Attention, Curiosity, and Memorability: Insights from Cognitive Science and AI

Thomas Espeseth,
Department of psychology, University of Oslo;
TRUST - The Norwegian Centre for Trustworthy AI;
Mentalese AS

Requirements for learning (Dehaene, 2020)

- **Attention**
- **Active engagement**
- **Error feedback**
- **Consolidation**

Requirements for learning (Dehaene, 2020)

- **Attention** (*Challenge:* easily distracted; difficult to sustain)
- **Active engagement** (*Challenge:* effortful and often resisted)
- **Error feedback** (*Challenge:* delayed, absent, or unclear)
- **Consolidation** (*Challenge:* fragile and easily disrupted)

Requirements for learning (Dehaene, 2020)

- **Attention** (*Challenge:* easily distracted; difficult to sustain)
- **Active engagement** (*Challenge:* effortful and often resisted)
- **Error feedback** (*Challenge:* delayed, absent, or unclear)
- **Consolidation** (*Challenge:* fragile and easily disrupted)

- I’d like to first demonstrate some of these challenges
- Then, I’d like to suggest one approach to meet them


WHODUNNIT?

Suppression of bottom-up attention - Change Blindness

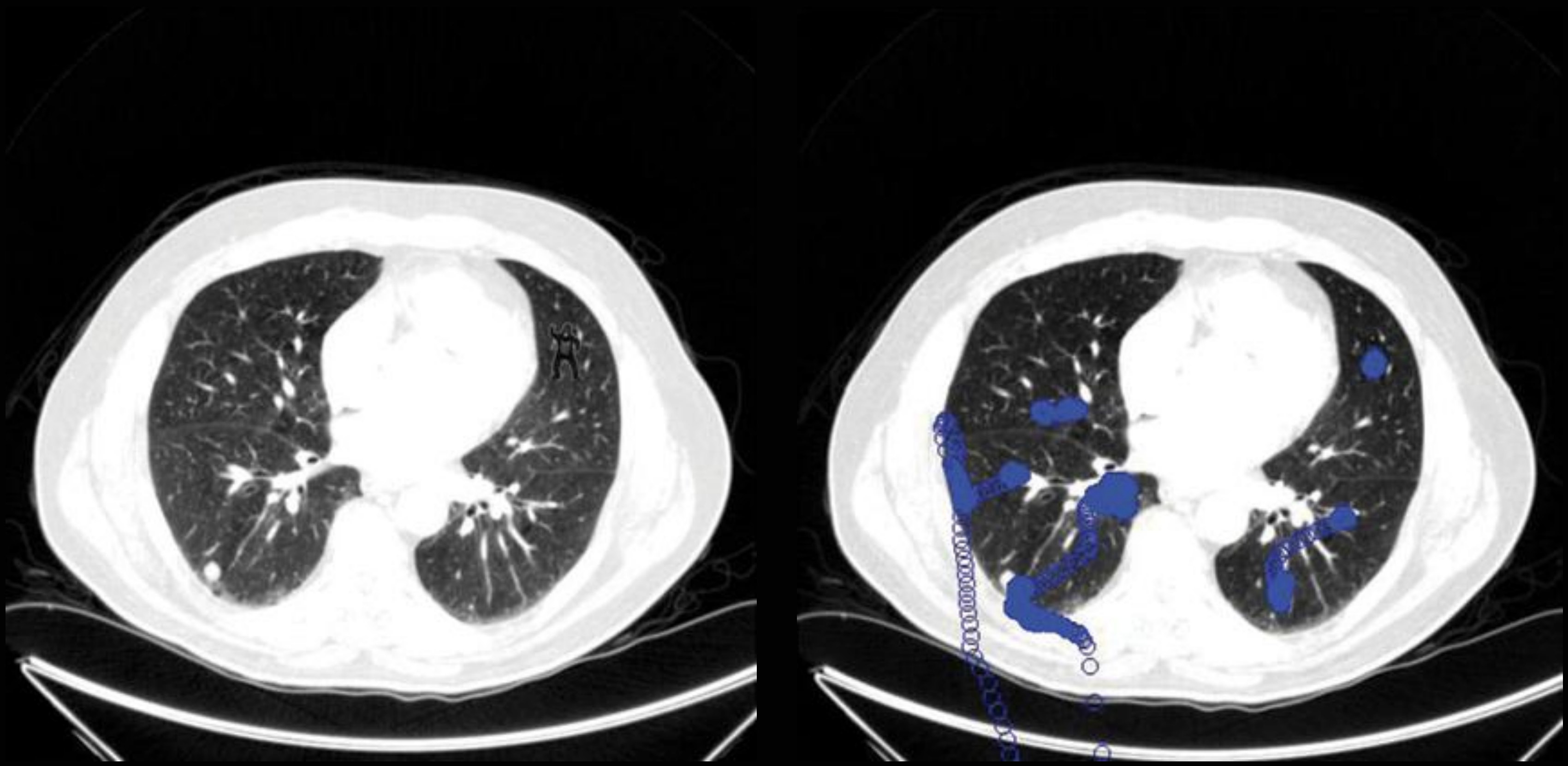
Suppression of bottom-up attention - Change Blindness

Inattentional blindness – the cost of focused attention

http://viscog.beckman.uiuc.edu/flashmovie/15.php

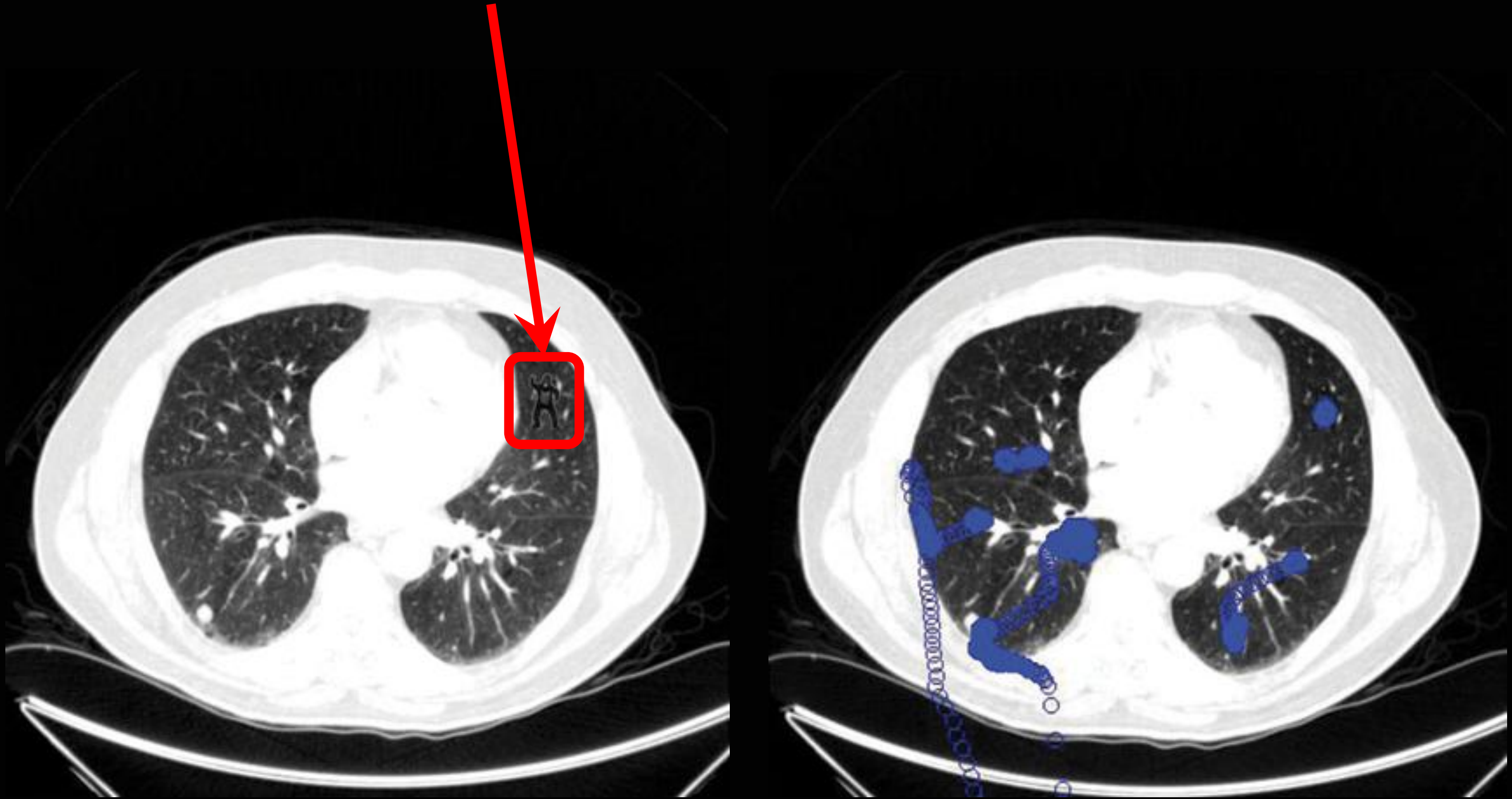
Source: Simons, 2000.

