Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neurobiology 130/230. Harvard College/GSAS 78454

Web site: http://tinyurl.com/visionclass (Class notes, readings, etc)

Location: Biolabs 1075

Time: Mondays 03:30 – 05:30

Dates: Friday 09/04*, Mondays 09/14, 09/21, 09/28, 10/05, 10/19, 10/26, 11/02, 11/09, 11/16, 11/23, 11/30, 12/07*

Lectures:

Faculty: Gabriel Kreiman and invited guests

Contact information:

Gabriel Kreiman

gabriel.kreiman@tch.harvard.edu

617-919-2530

Office Hours: After Class. Mon 05:30-06:30

Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neurobiology 230. Harvard College/GSAS 78454

Class 1. Sep-04	Introduction to pattern recognition. Why is vision difficult?
Class 2. Sep-14	Visual input. Natural image statistics. The retina.
Class 3. Sep-21	**Psychophysics of visual object recognition [Ken Nakayama]**
Class 4. Sep-28	Lesion studies in animal models. Neurological studies of cortical visual deficits in humans.
Class 5. Oct-05	Introduction to the thalamus and primary visual cortex [Camille Gomez-Laberge]
Oct-12	*Columbus Day. No class.*
Class 6. Oct-19	Adventures into *terra incognita*. Neurophysiology beyond V1 [Hanlin Tang]
Class 7. Oct-26	First steps into inferior temporal cortex [Carlos Ponce]
Class 8. Nov-02	From the highest echelons of visual processing to cognition [Leyla Isik]
Class 9. Nov-09	Correlation and causality. Electrical stimulation in visual cortex.
Class 10. Nov-16	Theoretical neuroscience. Computational models of neurons and neural networks.
Class 11. Nov-23	Computer vision. Towards artificial intelligence systems for cognition [Bill Lotter]
Class 12. Nov-30	Computational models of visual object recognition.
Class 13. Dec-07	[Extra class] Towards understanding subjective visual perception. Visual consciousness.

Psychophysics: The study of the dependencies of psychological experiences upon the physical stimuli that generate them

Basic measures:

- **Reaction time** — The time taken by subjects to perform a task or make a judgment can give an indication (or at least an upper bound) of how long the necessary psychological (and hence neural) processing takes.

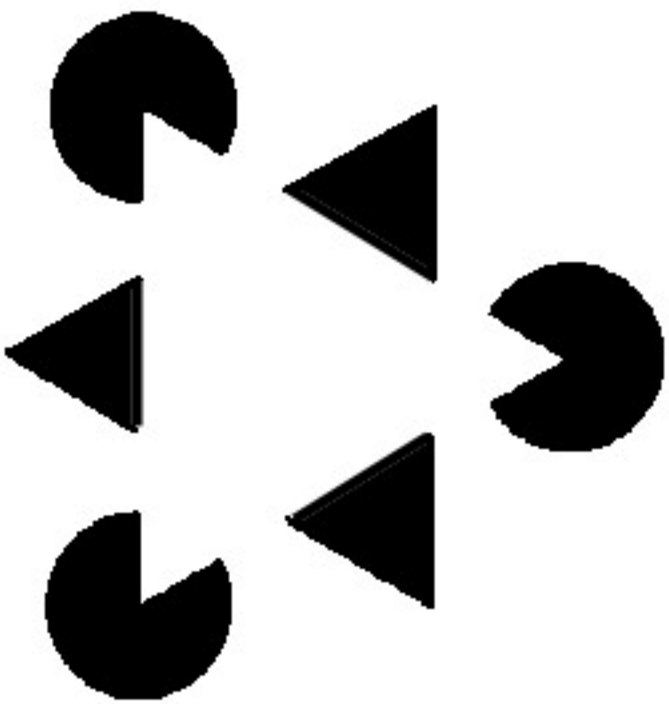
- **Performance** — Often inversely related to reaction time. There are techniques for mitigating response biases.

