Unit 2: Bayesian Learning

3. Rational Analysis

10/13/2020

Rational analysis

1. Rational analysis is a framework theory for modeling learning and cognition

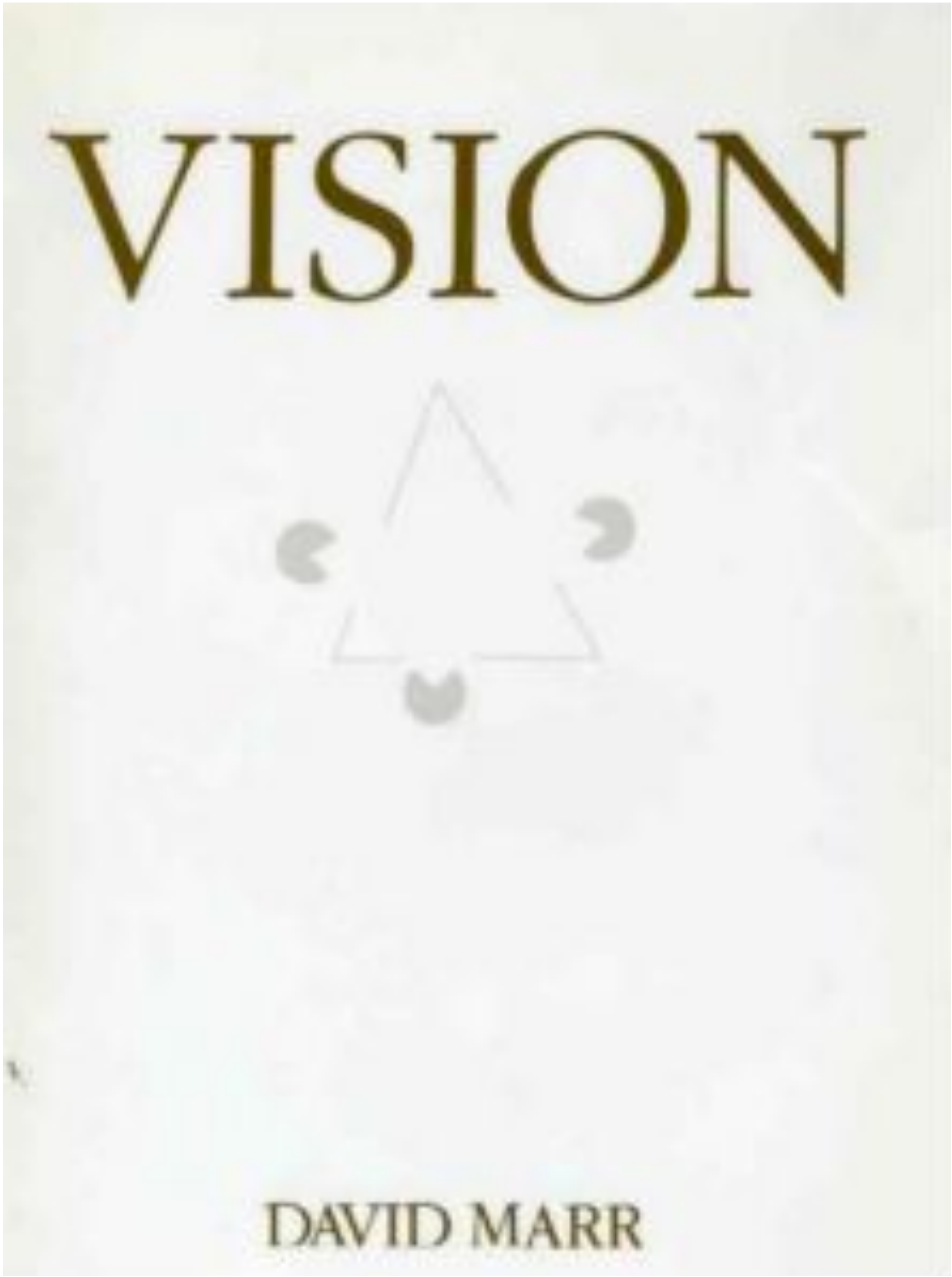
2. Memory retrieval can be modeled as optimal search

3. People track surprisingly precise frequency distributions

4. Rational analysis relies on characterizing the information in the environment

Marr’s levels of analysis

David Marr



Marr’s levels of analysis

Computational Theory

What is the goal of the computation? What is the logic of the strategy by which it can be carried out?

Representation and algorithm

What is the representation for the input and output, and what is the algorithm for the transformation?

Hardware implementation

How can the representation and algorithm be realized physically?

Marr (1982)

An example: The cash register

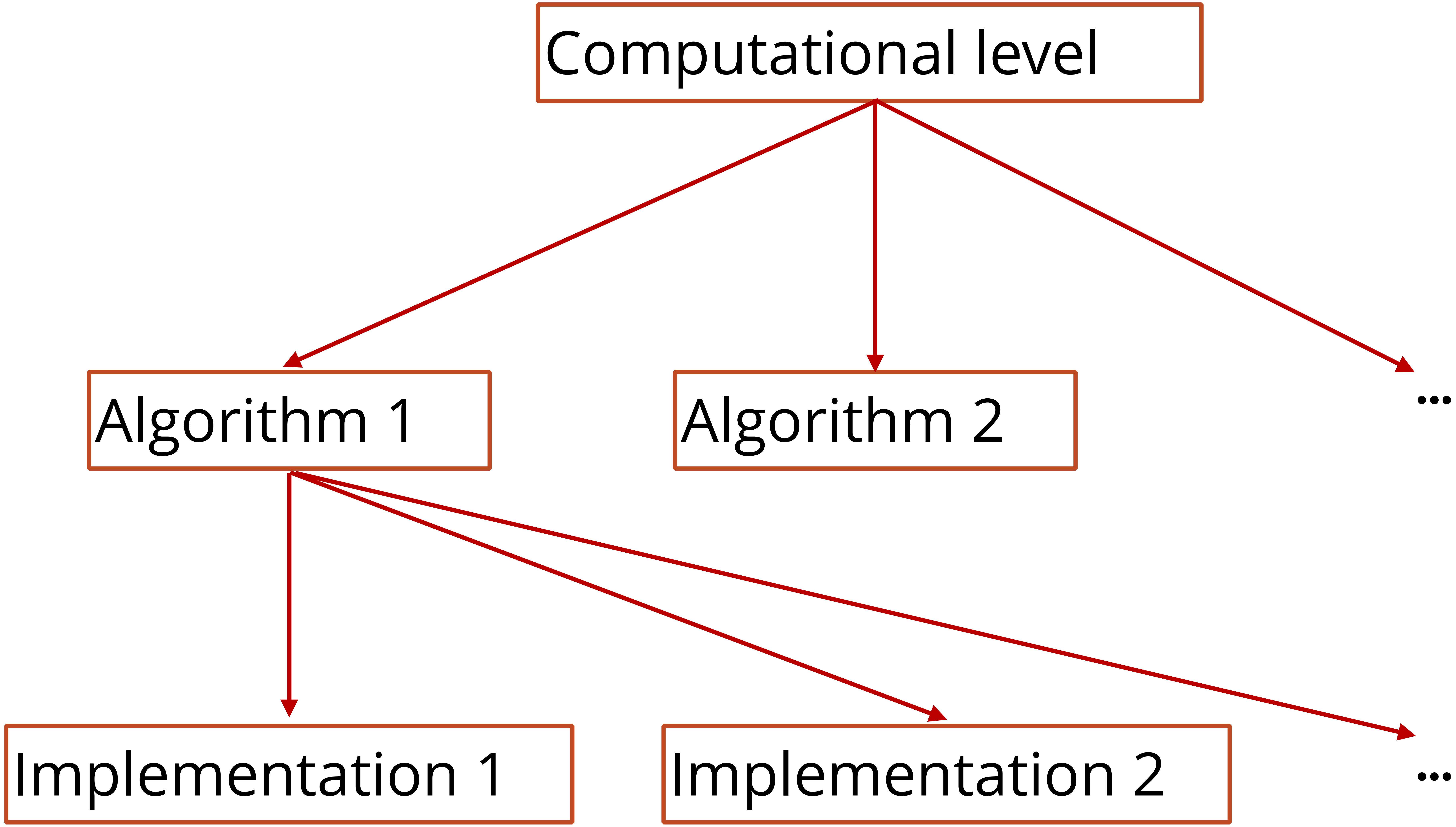
Computational level
Calculates sum of numbers using the theory of addition
So, it will be e.g commutative (order doesn't matter)

Algorithmic level
Uses fixed-point approximations and Arabic numerals
So, if numbers too big, it will fail (unlike theory of addition)

Implementation level
Uses physical buttons and gears
So, things can fail if these break.




Each level is multiply realizable in the next lower level



Constraints across levels

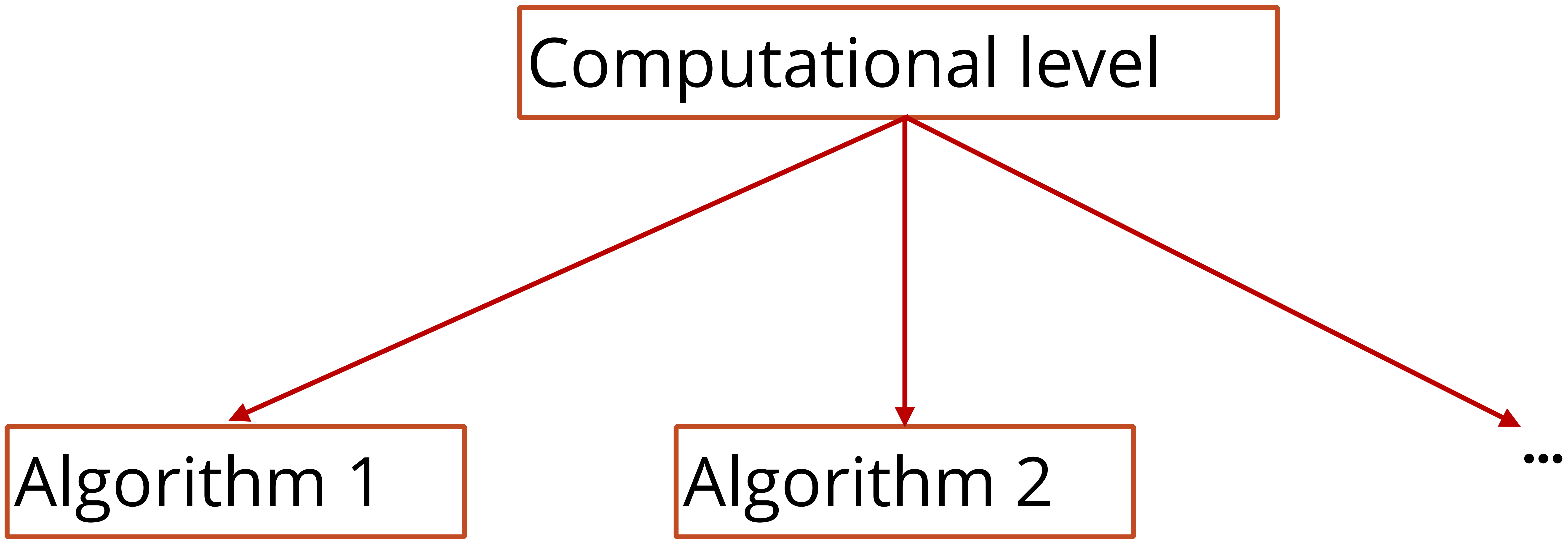
Every level constrains every other level.

But these constraints are *asymmetric*.
Higher levels constrain lower levels more.

So, you get more bang for your buck by working on higher levels.

Functionalism: Let’s not worry about implementation—forget about the brain.

Cognitive psychology (approximately)



Which algorithm do people use?

Constraints across levels

Every level constrains every other level.

But these constraints are *asymmetric*.
Higher levels constrain lower levels more.

So, you get more bang for your buck by working on higher levels.

Rational analysis:
Let's not worry about mechanism either.

Rational analysis

For a given computational problem, there is an *optimal solution*. Whatever it is, we have evolved to approximate it.

Figure out the optimal solution, and you'll know a lot about what people do.