The Invisible Gorilla Strikes Again



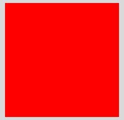
Radiologists better on lung nodules, but 83% did not see the gorilla Drew et al. 2013

The Invisible Gorilla Strikes Again



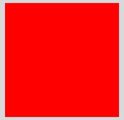
Radiologists better on lung nodules, but 83% did not see the gorilla   Drew et al. 2013

Attentional blink



Demo

Attentional blink



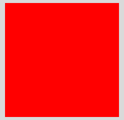
T2

Did you see an X?

B T D L G Q X ...

T2

Attentional blink



T1+T2

What was the identity of the white letter?

R

P

N

Y

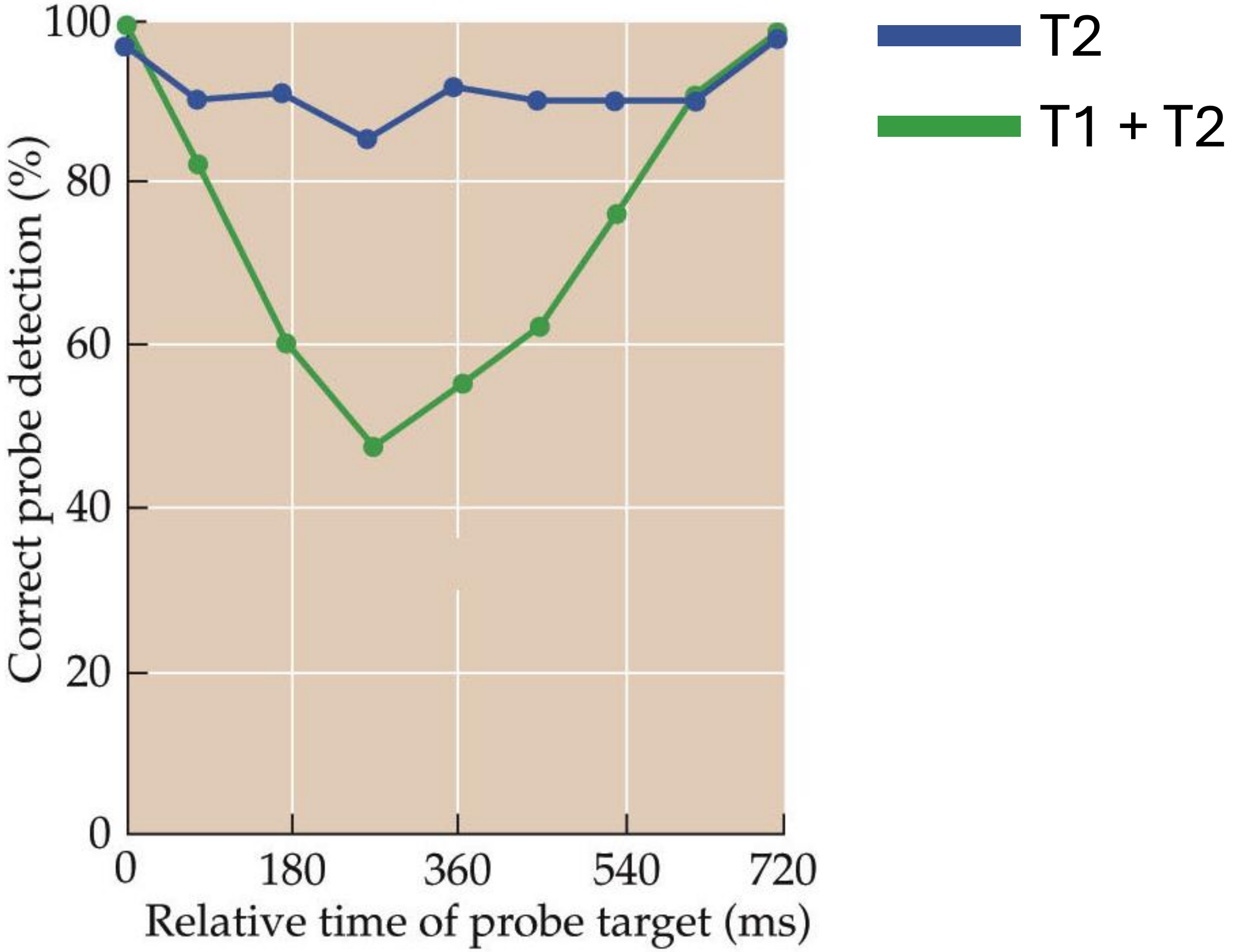
T1+T2

Did you see an X?

R P N Y J K X ...

T1+T2

Attentional blink



Data from Raymond et al. 1992; Figure from Purves et al. 2008

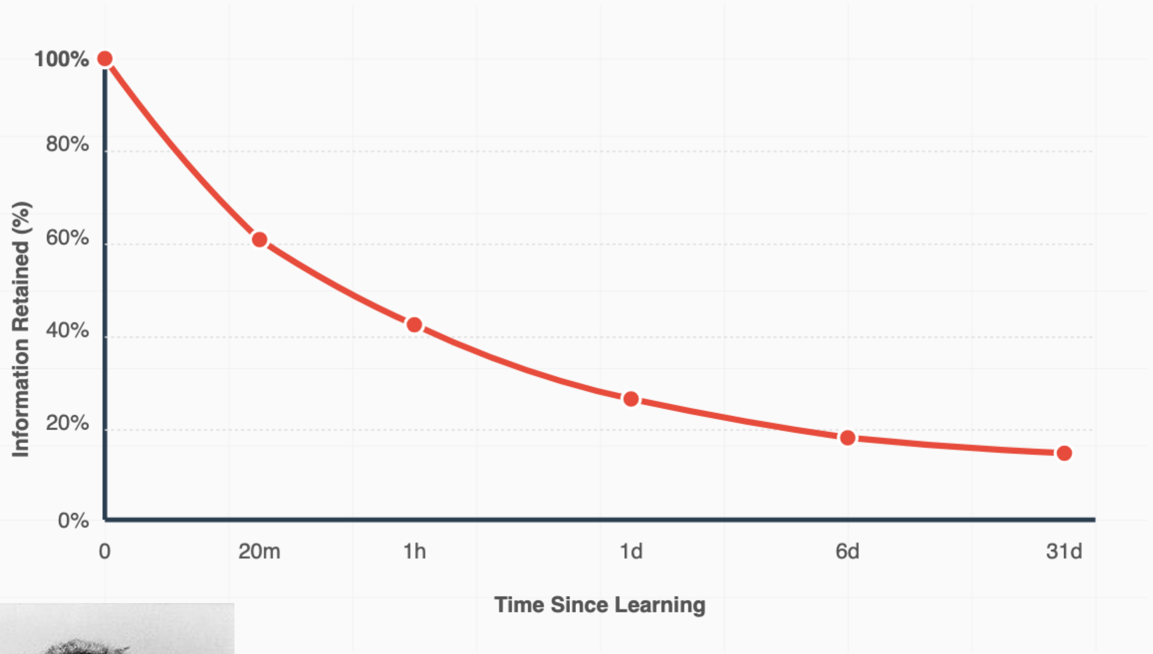
Attention: Nature and constraints

- Attention is a crucial component of consciousness, but it demands effort and depletes cognitive resources.
- Some people value mental effort more than others, which affects their attention allocation.
- What we attend to is often shaped more by stimulus properties and task demands than by deliberate choice

Attention is not sufficient for remembering

- Even if we pay attention, we (usually) forget:
- *The Seven Sins of Memory* (Schacter, 2001):
  1. **Transience** (i.e. forgetting)
  2. Absent-mindedness
  3. Blocking
  4. Misattribution
  5. Suggestibility
  6. Bias
  7. Persistence

