Human Memory Capacity for Object and Scene Representation

Aude Oliva

Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology

oliva@mit.edu
http://cvcl.mit.edu

In collaboration with

Timothy Brady, Talia Konkle & George Alvarez

brain+cognitive sciences

Research funded by NSF CAREER IIS # 0546262

Demo ...

What's the Capacity of Human Memory?

What we know…

Standing (1973)

10,000 images

83% Recognition

… people can remember thousands of images

What we don't know…

… what people are remembering for each item?

<u>According to Standing</u>

"Basically, my recollection is that we just separated the pictures into distinct thematic categories: e.g. cars, animals, single-person, 2-people, plants, etc.) Only a few slides were selected which fell into each category, and they were visually distinct."

Dogs
Playing Cards

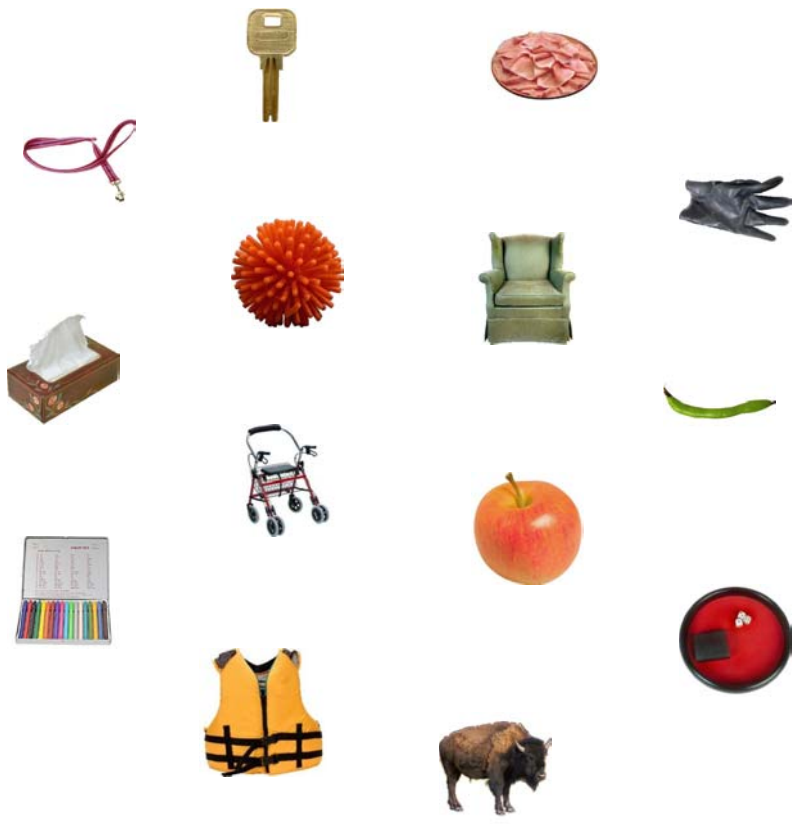
"Gist" Only

Sparse Details

Highly Detailed

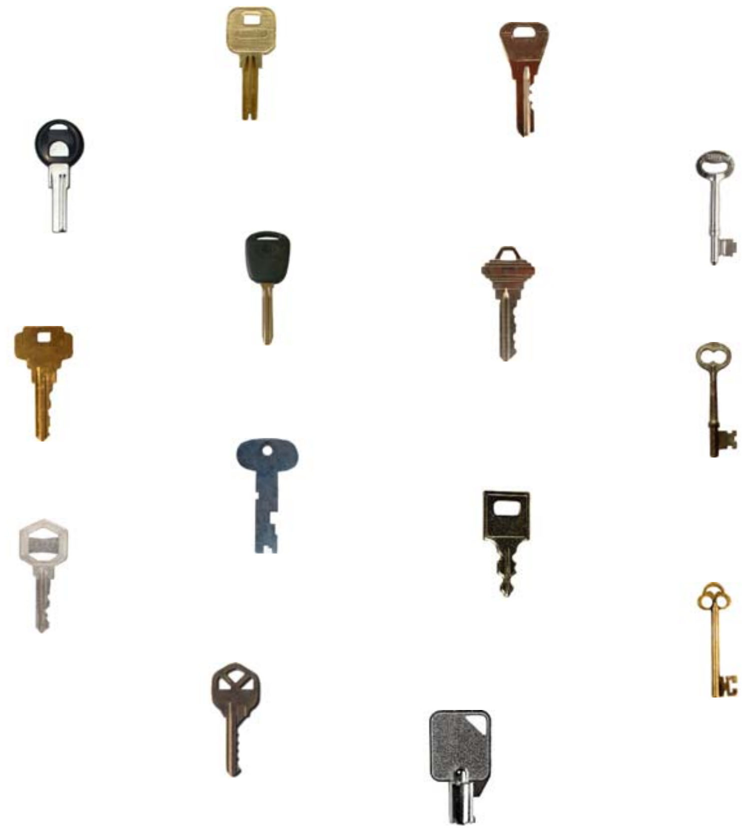
What’s the Capacity of Human Memory?

Quantity



Standing (1973): 82% of 10,000

Fidelity ?



Hollingworth (2004), Vogt & Magnussen (2007): 400 exemplars ~ 62% recognition

Massive Memory I: Thousands of objects

Massive Memory Experiment I

A stream of objects will be presented on the screen for ~ 1 second each.

Your primary task:

Remember them ALL!

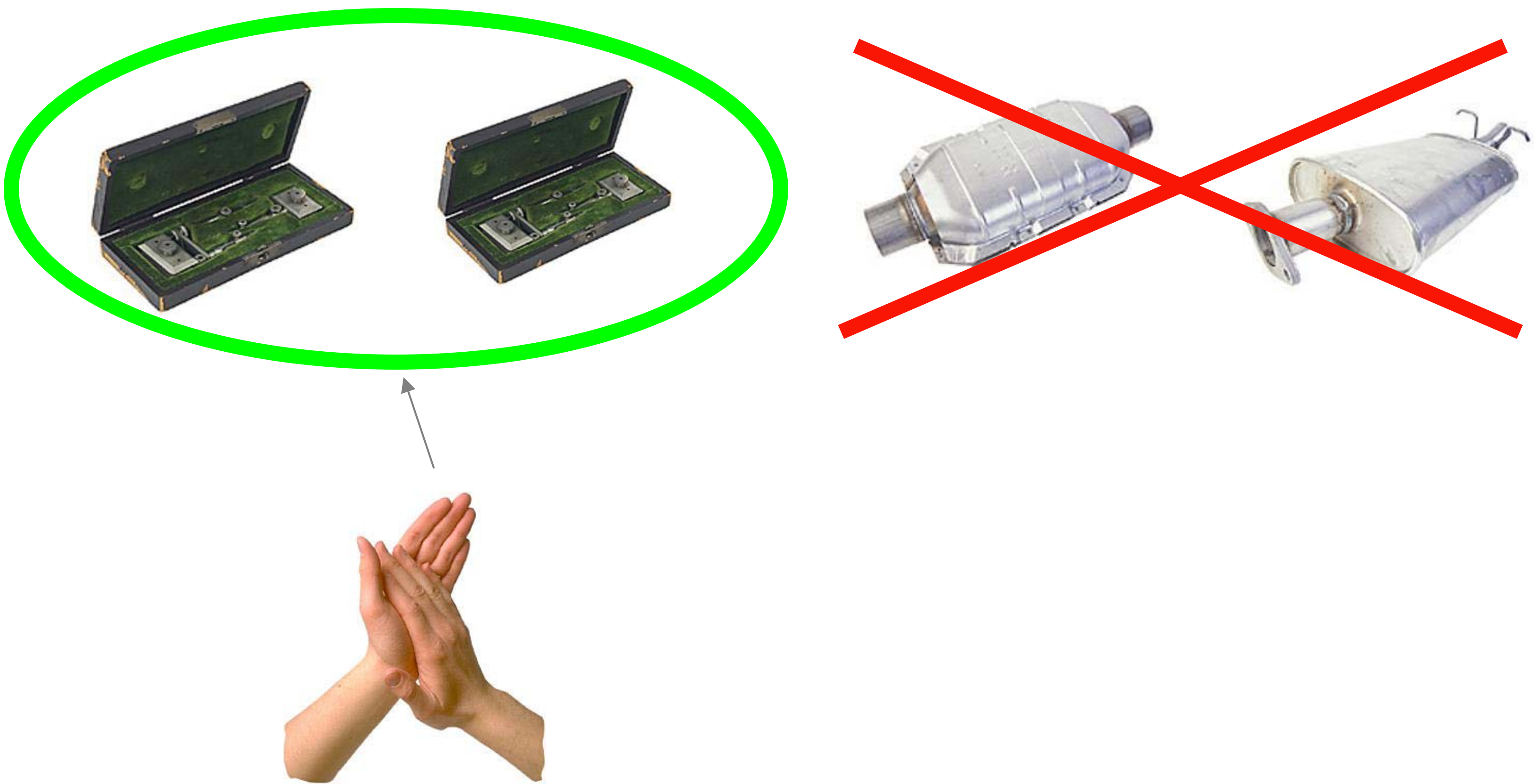
afterwards you will be tested with...

Completely different objects...

Different exemplars of the same kind of object...

