

# Computational Understanding of Image Memorability

A general overview



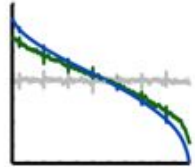
Joint work with Aude Oliva, Antonio Torralba, Phillip Isola, Michelle Borkin, and Melissa Le-Hoa Võ.

# Image Memorability



- objective and quantifiable measure of images
- consistent across observers
- measures utility of information
- filter for visual data

# Application areas



understand human memory



design better user interfaces



diagnose memory problems



design better logos



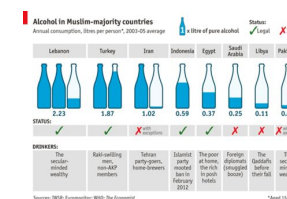
face memory accuracy, biases, and mistakes



design better educational material



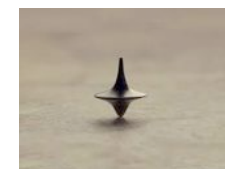
design mnemonic aids



design better data visualizations



smarter visual search



inception, image manipulation



summarize big data and videos



understand and predict cultural trends

Welcome to the

# Visual Memory Game

**A stream of images will be presented  
on the screen for 1 second each.**

Your task:

**Clap your hands** (press a key) anytime you see an image you saw before.

Be attentive, repeats may be separated by many images !

Whenever you press a key, you will get feedback:



You may exit the game at any time and you will be paid in proportion to your progress at that time

**Start Game!**



# Ready?



*(Seriously, get ready to clap. The images go by fast...)*





















**Repeat!**



























**Repeat!**







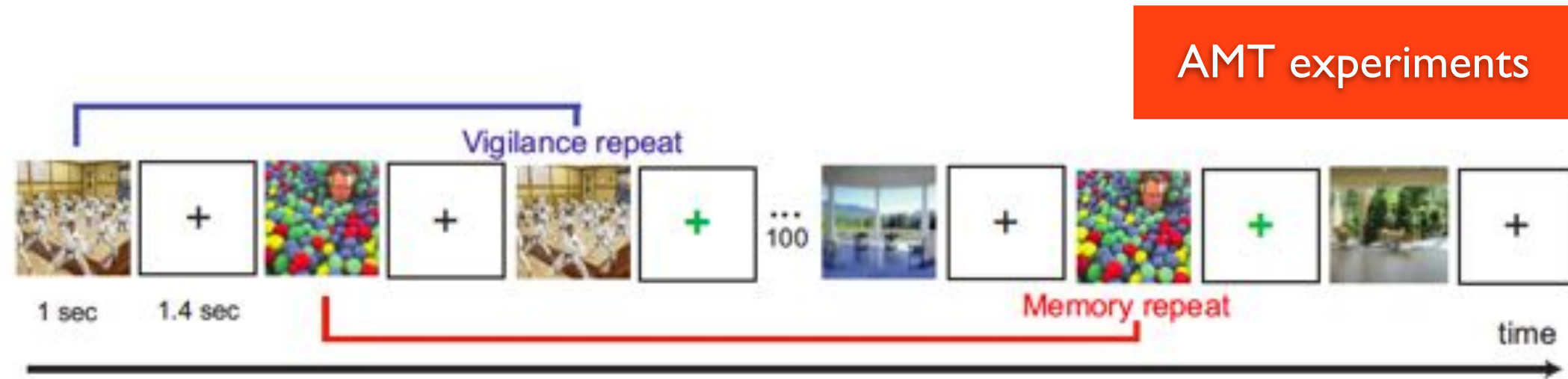






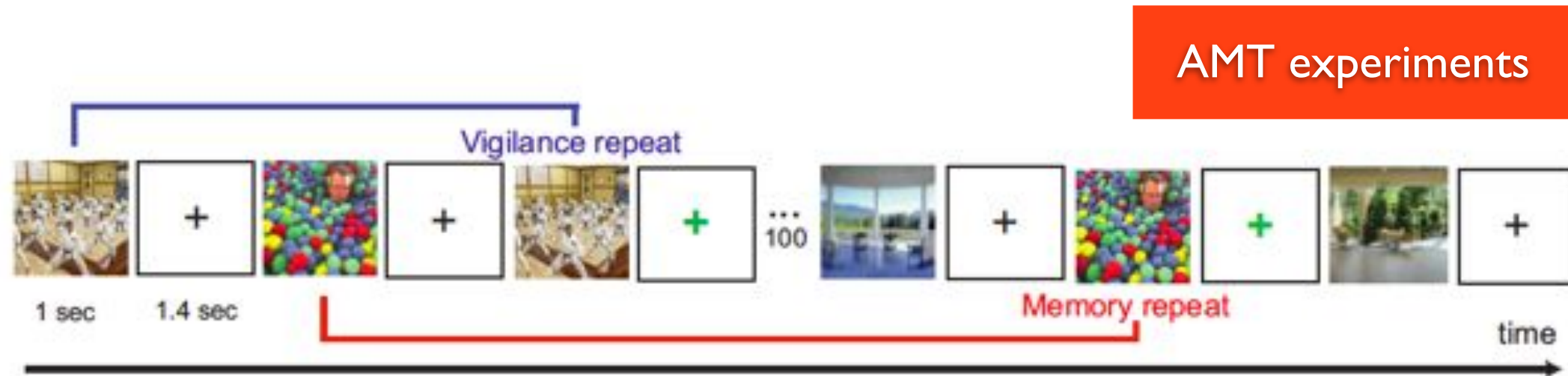
**Repeat!**

# Image Memorability



Isola, P., Xiao, J., Torralba, A., Oliva, A. What makes an image memorable? IEEE CVPR, 2011.

# Image Memorability



measuring memorability:

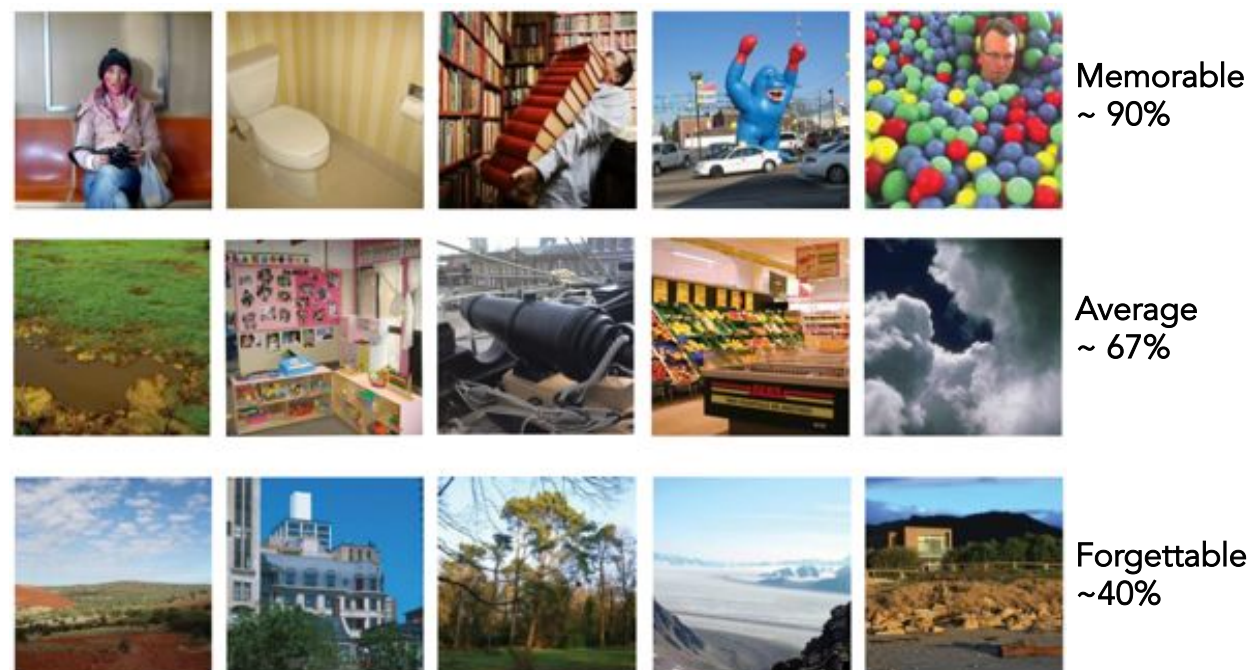
$$\text{HR}(I) = \frac{\text{hits}(I)}{\text{hits}(I) + \text{misses}(I)} \times 100\%$$

Isola, P., Xiao, J., Torralba, A., Oliva, A. What makes an image memorable? IEEE CVPR, 2011.



# Image Memorability

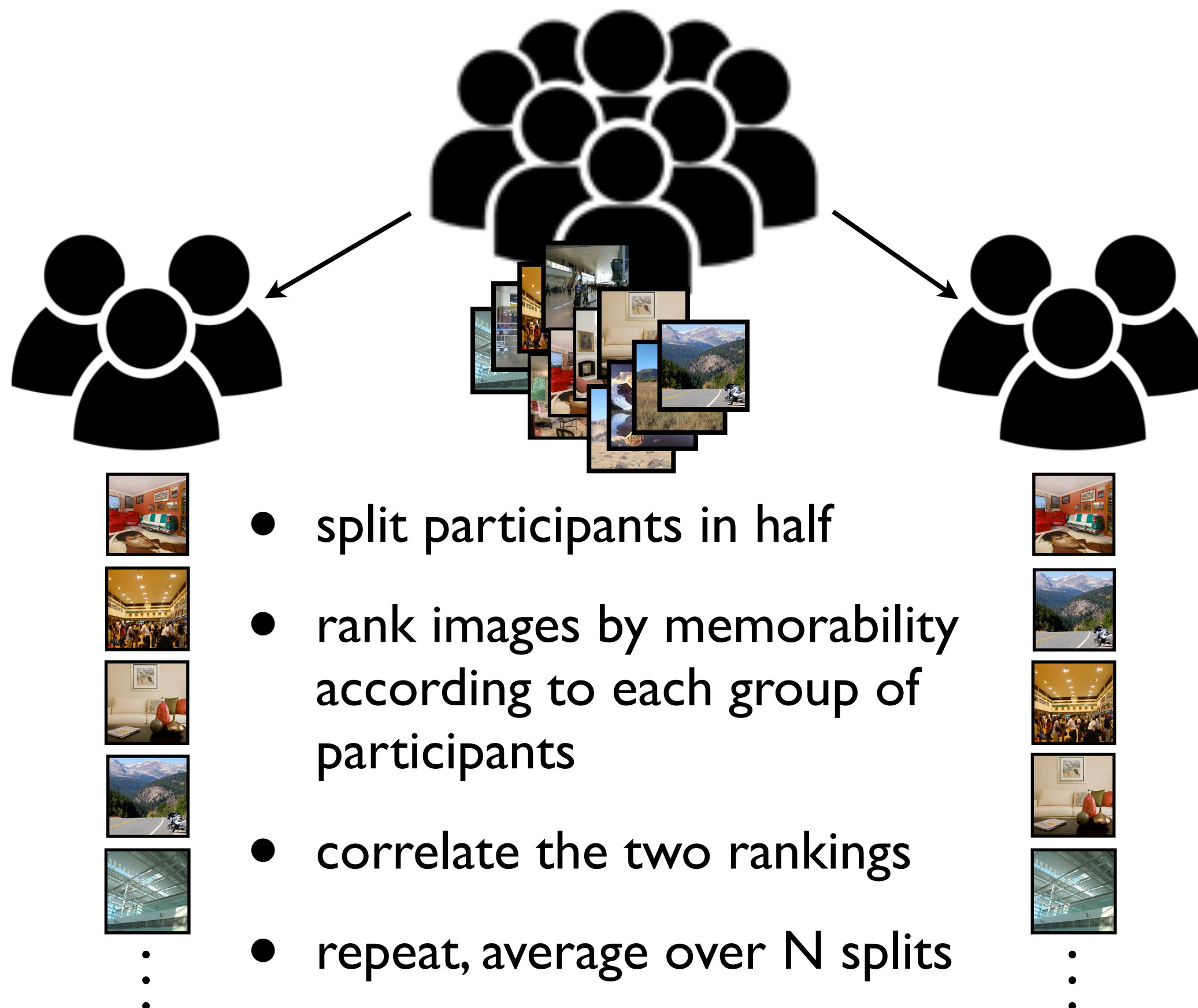
AMT experiments



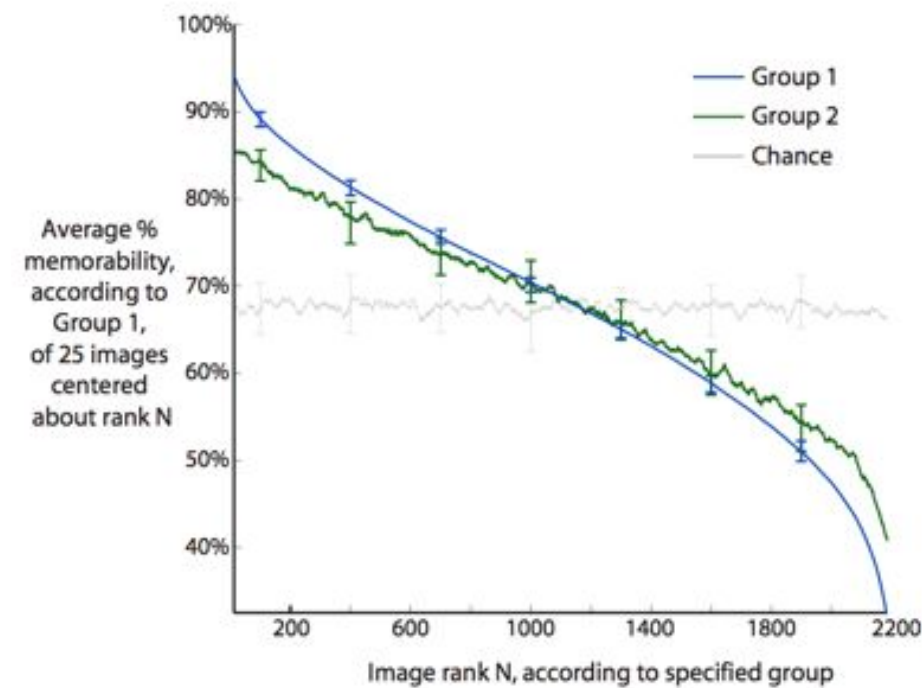
Isola, P., Xiao, J., Torralba, A., Oliva, A. What makes an image memorable? IEEE CVPR, 2011.



# Measuring consistency



# Measuring consistency



consistency  $r = 0.75$

Isola, P., Xiao, J., Torralba, A., Oliva, A. What makes an image memorable? IEEE CVPR, 2011.

# Previous findings on image memorability...

some images are consistently  
memorable, forgettable



a) Most memorable images (86%)



b) Typical images (74%)



c) Least memorable images (34%)

Isola, P., Xiao, J., Torralba, A., Oliva, A. What makes an image memorable? IEEE CVPR, 2011.



# Previous findings on image memorability...

some images are consistently  
memorable, forgettable



memorability is an **intrinsic**  
property of images,  
independent of observer



a) Most memorable images (86%)



b) Typical images (74%)



c) Least memorable images (34%)

Isola, P., Xiao, J., Torralba, A., Oliva, A. What makes an image memorable? IEEE CVPR, 2011.

# Previous findings on image memorability...

some images are consistently  
memorable, forgettable



memorability is an **intrinsic**  
property of images,  
independent of observer



can be computationally predicted



a) Most memorable images (86%)



b) Typical images (74%)



c) Least memorable images (34%)

# Outstanding questions

- does the consistency of human visual memory generalize?
- what factors can modulate image memorability?
- how do differences in memorability behave over time?

- **does the consistency of human visual memory generalize?**
- what factors can modulate image memorability?
- how do differences in memorability behave over time?