Forgetting over time

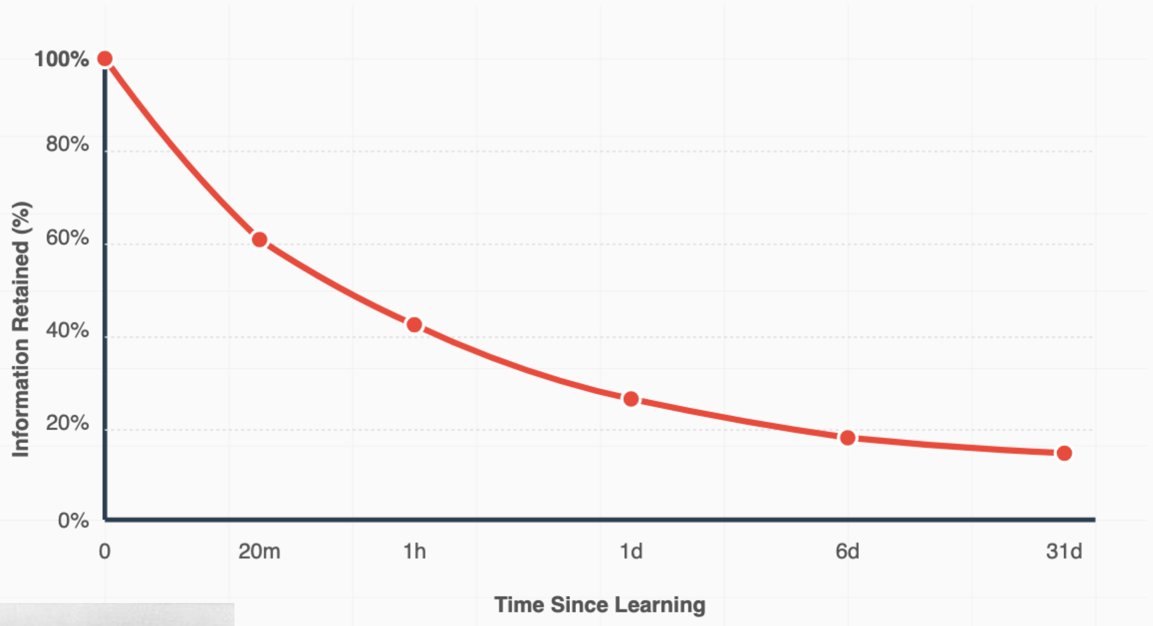


Hermann Ebbinghaus (1885)

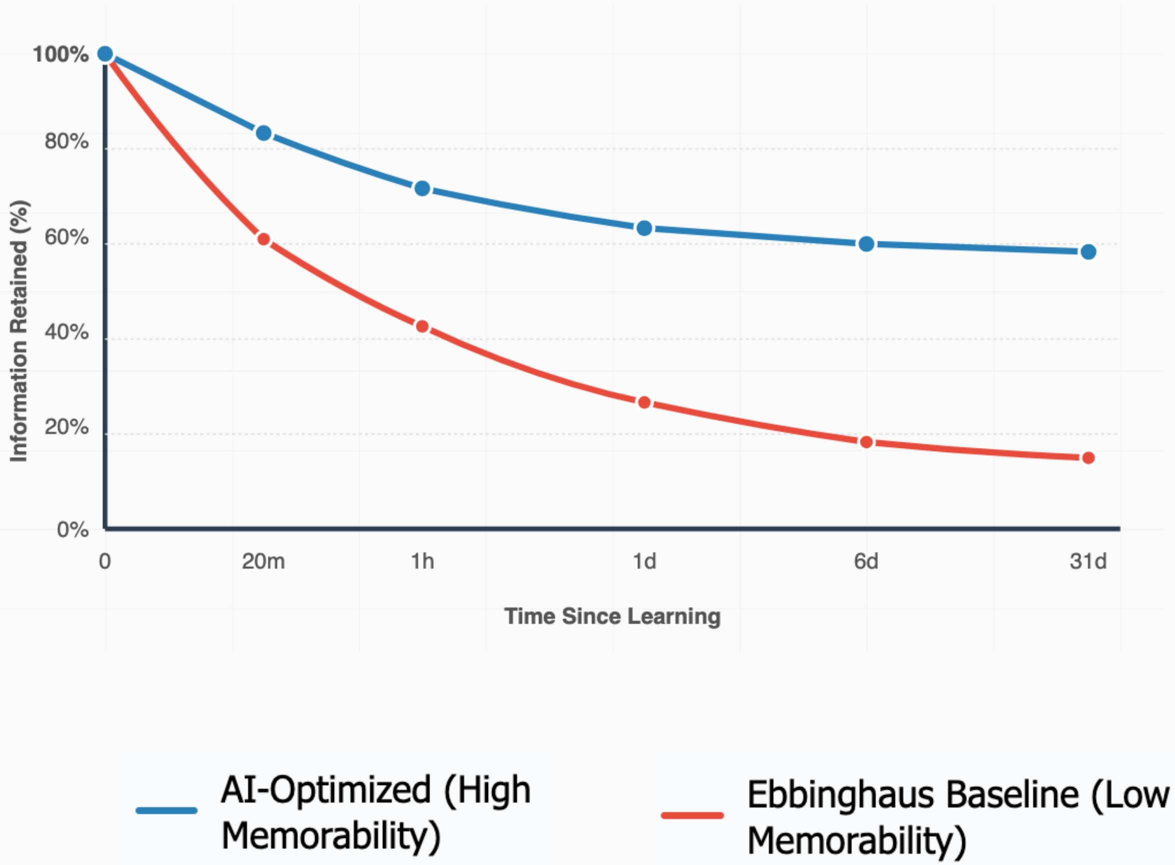
Can boosting *memorability* reshape the forgetting curve?

- Studies have shown that the extent to which a stimulus will be remembered is largely determined by features of the stimulus itself – not by the observer's traits.
- Memory performance in one group of people is a good predictor of memory performance in another group of people.
- If we train artificial neural networks on large-scale memory data from humans, we can generalize predictions to untrained content.

Forgetting over time

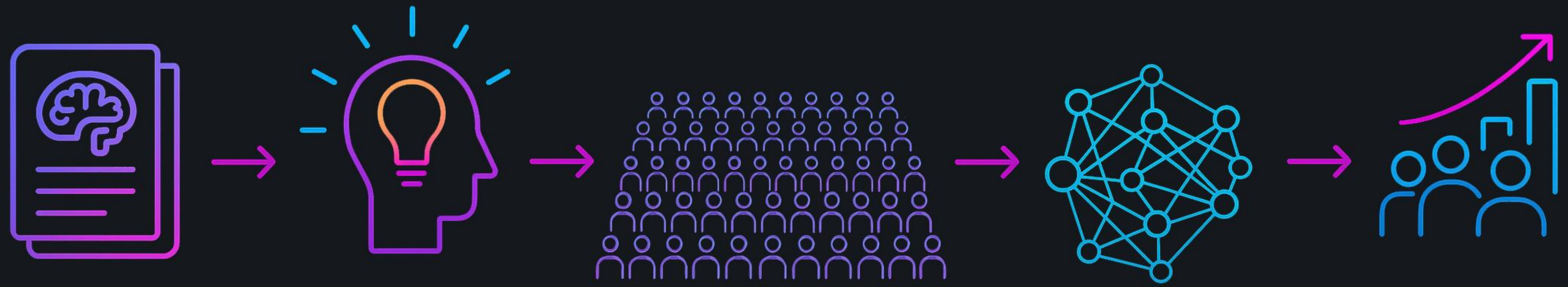


Hermann Ebbinghaus (1885)



Our Cognitive AI Strategy

Mentalese operates at the intersection of cognitive science and AI to bring you tools for optimizing your content. We've developed an AI-powered solution that optimizes text to align with the brain's natural processing patterns.