Massive Memory Experiment I

Your other task: **Detect exact repeats anywhere in the stream**



Massive Memory I: Methods

Showed 20 observers **2560** unique objects

from **480** different object categories

Number of objects per category varied from **1 to 16**

N-back, detect exact repeats, **2 to 1024 back**

Followed by 240 **2-alternative forced choice tests**

Bagel

Backpack

Bucket

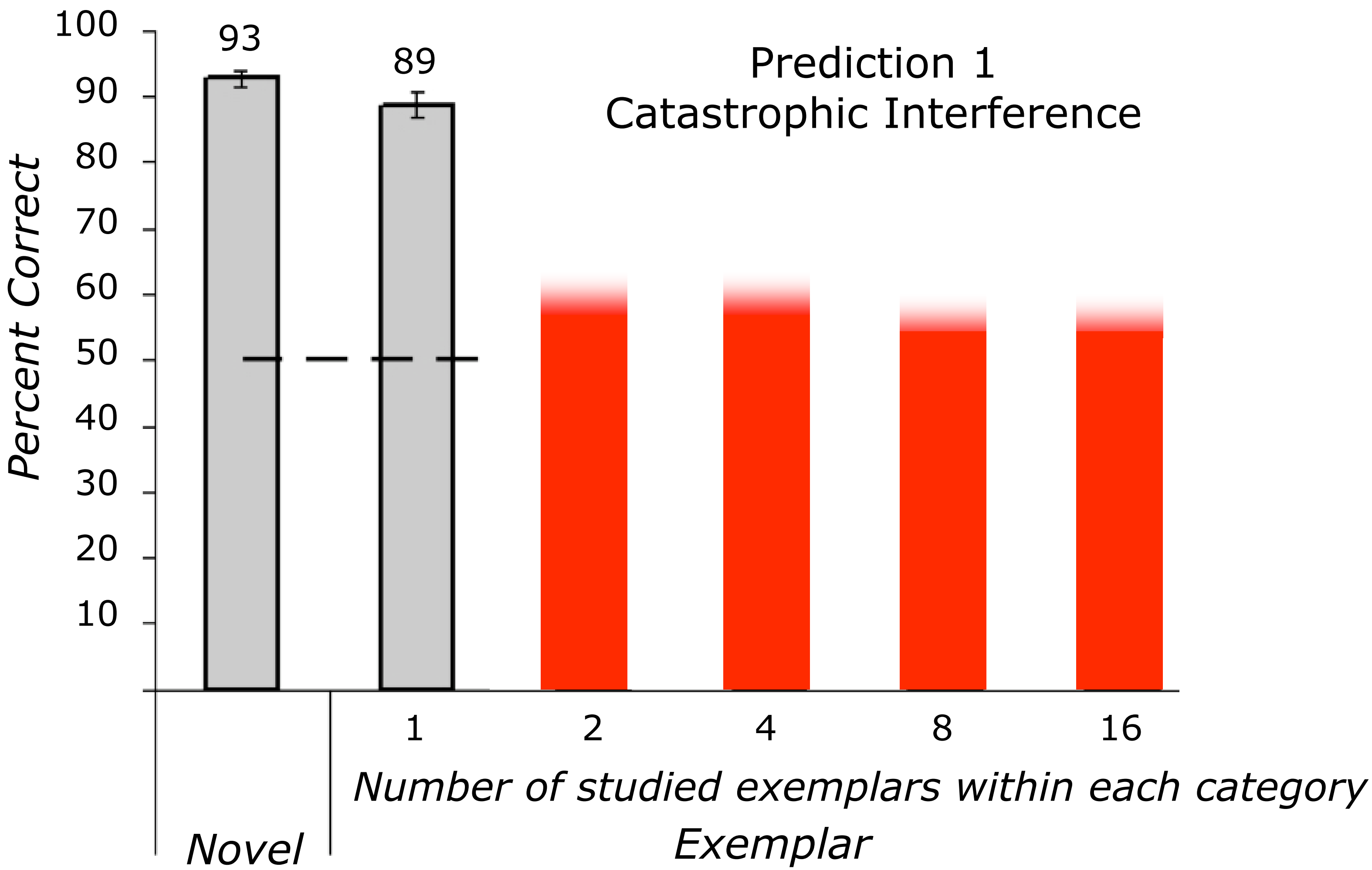
Camera

Cup saucer

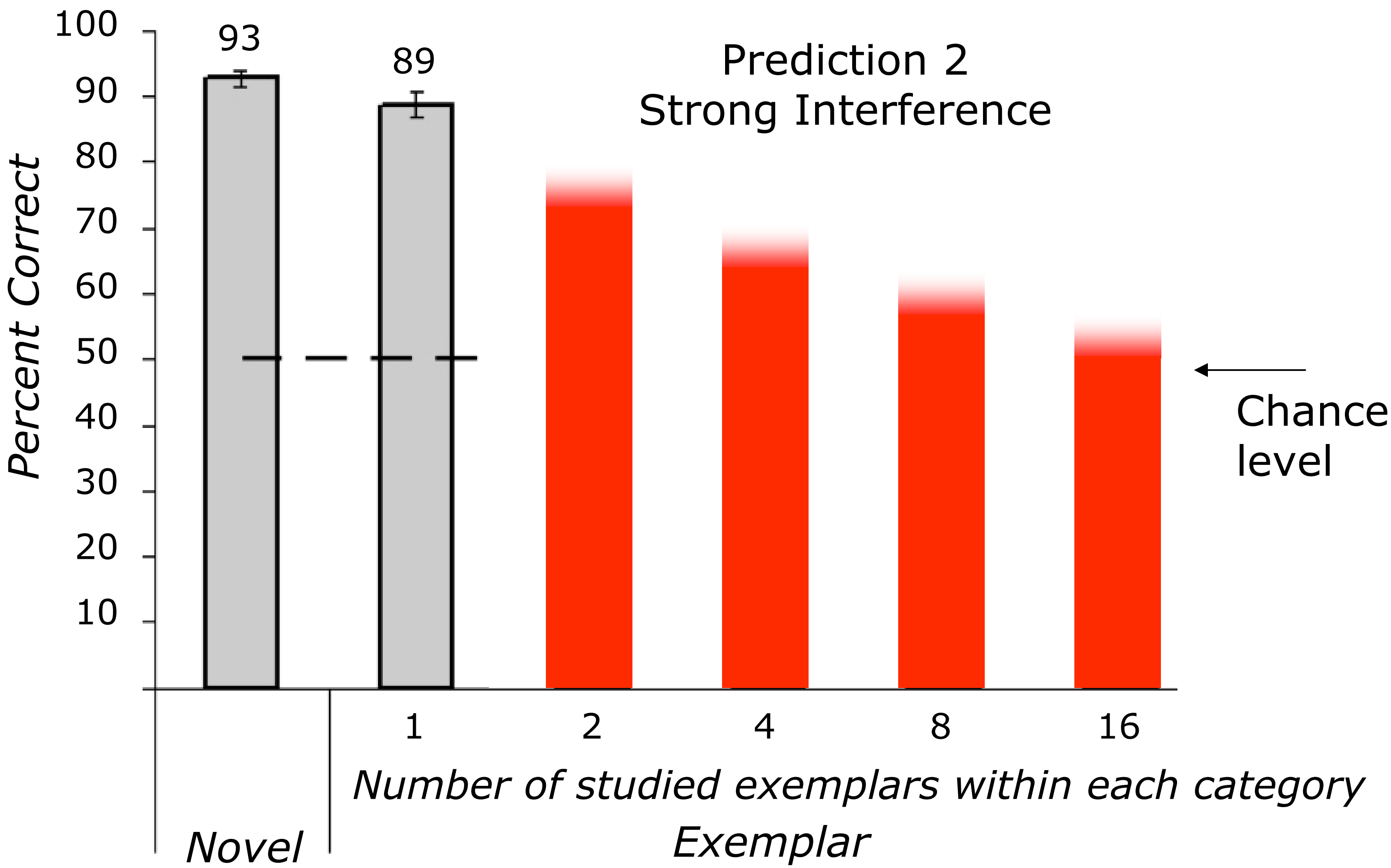
Tent

Watergun

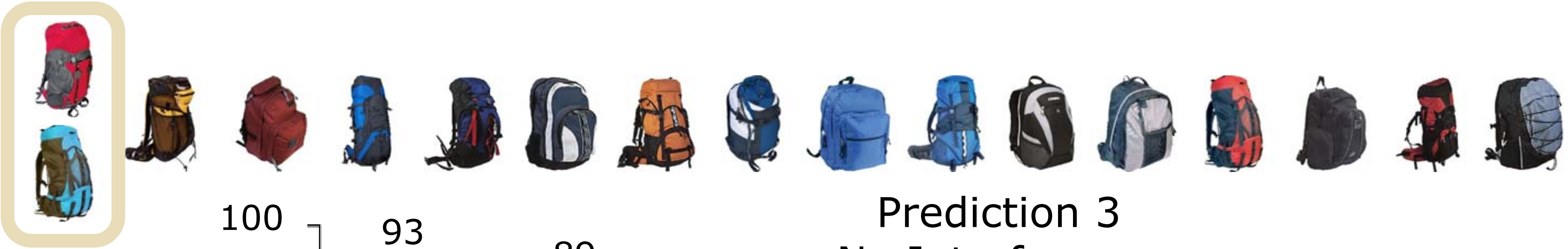
Recognition Memory Performance

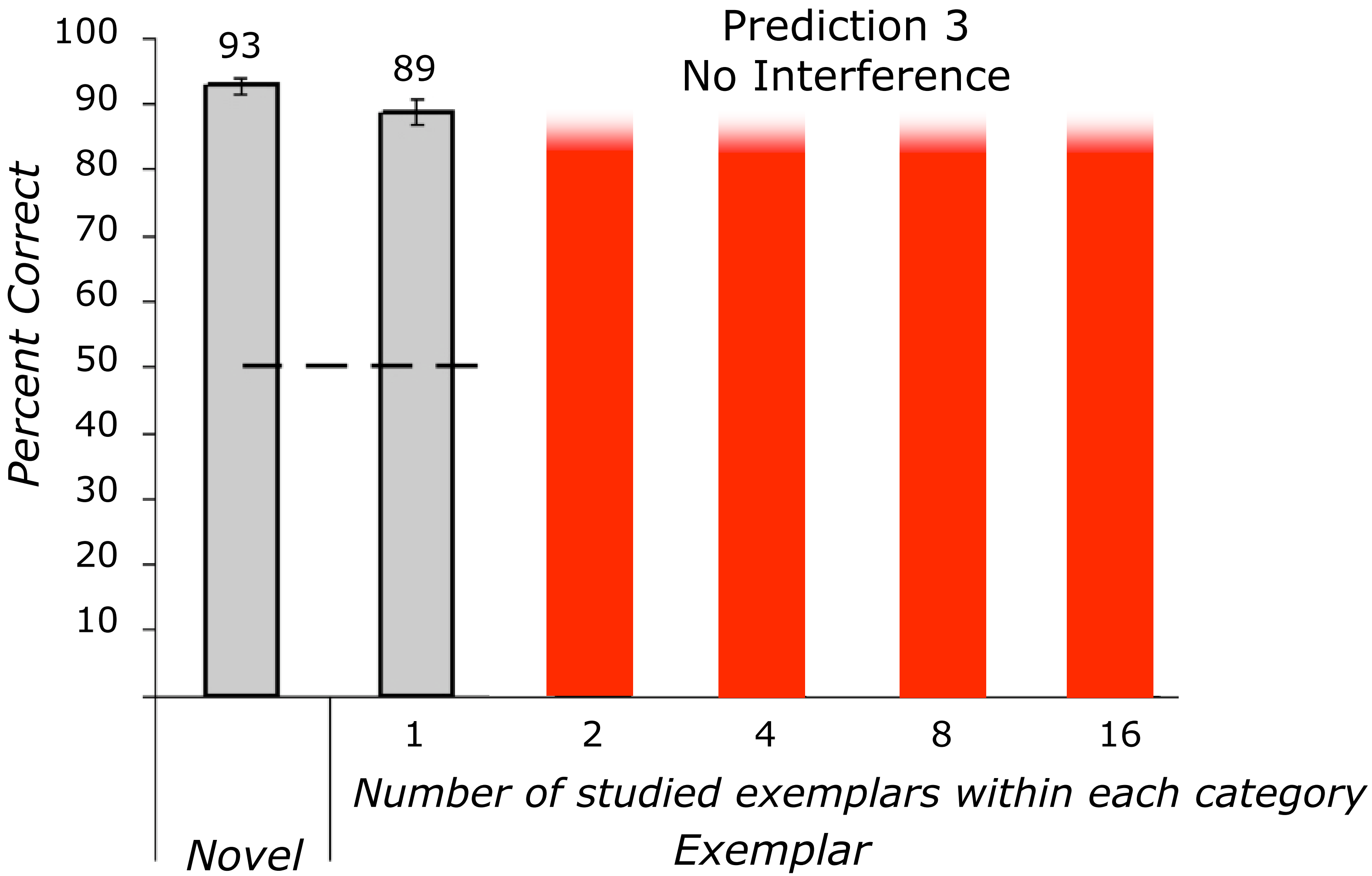


Recognition Memory Performance

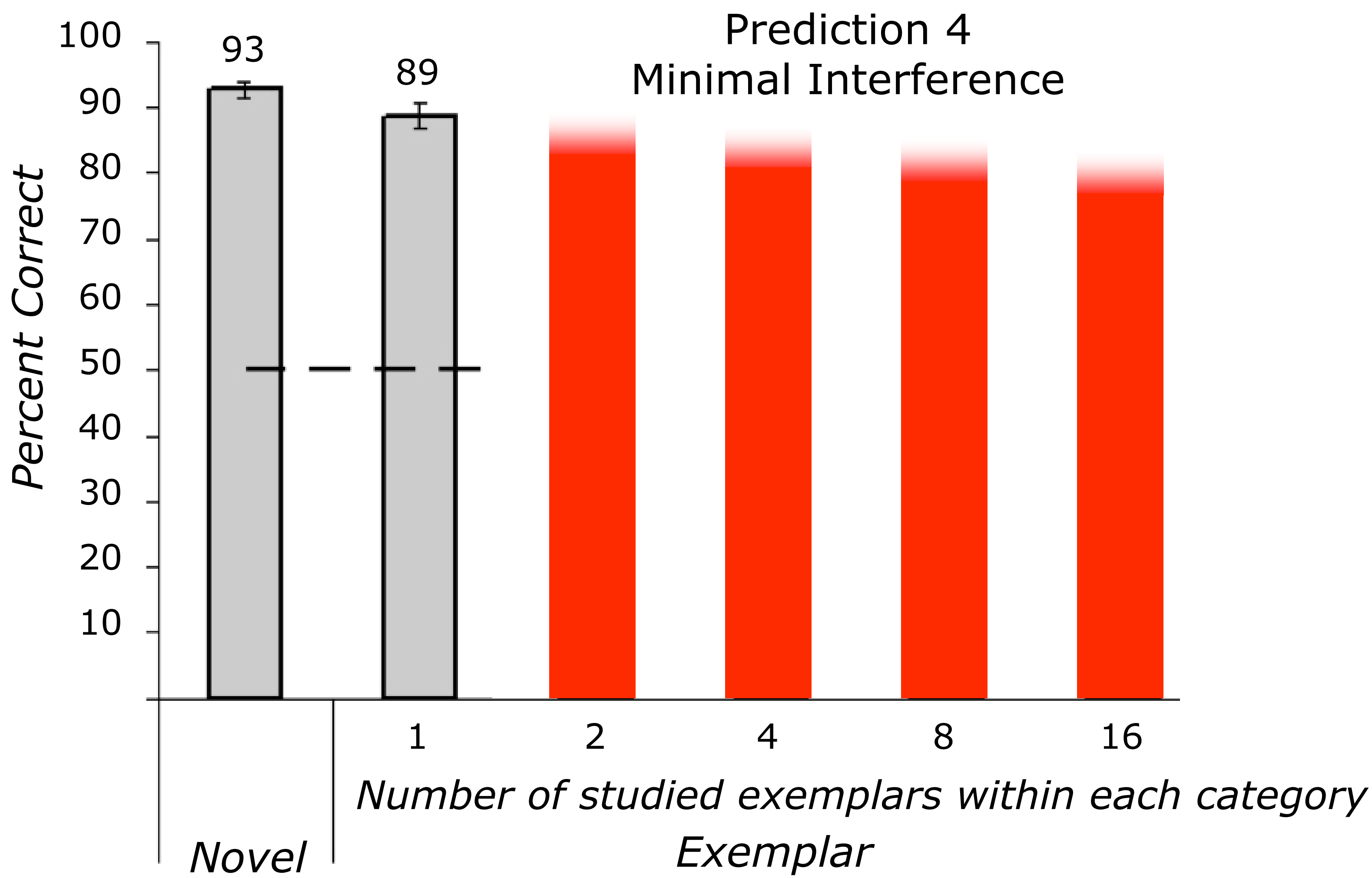


Recognition Memory Performance



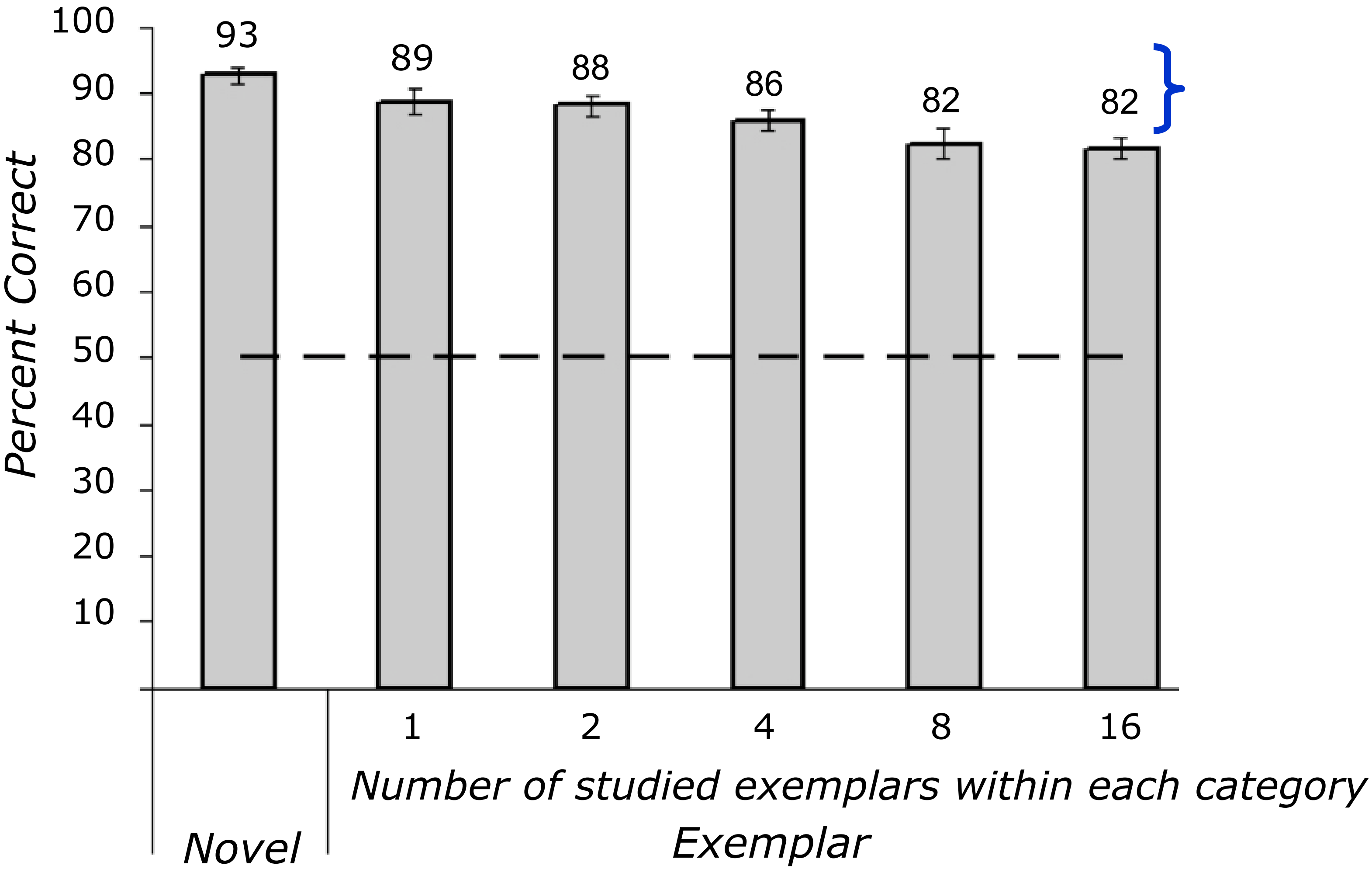


Recognition Memory Performance



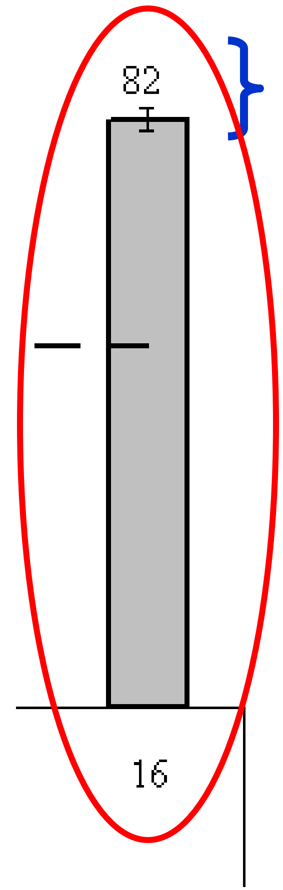
Recognition Memory Performance

Highly Detailed
Minor Interference



Recognition Memory Performance

Highly Detailed
Minor Interference



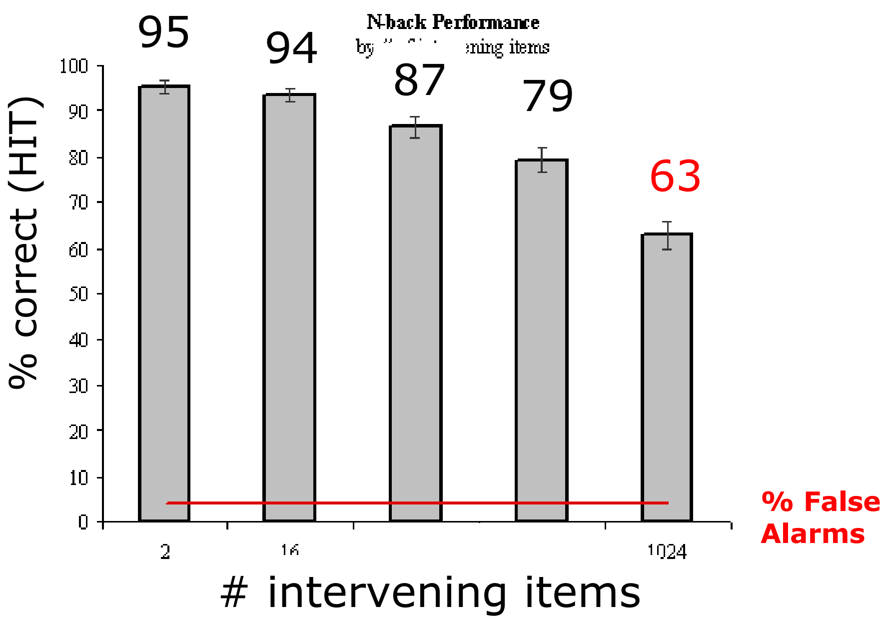
What about detection performances ?

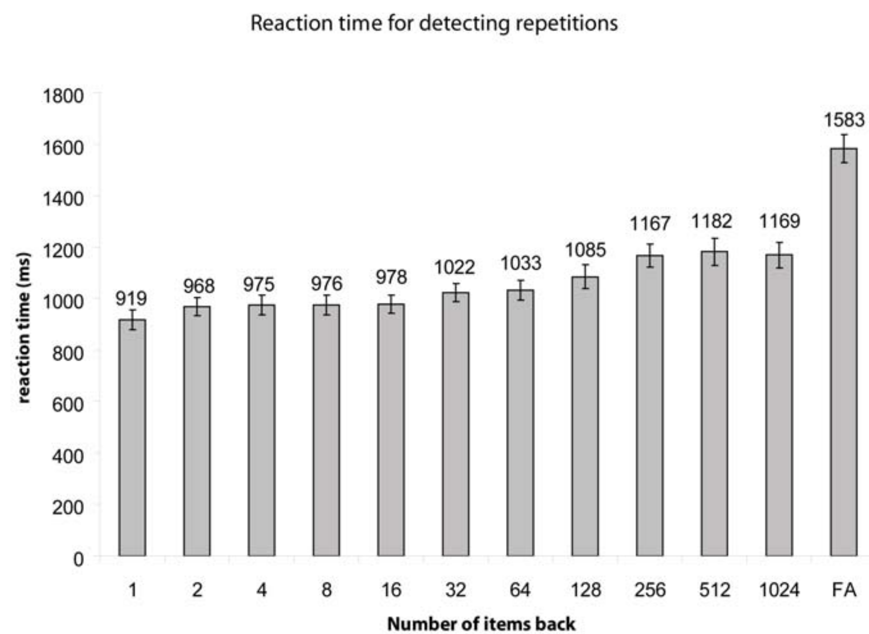
- During exposure phase, N-back repeat detection task probes recognition performances like an **old/new** task (e.g. *familiarity*)

- Have you seen that exact *same* image before in the stream ?

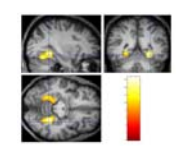
Repetition Detection Performances

N-back Performance by # of items back



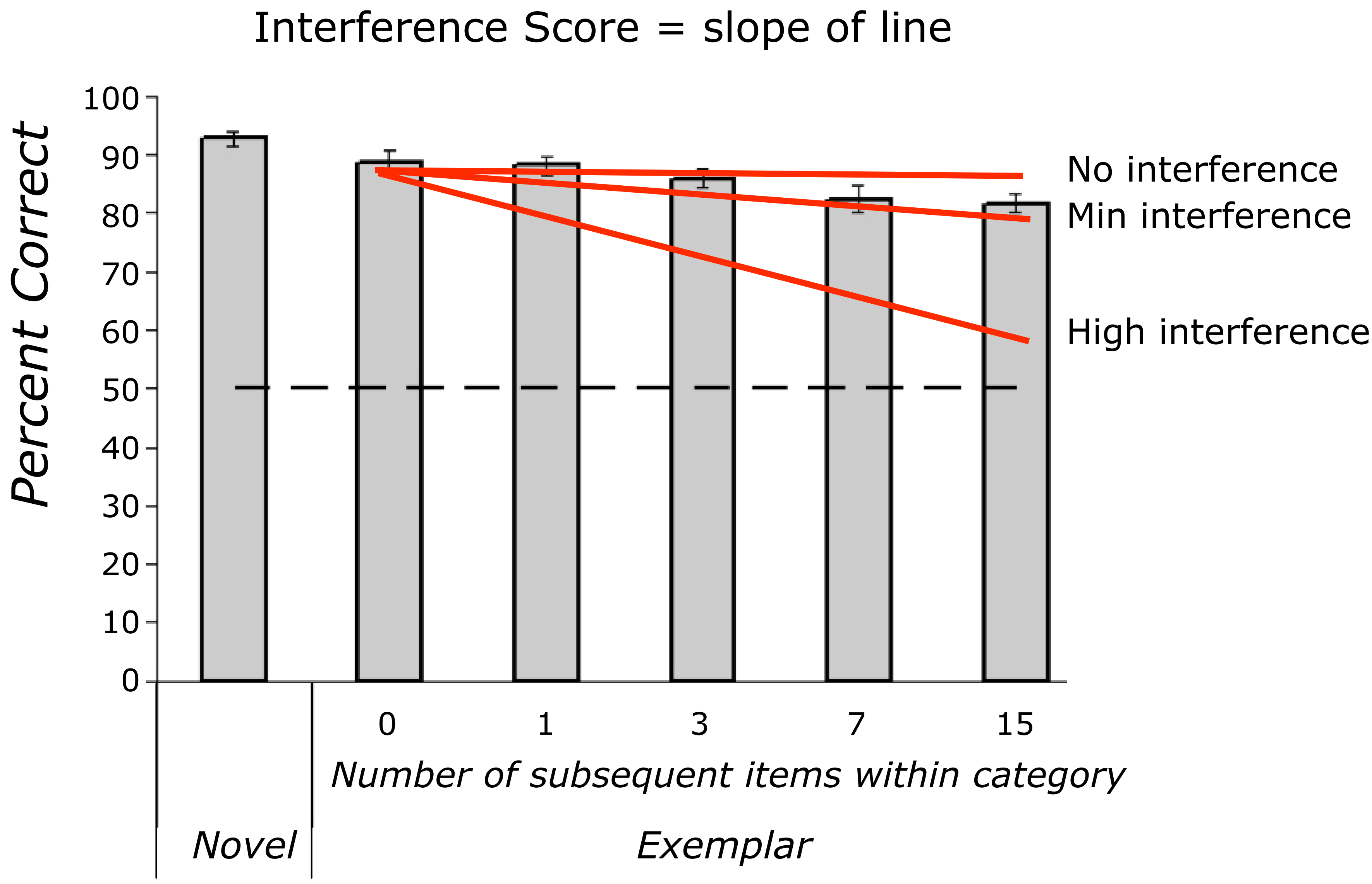


… high performance is not just about the 2-AFC memory test.