# FIGRIM data collection

21 SUN categories:  
8 indoor, 13 outdoor



airport terminal  
amusement park  
badlands  
bathroom  
bedroom  
bridge  
castle  
cockpit  
conference room  
dining room  
golf course  
highway  
house  
kitchen  
lighthouse  
living room  
mountain  
pasture  
playground  
skyscraper  
tower

those with at least 300  
700x700 images

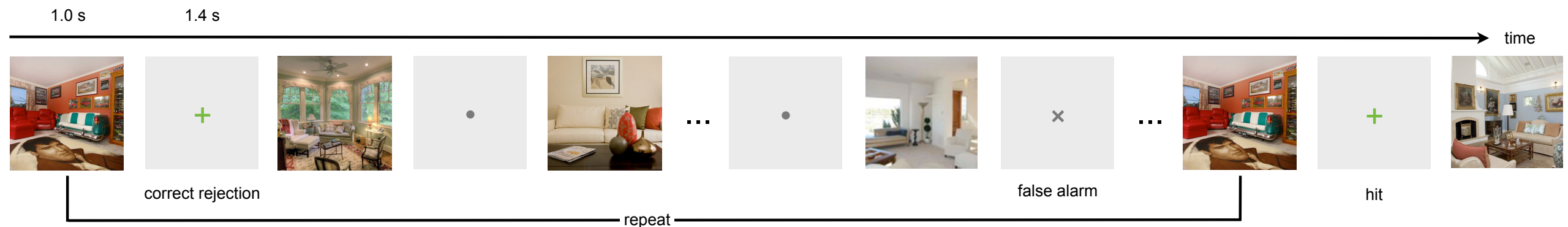


# FIGRIM data collection

21 SUN categories:  
8 indoor, 13 outdoor

ran 21 separate experiments:  
one per scene category

airport terminal  
amusement park  
badlands  
bathroom  
bedroom  
bridge  
castle  
cockpit  
conference room  
dining room  
golf course  
highway  
house  
kitchen  
lighthouse  
living room  
mountain  
pasture  
playground  
skyscraper  
tower

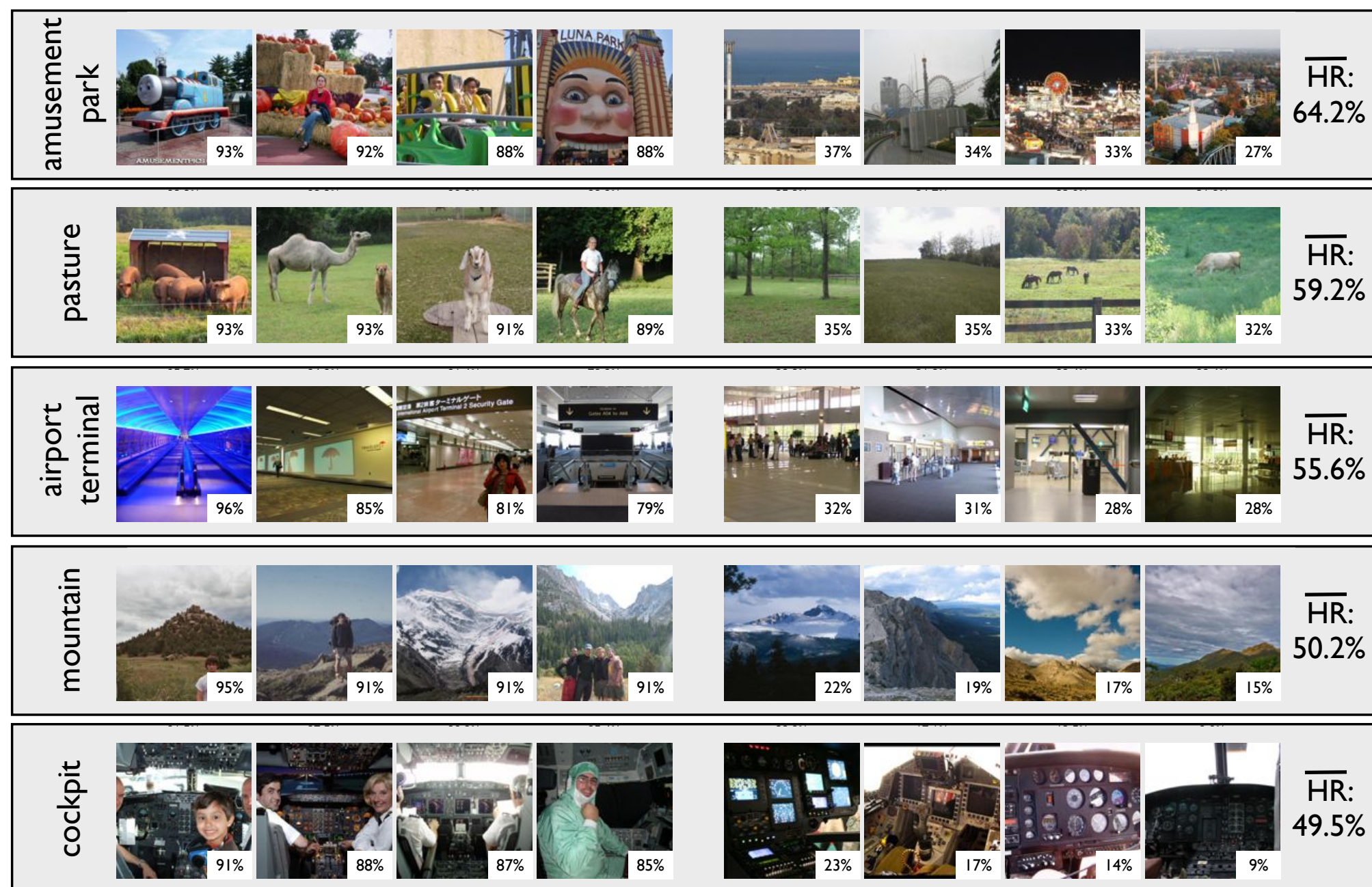


Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. Intrinsic and Extrinsic Effects on Image Memorability, Vision Research, 2015.

# sample of FIGRIM Dataset

highest memorability

lowest memorability



memorability scores for 1754 targets across 21 categories (7296 fillers)

# What about within individual scene categories?

92.9%



91.5%



95.7%



91.5%



36.8%



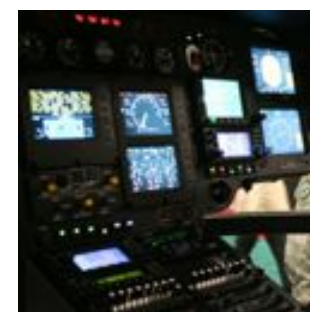
36.6%



32.3%



22.9%



- people are still consistent, even within categories!
- consistency ranges from  $r = 0.69$  to  $0.86$  for HR of individual categories

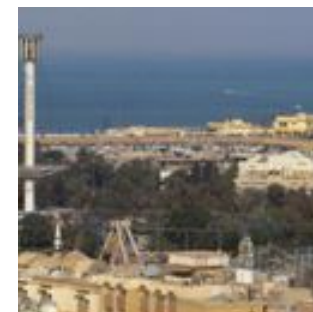


# All scene categories have intrinsically memorable and forgettable images

92.9%



36.8%



91.5%



36.6%



95.7%



32.3%



91.5%



22.9%



- people are still consistent, even within categories!
- consistency ranges from  $r = 0.69$  to  $0.86$  for HR of individual categories
- memorable instances have unusual objects, layouts, perspectives, or colors

will be quantified

- does the consistency of human visual memory generalize?      **yes, within scene categories**
- what factors can modulate image memorability?
- how do differences in memorability behave over time?

- does the consistency of human visual memory generalize?
- **what factors can modulate image memorability?**
- how do differences in memorability behave over time?

# FIGRIM data collection

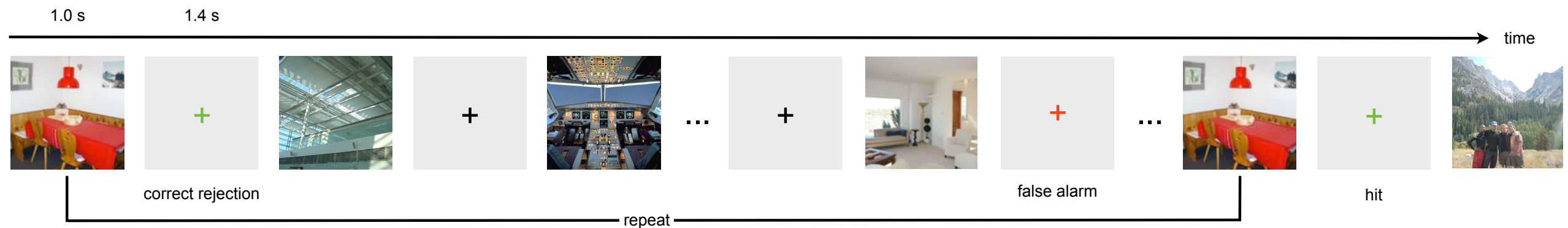
ran 21 separate experiments:  
one per scene category

21 SUN categories:  
8 indoor, 13 outdoor

AND

an additional experiment, combining  
images from all categories

airport terminal  
amusement park  
badlands  
bathroom  
bedroom  
bridge  
castle  
cockpit  
conference room  
dining room  
golf course  
highway  
house  
kitchen  
lighthouse  
living room  
mountain  
pasture  
playground  
skyscraper  
tower



# Data collection

ran 21 separate experiments:  
one per scene category

within-category  
experiment

AND

an additional experiment, combining  
images from all categories

across-category  
experiment



# Data collection

ran 21 separate experiments:  
one per scene category

**within-category  
experiment**

AND

an additional experiment, combining  
images from all categories

**across-category  
experiment**

**both sets of experiments contain the  
same images in different contexts.**

# Quantifying context effects on memorability

- large body of memory literature suggests that “distinct” or “unique” items (those that stand out from their context) are better remembered
- our goal is to quantify this intuition using our large set of natural scene images

# Quantifying context effects on memorability

- call an image **contextually distinct** if it is distinct with respect to the other images in its context



# Applying classical principles of information theory

- compute a probability distribution over the images in a given context (in some feature space)



Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. Intrinsic and Extrinsic Effects on Image Memorability, Vision Research, 2015.

# Applying classical principles of information theory

- compute a probability distribution over the images in a given context (in some feature space)
- estimate the likelihood of the image under this distribution



Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. Intrinsic and Extrinsic Effects on Image Memorability, Vision Research, 2015.

# Applying classical principles of information theory

feature mapping      image

$$f_i = F(I)$$

kernel distribution      Epanechnikov kernel

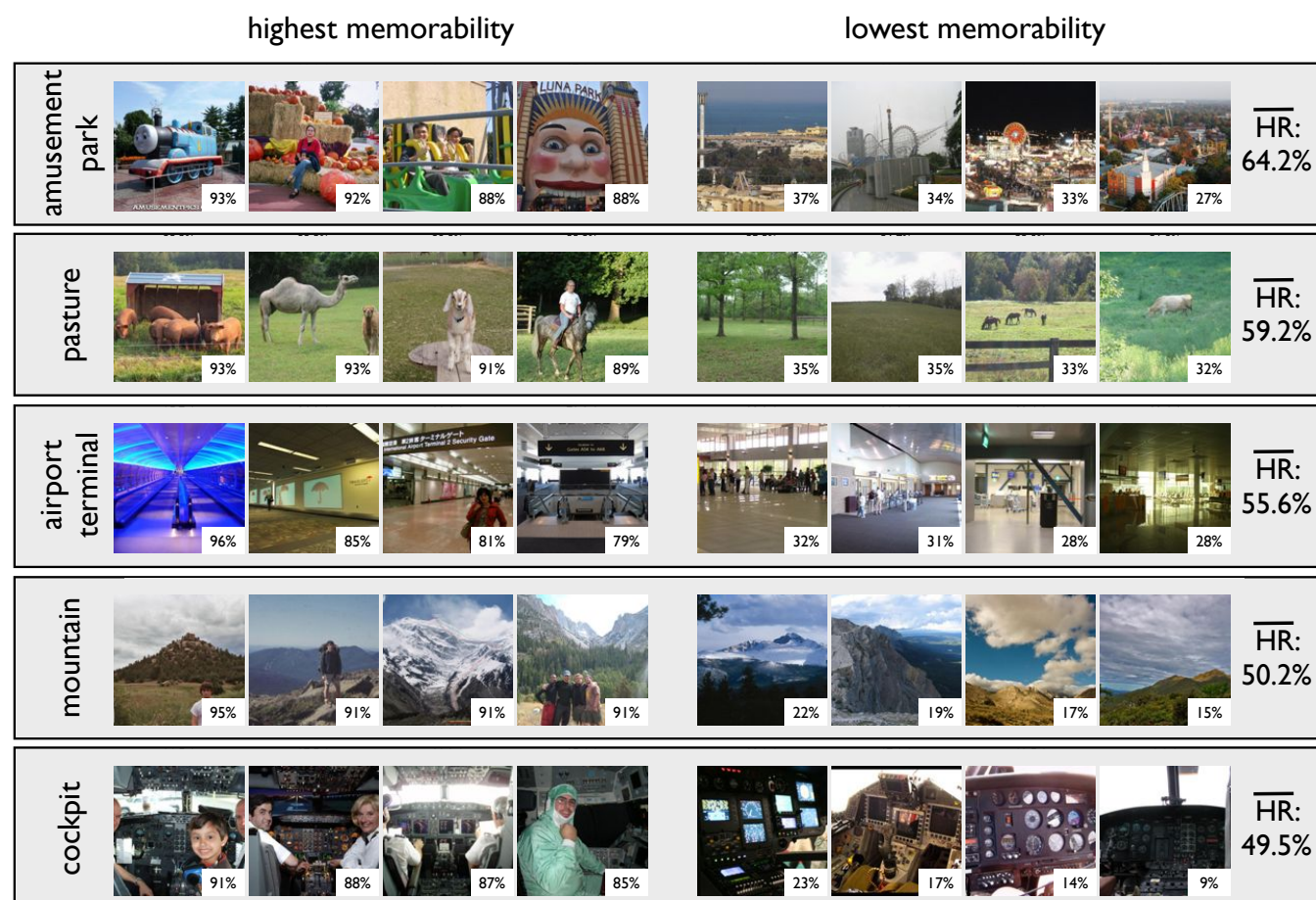
$$P_c(f_i) = \frac{1}{\|C\|} \sum_{j \in C} K(f_i - f_j)$$

distinctiveness of an image  
with respect to a context  $C$

$$D(I; C) = -\log P_c(f_i)$$



# Contextually distinct images are more memorable



Spearman  
correlation:  
0.24 ( $p < 0.01$ )

$$\text{HR}(I) = \frac{\text{hits}(I)}{\text{hits}(I) + \text{misses}(I)} \times 100\%$$

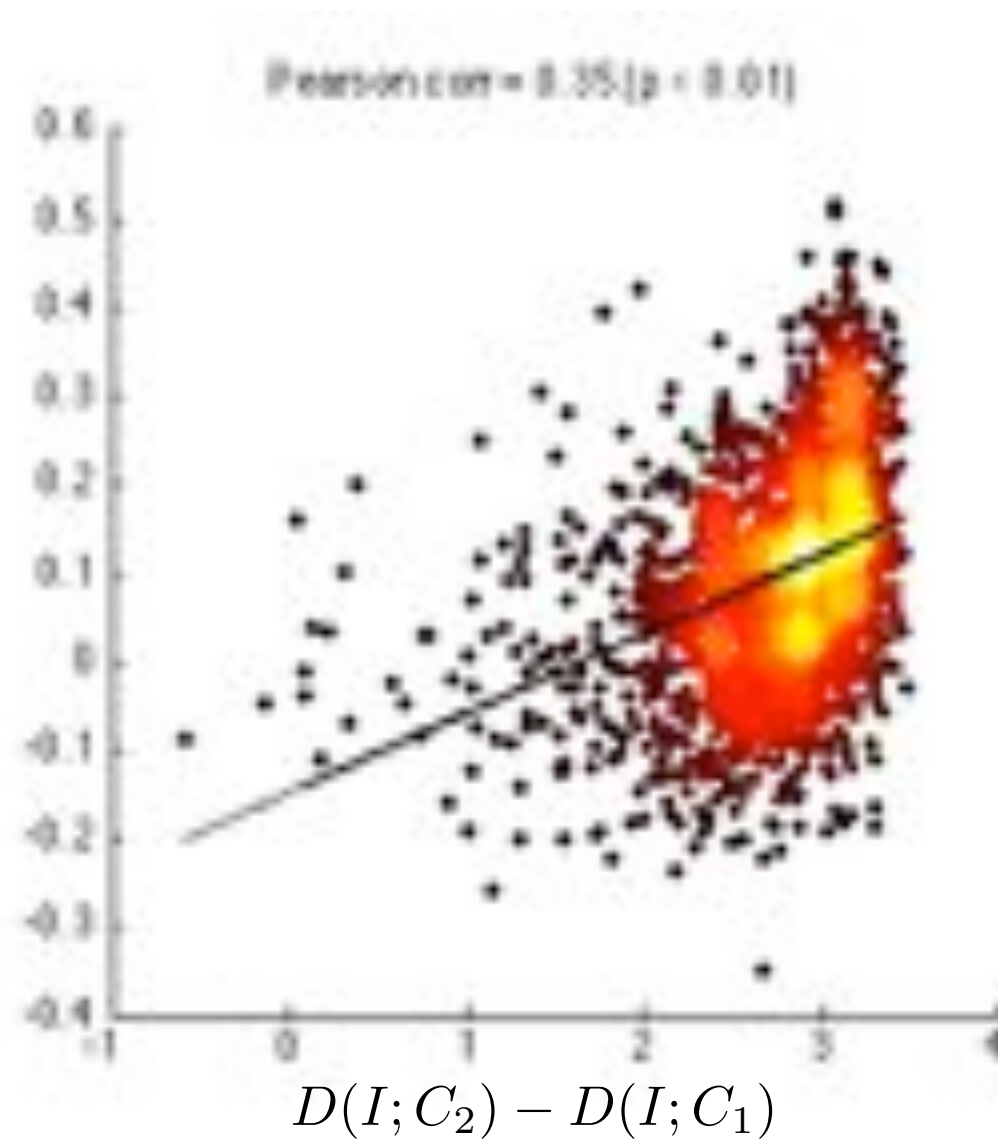
$$D(I; C) = -\log P_c(f_i)$$

using: Places-CNN  
features

# Changing context changes memorability



$$HR_{c_2}(I) - HR_{c_1}(I)$$



using: Places-CNN features



# Contextually distinct images are more memorable

memorable within-categories



# Contextually distinct images are more memorable

memorable within-categories



memorable across categories



want to select images like these  
with more stable memorability

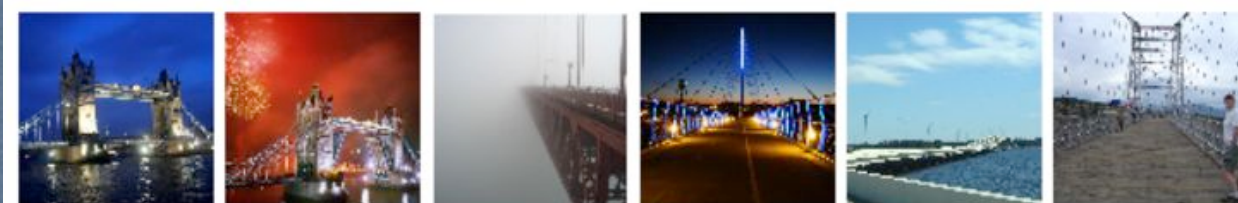
# Contextually distinct images are more memorable