"The predictions flow from the statistical structure of the environment and **not** the assumed structure of the mind." (Anderson, 1991)

Bayes' rule: The rule for rational analysis

Likelihood
(What the data say)

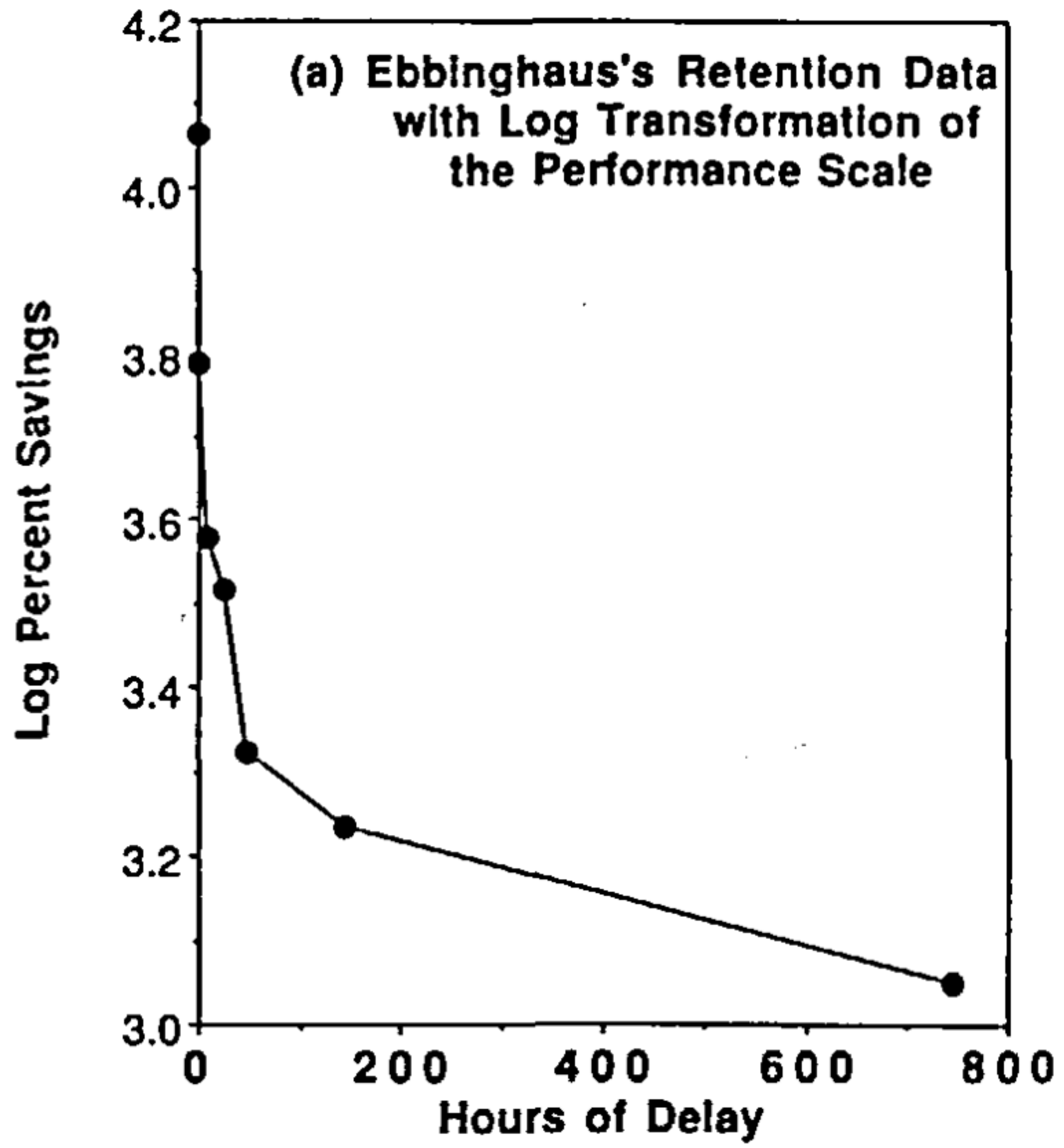
Prior probability
(What you used to believe)

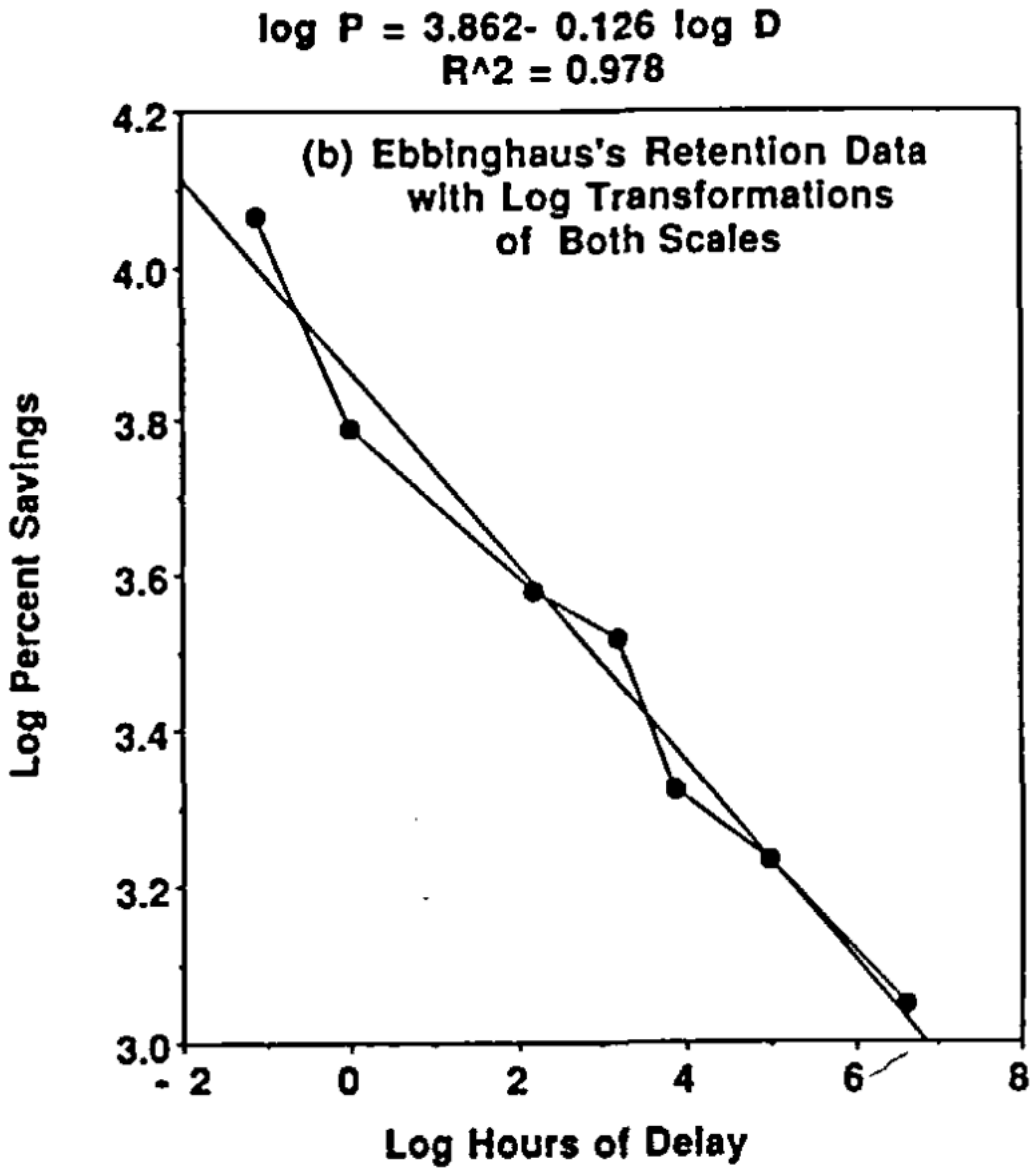
$$P\left(H|D\right) \propto P\left(D|H\right) P(H)$$

Posterior probability
(What you should believe
after seeing the data)

Memory retrieval as Bayesian inference (Anderson & Schooler, 1991)

