Architectures of Intelligence

Lecture 5

Declarative Memory I

university of groningen

Today

Today

- Declarative memory: basics
  - The rational basis for declarative memory
  - Declarative memory in ACT-R
  - Example model

Today

- Declarative memory: basics
  - The rational basis for declarative memory
  - Declarative memory in ACT-R
  - Example model

- Assignment

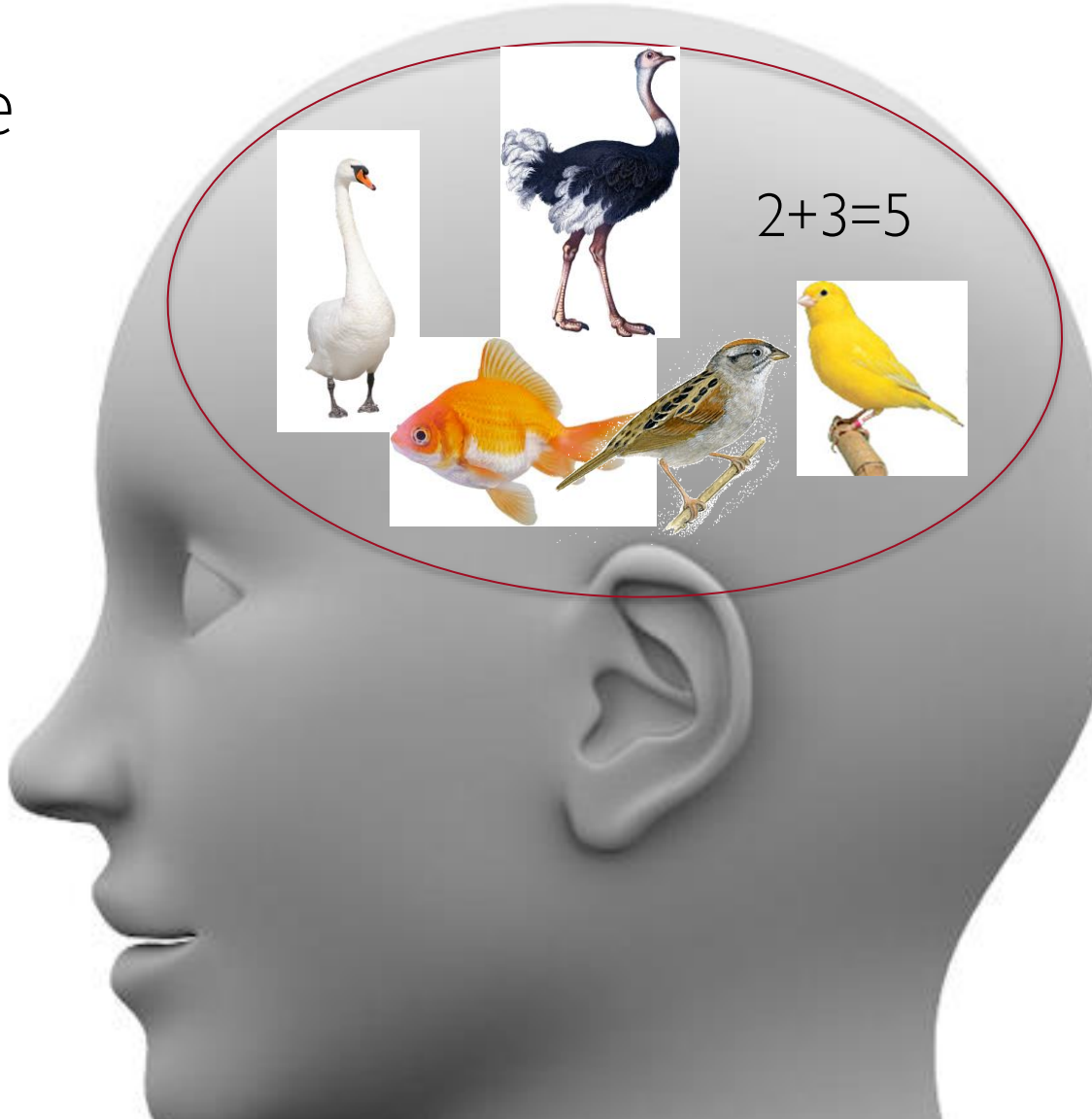
Declarative Memory: Rational Analysis

Declarative Memory

- Memory for facts, both semantic and episodic, e.g.:
  - 3+2=5
  - the past tense of go is went
  - Yesterday I had pizza for dinner

Our memory is full of facts

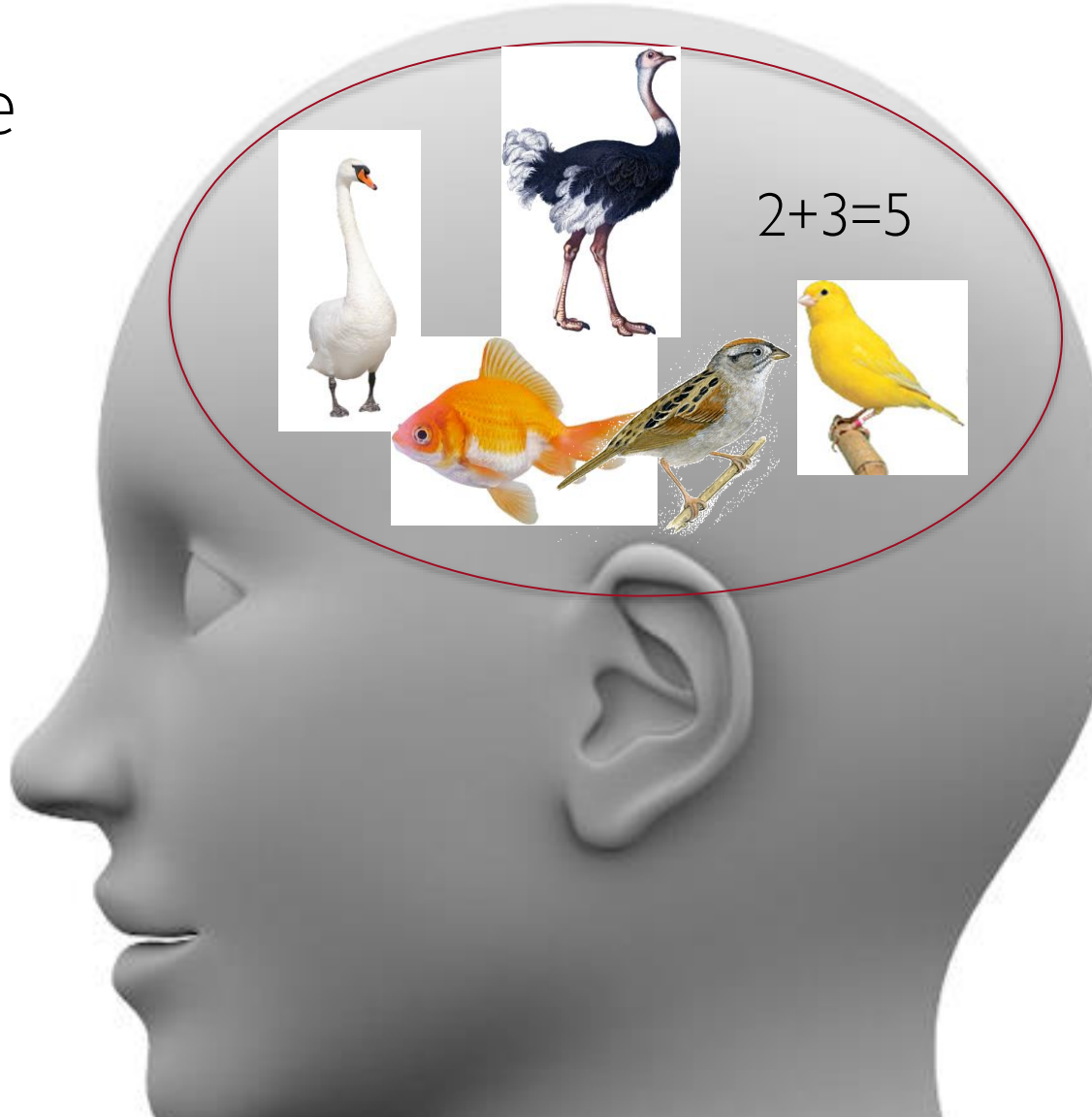
How can we retrieve
what we need,
when we need it?



Our memory is full of facts

How can we retrieve
what we need,
when we need it?

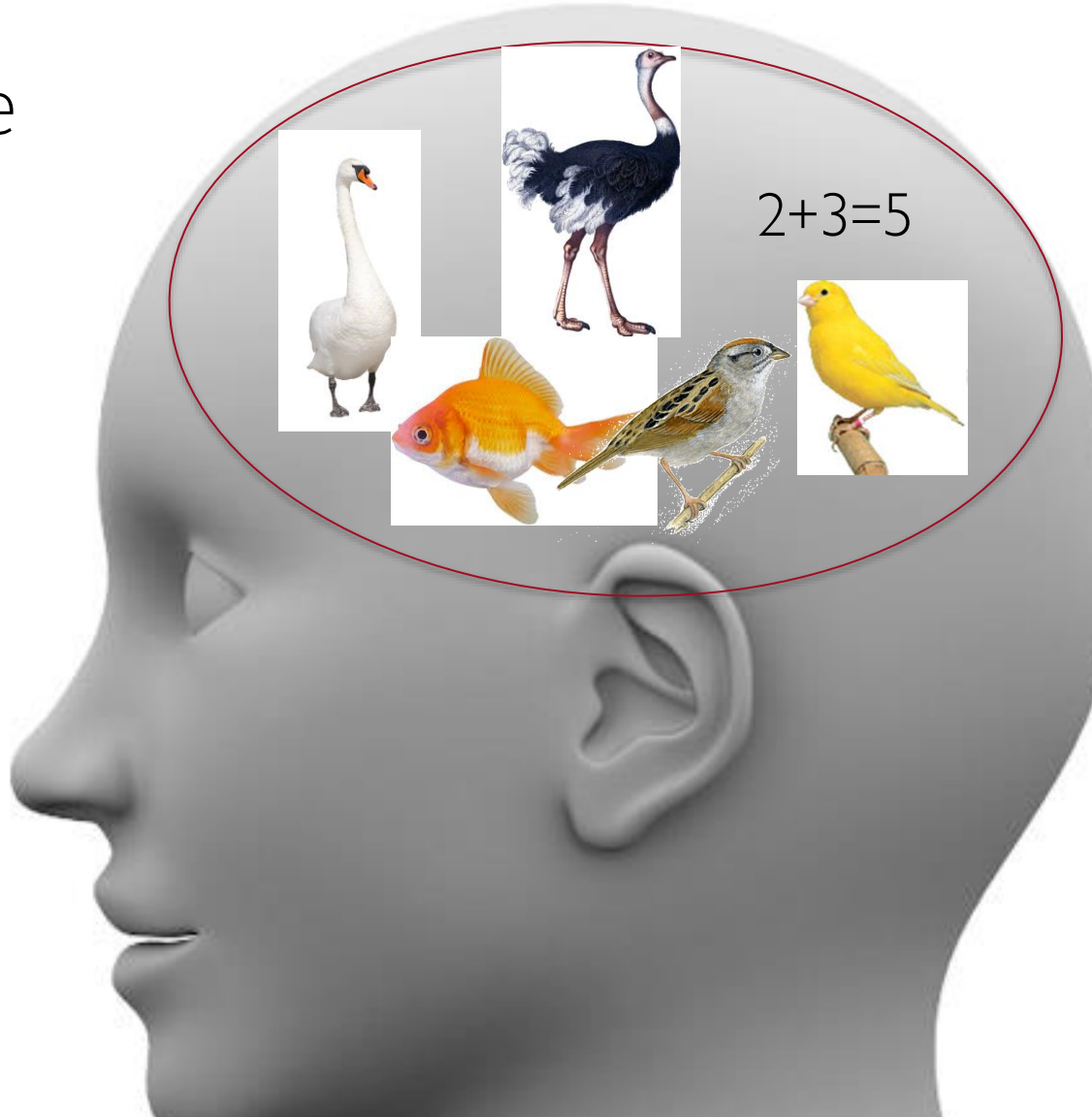
- What is available?



Our memory is full of facts

How can we retrieve
what we need,
when we need it?

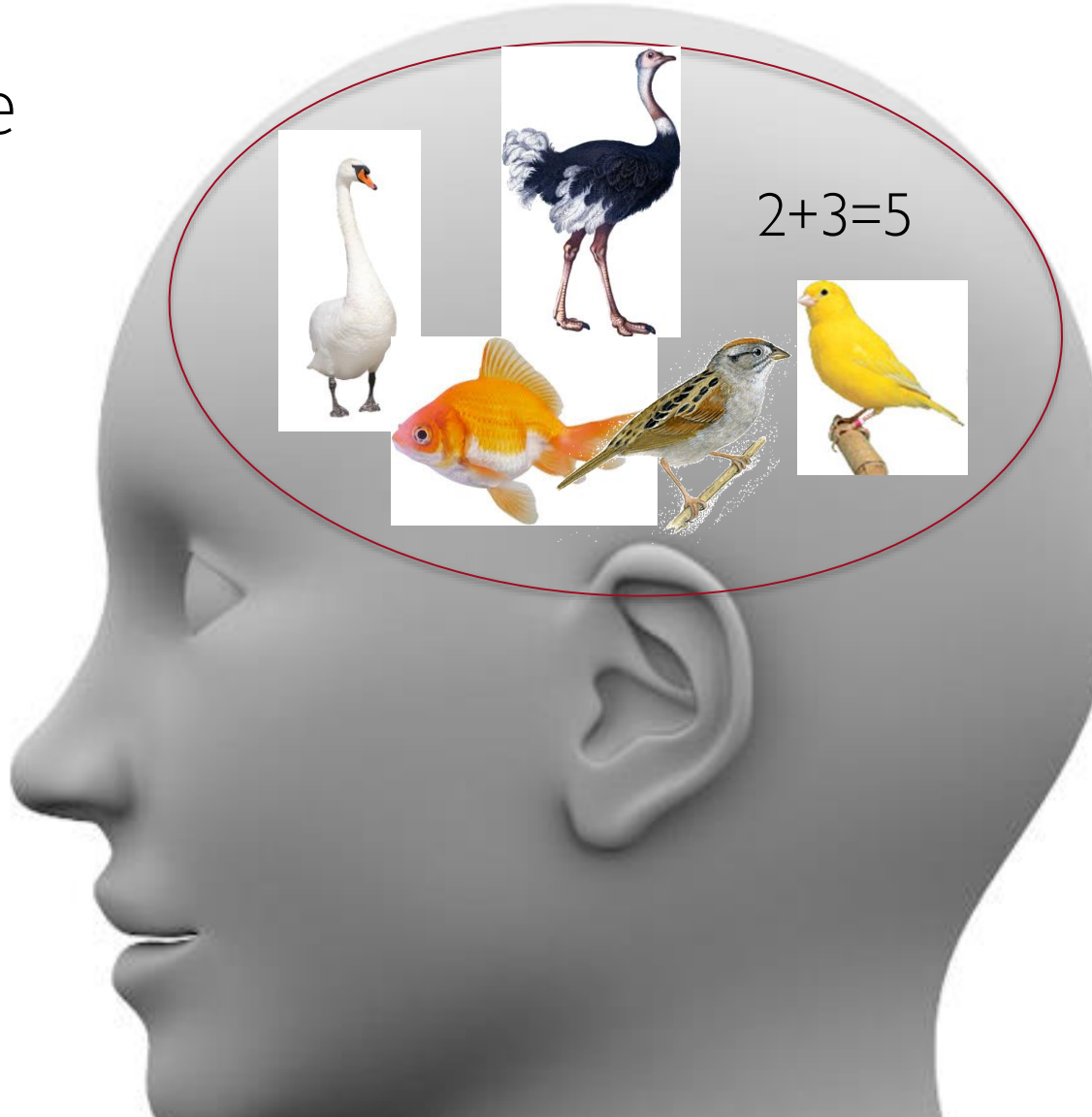
- What is available?
- Under what circumstances?



Our memory is full of facts

How can we retrieve what we need, when we need it?

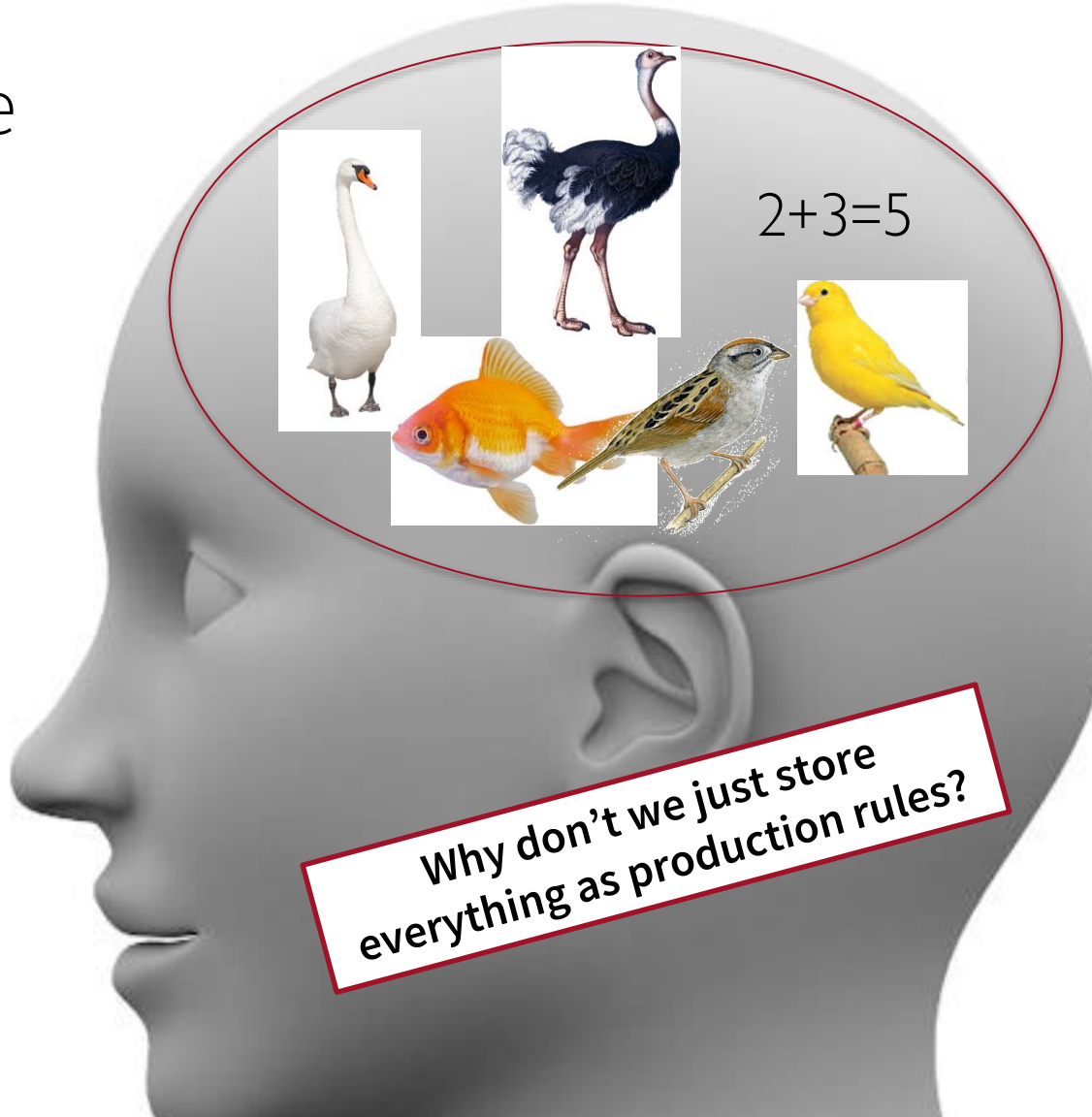
- What is available?
- Under what circumstances?
- When/how do you forget something?



Our memory is full of facts

How can we retrieve
what we need,
when we need it?

- What is available?
- Under what circumstances?
- When/how do you forget something?



“the memory system [. . .] makes most available those memories most likely to be needed”

(Anderson, 2007, p. 109)

"the memory system [. . .] makes most available those memories most likely to be needed"

(Anderson, 2007, p. 109)

How does it know which memories are most likely to be needed?

The rational basis for declarative memory

The rational basis for declarative memory

Suppose you are a librarian who has to move his library to a smaller building and toss away 50% of the books

The rational basis for declarative memory

Which of these two books would you keep?

Bring this book back before:

14 Jun 90

21 Feb 91

01 Mar 91

25 Sep 09

05 Jan 21

Bring this book back before:

04 Mar 20

10 Apr 20

05 May 20

21 Jun 20

05 Sep 20

21 Nov 20

The rational basis for declarative memory

Which of these two books would you keep?

Bring this book back before:

14 Jun 90

21 Feb 91

01 Mar 91

25 Sep 09

05 Jan 21

Frequency
Recency
Spacing

Bring this book back before:

Mar 20

Apr 20

ay 20

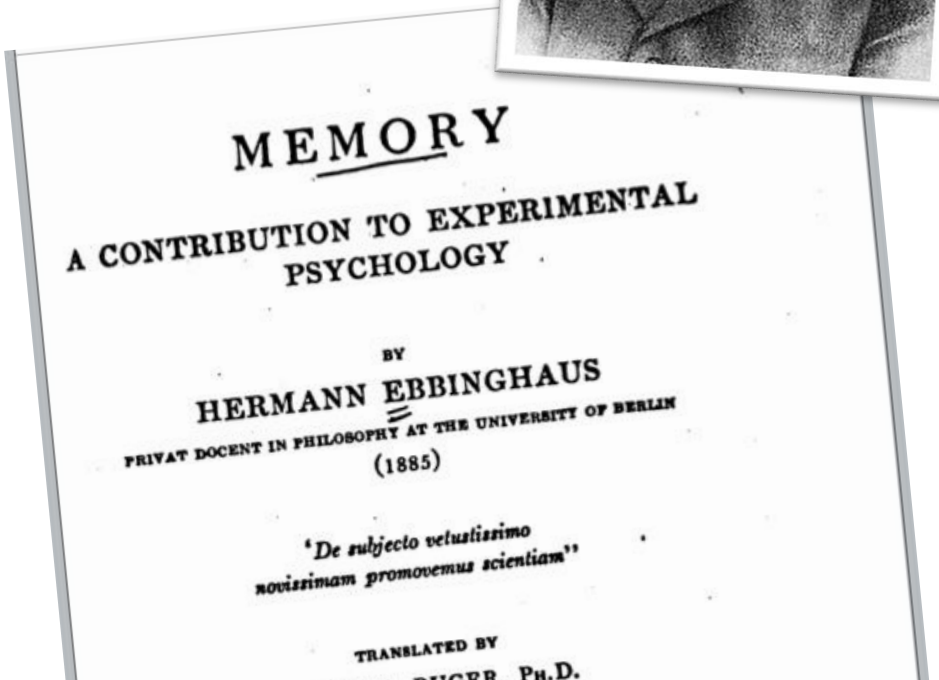
Jun 20

05 Sep 20

21 Nov 20

Availability of information in memory

Ebbinghaus (1850-1909)



MEMORY

A CONTRIBUTION TO EXPERIMENTAL PSYCHOLOGY

BY

HERMANN EBBINGHAUS

PRIVAT DOCENT IN PHILOSOPHY AT THE UNIVERSITY OF BERLIN

(1885)

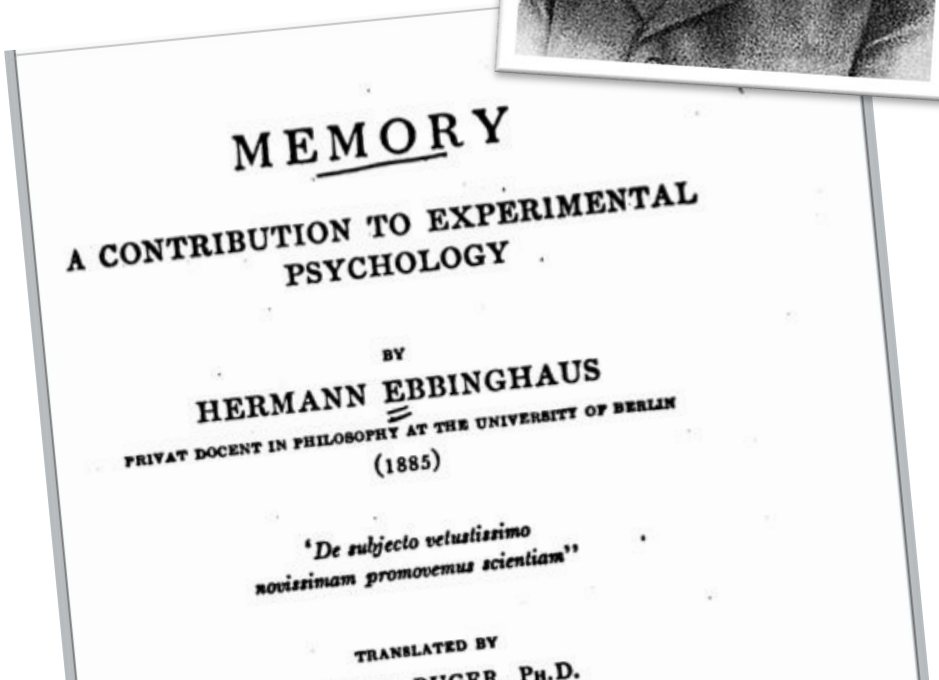
'De subjecto vetustissimo novissimam promovemus scientiam''

TRANSLATED BY

Availability of information in memory

Ebbinghaus (1850-1909)

- Studied memory with lists of "nonsense syllables"
  - BOK
  - YAT
  - KOJ
  - …



MEMORY

A CONTRIBUTION TO EXPERIMENTAL PSYCHOLOGY

BY

HERMANN EBBINGHAUS

PRIVAT DOCENT IN PHILOSOPHY AT THE UNIVERSITY OF BERLIN

(1885)

"De subjecto vetustissimo novissimam promovemus scientiam"

TRANSLATED BY

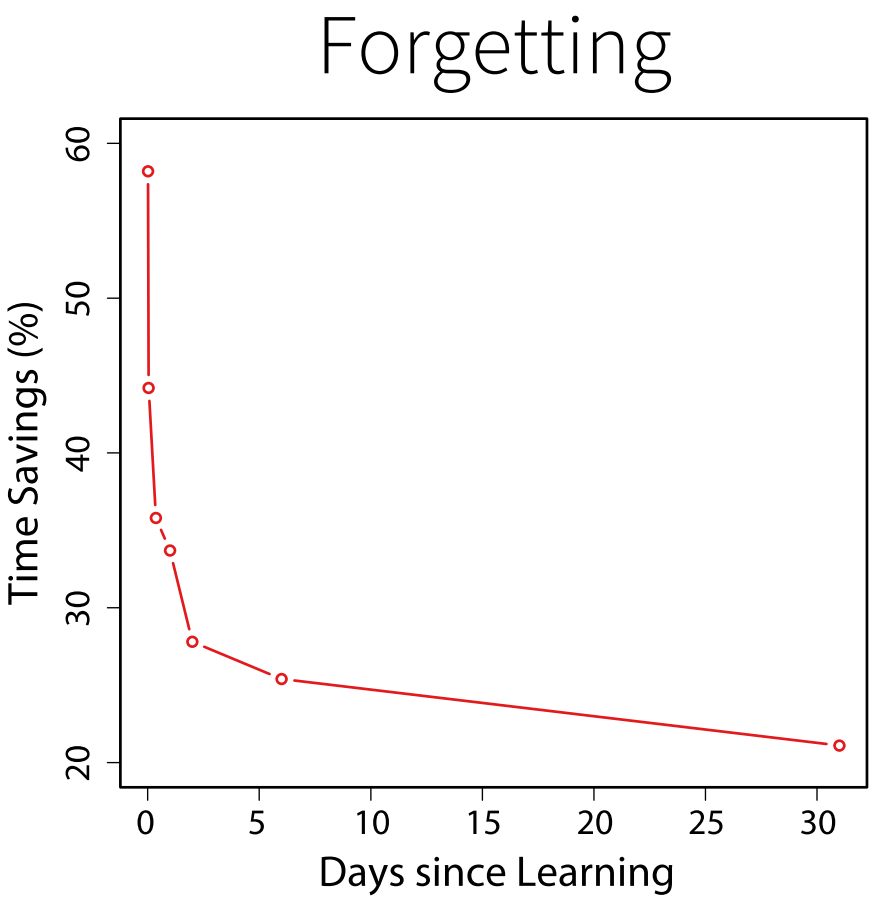
Availability of information in memory

Ebbinghaus (1850-1909)

Ebbinghaus (1885/1913)

Availability of information in memory

Ebbinghaus (1850-1909)

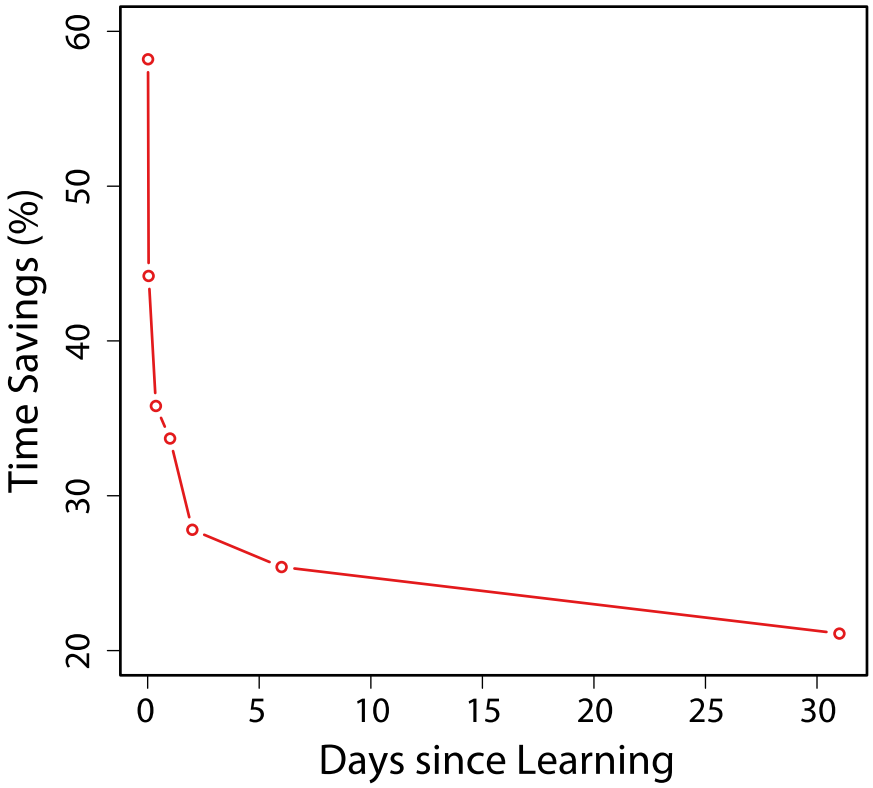


Ebbinghaus (1885/1913)