COGNITIVE NEUROSCIENCE RESEARCH

MENTALESE CONCEPT

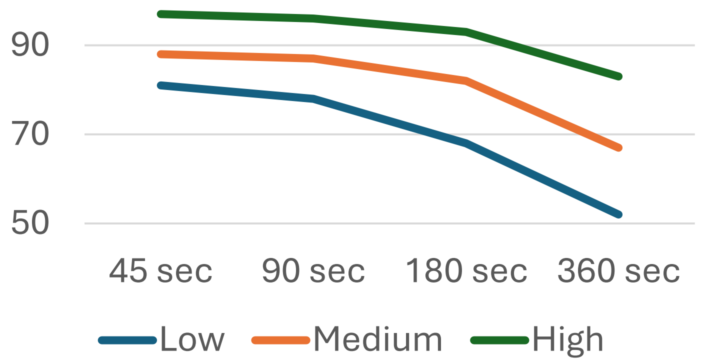
LARGE-SCALE EXPERIMENTS

AI MODELS

OPTIMIZED CONTENT IMPACT

High memorability images*:

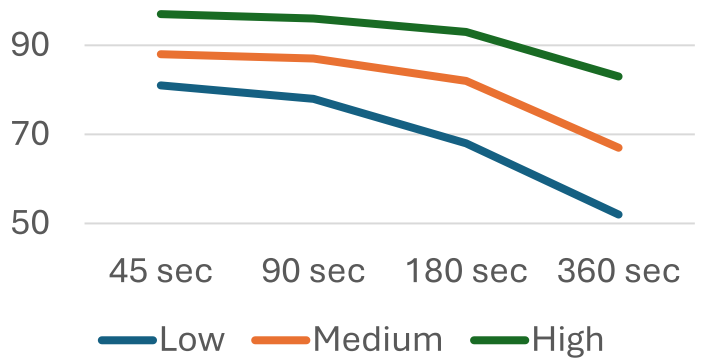
- Are remembered better:



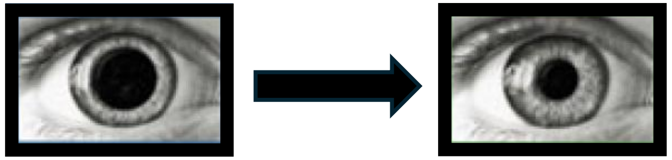
*Results from our lab at the University of Oslo

High memorability images*:

- Are remembered better:



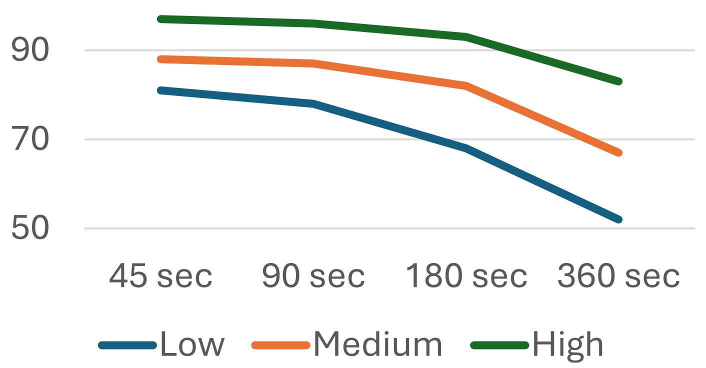
- Require less effort to retrieve:



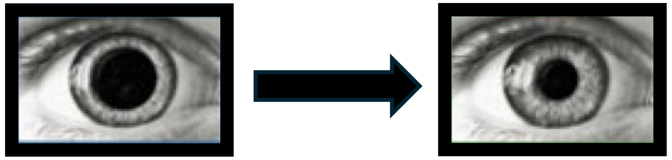
*Results from our lab at the University of Oslo

High memorability images*:

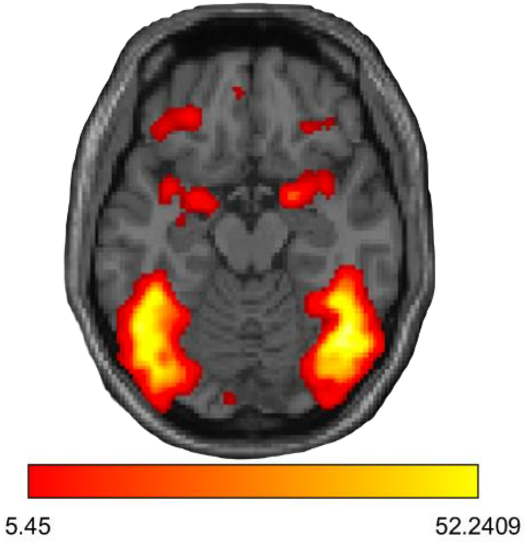
- Are remembered better:



- Require less effort to retrieve:

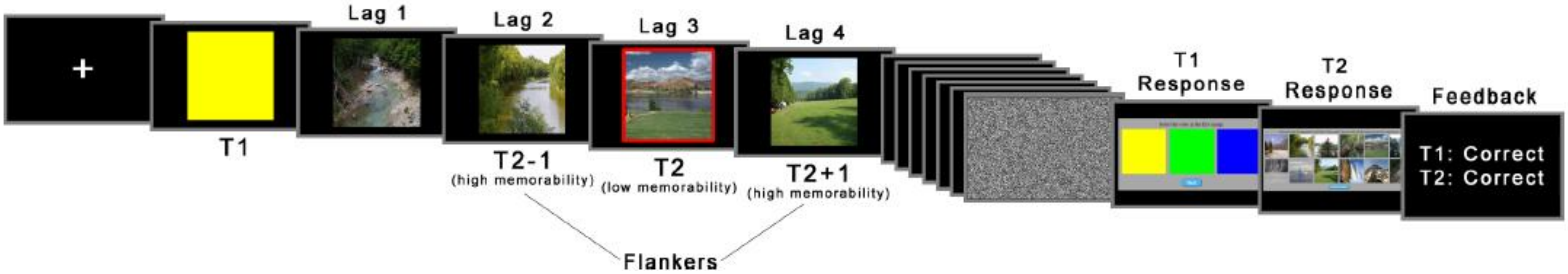


- Elicit stronger brain responses:



*Results from our lab at the University of Oslo

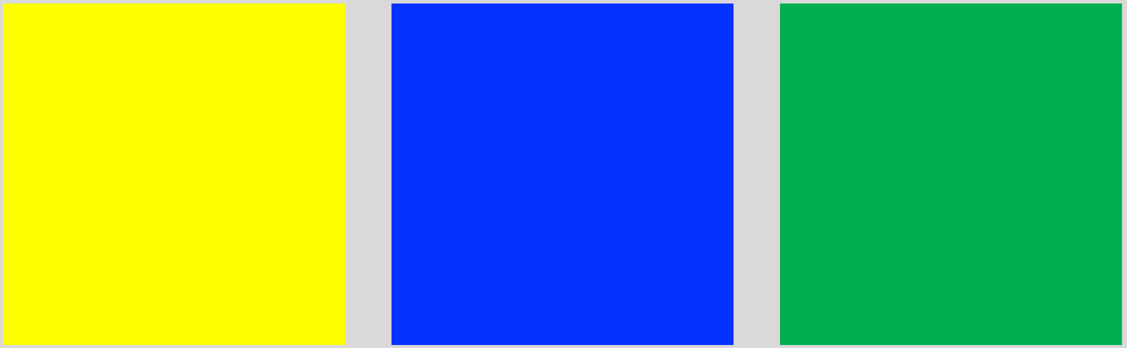
Can memorability interfere with the attentional blink?



Hagen & Espeseth, in prep

LHL

Which colored square was presented?



Which image was the target?

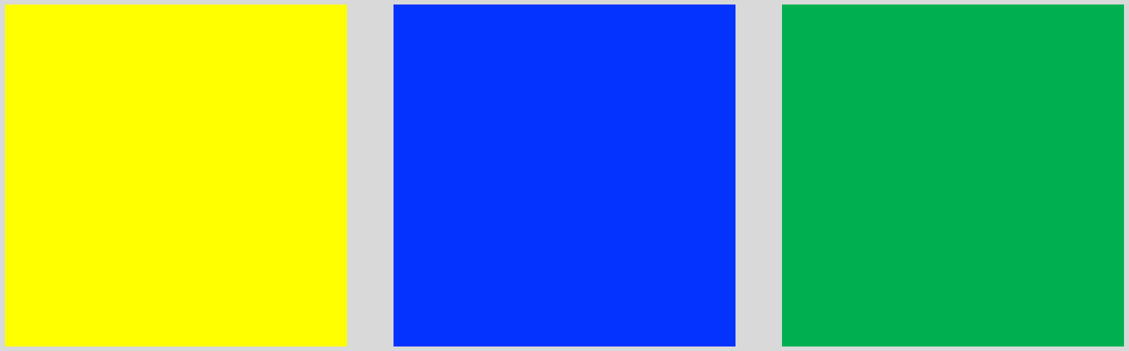


Which image was the target?