Experiments on CVC repetition



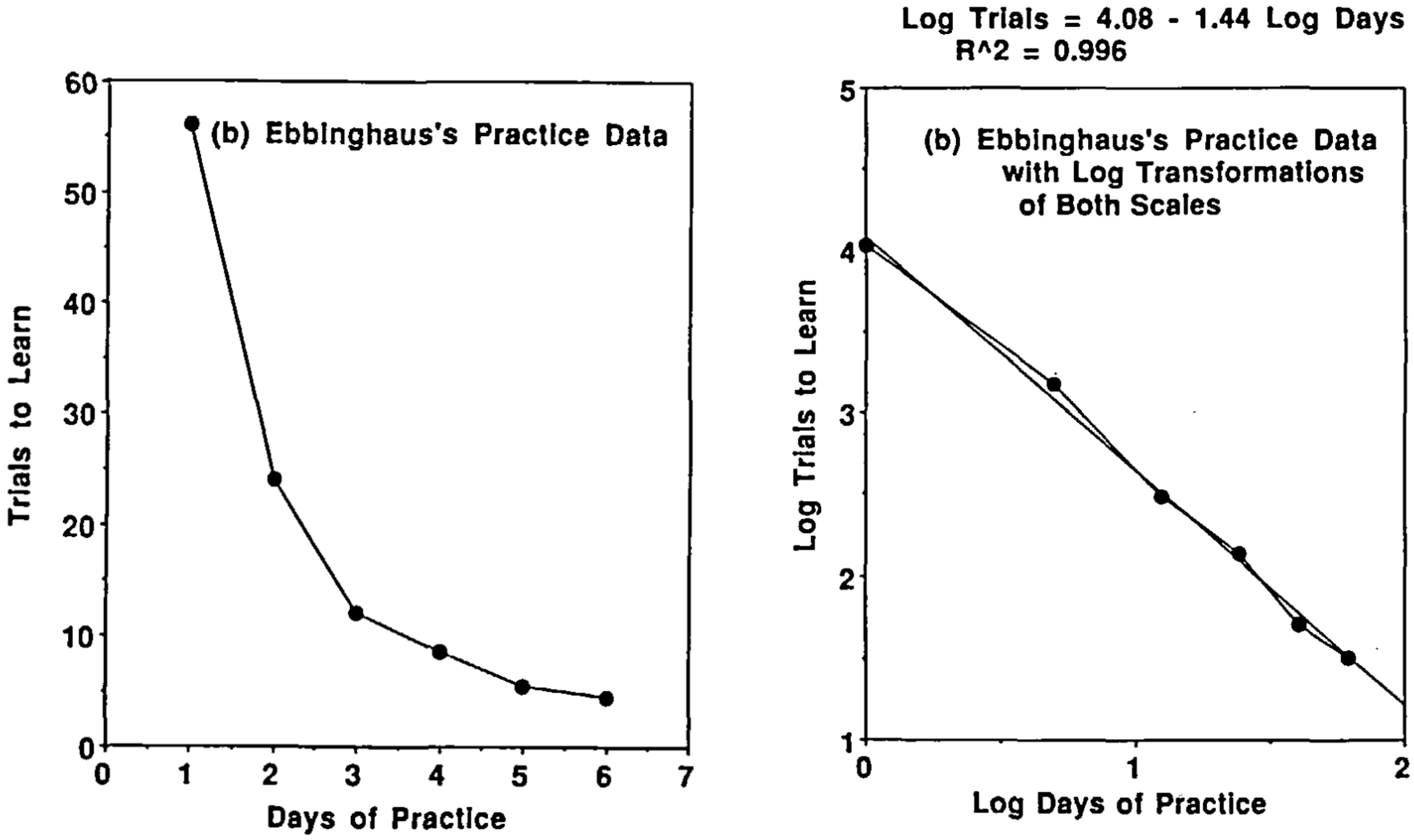


Hermann Ebbinghaus

$$P = AT^{-b}$$

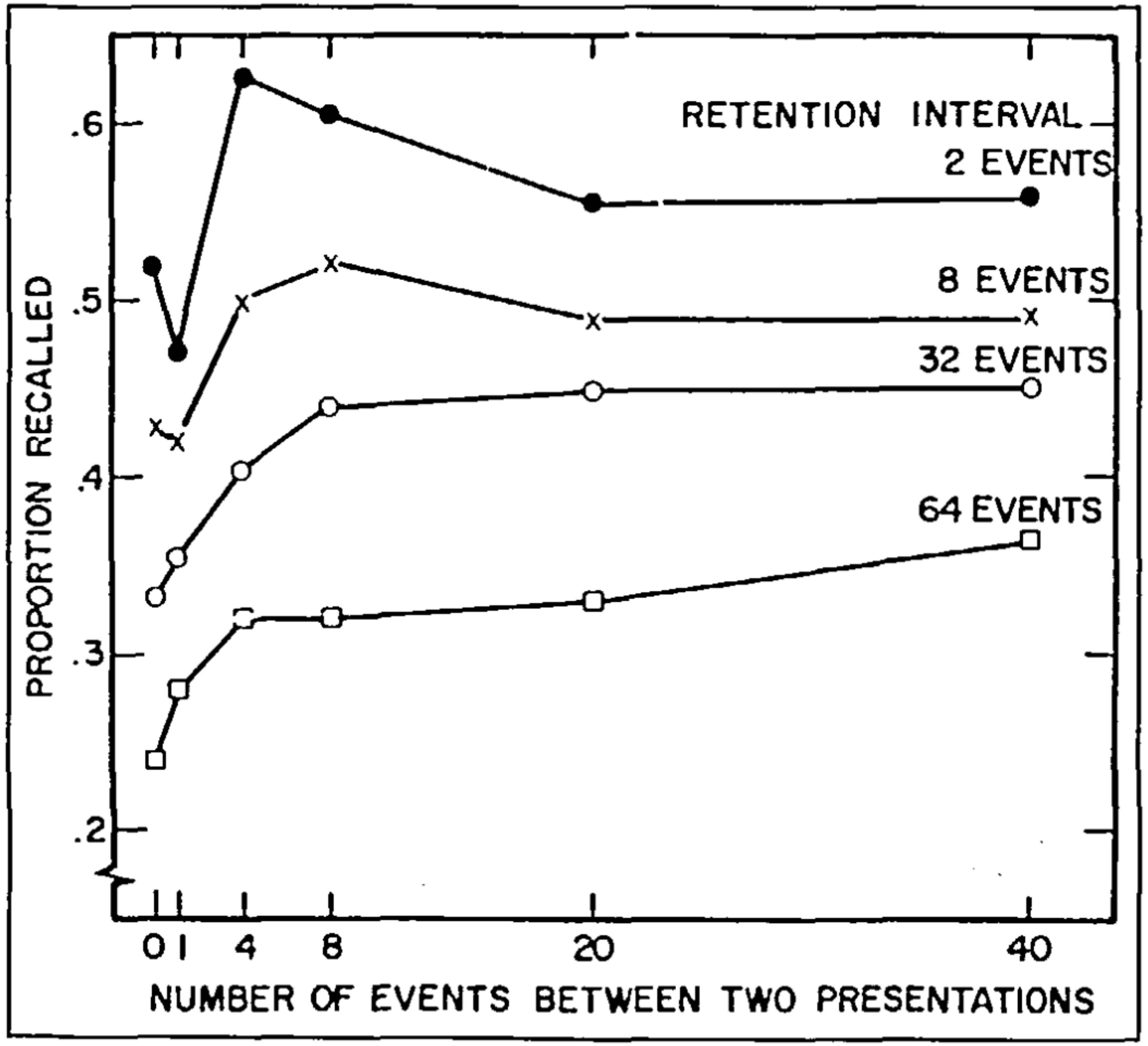
$$\log P = \log A - b \log T$$

Practice effect also follows the power law



Hermann Ebbinghaus

The spacing effect in memory



From Glenberg (1976)

S_1 S_2 T

Spacing between successive repetitions of an item affects how well it is remembered

AND interacts with delay between last study and test

What is happening here?

How would an optimal memory work?

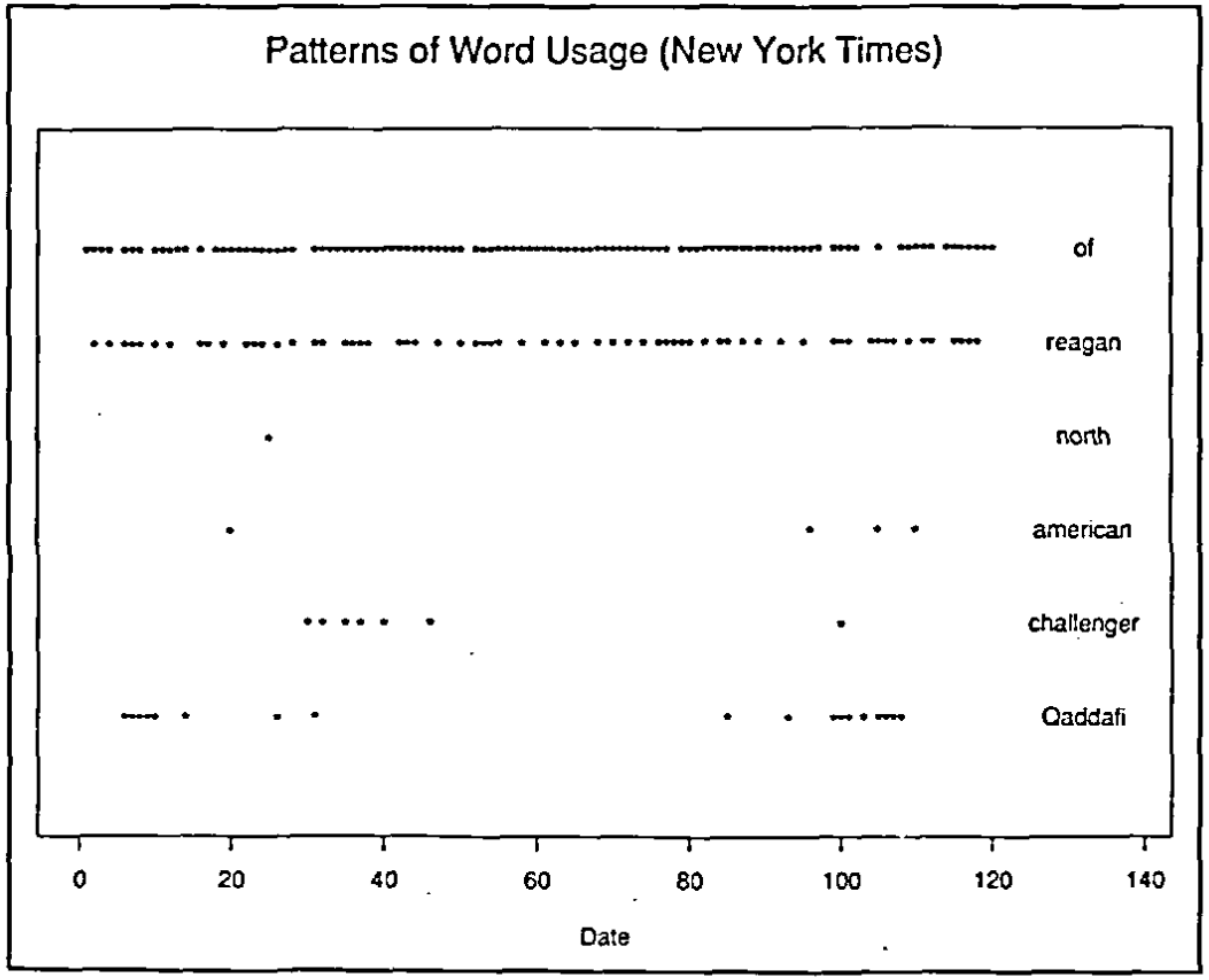
Imagine that memory is a list of entries corresponding to events/items/words/etc you've observed

You have to query memory serially, one thing at a time

Goal: Memories that are likely to be useful next should be kept at the front of the list

$$P(x) \propto H_x$$

What accounts for these effects in practice, retention, and spacing?



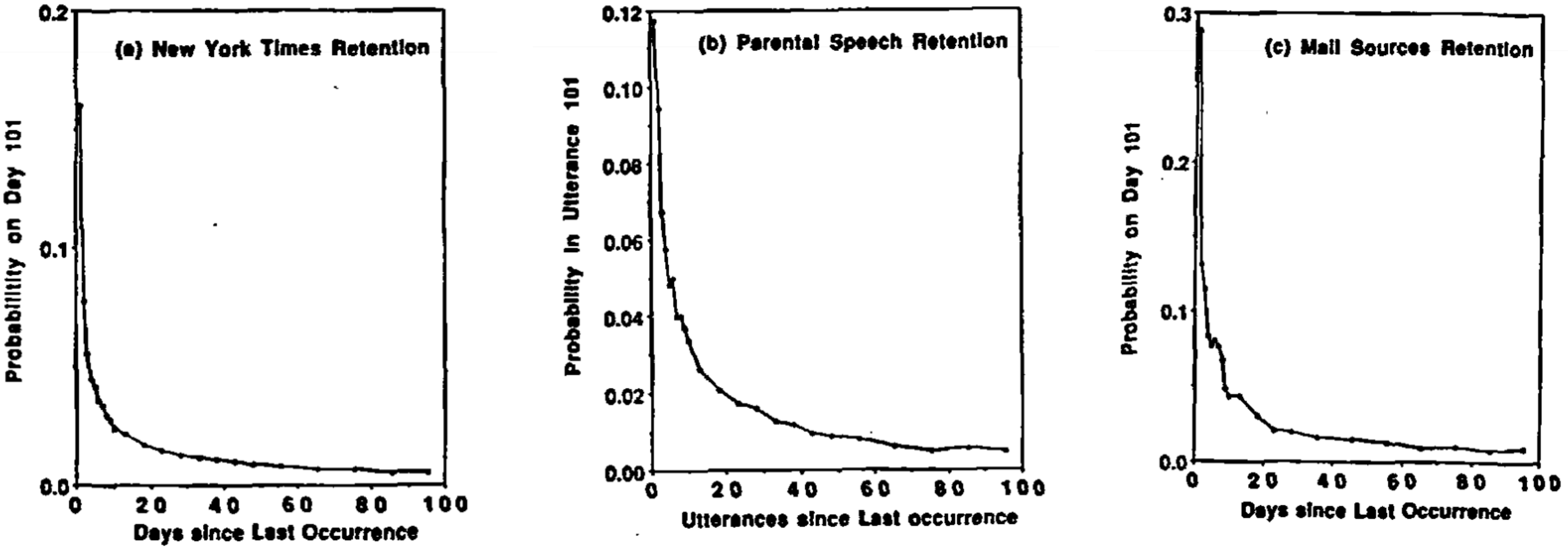
Memory dynamics reflect the statistics of the environment

$$f(x) = ax^{-k}$$

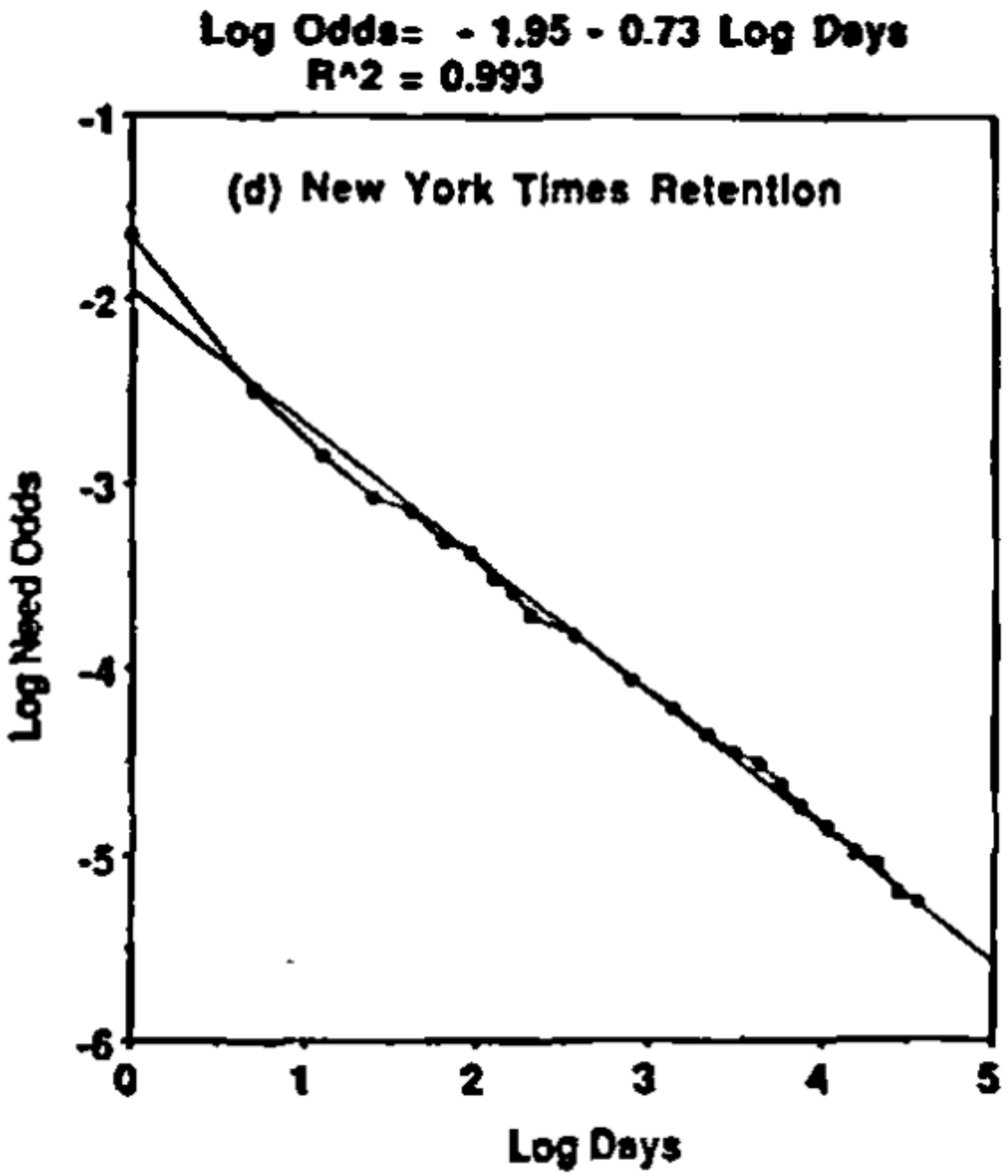
If words are queried in memory in order of frequency, time to query a word with frequency q is proportional to the number of words with frequency $> q$