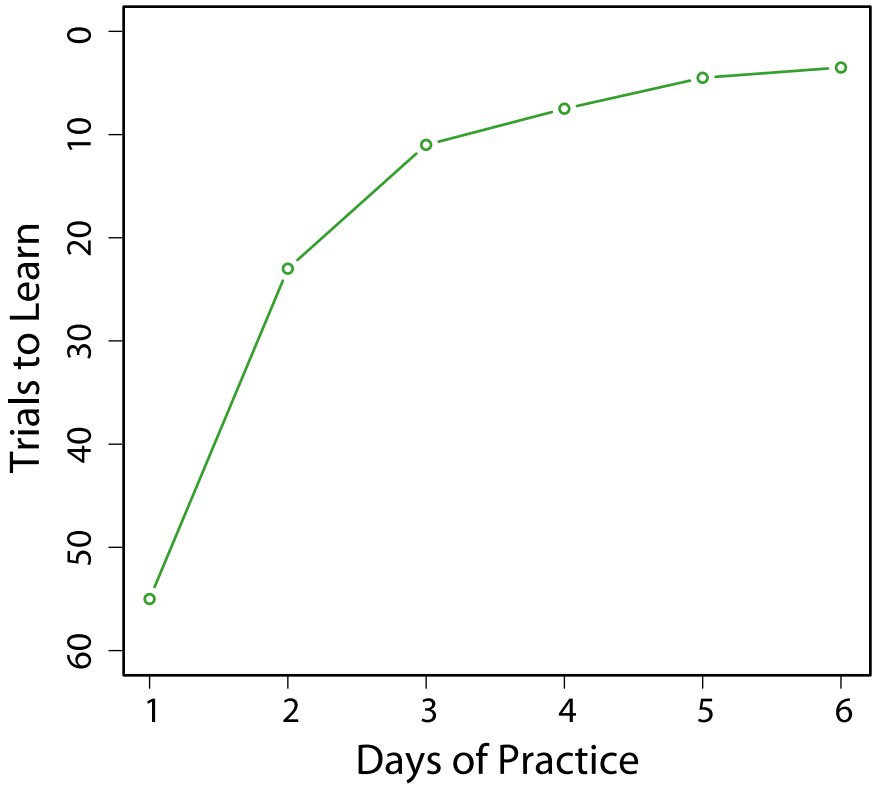
Availability of information in memory

Ebbinghaus (1850-1909)

Forgetting



Practice

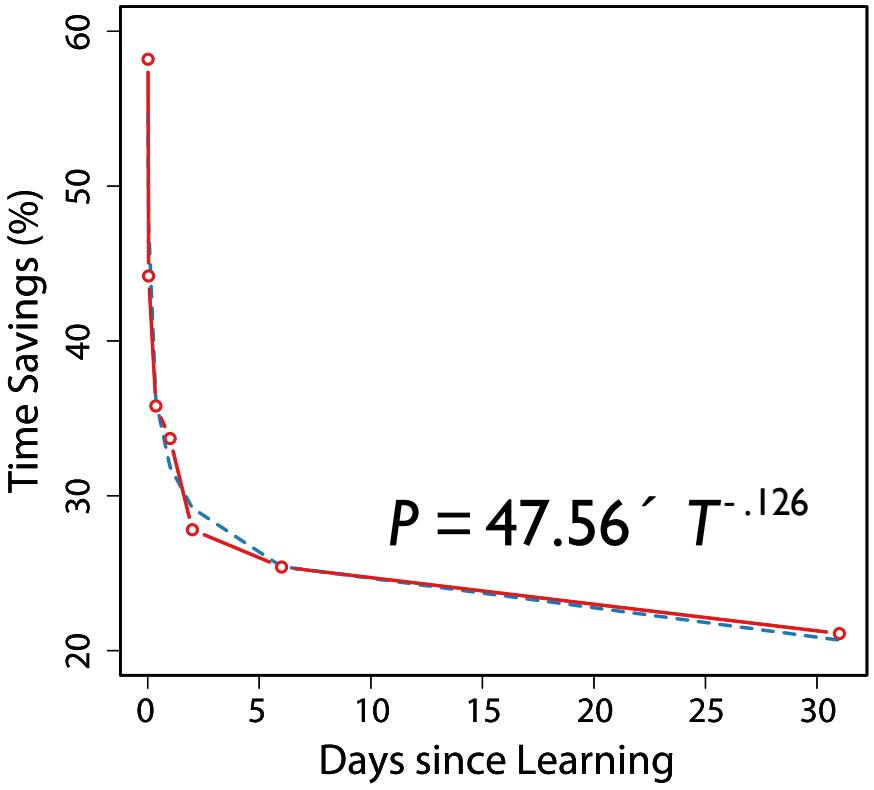


Ebbinghaus (1885/1913)

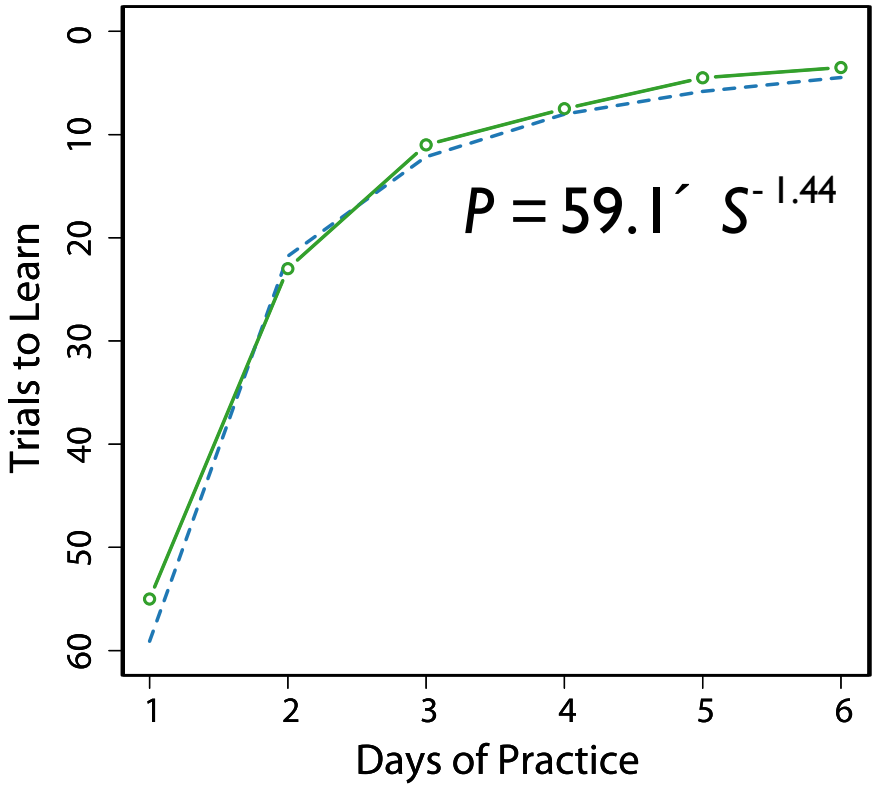
Availability of information in memory

Ebbinghaus (1850-1909)

Forgetting



Practice



Anderson & Schooler (1991)

Ebbinghaus (1885/1913)

Availability of information in the world

PSYCHOLOGICAL SCIENCE

Research Article

REFLECTIONS OF THE ENVIRONMENT IN MEMORY

John R. Anderson and Lael J. Schooler
Department of Psychology, Carnegie Mellon University

Abstract—*Availability of human memories for specific items shows reliable relationships to frequency, recency, and pattern of prior exposures to the item. These relationships have defied a systematic theoretical treatment. A number of environmental sources* (New York Times, *parental speech, electronic mail) are examined to show that the probability that a memory will be needed also shows reliable relationships to frequency, recency, and pattern of prior exposures. Moreover, the environmental relationships are the same as the memory relationships. It is argued that human memory has the form it does because it is adapted to these environmental relationships. Models for both the environment and human memory are described. Among the memory phenomena addressed are the practice function, the retention function, the effect of spacing of practice, and the relationship between degree of practice and retention.*

The title of our paper is inspired by the following remark in Shepard (1990): "We may look into that window on the mind as through a glass darkly, but what we are beginning to discern there looks very much like a reflection of the world" (p. 213). ... how the principles of perception are ... in which

current day? Memory would be behaving optimally if it made this memory less available than memories that were more likely to be used but made it more available than less likely memories.

In this paper we examine a number of environmental sources to determine how probability of a memory being needed varies with pattern of past use. However, we first review how availability in human memory varies with pattern of past use. Some aspects of this problem have been extensively studied in empirical studies of human memory.

FORM OF THE MEMORY FUNCTIONS

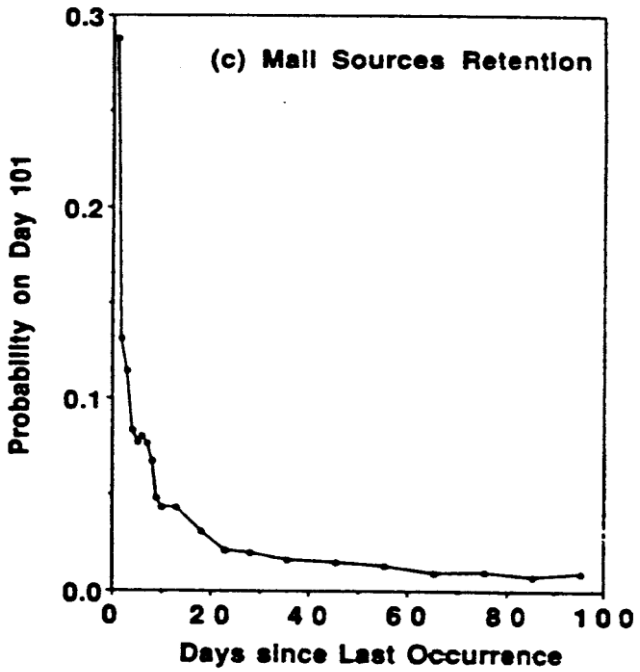
Two of the most basic statistics we might gather about pattern of past use are how often a memory has been practiced and how long it has been since it was last practiced. Learning functions and retention functions to describe these two aspects of human memory have been collected since the original experiments of Ebbinghaus (1885/1964). Figure 1 shows the retention function and practice function obtained by Ebbinghaus.

The Retention Function

... in terms of the percent sav-

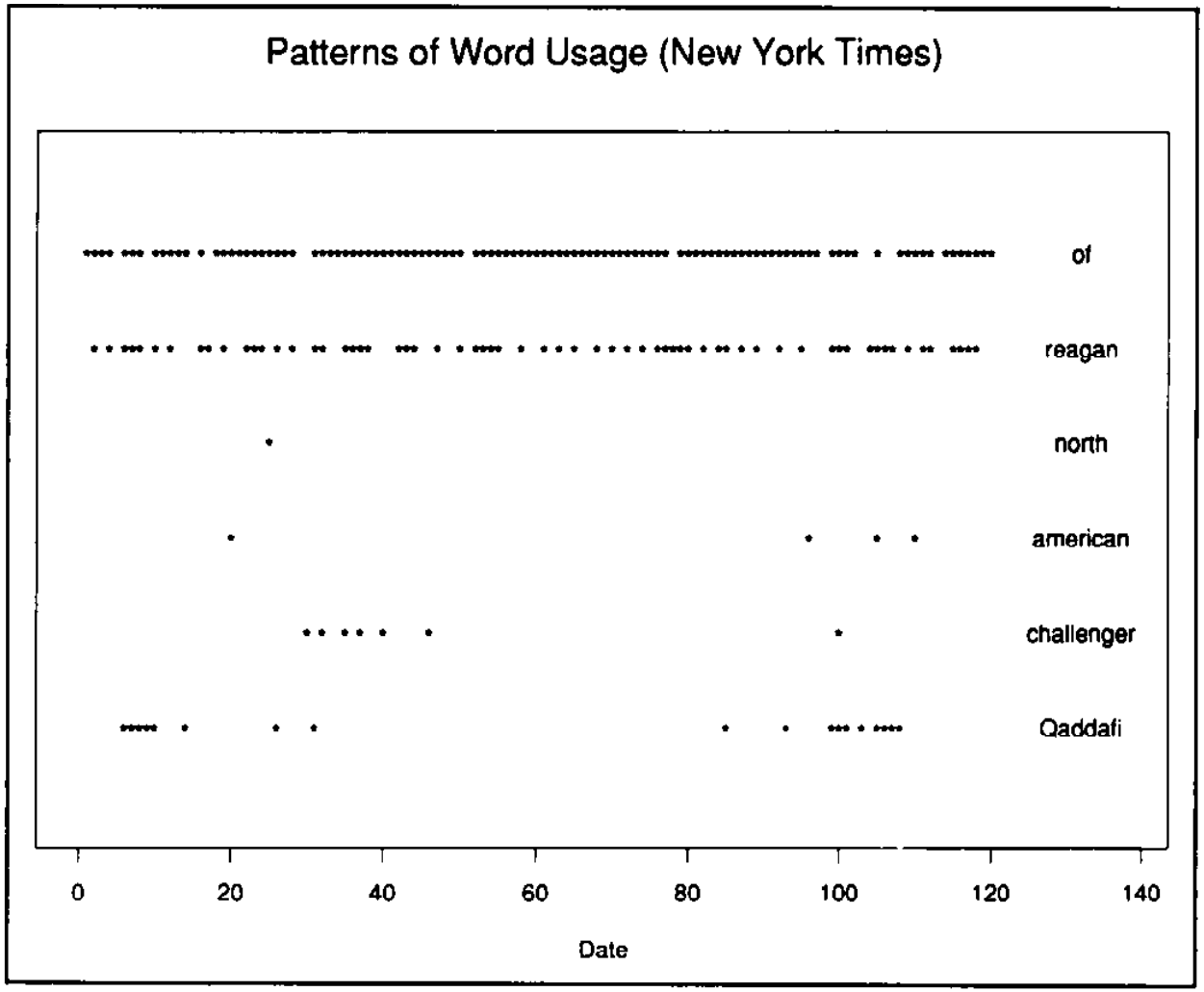
Availability of information in the world

Probability that you will send me an email today as a function of how long ago you last sent me an email

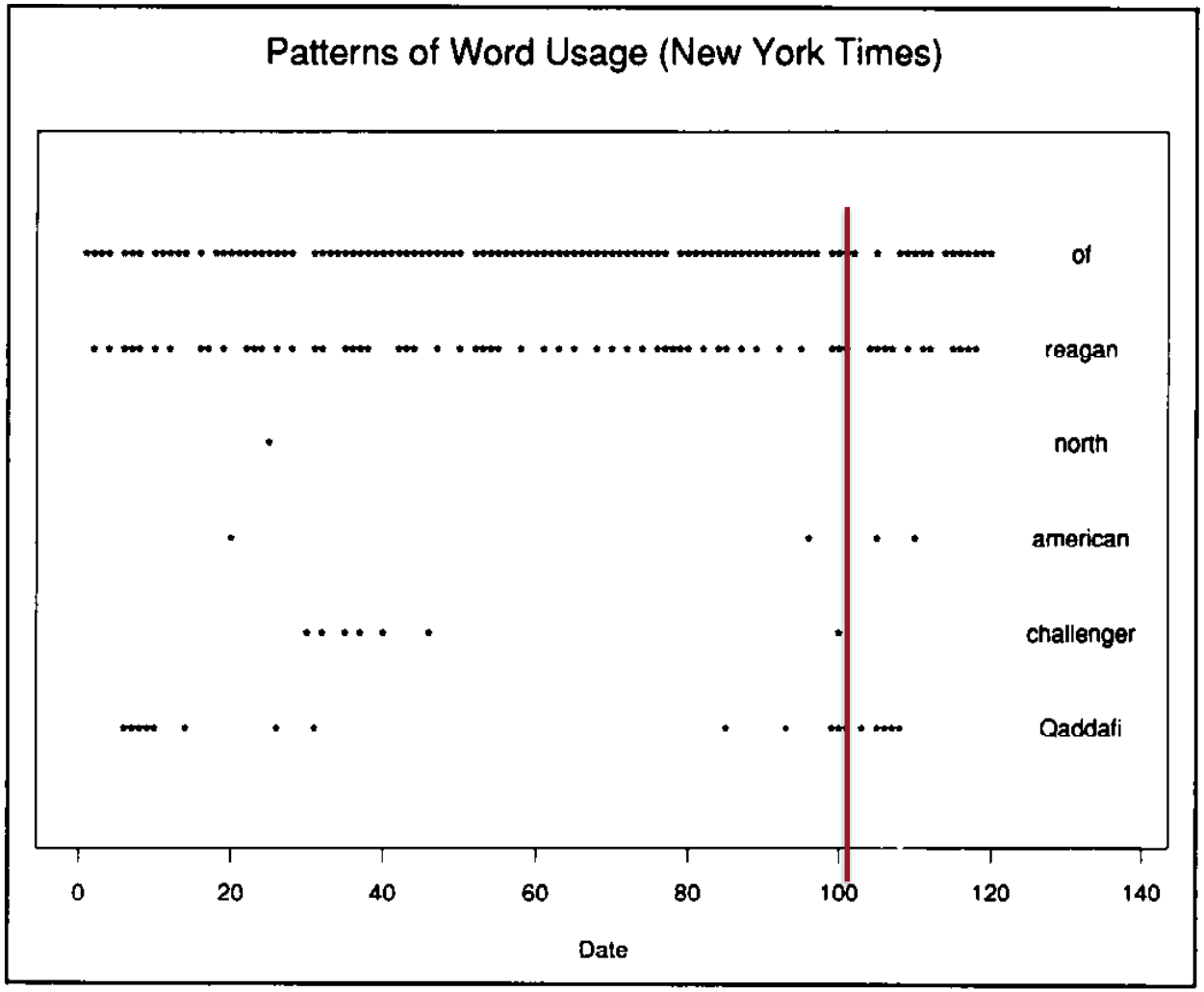


How likely is it that a word appears in the headlines of the New York times?

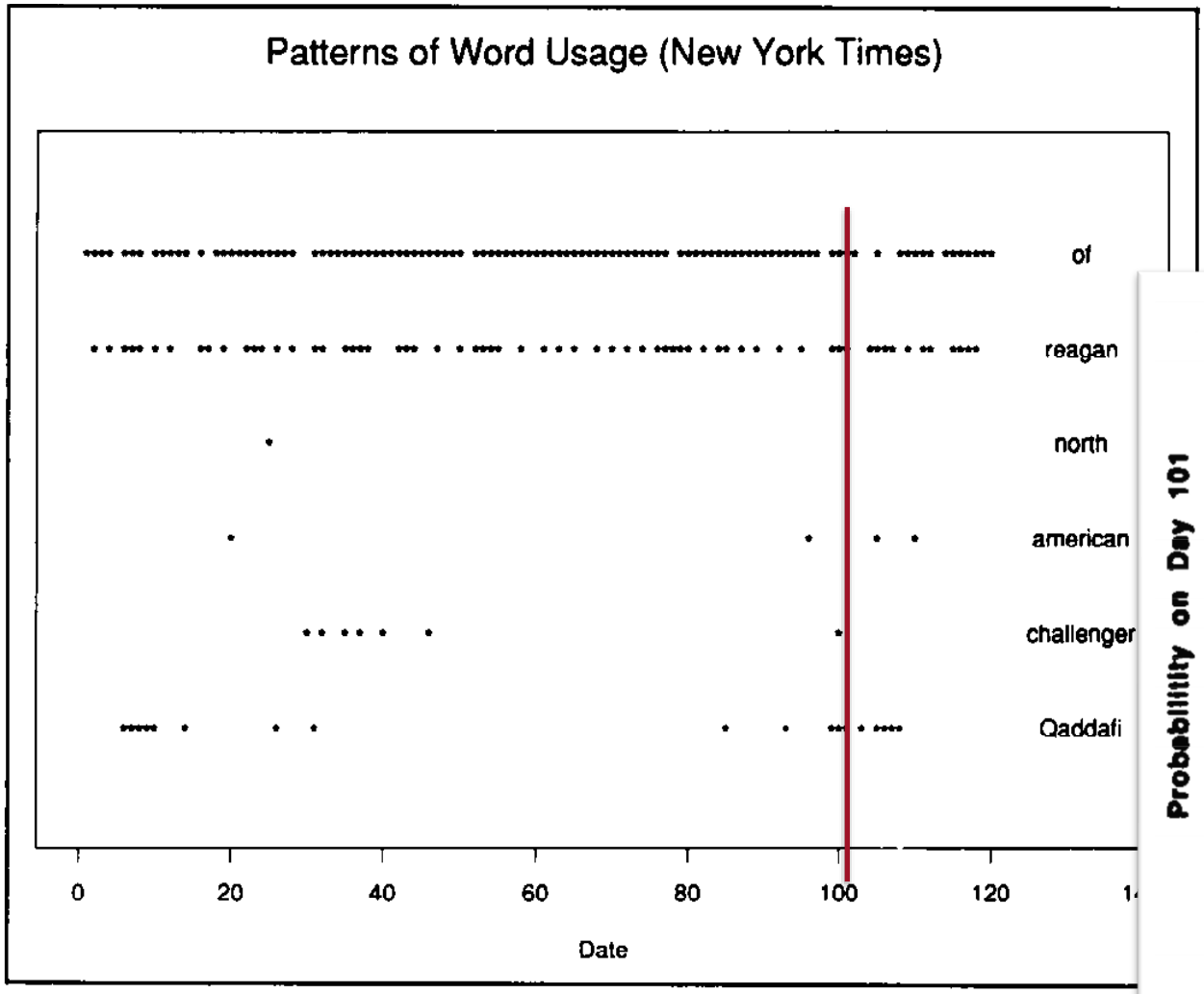
How likely is it that a word appears in the headlines of the New York times?

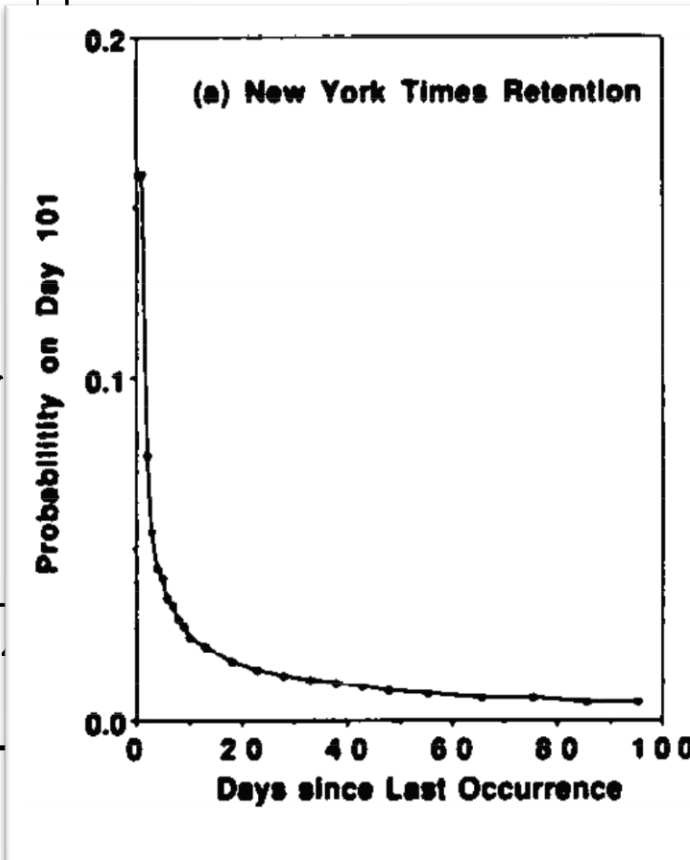


How likely is it that a word appears in the headlines of the New York times?

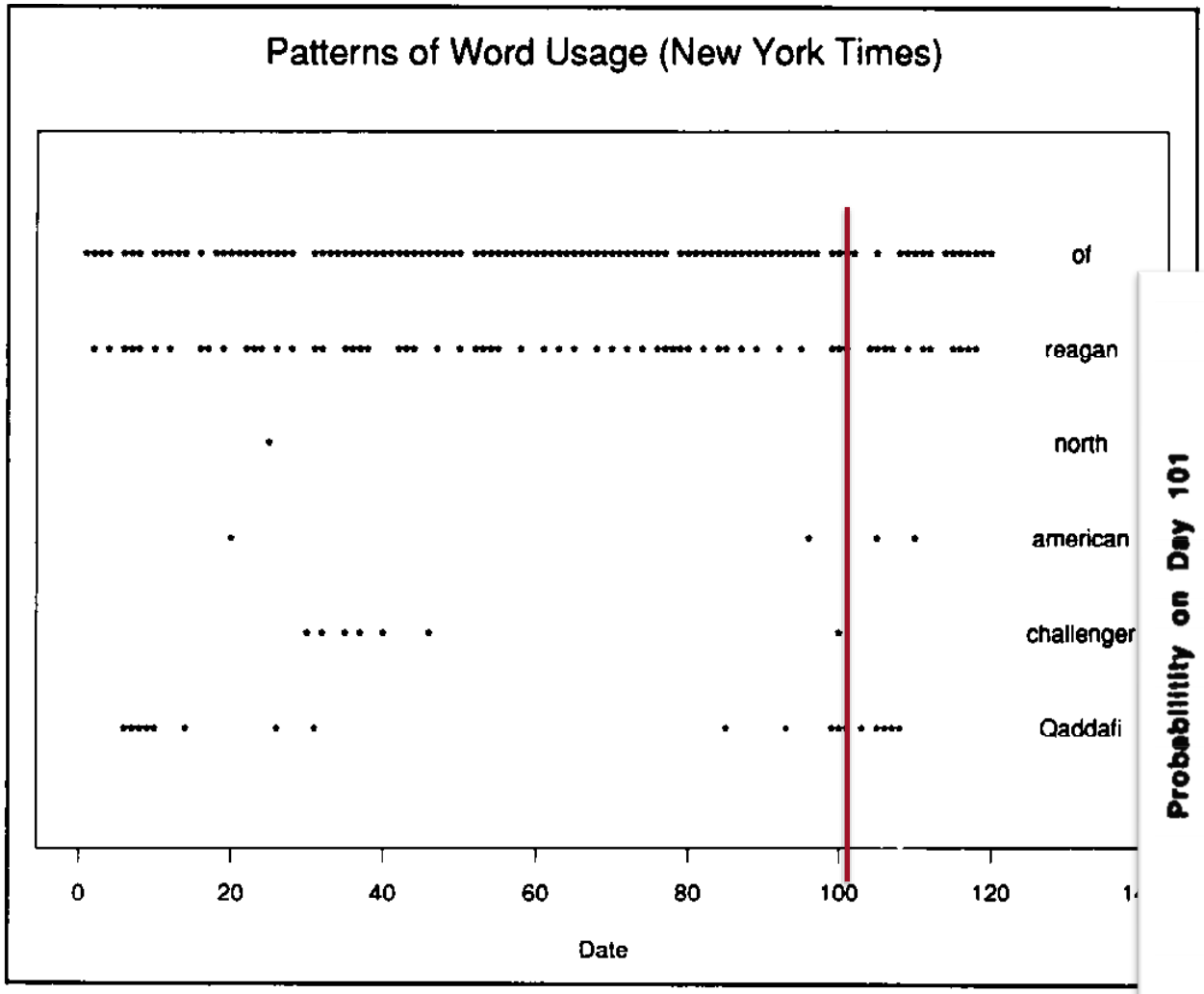


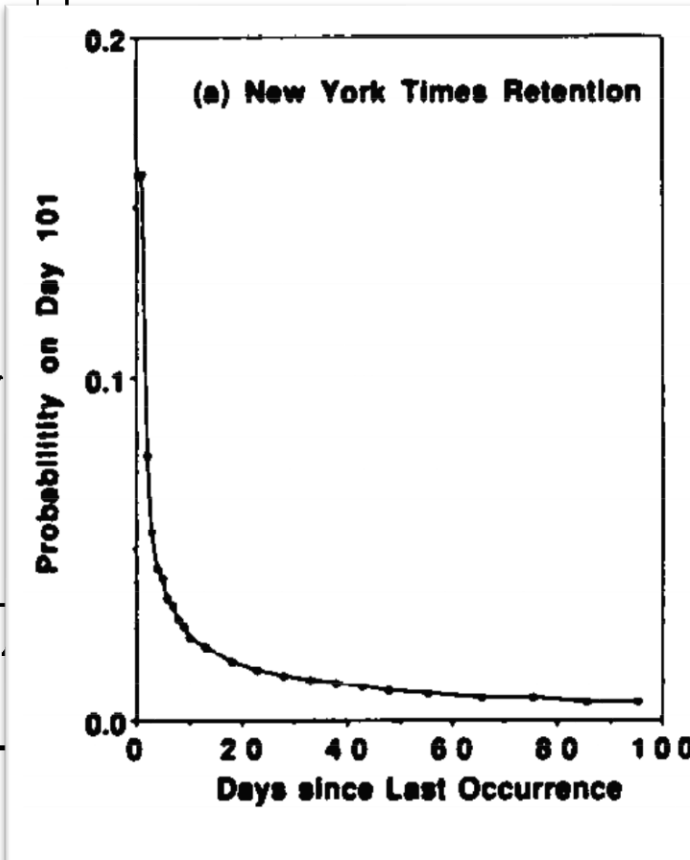
How likely is it that a word appears in the headlines of the New York times?

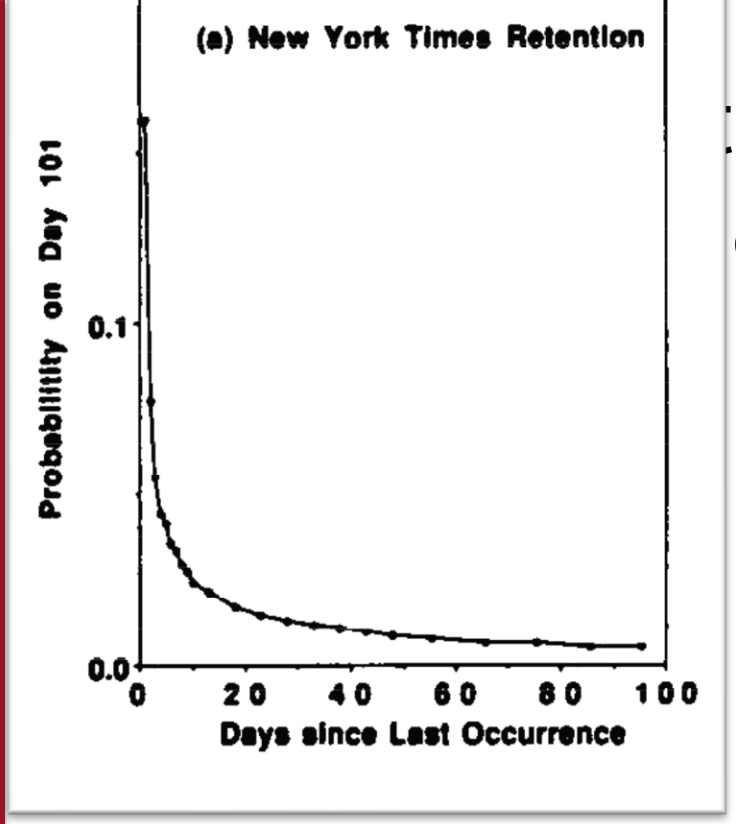




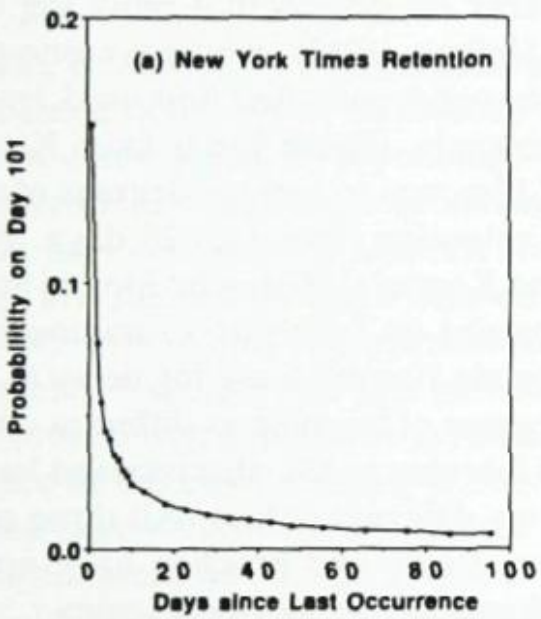
How likely is it that a word appears in the headlines of the New York times?