## memorable within-categories



corr. b/w HR and  $D(I;AMT1) = 0.26$  ( $p < 0.01$ )

## memorable across categories



corr. b/w HR and  $D(I;AMT2) = 0.24$  ( $p < 0.01$ )

using: Places-CNN features

# Images most affected by context

amusement parks that dropped in memorability  
when combined with other categories

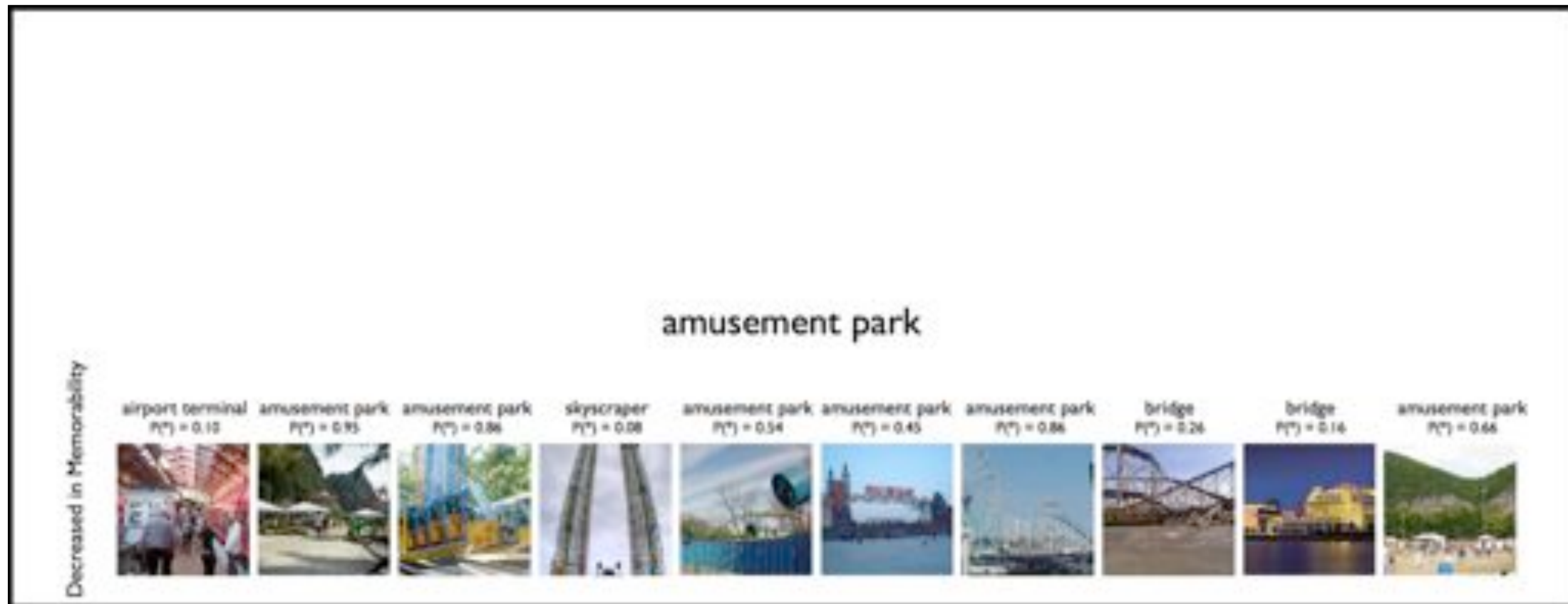


bedrooms that dropped in memorability when  
combined with other categories



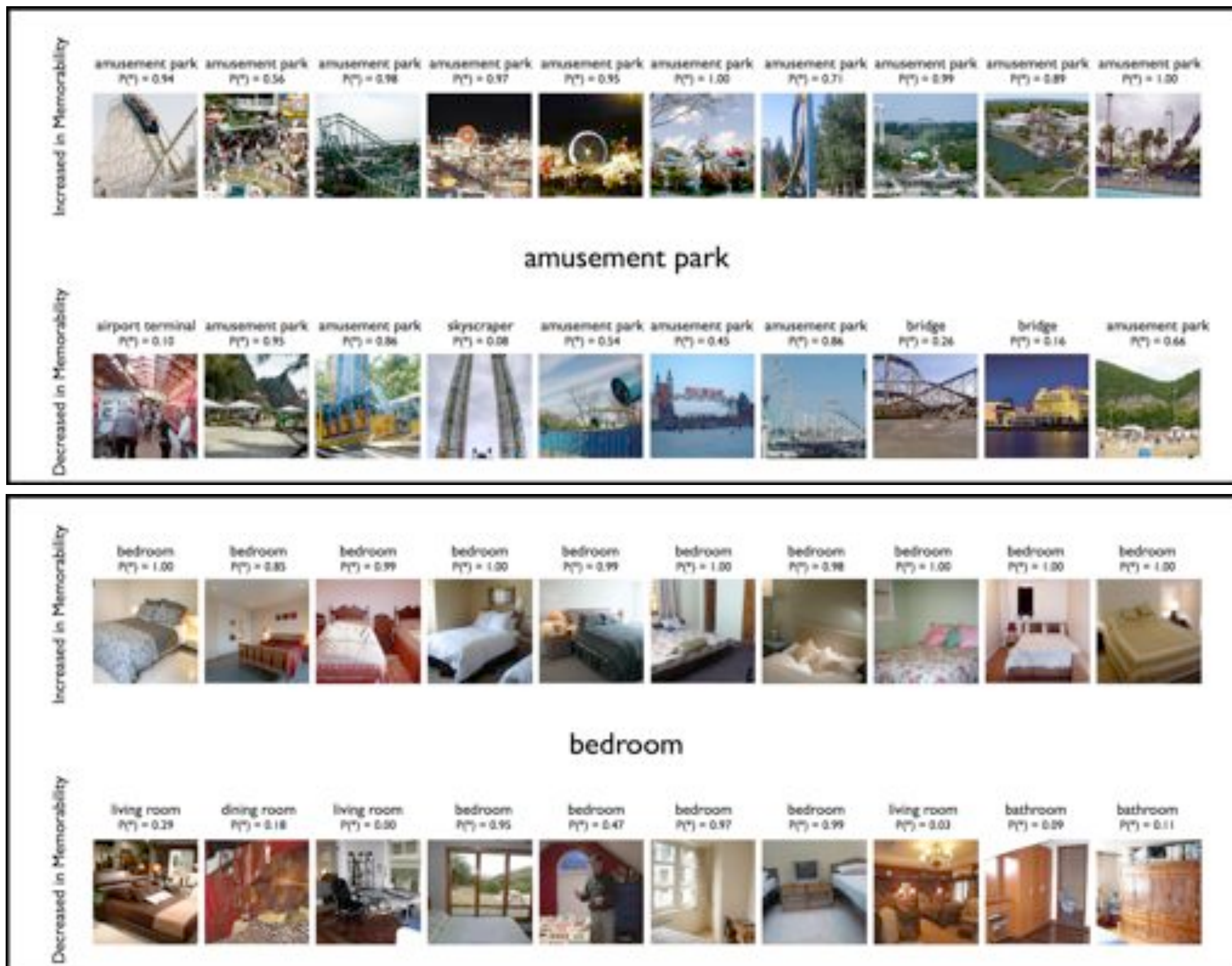


# What does a classifier think?



Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. Intrinsic and Extrinsic Effects on Image Memorability, Vision Research, 2015.

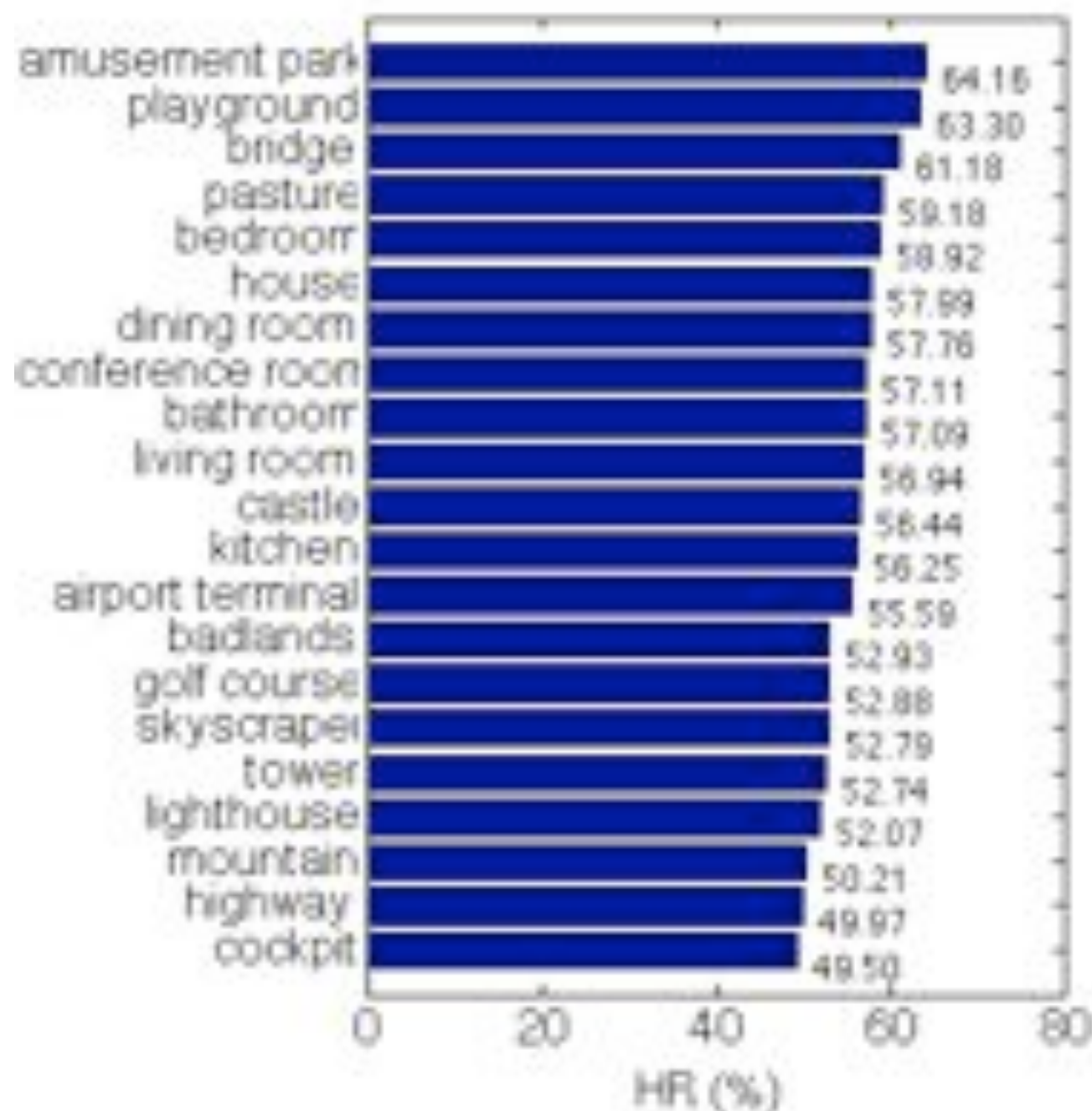
# What does a classifier think?



Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. Intrinsic and Extrinsic Effects on Image Memorability, Vision Research, 2015.



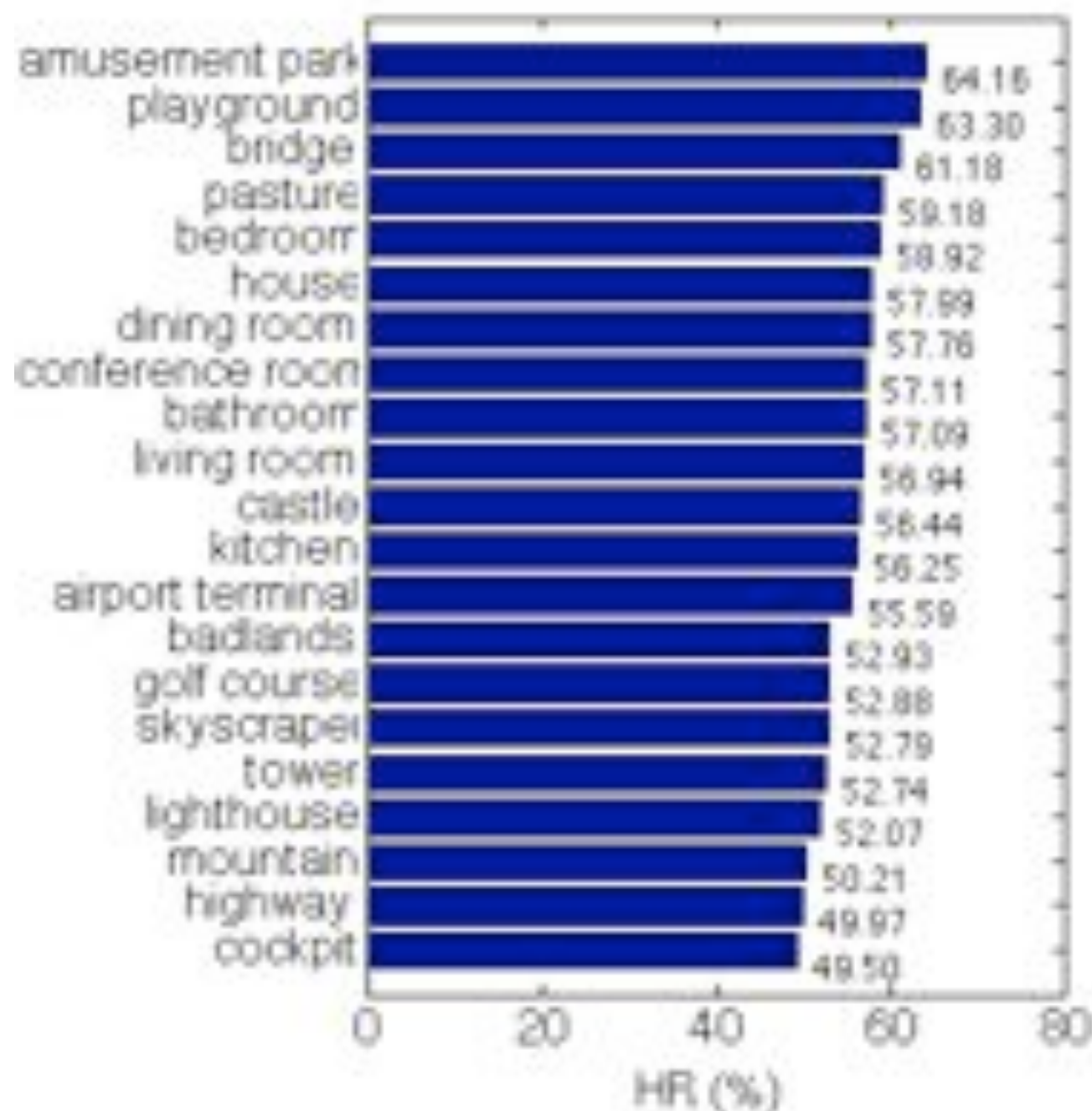
# Memorability of a context



- some scene categories are intrinsically more memorable
- relative ranking of the scene categories is stable (correlation of 0.68 across different splits of images)



# Memorability of a context



- some scene categories are intrinsically more memorable will be quantified
- they contain more variety
- relative ranking of the scene categories is stable (correlation of 0.68 across different splits of images)