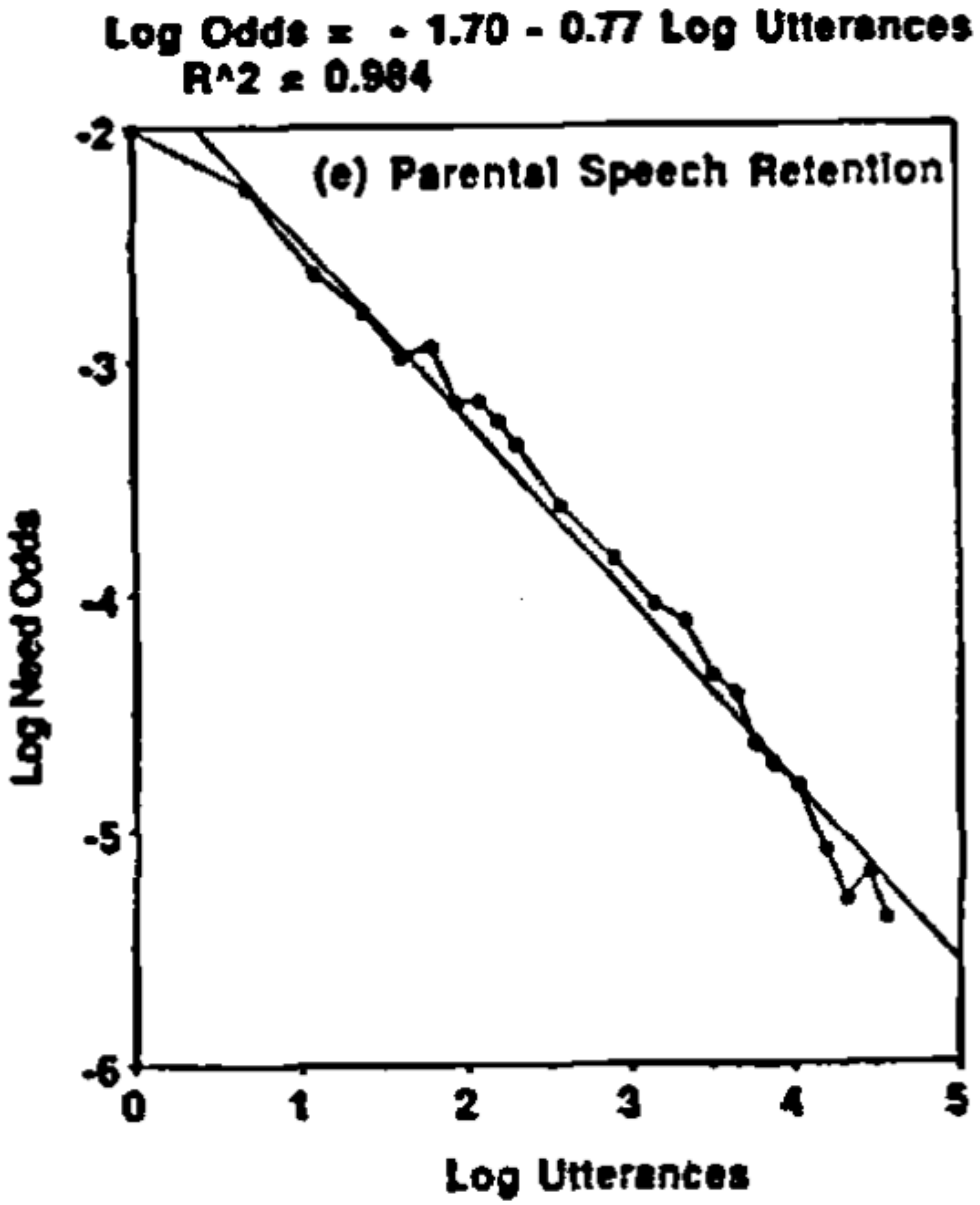
$$\int_q ax^{-k}dx = \frac{a}{k-1}q^{-(k-1)}$$

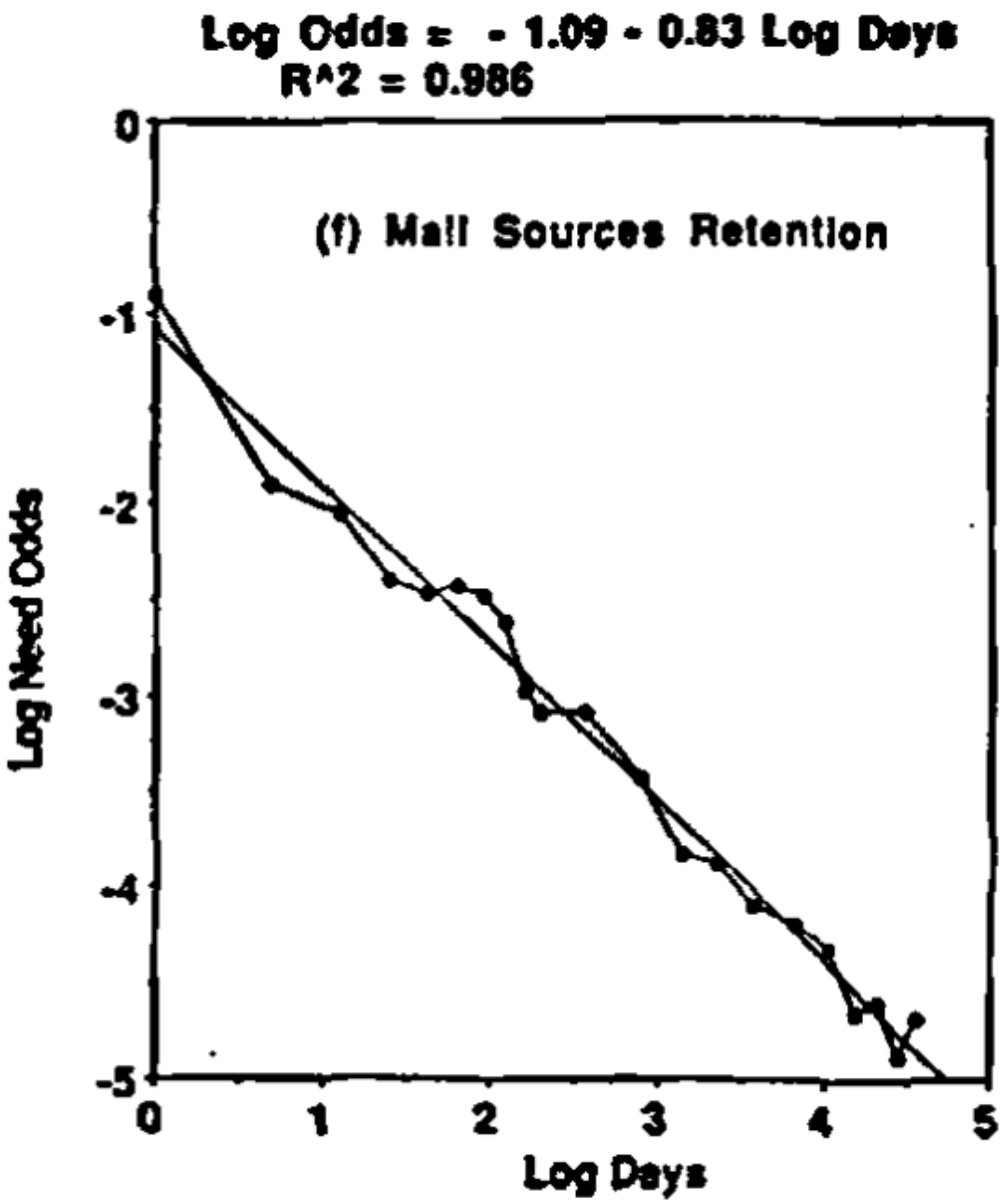
Time since last occurrence predicts occurrence - The power law of retention



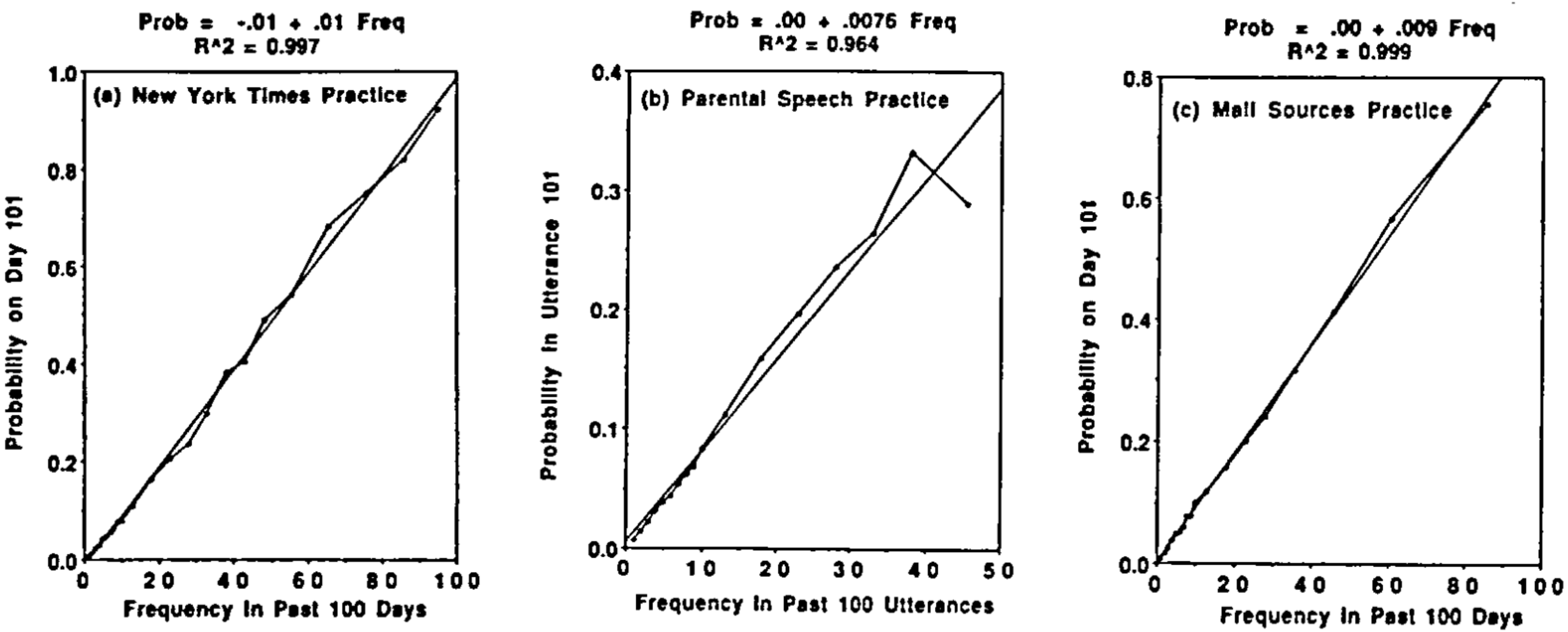
Time since last occurrence predicts occurrence - The power law of retention