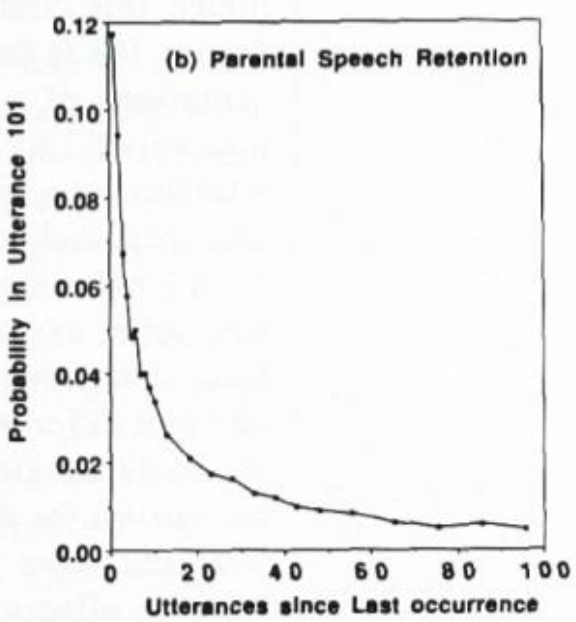


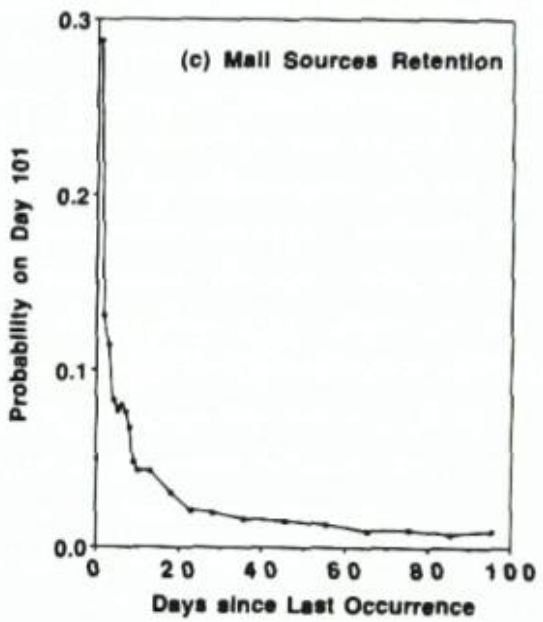


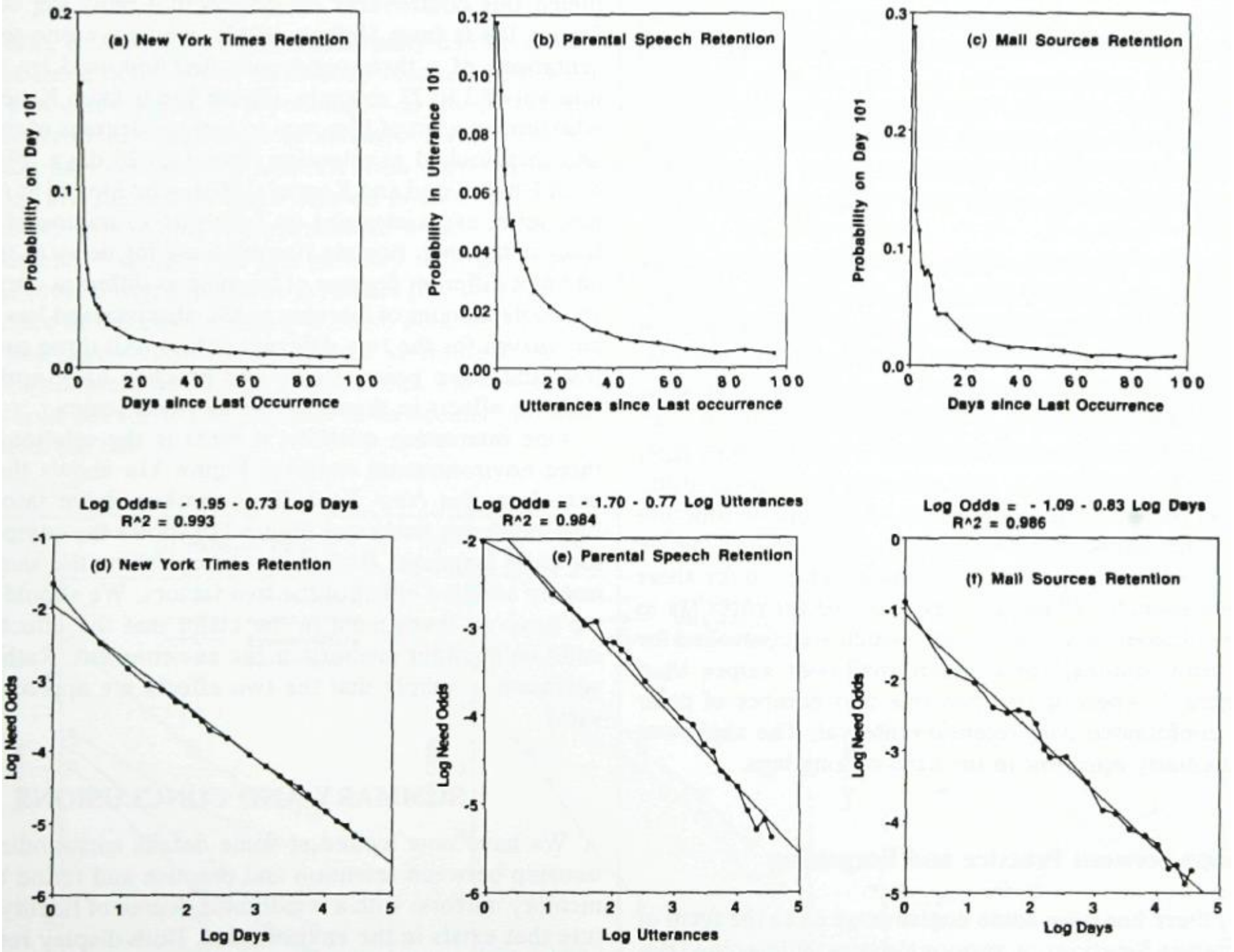


that a word appears in the
of the New York times?

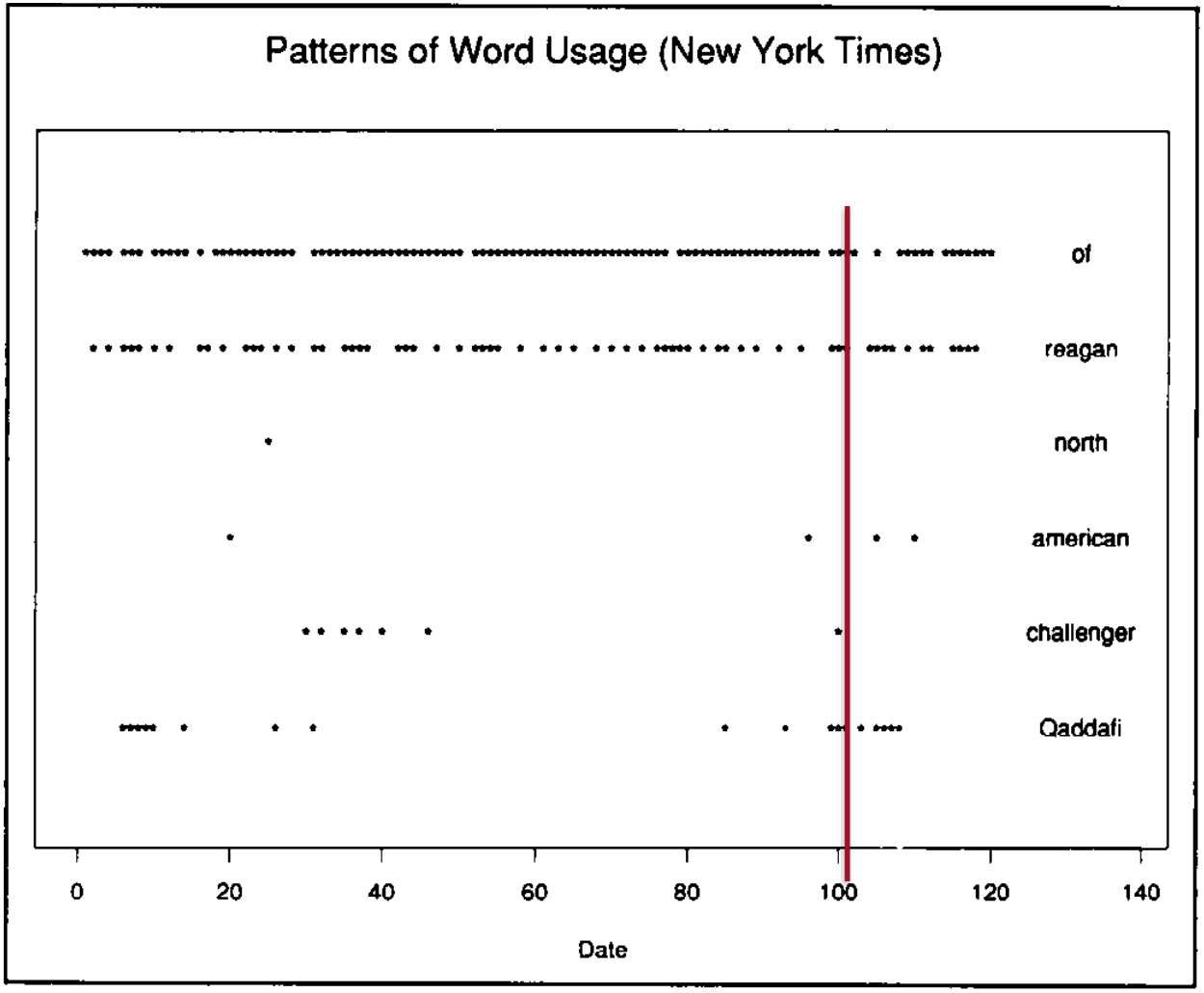

(a) New York Times Retention
Probability on Day 101
0.2
0.1
0.0
0
20
40
60
80
100
Days since Last Occurrence


(b) Parental Speech Retention
Probability in Utterance 101
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
20
40
60
80
100
Utterances since Last occurrence

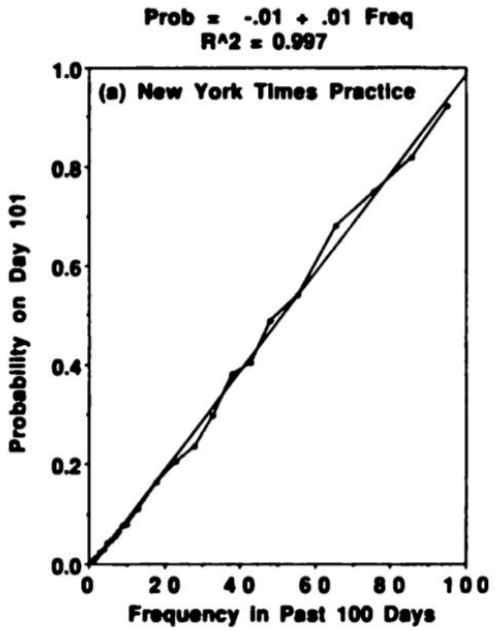

(c) Mail Sources Retention
Probability on Day 101
0.3
0.2
0.1
0.0
0
20
40
60
80
100
Days since Last Occurrence

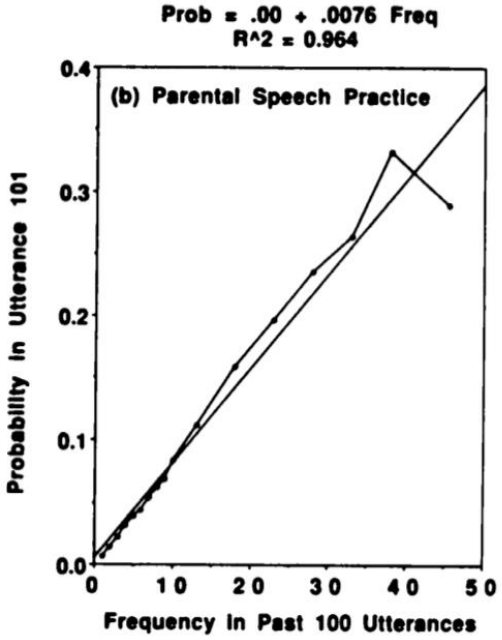

(a) New York Times Retention
Probability on Day 101
Days since Last Occurrence
(b) Parental Speech Retention
Probability in Utterance 101
Utterances since Last occurrence
(c) Mail Sources Retention
Probability on Day 101
Days since Last Occurrence
Log Odds= - 1.95 - 0.73 Log Days
R^2 = 0.993
(d) New York Times Retention
Log Need Odds
Log Days
Log Odds = - 1.70 - 0.77 Log Utterances
R^2 = 0.984
(e) Parental Speech Retention
Log Need Odds
Log Utterances
Log Odds = - 1.09 - 0.83 Log Days
R^2 = 0.986
(f) Mail Sources Retention
Log Need Odds
Log Days

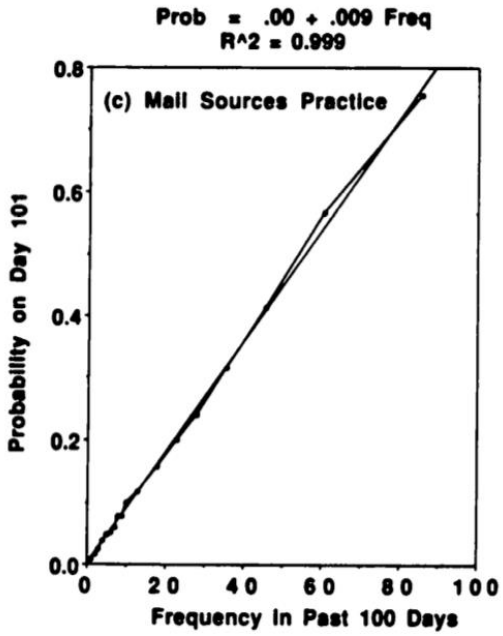
Frequency


Patterns of Word Usage (New York Times)
of
reagan
north
american
challenger
Qaddafi
0
20
40
60
80
100
120
140
Date

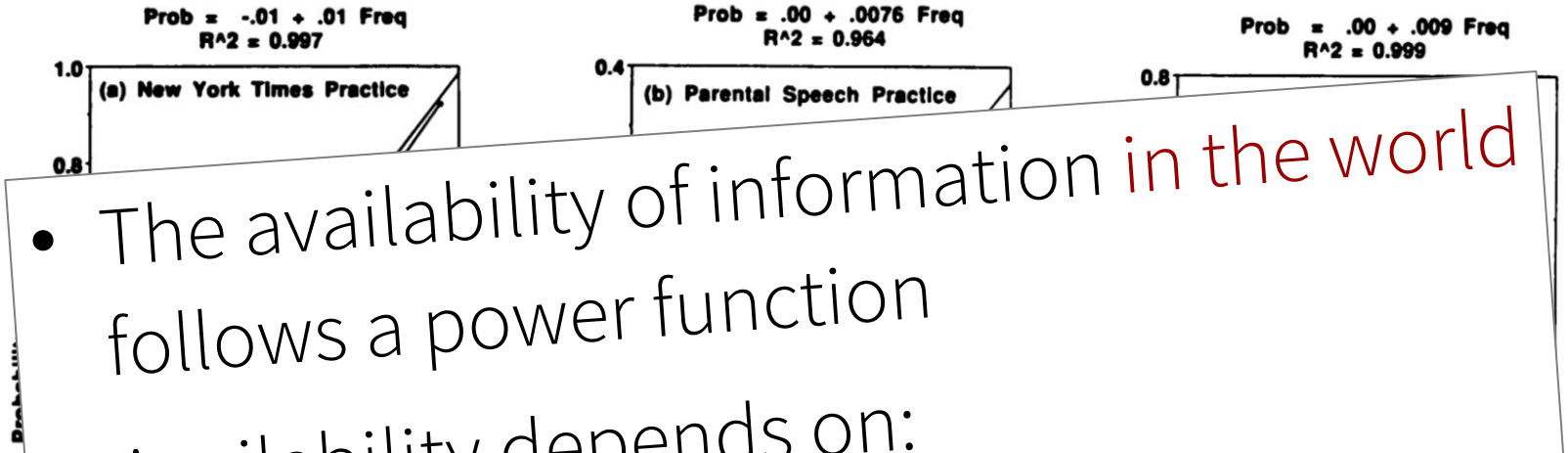
Frequency


Prob = -.01 + .01 Freq
R^2 = 0.997
(a) New York Times Practice
Probability on Day 101
Frequency in Past 100 Days
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100


Prob = .00 + .0076 Freq
R^2 = 0.964
(b) Parental Speech Practice
Probability in Utterance 101
Frequency in Past 100 Utterances
0.0
0.1
0.2
0.3
0.4
0
10
20
30
40
50


Prob = .00 + .009 Freq
R^2 = 0.999
(c) Mail Sources Practice
Probability on Day 101
Frequency in Past 100 Days
0.0
0.2
0.4
0.6
0.8
0
20
40
60
80
100

Frequency



- The availability of information in the world follows a power function
- Availability depends on:
  - Recency: power-law drop-off
  - Frequency: encounters are additive

Summary

Summary

- The availability of information **in memory** follows **a power function** (Ebbinghaus)

Summary

- The availability of information **in memory** follows **a power function** (Ebbinghaus)
- The availability of information **in the world** follows **a power function** (Anderson & Schooler)

Summary

- The availability of information **in memory** follows **a power function** (Ebbinghaus)
- The availability of information **in the world** follows **a power function** (Anderson & Schooler)

➢ Anderson & Schooler imply that this is a causal relationship: **our memory is optimally suited for the environment it lives in**

Summary

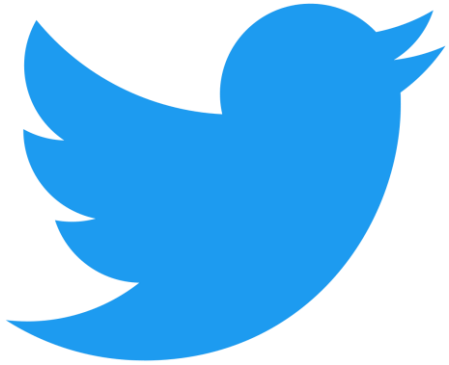
- The availability of information **in memory** follows **a power function** (Ebbinghaus)
- The availability of information **in the world** follows **a power function** (Anderson & Schooler)

➢ Anderson & Schooler imply that this is a causal relationship: **our memory is optimally suited for the environment it lives in**

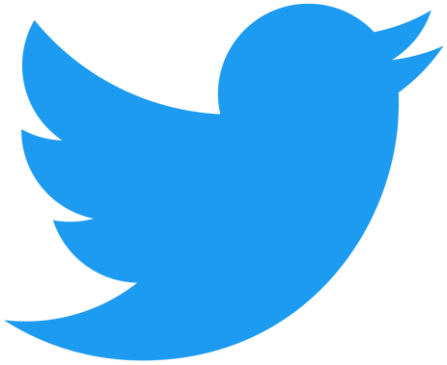
- Availability in both cases depends on the **recency & frequency** of prior experiences

2022 update

2022 update



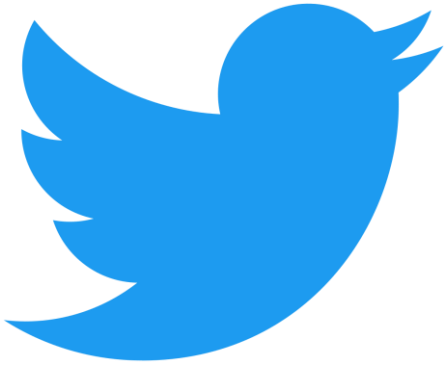
2022 update



- Tweets between 2007-2014 from top 500 tweeters

=> 1,038,632 tweets with on average 7.5 unique strings

2022 update

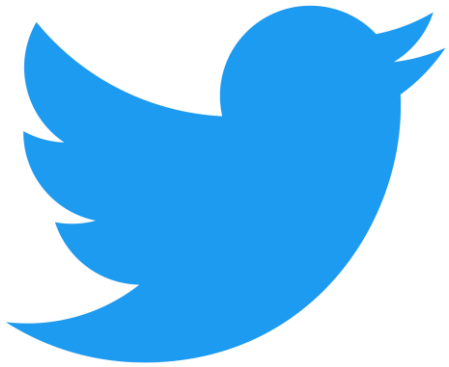


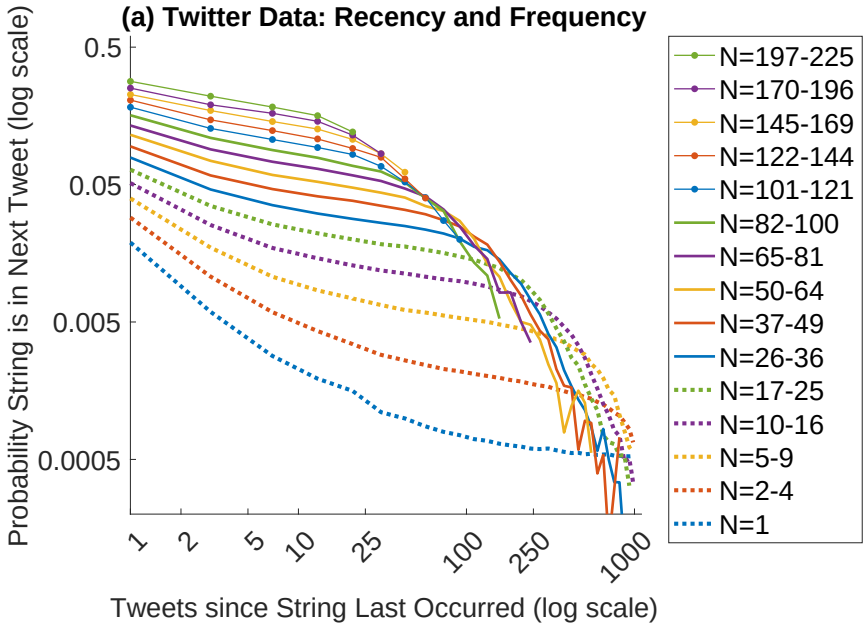
- Tweets between 2007-2014 from top 500 tweeters

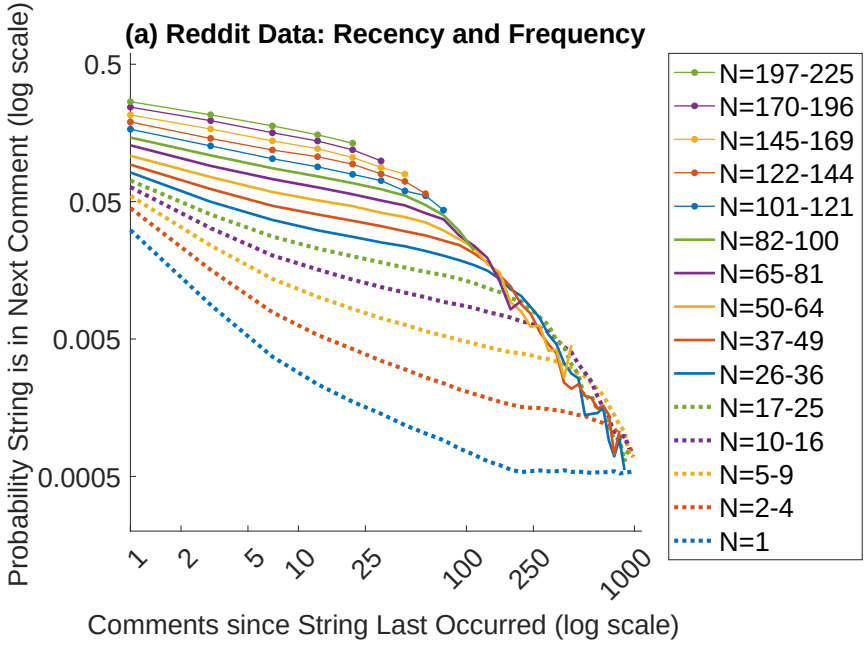
=> 1,038,632 tweets with on average 7.5 unique strings

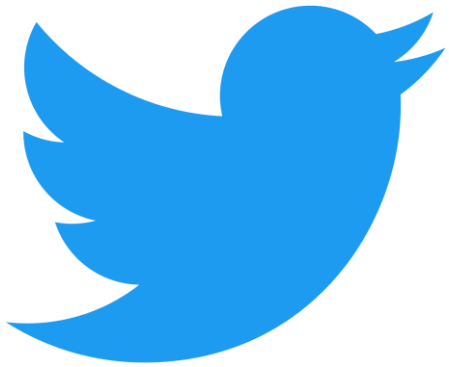
- 200 comments in 25 discussions of top 501 subreddits on a single day, for April 23 and May 5, 2021

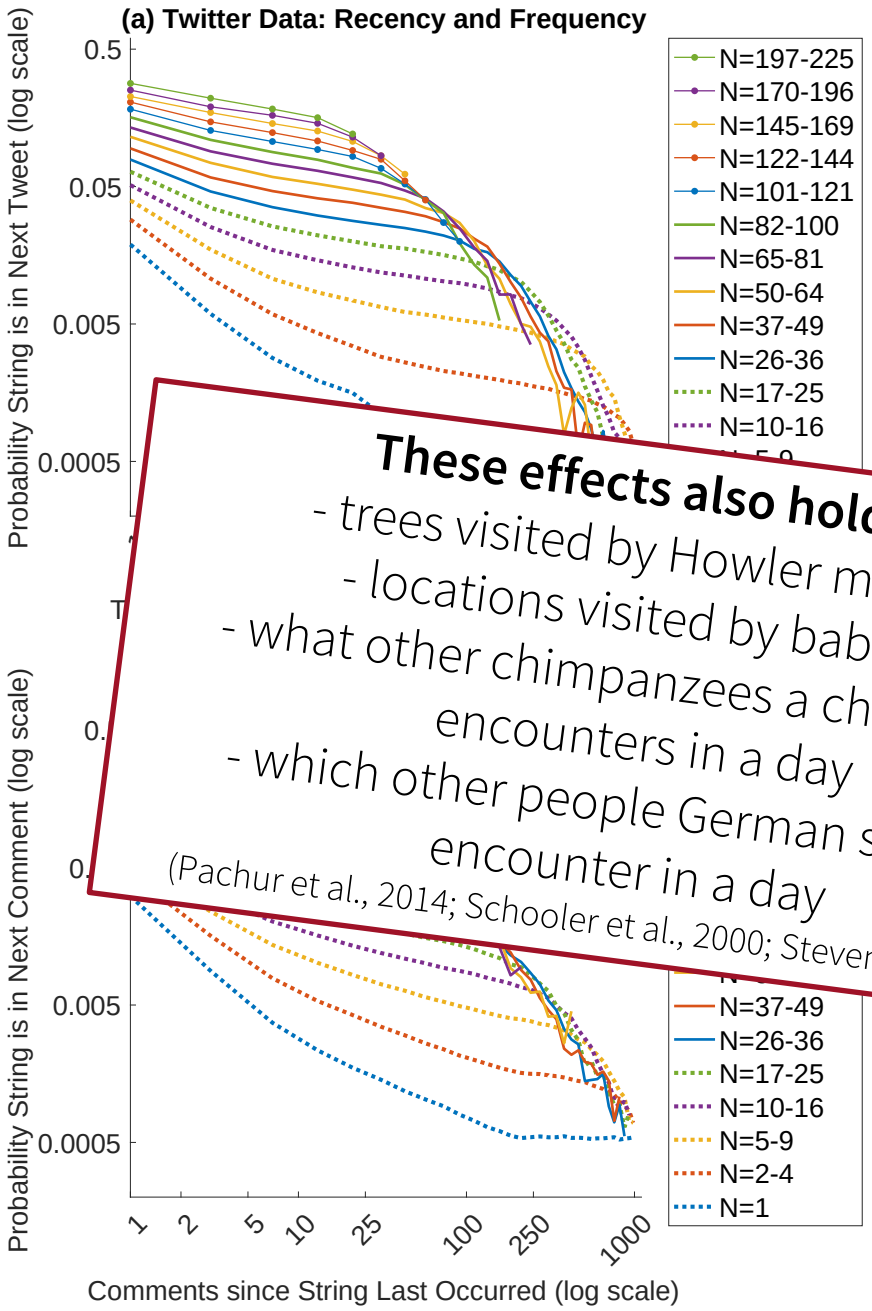
=> 1,133,182 comments, averaging 14.85 strings




(a) Twitter Data: Recency and Frequency
Probability String is in Next Tweet (log scale)
0.5
0.05
0.005
0.0005
1
2
5
10
25
100
250
1000
Tweets since String Last Occurred (log scale)
N=197-225
N=170-196
N=145-169
N=122-144
N=101-121
N=82-100
N=65-81
N=50-64
N=37-49
N=26-36
N=17-25
N=10-16
N=5-9
N=2-4
N=1


(a) Reddit Data: Recency and Frequency
Probability String is in Next Comment (log scale)
0.5
0.05
0.005
0.0005
1
2
5
10
25
100
250
1000
Comments since String Last Occurred (log scale)
N=197-225
N=170-196
N=145-169
N=122-144
N=101-121
N=82-100
N=65-81
N=50-64
N=37-49
N=26-36
N=17-25
N=10-16
N=5-9
N=2-4
N=1



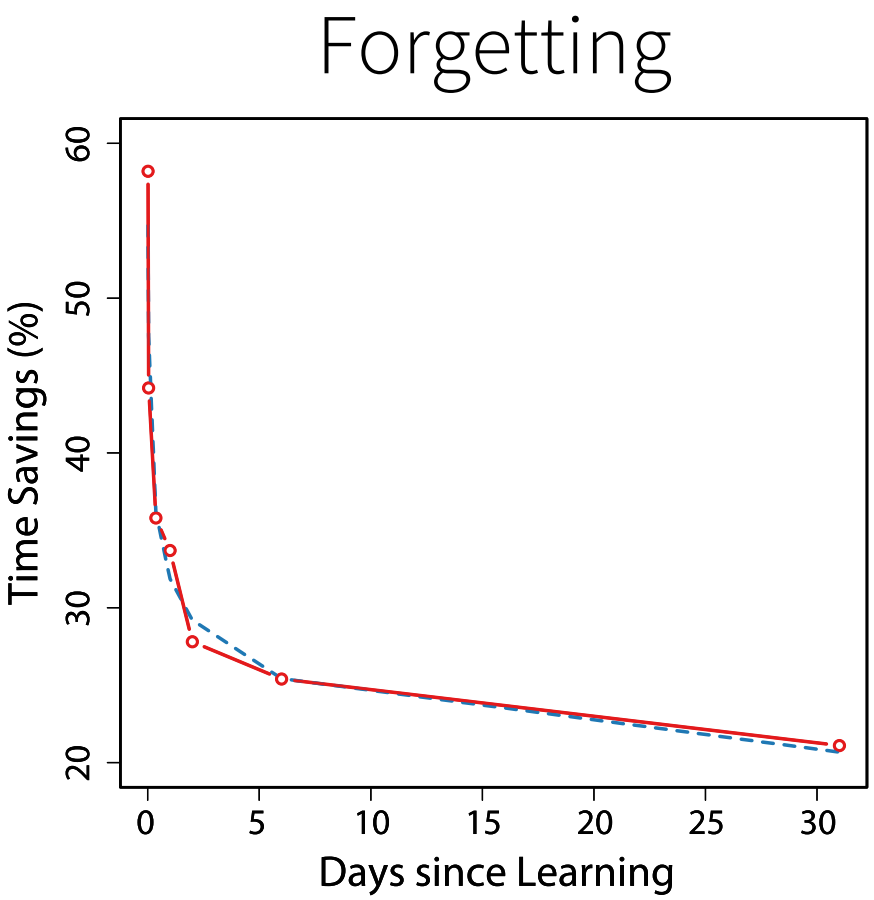


These effects also hold for:

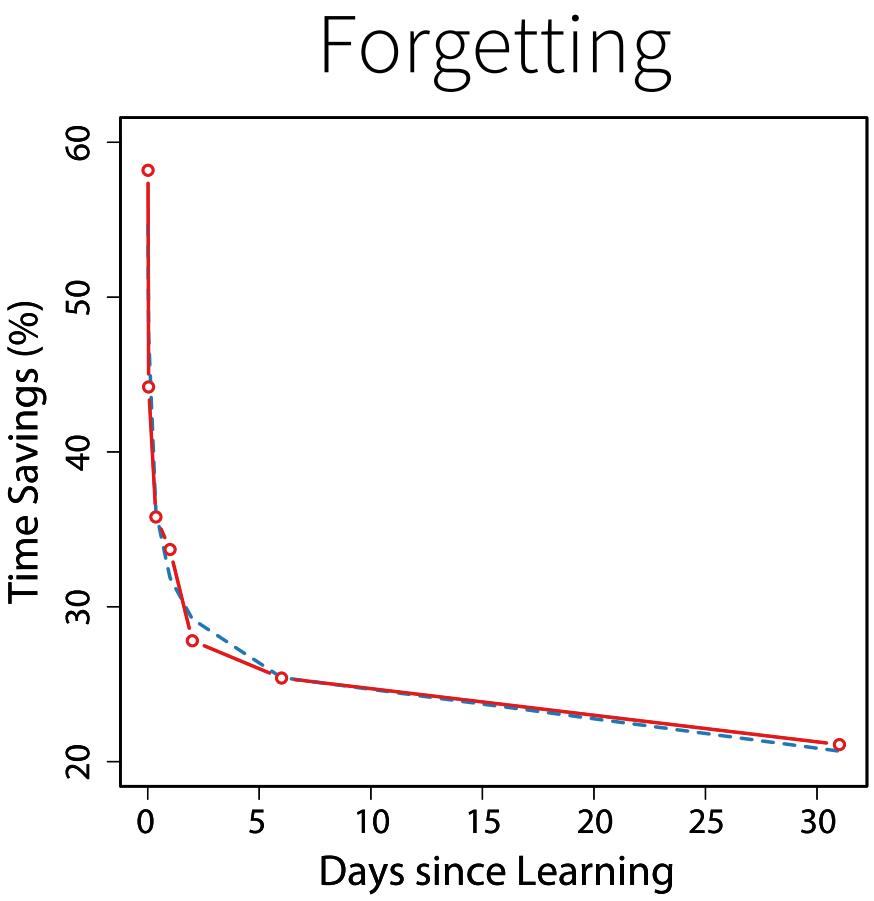
- trees visited by Howler monkeys
- locations visited by baboons
- what other chimpanzees a chimpanzee encounters in a day
- which other people German students encounter in a day

(Pachur et al., 2014; Schooler et al., 2000; Stevens et al., 2016)

How does forgetting work?