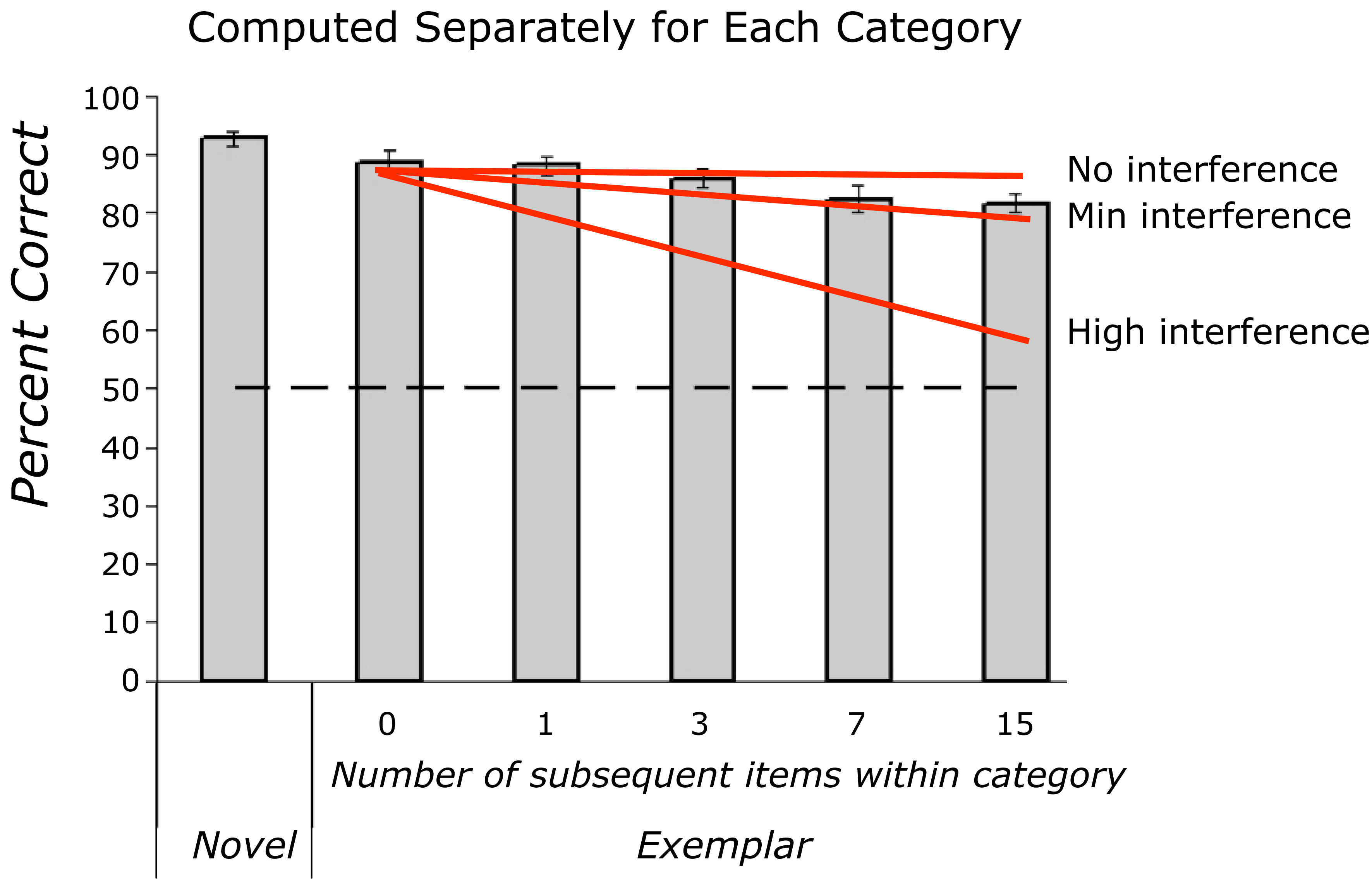


What about object similarity?

Category Interference Measure



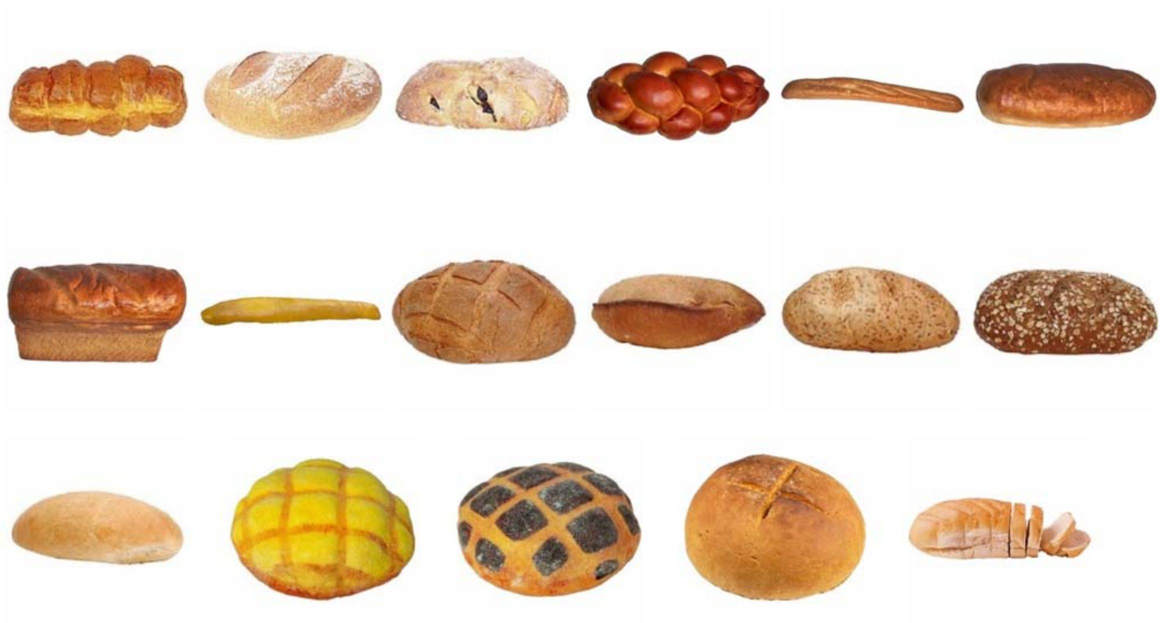
Category Interference Measure



Does **Distinctiveness** in the category make it easier to remember more items?

I – Measure of **Conceptual** Distinctiveness

Are There Few or Many **Kinds**?



Each category was judged by 12 observers on a 1-5 scale

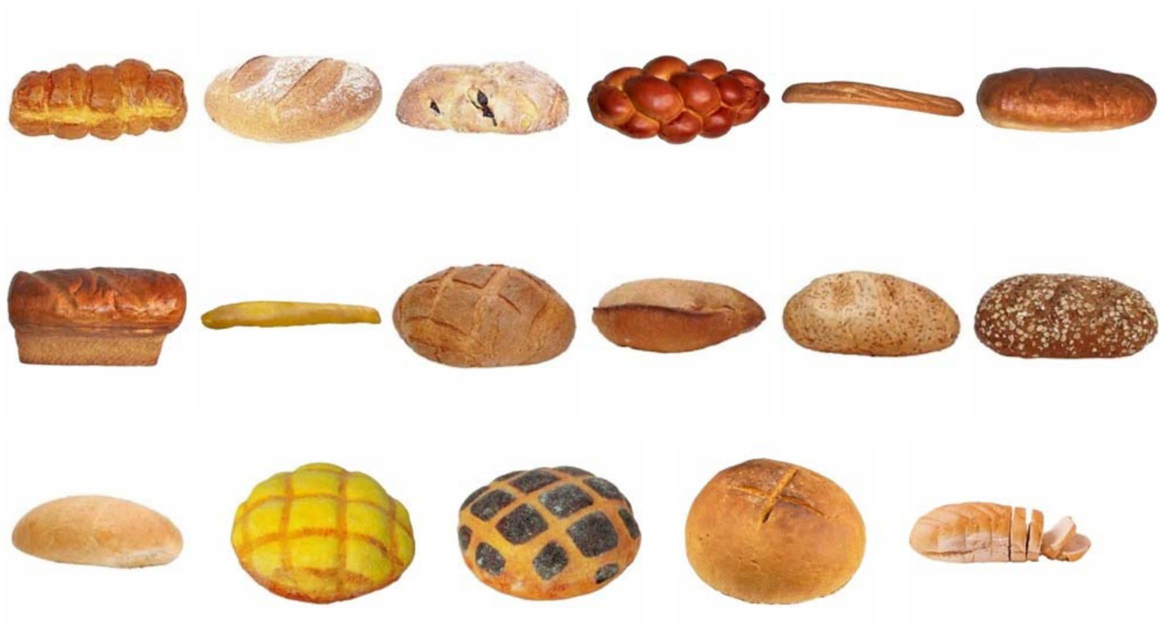
I – Measure of **Conceptual** Distinctiveness

Similar

Distinctive

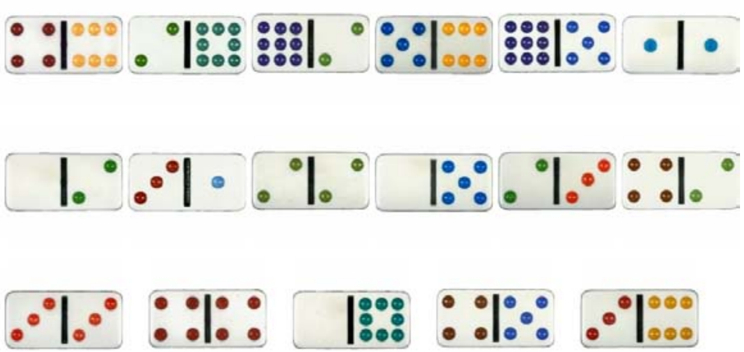
II – Measure of **Shape** Distinctiveness

How Similar or Different are their **shapes**?



II – Measure of **Shape** Distinctiveness

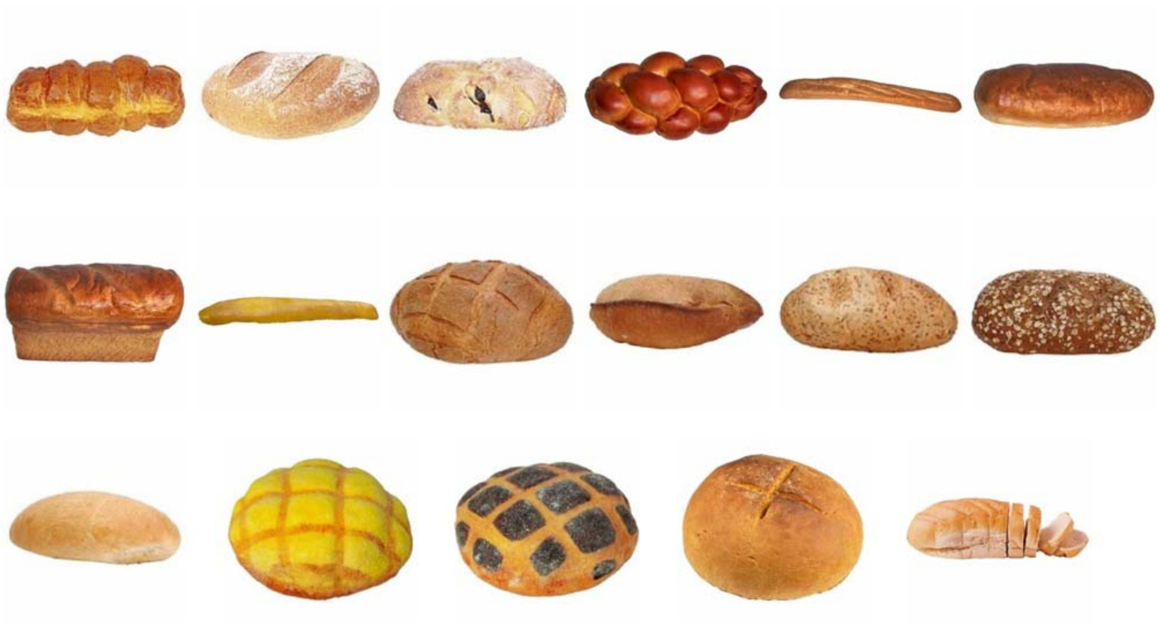
Similar



Distinctive

III – Measure of **Color** Distinctiveness

How Similar or Different are their **colors**?

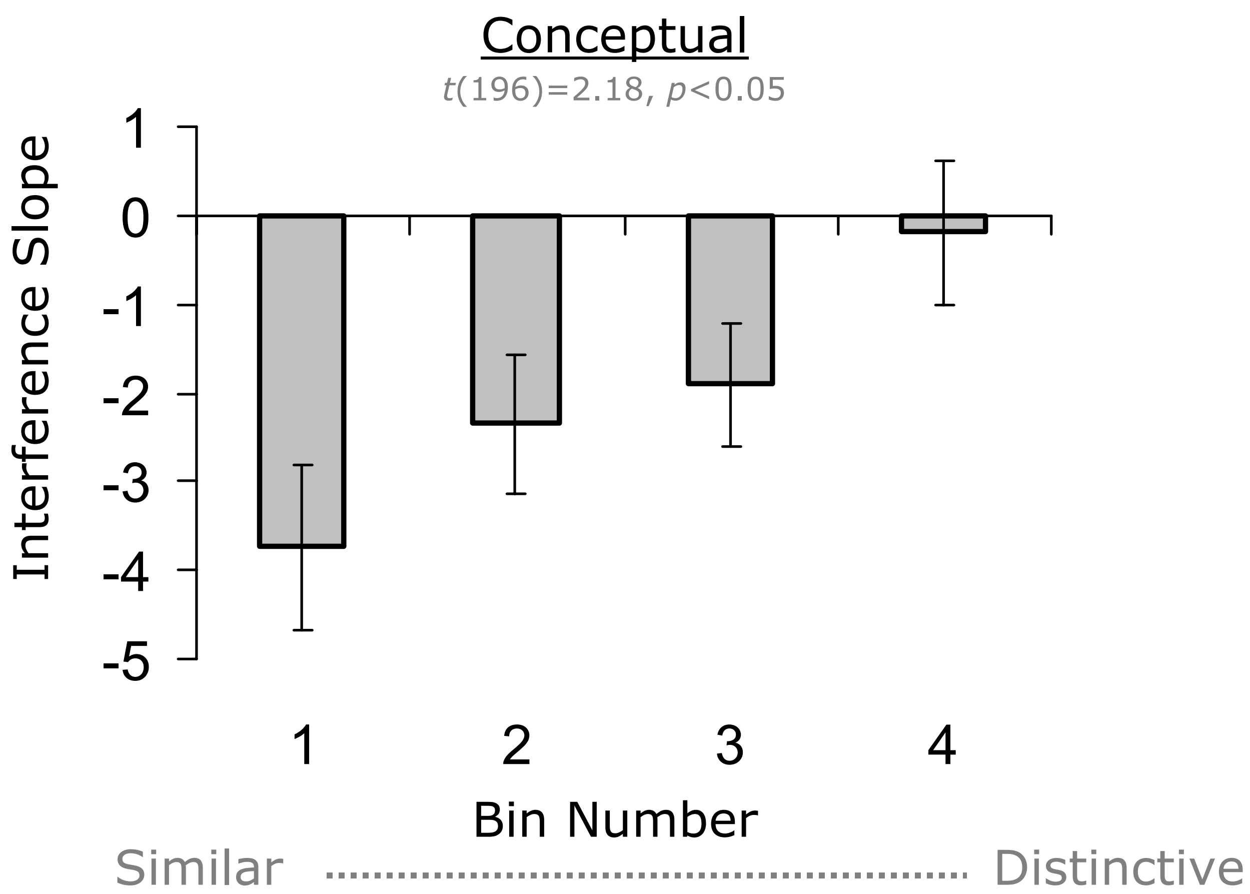


III – Measure of **Color** Distinctiveness

Similar

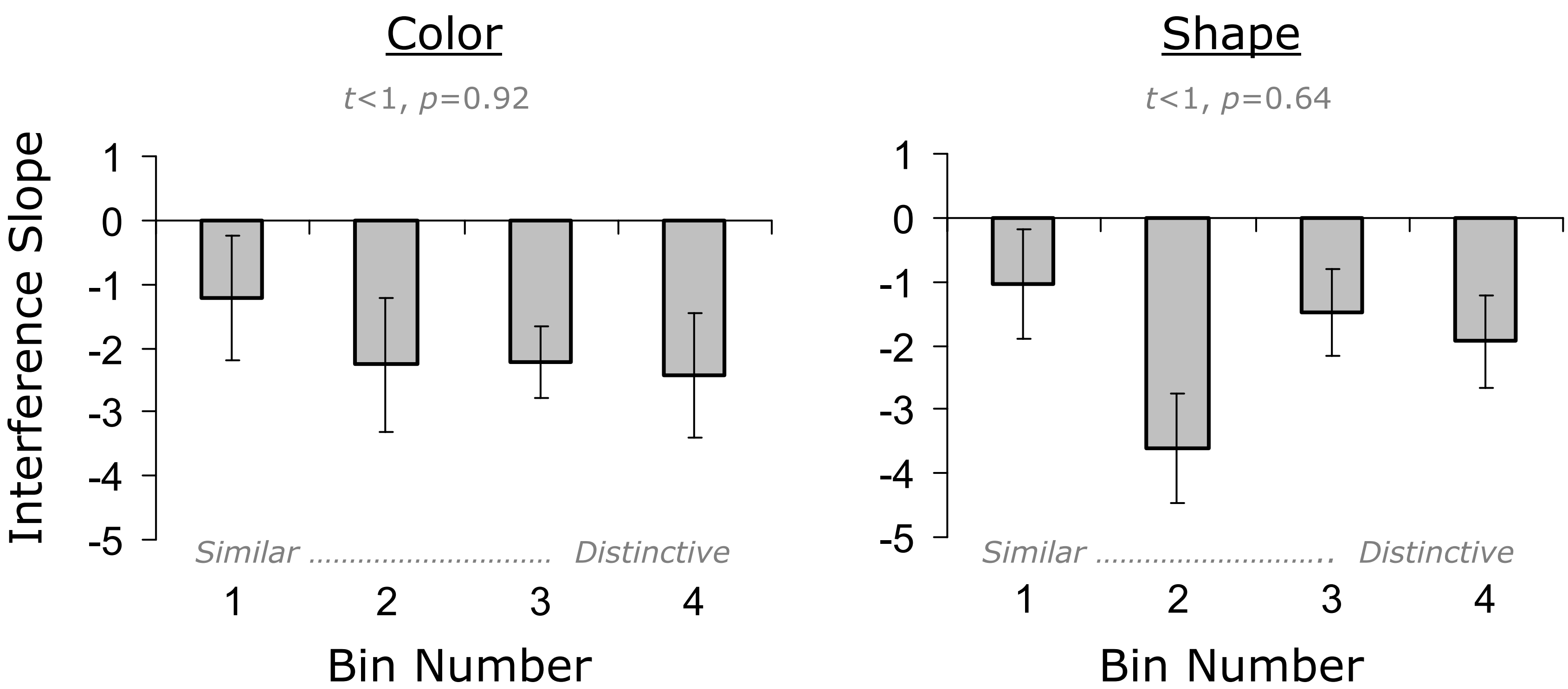
Distinctive

Distinctiveness vs. Interference

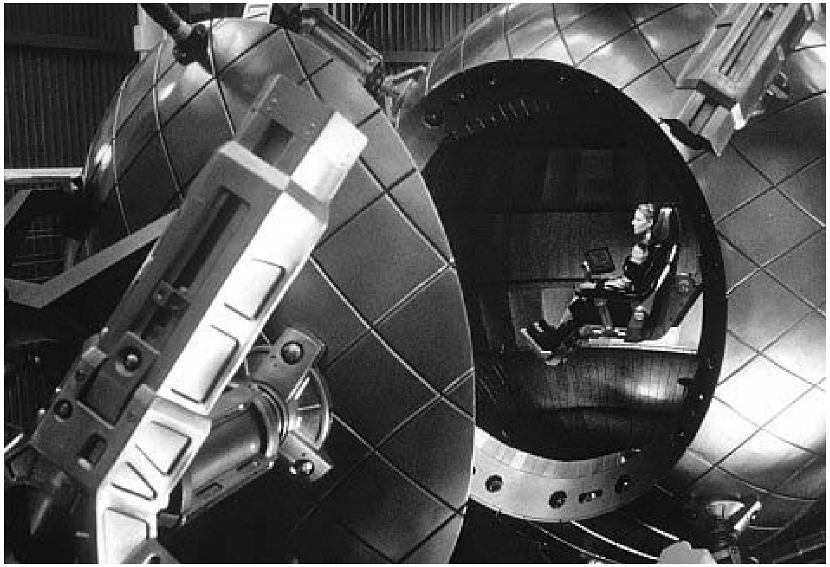


Conceptual Distinctiveness helps you remember

Distinctiveness vs. Interference



No Effect of Perceptual Distinctiveness



Contact (1997)

Why have one Massive Memory Experiment,

when you can have **two**?

how far can we push the fidelity of visual LTM representation ?

