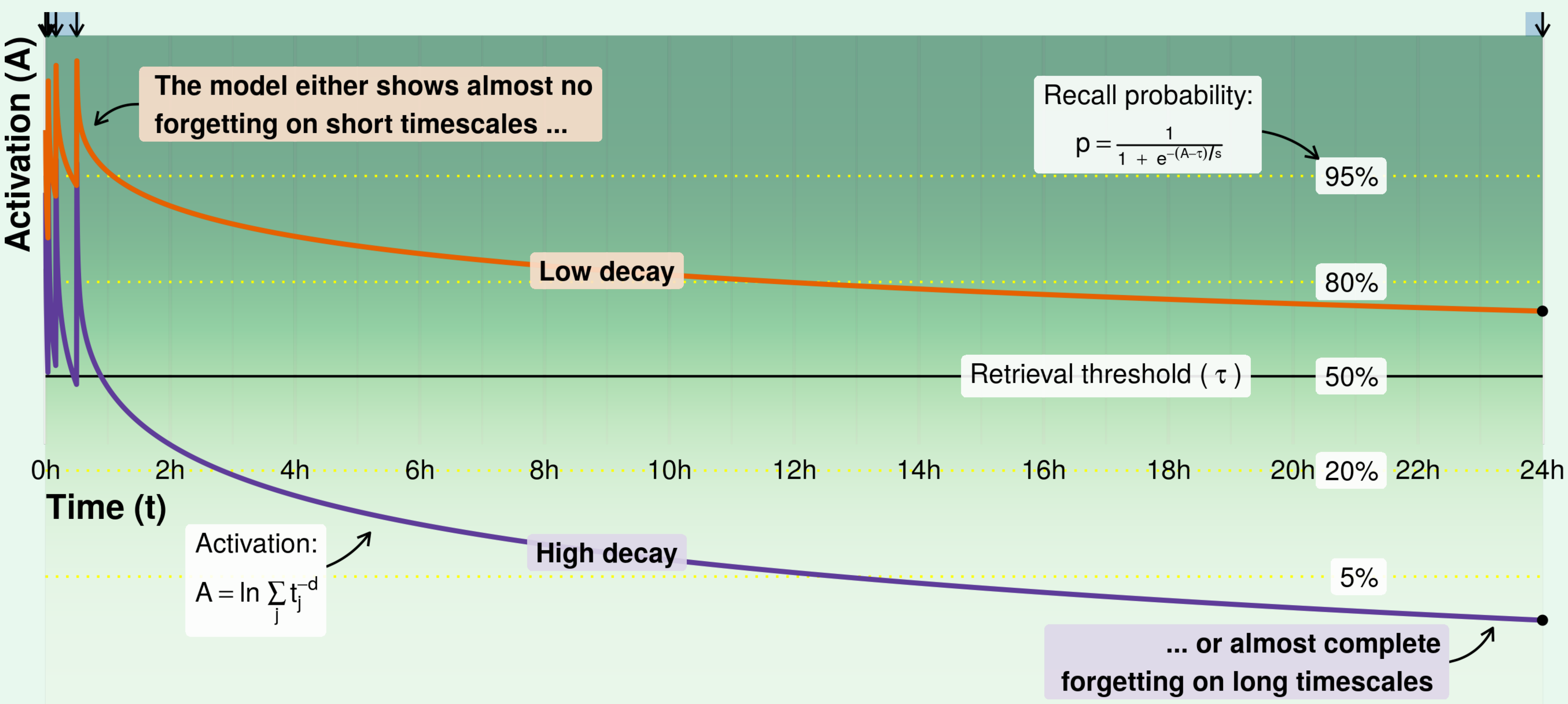# Explaining forgetting at different timescales requires a time-variant forgetting function

Maarten van der Velde , Florian Sense , Jelmer Borst , and Hedderik van Rijn

Problem: a simple trace-decay function does not work across all timescales simultaneously

The model either shows almost no forgetting on short timescales ...