- **Threshold** — Stimuli can be varied to determine the threshold for detection, discrimination, or some more complex psychological phenomenon.

Gestalt laws of grouping
Basic phenomenological constraints

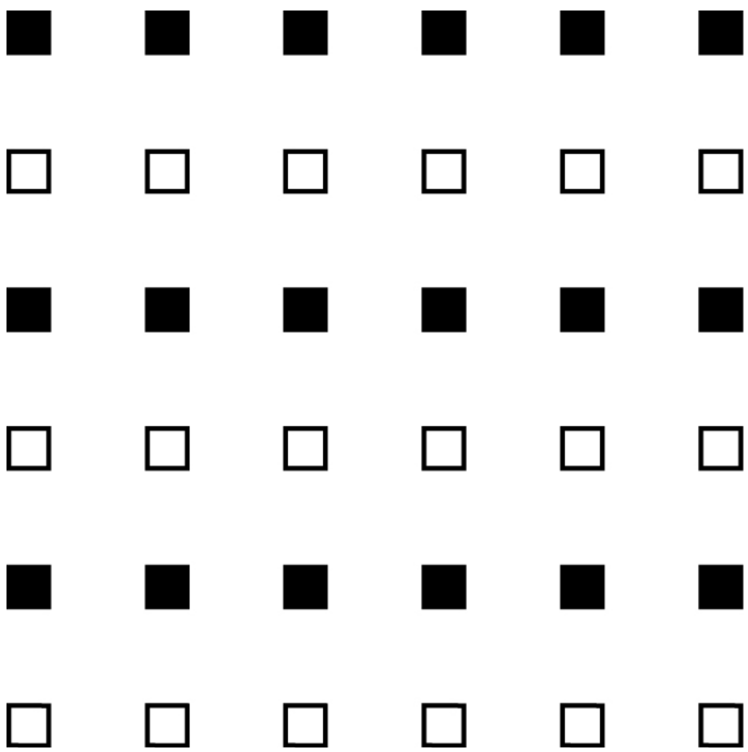
- Law of Closure — The mind may experience elements it does not perceive through sensation, in order to complete a regular figure (that is, to increase regularity).
- Law of Similarity — The mind groups similar elements into collective entities or totalities. This similarity might depend on relationships of form, color, size, or brightness.
- Law of Proximity — Spatial or temporal proximity of elements may induce the mind to perceive a collective or totality.
- Law of Symmetry (Figure ground relationships)— Symmetrical images are perceived collectively, even in spite of distance.
- Law of Continuity — The mind continues visual, auditory, and kinetic patterns.
- Law of Common Fate — Elements with the same moving direction are perceived as a collective or unit.

Law of closure: perceiving objects as whole even if they are not complete



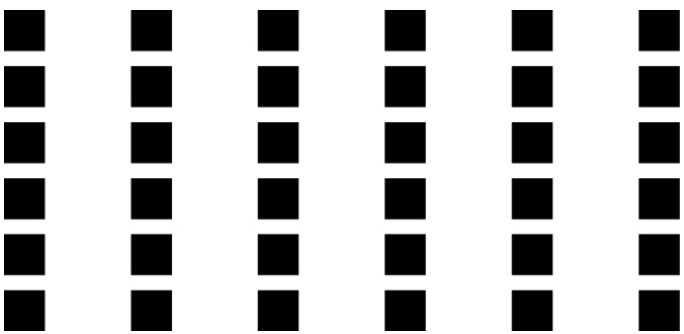
The mind may experience elements it does not perceive through sensation, in order to complete a regular figure (that is, to increase regularity)