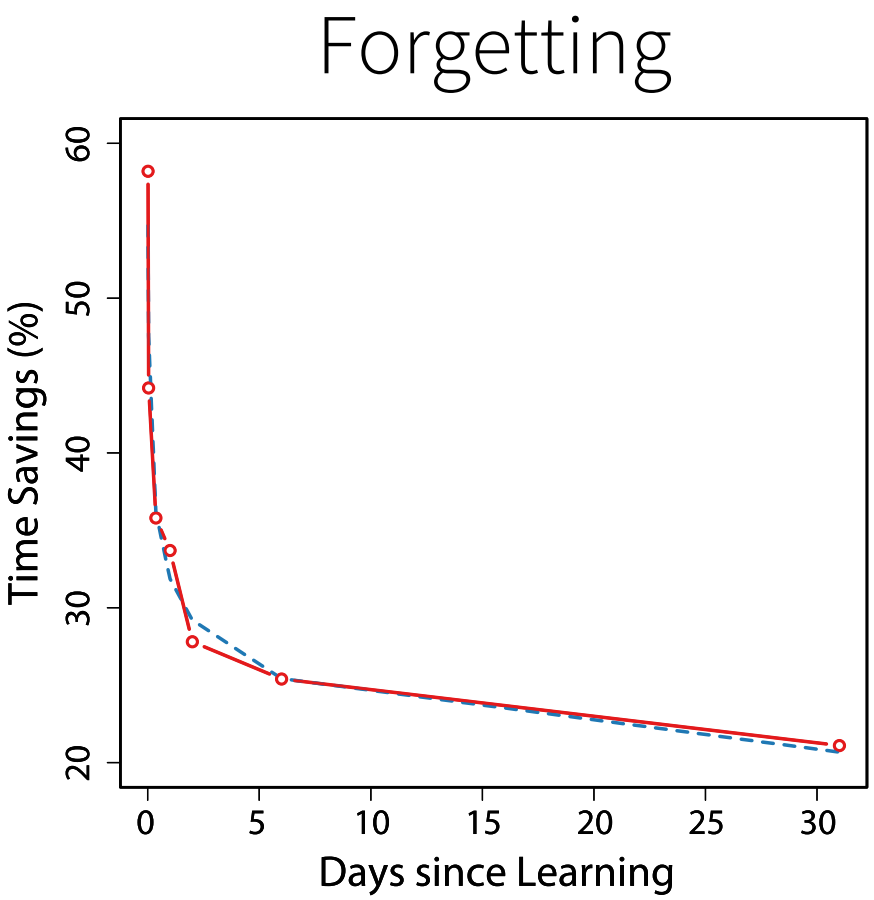


How does forgetting work?



Two theories (mechanisms?) of forgetting

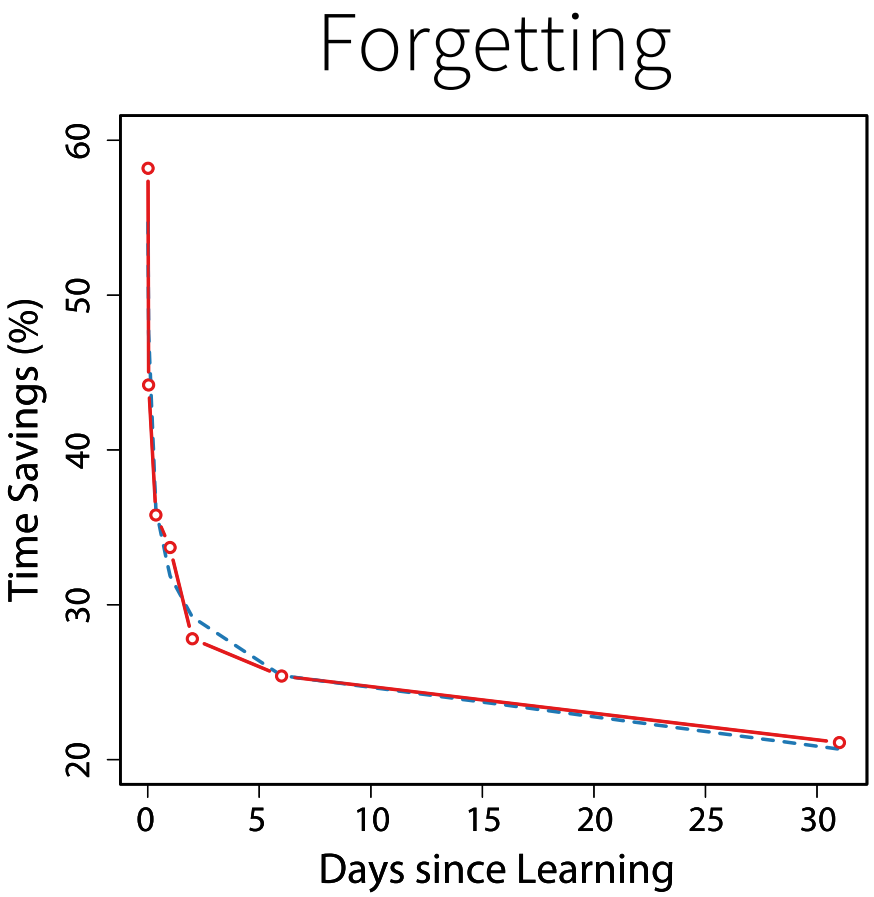
How does forgetting work?



Two theories (mechanisms?) of forgetting

- Decay: elements in memory fade away if they are not used (ACT-R)

How does forgetting work?

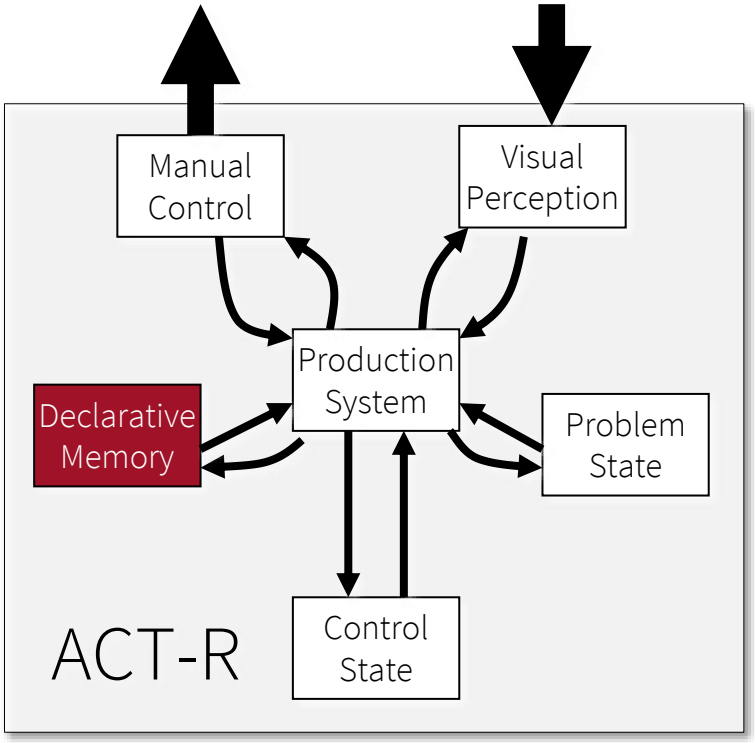


Two theories (mechanisms?) of forgetting

- Decay: elements in memory fade away if they are not used (ACT-R)
- Interference: facts are overwritten by similar facts (nengo)

Today

- Declarative memory: basics
  - The rational basis for declarative memory
  - Declarative memory in ACT-R
  - Example model
- Assignment


Manual Control
Visual Perception
Production System
Declarative Memory
Problem State
Control State
ACT-R

DM in ACT-R

Declarative Memory

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8
```

Declarative Memory

Psychologically meaningless

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8
```

Declarative Memory

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8
```

```
+retrieval>
    isa addition-fact
    addend1 1
```

Declarative Memory

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8
```

```
+retrieval>
    isa addition-fact
    addend1 1
```

How long does it take to retrieve a fact?

Declarative Memory

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8
```

```
+retrieval>
    isa addition-fact
    addend1 1
```

How long does it take to retrieve a fact?

Can it be retrieved at all?

Declarative Memory

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8

fact13
    isa addition-fact
    addend1 1
    addend2 3
    sum 4
```

```
+retrieval>
    isa addition-fact
    addend1 1
```

How long does it take to retrieve a fact?

Can it be retrieved at all?

Which one is retrieved?

Declarative Memory

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8
```

```
fact13
    isa addition-fact
    addend1 1
    addend2 3
    sum 4
```

```
+retrieval>
    isa addition-fact
    addend1 1
```

How long does it take to retrieve a fact?

Can it be retrieved at all?

Which one is retrieved?

How do we learn those chunks?

Declarative Memory

