Brown-Peterson Demo

Pontiac Lexus Mazda
Brown-Peterson Demo

9853
Brown-Peterson Demo

Subaru Honda Volvo
Brown-Peterson Demo

Isuzu Buick Nissan
Brown-Peterson Demo

1533
Brown-Peterson Demo

Isuzu Buick Nissan
Brown-Peterson Demo

Volkswagen Pontiac Mitsubishi
Brown-Peterson Demo

2206
Brown-Peterson Demo

Volkswagen Pontiac Mitsubishi
Brown-Peterson Demo

Acura Ferrari Oldsmobile
Brown-Peterson Demo

Acura Ferrari Oldsmobile
Brown-Peterson Demo

Chevrolet Saab Cadillac
Brown-Peterson Demo

6612
Brown-Peterson Demo

Chevrolet Saab Cadillac
sparrow eagle bluejay
Brown-Peterson Demo

3486
Brown-Peterson Demo

sparrow eagle bluejay
Important BP Findings

- Improved performance with category switch
  - Suggests that proactive interference is contributing to forgetting

- Some studies have looked at within-subject delay manipulations: short delay vs long delay

Increasing delay between study & test hurts performance

  Evidence for decay....???

However, no effect of delay on the first trial

  Suggests that decay is not occurring
Important BP Findings

If decay isn’t contributing to forgetting, then why does increasing the retention interval hurt performance (on trials after the first one)?

Put another way: How can we get time-dependent forgetting in the absence of decay?
The constellation of thoughts in our heads drifts over time...
Items presented at a particular time get associated with active contextual elements.

If you study “printer” at 9:42, it gets associated with “hungry”, “work”, and “switch at 96th St”
If you want to figure out what you studied recently, cue memory with the current set of contextual elements.

If, at 9:43, you want to recall recently studied items, cue with “hungry” and “switch at 96th St.”
Because (earlier) you associated “printer” with “hungry” and “switch at 96th St”, cuing with “hungry” and “switch at 96th St” should trigger recall of “printer”
In the Brown-Peterson test, we have multiple study and test phases (S1, T1, S2, T2, etc.)

Each study list gets associated with the currently active contextual elements

e.g., S1 gets associated with “itchy”, “Work”, and “switch at 96th St”
Goal: you want to recall the most recent study list
T2: Cue with the current context “Work, hungry, 96th St”

There is a competition between S2 and S1:
S2 was associated with 3/3 current contextual elements
S1 was associated with 2/3 current contextual elements
... so S2 wins the competition
Now we can begin to address delay effects in the Brown-Peterson paradigm
This diagram illustrates the short-delay condition

S2 is recalled well because it matches the current context better than S1
In this diagram, T2 is a long-delay trial

If you wait long enough, most of the active contextual elements will change

In this case, both S1 and S2 are equally poor matches to the current context (1/3 features), so recall will be poor
Summary:
Items are associated with contextual features at study
We use the current context as a retrieval cue

According to the theory presented here, there is no effect of the passage of time *per se*: traces do not passively decay
Still, we end up getting better recall for recently presented items because the passage of time is correlated with contextual match.

Events that occurred more recently will have more contextual features in common with the retrieval cue, so their memory traces will be more strongly activated by the cue.
In the short-delay condition, contextual match to T2 is better for S2 vs. S1, so you get good recall
In the long-delay condition, contextual match to T2 is equivalently poor for S2 vs. S1, so you get poor recall.

Telephone poles analogy (Crowder, 1976)
The fact that delay does not affect recall on the first trial can be explained in terms of the idea that recall is a competitive process. Because memory is a competitive process, what matters most is the relative contextual match between S1 and T, rather than the absolute level of match.
Increasing delay reduces the **absolute** amount of study-test contextual match (in this case, from 3 features to 2).

However, **relative** to other traces in your brain (relating to other things, besides Brown-Peterson) S1 is still by and far away the best-matching trace, so recall should be good.
- You can get effects of temporal delay without a decay process, if you assume that contextual drift occurs with time.

- However, we also still know that post-encoding processes that enhance memory stabilization DO occur (consolidation)....
Overall Summary

- Reasons for forgetting over time:
  - **Interference**: Learning about one thing can mess up your memory for other, similar things
  - **Cue Overload Principle**: The more competitors attached to a cue, the worse recall for any one item becomes
  - **Proactive Interference**: Previously encoded memories can disrupt retrieval of recently learned ones
  - **Retroactive Interference**: New learning can impair retrieval of older memories, either by *structurally damaging* these older memories or by *competing with* these older memories at retrieval
  - **Poor match between retrieval cues (at test) at processing at study**
    - **Contextual fluctuation**: Subjects’ mental context is constantly changing; if your mental context now is very different from the way it was at study, this will hurt recall.
Factors that lead to good memory

- Elaborative/distinctive encoding, which reduces interference and ensures that memory traces are available for recall at test.
- Good match between retrieval cues and stored traces
  - The mnemonic techniques that we discussed use a structured retrieval framework to ensure that subjects can find (access) the memory traces that they formed at study...
5 Context and Retrieval
Memory Search

• Free recall

etc…

…delay…

Cartoon by Sean Polyn
Reinstating Context

96th St

itchy

exams

hungry

9:40 9:41 9:42 9:43
Reinstating Context

- 96th St
- itchy
- exams
- hungry
- shark
- computer
- bat
- test
Initiate recall by cuing with the current context: “hungry, 96th St”

Given this cue, you end up recalling “bat”

You also recall **other contextual elements** associated with bat: “exams”
Reinstating Context