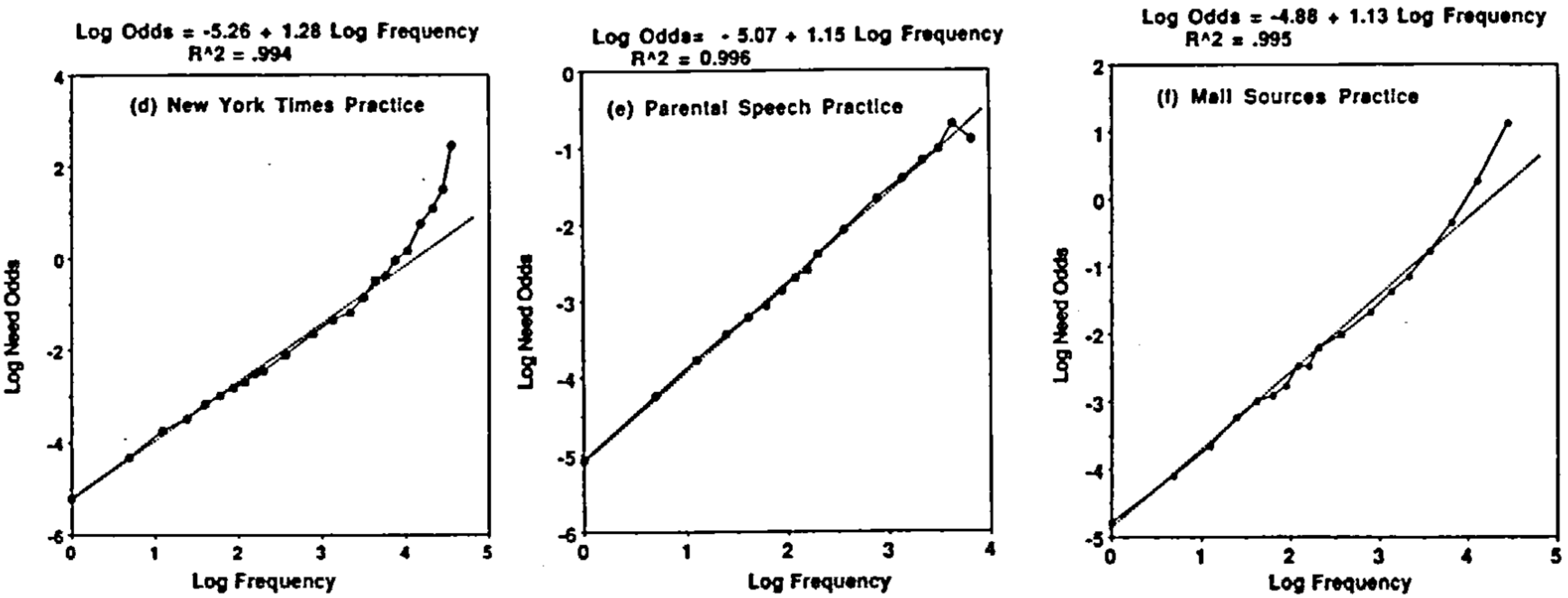




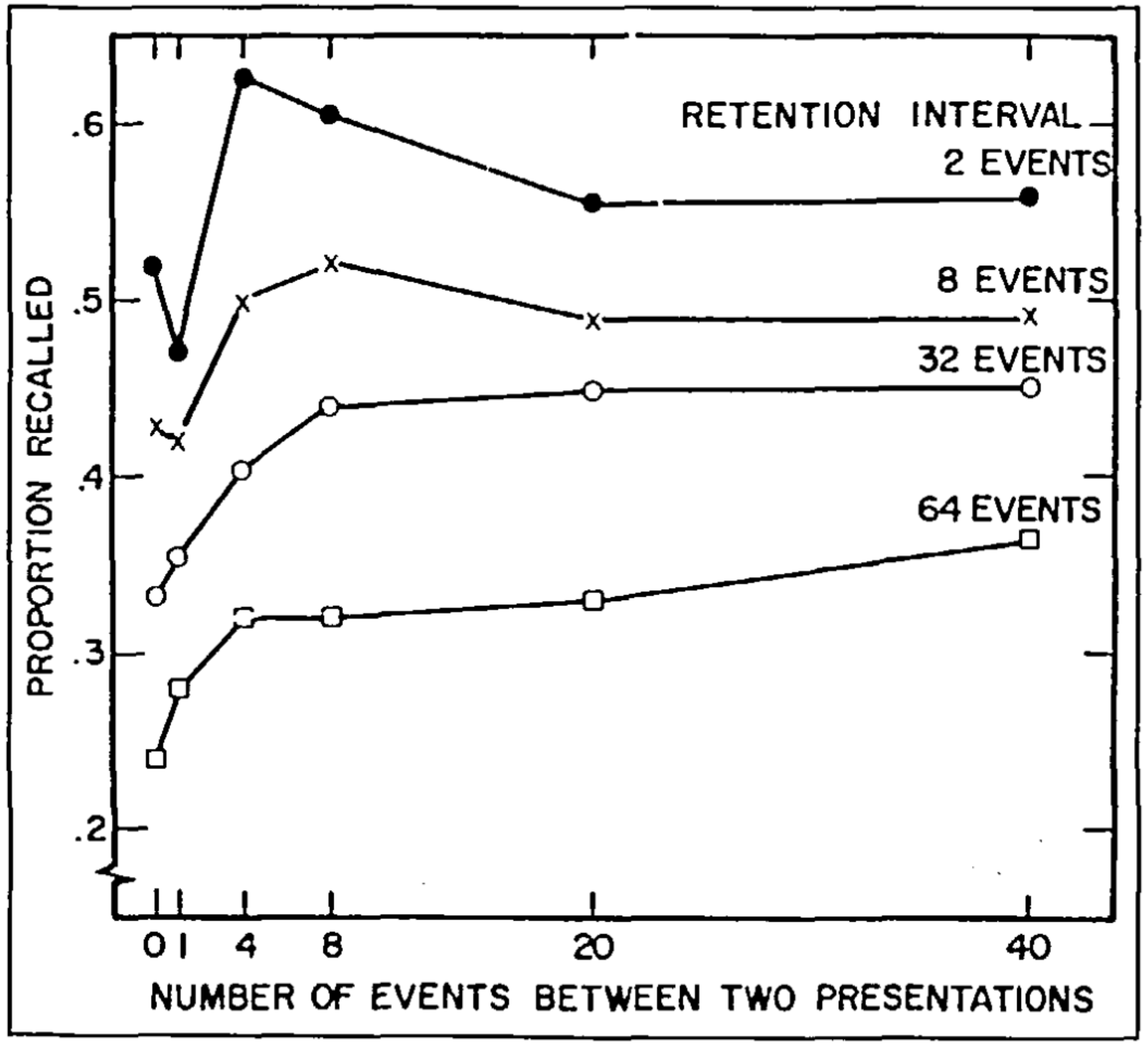
Past frequency predicts need frequency - The power law of practice

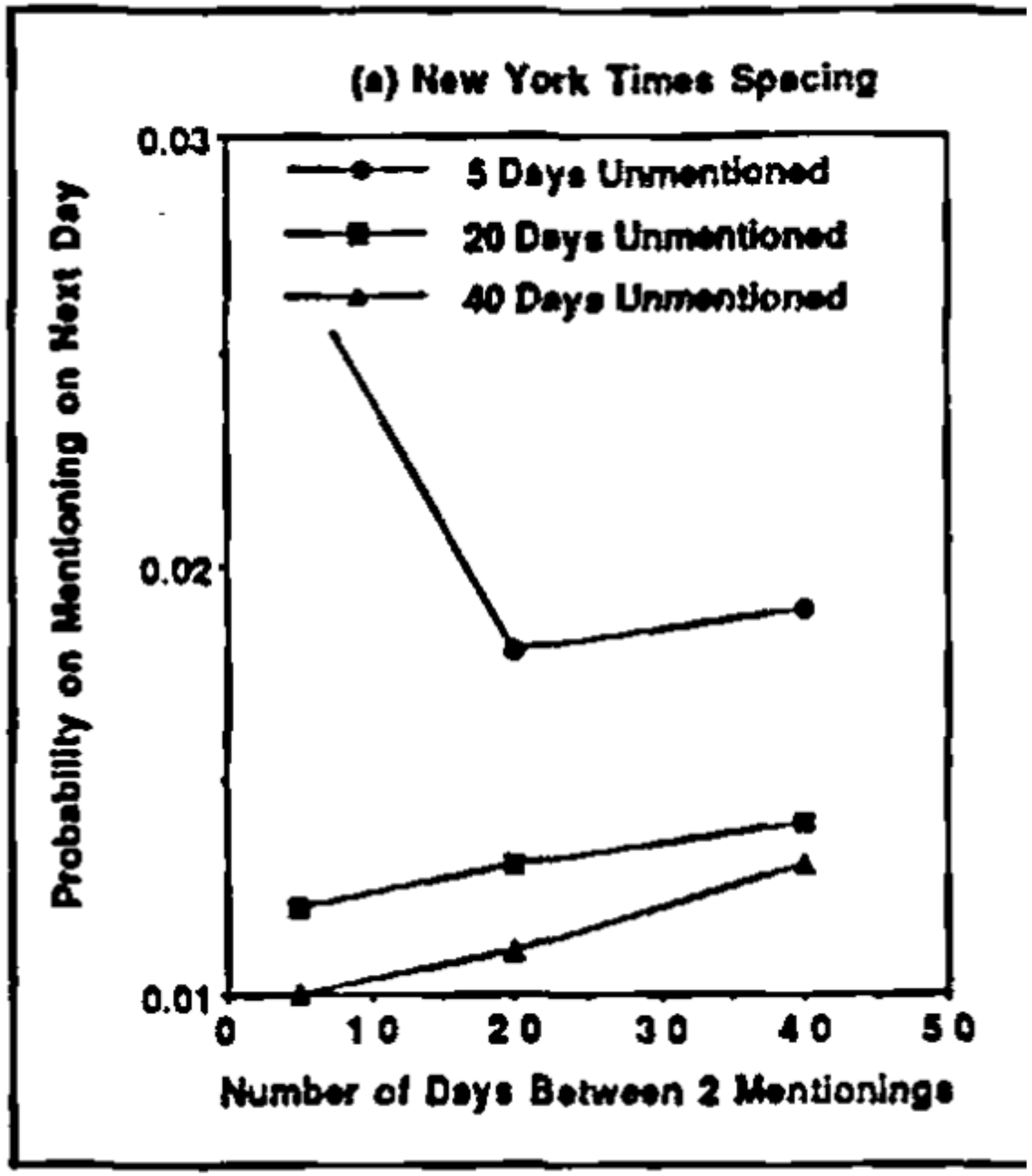


Past frequency predicts need frequency - The power law of practice



Total frequency and recent use interact - the spacing effect





S_1 S_2 T

Does this count as an explanation?

There is a lawful relationship between odds of something recurring and memory for that item

The explanation is: The mind is optimized to make “desirable” memories readily available

1. memories vary in desirability, and this desirability affects rates of use (retention and practice)

2. memories can rise and fall in desirability, and memory tracks this volatility (spacing)

Where do environmental frequencies come from?

Do you really track all of this information?

Every day prediction problems (Griffiths & Tenenbaum, 2006)

1. You read that a move has made $60 million to date.
   How much money will it make in total?

2. You see that something has been baking for 34 minutes.
   How long until it's ready?

3. You meet someone who is 78 years old.
   How long will they live?

4. Your friend quotes to you from line 17 of her favorite poem.
   How long is the poem?

5. You see cab #107 pull up to the curb at the airport.
   How many cabs in this city?

Making predictions

You encounter a phenomenon that has existed for t_{past} time.

How long will it continue into the future? (What is t_{total} ?)

Works the same for any quantity being estimate
(money made, poem length, number of cabs, etc.)

Making predictions by Bayesian inference

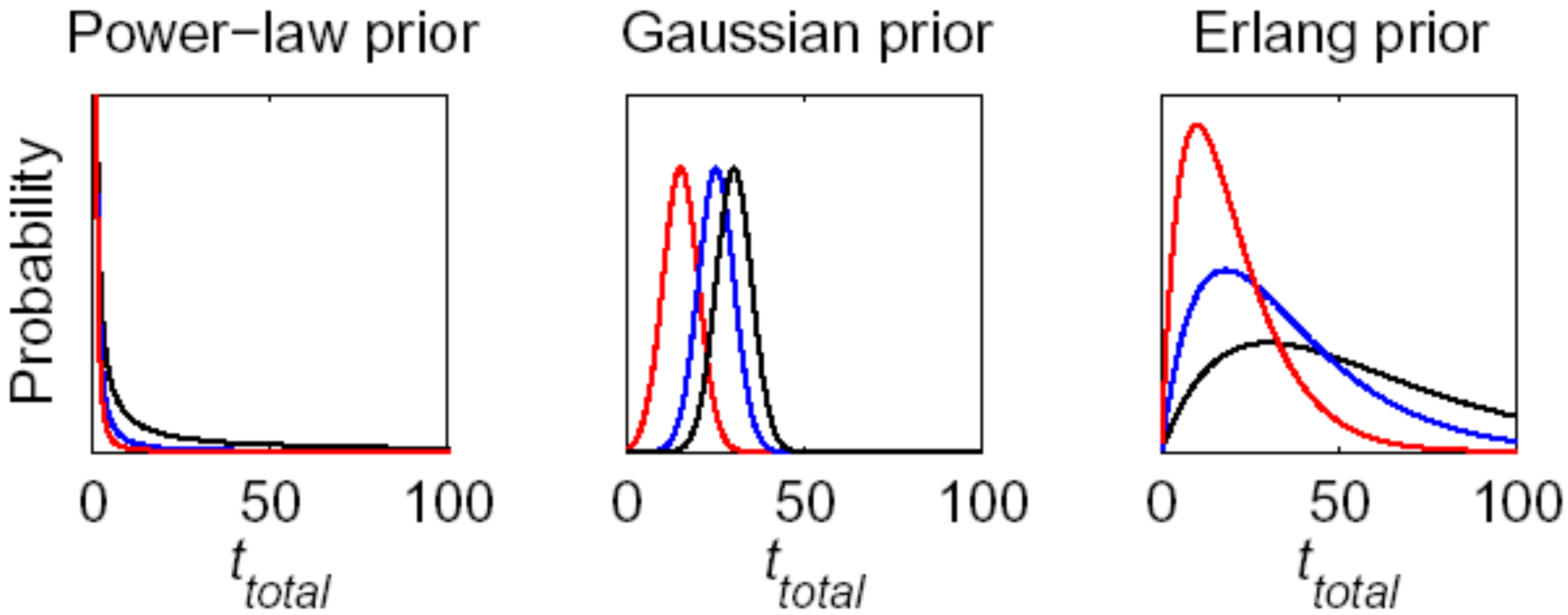
$$P\left(t_{total} \mid t_{past}\right) \propto P\left(t_{past} \mid t_{total}\right) P\left(t_{total}\right)$$

Assume you are equally likely to encounter an event at any point in duration

$$\frac{1}{t_{total}}$$

But how do you get this prior?