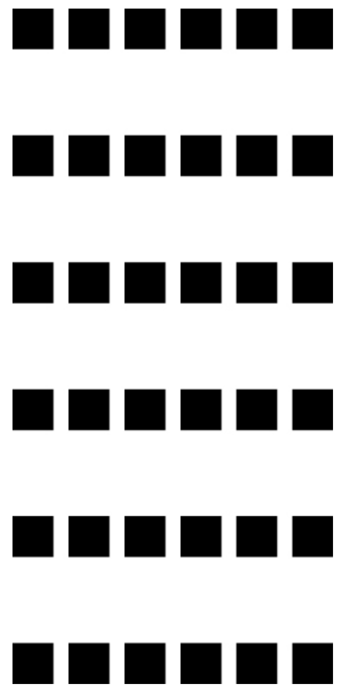
Law of similarity



The mind groups similar elements into collective entities or totalities. This similarity might depend on relationships of form, color, size, or brightness

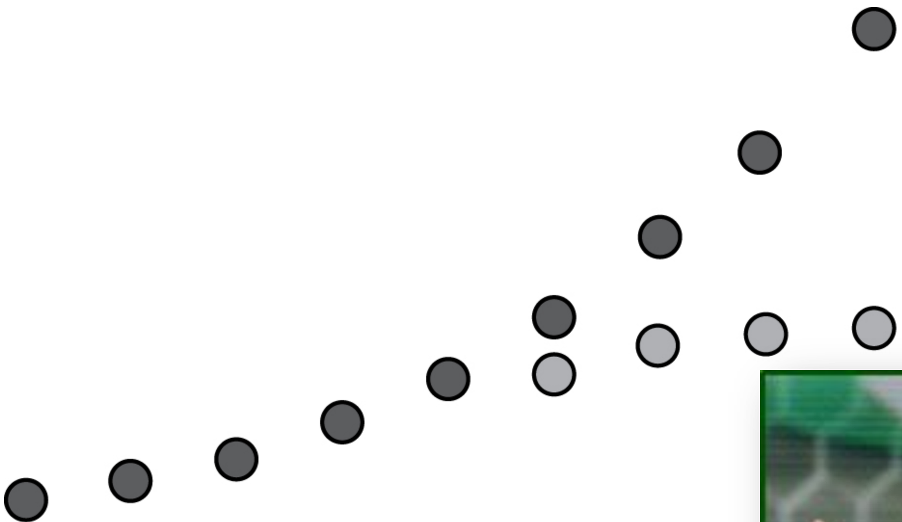
Law of proximity





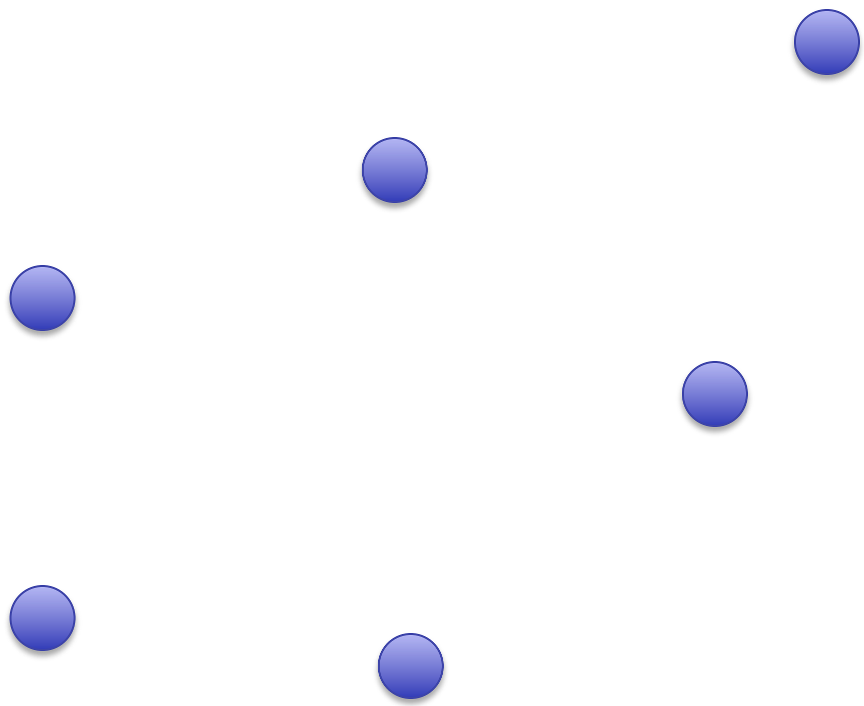
•Spatial or temporal proximity of elements may induce the mind to perceive a collective or totality.