Recall probability:
$$p = \frac{1}{1 + e^{-(A-\tau)/s}}$$

Low decay — 80%, 95%

Retrieval threshold (τ) — 50%

Activation:
$$A = \ln \sum_j t_j^{-d}$$

High decay — 5%
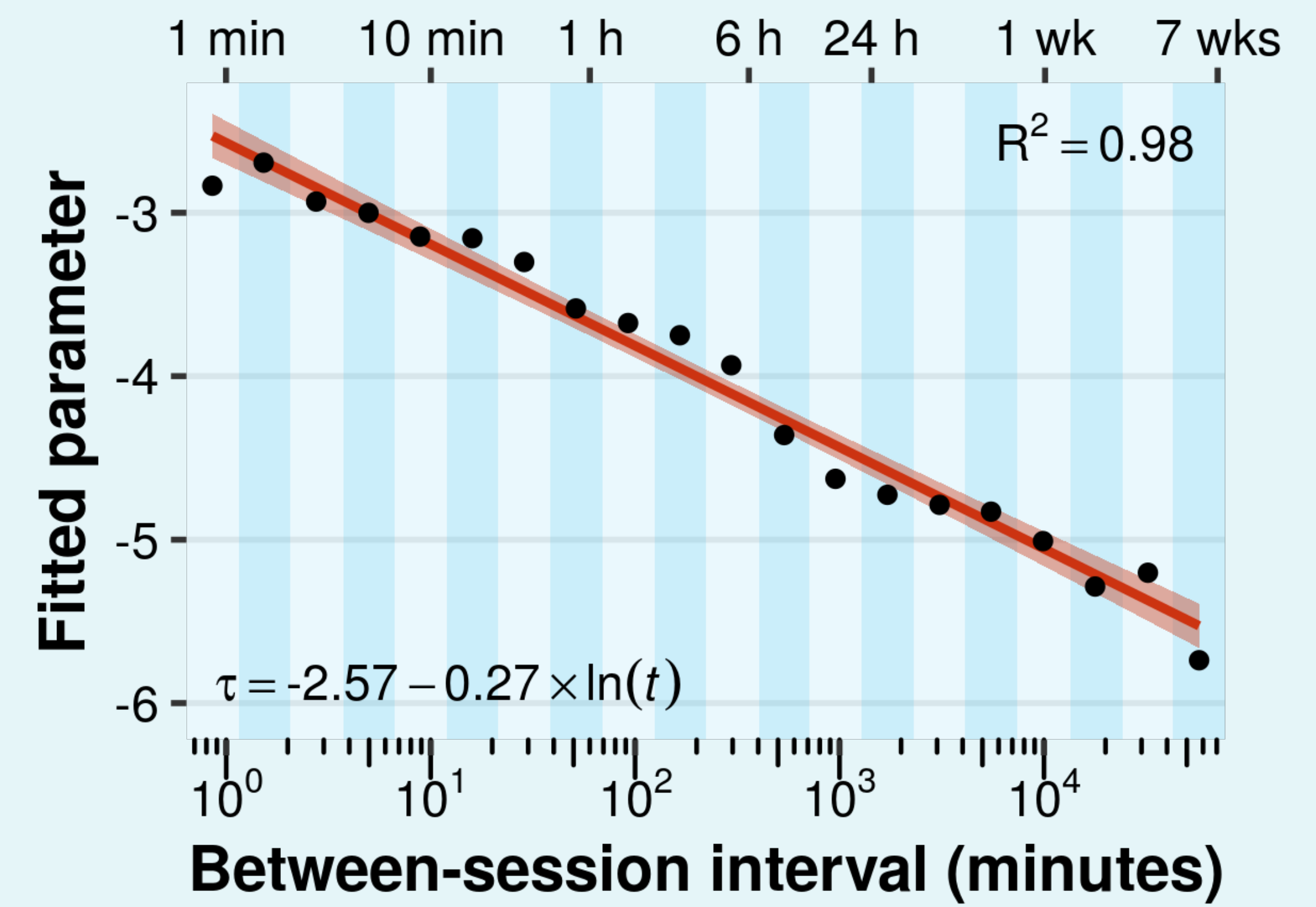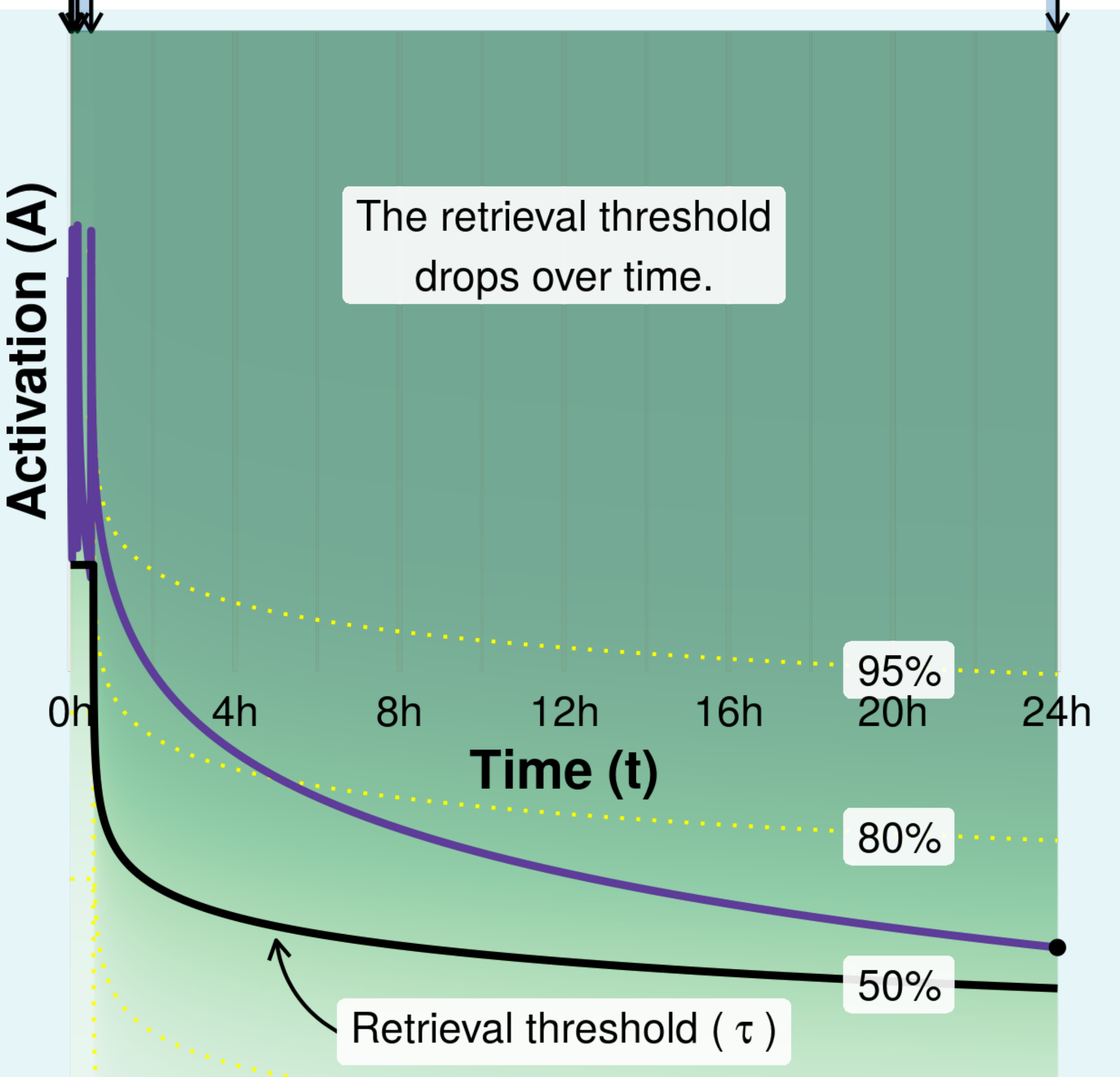
... or almost complete forgetting on long timescales