Same object, different states

Massive Memory II:
2500 unique object **categories**

Followed by 300 2-alternative forced choice tests

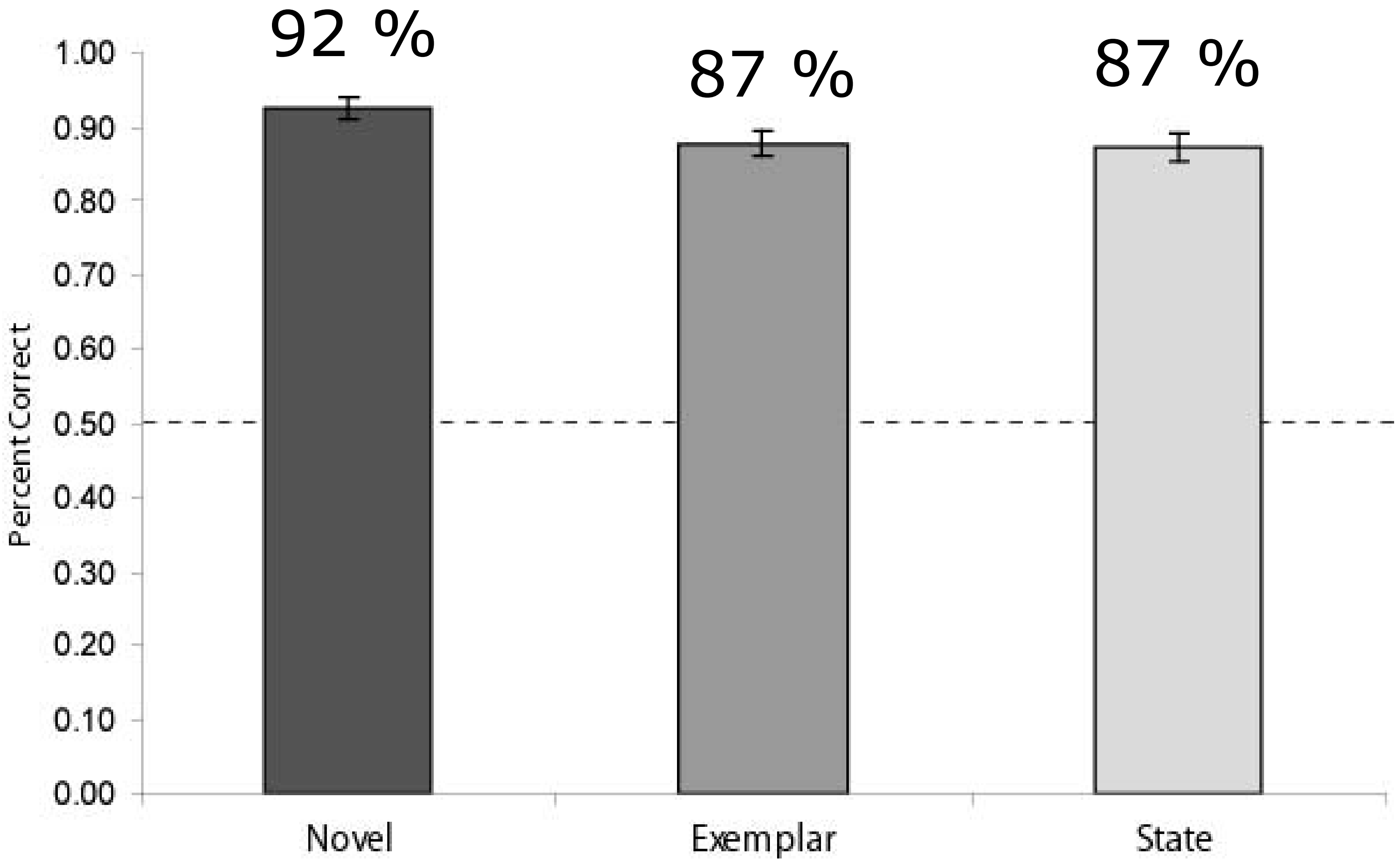
100 novel pairs
100 exemplar pairs
100 state pairs

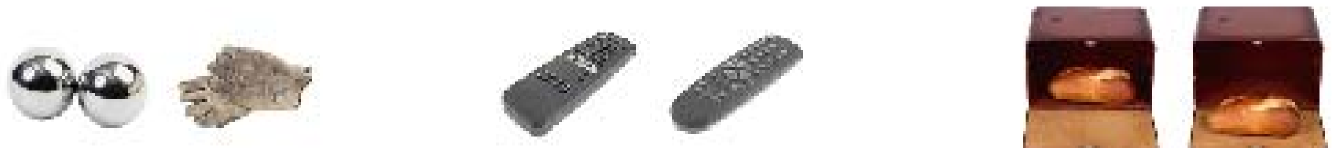
Examples of Exemplars in memory test

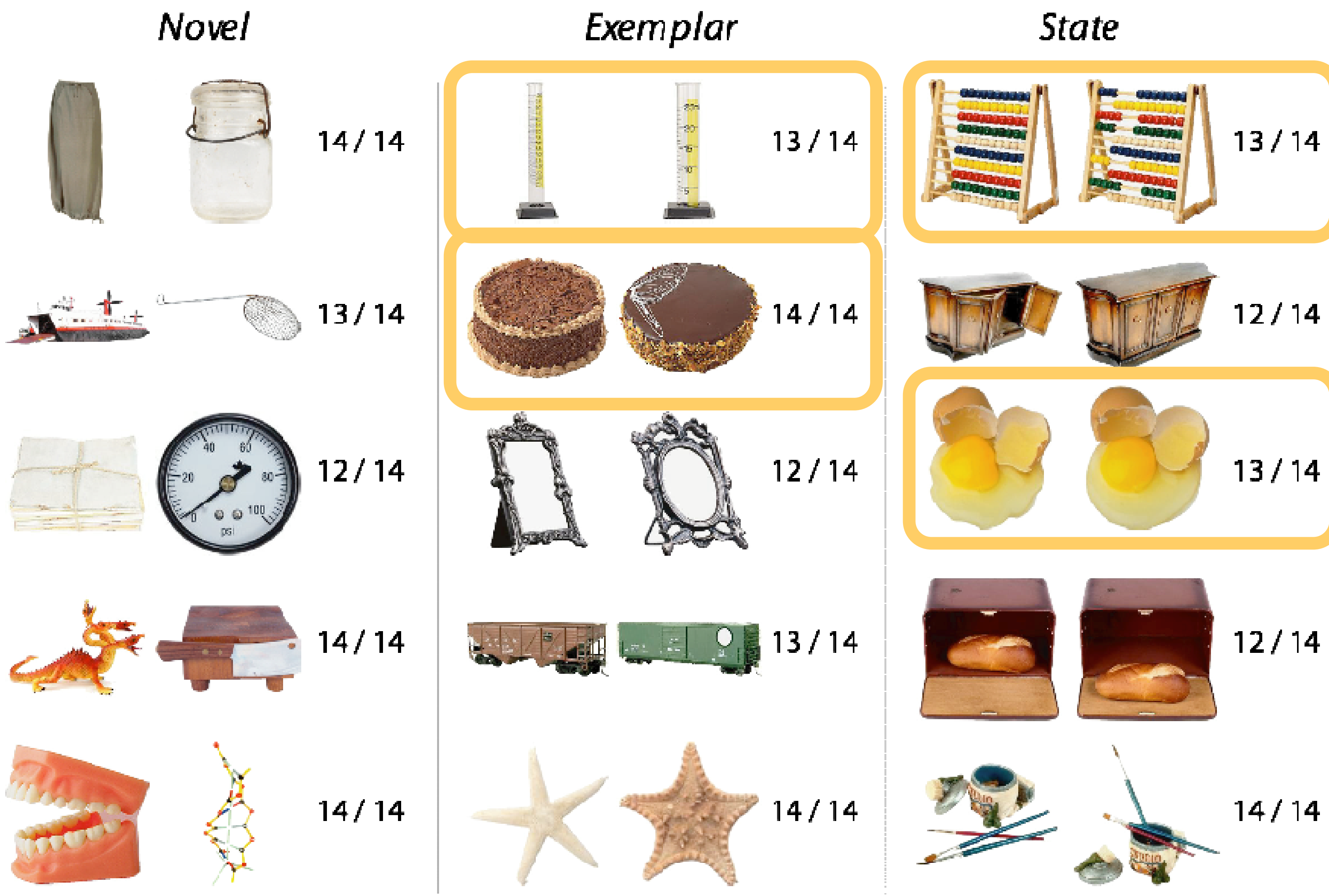
Examples of *states* in the memory test

Results Memory Test

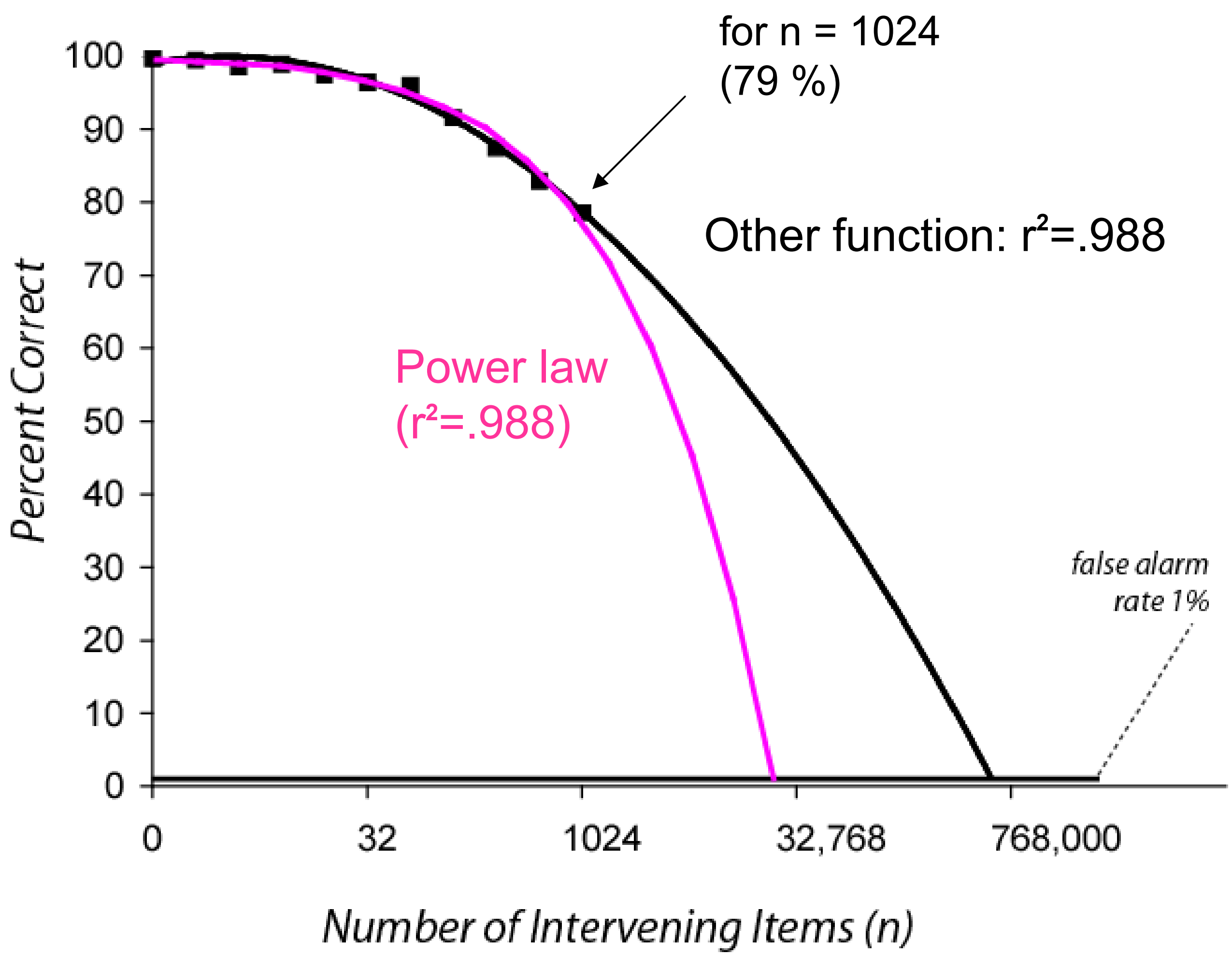
No interference !



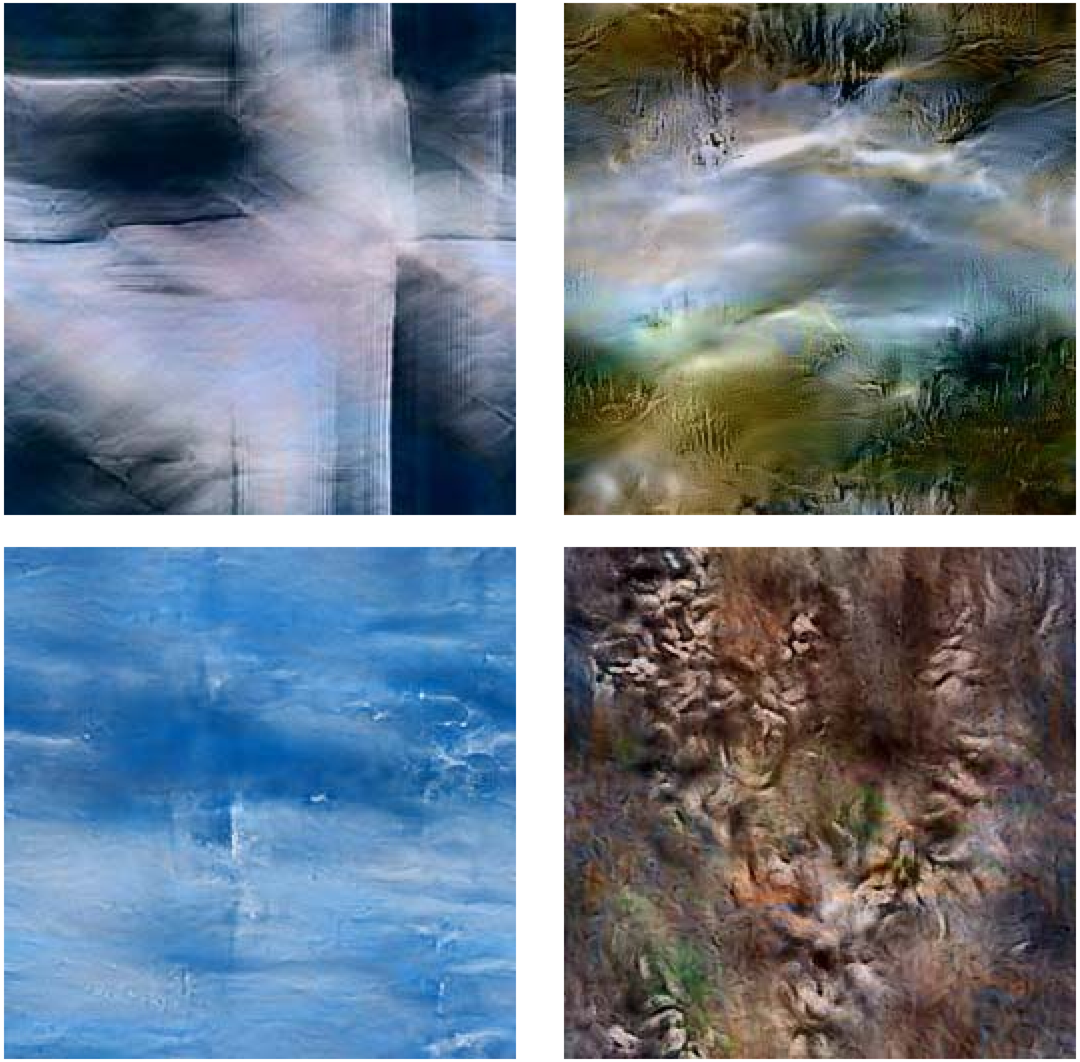



Novel
Exemplar
State
14 / 14
13 / 14
13 / 14
13 / 14
14 / 14
12 / 14
12 / 14
12 / 14
13 / 14
14 / 14
13 / 14
12 / 14
14 / 14
14 / 14
14 / 14

How many different images can you see before loosing familiarity ?



What about distinct textures?



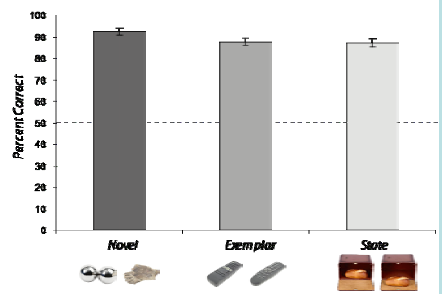
d′ ~ 0

Concluding Remarks – Part I

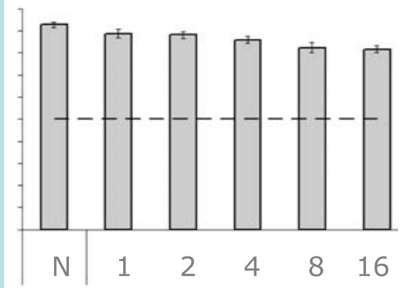
Capacity of Human Representation

- Can be massive and detailed
- details are not by necessity discarded through visual transformations

Massive Memory II



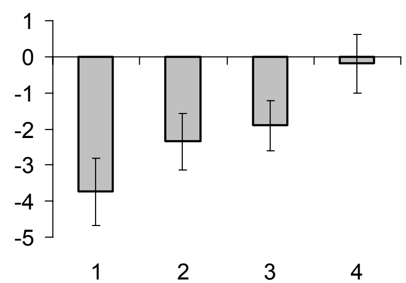
Massive Memory I



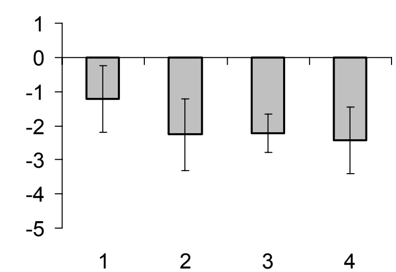
Structure of Human Memory

-Memory for “visual” details is linked more to conceptual knowledge rather than perceptual similarity

Concepts & Memory



Percepts & Memory



How detailed are visual scene representations?