Different priors are appropriate for different domains

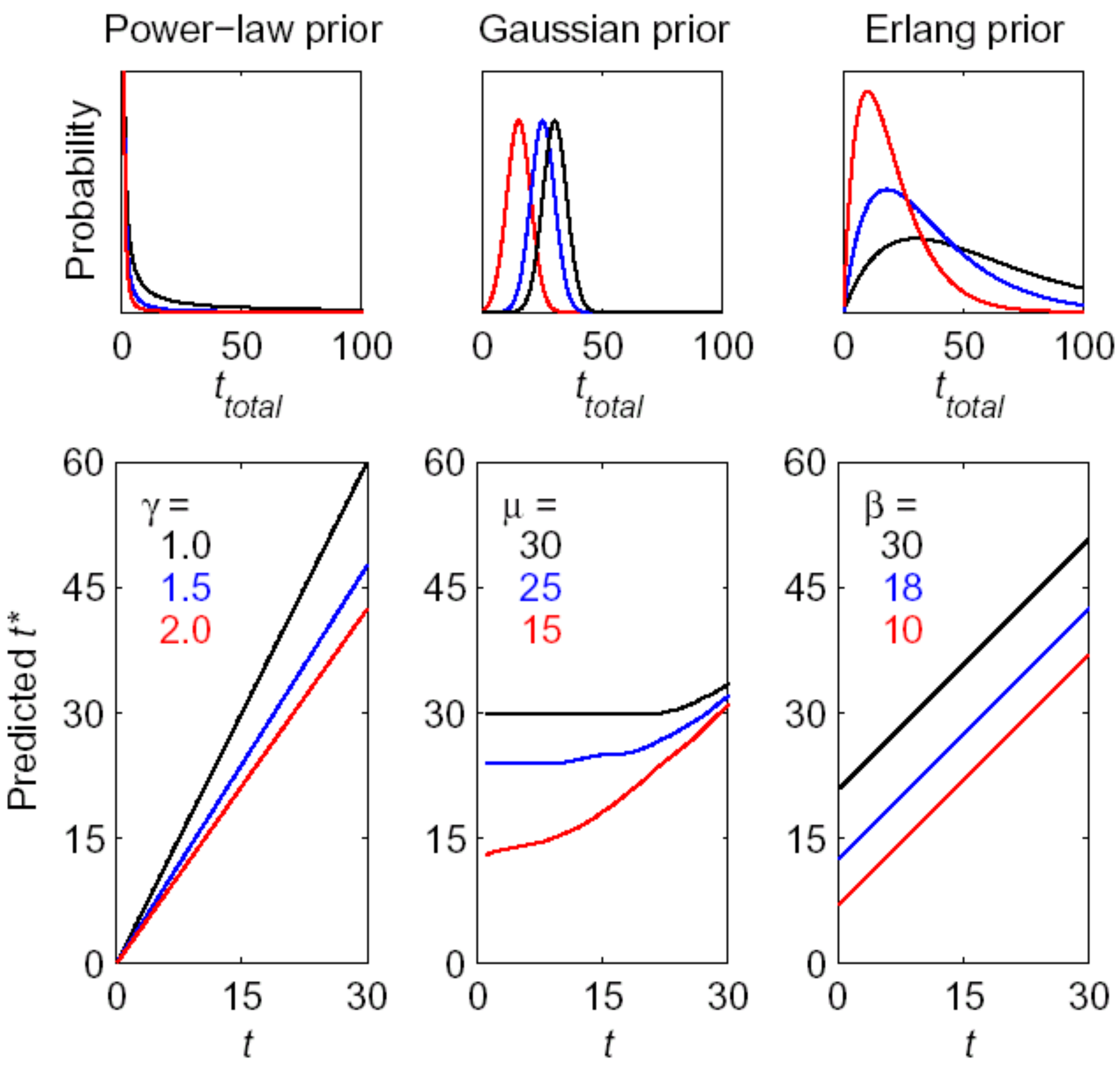


e.g. wealth,
friends

e.g. height,
lifespan

e.g. years in office,
reigns of Pharaohs

The effect of different priors



Evaluating peoples' predictions

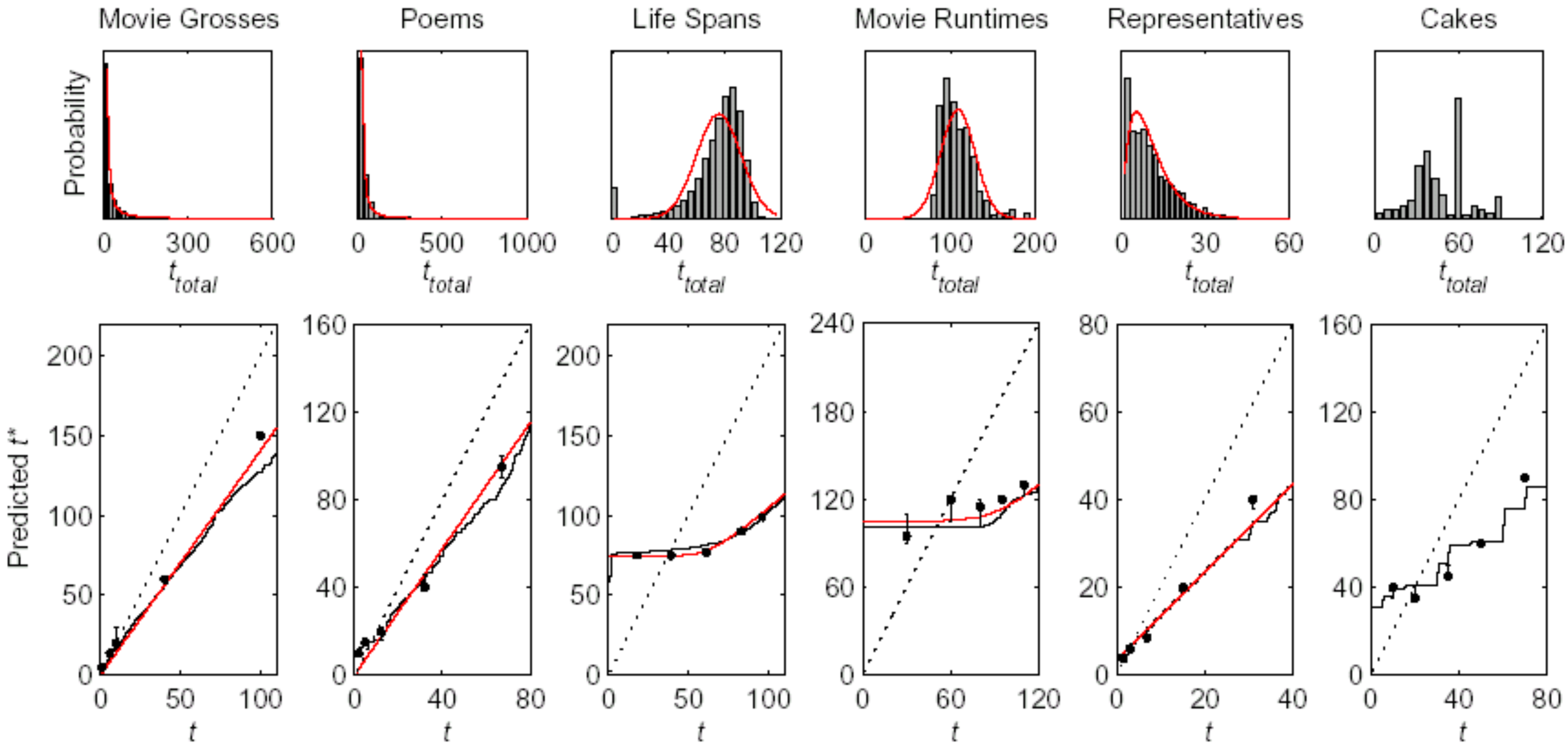
Gave people 5 values of t_{past} in different domains

Had people predict t_{total}

Sources of Data for Estimating Prior Distributions

Data set	Source (number of data points)
Movie grosses	http://www.worldwideboxoffice.com/ (5,302)
Poem lengths	http://www.emule.com/ (1,000)
Life spans	http://www.demog.berkeley.edu/wilmoth/mortality/states.html (complete life table)
Movie run times	http://www.imdb.com/charts/usboxarchive/ (233 top-10 movies from 1998 through 2003)
U.S. representatives' terms	http://www.bioguide.congress.gov/ (2,150 members since 1945)
Cake baking times	http://www.allrecipes.com/ (619)
Pharaohs' reigns	http://www.touregypt.com/ (126)

Evaluating peoples' predictions



People make optimal predictions in a variety of domains

People’s predictions across domains follow the principle of Bayesian inference

This means:

1. People track the shape of frequency distributions

2. People track the specifics of distributions across domains

3. People perform approximately optimal inference from these distributions

Rational analysis

For a given computational problem, there is an *optimal solution*. Whatever it is, we have evolved to approximate it.

Figure out the optimal solution, and you'll know a lot about what people do.

“The predictions flow from the statistical structure of the environment and **not** the assumed structure of the mind.“ (Anderson, 1991)

Does the structure in people's minds reflect the structure of the environment?

1. People's priors are merely a reflection of the environment
2. What people already know changes what they learn next
   (e.g. shape bias, semantics, transitional probabilities, etc.)

How do you choose what you should learn next?

How do you allocate your attention?

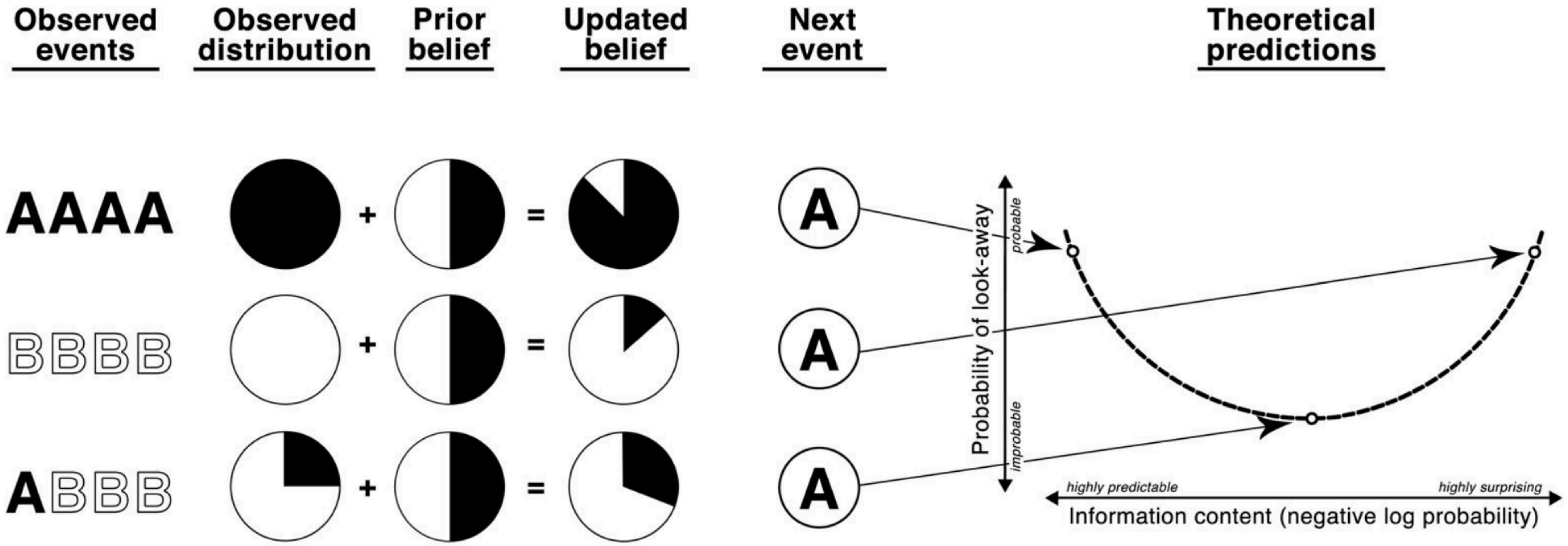
A rational analysis of attention allocation (Kidd, Piantadosi, & Aslin, 2012)