```
fact17
    isa addition-fact
    addend1 1
    addend2 7
    sum 8
```

```
fact13
    isa addition-fact
    addend1 1
    addend2 3
    sum 4
```

```
+retrieval>
    isa addition-fact
    addend1 1
```

How long does it take to retrieve a fact?

Can it be retrieved at all?

Which one is retrieved?

How do we learn those chunks?

How do we forget them?

Activation

Activation

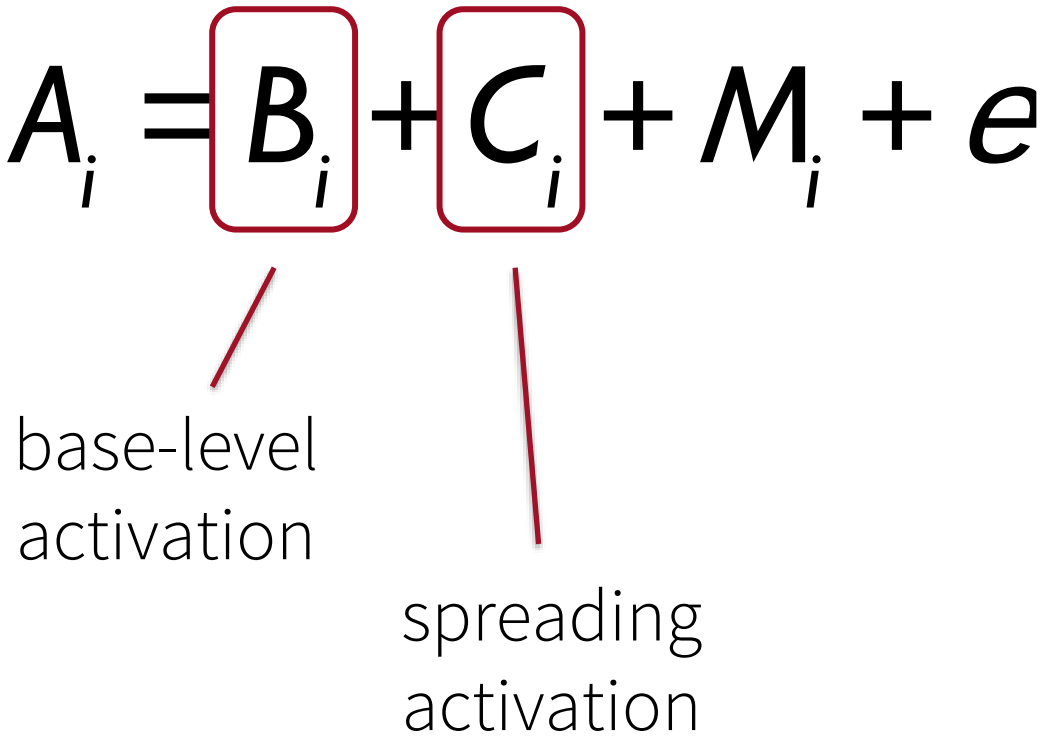
$$A_i = B_i + C_i + M_i + e$$

Activation

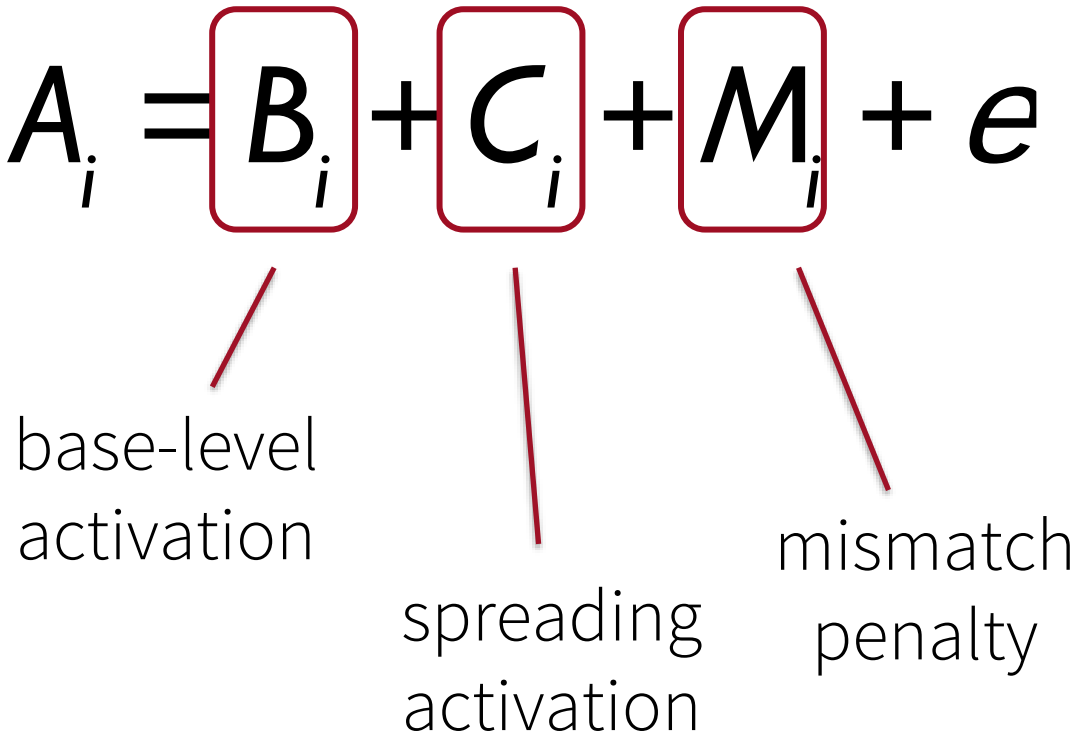
$$A_i = B_i + C_i + M_i + e$$

base-level activation

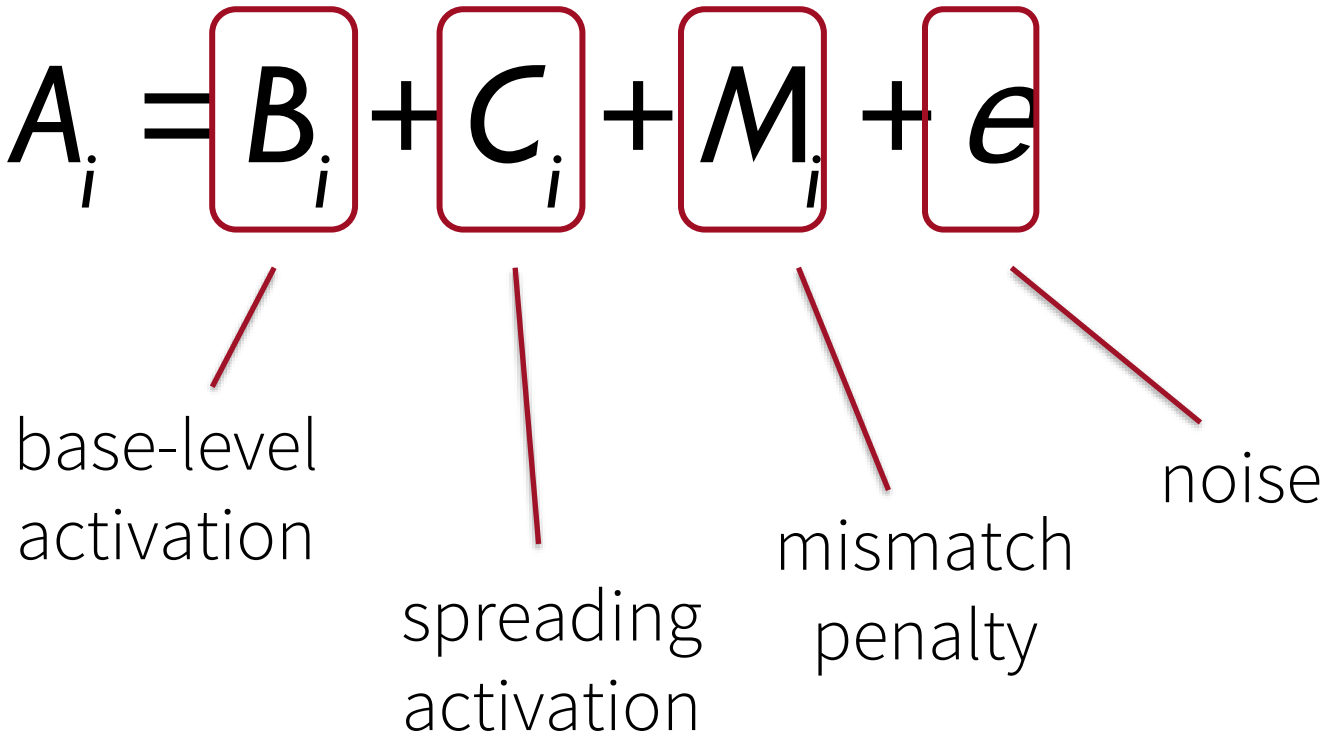
Activation



Activation


$A_i = B_i + C_i + M_i + e$
base-level activation
spreading activation
mismatch penalty

Activation


$A_i = B_i + C_i + M_i + \epsilon$
base-level activation
spreading activation
mismatch penalty
noise

Activation

For now:

$$A_i = B_i + e$$

Activation

For now:

A_i = activation of chunk *i*

Activation

A_i = activation of chunk *i*

Activation

A_i = activation of chunk *i*

Activation

A_i = activation of chunk i

- Determines:
  - if a chunk can be retrieved: $A_i > t$

Activation

A_i = **activation of chunk *i***

- Determines:
  - if a chunk can be retrieved: $A_i > t$

retrieval threshold

Activation

A_i = activation of chunk *i*

- Determines:
  - if a chunk can be retrieved: $A_i > t$
  - which chunk is retrieved: highest A_i

retrieval threshold

Activation

A_i = activation of chunk *i*

- Determines:
  - if a chunk can be retrieved: $A_i > t$
  - which chunk is retrieved: highest A_i
  - retrieval time: $RT = Fe^{-A}$

retrieval threshold

Activation

A_i = activation of chunk i

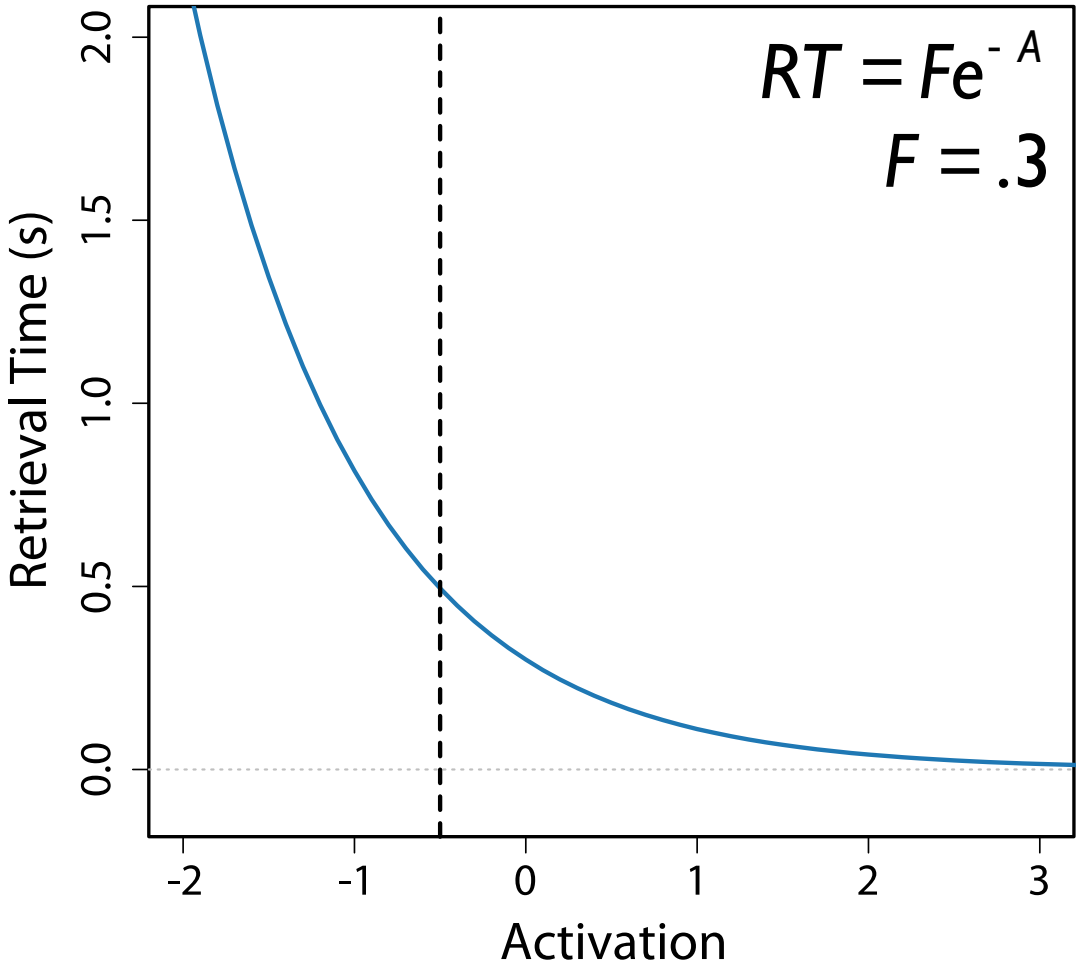
- Determines:
  - if a chunk can be retrieved: $A_i > t$
  - which chunk is retrieved: highest A_i
  - retrieval time: $RT = Fe^{-A}$

retrieval threshold

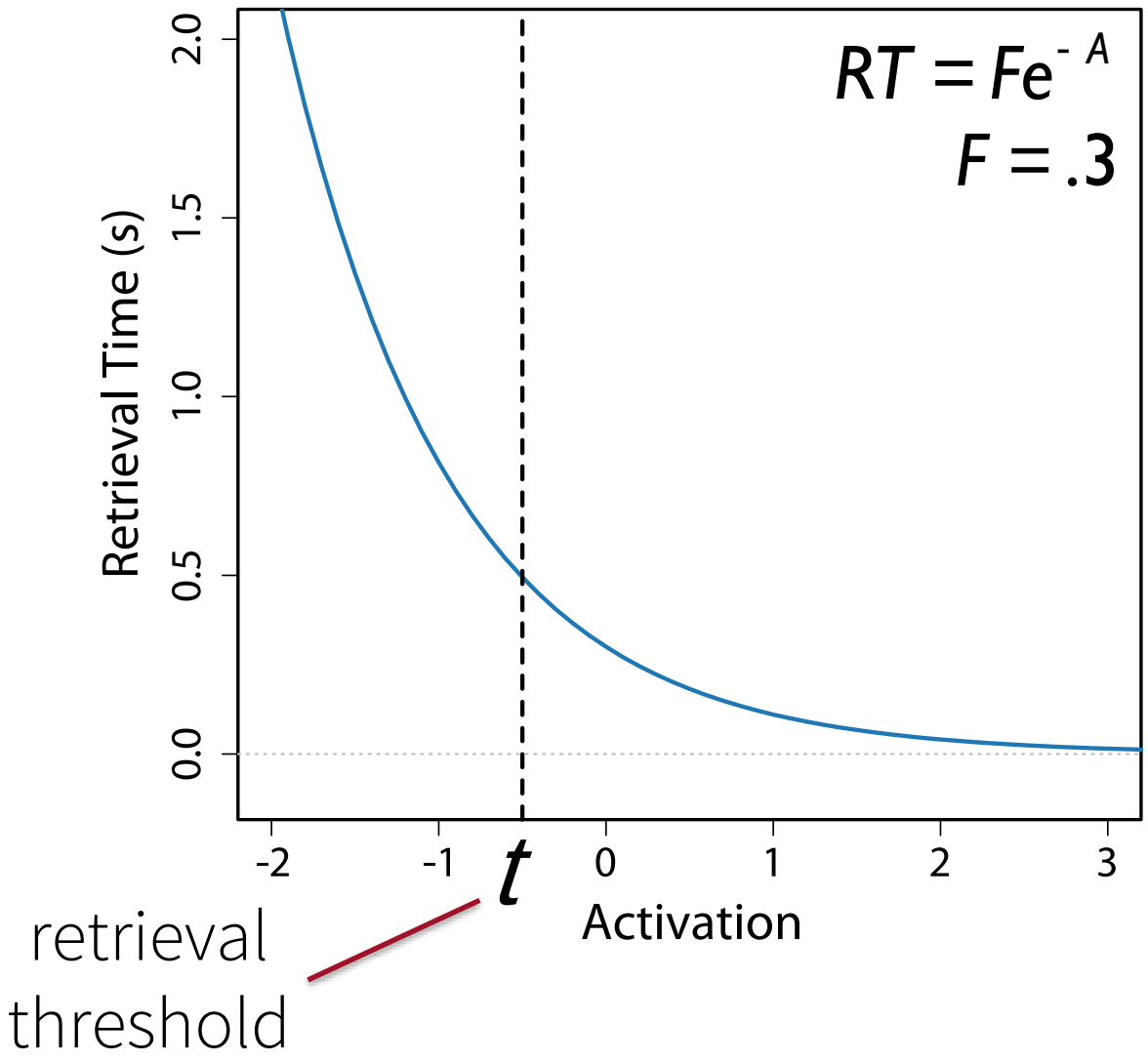
latency factor

Retrieval Time

Retrieval Time



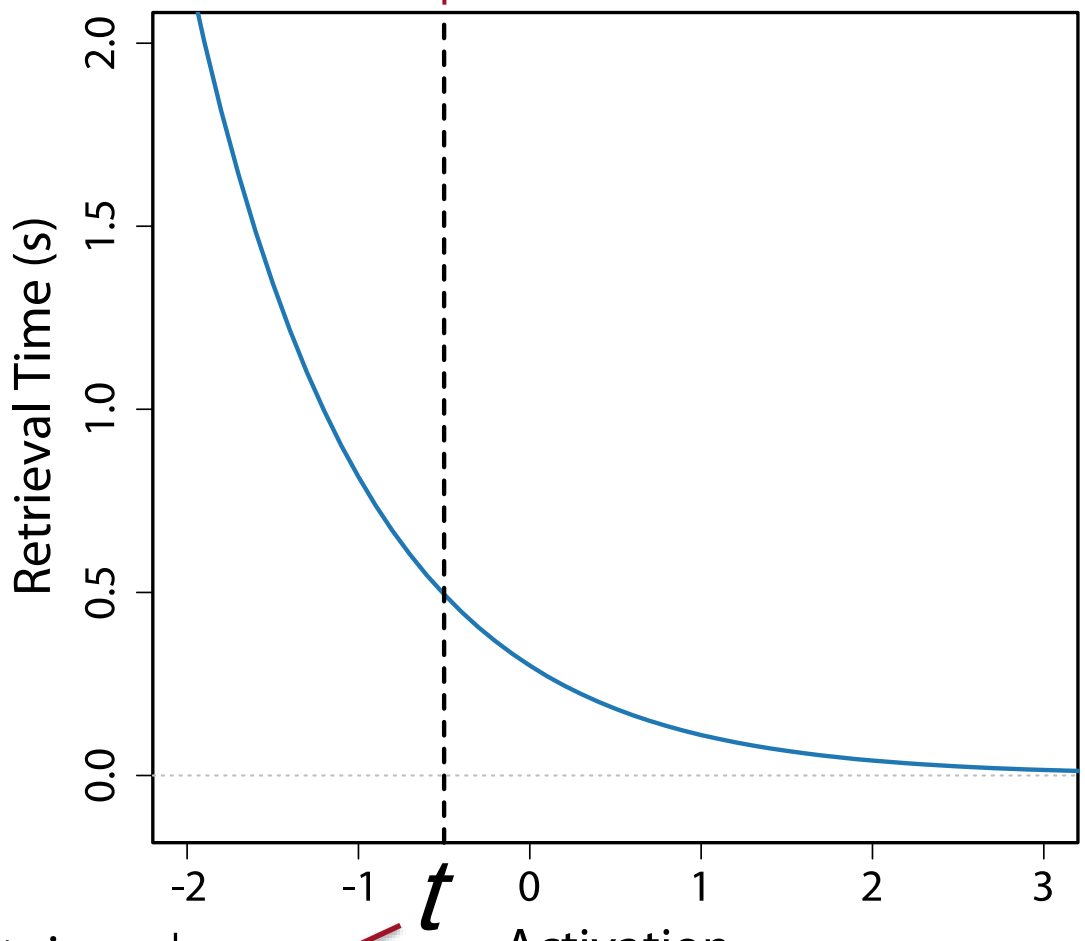
Retrieval Time



$A_i > t$

retrieval success

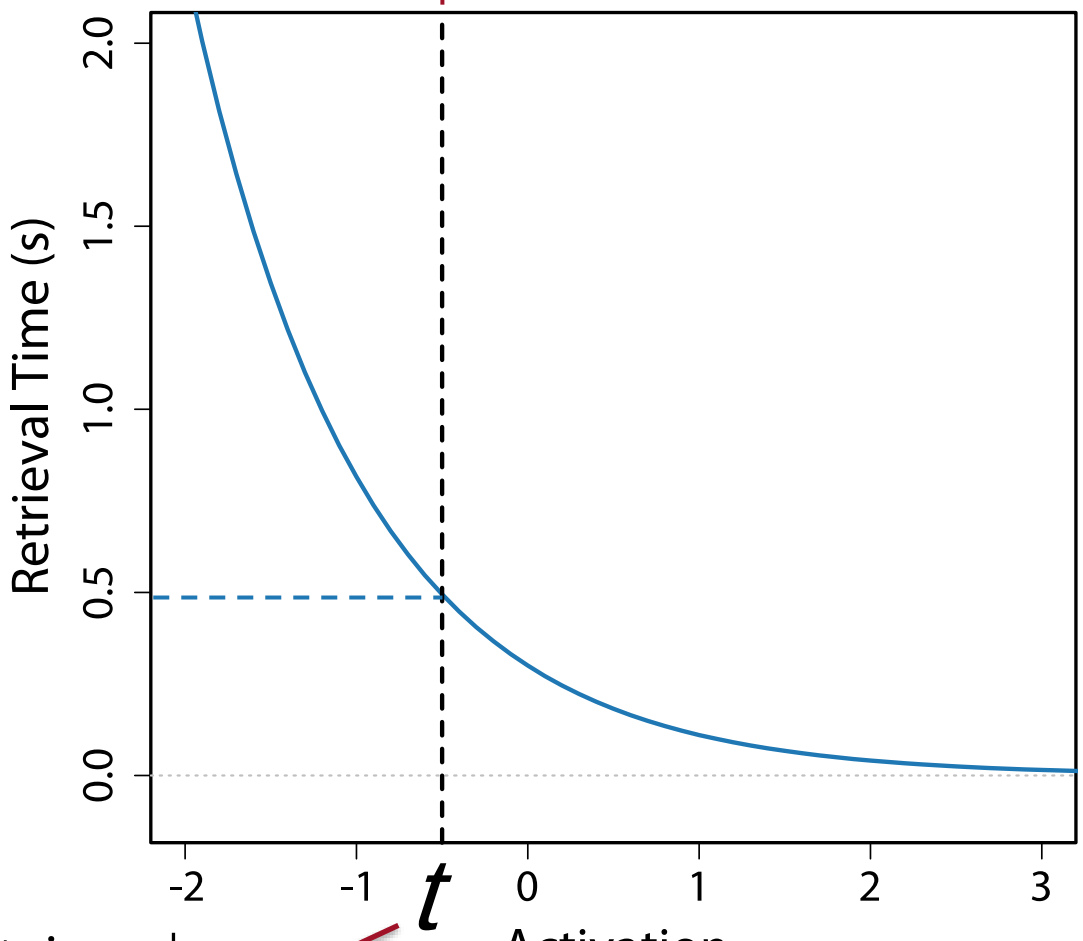
$RT_{success} = Fe^{-A}$



retrieval threshold

$RT_{failure} = Fe^{-t}$ $A_i < t$ retrieval failure

$A_i > t$ retrieval success $RT_{success} = Fe^{-A}$



retrieval threshold

Activation

A_i = activation of chunk *i*

Activation

A_i = activation of chunk *i*

- A_i estimates the odds that a fact is needed in the current context

Activation

A_i = **activation of chunk *i***

- A_i estimates the odds that a fact is needed in the current context

$$A_i = B_i + C_i + M_i + e$$

Activation

A_i = **activation of chunk *i***

- A_i estimates the odds that a fact is needed in the current context

$$A_i = B_i + C_i + M_i + e$$

base-level activation: odds regardless of context

Activation

A_i = **activation of chunk *i***

- A_i estimates the odds that a fact is needed in the current context

$$A_i = B_i + C_i + M_i + e$$

base-level activation: odds regardless of context

influence of current context

Base-Level Activation

Base-Level Activation

- Reflects the **odds** that you need a chunk based on frequency and recency

Odds

Odds

- Odds is the probability of an event occurring compared to it not occurring:

$$Odds = \frac{p}{1 - p}$$

Odds

- Odds is the probability of an event occurring compared to it not occurring:

$$Odds = \frac{p}{1 - p}$$

- Odds of 3 means that an event is 3 times as likely to occur than not to occur

Odds

- Odds is the probability of an event occurring compared to it not occurring:

$$Odds = \frac{p}{1 - p}$$

- Odds of 3 means that an event is 3 times as likely to occur than not to occur
- Odds of 1 is 50/50

Base-Level Activation

- Reflects the **odds** that you need a chunk based on frequency and recency

Base-Level Activation

- Reflects the **odds** that you need a chunk based on frequency and recency
- For a single presentation of a chunk:

$$Odds_i(t) = (t - t_i)^{-d}$$

Base-Level Activation

- Reflects the **odds** that you need a chunk based on frequency and recency
- For a single presentation of a chunk:

$$Odds_i(t) = (t - t_i)^{-d}$$

time since last use

Base-Level Activation

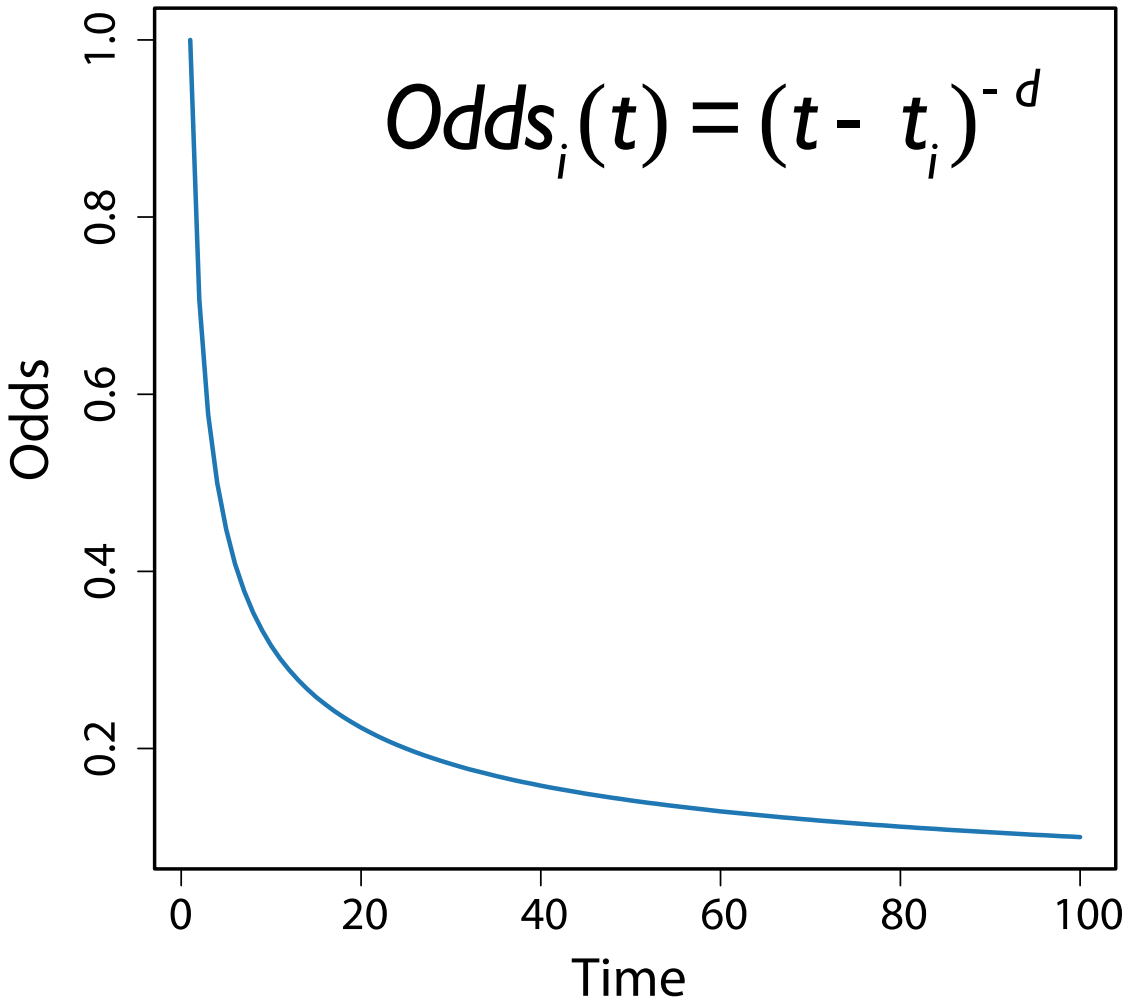
- Reflects the **odds** that you need a chunk based on frequency and recency
- For a single presentation of a chunk:

$$Odds_i(t) = (t - t_i)^{-d}$$

time since last use

decay parameter

Base-Level Activation



Base-Level Activation

- Reflects the **odds** that you need a chunk based on frequency and recency
- For a single presentation:

$$Odds_i(t) = (t - t_i)^{-d}$$

Base-Level Activation

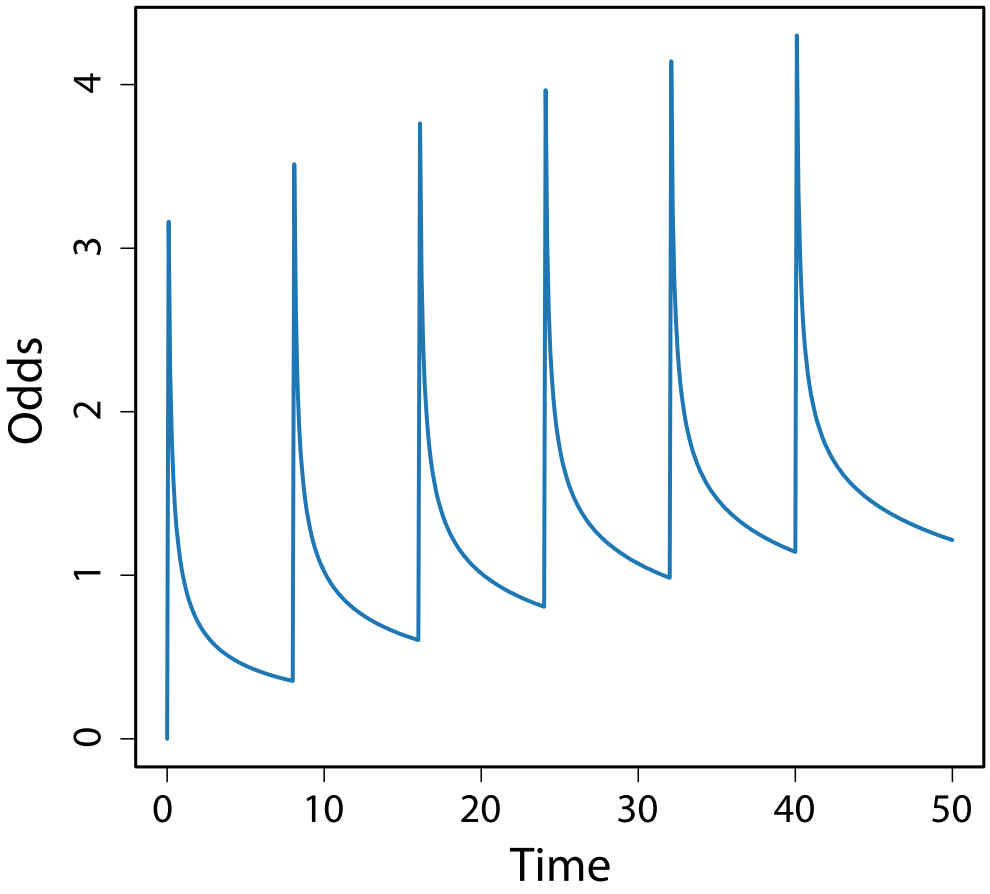
- Reflects the **odds** that you need a chunk based on frequency and recency
- For a single presentation:

$$Odds_i(t) = (t - t_i)^{-d}$$

- Presentations are additive (practice):

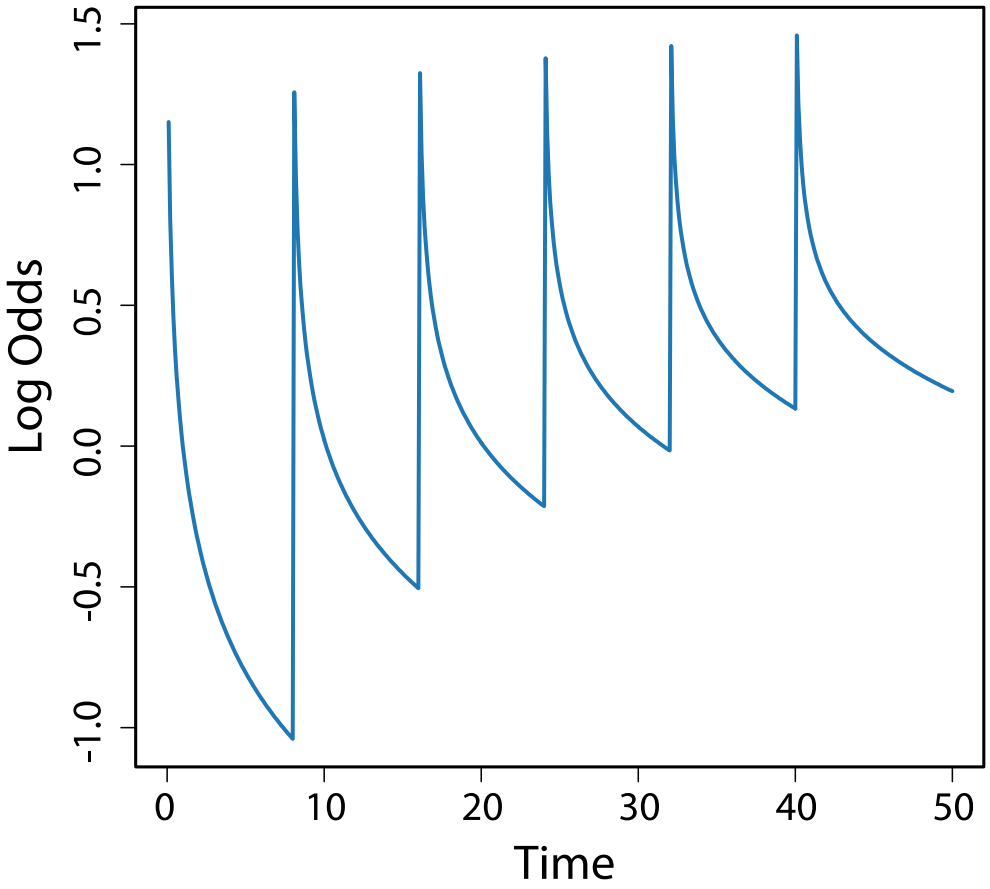
$$Odds_i(t) = \sum_{k=1}^{n} (t - t_k)^{-d}$$

Base-Level Activation



$$Odds_i(t) = \sum_{k=1}^{n} (t - t_k)^{-d}$$

Base-Level Activation



$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Base-Level Activation

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Base-Level Activation

n presentations

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Base-Level Activation

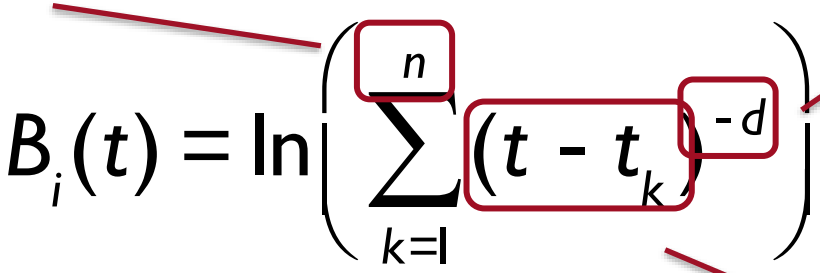
n presentations

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

time since *n*th presentation

Base-Level Activation

n presentations



decay parameter

time since *n*th presentation

What is a presentation?

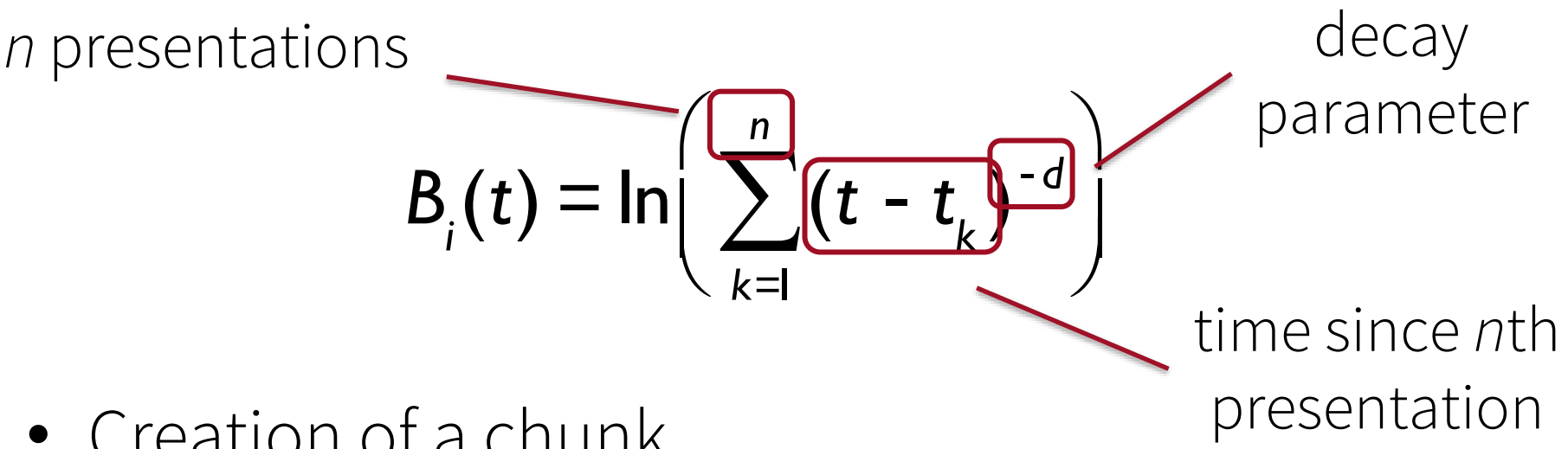
n presentations

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

decay parameter

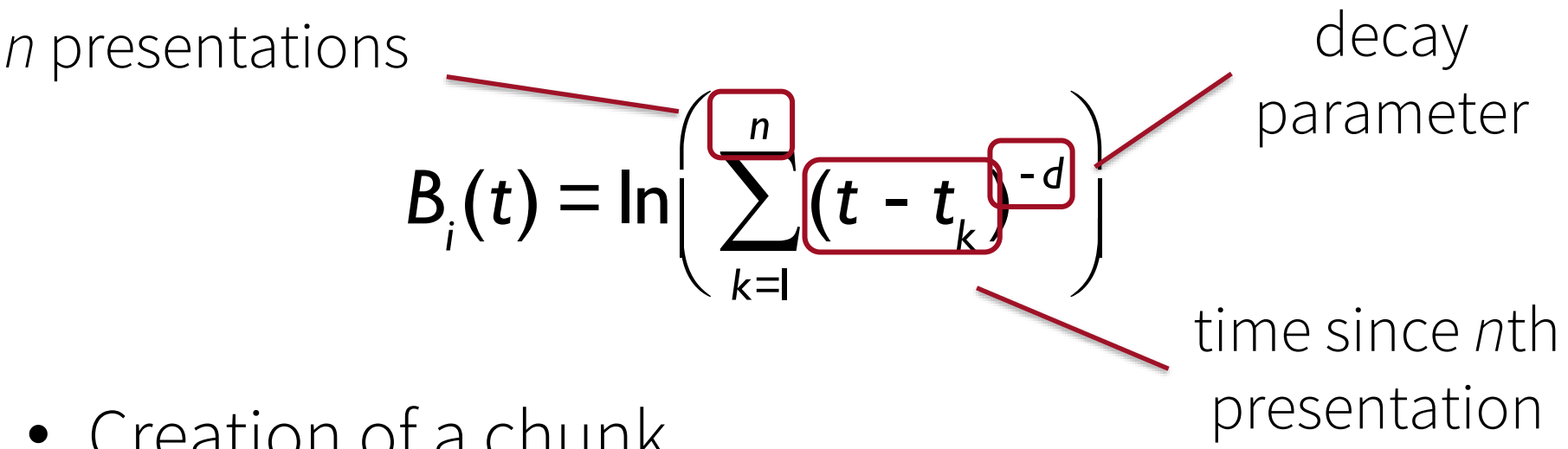
time since *n*th presentation

What is a presentation?



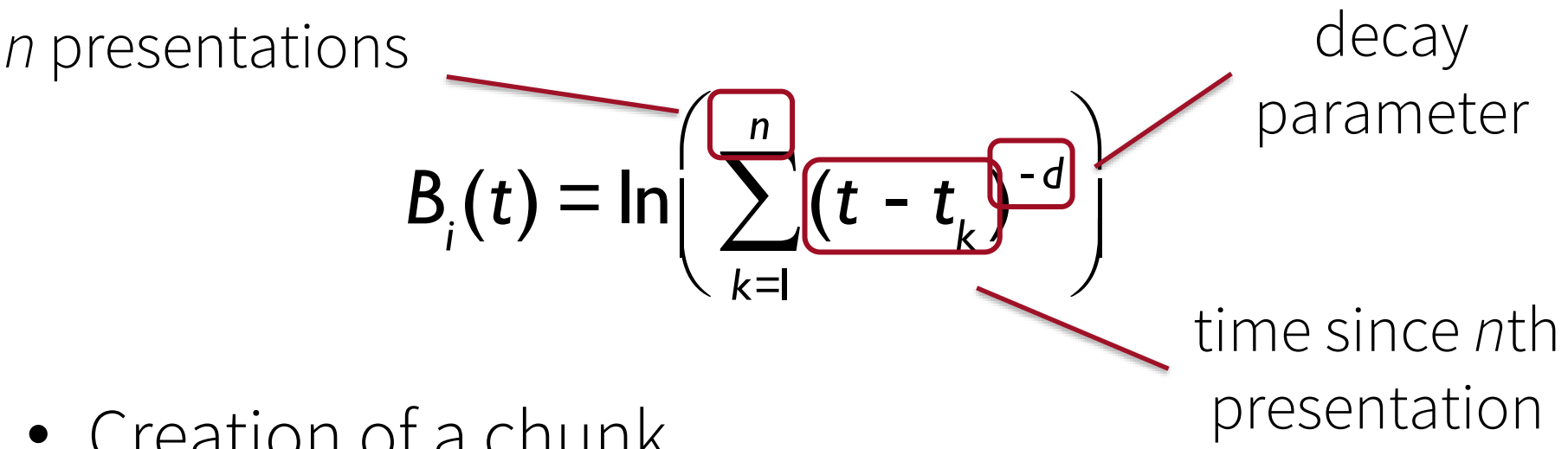
- Creation of a chunk

What is a presentation?



- Creation of a chunk
- Buffer clearing / merging

What is a presentation?



- Creation of a chunk
- Buffer clearing / merging
- Buffer clearing after retrieval: rehearsal!

What is a presentation?

$$B_i(t) = \ln\left(\sum_{k=1}^{n} (t - t_k)^{-d} \right)$$

What is a presentation?

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

- For every retrieval, ACT-R has to calculate the activity of matching chunks by summing over all presentations