# Quantifying context entropy

distinctiveness of an image  
with respect to a context  $C$



$$D(I; C) = -\log P_c(f_i)$$

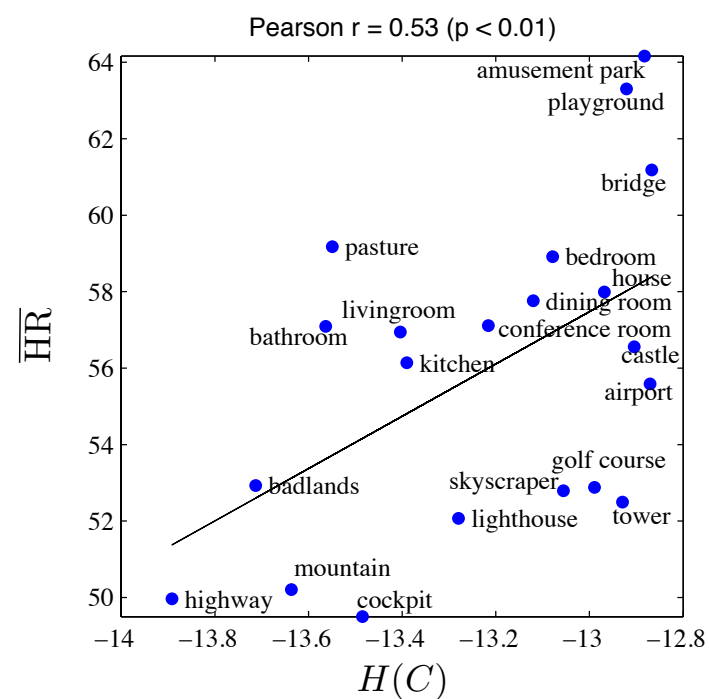
context entropy



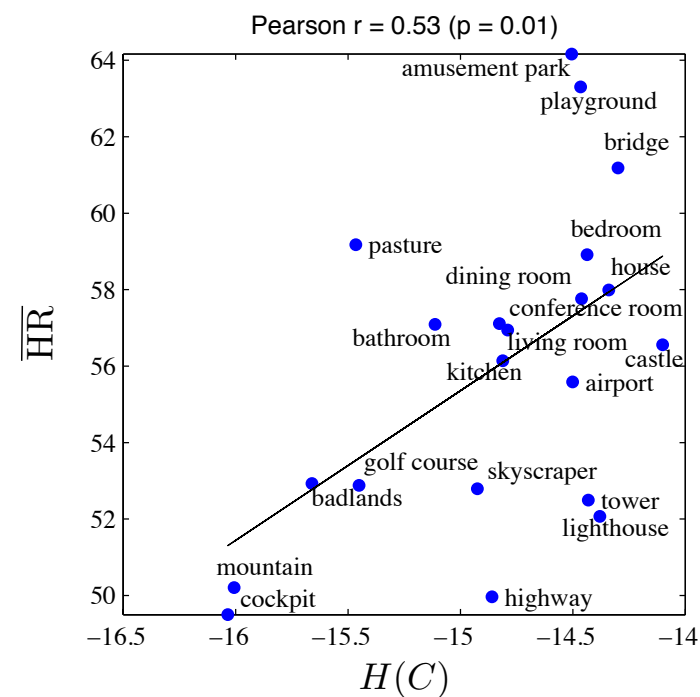
$$H(C) = \mathbb{E}_c[-\log P_c(f_i)]$$

# Memorability increases with context variability

## GIST features

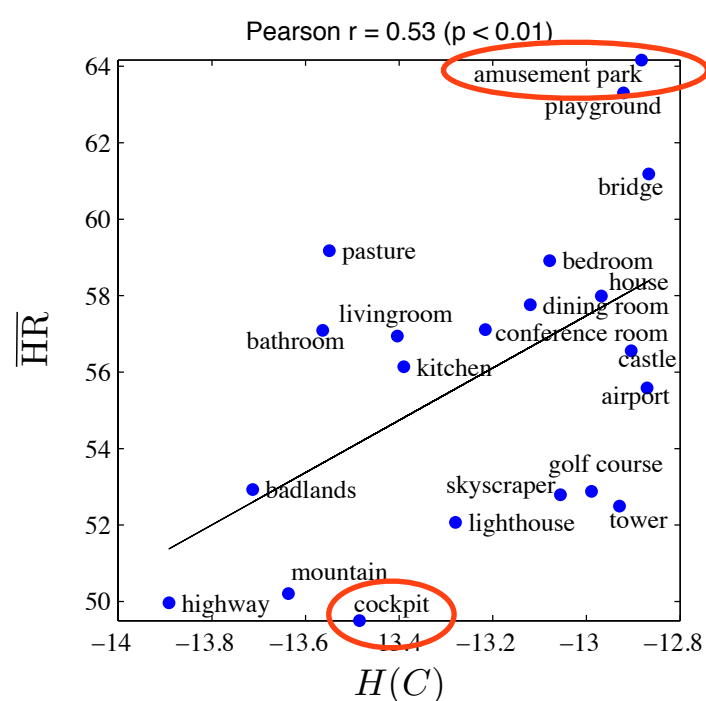


## CNN features

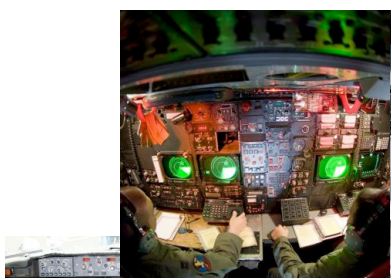
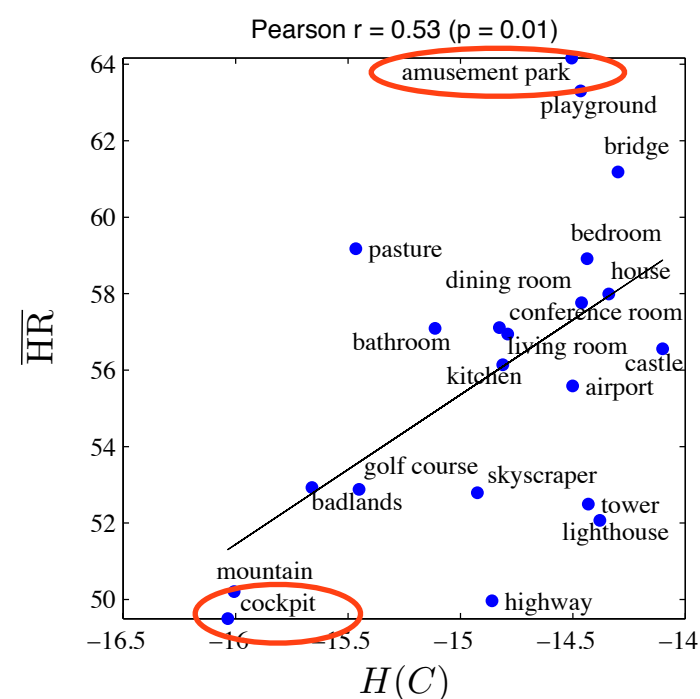


# Memorability increases with context variability

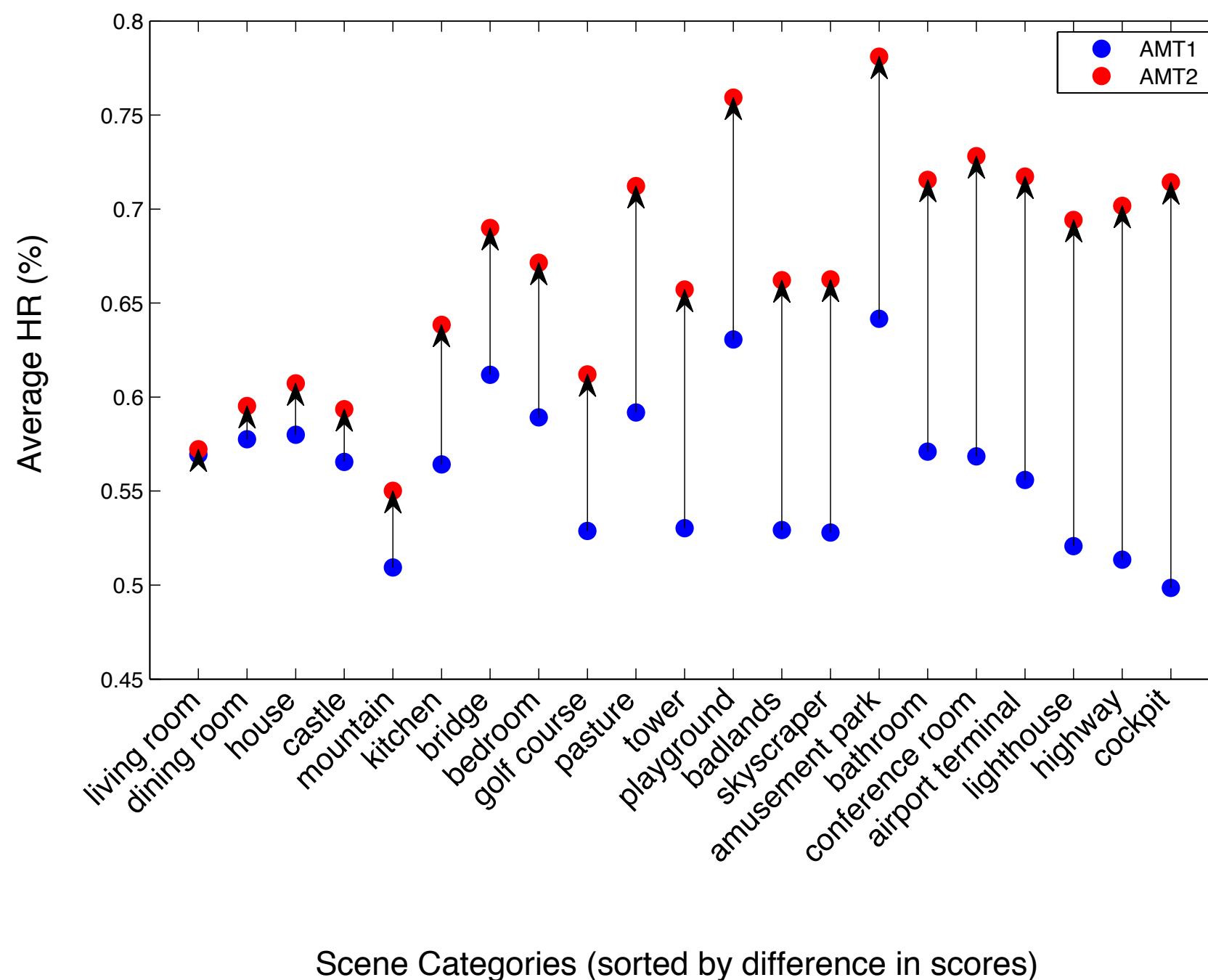
## GIST features



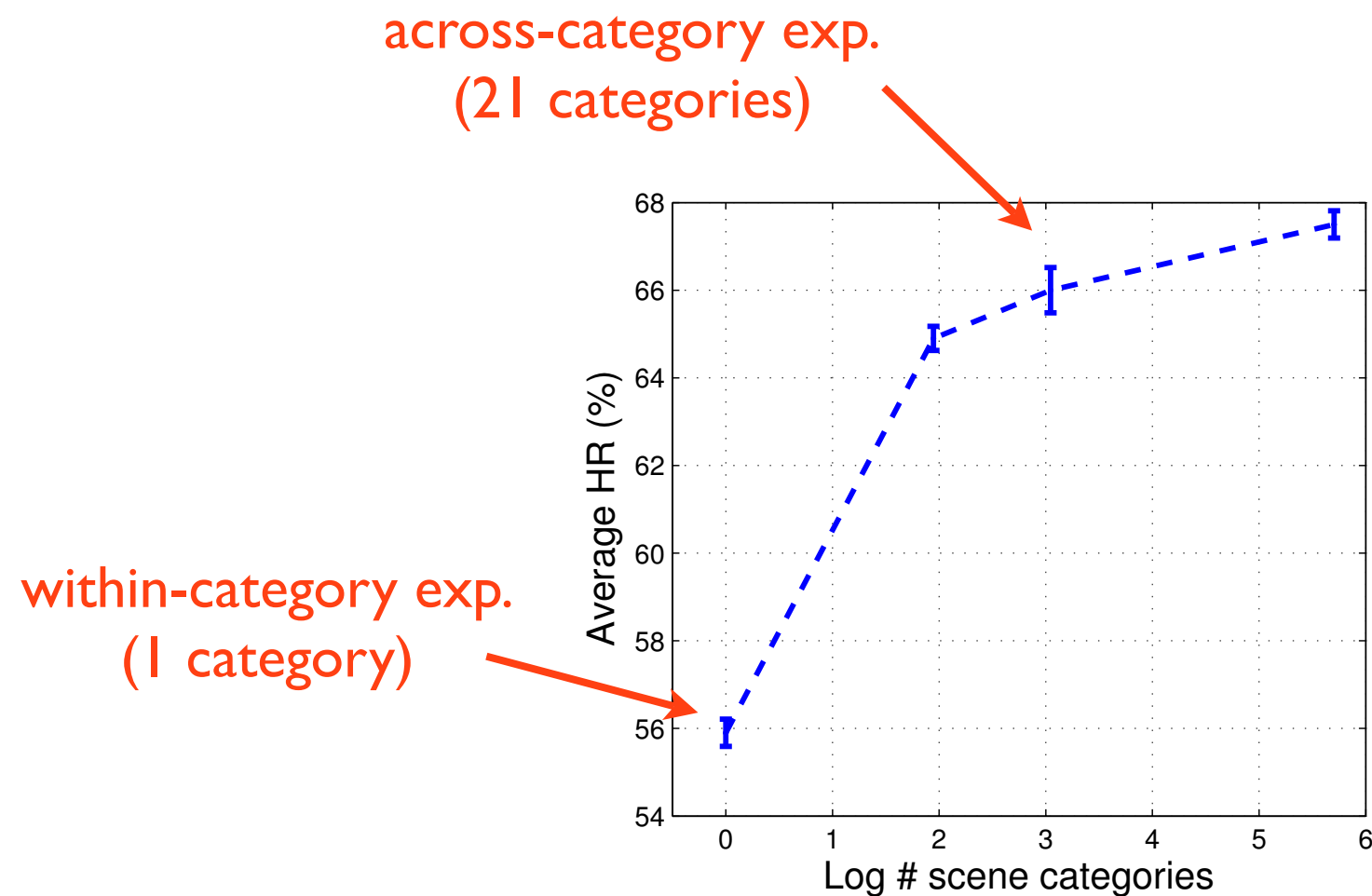
## CNN features



# Memorability increases with context variability

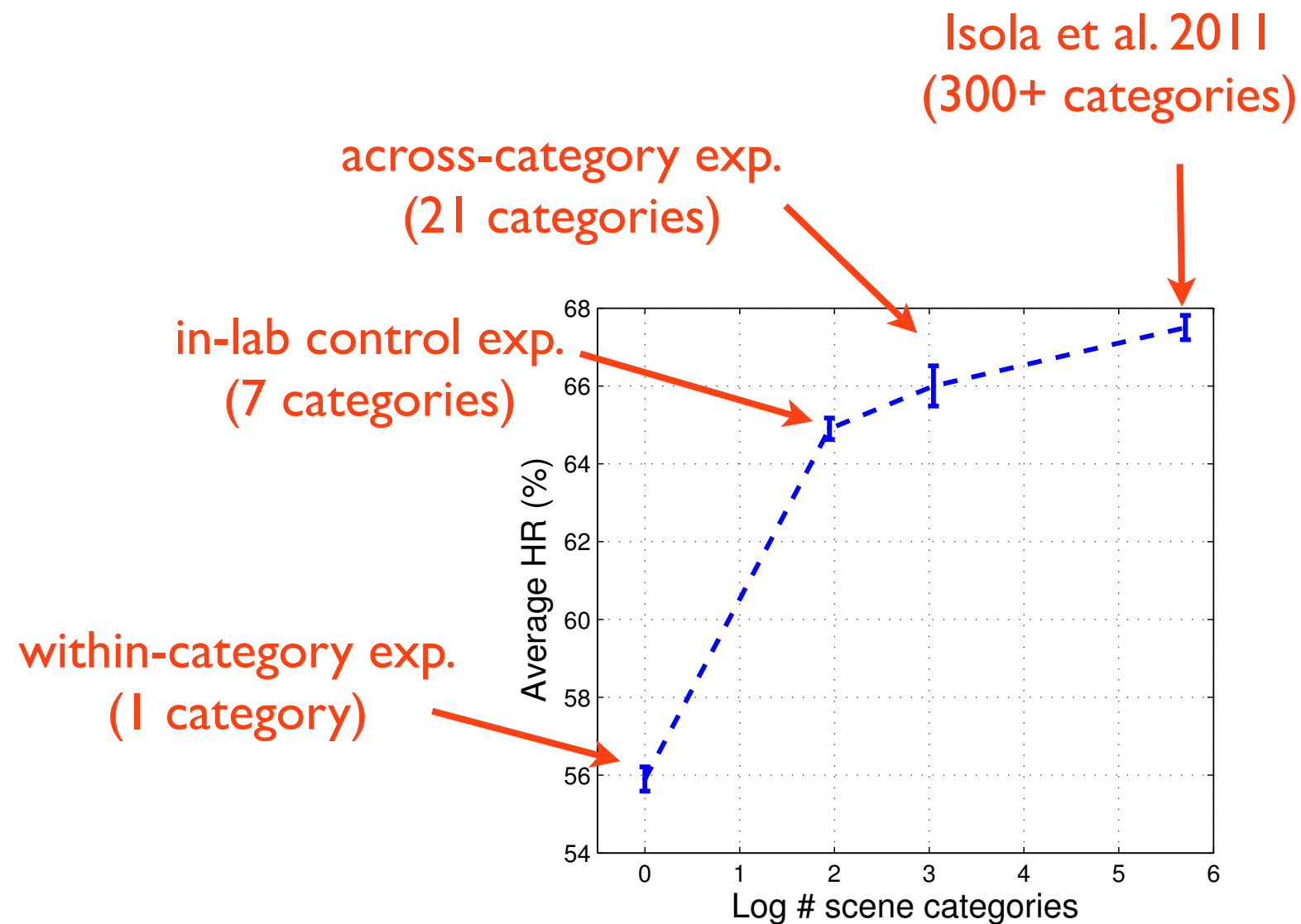


# Memorability increases with context variability



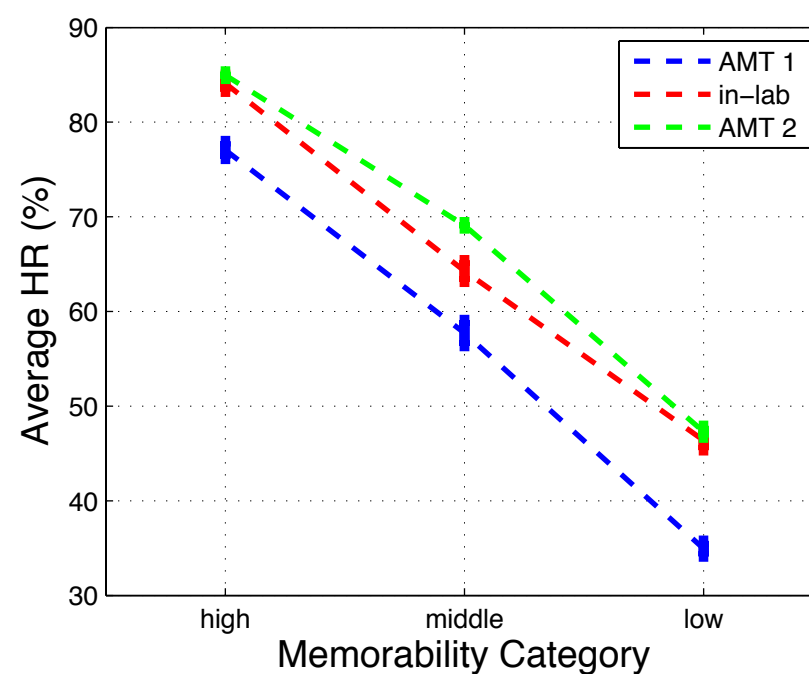


# Memorability increases with context variability



supports classical results in memory research!