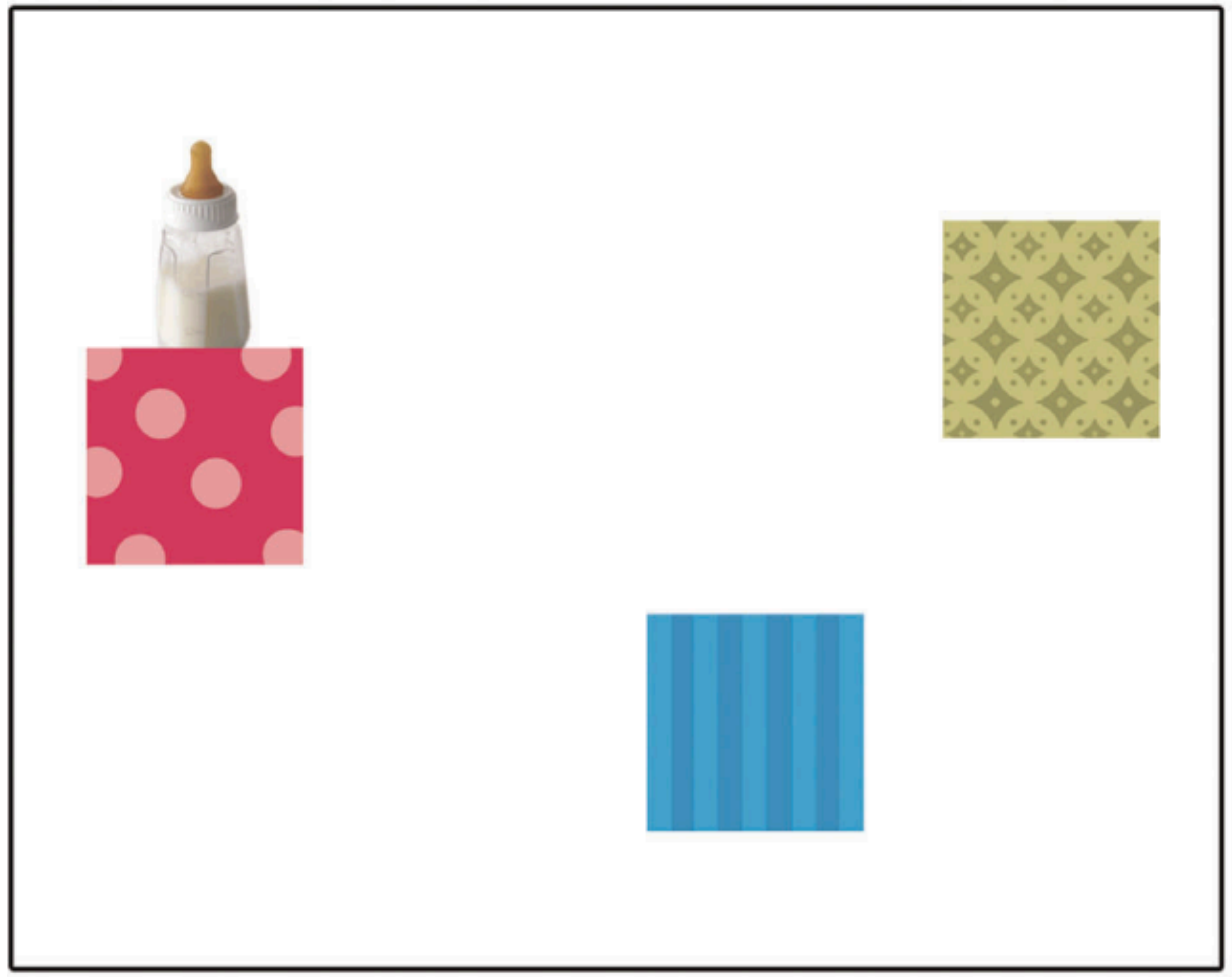
Goal: You should attend to information that is most likely to lead to learning

Strategy: You should attend to things that are surprising. But not *too* surprising.

- If something is not surprising enough, you probably already know it so there's nothing to learn
- If something is too surprising, you might not know enough about it to learn anything from it

The Goldilocks effect in infant attention





People make optimal predictions in a variety of domains

Modeling surprisal

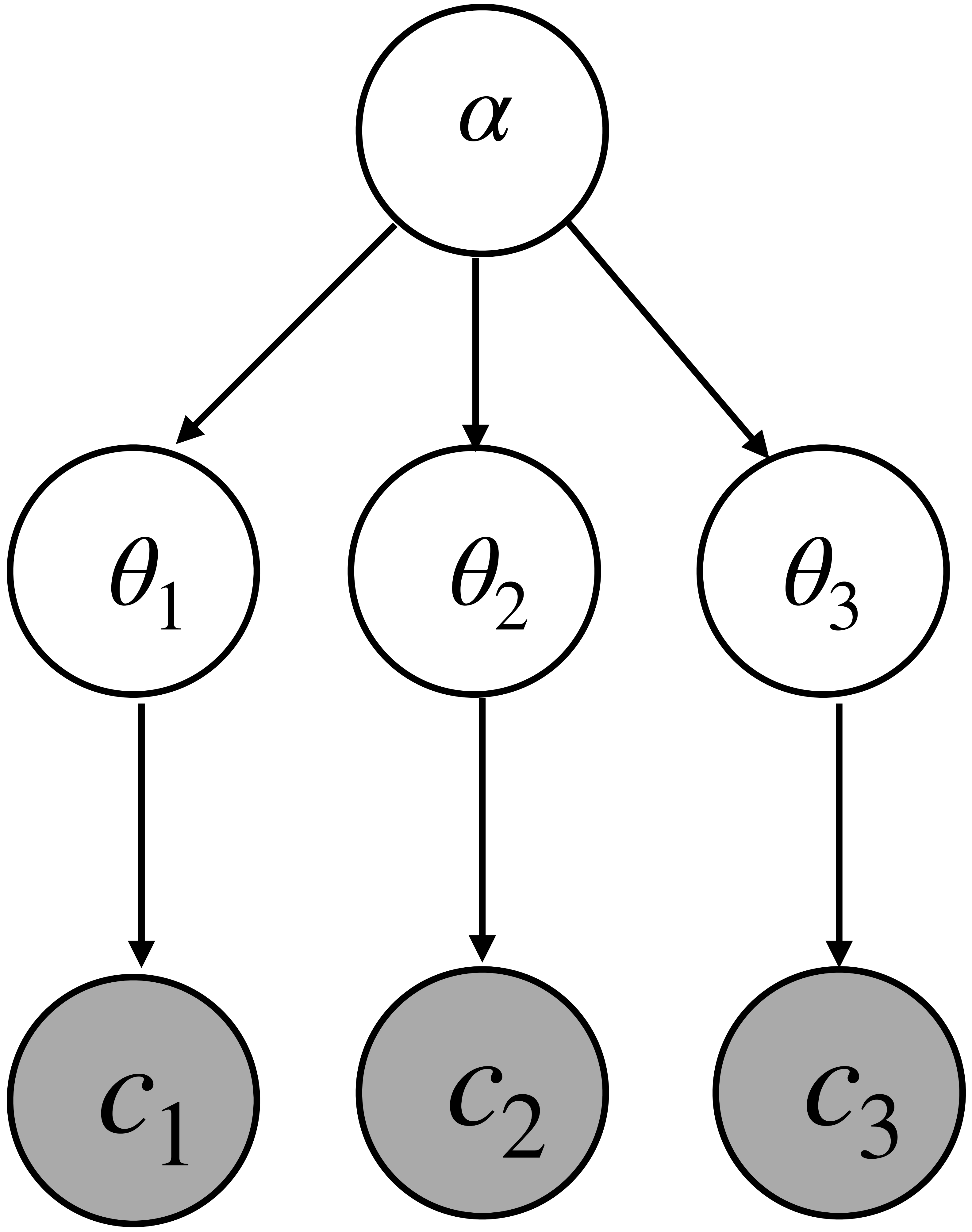
Goal: Learn a model θ that predicts the probability that an object pops out of each box

$\alpha = 1$ (weak Uniform prior)

$\theta \sim \text{Dirichlet}\,(\alpha)$

$c \sim \text{MultiNomial}\,(\theta)$

$$P\left(\theta \,|\, c_1, c_2, c_3, \alpha\right) = \frac{1}{B}\prod_{i=1}^{3}\theta_i^{\alpha+c_i-1}$$



This is just a generalization of the coinflip model you already saw!

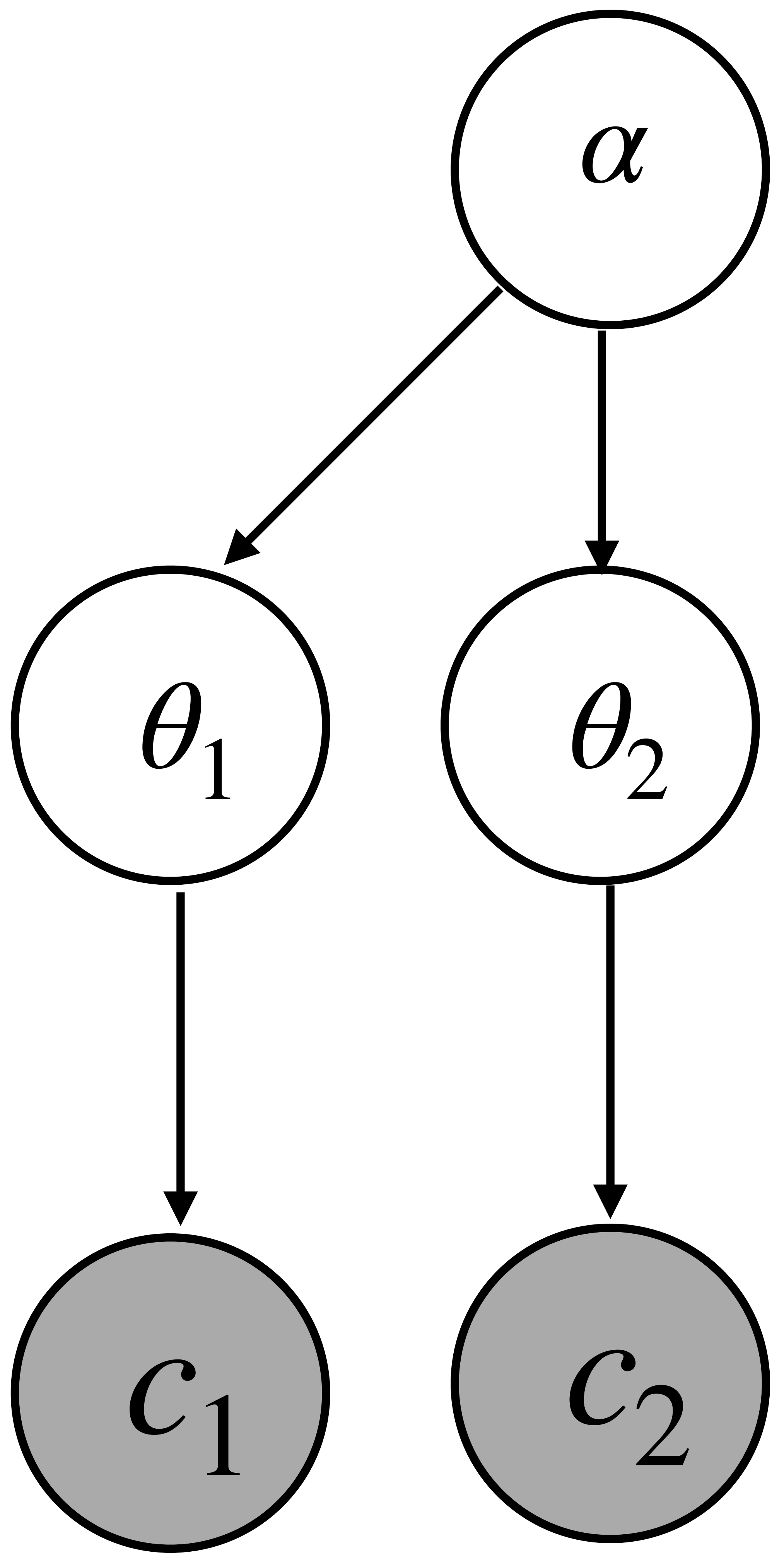
Goal: Learn a model θ that predicts the probability that a coin comes up heads c_1 times and tails c_2 times