What is a presentation?

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

- For every retrieval, ACT-R has to calculate the activity of matching chunks by summing over all presentations
- Computationally expensive

What is a presentation?

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

- For every retrieval, ACT-R has to calculate the activity of matching chunks by summing over all presentations
- Computationally expensive
- Implies that we store all encounters

Optimized Learning

$$B_i(t) = \ln\left(\sum_{k=1}^{n} (t - t_k)^{-d} \right)$$

Optimized Learning

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Optimized learning:

$$B_i(t) = \ln\left(\frac{n}{1 - d}\right) - d \times \ln(t - t_1)$$

Optimized Learning

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Optimized learning:

$$B_i(t) = \ln\left(\frac{n}{1 - d}\right) + d \times \ln(t - t_1)$$

decay parameter

Optimized Learning

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Optimized learning:

n presentations

$$B_i(t) = \ln\left(\frac{n}{1 - d}\right) + d \times \ln(t - t_1)$$

decay parameter

Optimized Learning

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Optimized learning:

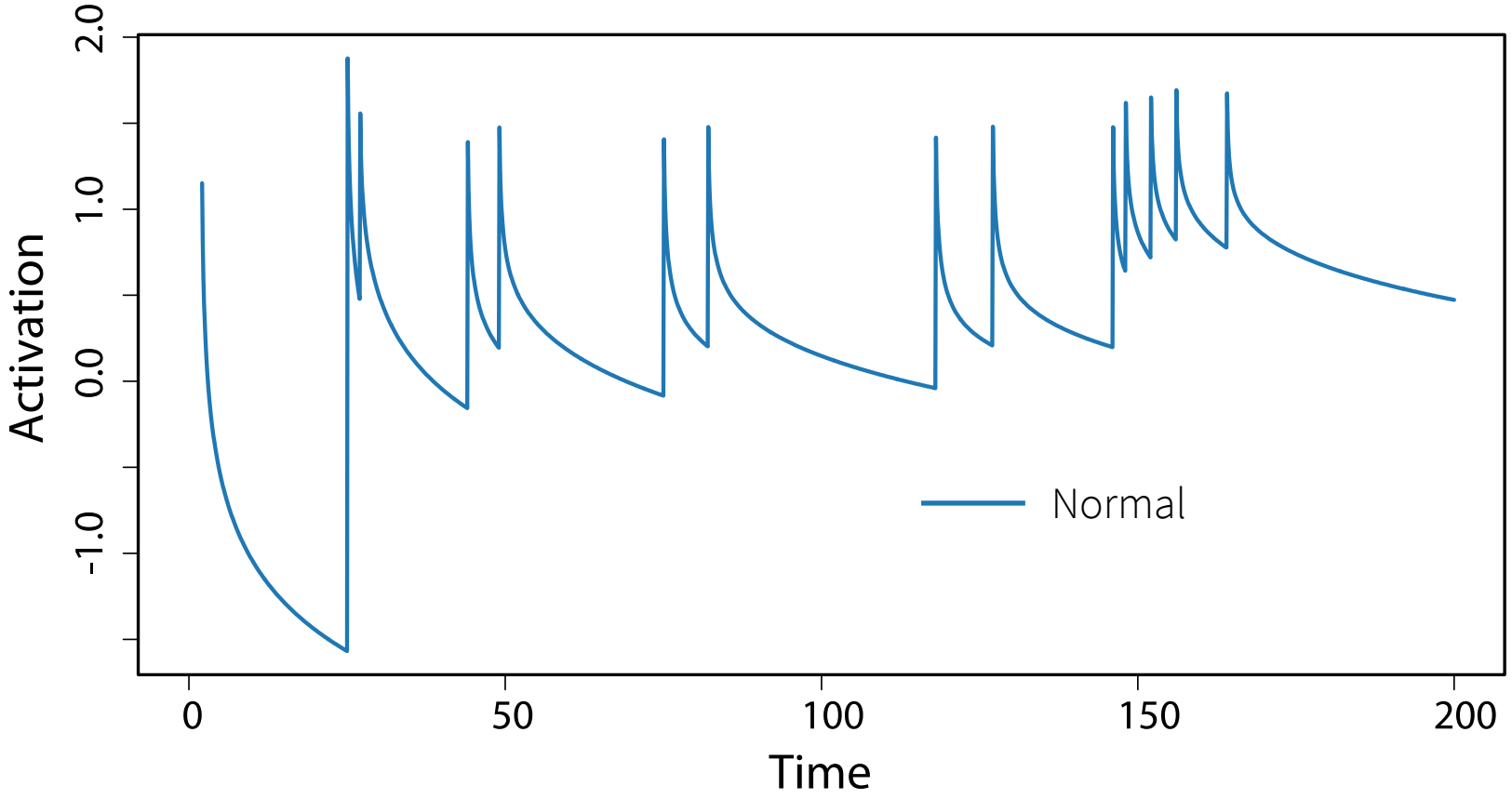
$$B_i(t) = \ln\left(\frac{n}{1 - d}\right) - d \times \ln(t - t_1)$$

n presentations

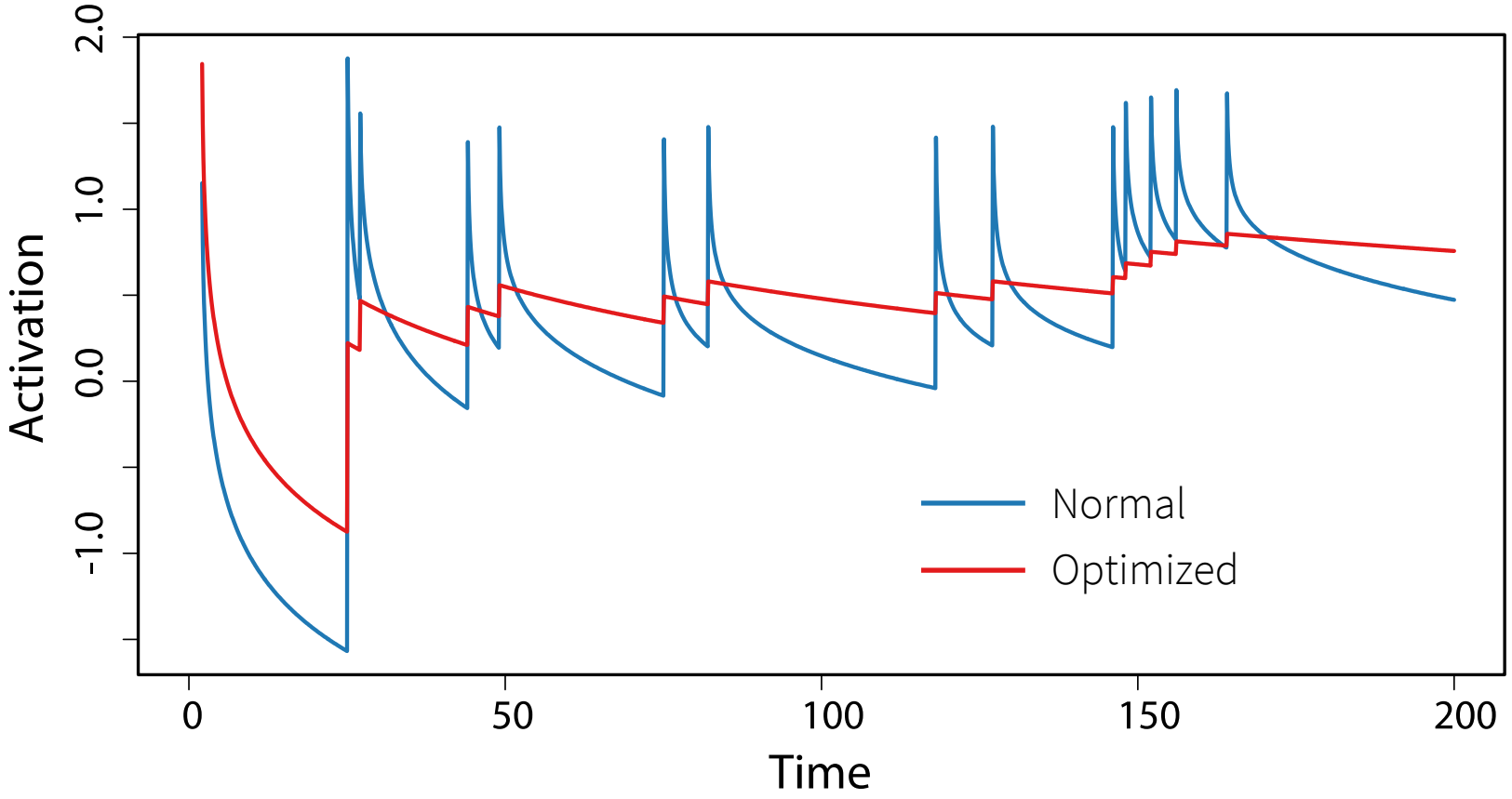
Time since first presentation

decay parameter

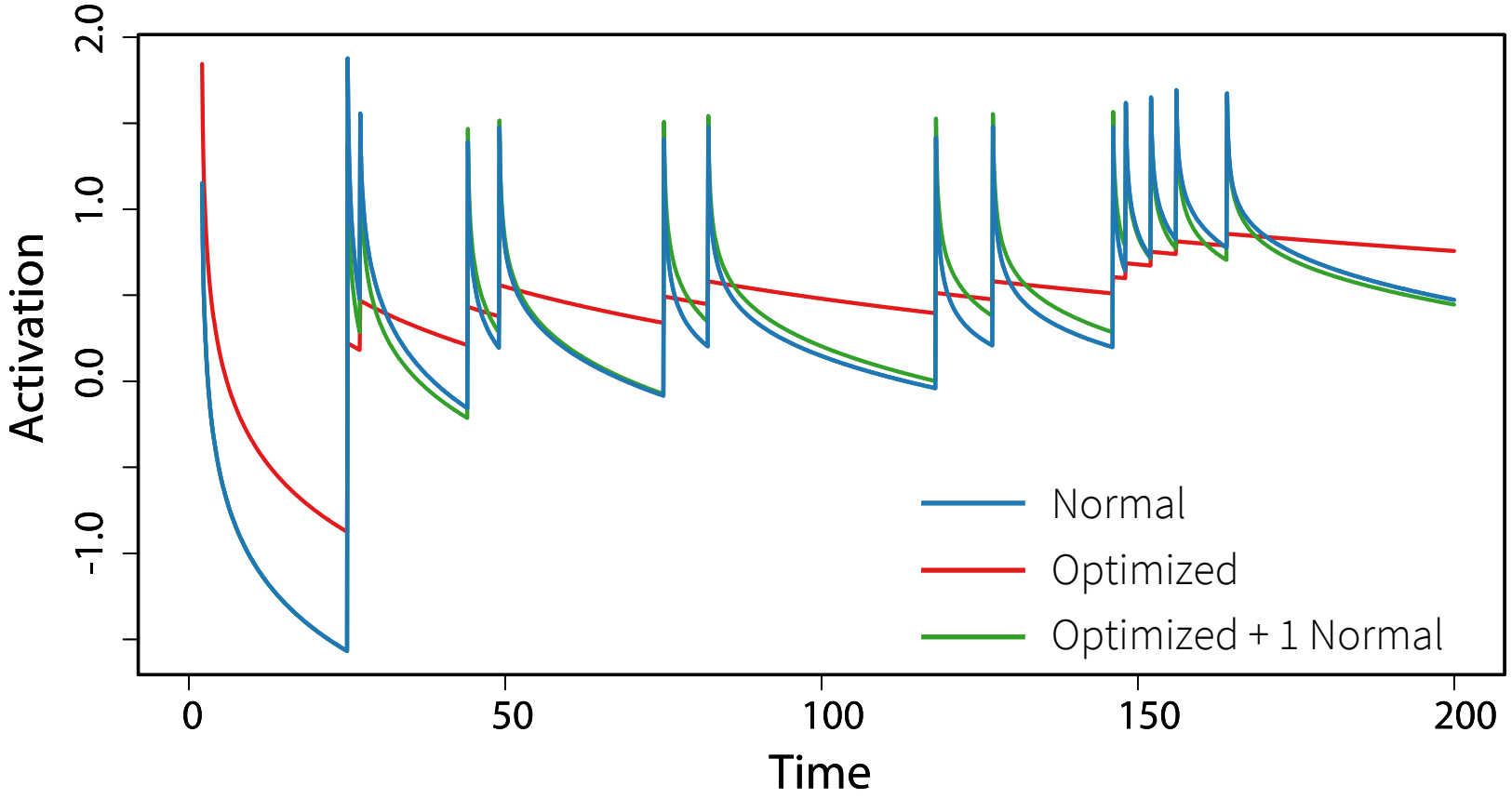
Optimized Learning



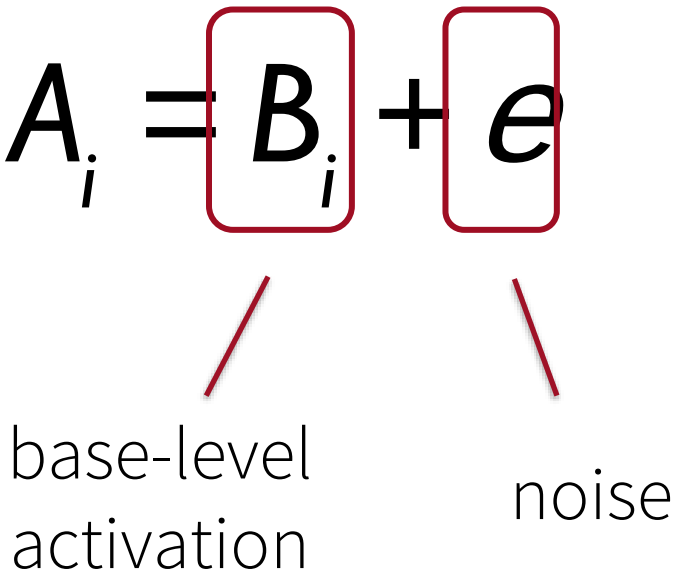
Optimized Learning



Optimized Learning



Activation



Activation

$$A_i = B_i + \epsilon$$

noise

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Noise

$$e = e_{instantanious} + e_{permanent}$$

Noise

$$e = e_{instantanious} + e_{permanent}$$

added at retrieval

Noise

$$e = e_{instantanious} + e_{permanent}$$

added at retrieval

associated with a chunk

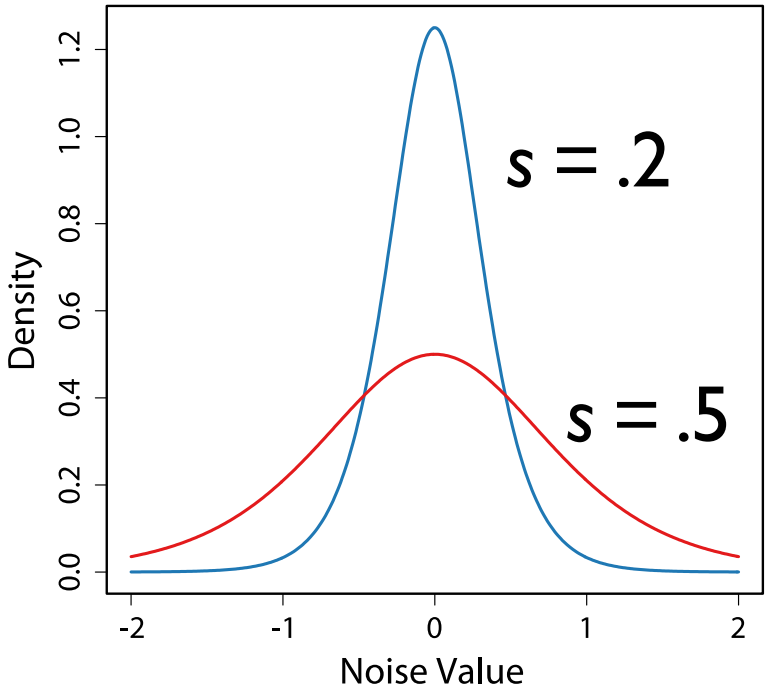
Noise

$$e = e_{instantanious} + e_{permanent}$$

Noise

$$e = e_{instantanious} + e_{permanent}$$

Logistic Distribution



$$m = 0$$

$$s^2 = \frac{p^2}{3} s^2$$

Activation

$$A_i = B_i + \epsilon$$

noise

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

Activation

$$A_i = B_i + e$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

$$e = e_{instantanious} + e_{permanent}$$

Recall Probability

Recall Probability

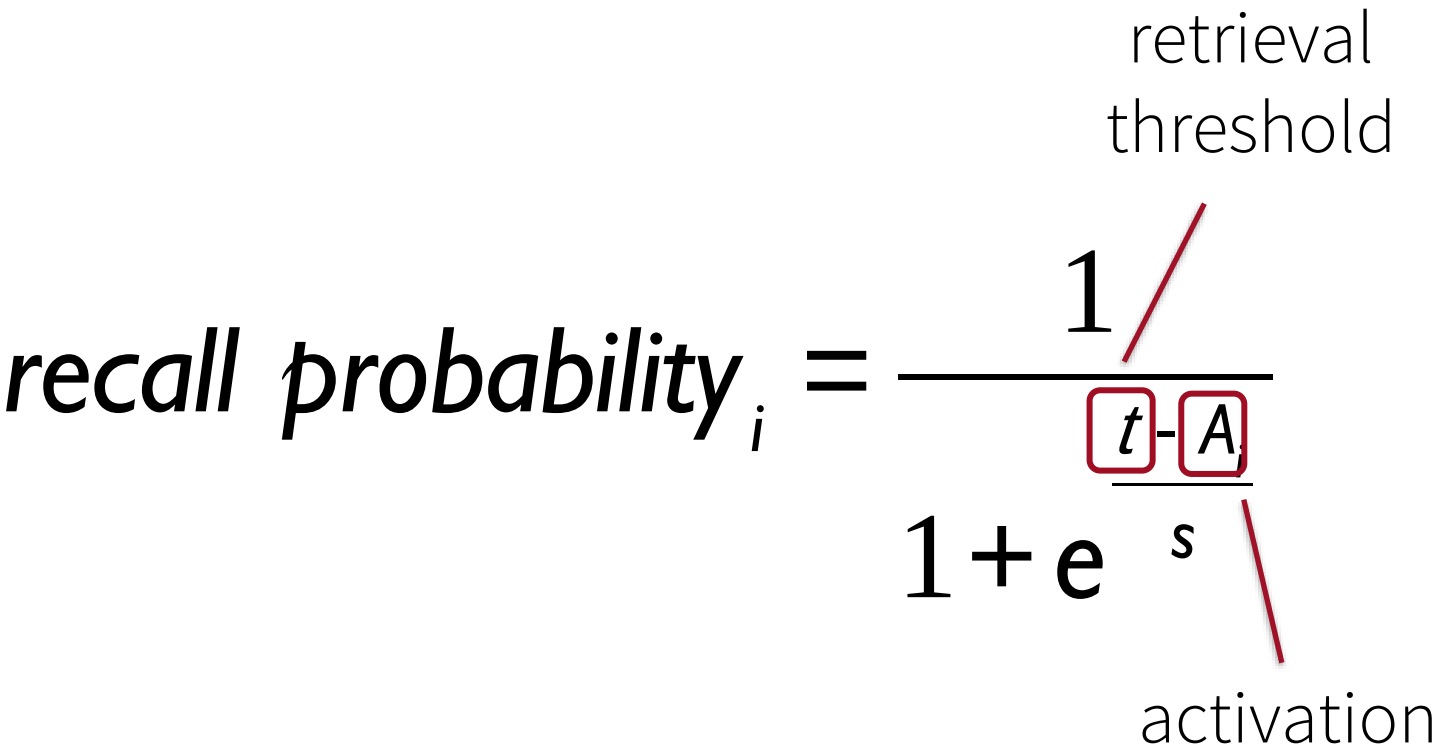
$$recall\ probability_i = \frac{1}{1 + e^{\frac{t - A_i}{s}}}$$

Recall Probability

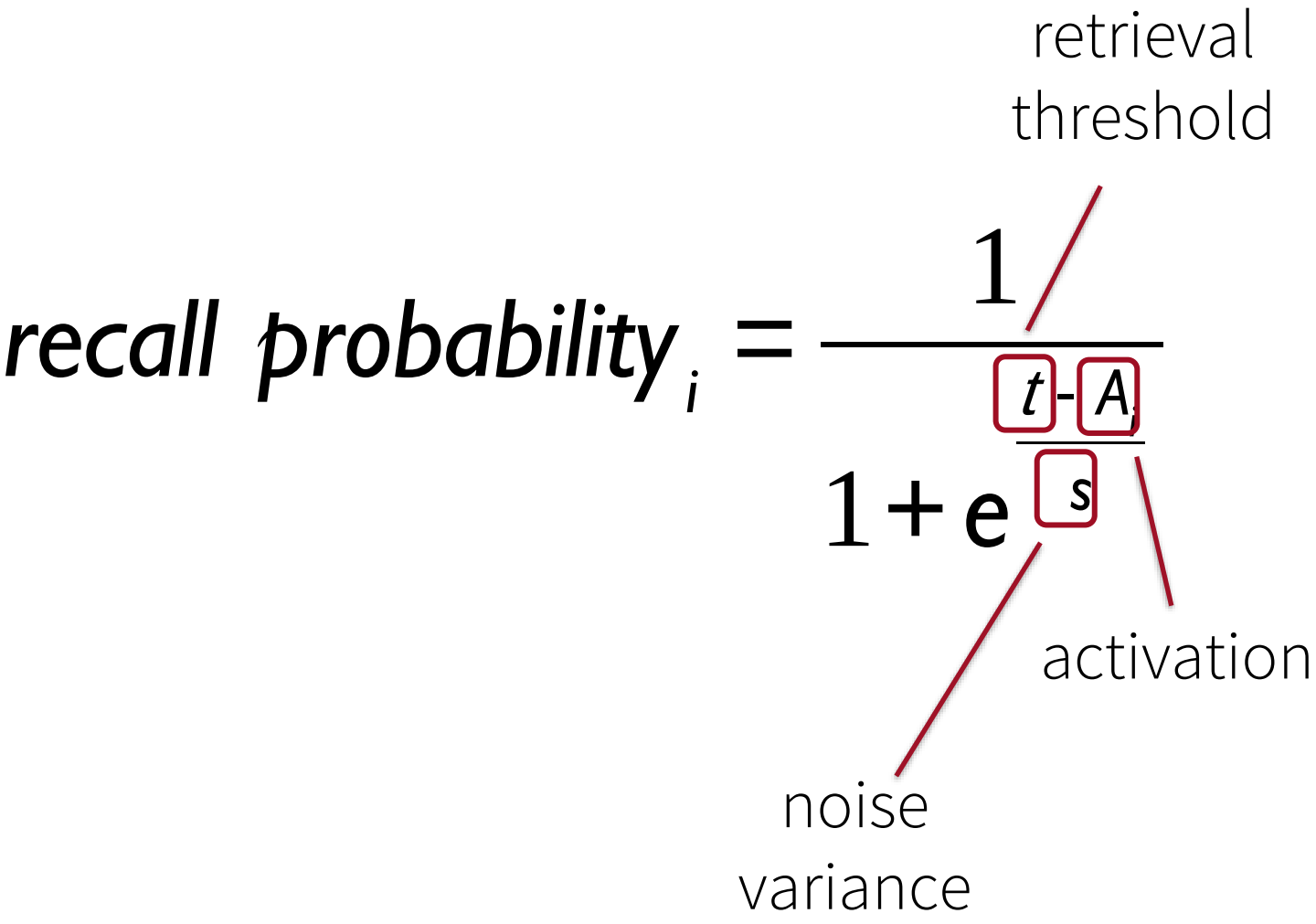
retrieval threshold

$$recall\ probability_i = \frac{1}{1 + e^{\frac{t - A_i}{s}}}$$

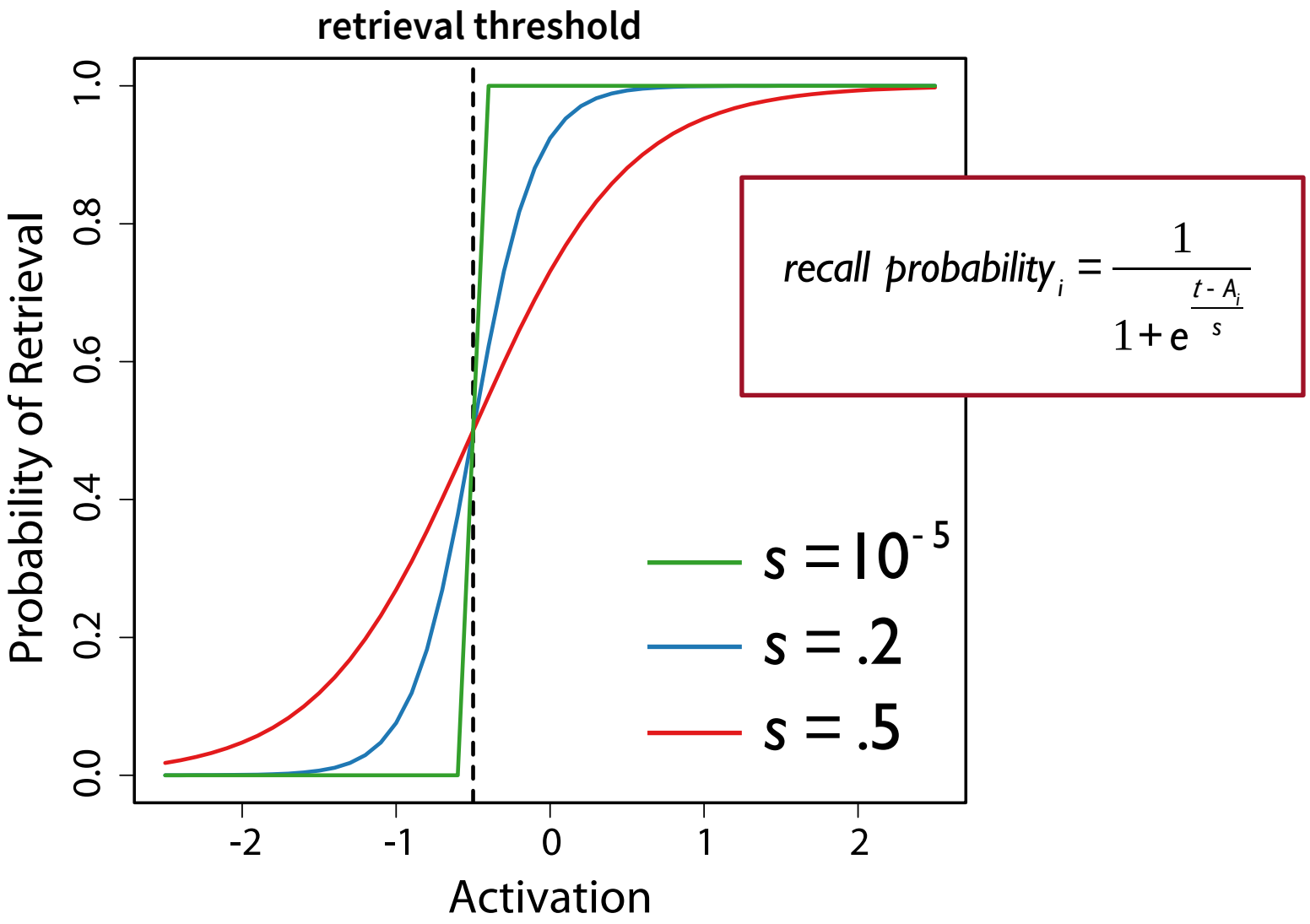
Recall Probability



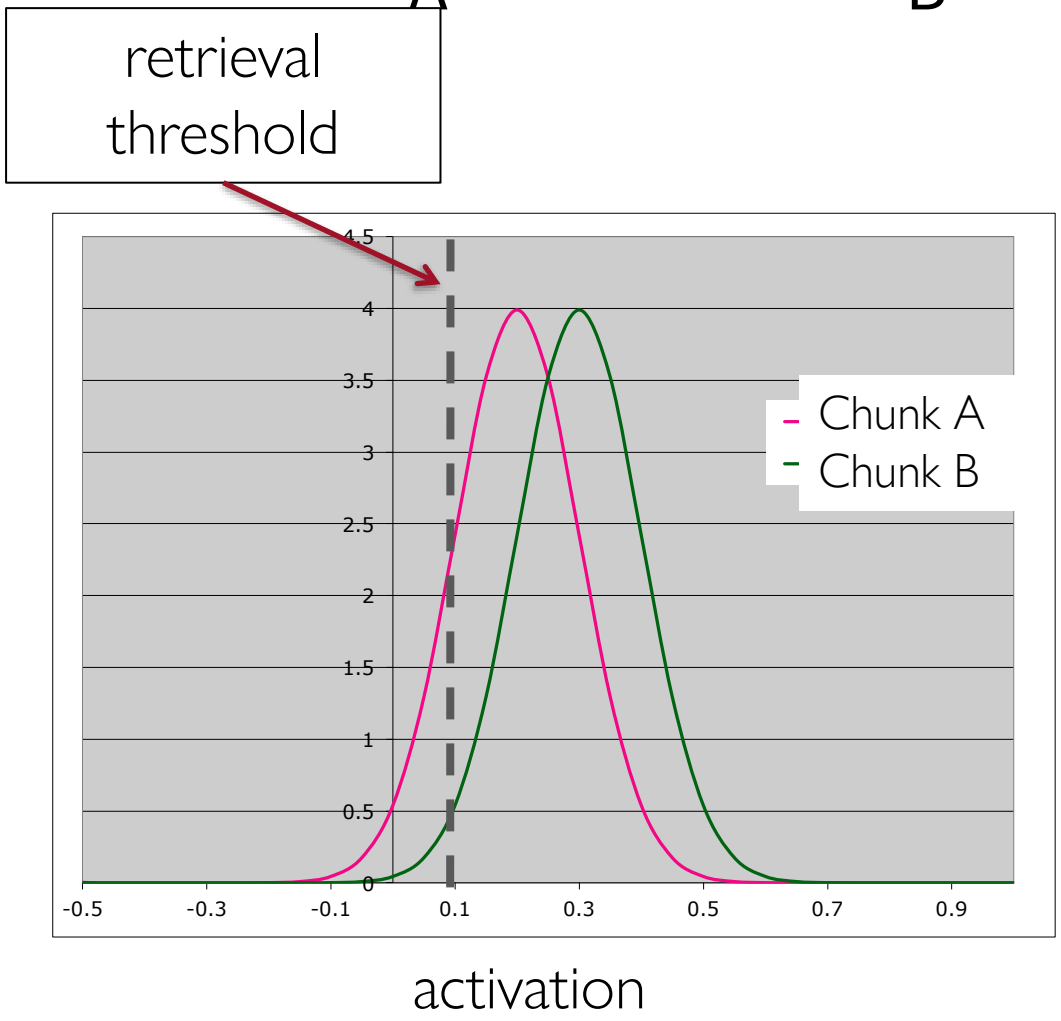
Recall Probability



Recall Probability

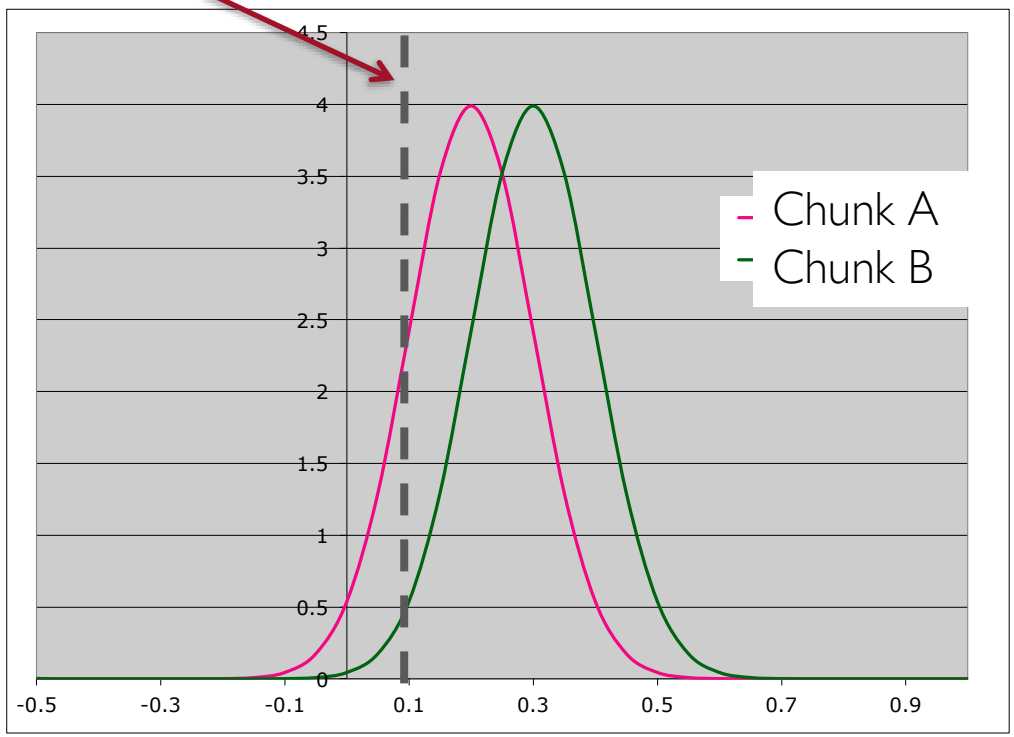


Example: s=0.1, two chunks with A_A=0.2 and A_B=0.3



Example: s=0.1, two chunks with A_A=0.2 and A_B=0.3

retrieval threshold

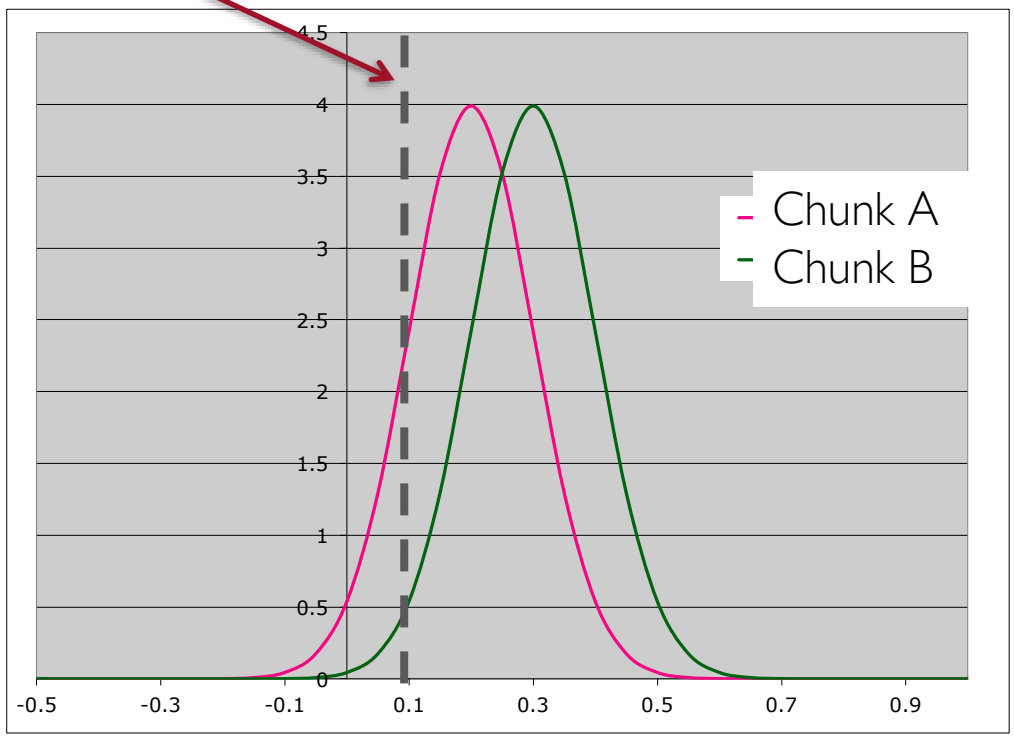


activation

Which chunk is more likely to be retrieved?

Example: s=0.1, two chunks with A_A=0.2 and A_B=0.3

retrieval threshold



activation

Which chunk is more likely to be retrieved?

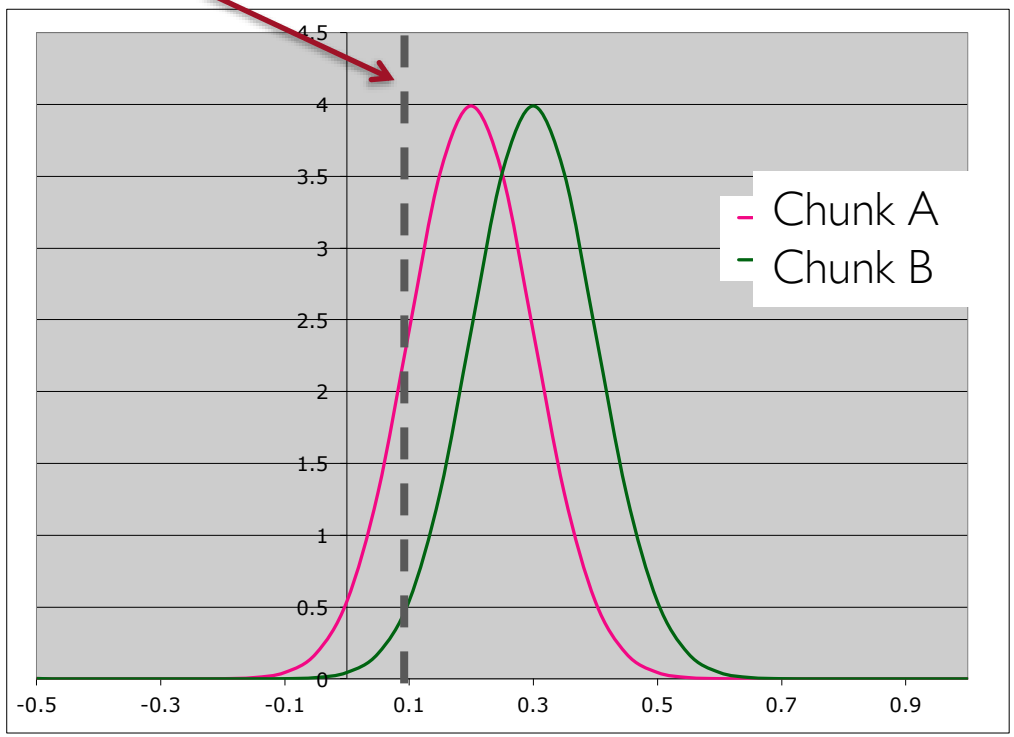
P(A)=0.31

P(B)=0.51

P(failure)=0.19

Example: s=0.1, two chunks with A_A=0.2 and A_B=0.3

retrieval threshold



activation

Which chunk is more likely to be retrieved?

P(A)=0.31
P(B)=0.51
P(failure)=0.19

How does noise affect the shape of these curves?

Parameters

Parameters

$$A_i = B_i + e$$

Parameters

$$A_i = B_i + e$$

- Enable subsymbolic computation :esc nil or t

Parameters

:bll

nil or .5

$$A_i = B_i + e$$

- Enable subsymbolic computation :esc nil or t

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

- Enable subsymbolic computation :esc nil or t

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0
- Retrieval treshold :rt #

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right) \qquad e = e_{instantanious} + e_{permanent}$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0
- Retrieval treshold :rt #

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

:ans
nil or >0

:pas
nil or >0

$$e = e_{instantanious} + e_{permanent}$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0
- Retrieval treshold :rt #

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

:ans
nil or >0

:pas
nil or >0

$$e = e_{instantanious} + e_{permanent}$$

$$s^2 = \frac{\rho^2}{3}s^2$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0
- Retrieval treshold :rt #

Parameters

:bll
nil or .5

$$A_i = B_i + e \qquad RT = Fe^{-A}$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

:ans
nil or >0

:pas
nil or >0

$$e = e_{instantanious} + e_{permanent}$$

$$s^2 = \frac{p^2}{3}s^2$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0
- Retrieval treshold :rt #

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

:lf >0

$$RT = Fe^{-A}$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

:ans
nil or >0

:pas
nil or >0

$$e = e_{instantanious} + e_{permanent}$$

$$s^2 = \frac{\rho^2}{3}s^2$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0
- Retrieval treshold :rt #

Parameters

:bll
nil or .5

$$A_i = B_i + e$$

:lf >0

$$RT = Fe^{-A}$$

$$B_i(t) = \ln\left(\sum_{k=1}^{n}(t - t_k)^{-d}\right)$$

:ans
nil or >0

:pas
nil or >0

$$e = e_{instantanious} + e_{permanent}$$

$$s^2 = \frac{p^2}{3}s^2$$

- Enable subsymbolic computation :esc nil or t
- Optimized Learning :ol nil, t, >0
- Retrieval treshold :rt #
- Show all parameters: (sgp)

Today

- Declarative memory: basics
  - The rational basis for declarative memory
  - Declarative memory in ACT-R
  - Example model
- Assignment

Example Model: Paired Associates

Anderson (1981)

Experiment

tent

2

zinc

9

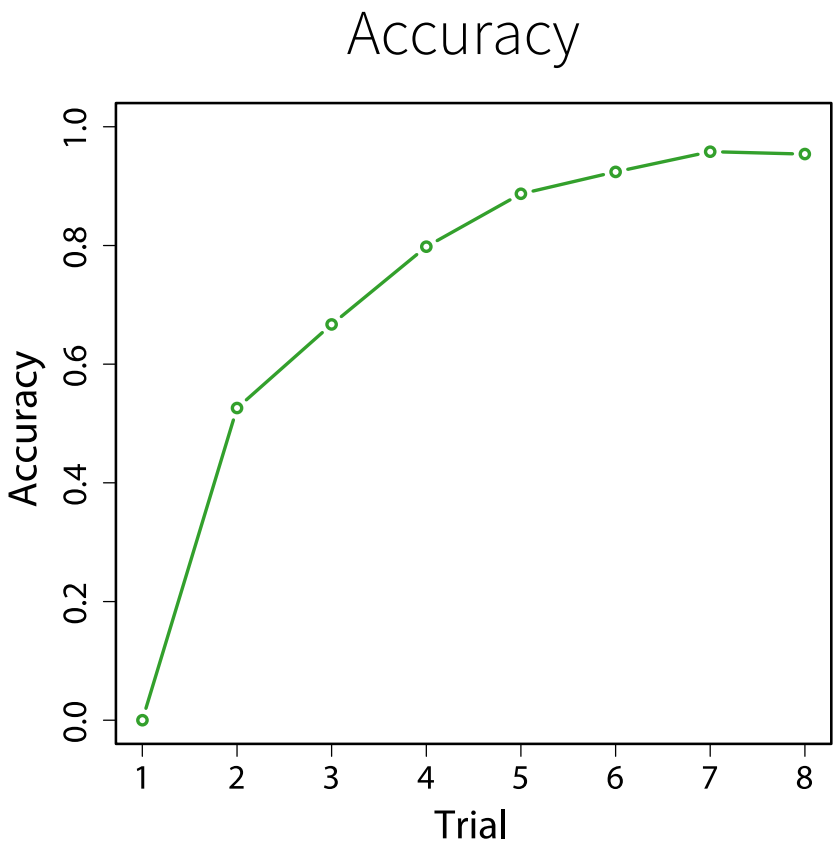
tent

Results

- 20 paired associates, 8 repetitions

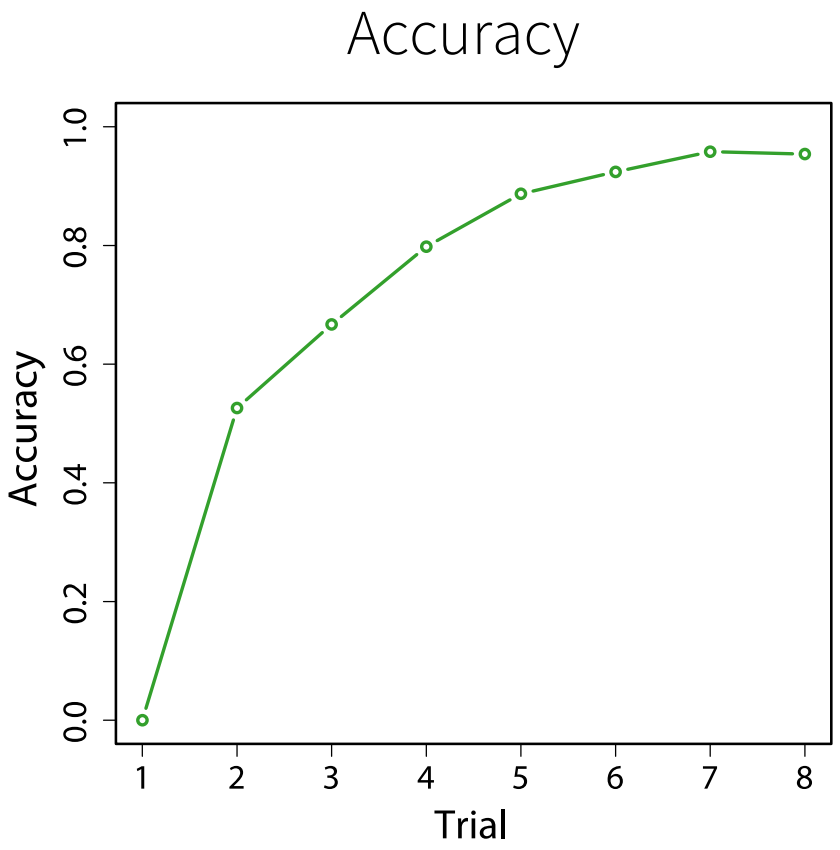
Results

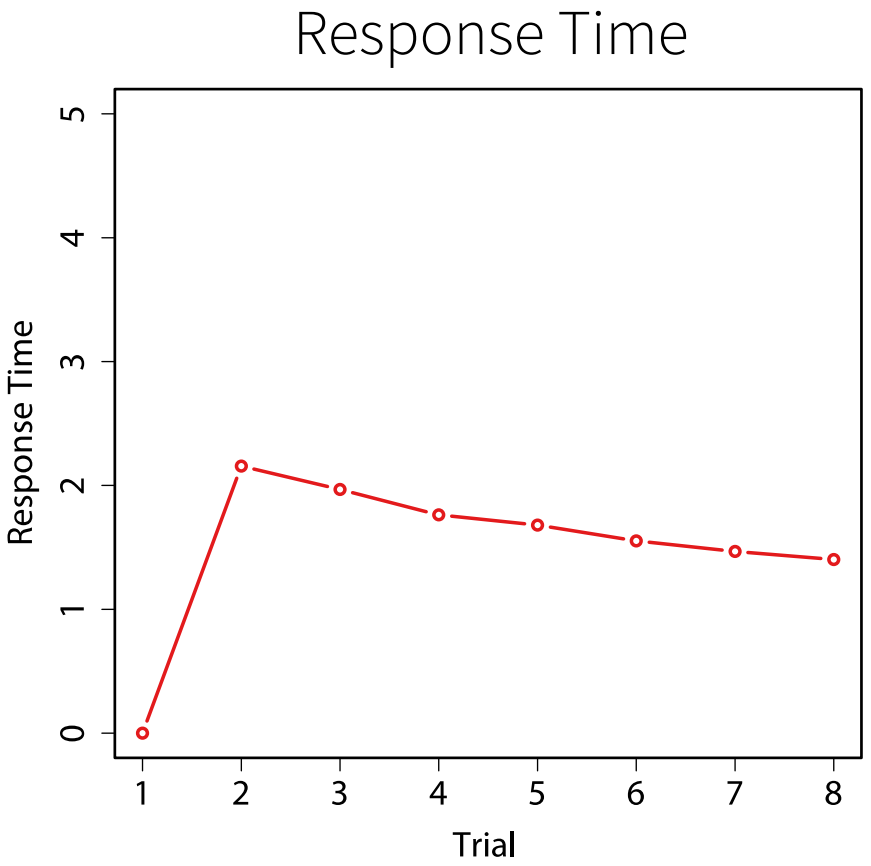
- 20 paired associates, 8 repetitions



Results

- 20 paired associates, 8 repetitions





Memorizing Paired Associates

Memorizing Paired Associates

Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val


    =goal>
      state     testing
)
```