Law of continuity



The mind continues visual, auditory, and kinetic patterns

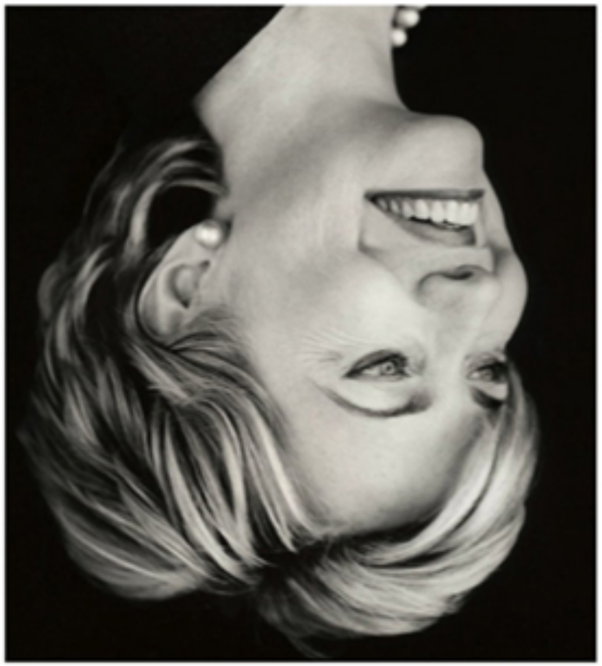
Law of common fate



Holistic representation of faces

A Thatcher illusion

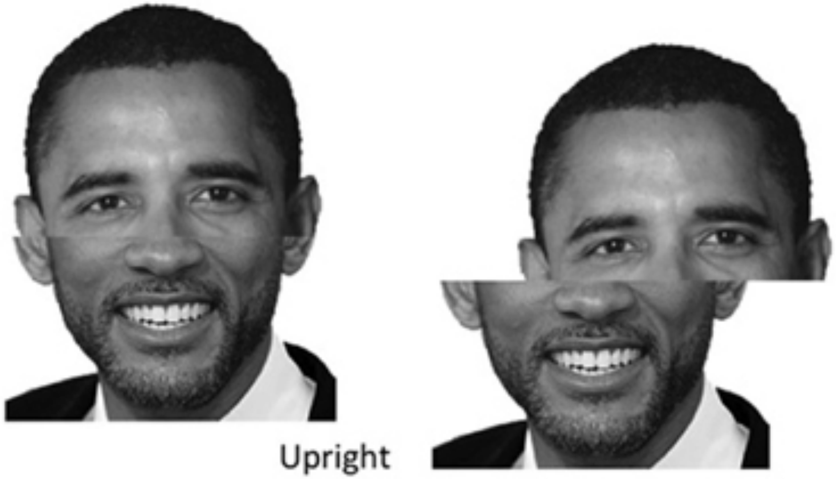
Inverted



McKone et al, Frontiers in Psychology, 2013