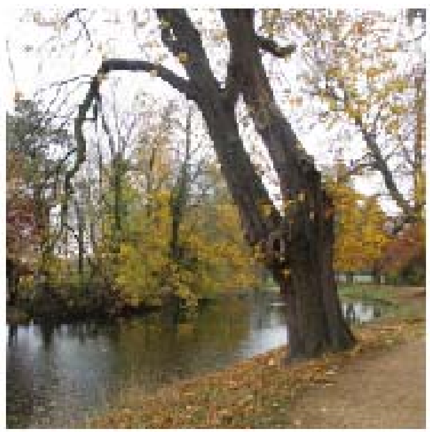
At a glance ... You remember the category and the layout but you have lost some object details

You have seen these pictures

You were tested with these pictures [average false alarms ~ 30%]





1

2

3

4

5

6

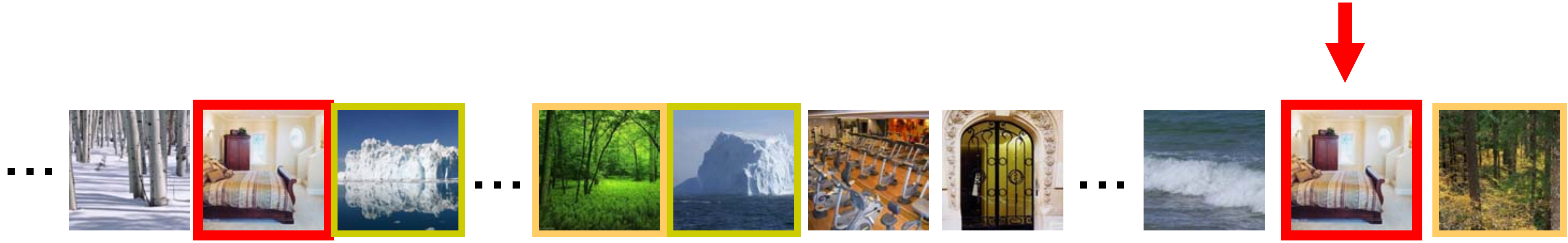
7

8

What about memory for thousands of scenes?

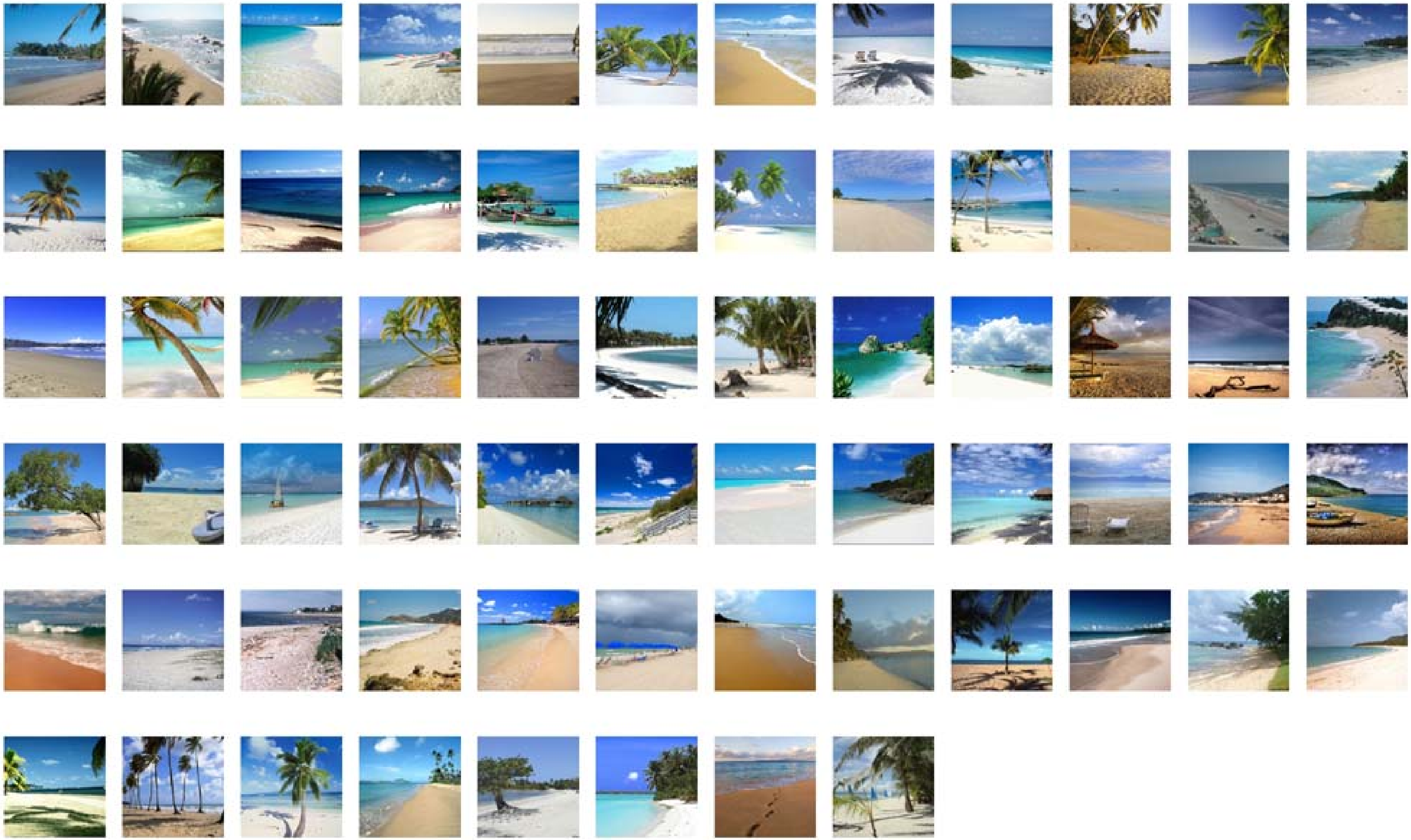
128 unique semantic categories of natural images