Data: 25,843 learning sequences

University students doing spaced retrieval practice across multiple sessions, with intervals ranging from seconds to weeks.
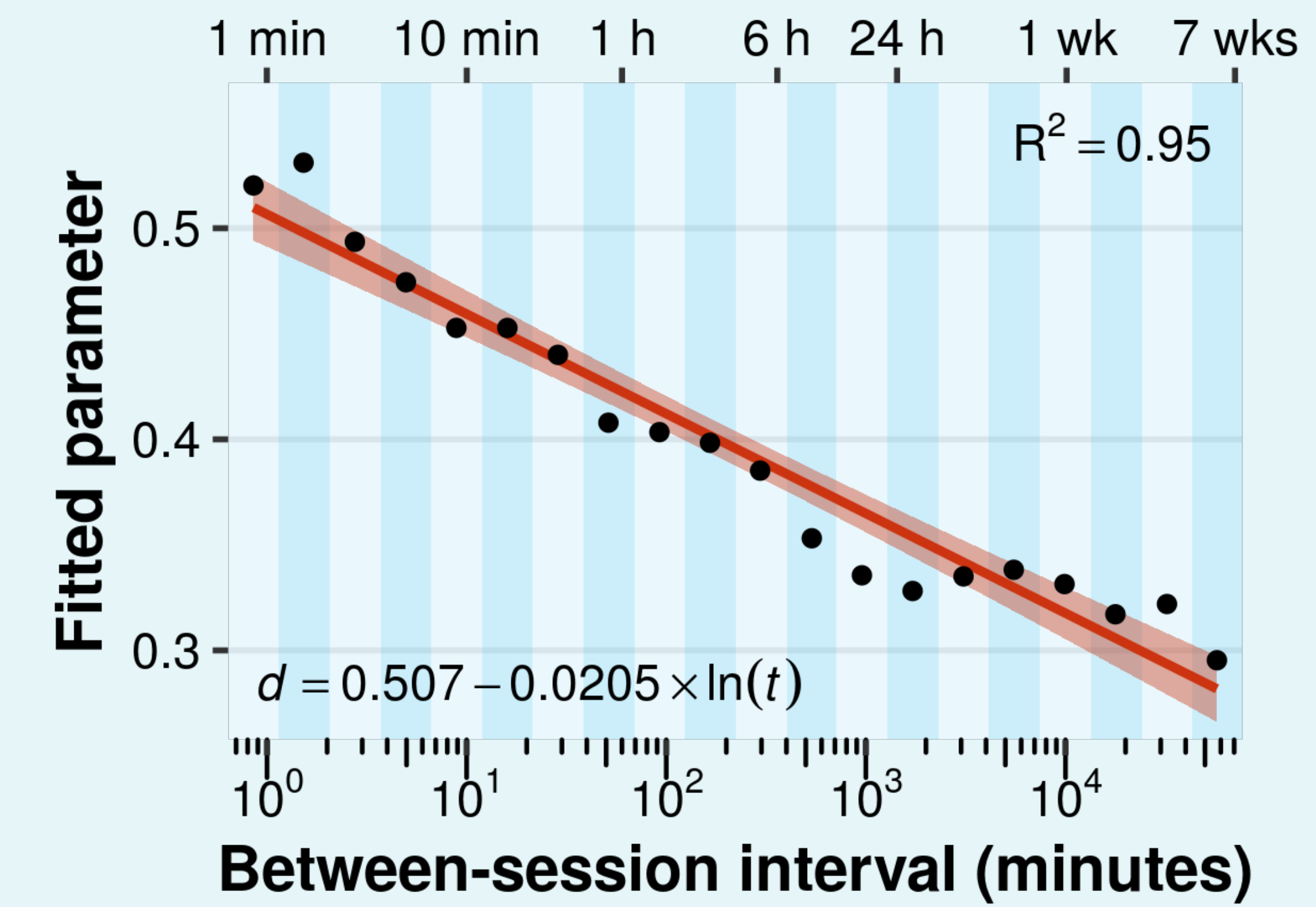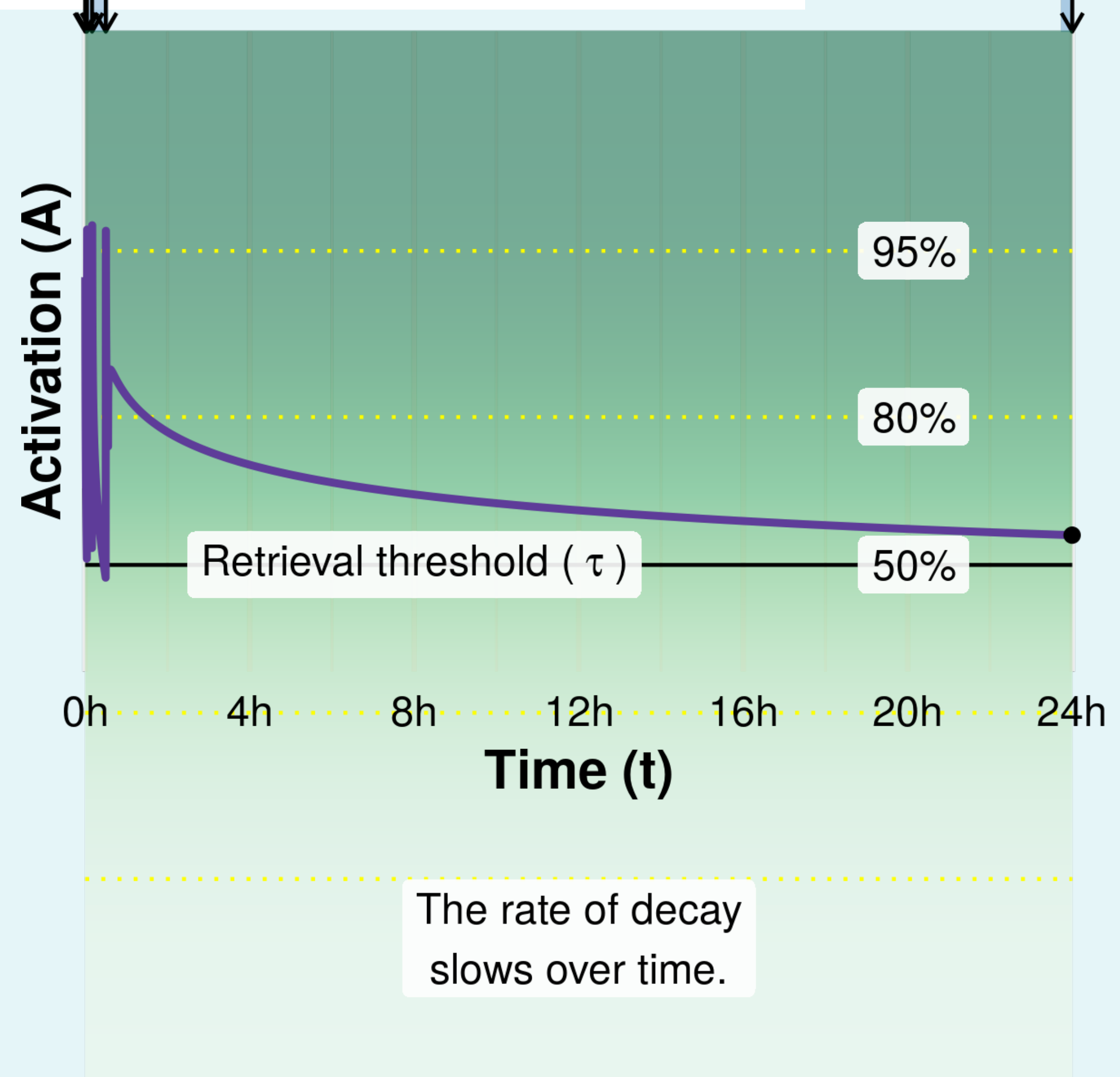
Data
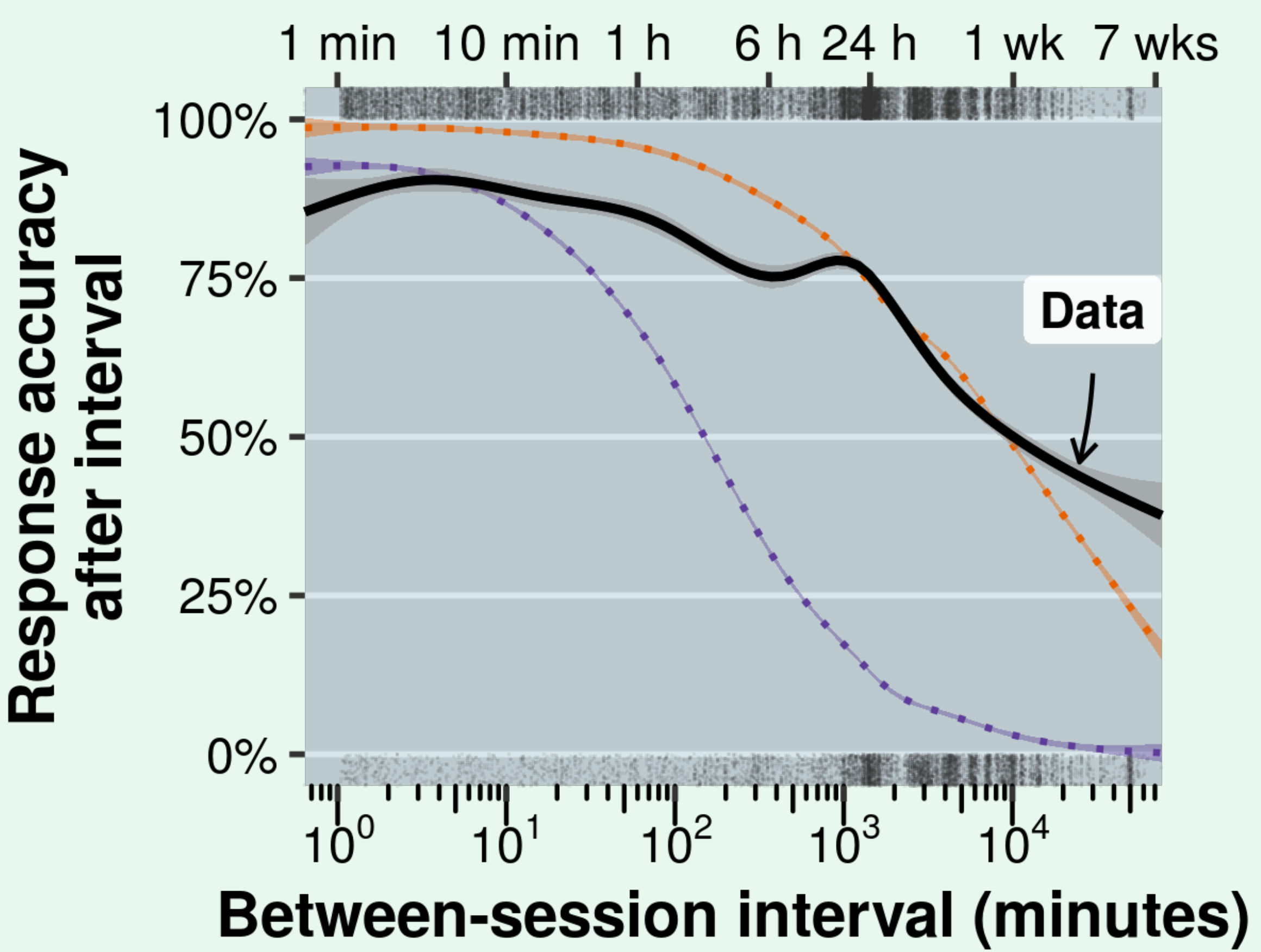