Holistic representation of faces

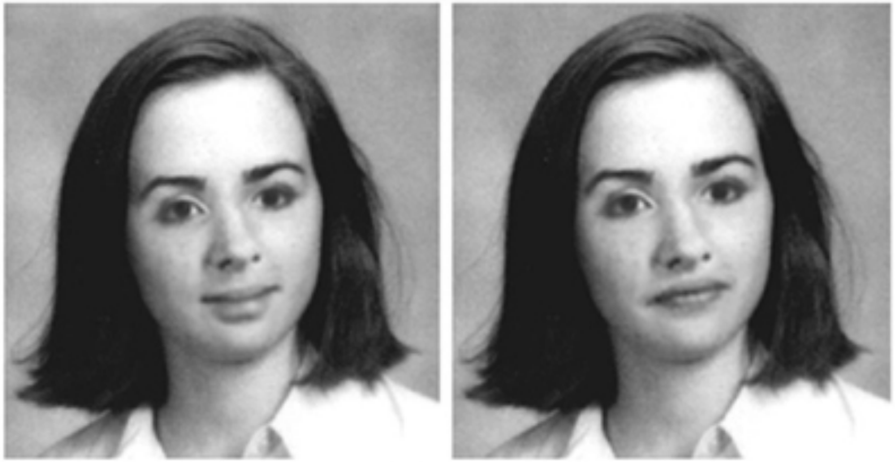
Composite illusion



McKone et al, Frontiers in Psychology, 2013

Holistic representation of faces

C **Part-whole illusion**



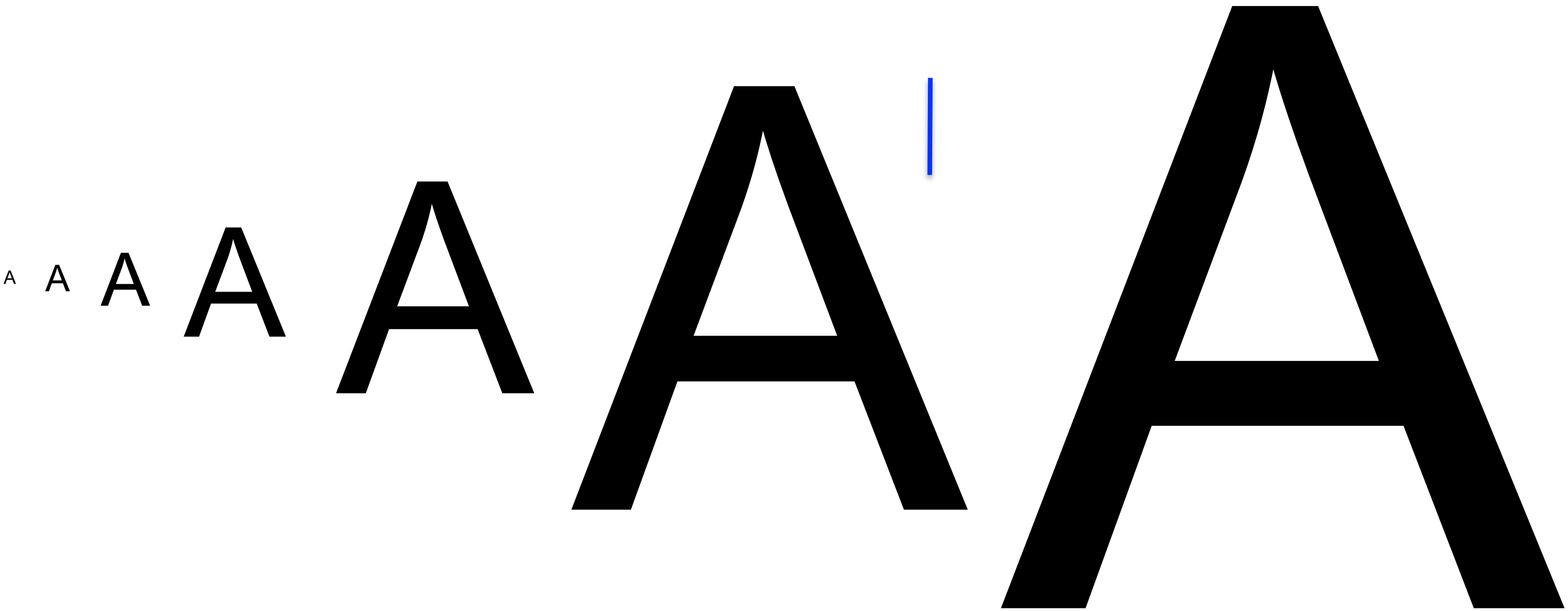
Upright

McKone et al, Frontiers in Psychology, 2013