Presentation: 3 seconds each

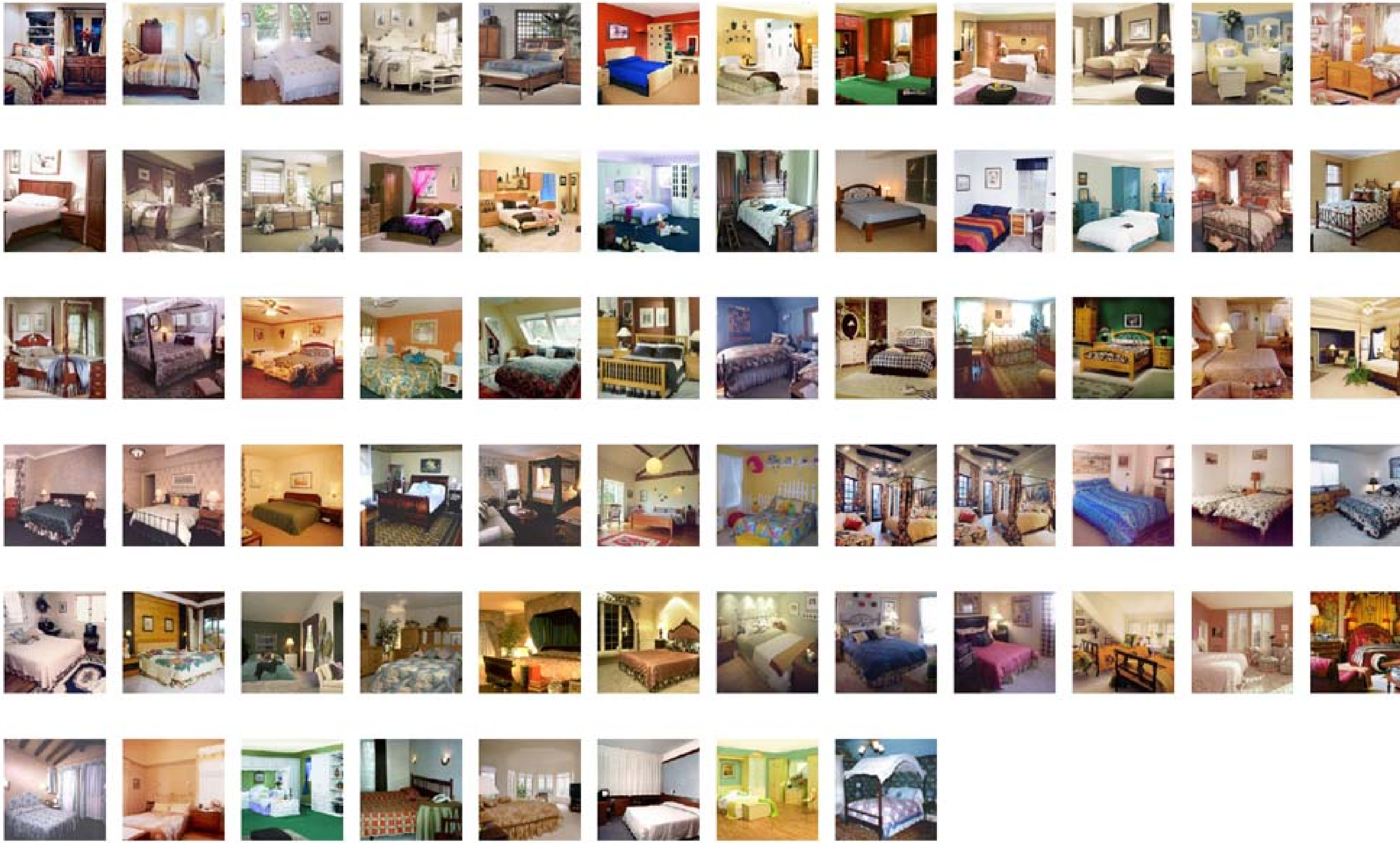


Barn

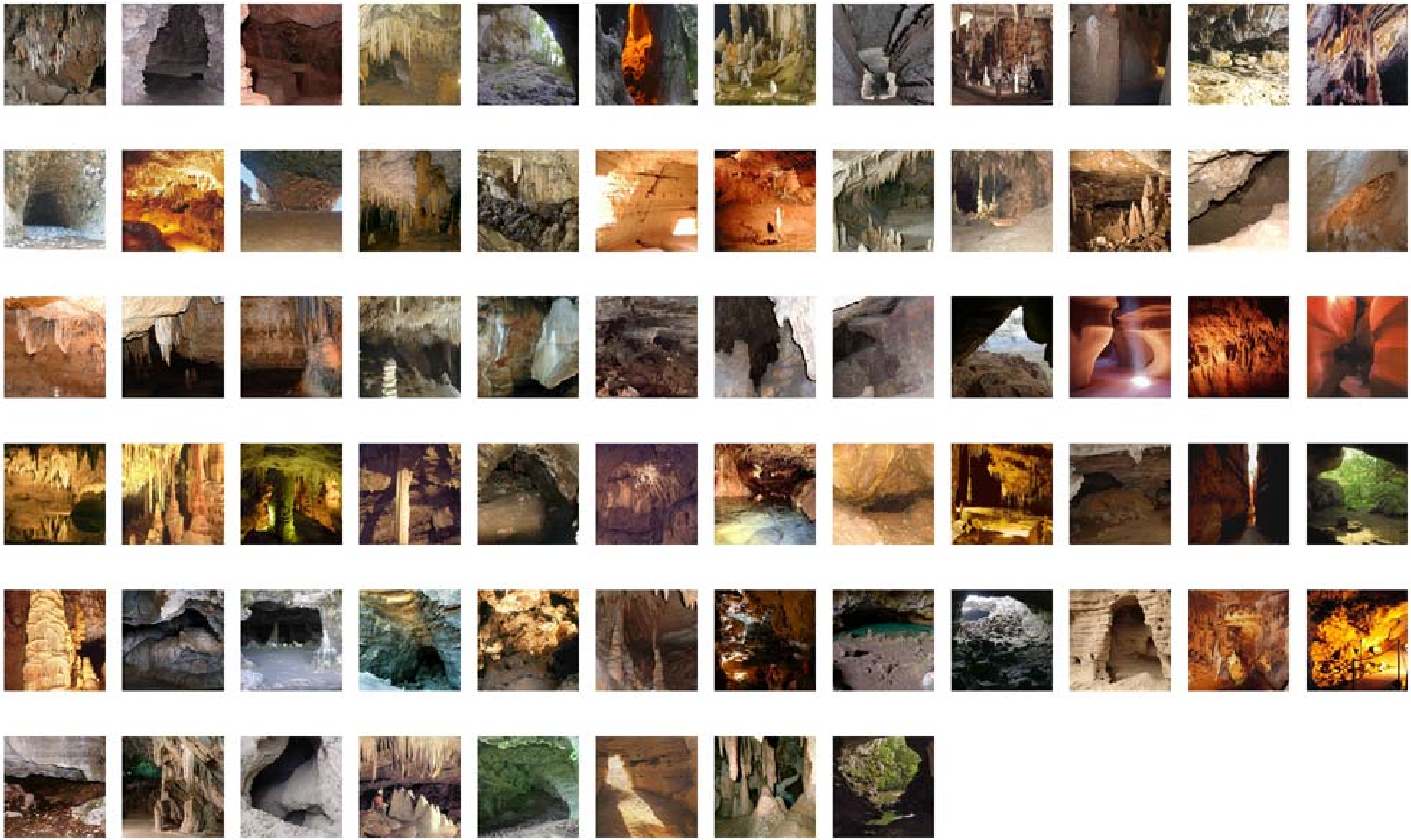
Beach



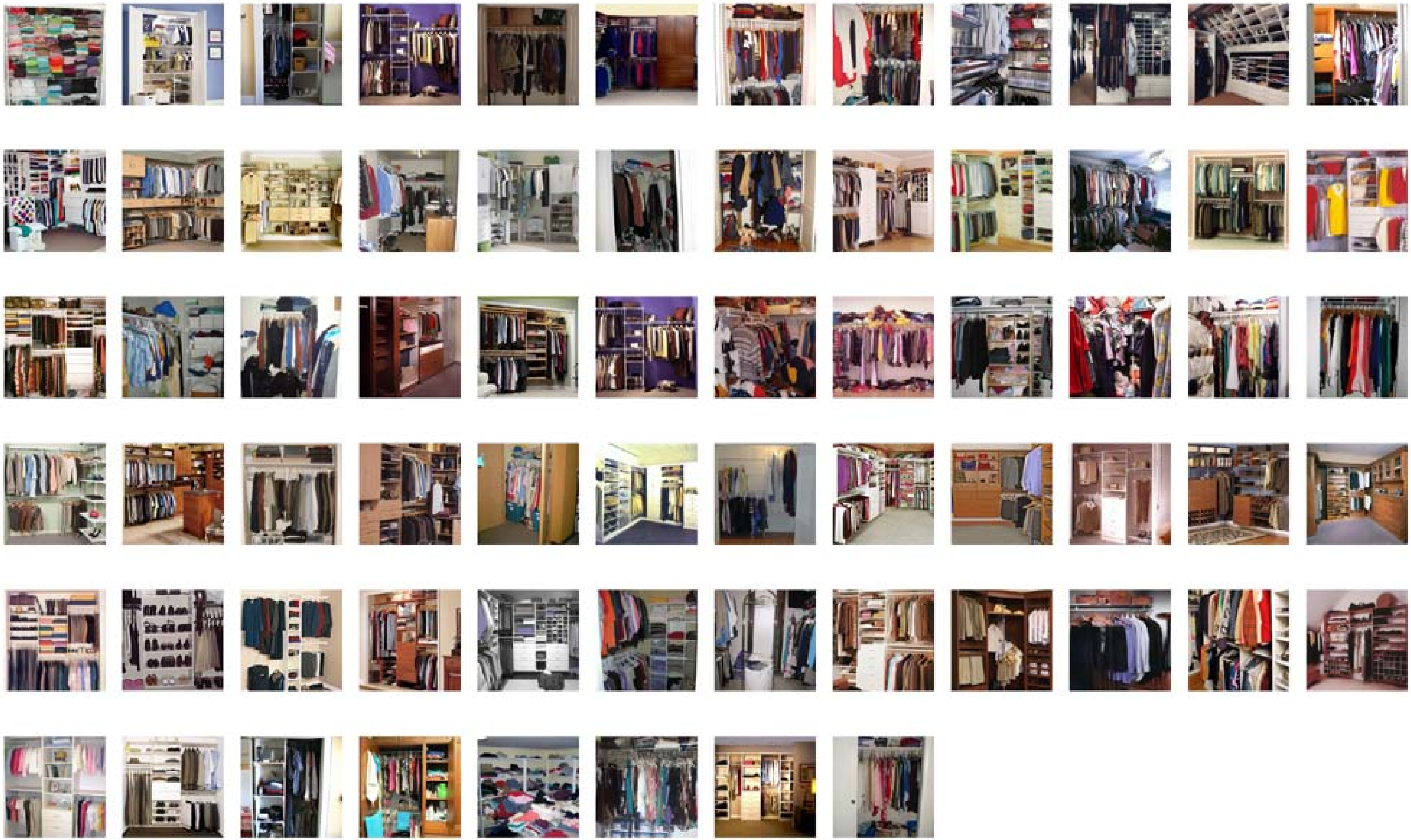
Bedroom



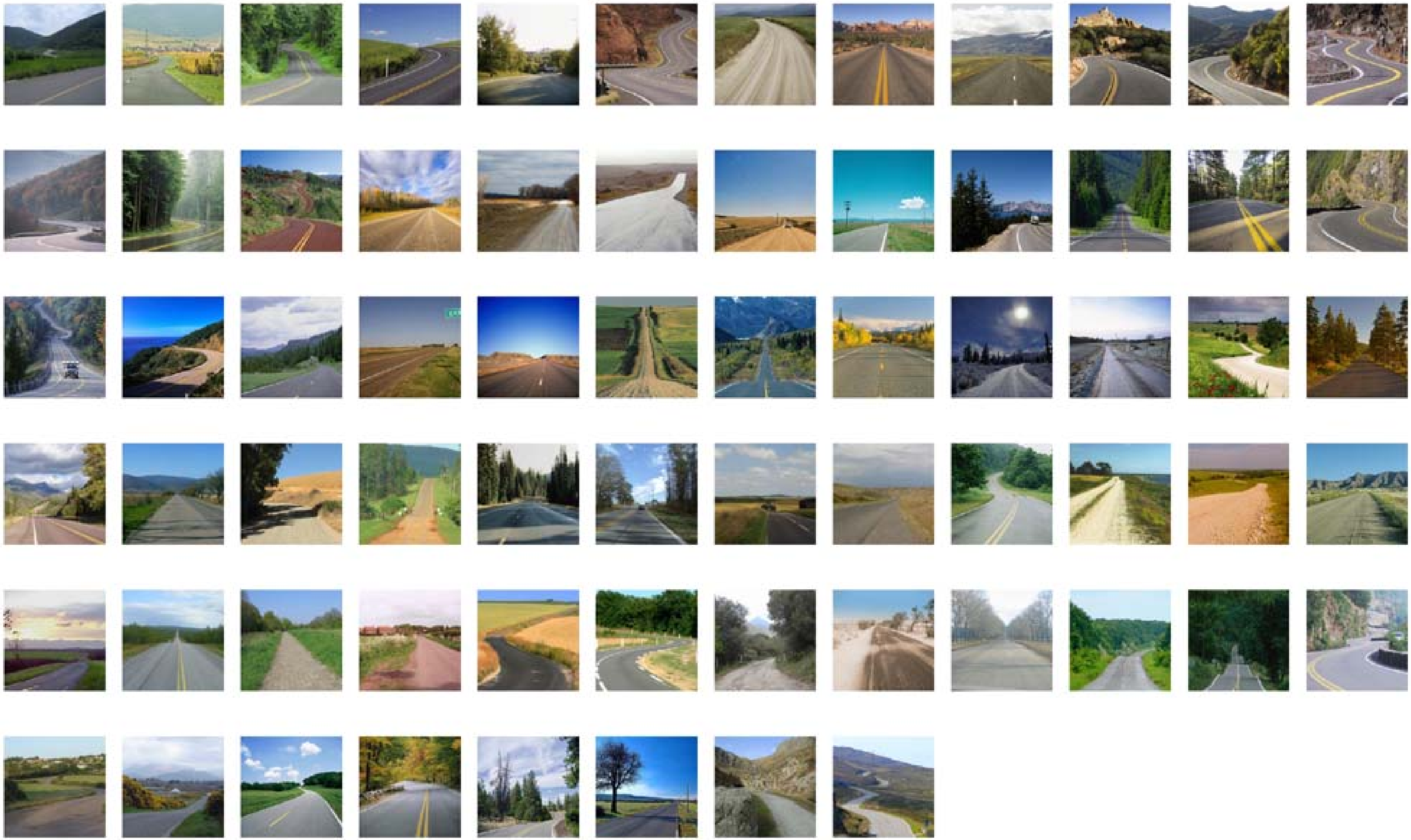
Cavern



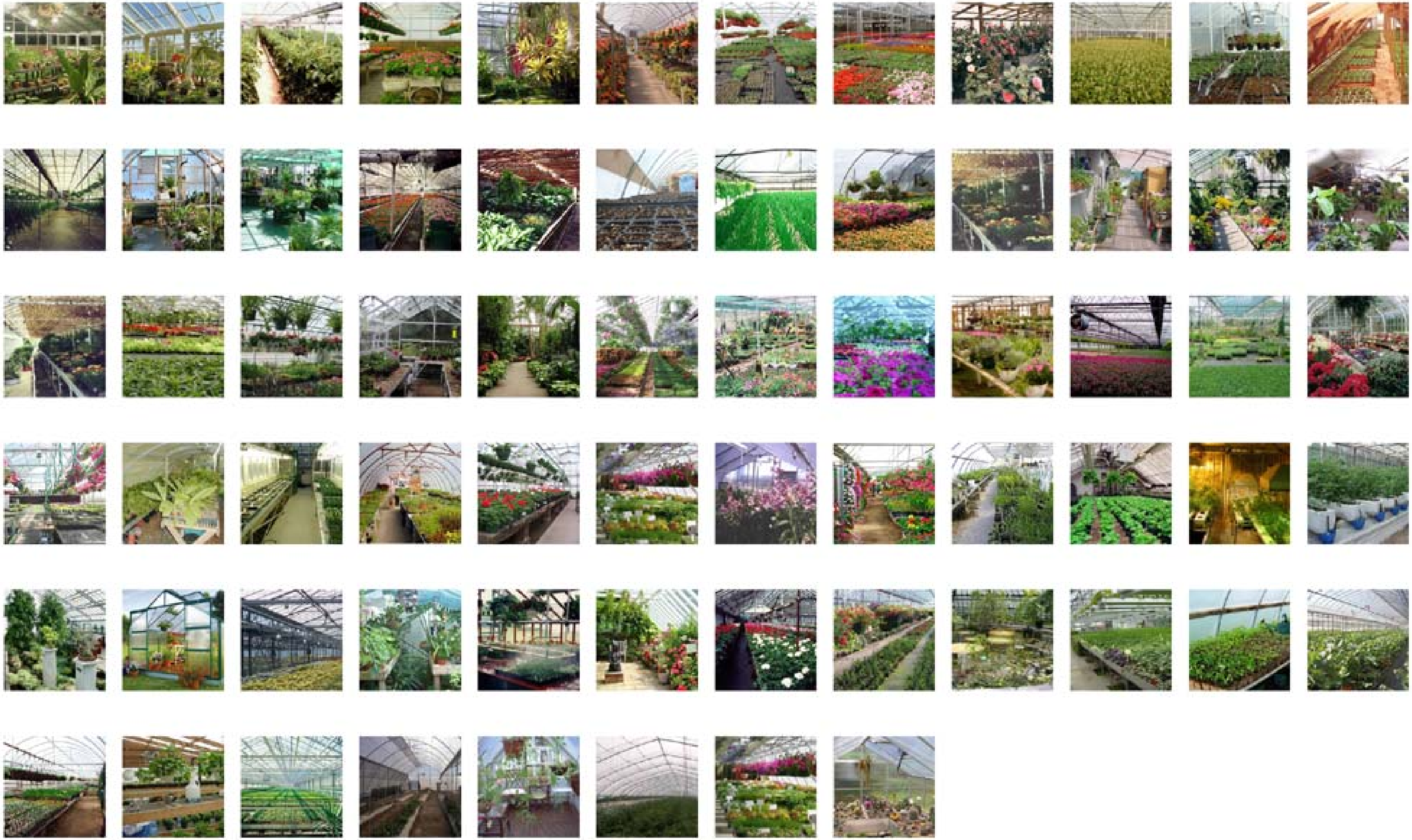
Closet



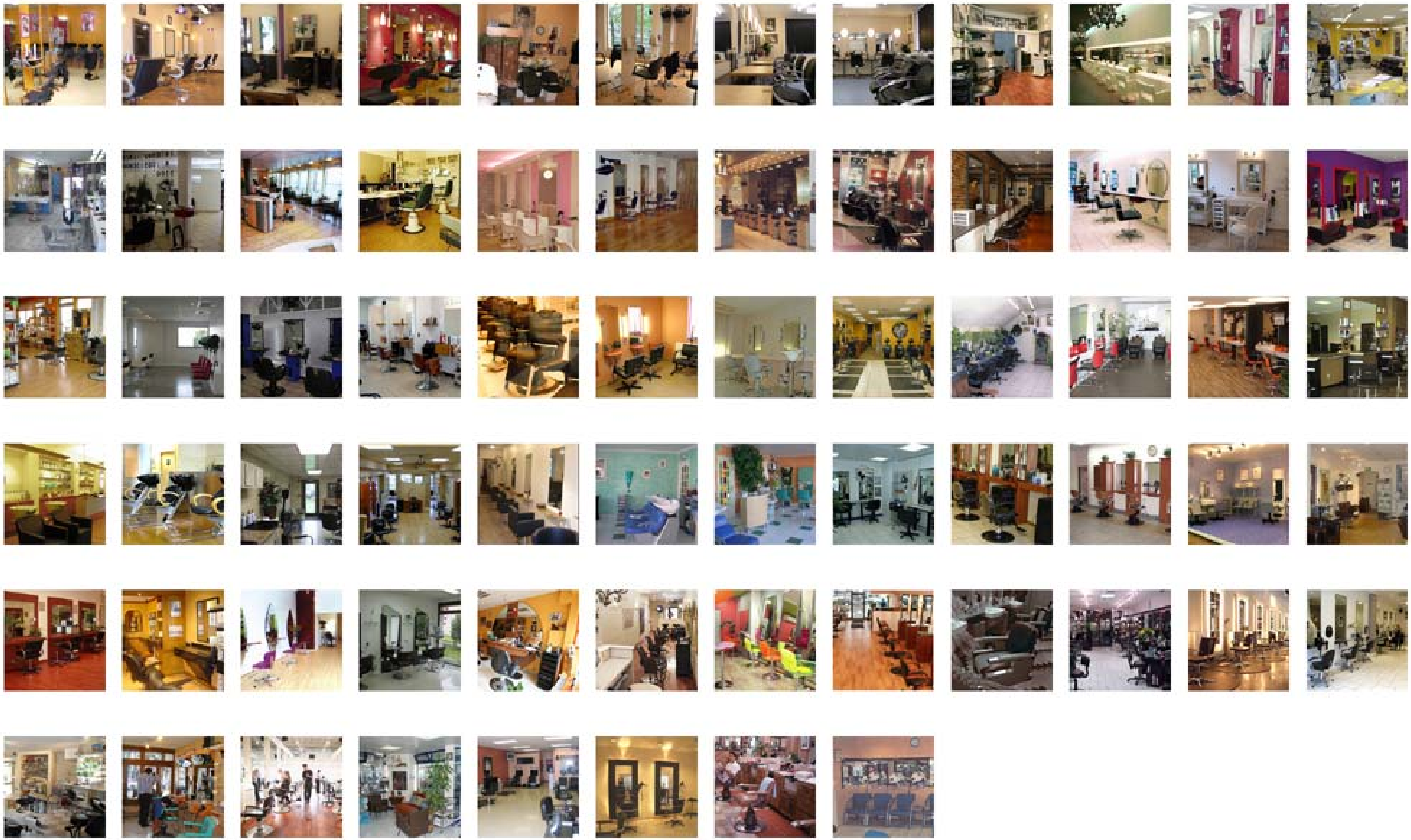
Countryroad



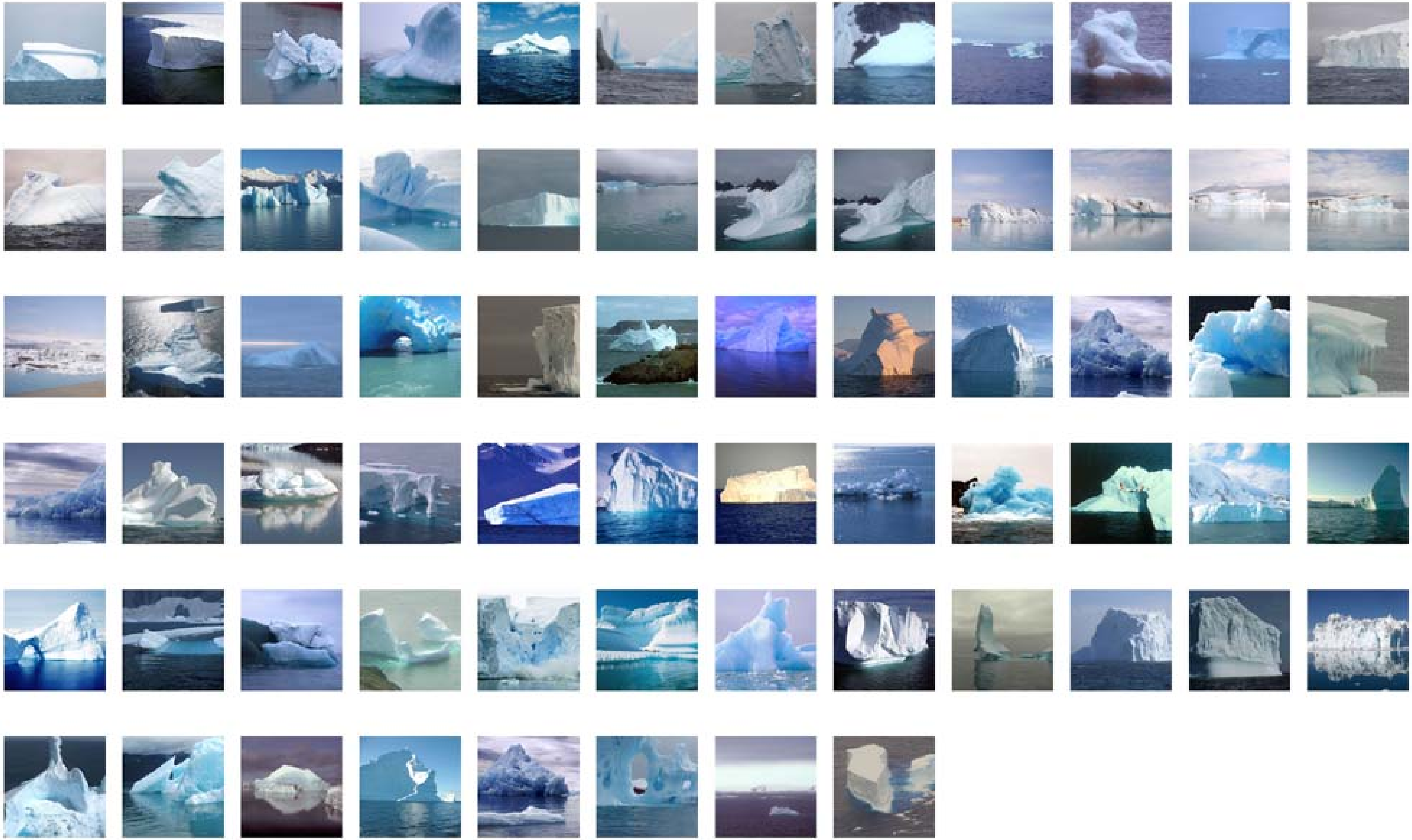
Greenhouse



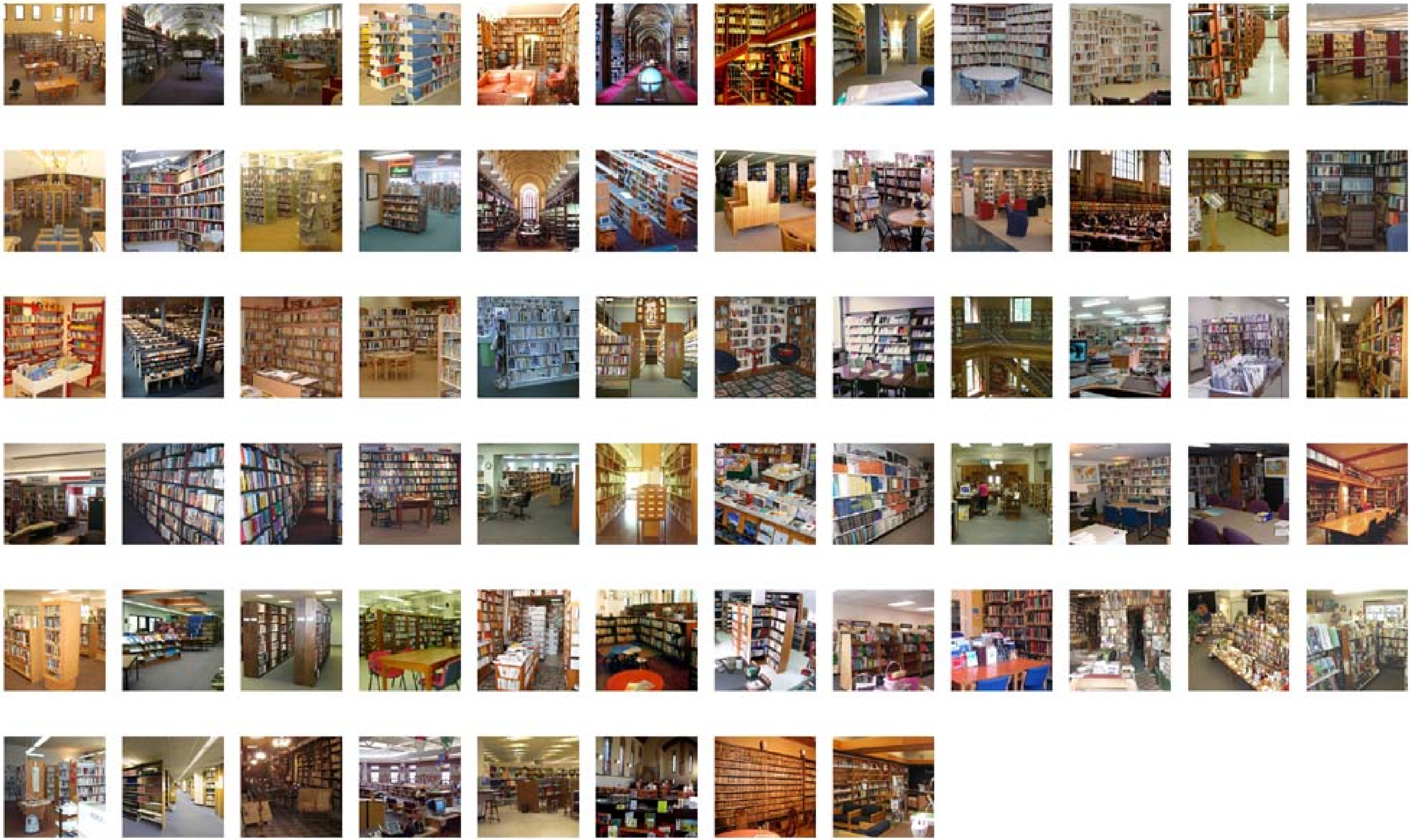
Hair salon



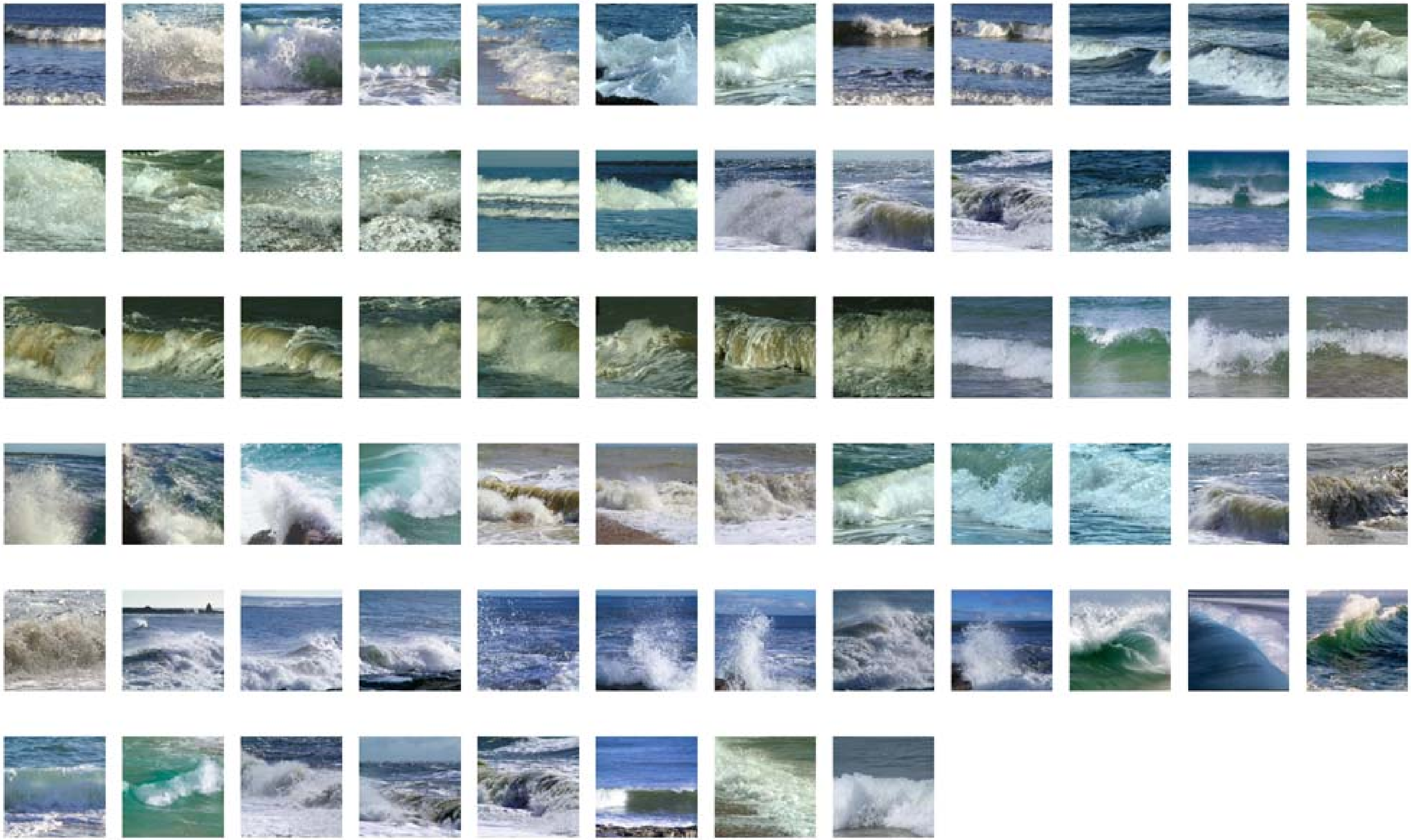
Iceberg



Library



Waves

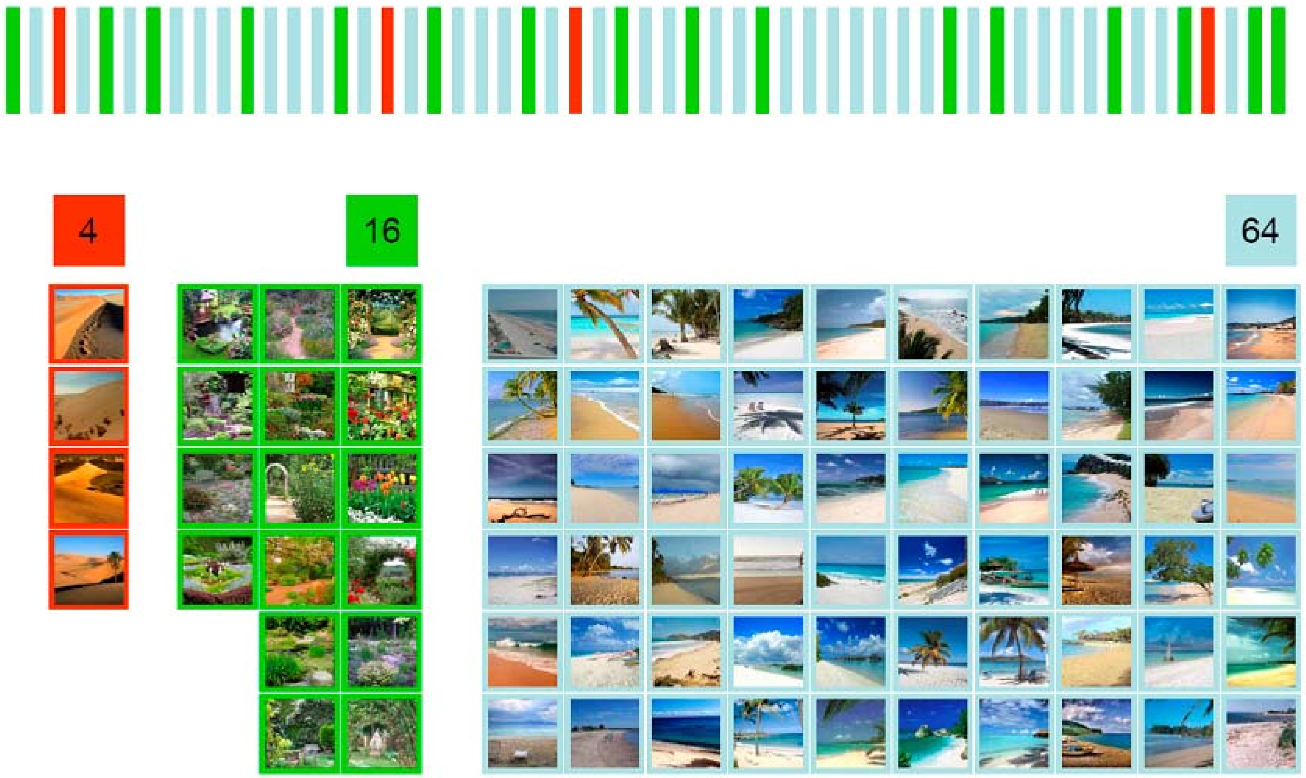


Methods – The Study Stream

128 unique semantic categories of natural images