Memorizing Paired Associates

Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val

    =goal>
      state     testing
)
```

Memorizing Paired Associates

Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val

    =goal>
      state     testing
)
```

Number

```
(p associate
    =goal>
      isa       goal
      state     attending-target
    =visual>
      isa       text
      value     =val
   =imaginal>
      isa       pair
==>
   =imaginal>
      answer    =val
   -imaginal>
   =goal>
      state     start
   +visual>
      isa       clear
)
```

Memorizing Paired Associates

Word

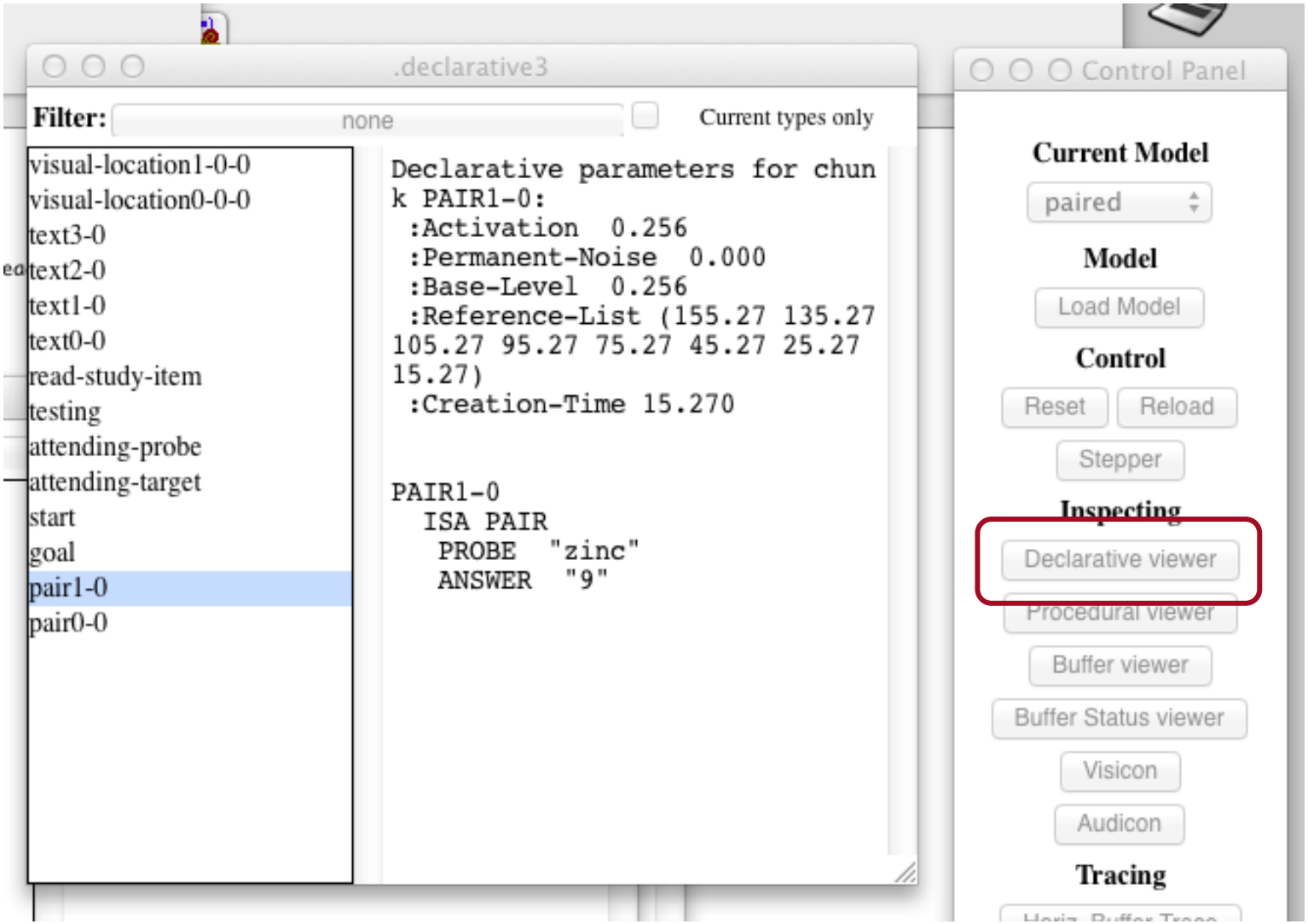
```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val

    =goal>
      state     testing
)
```

Number

```
(p associate
    =goal>
      isa       goal
      state     attending-target
    =visual>
      isa       text
      value     =val
   =imaginal>
      isa       pair
==>
   =imaginal>
      answer    =val
   -imaginal>
   =goal>
      state     start
   +visual>
      isa       clear
)
```

Memorizing Paired Associates



Retrieving Paired Associates

Retrieving Paired Associates

Read Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val
    +retrieval>
      isa       pair
      probe     =val
    =goal>
      state     testing
)
```

Retrieving Paired Associates

Read Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val
    +retrieval>
      isa       pair
      probe     =val
    =goal>
      state     testing
)
```

Retrieving Paired Associates

Read Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val
    +retrieval>
      isa       pair
      probe     =val
    =goal>
      state     testing
)
```

Retrieval

```
(p recall
    =goal>
      isa       goal
      state     testing
    =retrieval>
      isa       pair
      answer    =ans
    ?manual>
      state     free
==>
    +manual>
      isa       press-key
      key       =ans
    =goal>
      state     read-study-item
    +visual>
      isa       clear
)
```

Retrieving Paired Associates

Read Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val
    +retrieval>
      isa       pair
      probe     =val
    =goal>
      state     testing
)
```

Retrieval

```
(p recall
    =goal>
      isa       goal
      state     testing
    =retrieval>
      isa       pair
      answer    =ans
    ?manual>
      state     free
==>
    +manual>
      isa       press-key
      key       =ans
    =goal>
      state     read-study-item
    +visual>
      isa       clear
)
```

Retrieving Paired Associates

Read Word

```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val
    +retrieval>
      isa       pair
      probe     =val
    =goal>
      state     testing
)
```

Retrieval Error

```
(p cannot-recall
    =goal>
      isa       goal
      state     testing
    ?retrieval>
      state     error
   ==>
    =goal>
      state     read-study-item
    +visual>
      isa       clear
)
```

Retrieving Paired Associates

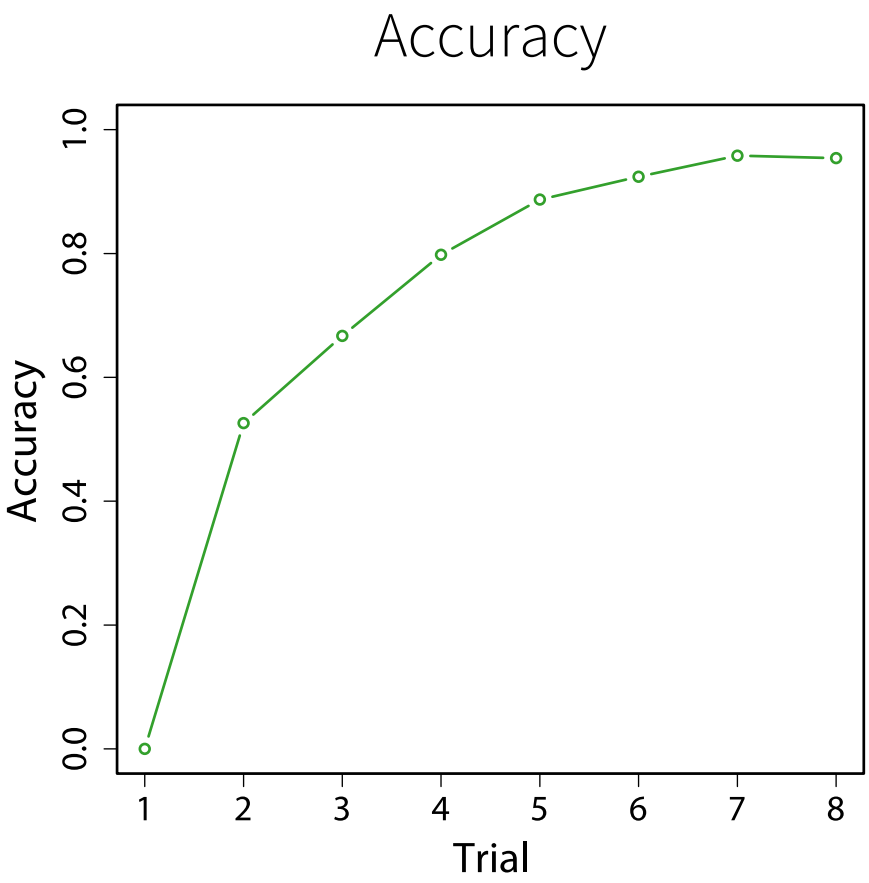
Read Word

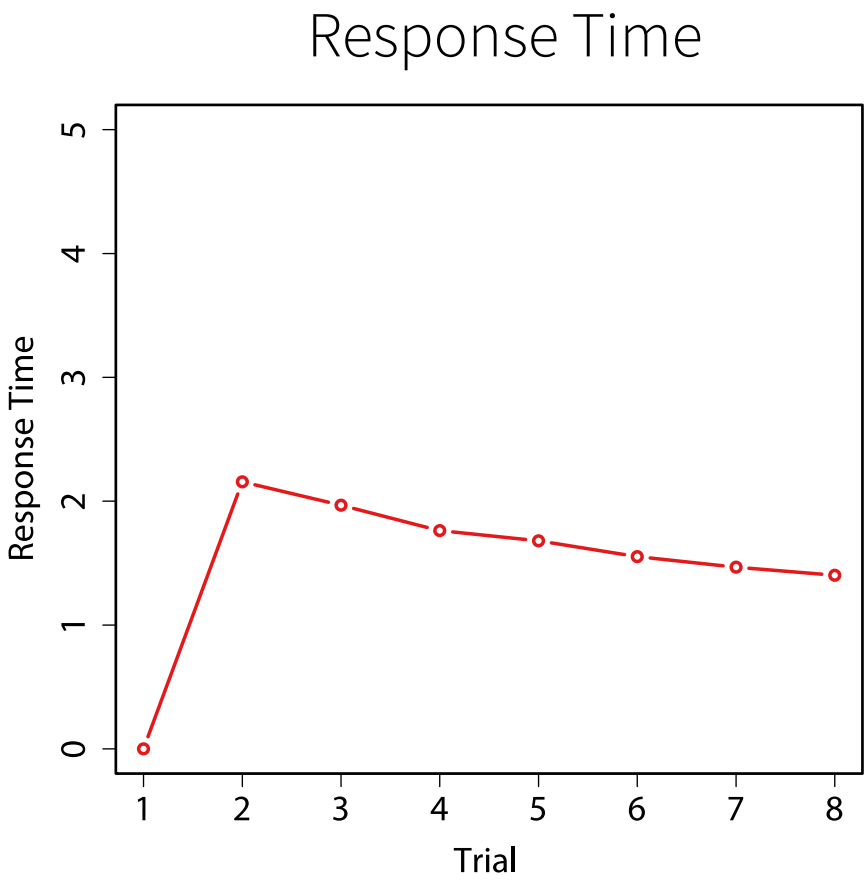
```
(p read-probe
    =goal>
      isa       goal
      state     attending-probe
    =visual>
      isa       text
      value     =val
==>
    +imaginal>
      isa       pair
      probe     =val
    +retrieval>
      isa       pair
      probe     =val
    =goal>
      state     testing
)
```

Retrieval Error

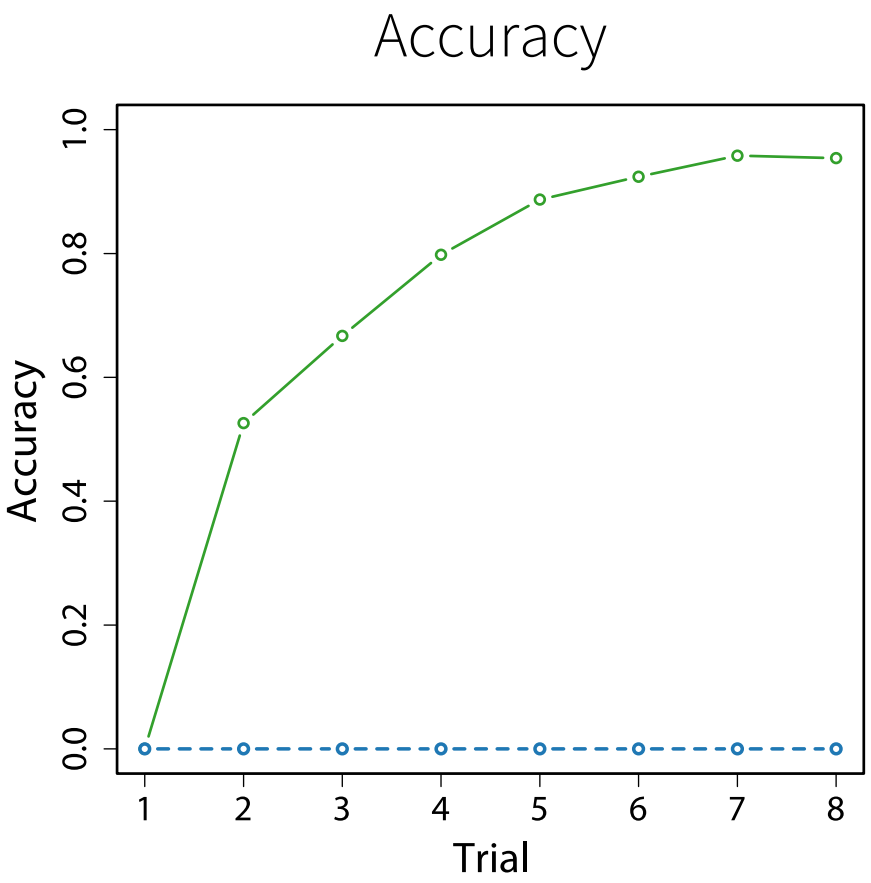
```
(p cannot-recall
    =goal>
      isa       goal
      state     testing
    ?retrieval>
      state     error
   ==>
    =goal>
      state     read-study-item
    +visual>
      isa       clear
)
```