Tolerance to image transformations

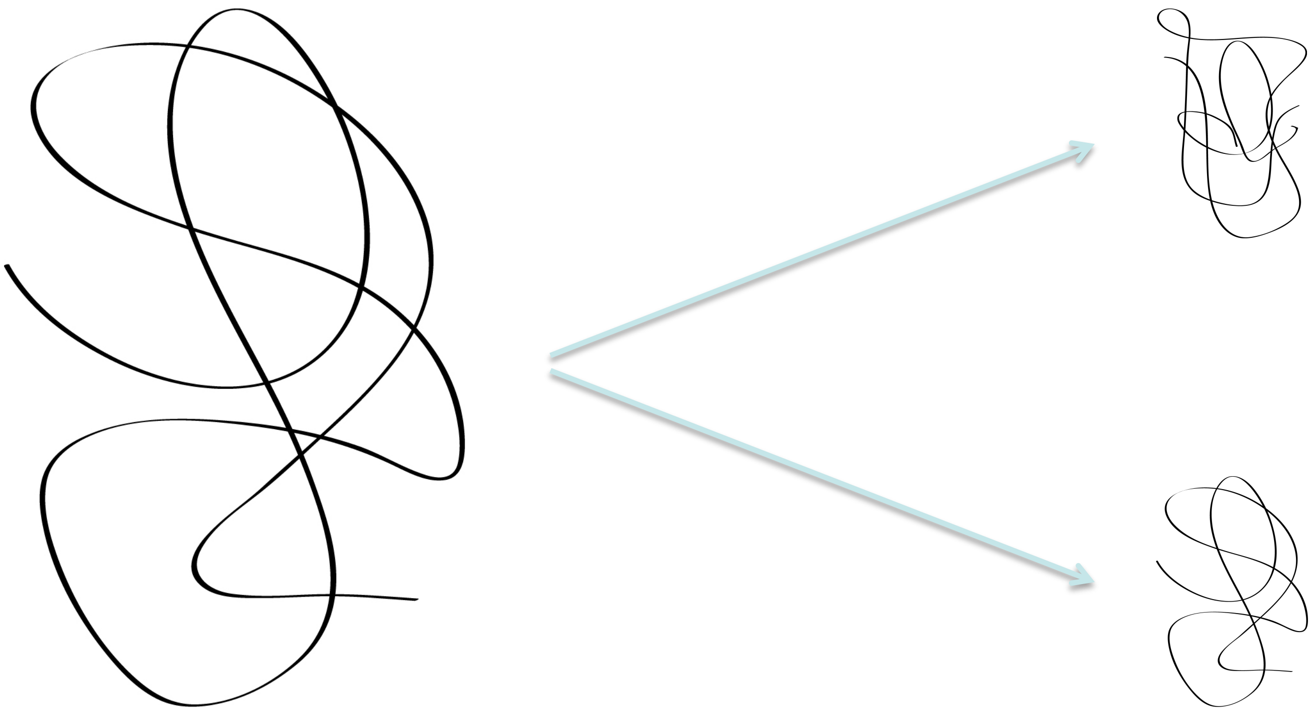
Scale
Position
Rotation (2D)
Rotation (3D) – viewpoint
Color
Illumination
Cues
Clutter
Occlusion
Other non-rigid transformations (aging, expressions, etc)

Scale tolerance



One-shot learning for scale tolerance

Which one is it?