# Consistency across memorability experiments



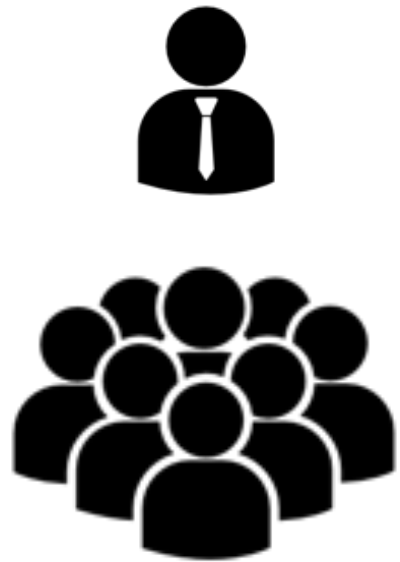
Dataset	targets	fillers	datapoints/target	mean HR (%)	mean FAR (%)	HR cons. ( $\rho$ )	FAR cons. ( $\rho$ )
FIGRIM	1754	7296	74	66.0 (SD: 13.9)	11.1 (SD: 9.5)	0.74	0.72
Isola [2]	2222	8220	78	67.5 (SD: 13.6)	10.7 (SD: 7.6)	0.75	0.66
Faces [4]	2222	6468	82	51.6 (SD: 12.6)	14.4 (SD: 8.7)	0.68	0.69
Visualizations [5]	410	1660	87	55.4 (SD: 16.5)	13.2 (SD: 10.7)	0.83	0.78

- does the consistency of human memory generalize?
- **what factors can modulate image memorability?**
- how do differences in memorability behave over time?

*Goal:* to make memory predictions for a  
specific individual on a specific trial



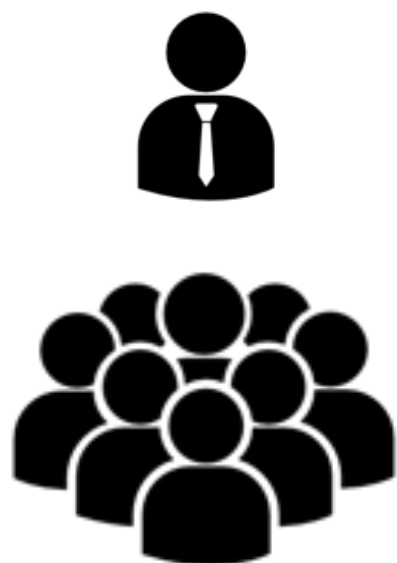
# Procedure



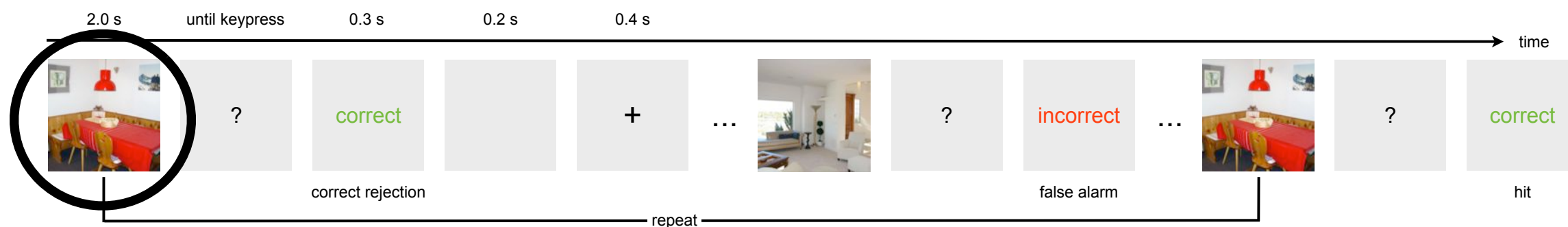
- leave one user out
- to predict if that user will remember a particular image...
- we use the rest of the crowd to make predictions on a per-image, per-trial basis



# Procedure



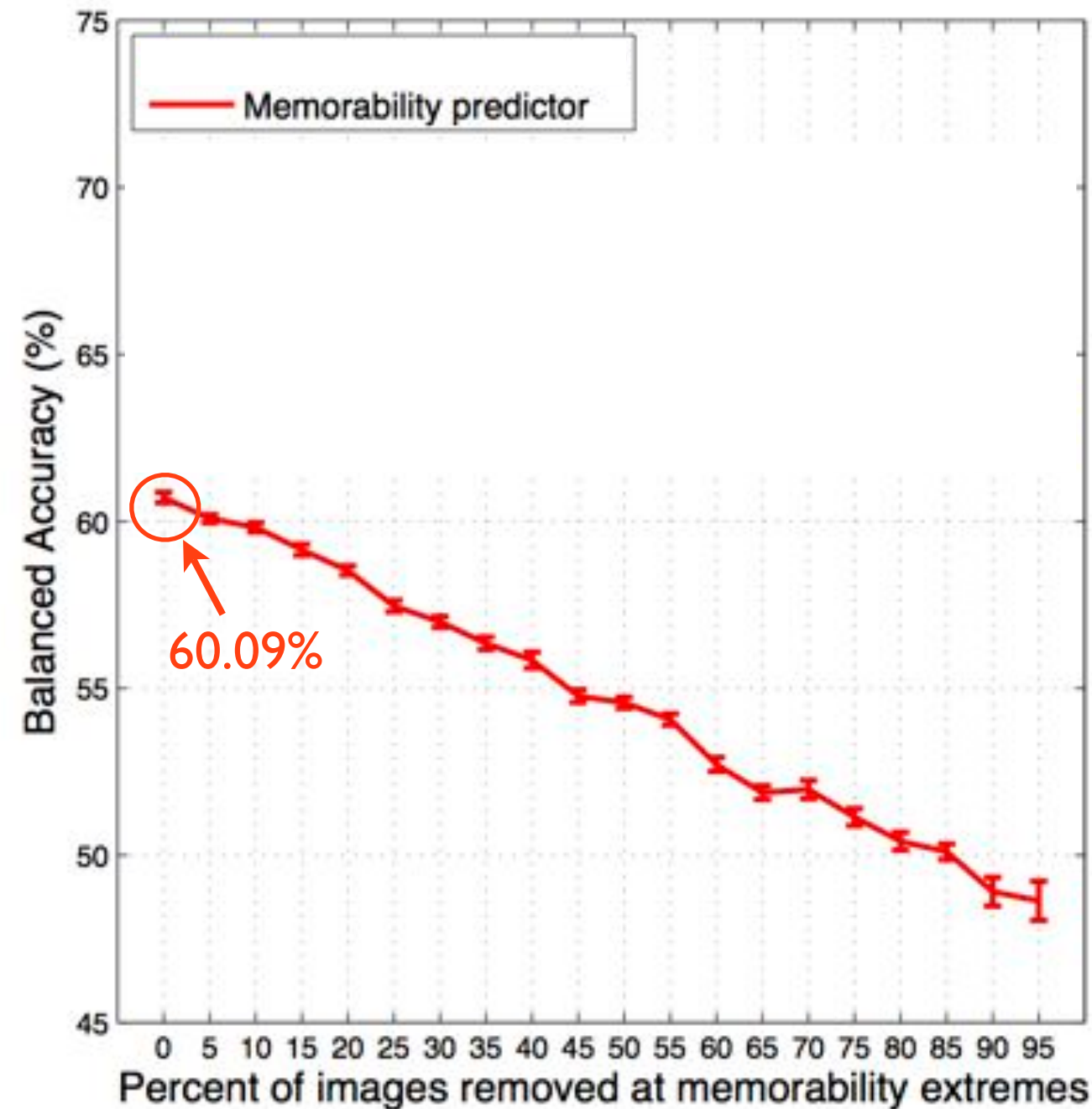
- leave one user out
- to predict if that user will remember a particular image...
- we use the rest of the crowd to make predictions on a per-image, per-trial basis



$$\text{balanced accuracy} = \frac{0.5 \times \text{true positives}}{\text{true positives} + \text{false negatives}} + \frac{0.5 \times \text{true negatives}}{\text{true negatives} + \text{false positives}}$$

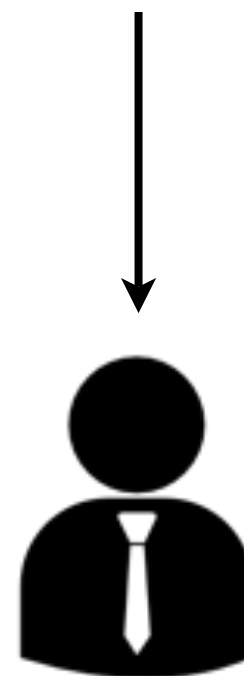


# How well can we predict if people will remember an image?



Need to make better memory predictions for a specific user on a specific trial

Consider the behavior,  
attentional biases of the user



We use eye-movements to predict  
if an image will be remembered



# Eye-tracking experiments

We use eye-movements to predict if an image will be remembered

note: can only respond after image,  
forced-choice keypress



- 630 targets/experiment
- 42 participants in total (~16.2 per image)

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. Intrinsic and Extrinsic Effects on Image Memorability, Vision Research, 2015.

# Eye-tracking model

**Idea:** for a set of fixations on an image, evaluate how likely it is that they came from this image

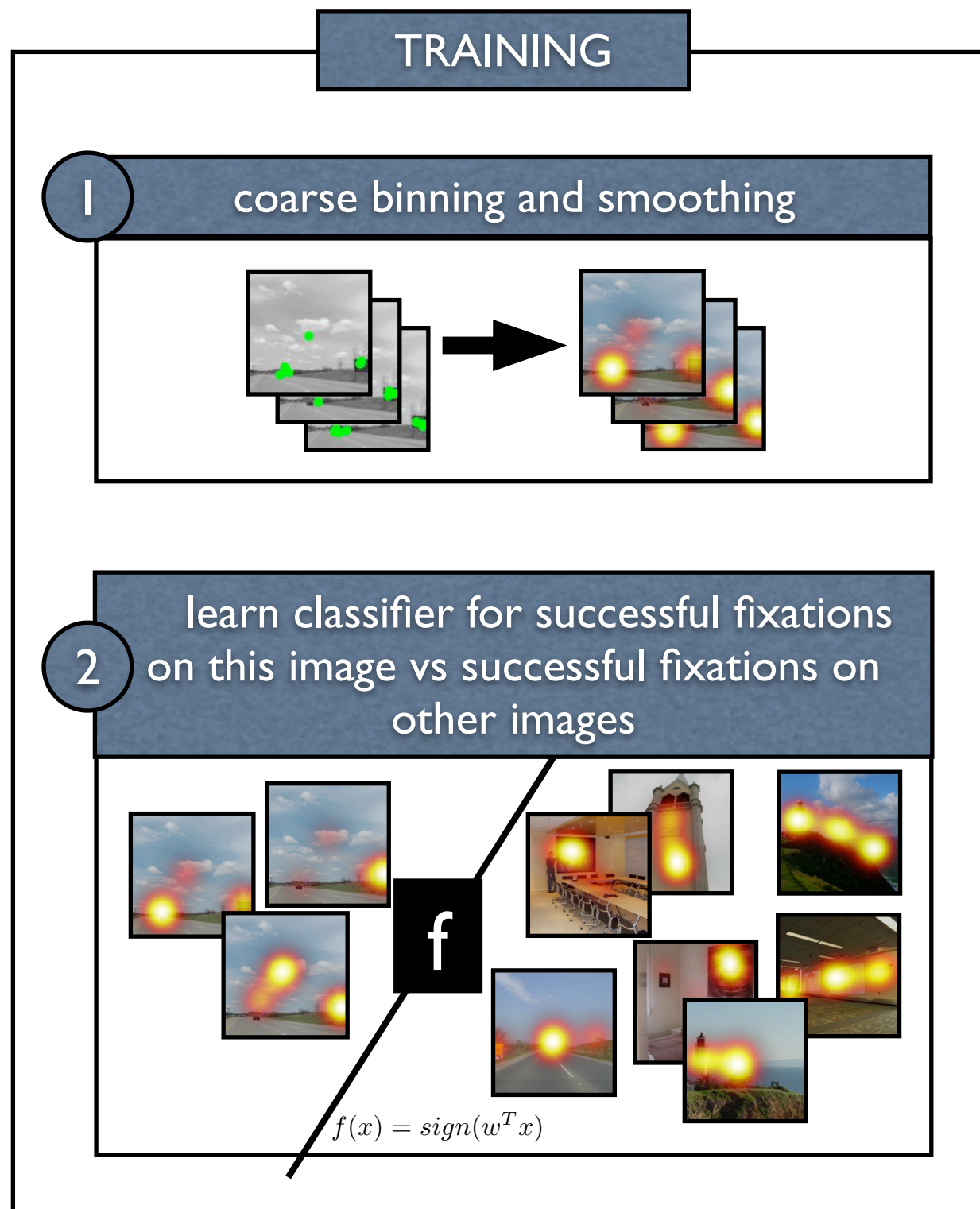
if very likely, assume user was paying attention to the image, successfully encoded it, and will remember it later



if not very likely, assume user was not paying attention to the image, failed to encode it, and will not remember it later



# Classification of fixations

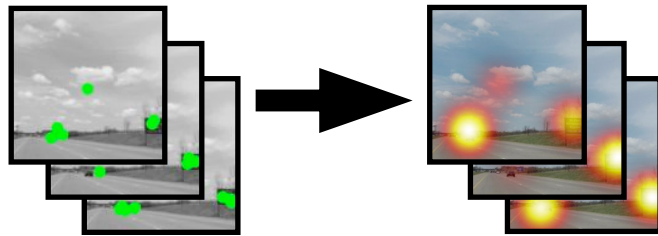




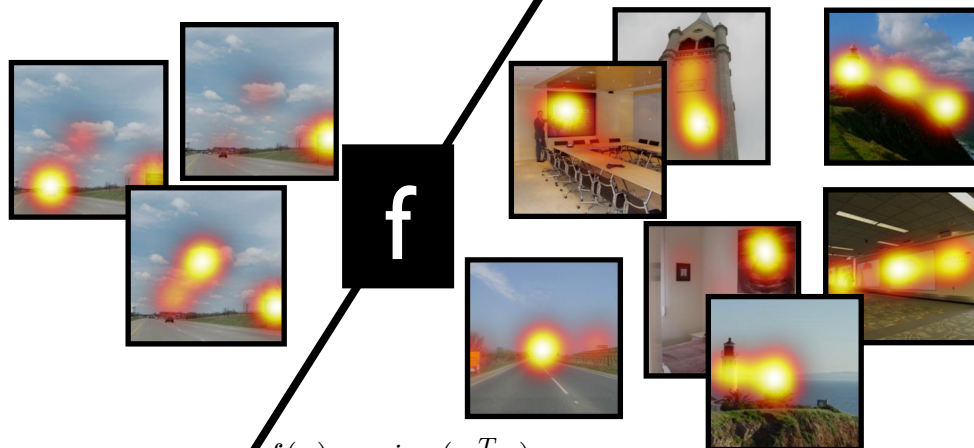
# Classification of fixations

## TRAINING

### 1 coarse binning and smoothing



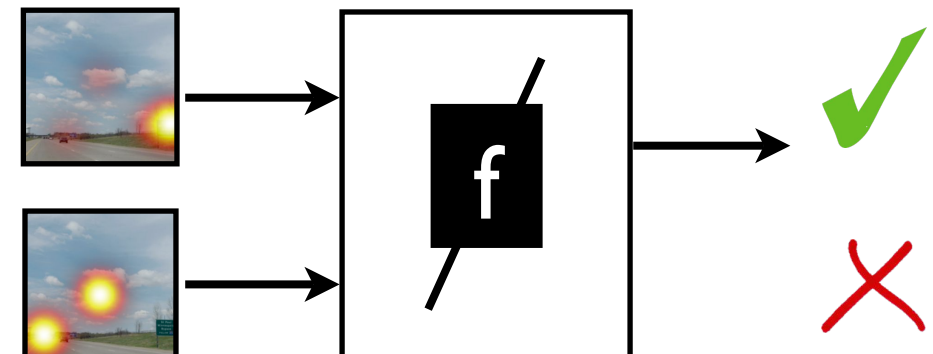
### 2 learn classifier for successful fixations on this image vs successful fixations on other images