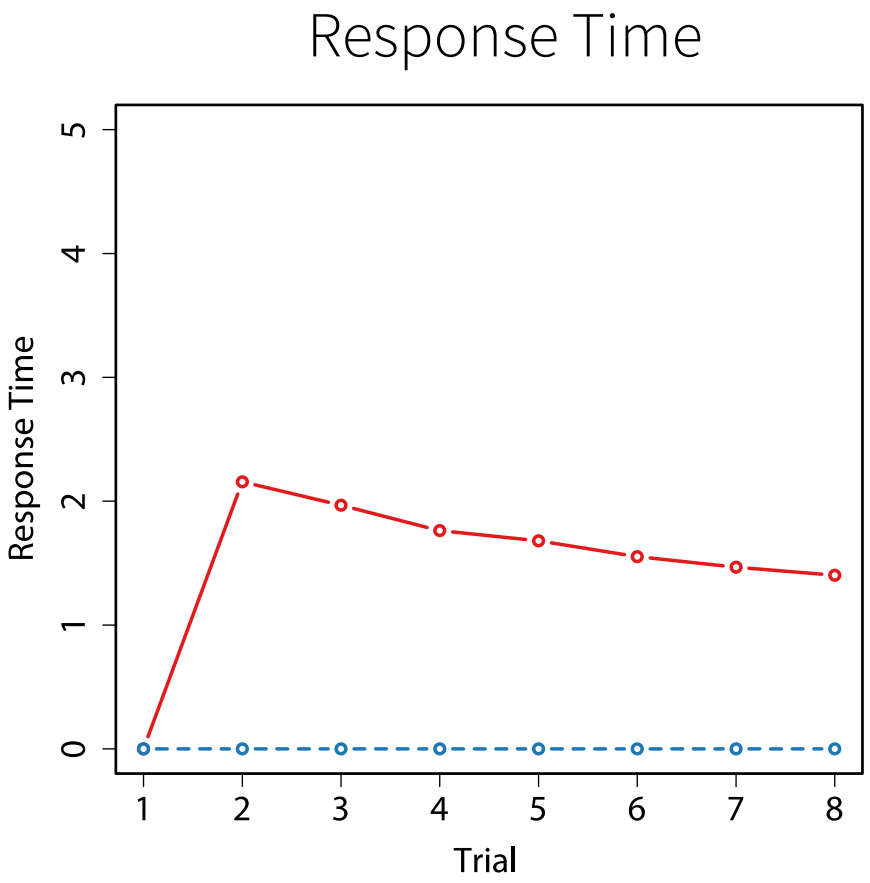
Fitting the Model



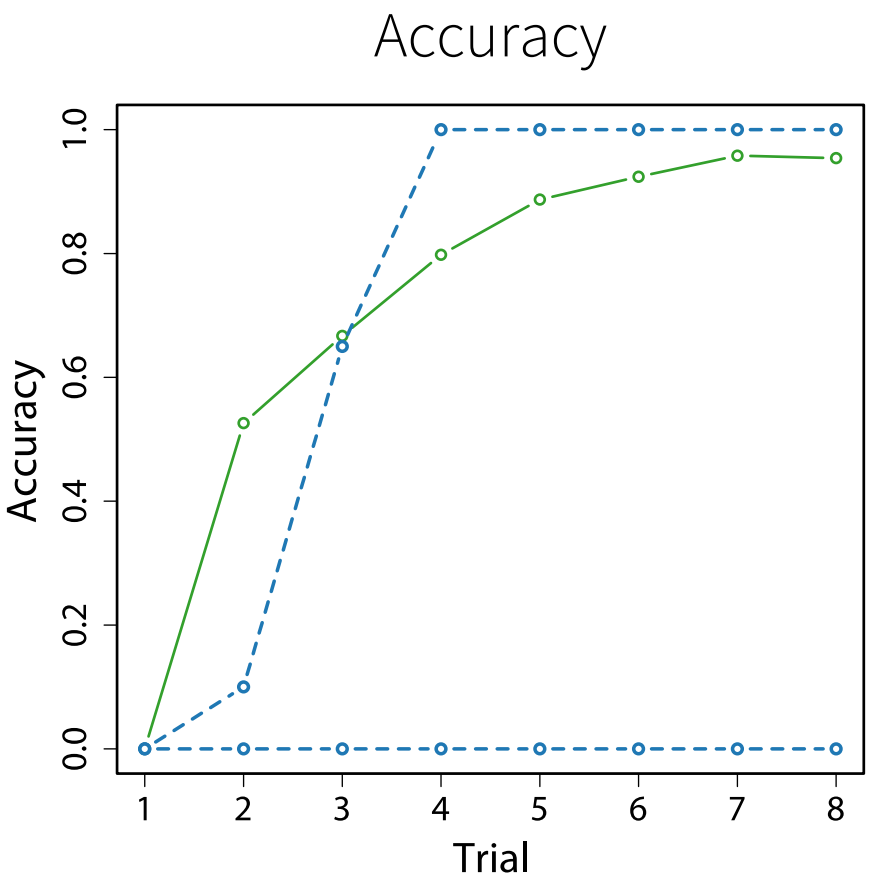


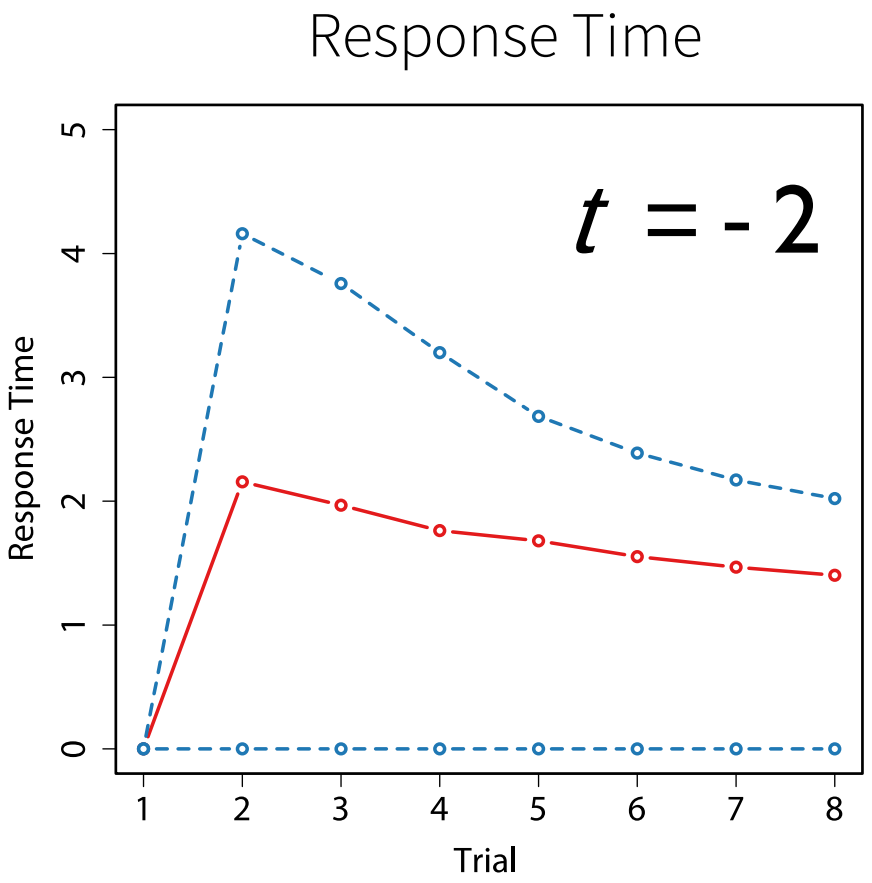
Fitting the Model



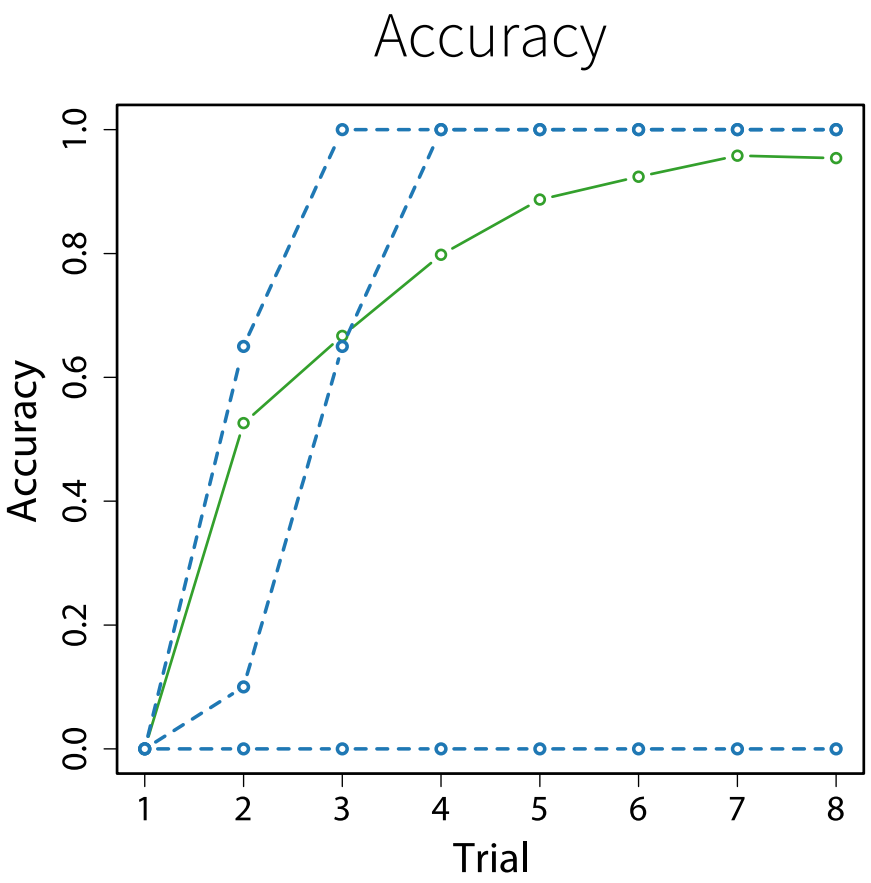


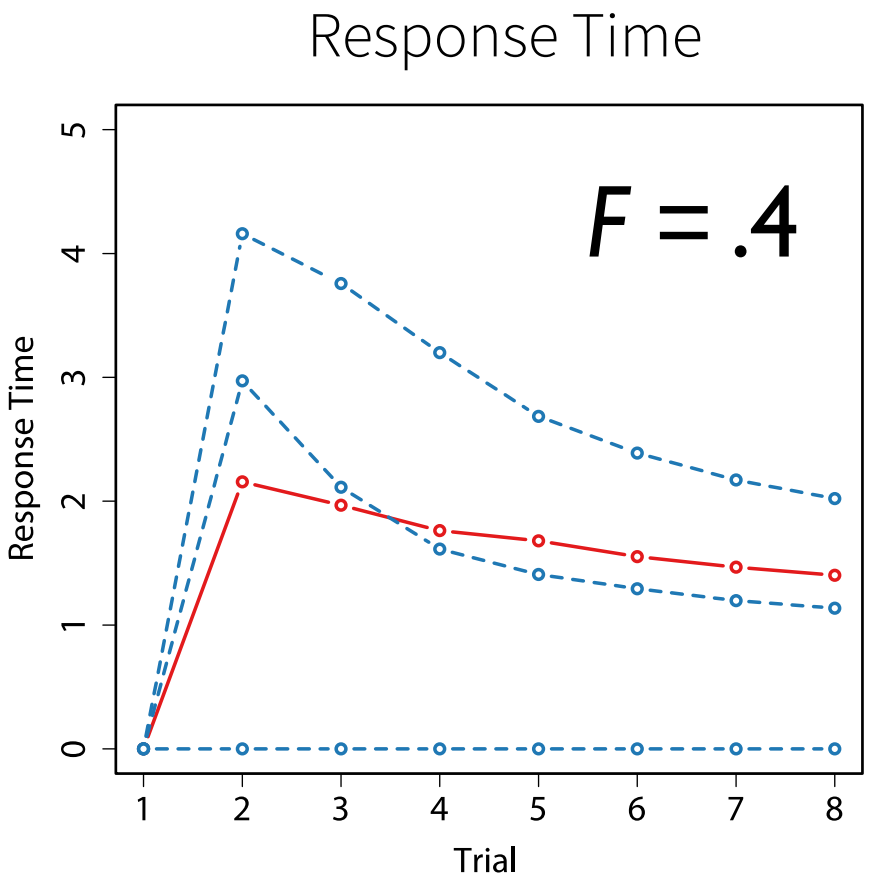
Fitting the Model



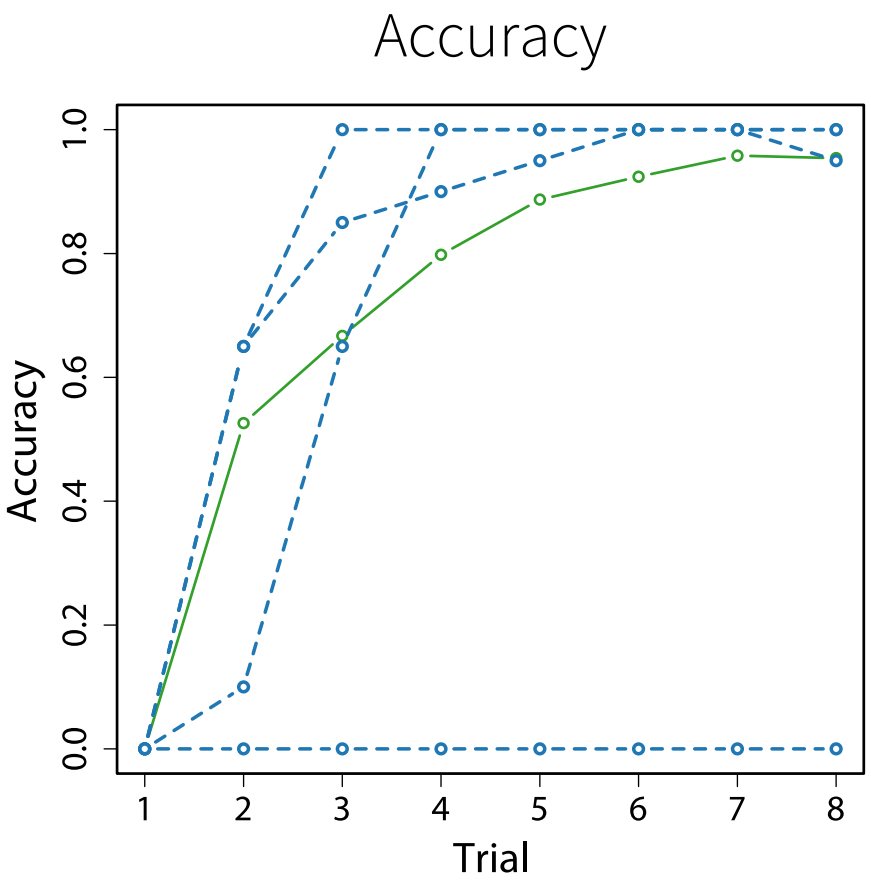


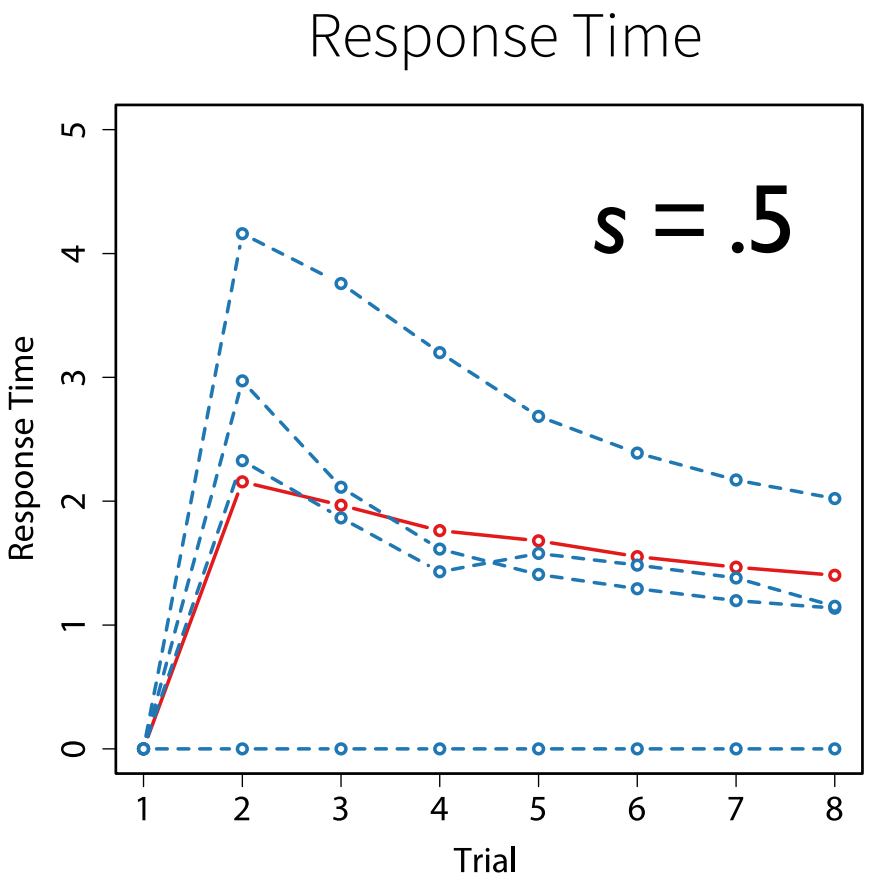
Fitting the Model





Fitting the Model





Fitting the Model

Fitting the Model

- Activation trace: :act t

Fitting the Model

- Activation trace: :act t

```
     0.185   DECLARATIVE            START-RETRIEVAL
No matching chunk found retrieval failure
     0.185   PROCEDURAL             CONFLICT-RESOLUTION
     ...
     3.141   DECLARATIVE            RETRIEVAL-FAILURE
     ...
    10.185   DECLARATIVE            START-RETRIEVAL
Chunk PAIR0-0 matches
Computing activation for chunk PAIR0-0
Computing base-level
Starting with blc: 0.0
Computing base-level from 1 references (5.185)
  creation time: 5.185 decay: 0.5  Optimized-learning: T
base-level value: -0.111571856
Total base-level: -0.111571856
Adding transient noise -0.7636829
Adding permanent noise 0.0
Chunk PAIR0-0 has an activation of: -0.87525475
Chunk PAIR0-0 has the current best activation -0.87525475
Chunk PAIR0-0 with activation -0.87525475 is the best
     ...
    11.145   DECLARATIVE            RETRIEVED-CHUNK PAIR0-0
```

Fitting the Model

- Activation trace: :act t

```
      0.185   DECLARATIVE             START-RETRIEVAL
No matching chunk found retrieval failure
      0.185   PROCEDURAL              CONFLICT-RESOLUTION
      ...
      3.141   DECLARATIVE             RETRIEVAL-FAILURE
      ...
     10.185   DECLARATIVE             START-RETRIEVAL
Chunk PAIR0-0 matches
Computing activation for chunk PAIR0-0
Computing base-level
Starting with blc: 0.0
Computing base-level from 1 references (5.185)
  creation time: 5.185 decay: 0.5  Optimized-learning: T
base-level value: -0.111571856
Total base-level: -0.111571856
Adding transient noise -0.7636829
Adding permanent noise 0.0
Chunk PAIR0-0 has an activation of: -0.87525475
Chunk PAIR0-0 has the current best activation -0.87525475
Chunk PAIR0-0 with activation -0.87525475 is the best
      ...
     11.145   DECLARATIVE             RETRIEVED-CHUNK PAIR0-0
```

Fitting the Model

- Activation trace: :act t

```
     0.185   DECLARATIVE             START-RETRIEVAL
No matching chunk found retrieval failure
     0.185   PROCEDURAL              CONFLICT-RESOLUTION
     ...
     3.141   DECLARATIVE             RETRIEVAL-FAILURE
     ...
    10.185   DECLARATIVE             START-RETRIEVAL
Chunk PAIR0-0 matches
Computing activation for chunk PAIR0-0
Computing base-level
Starting with blc: 0.0
Computing base-level from 1 references (5.185)
  creation time: 5.185 decay: 0.5  Optimized-learning: T
base-level value: -0.111571856
Total base-level: -0.111571856
Adding transient noise -0.7636829
Adding permanent noise 0.0
Chunk PAIR0-0 has an activation of: -0.87525475
Chunk PAIR0-0 has the current best activation -0.87525475
Chunk PAIR0-0 with activation -0.87525475 is the best
     ...
    11.145   DECLARATIVE             RETRIEVED-CHUNK PAIR0-0
```

Today

- Declarative memory: basics
  - The rational basis for declarative memory
  - Declarative memory in ACT-R
  - Example model

- Assignment

B + 2 = D

$$C + 4 = F$$

Alphabet Arithmetic

- Verify equations with letters, e.g.

A + 3 = D ✔

C + 2 = F ✗

(Zbrodoff, 1995)

Alphabet Arithmetic

- Verify equations with letters, e.g.

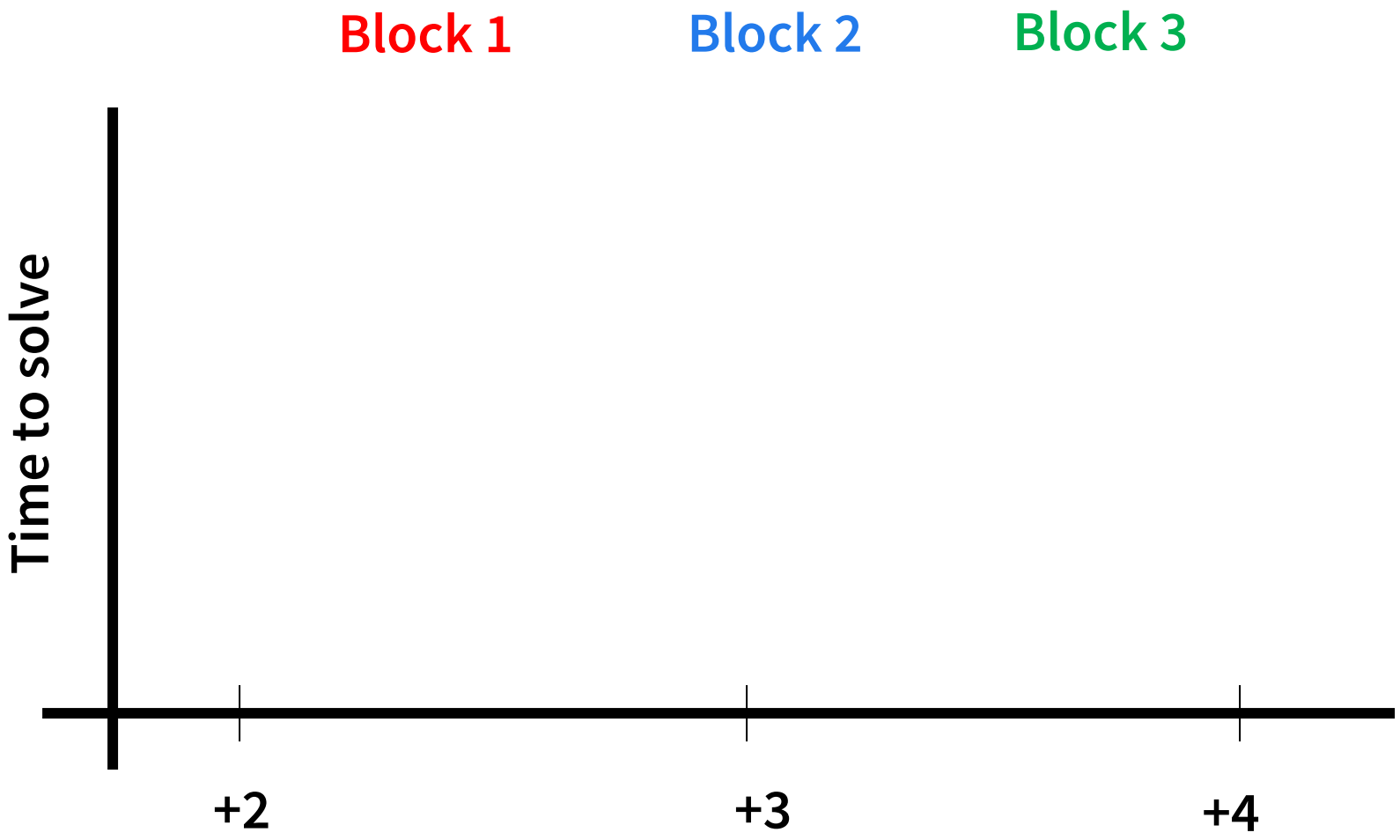
A + 3 = D ✔

C + 2 = F ✗

- Participants solved 3 blocks of 192 problems each:
  - 8 repetitions of each problem in a block
  - total of 576 trials

(Zbrodoff, 1995)

Predict the outcome



How can we solve those problems?

How can we solve those problems?

- Verify answer through counting

How can we solve those problems?

- Verify answer through counting

- Verify answer through memory retrieval

How can we solve those problems?

- Verify answer through counting
  - Is slow, but always works

- Verify answer through memory retrieval

How can we solve those problems?

- Verify answer through counting
  - Is slow, but always works
  - Takes more time as the addend grows (+2, +3, +4)

- Verify answer through memory retrieval

How can we solve those problems?

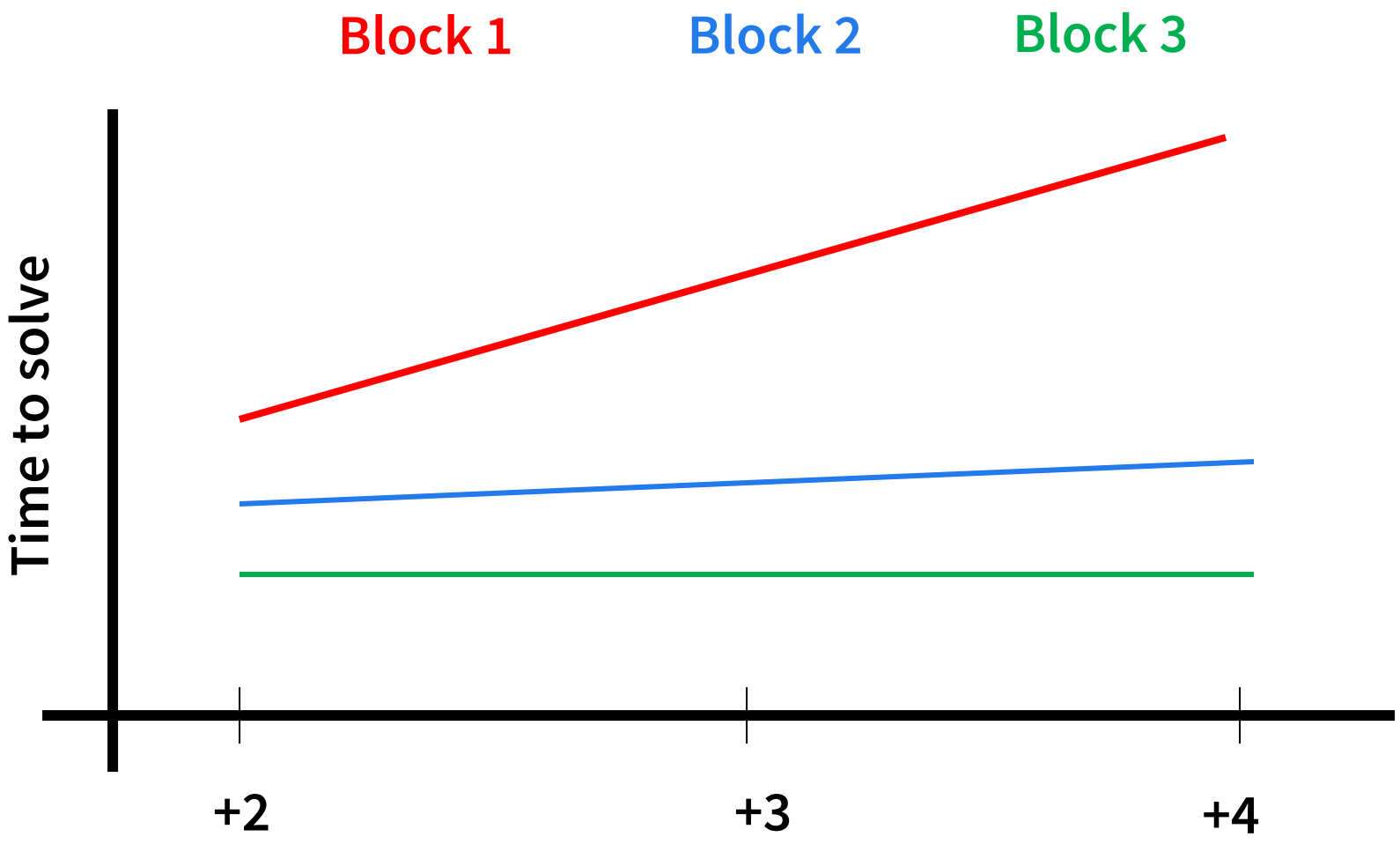
- Verify answer through counting
  - Is slow, but always works
  - Takes more time as the addend grows (+2, +3, +4)
- Verify answer through memory retrieval
  - Is fast, but requires a fact in memory

How can we solve those problems?

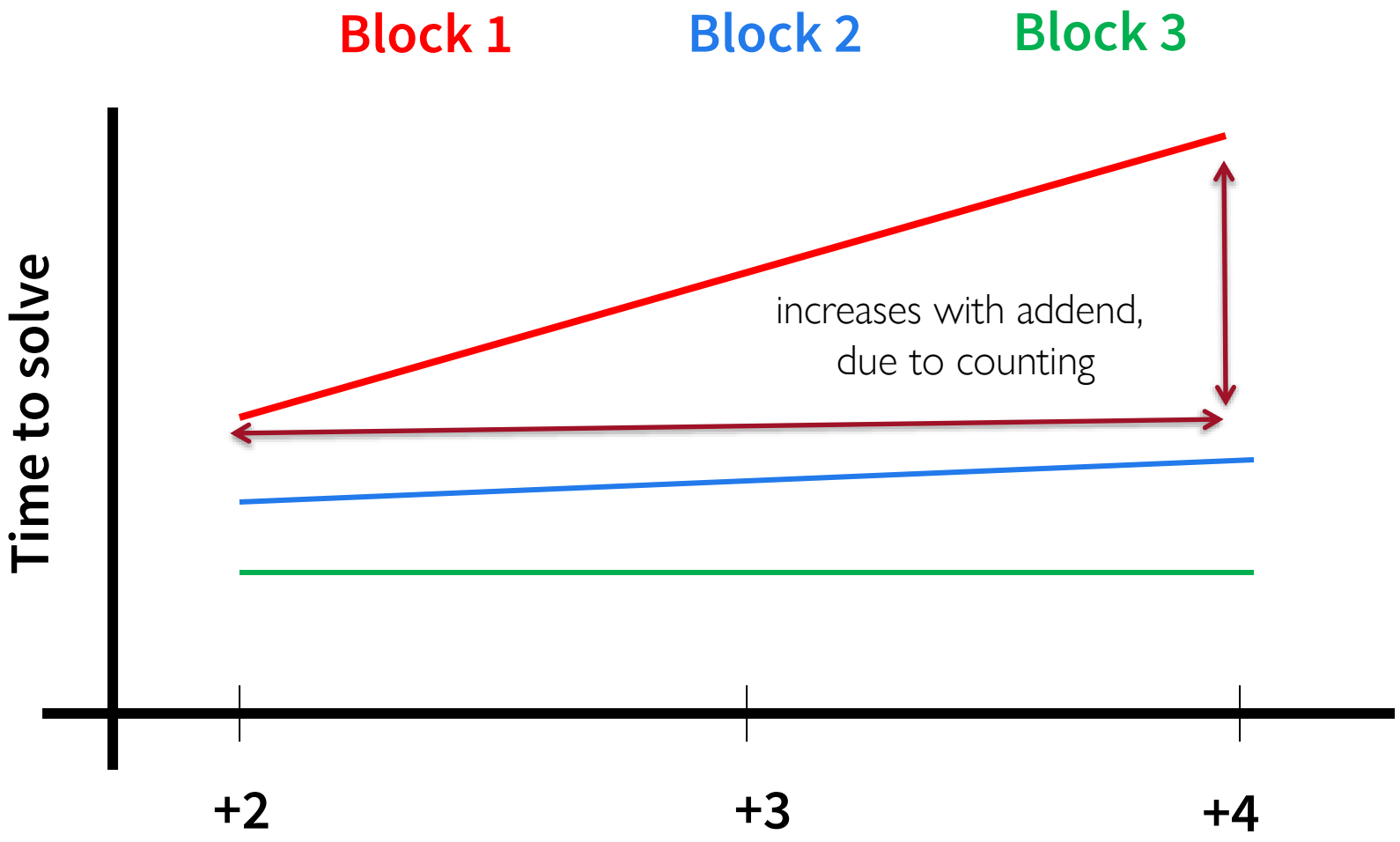
- Verify answer through counting
  - Is slow, but always works
  - Takes more time as the addend grows (+2, +3, +4)

- Verify answer through memory retrieval
  - Is fast, but requires a fact in memory
  - Speed is based on prior frequency, and not on size of the addend

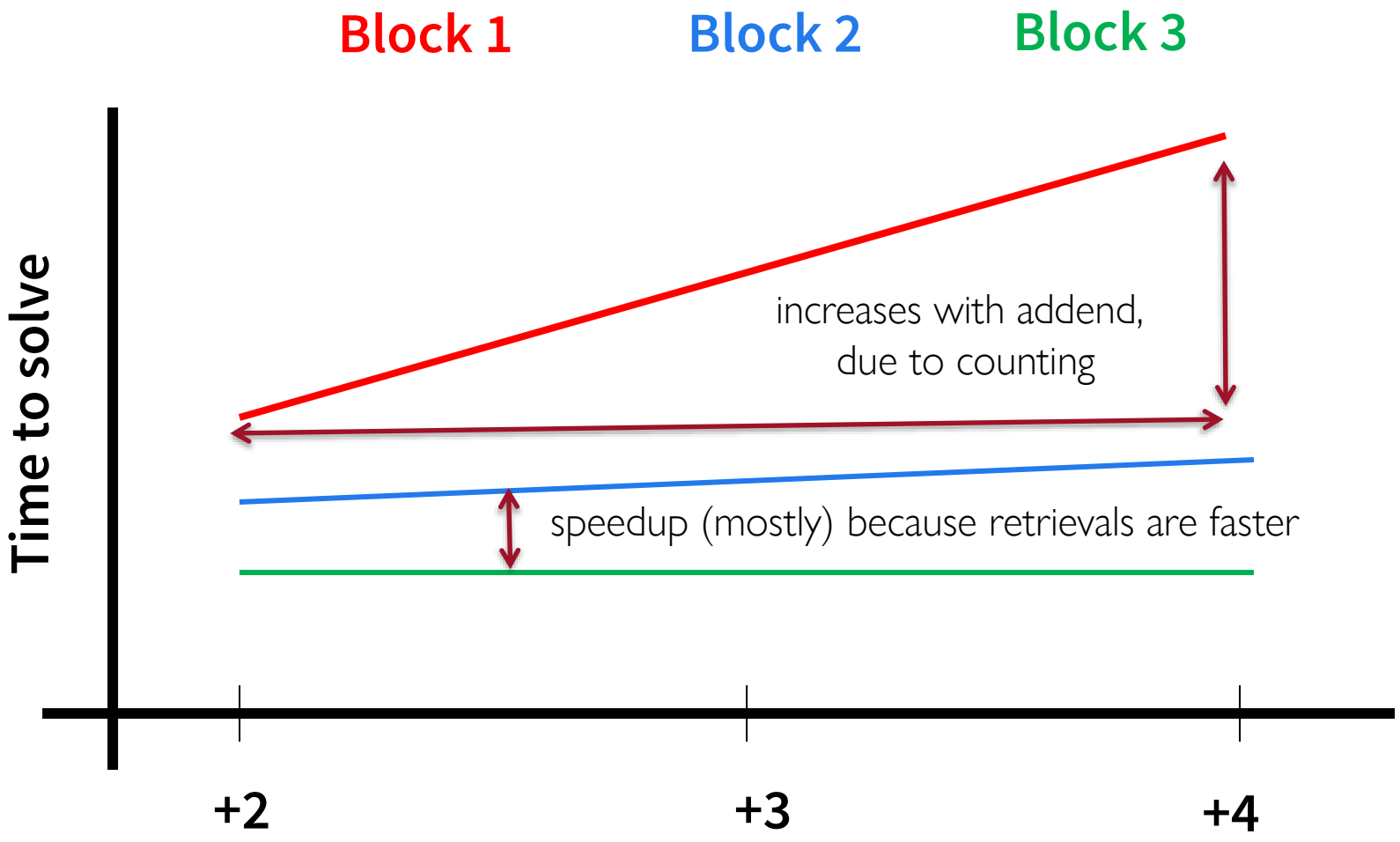
Predict the outcome



Predict the outcome



Predict the outcome



The Assignment

The Assignment

- Unit 4 has a model that implements the counting strategy

The Assignment

- Unit 4 has a model that implements the counting strategy
- Your task:
  - Extend the model with the retrieval strategy
    - To run the experiment call (zbrodoff n)
    - To run a single trial call e.g., (zbrodoff-trial "a" "2" "c" "k" t)

The Assignment

- Unit 4 has a model that implements the counting strategy
- Your task:
  - Extend the model with the retrieval strategy
    - To run the experiment call (zbrodoff n)
    - To run a single trial call e.g., (zbrodoff-trial "a" "2" "c" "k" t)
  - Adjust parameters to fit the human data

The Assignment

- Unit 4 has a model that implements the counting strategy
- Your task:
  - Extend the model with the retrieval strategy
    - To run the experiment call (zbrodoff n)
    - To run a single trial call e.g., (zbrodoff-trial "a" "2" "9" "k" t)
  - Adjust parameters to fit the human data

But first make sure that your model works!

The Assignment

- Unit 4 has a model that implements the counting strategy
- Your task:
  - Extend the model with the retrieval strategy
    - To run the experiment call (zbrodoff n)
    - To run a single trial call e.g., (zbrodoff-trial "a" "2" "b" 'k 't)
  - Adjust parameters to fit the human data

But first make sure that your model works!

→How precisely should your model fit?

The Assignment

- Unit 4 has a model that implements the counting strategy
- Your task:
  - Extend the model with the retrieval strategy
    - To run the experiment call (zbrodoff n)
    - To run a single trial call e.g., (zbrodoff-trial "a" "2" "a" "4" t)
  - Adjust parameters to fit the human data

But first make sure that your model works!

→How precisely should your model fit?

→ Aim at getting the correlation above .98 and the deviation below .160

The Assignment

- Unit 4 has a model that implements the counting strategy
- Your task:
  - Extend the model with the retrieval strategy
    - To run the experiment call (zbrodoff n)
    - To run a single trial call e.g., (zbrodoff-trial "a" "2" "a" "2" nil t)
  - Adjust parameters to fit the human data

given the noise, use n =5 when fitting

But first make sure that your model works!

→How precisely should your model fit?

→ Aim at getting the correlation above .98 and the deviation below .160

Today

- Declarative memory: basics
  - The rational basis for declarative memory
  - Declarative memory in ACT-R
  - Example model

- Assignment

Good luck with the assignment!

university of groningen