HLH

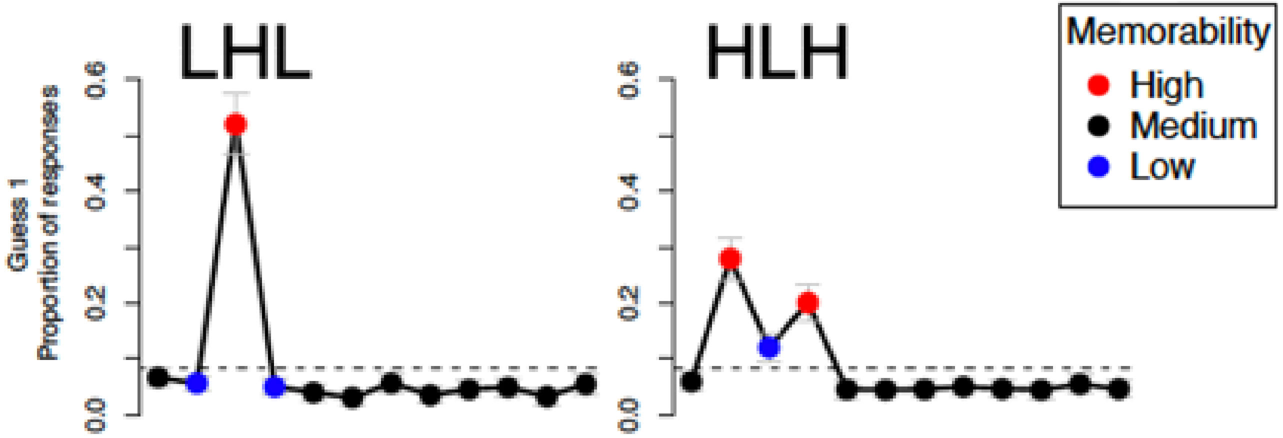
Which colored square was presented?



Which image was the target?

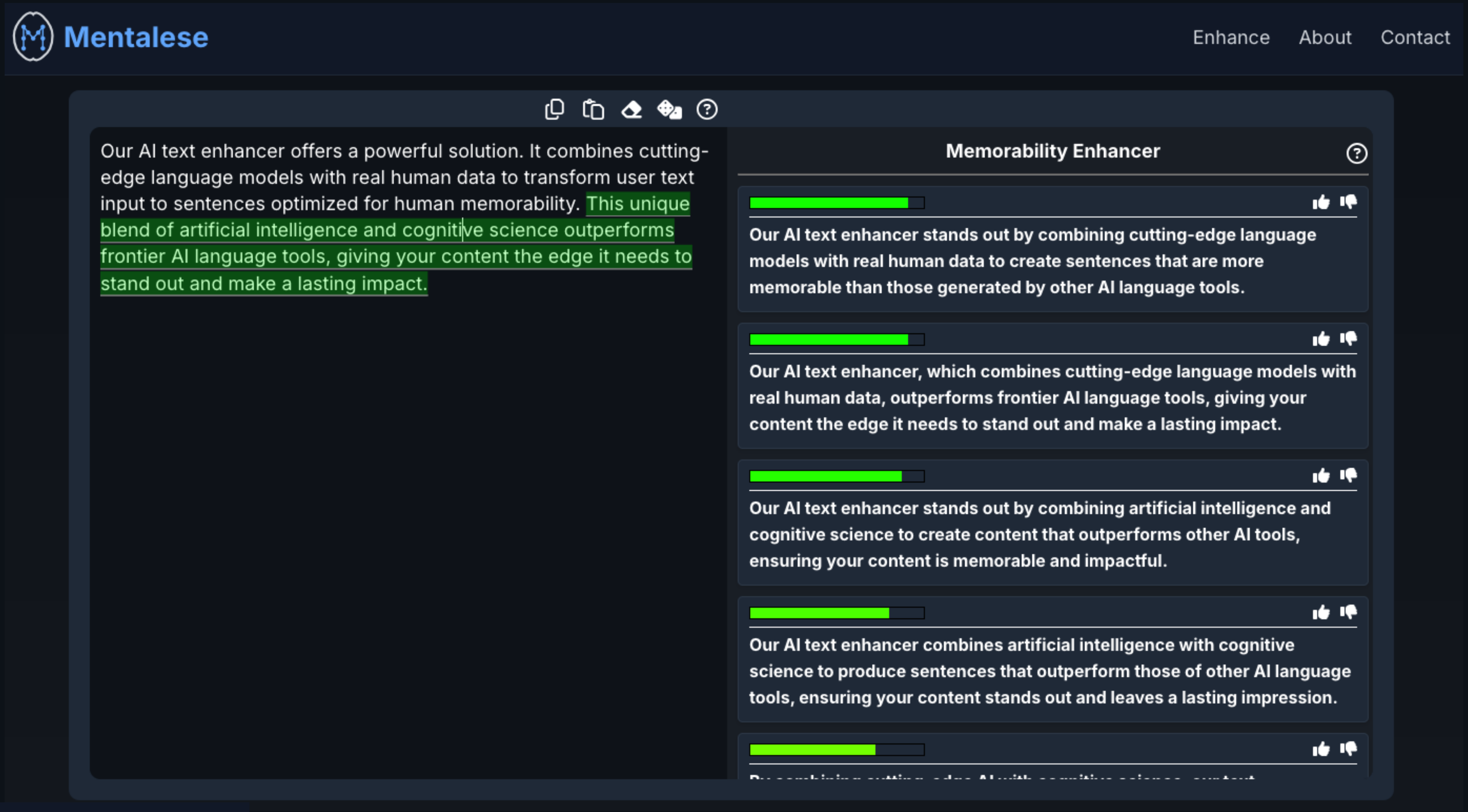
Which image was the target?

Memorability counteracts the attentional blink



Hagen & Espeseth, in prep

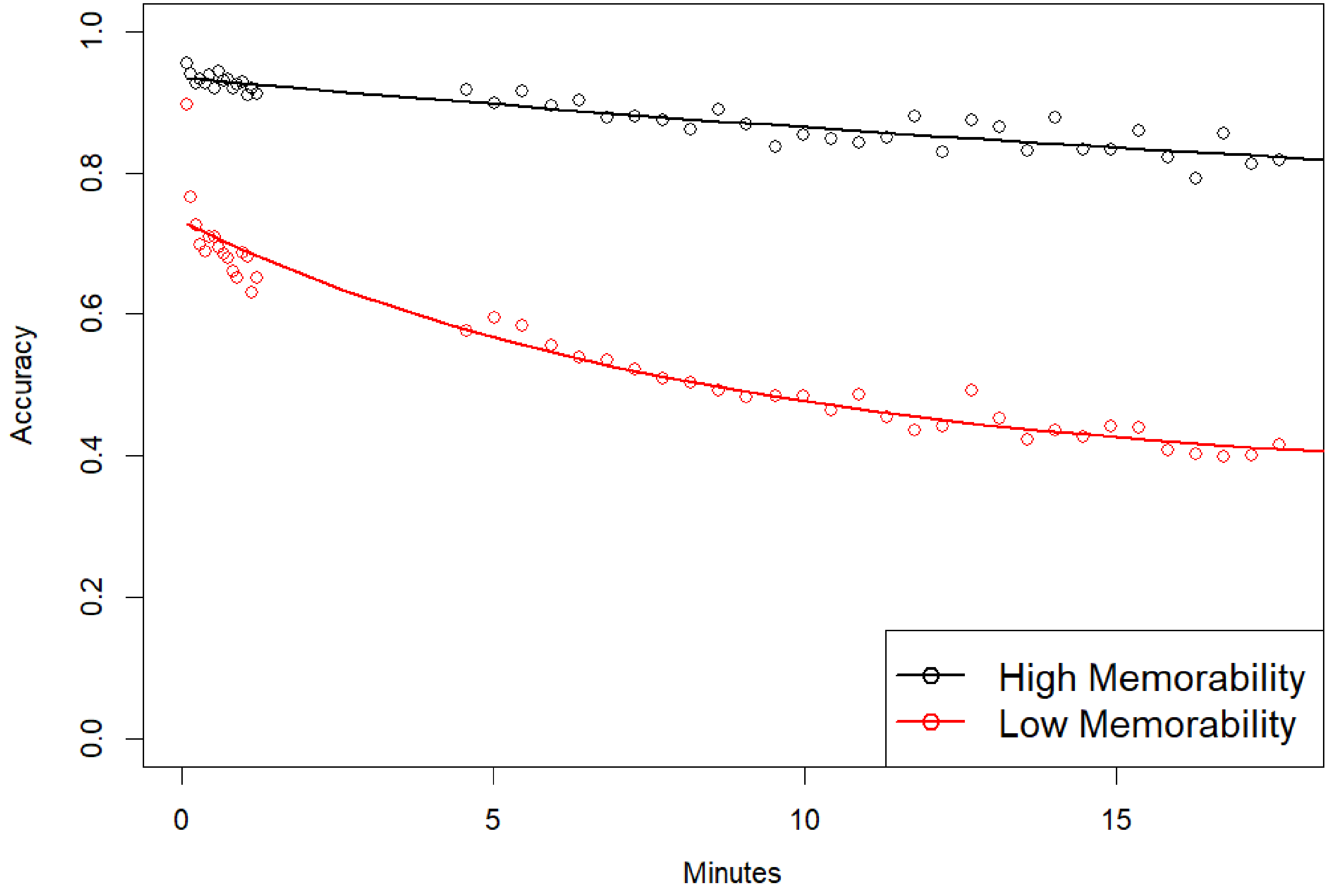
Predicting and Enhancing Text Memorability