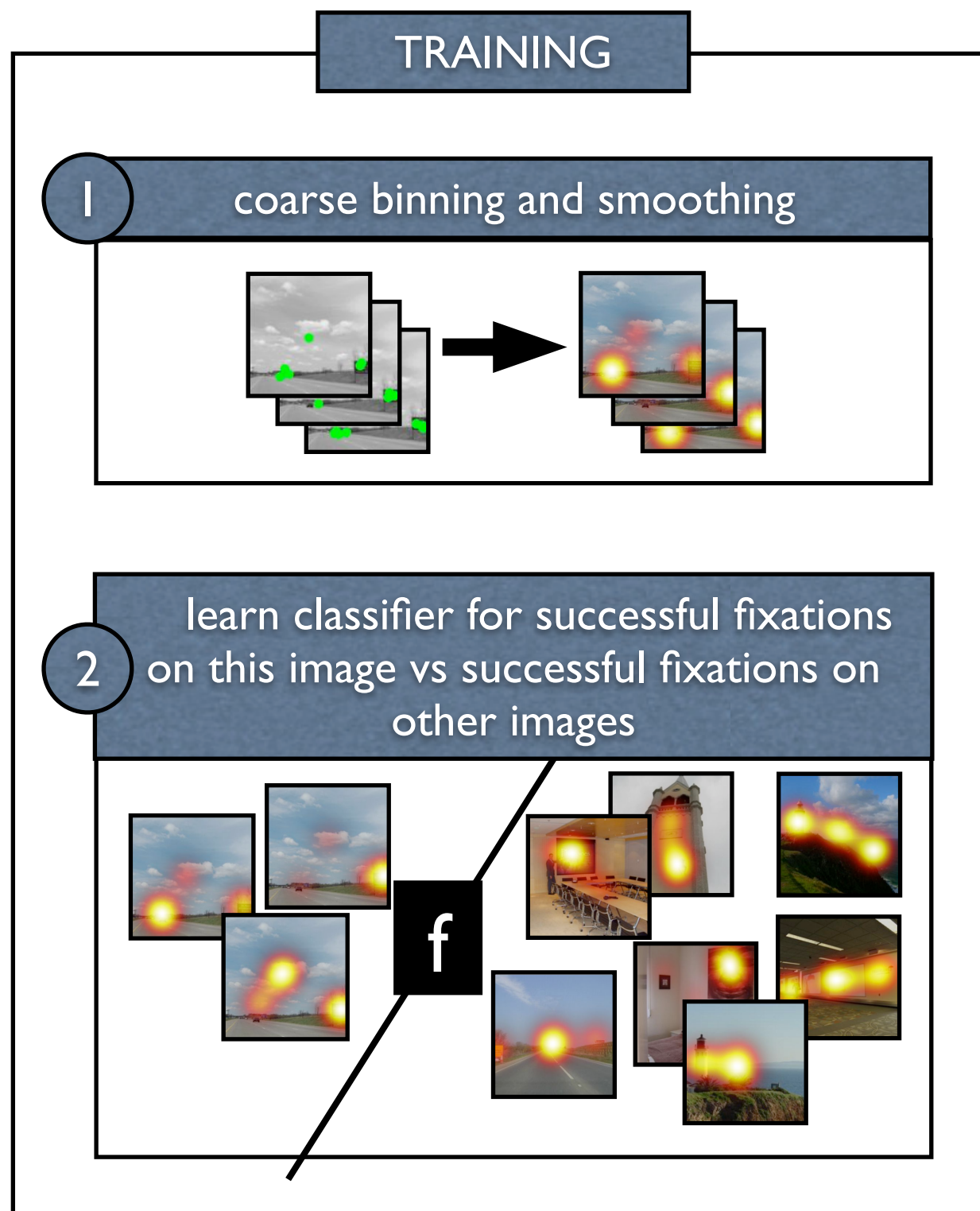
$$f(x) = \text{sign}(w^T x)$$

## TESTING

compute confidence of new fixation map under classifier for image

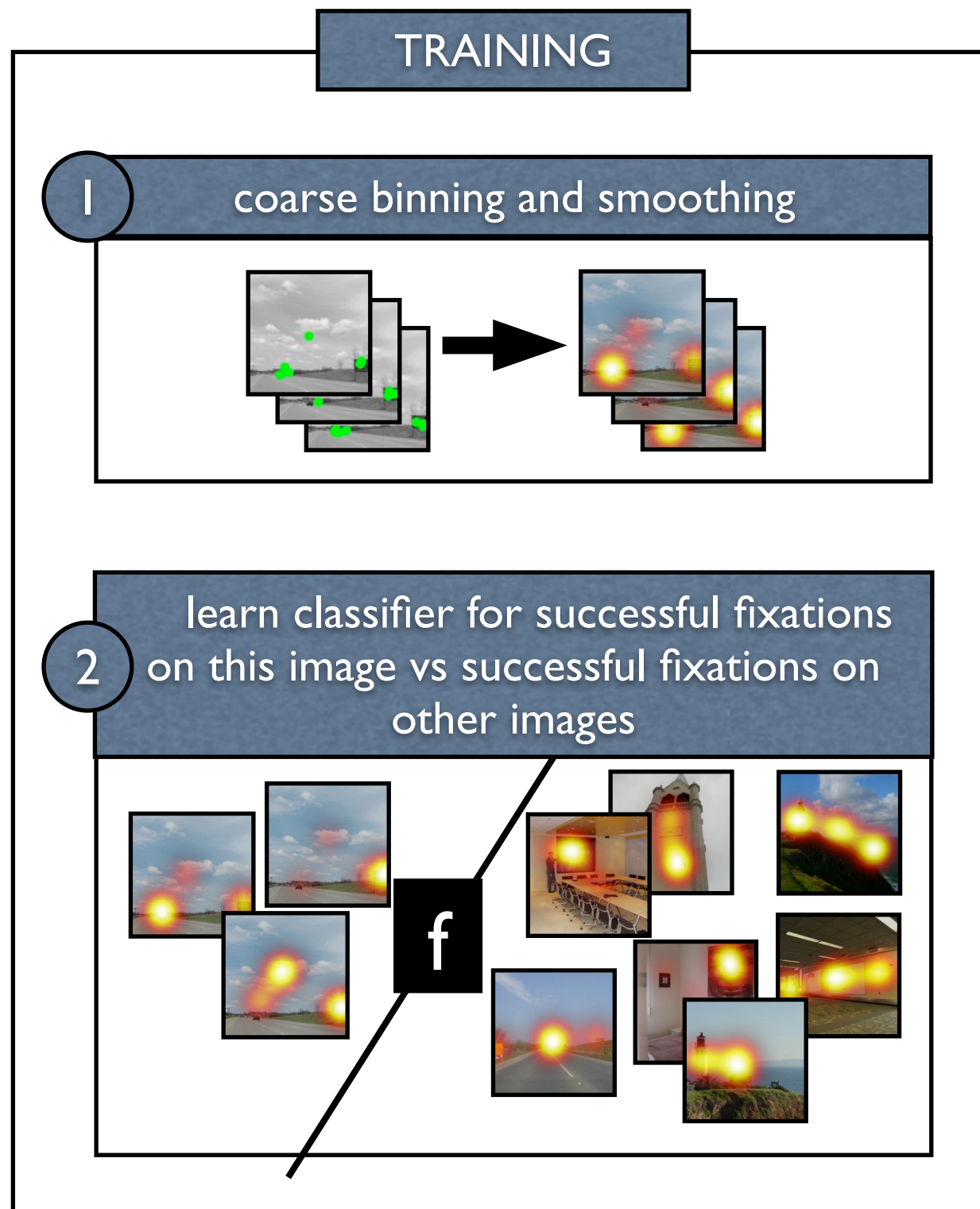


# Classification of fixations



Just the task of classifying whether fixations come from this VS other image achieves **79.7%** balanced accuracy

# Classification of fixations



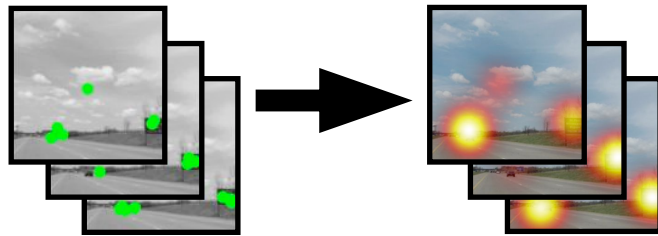
But want to solve a harder problem:  
 successful fixations  
 on this image VS  
 unsuccessful  
 fixations on this  
 image



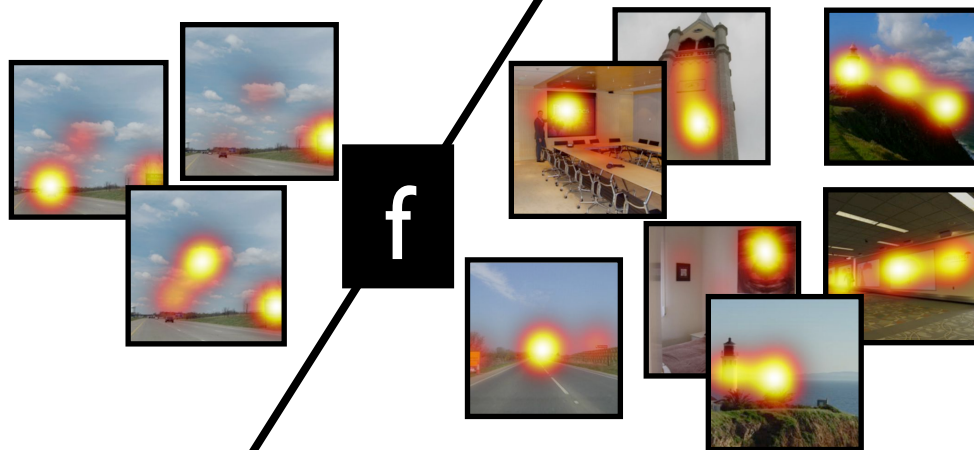
# Classification of fixations

## TRAINING

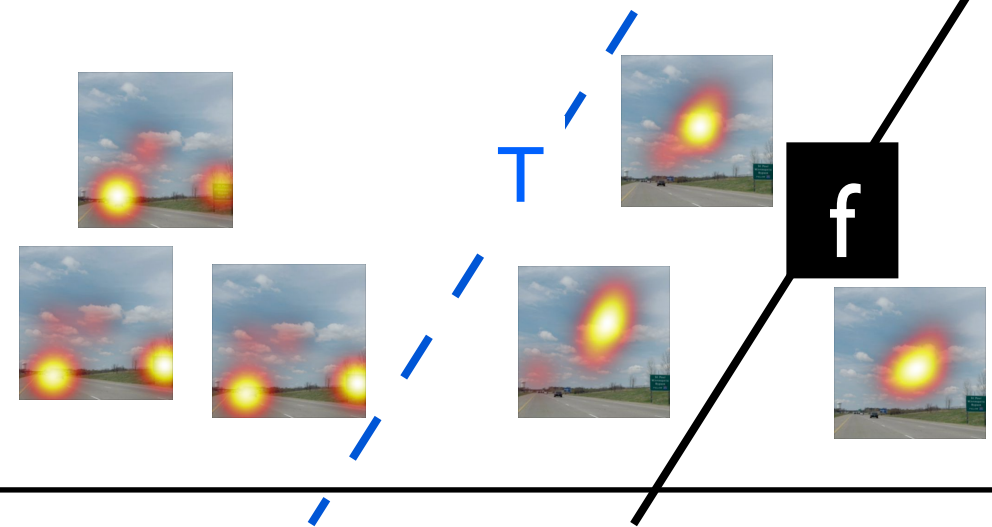
1 coarse binning and smoothing



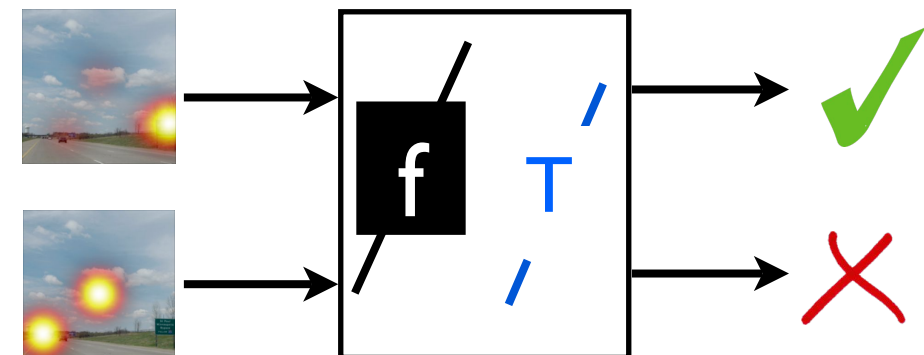
2 learn classifier for successful fixations on this image vs successful fixations on other images



3 find threshold for separating successful vs unsuccessful fixations on this image



## TESTING



## *Model details:*

representation:

- fixation map computed by coarse binning and smoothing (20x20 grid,  $\sigma=2$ )

positives:

- successful encoding fixations on image I

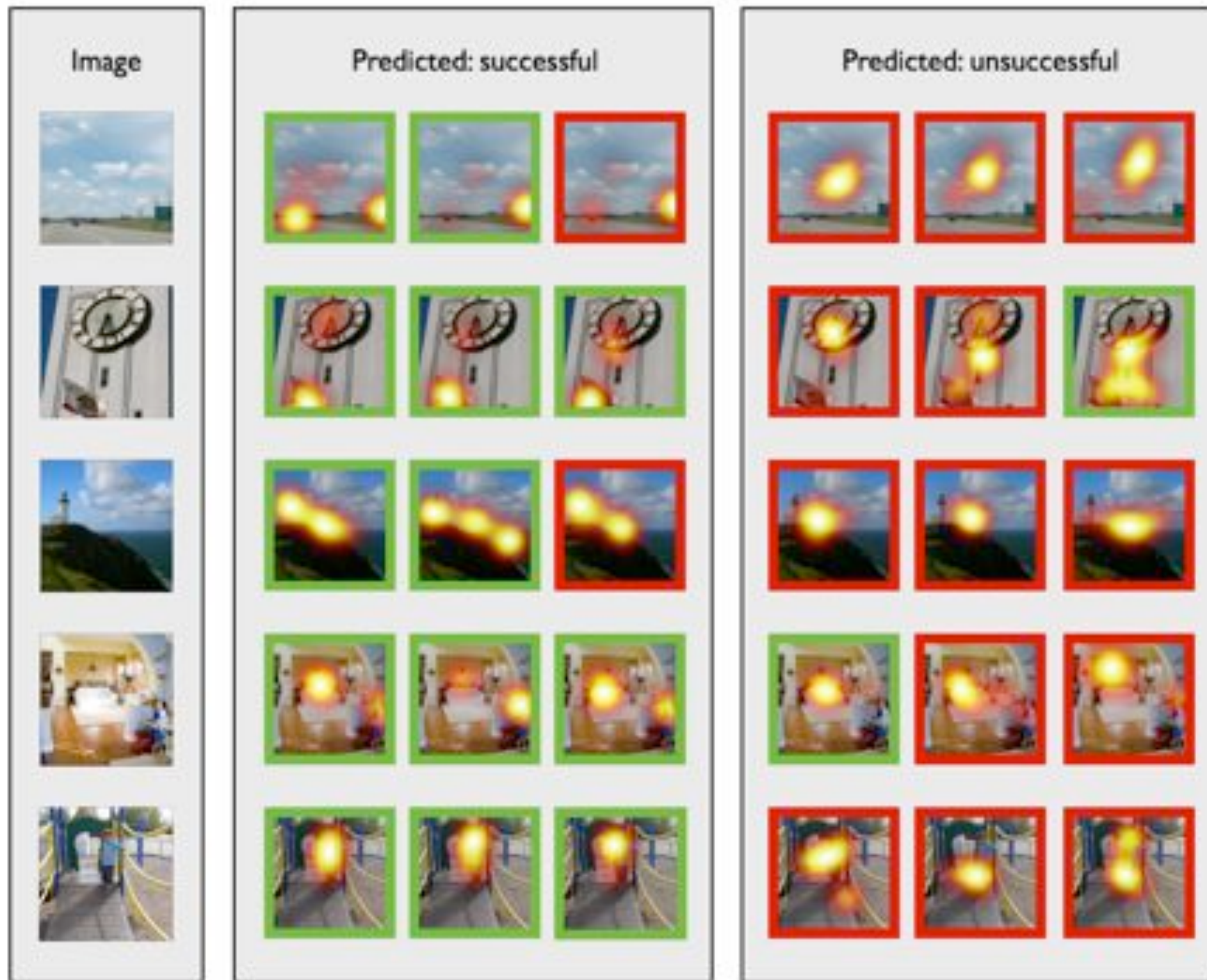
negatives:

- successful encoding fixations on other images

to handle class imbalance:

- classifier: RUSboost
- performance metric: balanced accuracy

# Classifying fixations as successful or unsuccessful



Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. Intrinsic and Extrinsic Effects on Image Memorability, Vision Research, 2015.