2912 natural images shown in the stream (3 seconds each, 800 msec ISI)

Number of exemplars per category: 4, 16, or **64 !**

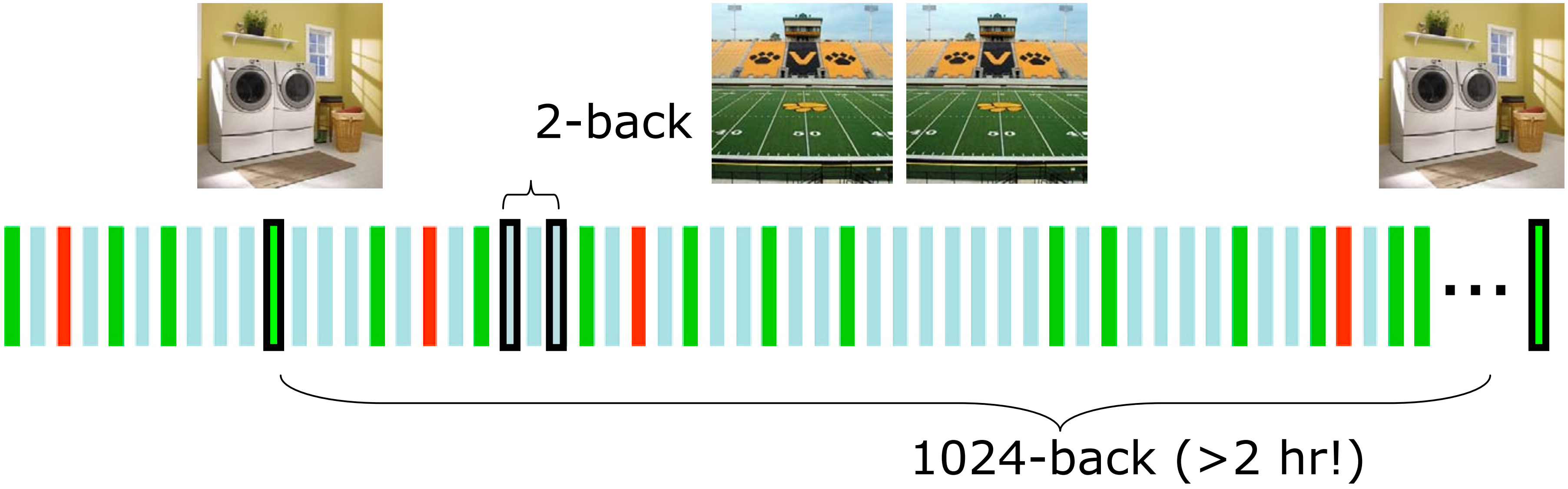


N= 24 observers

Methods – The Study Stream

Online Task: **Detect Exact Repeats**

Repeats could be 2 to 1024 back in the stream



Methods – The Memory Test

Followed by 224 **2-alternative forced choice tests**

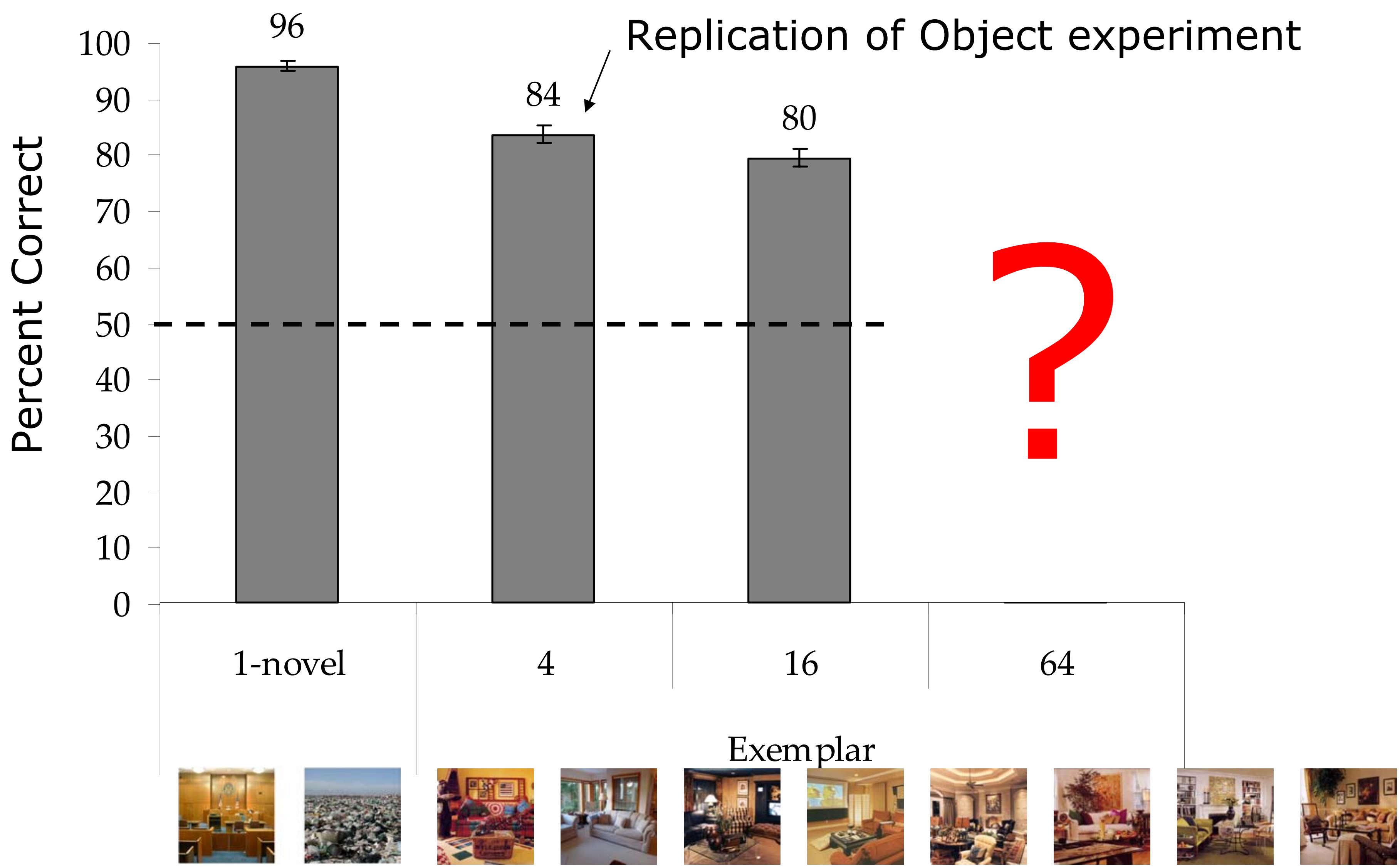
Novel

Exemplar

Test Pairs were always the same for all subjects (4 test pairs for each scene category)

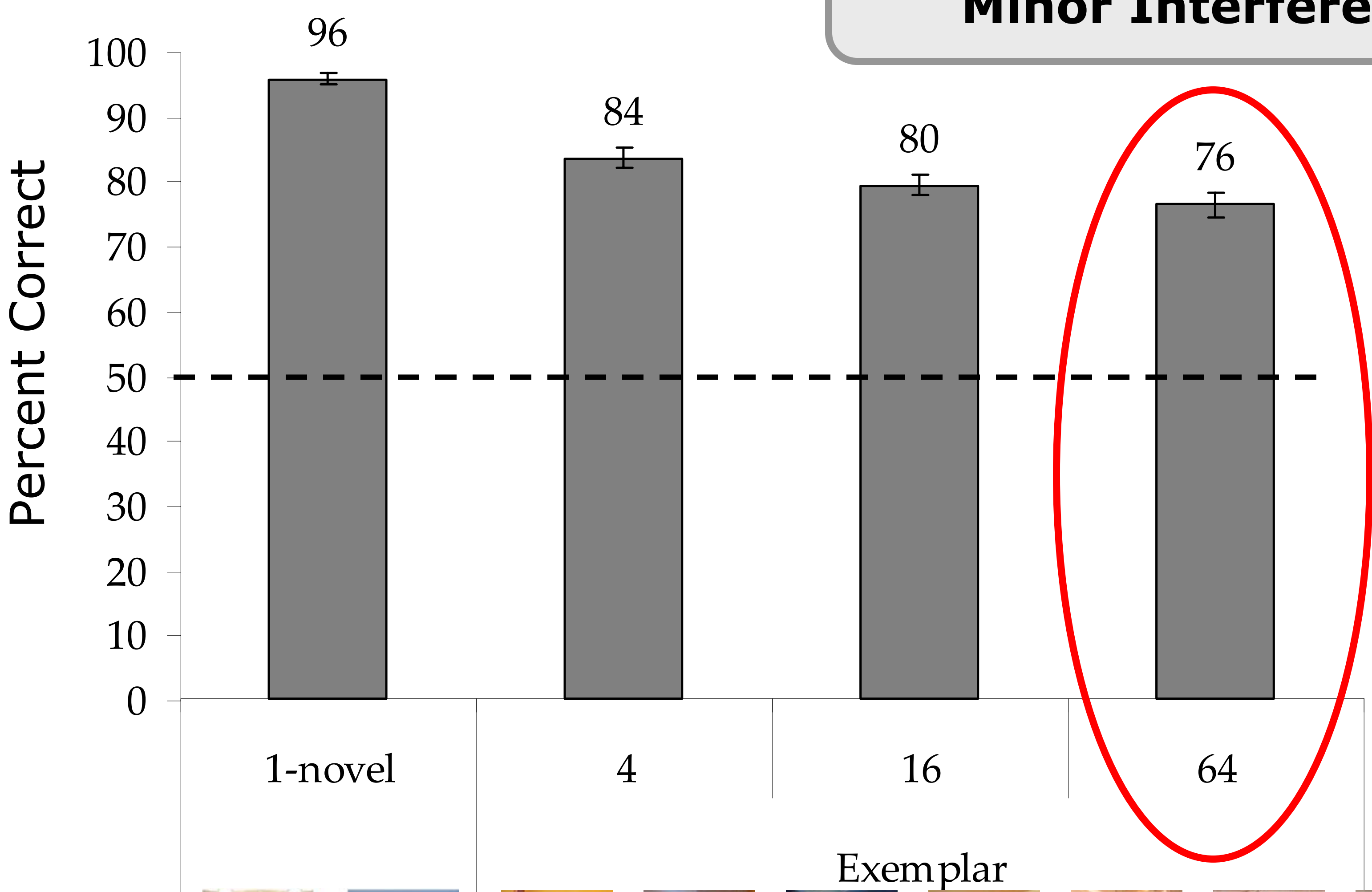
Any effect of interference is due to the additional exemplars seen in the stream