Our text memorability tool could be a game-changer for communication and education

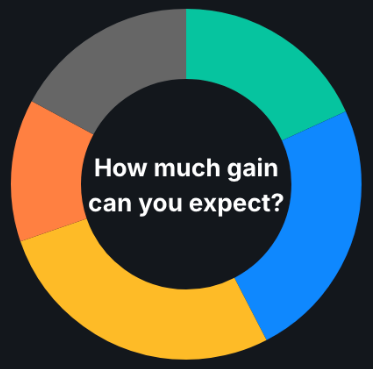
Behavioral data from our text memory games

- 4000 sentences
- 2500 participants



Proven Results Distribution

See how our AI enhancement performs.



Major (50%+)	18.2%
Significant (30-50%)	24.2%
Moderate (10-30%)	27.3%
Minor (1-10%)	13.1%
No Improvement	17.2%

28.4%
Average Improvement

82.8%
Success Rate

Impact Example

For content with 1,000 views:

Original content remembered by **330 people**

Enhanced content remembered by **424 people**

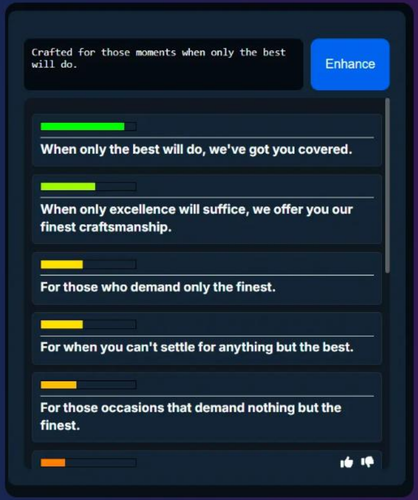
+94 more people (+28.4%)
remember your message

Use case in advertising

Before

Low memorability

AI Optimization



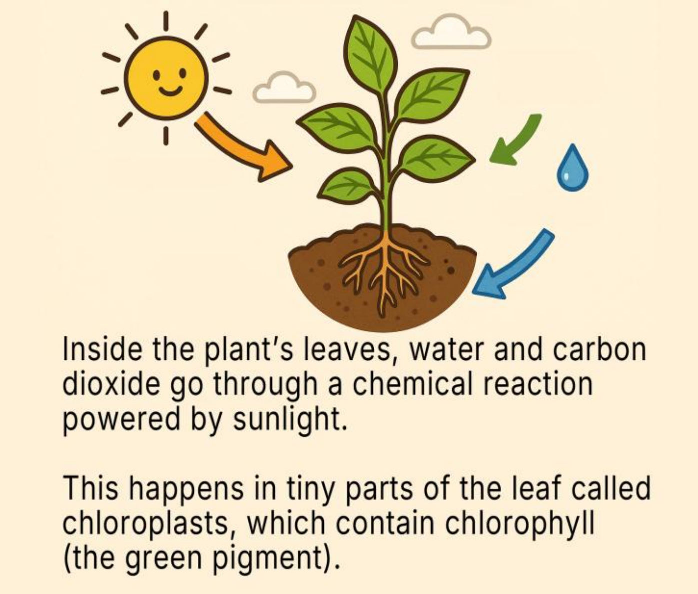
Enhancing memorability

After

High memorability

Use case in education

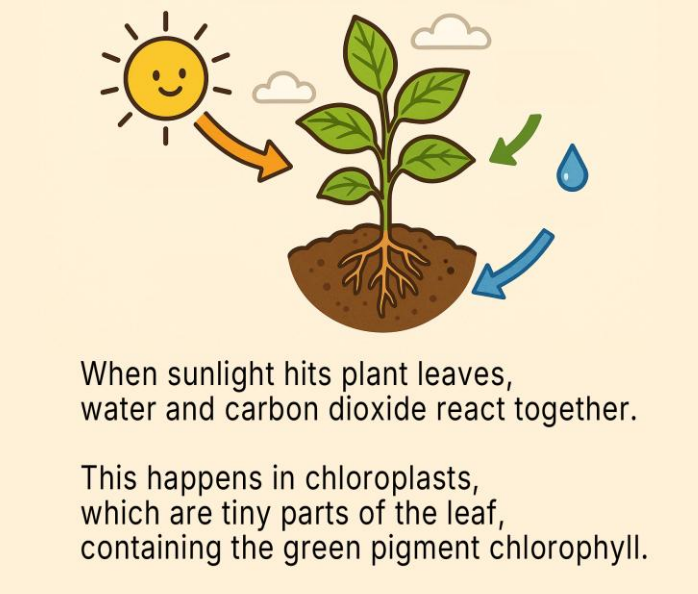
Before



Low memorability

AI Optimization

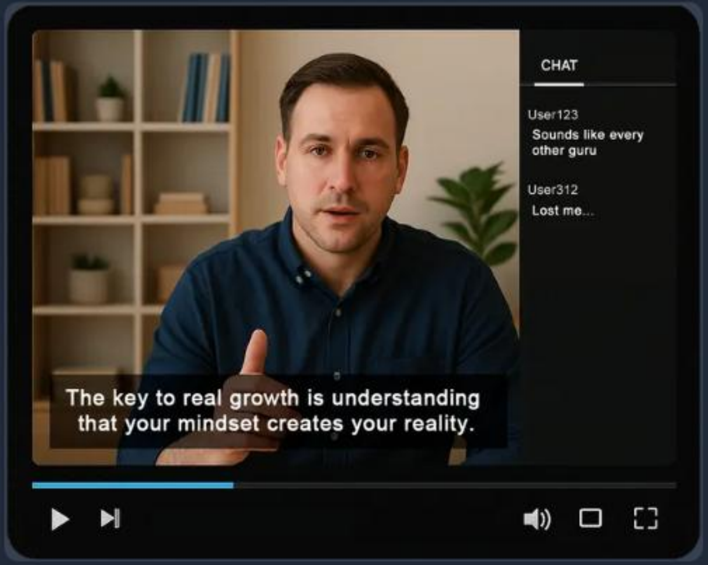
After



High memorability

Use case in content production

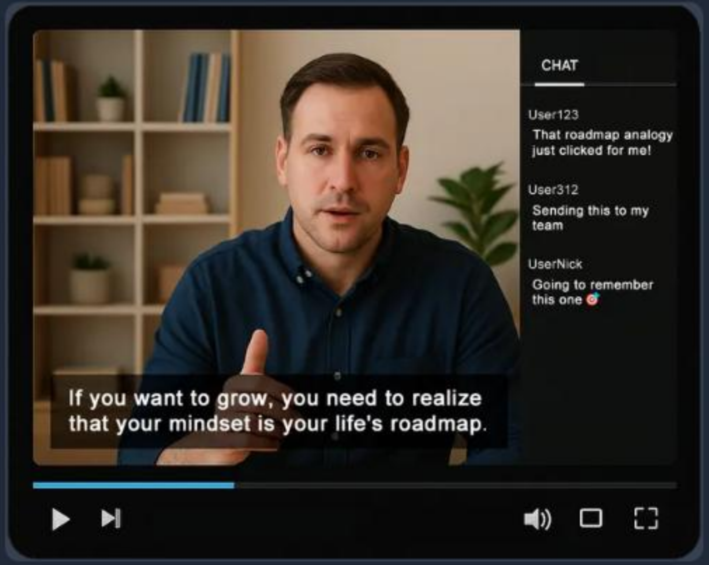
Before



Low memorability

AI Optimization

After



High memorability

Conclusion

- Attention has limited capacity and memory is transient
- Boosting content memorability **modulates** forgetting by elevating retention early and decelerating its decline over time
- Training AI on large-scale memory data has great potential for prediction and generation of high impact content
- A promising approach complementary to improving learning strategies