96th St

itchy

exams

hungry

9:40 9:41 9:42 9:43

shark computer bat test
Step 2: Take retrieved contextual elements and incorporate them into your retrieval cue: “hungry, exams, 96th St”

With “exams” in your retrieval cue, you can now recall “computer”, plus a new contextual element “itchy”
Step 3: Incorporate “itchy” in your retrieval cue

Now you can recall “shark”

Using **retrieved context** as a retrieval cue allows you to bootstrap your way backwards in time...
Contextual Reinstatement

- What is the empirical evidence for contextual reinstatement?
Conditional Response Probability

1st     2nd     3rd     4th     5th     6th     7th

notebook  skull  leaf  watch  shark  bat  computer
Conditional Response Probability

1st 2nd 3rd 4th 5th 6th 7th

notebook skull leaf watch shark bat computer
Conditional Response Probability

1st  2nd  3rd  4th  5th  6th  7th

notebook  skull  leaf  watch  shark  bat  computer

Given that you just recalled the 4th item on the study list, what are the odds that the next item you recall will be the 3rd, or 5th, or 2nd, or 6th...?
Conditional Response Probability

lag: -3 -2 -1 0 +1 +2 +3
1st 2nd 3rd 4th 5th 6th 7th

notebook skull leaf watch shark bat computer

(Graph from Kahana, 1996)
Conditional Response Probability

key finding:
Successive recalls come from nearby serial positions, in both the forward and backward directions

(Graph from Kahana, 1996)
• Can this pattern be explained in terms of conjoint rehearsal?
Conditional Response Probability

- Can this pattern be explained in terms of conjoint rehearsal?
  - No, it is found even when a long, arithmetic-filled delay is placed between items at study
  - So what is going on?
Reinstating Context

96th St

itchy
exams
hungry
weather

9:40  9:41  9:42  9:43

notebook  skull  leaf  watch  shark  bat  computer
Reinstating Context

96th St

itchy

exams

hungry

weather

9:40 9:41 9:42 9:43

notebook skull leaf watch shark bat computer
Reinstating Context

96th St

itchy

exams

hungry

weather

9:40

9:41

9:42

9:43

notebook skull leaf watch shark bat computer
• Recency effect: Better recall of items from the end of the list

• Hypothesis: This occurs because **contextual match** is better for items from the end of the list
Recency Judgments

96th St

itchy

exams

hungry

weather

9:40  9:41  9:42  9:43

notebook  skull  leaf  watch  shark  bat  computer  test

A slightly different task: judgments of recency
Which was presented more recently: shark or skull?

Use the words to cue for contextual info
shark retrieves “hungry, weather, exams, 96th St”
skull retrieves “exams, itchy, 96th St”

**compare retrieved context to current context:**
“weather, 96th St”
shark: “hungry, weather, exams, 96th St”
skull: “exams, itchy, 96th St”
current: “weather, 96th St”

**Contextual match** is higher for shark, so shark is probably more recent...
• Rats can be trained to do recency judgments
• Recording brain data from rats can give us a window into the role that contextual drift plays in recency judgments...

Figure 8  Sequential order and recognition tasks. (a) On each trial the animal was presented with a series of five odors (e.g., odors A through E). The animal was then either probed for its memory of the order of the items in the series (top) or its memory of the items presented (bottom). +, rewarded odor; −, nonrewarded odor. (b) Hippocampal animals were impaired on all sequential order probes. Performances on different probes are grouped according to the lag (number of intervening elements). (c) Hippocampal animals performed as well as controls on the recognition probes. ‘X’ designates a randomly selected odor that was not presented in the series and used as the alternative choice. *, p < .05. (a–c) Adapted from Fortin NJ, Agster KL, and Eichenbaum H (2002) Critical role of the hippocampus in memory for sequences of events. Nat. Neurosci. 5: 458–462, with permission from Macmillan Publishers Ltd.
Show rats a series of odors

Train rats to perform recency judgments

Record multi-unit activity from region CA1 of the hippocampus
Treat this neural recording data as the animal’s “mental context”

Measure how much the neural activity pattern drifts during the encoding phase

Use this to predict accuracy

Intuitively: The more that mental context changes between item presentations, the more temporally discriminable the items will be
Here, context changes sharply between shark and computer.

In this situation, the shark vs. computer recency judgment is easy.
Here, context does not change at all between shark and computer

In this situation, the shark vs. computer recency judgment is hard
Distance related to memory

• The experiment works!

• Increased “neural drift” between items predicts increased accuracy on the recency judgment

• Measuring neural drift gives us a window into how much the animal’s mental context is drifting

• Which allows us to predict memory performance
Cognitive memory is our everyday memory for facts and events - unlike other forms of memory, it is declarative, accessible to consciousness and can be expressed flexibly.

Cognitive memory exists across species and depends on the same brain systems.

Encoding can be bolstered by imagery which improves distinctiveness, prior knowledge which increases associative binding but also contextual match.

Retention of memories diminishes according to the Power Law. But there is compelling evidence that interference, not decay, is the primary element which determines whether a memory will be remembered or forgotten.

Retrieval is facilitated by the appropriate retrieval cues - and the more of them. Context reinstatement facilitates retrieval.

Hippocampus! (communication with cortex)
Key Principals for the Semester

- Learning and memory are closely related and intertwined states of information processing.

- Major insights about learning and memory have come from studies of the brain.

- The concept of multiple memory systems unifies the study of learning and memory.

- The underlying bases of learning and memory are the same in humans and animals.

- Our theoretical approaches to studying learning are always closely tied to technological advances that are unfolding in general society (e.g., today - machine learning).
The End