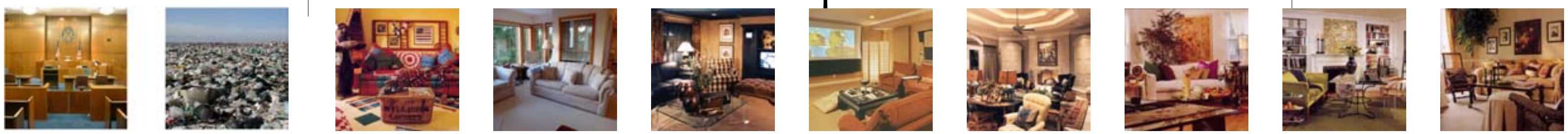
Results – Memory Test Performance



Results

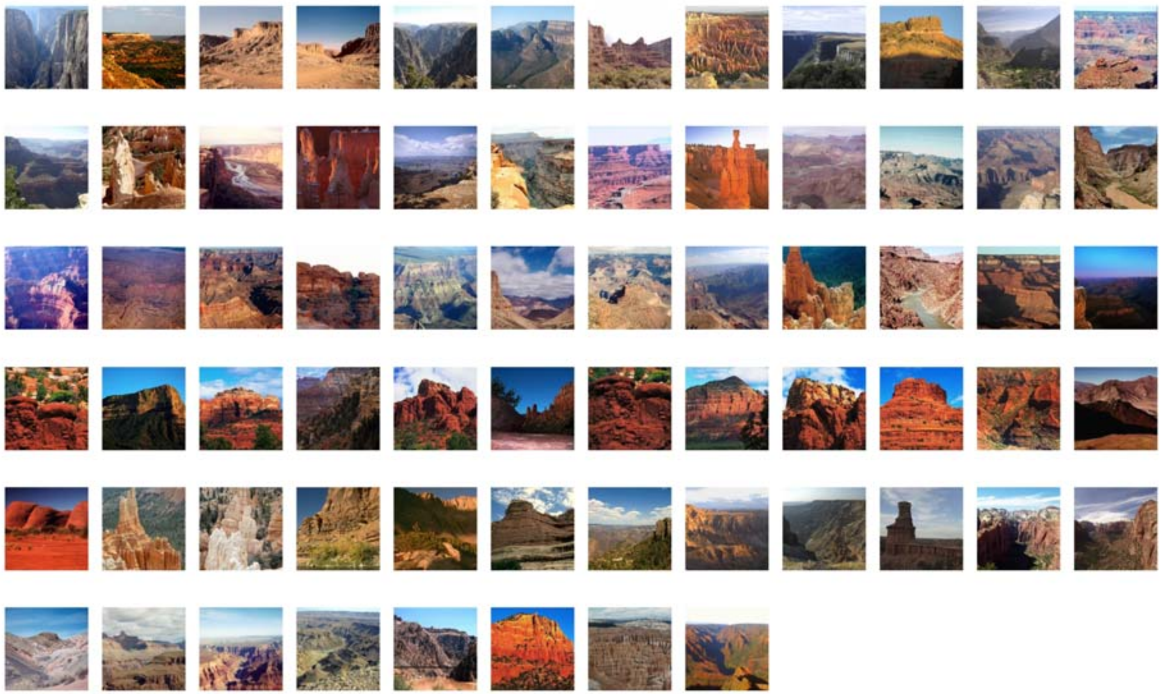
Highly Detailed
Minor Interference

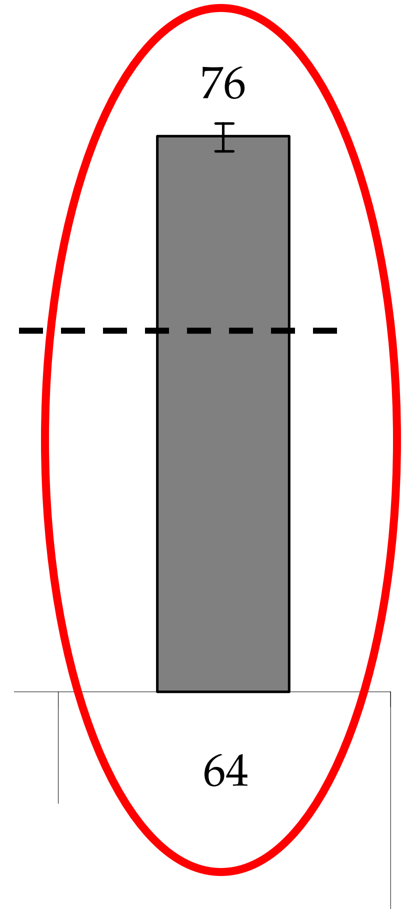


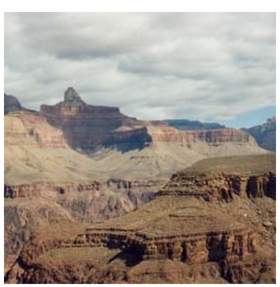


Highly Detailed

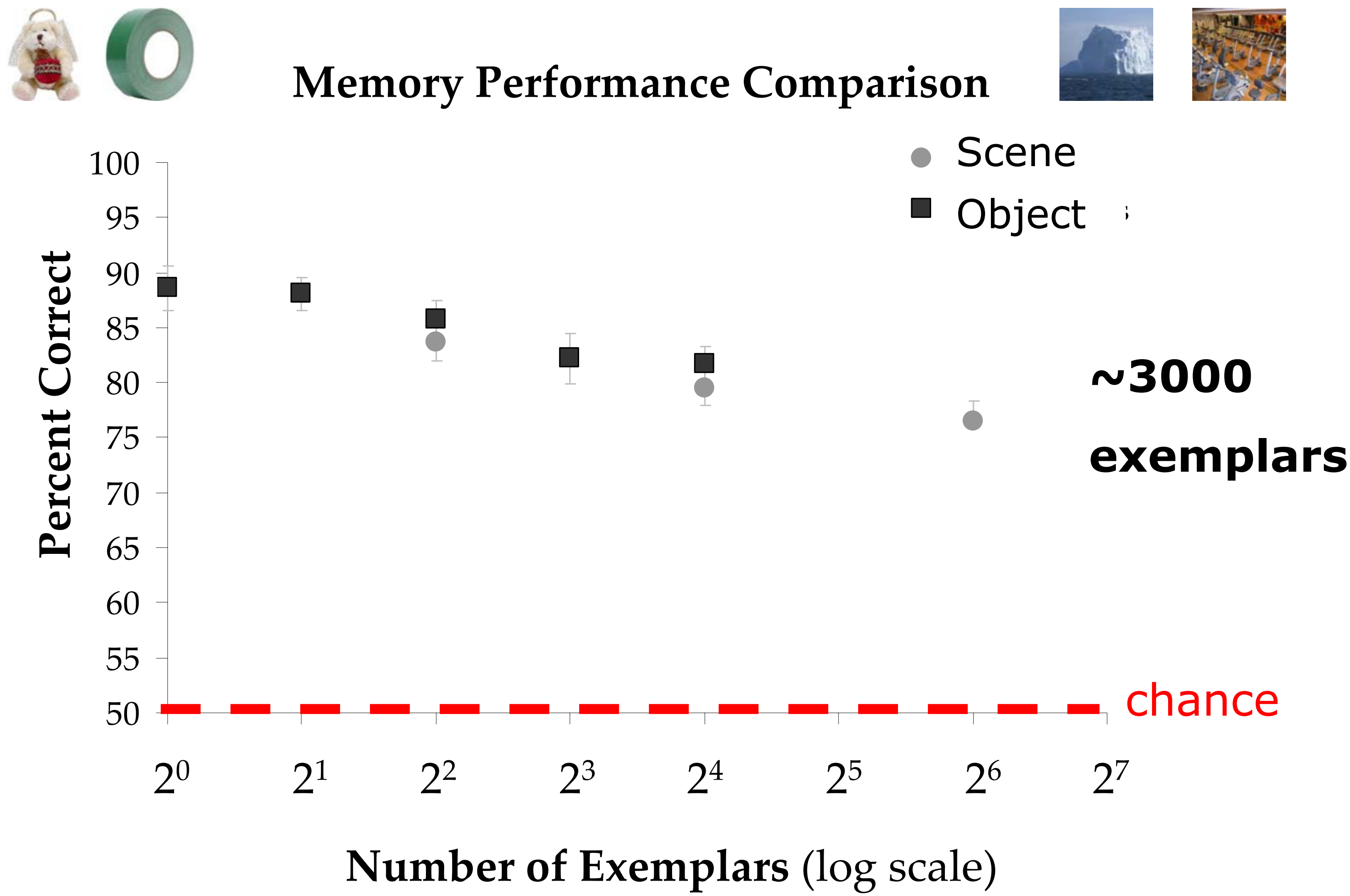
Minor Interference





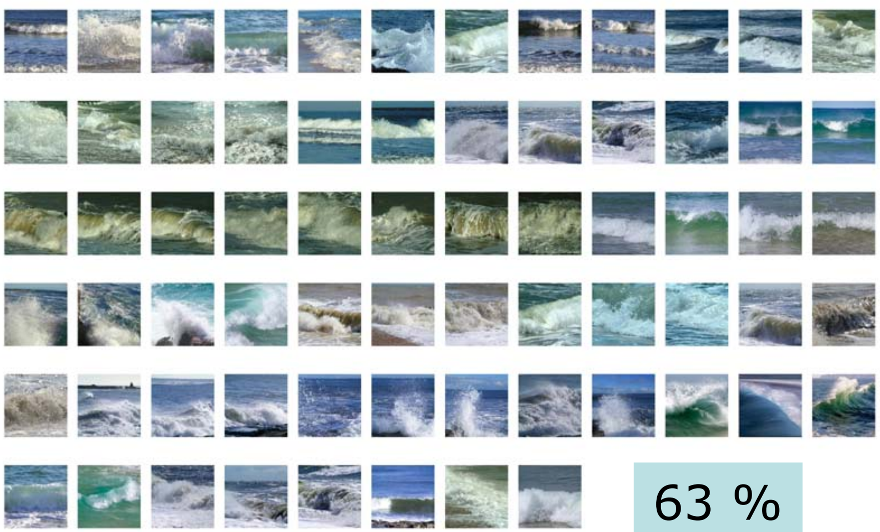


Massive Memory for Scenes and Objects



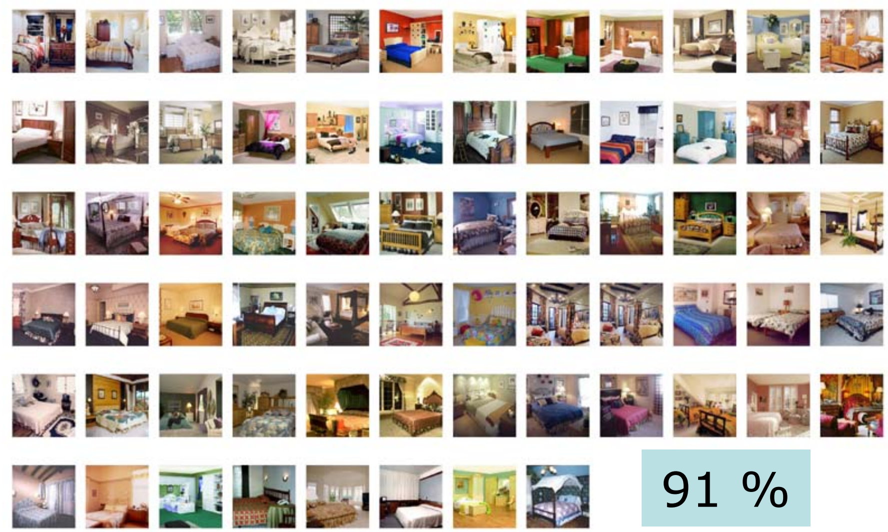
Distinctiveness among exemplars

High homogeneity



63 %

High Heterogeneity

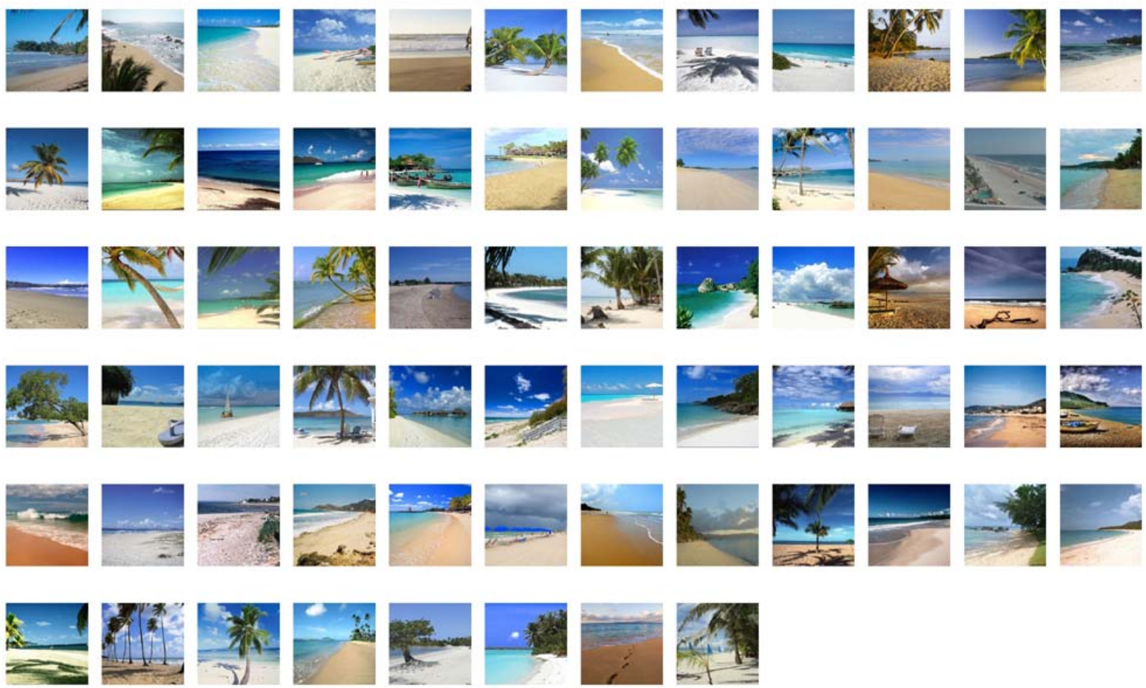


91 %

Standing out details… the novelty factor

Accuracy: 100 %

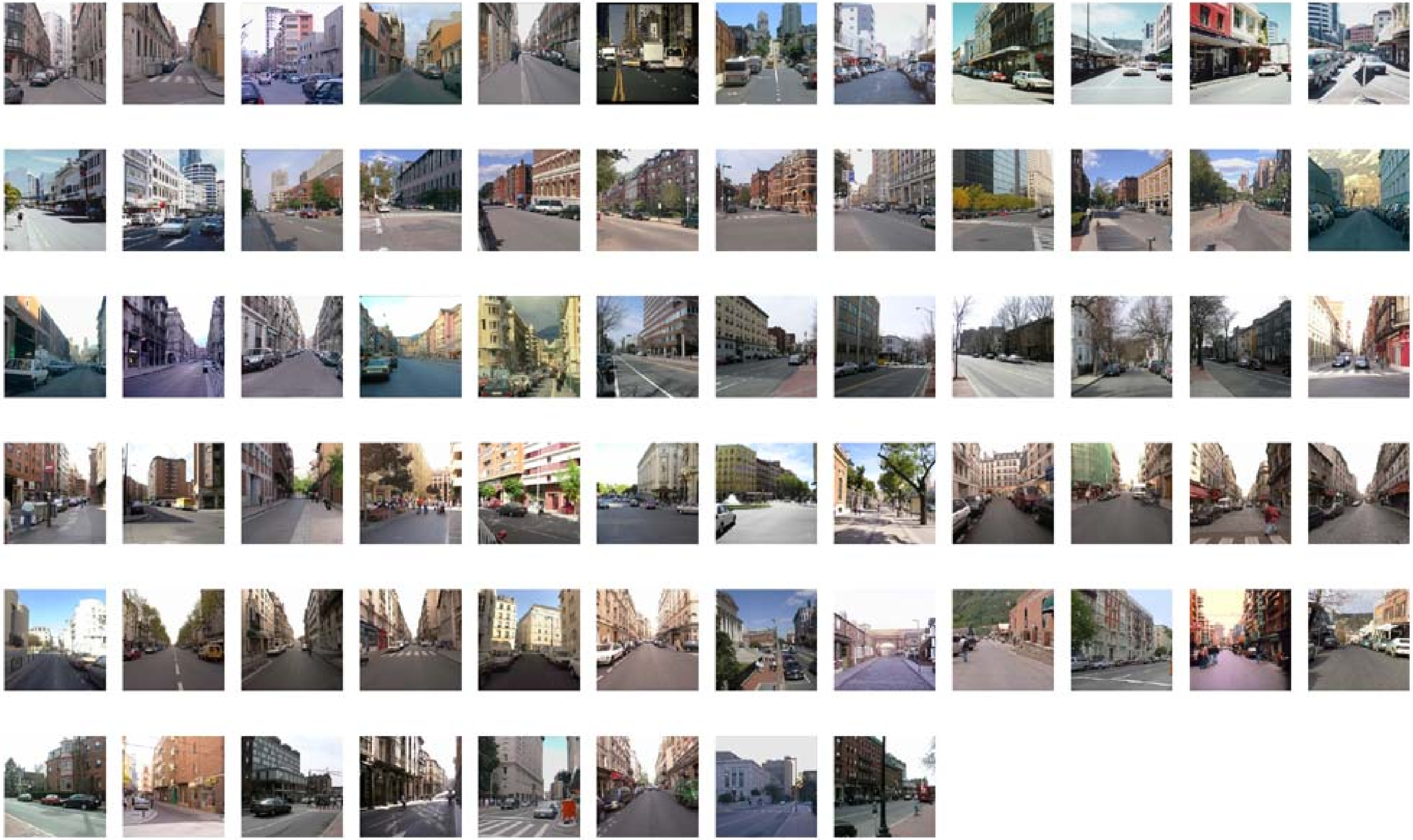
Accuracy: 70 %



Confusion from the Mean …

Accuracy: 62 %

Very typical images are confused



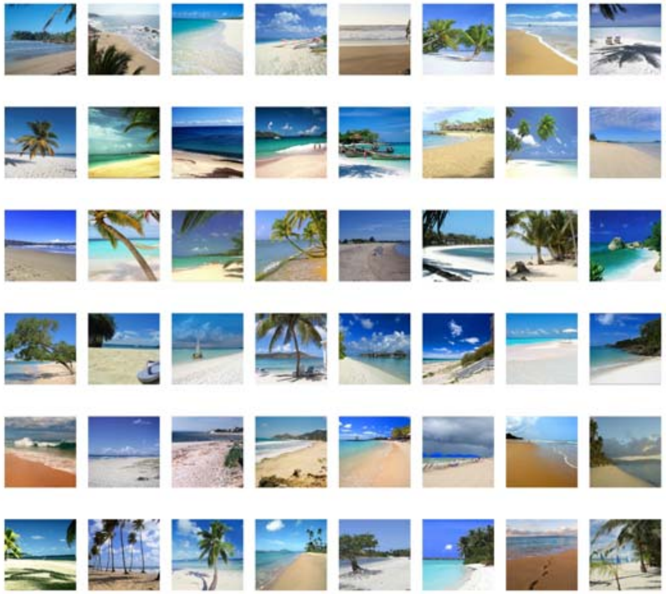
What is this massive visual memory capacity good for?

Models of Object Recognition

- A massive memory for details lend credence to object recognition approaches that require brute force storage of *multiples* viewpoints and exemplars (and image alignment approaches)

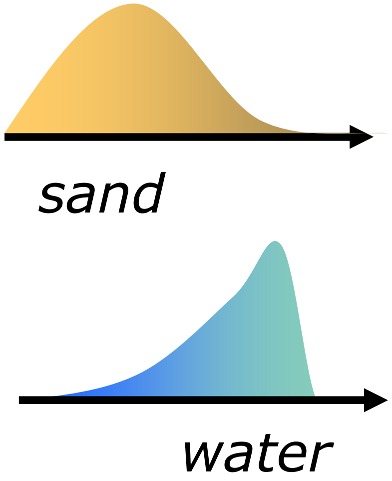
Recognizing the gist of a scene

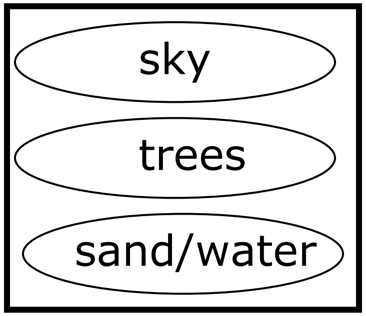
Proposal : Massive memory capacity is the **infrastructure** of **scene gist** recognition



The brain perceives ~ 60 millions *diagnostic* inputs per year (3 samples per second)

A robust representation of natural images require **accumulated** information about the **details**





… the challenge for natural image recognition systems is to find the **relevant regularities** to encode

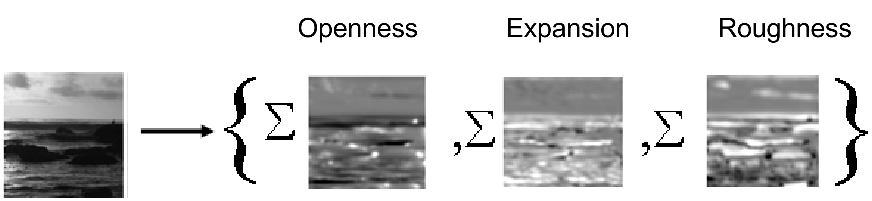
Human Scene Understanding: What are the rules of memory distortion?

Photographic Memory = A unique code per image

Image perceived

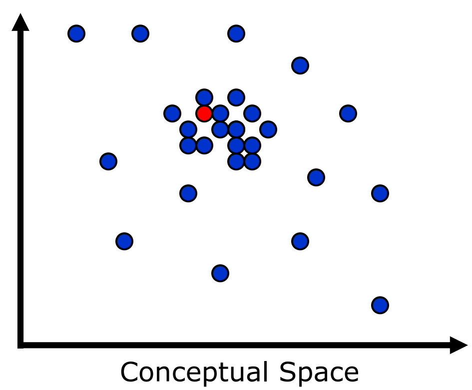
Memory Distortion
Compression - Reconstruction

Image Retrieved



"ocean"

Basis of spatial layout properties



Oliva & Torralba(2001,2002,2006,2007), Torralba & Oliva (2002,2003)
Greene & Oliva (2006, 2007, submitted)

Conclusion

Memory capacity for natural images is of an order of magnitude higher than previously believed

Fidelity of storage of visual details is very high

A unique “conceptual hook” permit to store images with preserved featural details