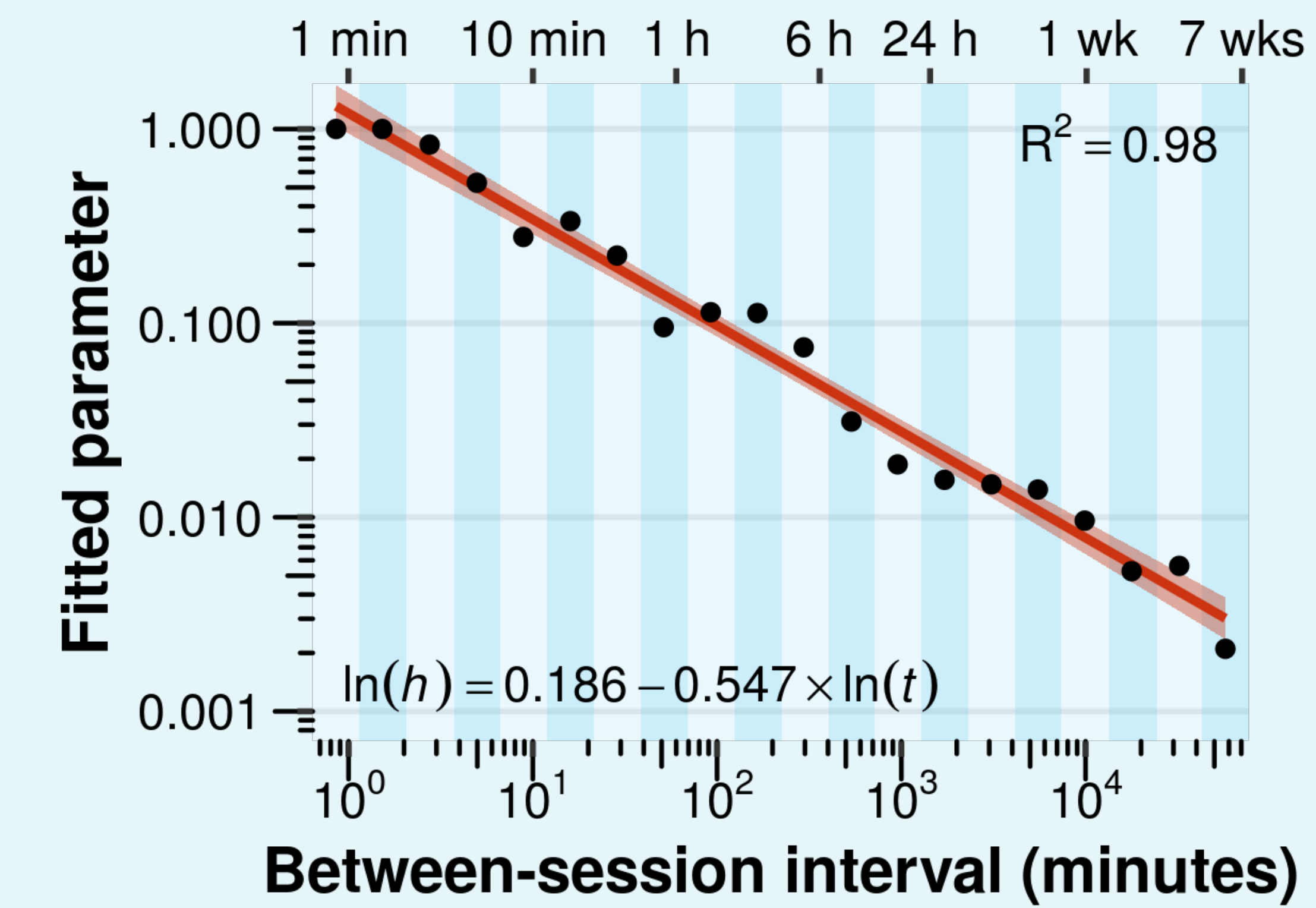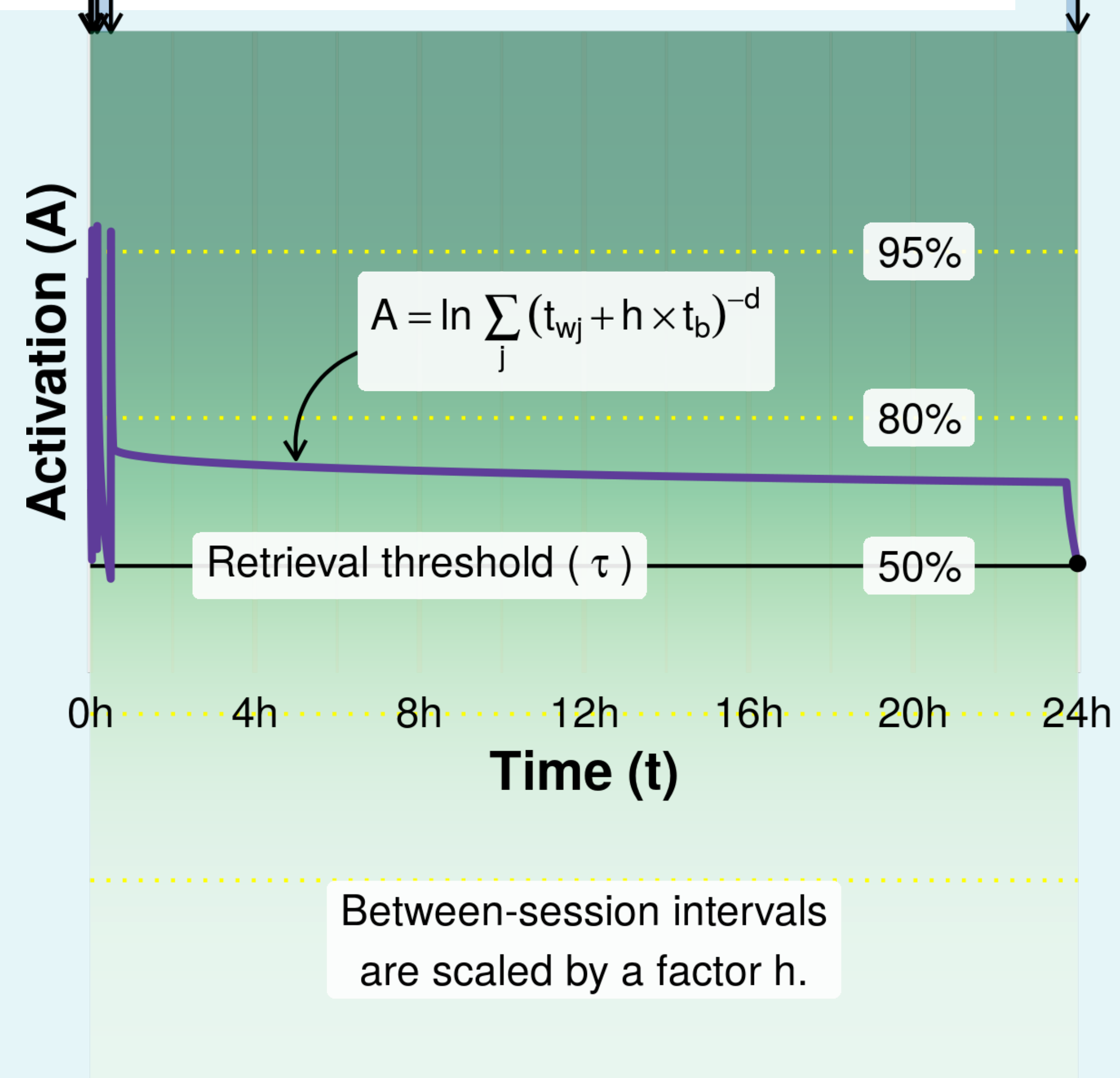## Solution: introduce a time-variant component to the forgetting function

Option 1: time-variant retrieval threshold

The retrieval threshold drops over time.

Retrieval threshold (τ)

$R^2 = 0.98$

$$\tau = -2.57 - 0.27 \times \ln(t)$$

Option 2: time-variant decay

The rate of decay slows over time.

Retrieval threshold (τ)

$R^2 = 0.95$

$$d = 0.507 - 0.0205 \times \ln(t)$$

Option 3: time-variant scaling factor

$$A = \ln \sum_j (t_{wj} + h \times t_b)^{-d}$$

Retrieval threshold (τ)

Between-session intervals are scaled by a factor h.

$R^2 = 0.98$

$$\ln(h) = 0.186 - 0.547 \times \ln(t)$$

Conclusions

- Capturing memory retention across timescales using a common trace-decay function requires a **time-variant component**.
- Using naturalistic data, we show that there are **multiple solutions**, each with its own theoretical commitments and implications.

Practical implications

- For education: a memory model that is accurate both within and between learning sessions can **improve learning outcomes**.
- Optimal parameters change predictably, making it possible to use a single model for a wide range of **educationally relevant intervals**.

More information

Read the paper: **psyarxiv.com/d58n4**
Get in touch:
✉ m.a.van.der.velde@rug.nl
🐦 @mavdvelde