Image Memorability Prediction with Vision Transformers

Thomas Hagen[1,✉] **and Thomas Espeseth**[1,2]

[1]Department of Psychology, University of Oslo, Oslo, Norway
[2]Department of Psychology, Oslo New University College, Oslo, Norway

Behavioral studies have shown that the memorability of images is similar across groups of people, suggesting that memorability is a function of the intrinsic properties of images, and is unrelated to people's individual experiences and traits. Deep learning networks can be trained on such properties and be used to predict memorability in new data sets. Convolutional neural networks (CNN) have pioneered image memorability prediction, but more recently developed vision transformer (ViT) models may have the potential to yield even better predictions. In this paper, we present the ViTMem, a new memorability model based on ViT, and evaluate memorability predictions obtained by it with state-of-the-art CNN-derived models. Results showed that ViTMem performed equal to or better than state-of-the-art models on all data sets. Additional semantic level analyses revealed that ViTMem is particularly sensitive to the semantic content that drives memorability in images. We conclude that ViTMem provides a new step forward, and propose that ViT-derived models can replace CNNs for computational prediction of image memorability. Researchers, educators, advertisers, visual designers and other interested parties can leverage the model to improve the memorability of their image material.

stimulus set to predict memory performance in a new group of participants.

These results have been replicated and extended in a number of studies, revealing that similar findings are obtained with different memory tasks (2), different retention times (1, 2), different contexts (3), and independent of whether encoding is intentional or incidental (4). However, although image memorability has proven to be a robust and reliable phenomenon, it has not been straightforward to pinpoint the image properties that drive it. What seems clear though, is that memorability is multifaceted (5, 6). One way to characterize the underpinnings of memorability is to investigate the contribution from processes at different levels of the visual processing stream. For example, at the earliest stages of processing of a visual scene, visual attributes such as local contrast, orientation, and color are coded. At an intermediate level, contours are integrated, surfaces, shapes, and depth cues are segmented, and foreground and background are distinguished. At a higher level, object recognition is conducted

Table 2. Model performance on LaMem and MemCat combiend dataset

Model	MSE Loss ↓	Spearman ρ ↑
ResMem	0.009	0.67
ViTMem	0.005	0.77

Hagen & Espeseth, 2023

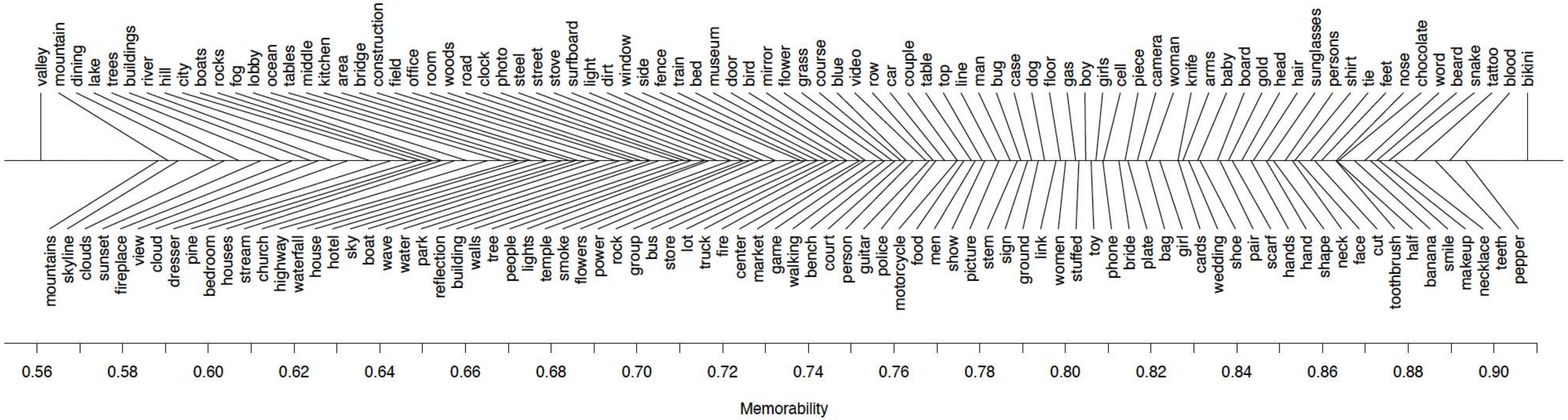


Fig. 1. Average behavioral image memorability scores for nouns that were extracted from images in the LaMem and MemCat data sets. The nouns shown are those that occurred most frequently or that are more frequent in the English language (38).

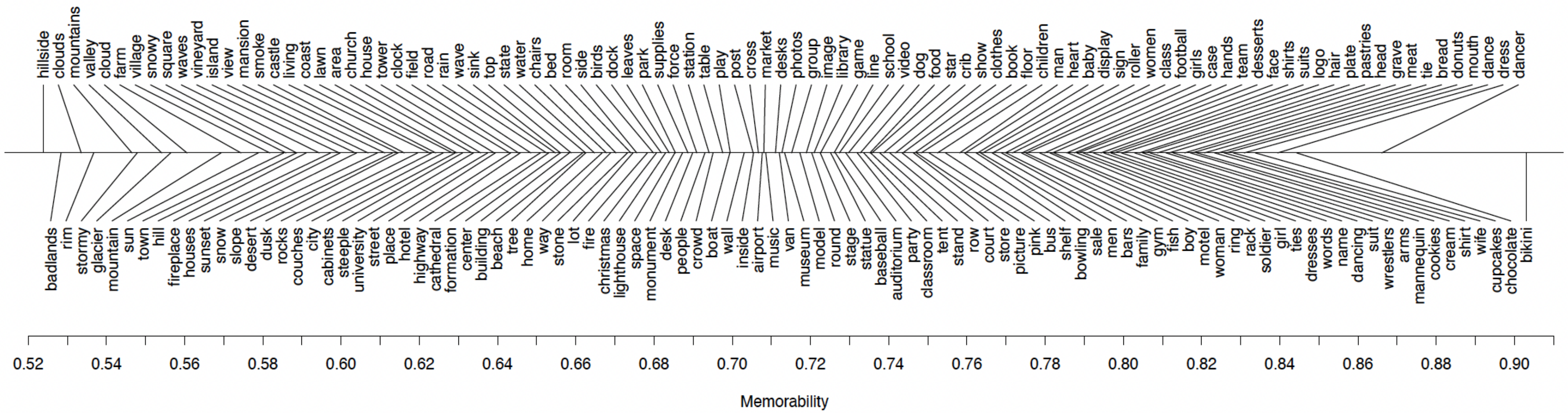


Fig. 2. Average ViTMem predicted image memorability scores for nouns that were extracted from images in the Places205 data set. The nouns shown are those that occurred most frequently or that are more frequent in the English language (38).

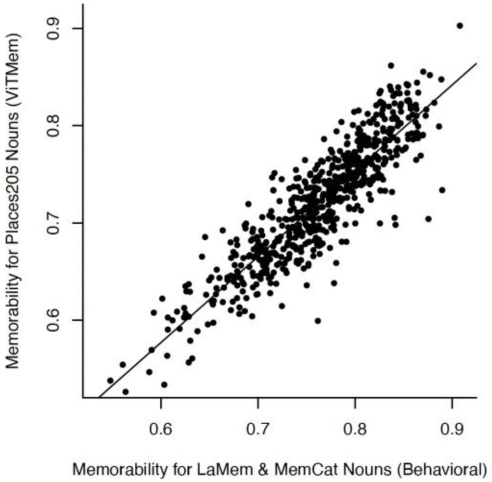


Fig. 3. Average memorability scores for images with matching nouns in different data sets. The y-axis shows average predicted memorability scores from ViTMem on the Places205 data set. The x-axis shows average behavioral memorability scores on the combined LaMem and MemCat data set.