$$\alpha = 1 \quad \text{(weak Uniform prior)}$$

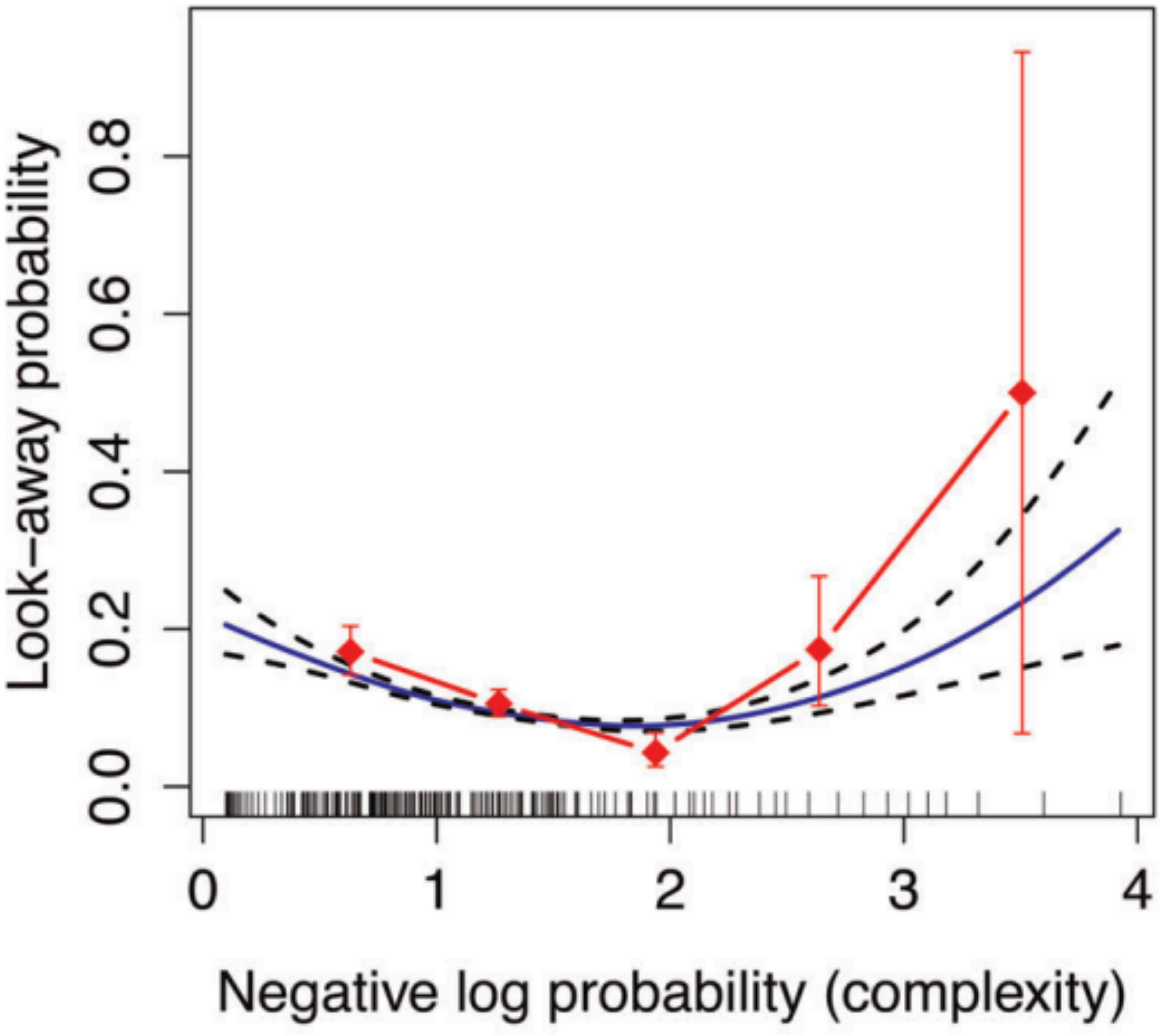
$$\theta \sim \text{Beta}\,(\alpha)$$

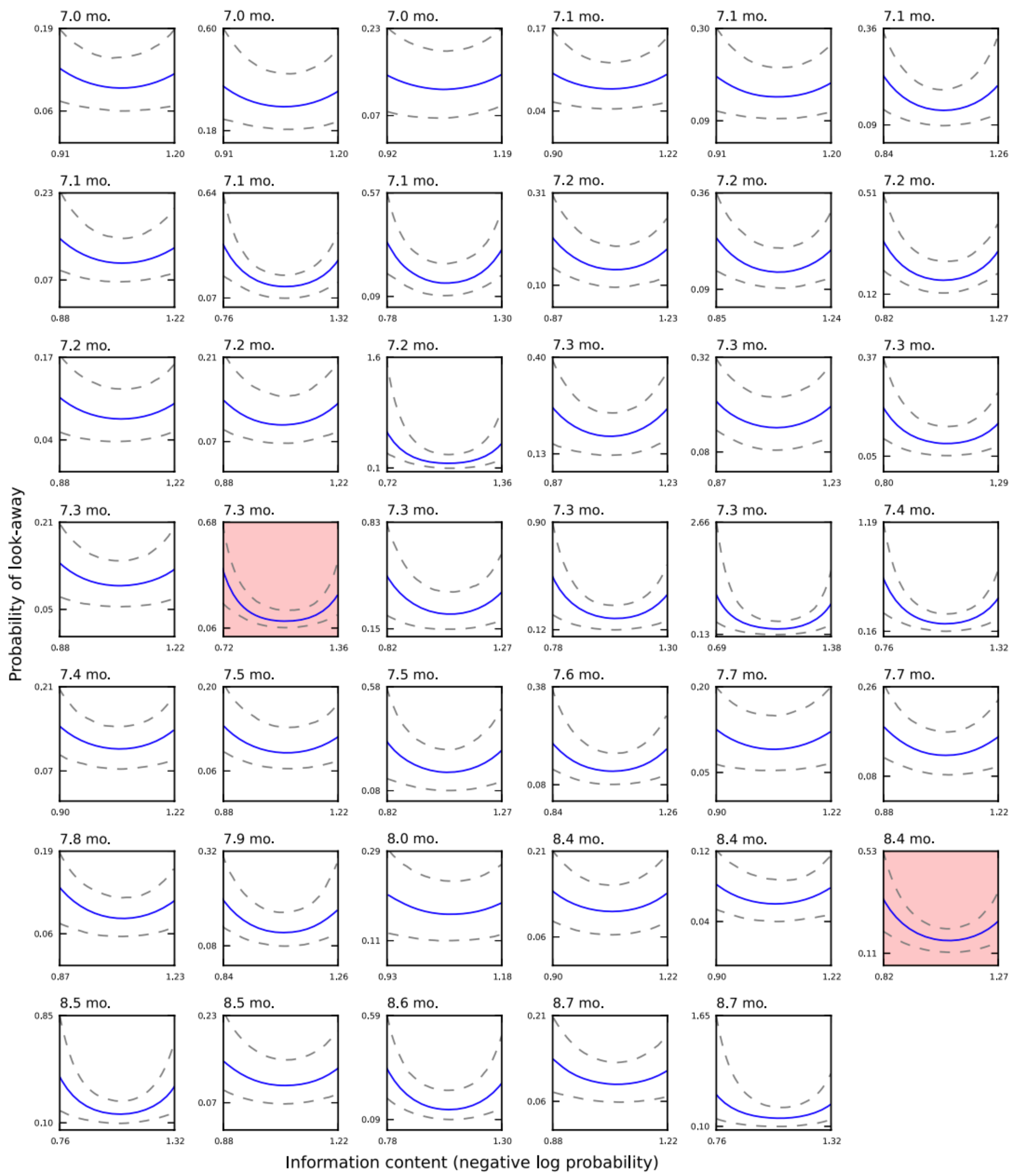
$$c \sim \text{Binomial}\,(\theta)$$

$$P\left(\theta \,|\, c_1, c_2, \alpha\right) = \frac{1}{B}\prod_{i=1}^{2}\theta_i^{\alpha + c_i - 1}$$

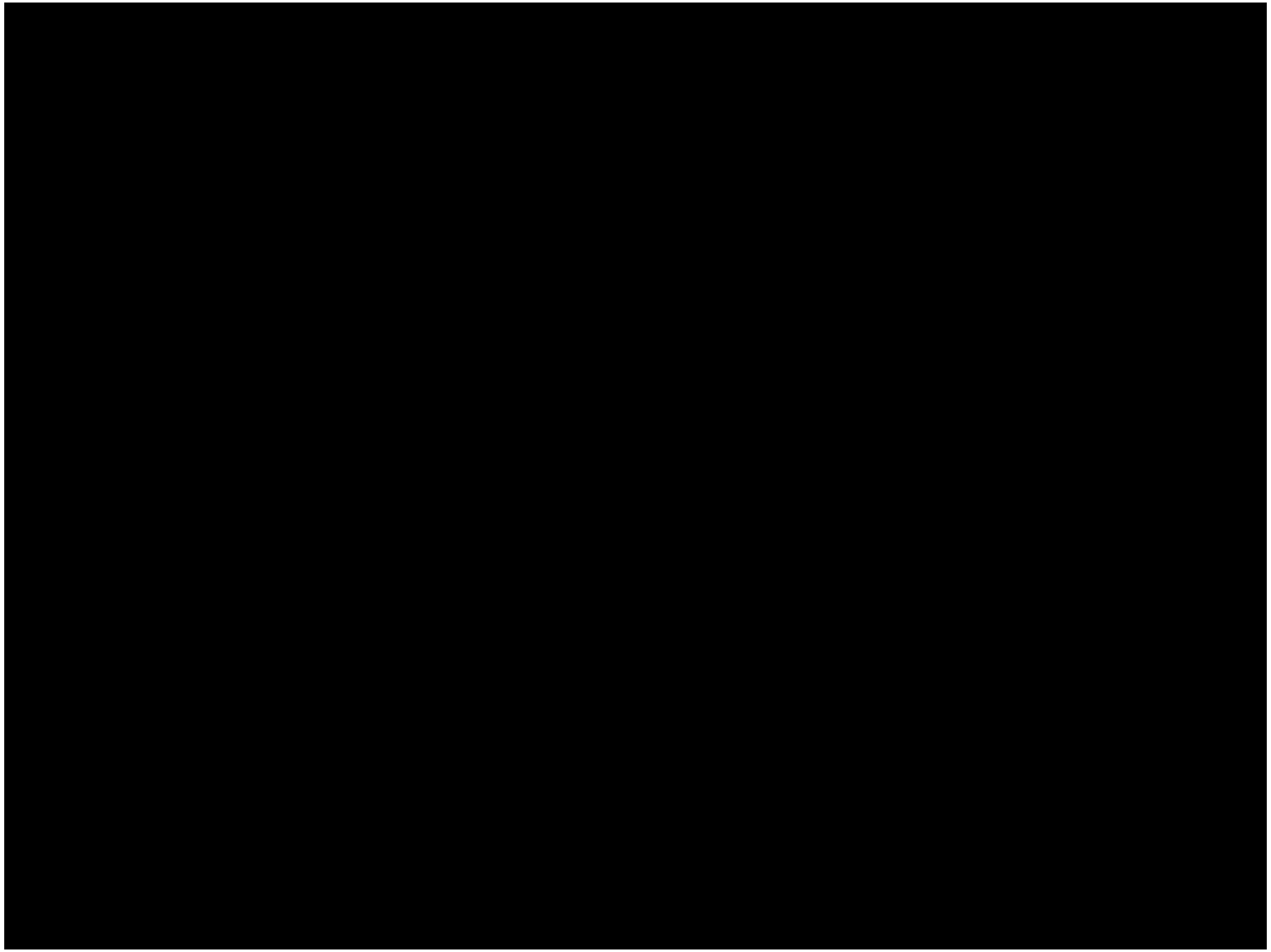


Infants look away when events are either too surprising or not surprising enough





How rational analysis can go wrong



Rational analysis

1. Rational analysis is a framework theory for modeling learning and cognition

2. Memory retrieval can be modeled as optimal search

3. People track surprisingly precise frequency distributions

4. Rational analysis relies on characterizing the information in the environment

The next few days

October 15	**Sampling methods**	
October 20	**Bayesian associative learning**	
October 21	The number game	