## *Model details:*

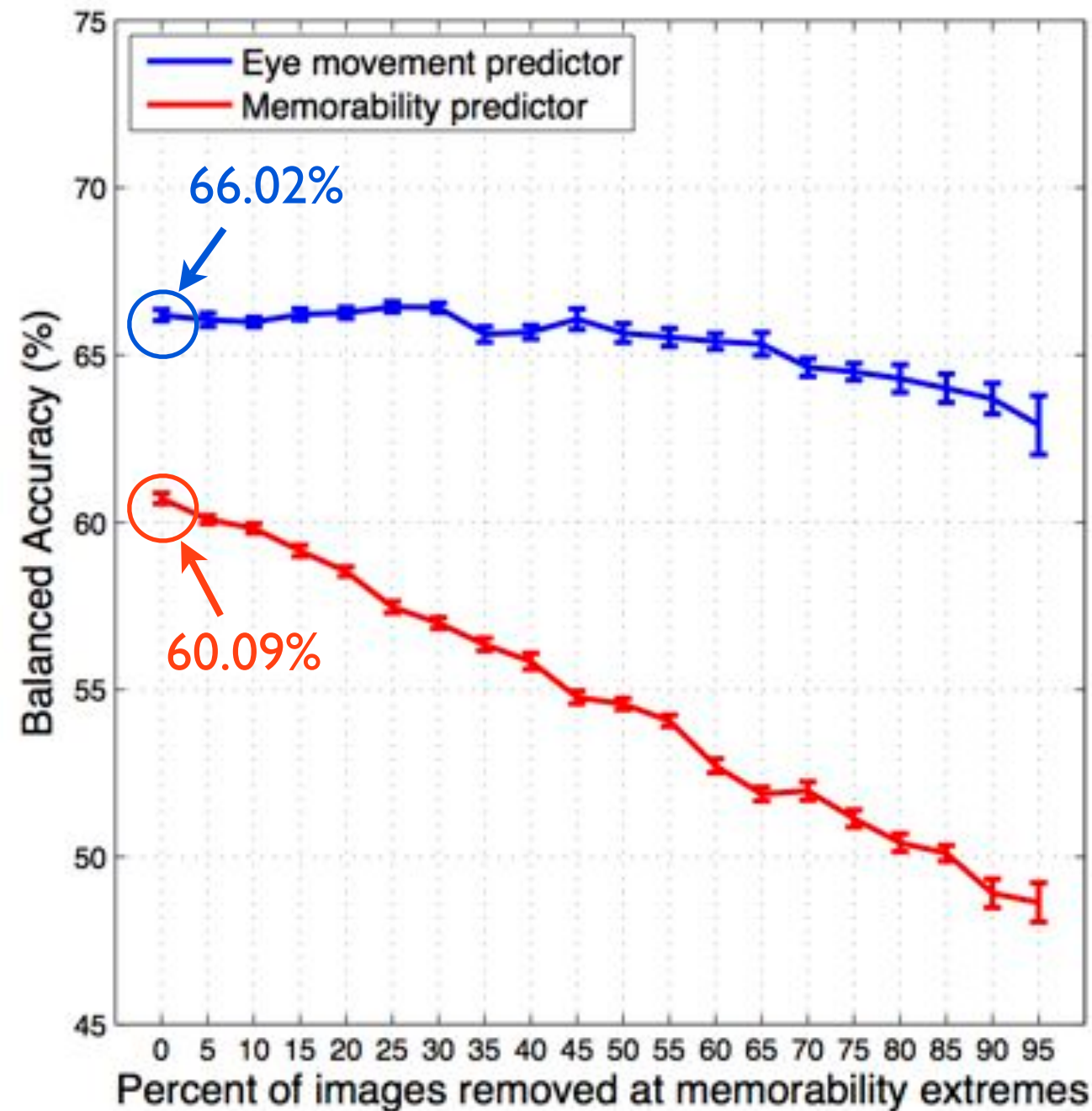
During training (on training participants):

- learn a classifier for each image  $I$  to differentiate fixations on this image VS fixations on other image
- for each participant, evaluate score (of fixation map) under the classifier
- learn a threshold that optimizes balanced accuracy on the task: successful VS failed fixations on this image

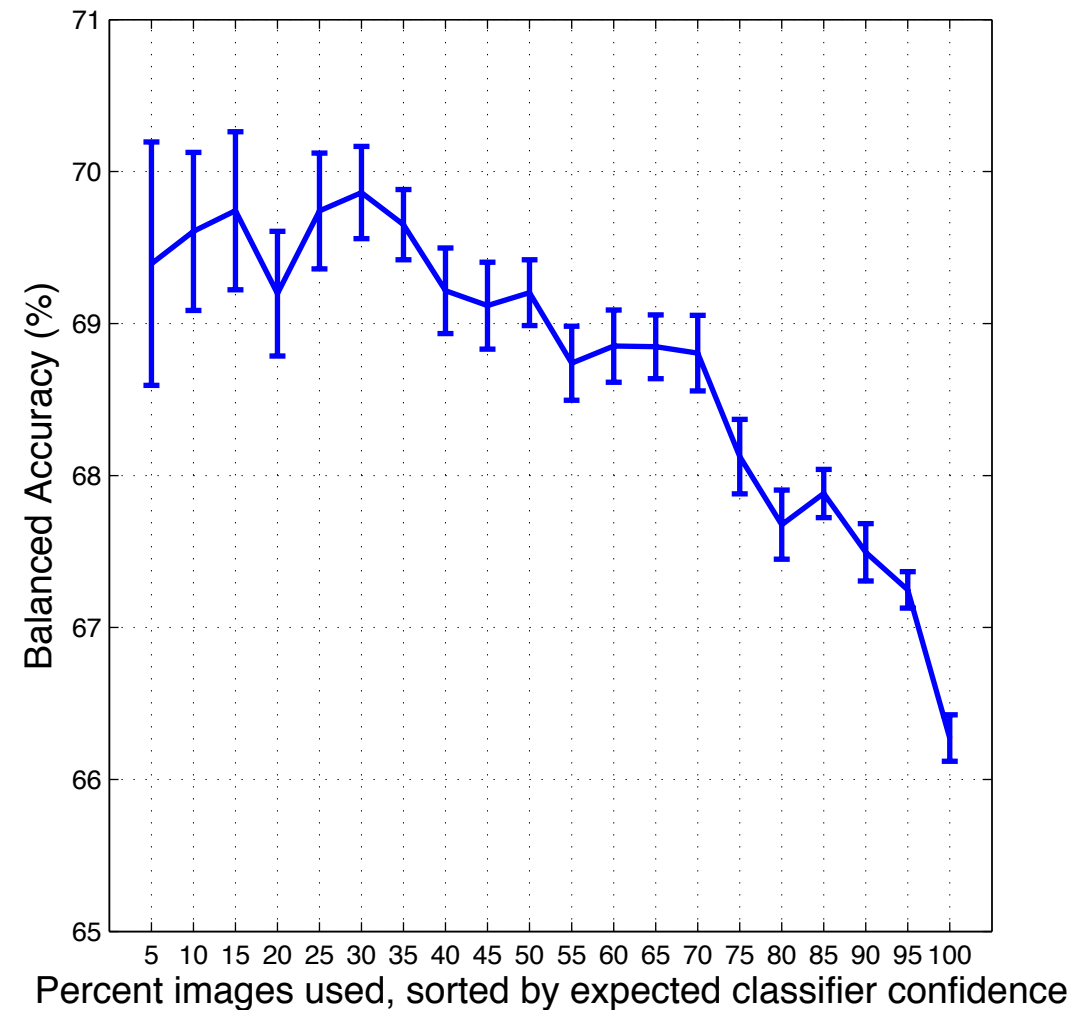
During testing (on novel participants):

- evaluate fixation map on image  $I$  under classifier
- use learned threshold as a cut-off to label fixations as successful or not

# How well can we predict if people will remember an image?



# Not all images are equally predictable



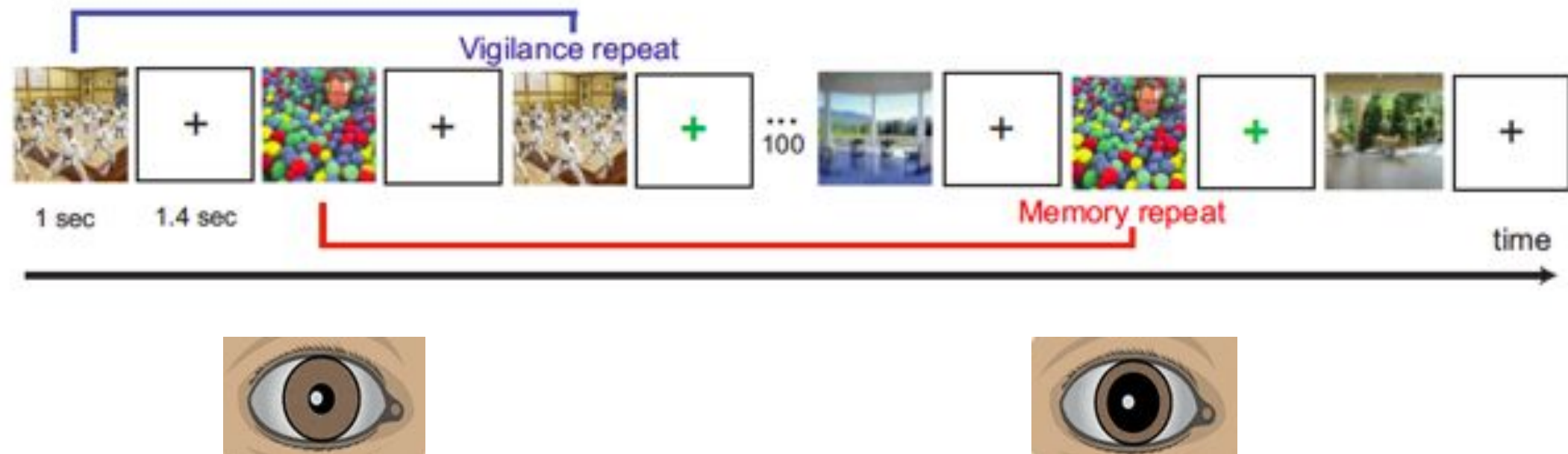
images sorted by the  
average score  
(confidence) of the  
positive training  
examples under the  
classifier

- does the consistency of human memory generalize?
- what factors can modulate image memorability? context, attentional biases (e.g. eye movements), ...
- how do differences in memorability behave over time?



- pupils and blinks have been linked to cognitive load

↑ cognitive effort    ↑ pupil dilations    ↓ blink rates



- **PONE**: difference in pupillary responses b/w hits and correct rejections
- **BONE**: difference in blink rates b/w hits and correct rejections



## Image memorability categories

high



med



low



80 images with highest, middle, and lowest memorability scores from the dataset of Isola et al. 2011

## Image memorability categories

high



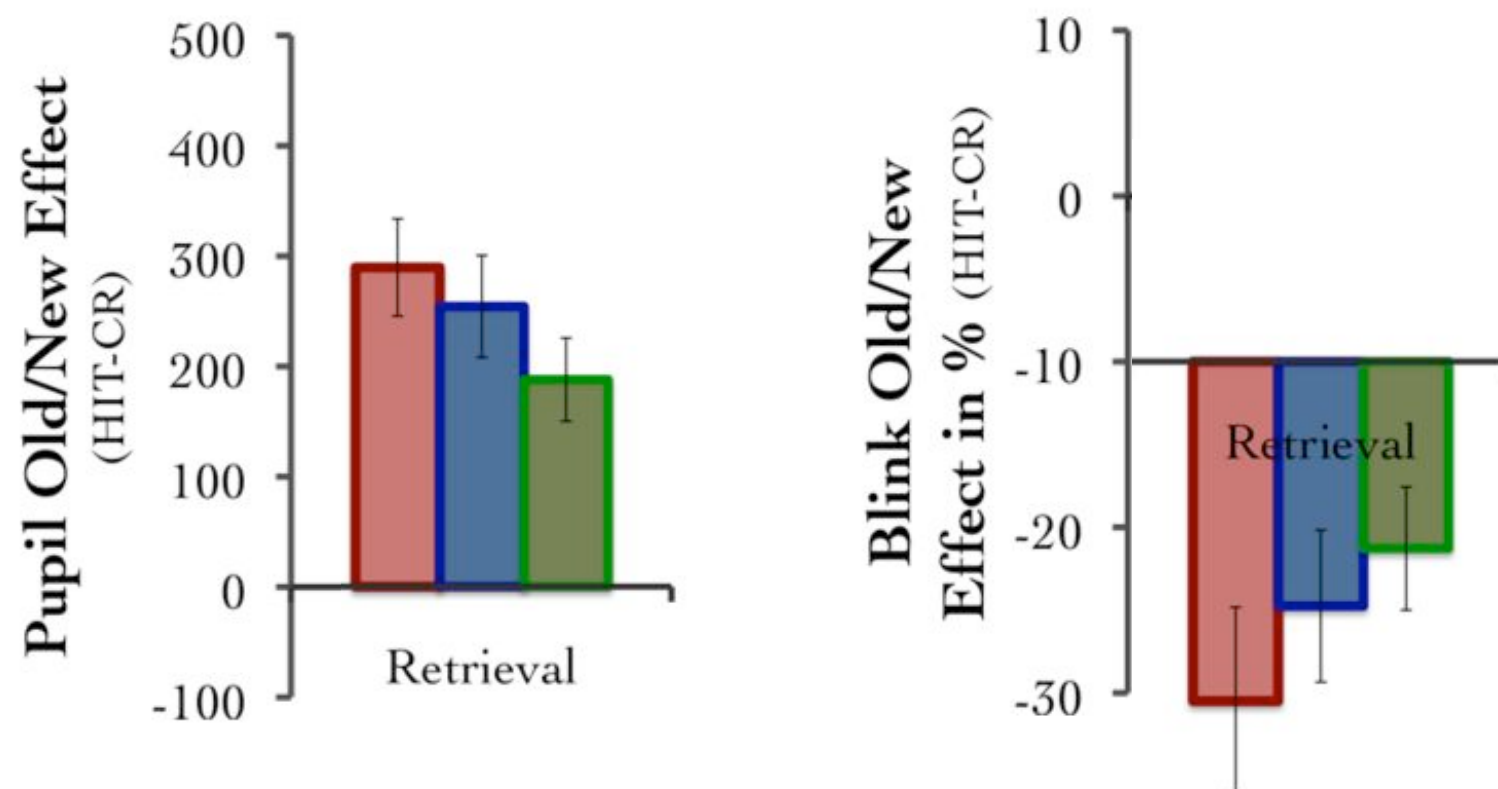
med



low



80 images with highest, middle, and lowest memorability scores from the dataset of Isola et al. 2011