Hagen & Espeseth, 2023

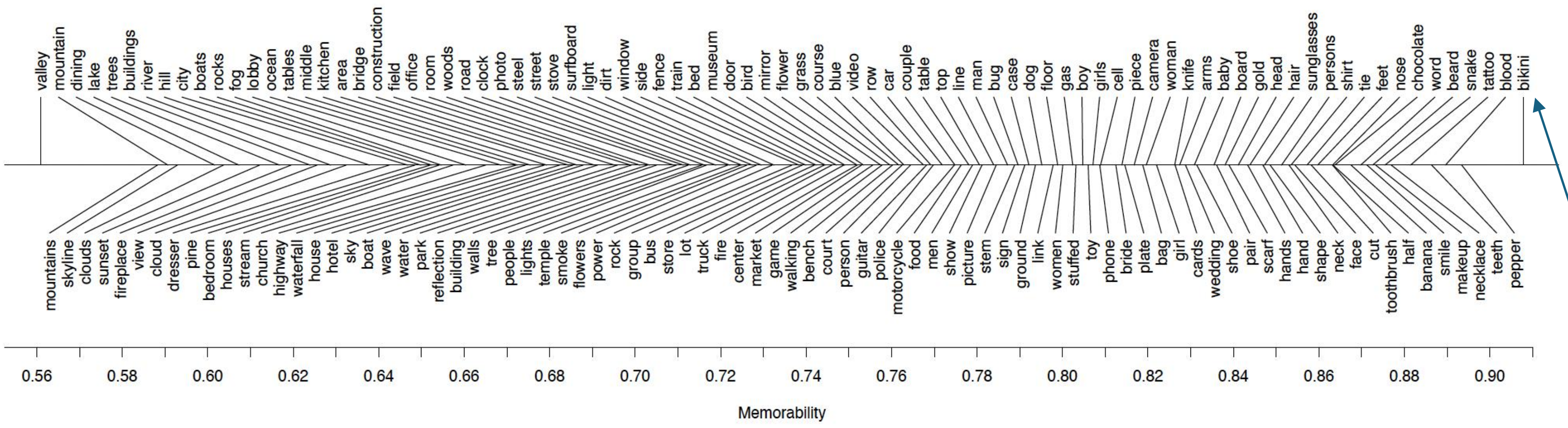


Fig. 1. Average behavioral image memorability scores for nouns that were extracted from images in the LaMem and MemCat data sets. The nouns shown are those that occurred most frequently or that are more frequent in the English language (38).

bikini

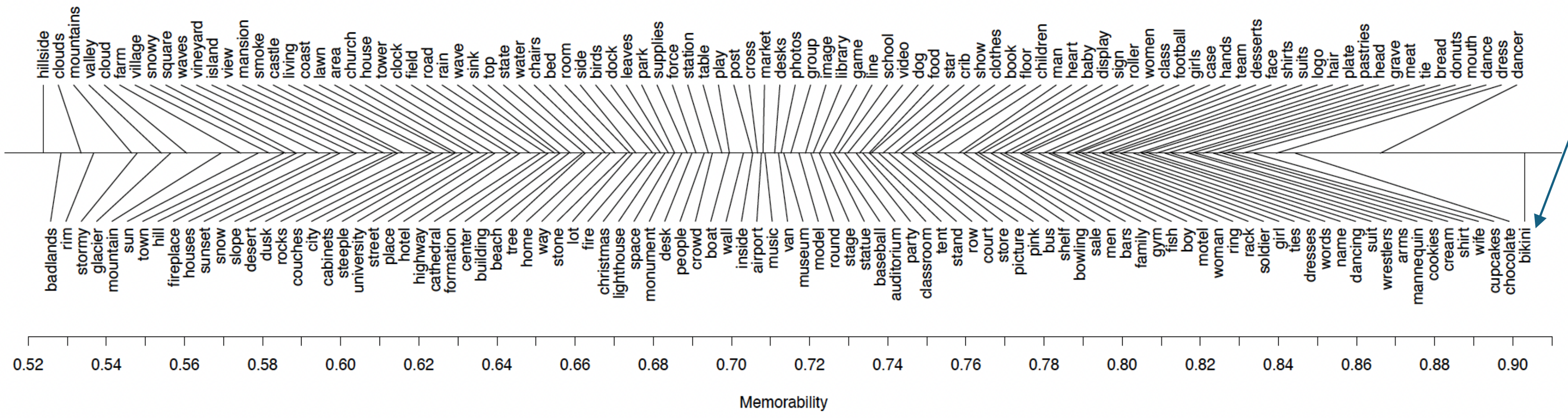


Fig. 2. Average ViTMem predicted image memorability scores for nouns that were extracted from images in the Places205 data set. The nouns shown are those that occurred most frequently or that are more frequent in the English language (38).

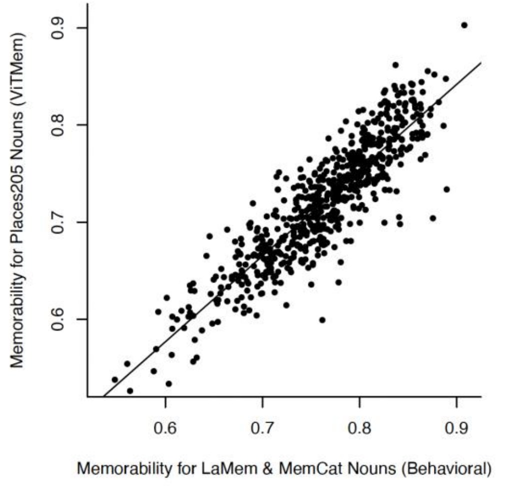


Fig. 3. Average memorability scores for images with matching nouns in different data sets. The y-axis shows average predicted memorability scores from ViTMem on the Places205 data set. The x-axis shows average behavioral memorability scores on the combined LaMem and MemCat data set.

Hagen & Espeseth, 2023

Memorability

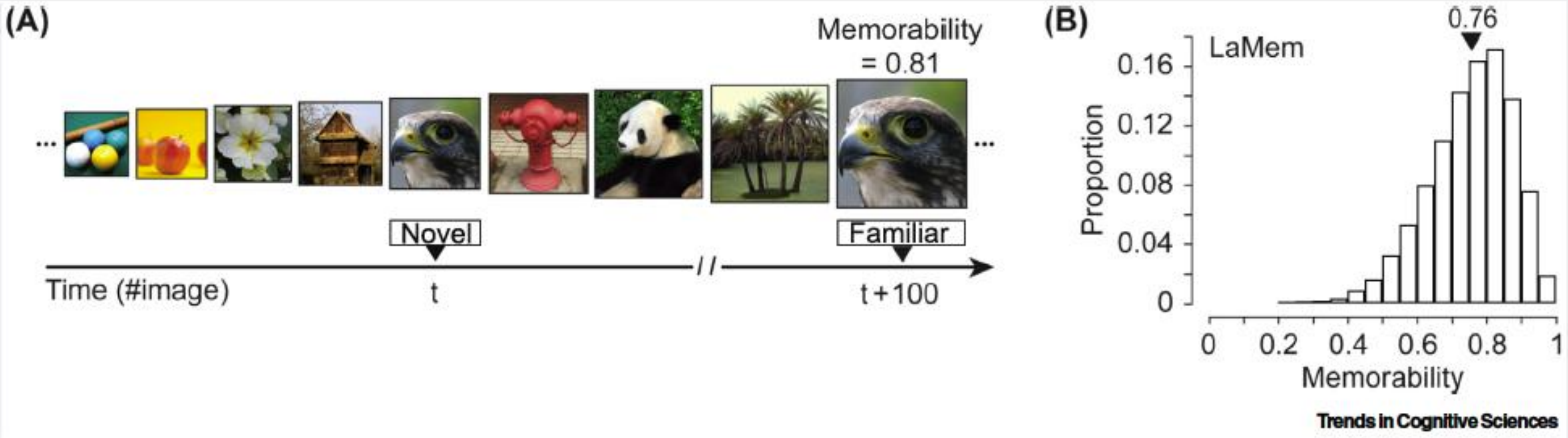


Figure I. Quantifying Image Memorability. (A) The visual recognition memory task used to compute image memorability scores. On each trial, subjects judge whether images are novel or familiar. Memorability scores are computed based on the subject-average performance for familiar images, corrected for false alarms. (B) Distribution of memorability scores for the LaMem data set, ~60 000 images pulled from a diversity of sources [5].

Rust & Mehrpour, 2020