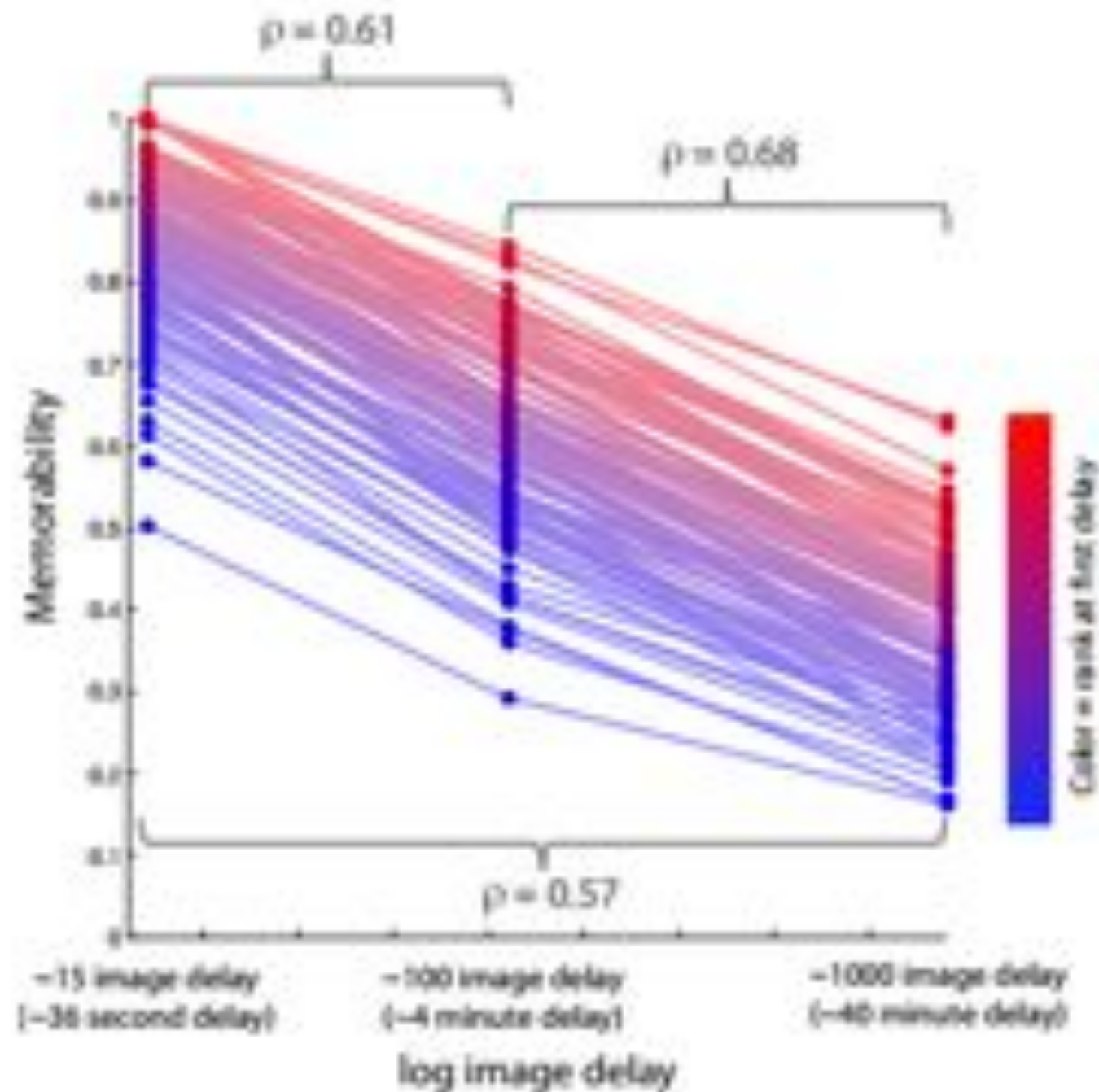
Võ, M., Bylinskii, Z., Oliva, A. Image memorability in the eye of the beholder: tracking the decay of visual scene representations (in prep).

- does the consistency of human memory generalize?
- what factors can modulate image memorability? context, attentional biases (e.g. eye movements, pupils, blinks), ...
- how do differences in memorability behave over time?

- what factors can modulate image memorability?
  - context
  - observer's attention, behavior (measured as: eye movements, pupils, blinks)
  - observer's expertise
  - time for task
  - familiarity
  - utility

- does the consistency of human memory generalize?
- what factors can modulate image memorability?
- **how do differences in memorability behave over time?**

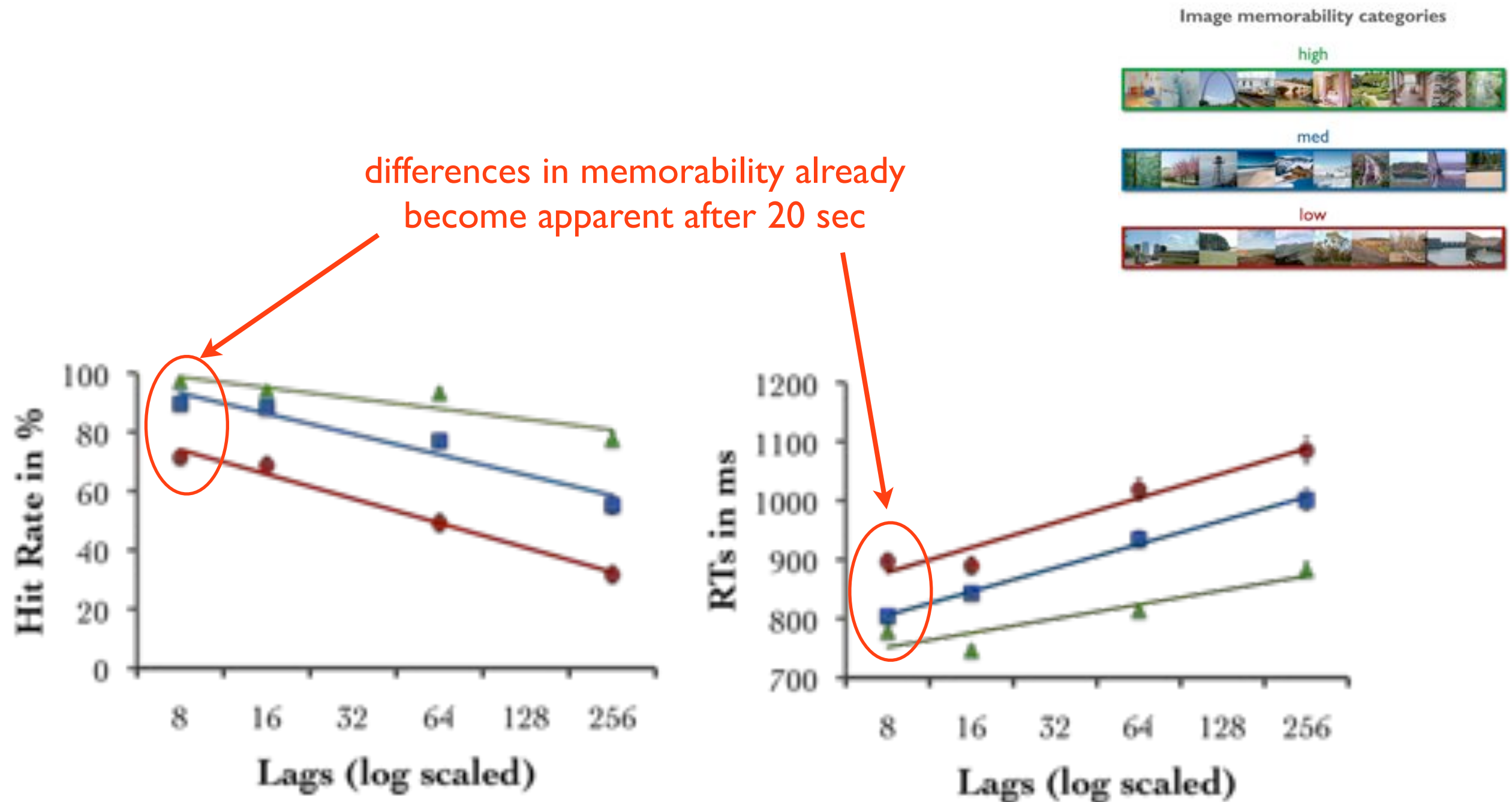
# Ranks of memorability are conserved over time



Isola, P., Xiao, J., Parikh, D., Torralba, A., Oliva, A. What makes a photograph memorable? PAMI, 2013



# Differences in memorability show up early

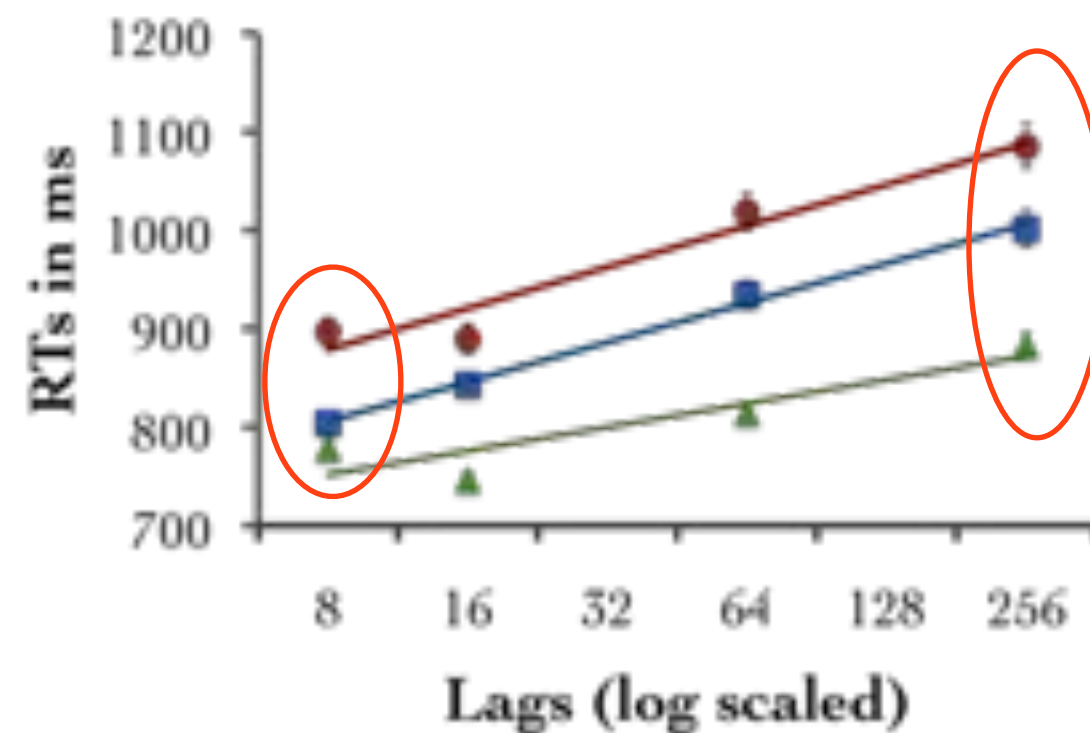
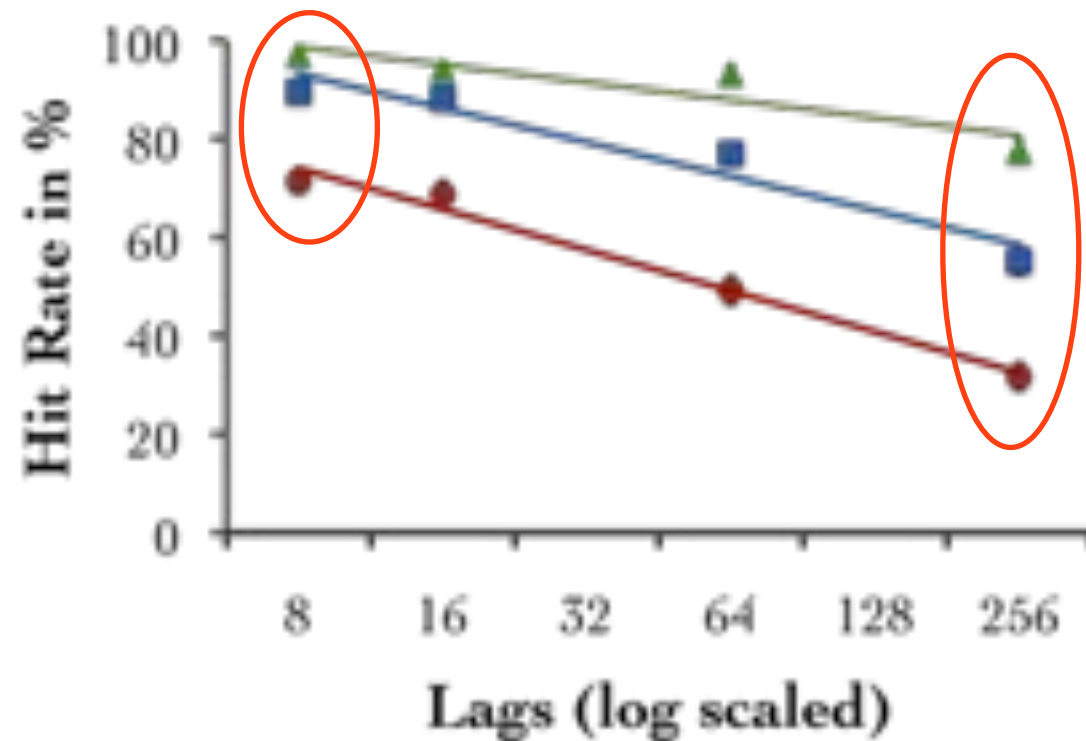


Võ, M., Bylinskii, Z., Oliva, A. Image memorability in the eye of the beholder: tracking the decay of visual scene representations (in prep).

# Differences in memorability increase over time

memorable images stay memorable for longer, forgettable images are forgotten even faster

Image memorability categories



- does the consistency of human memory generalize?
- what factors can modulate image memorability?
- how do differences in memorability behave over time? ranks conserved, differences show up early, and increase over time

- **does the consistency of human memory generalize?**
- what factors can modulate image memorability?
- how do differences in memorability behave over time?

# Faces



Forgettable Faces



Memorable Faces

Bainbridge, W., Isola, P., Oliva, A. The Intrinsic Memorability of Face Photographs. J. Exp Psych: General, 2013



# Visualizations



Distinct visualizations are more memorable.



Borkin, M., Vo, A., Bylinskii, Z., Isola, P., Sunkavalli, S. Oliva, A., Pfister, H. What makes a visualization memorable? InfoVis, 2013



- does the consistency of human memory generalize?  
different scene categories,  
faces, visualizations,...
- what factors can modulate image memorability?
- how do differences in memorability behave over time?

# Questions answered

- **does the consistency of human visual memory generalize?**
  - to fine-grained categories
  - to scenes, faces and visualizations
  - across experiments and populations






# Questions answered

- **does the consistency of human visual memory generalize?**
  - to fine-grained categories
  - to scenes, faces and visualizations
  - across experiments and populations
- **what factors can modulate image memorability?**
  - image context
  - observer behavior and attentional biases



# Questions answered

- **does the consistency of human visual memory generalize?**
  - to fine-grained categories
  - to scenes, faces and visualizations
  - across experiments and populations
- **what factors can modulate image memorability?**
  - image context
  - observer behavior and attentional biases
- **how do differences in memorability behave over time?**
  - memorability ranks are conserved
  - differences in memorability show up early and increase over time

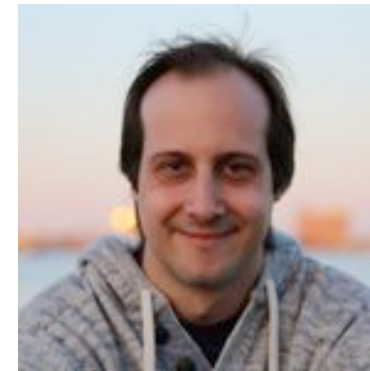
# Questions remaining

- how generalizable is the consistency of human visual memory?
  - other natural and artificial visual stimuli
  - expert domains
- what factors can modulate image memorability?
  - observer's expertise
  - time for task
  - familiarity
  - utility
- how do differences in memorability behave over time?
  - longer time course
  - interference in memory

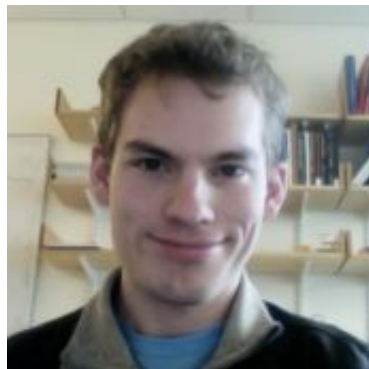
# Thank you



Aude  
Oliva



Antonio  
Torralba



Phillip  
Isola



Michelle  
Borkin



Melissa  
Le-Hoa Võ



Constance  
Bainbridge



Wilma  
Bainbridge



Bolei  
Zhou



# Questions answered

- **does the consistency of human visual memory generalize?**

- to fine-grained categories

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. [Intrinsic and Extrinsic Effects on Image Memorability](#). Vision Research, 2015

- to scenes, faces and visualizations

Bainbridge, W., Isola, P., Oliva, A. [The Intrinsic Memorability of Face Photographs](#). J. Exp Psych: General, 2013

Borkin, M., Vo, A., Bylinskii, Z., Isola, P., Sunkavalli, S. Oliva, A., Pfister, H. [What makes a visualization memorable?](#) InfoVis, 2013

- across experiments and populations

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. [Intrinsic and Extrinsic Effects on Image Memorability](#). Vision Research, 2015

- **what factors can modulate image memorability?**

- image context

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. [Intrinsic and Extrinsic Effects on Image Memorability](#). Vision Research, 2015

- observer behavior and attentional biases

Võ, M., Bylinskii, Z., Oliva, A. [Image memorability in the eye of the beholder: tracking the decay of visual scene representations](#). (in prep).

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., Oliva, A. [Intrinsic and Extrinsic Effects on Image Memorability](#). Vision Research, 2015

- **how do differences in memorability behave over time?**

- memorability ranks are conserved

Isola, P., Xiao, J., Parikh, D., Torralba, A., Oliva, A. (2013) [What makes a photograph memorable?](#) PAMI

- differences in memorability show up early and increase over time

Võ, M., Bylinskii, Z., Oliva, A. [Image memorability in the eye of the beholder: tracking the decay of visual scene representations](#). (in prep).