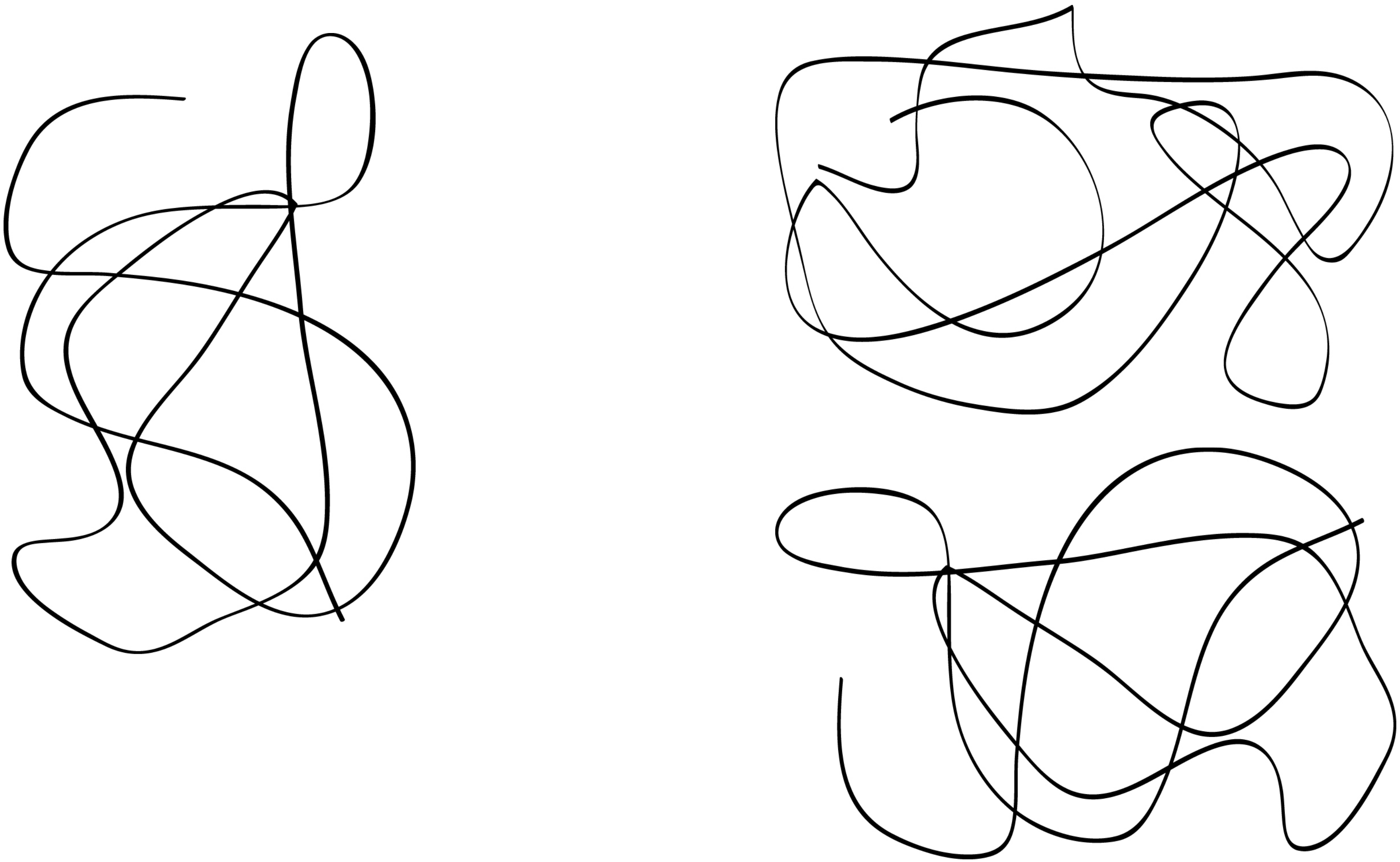


Position tolerance

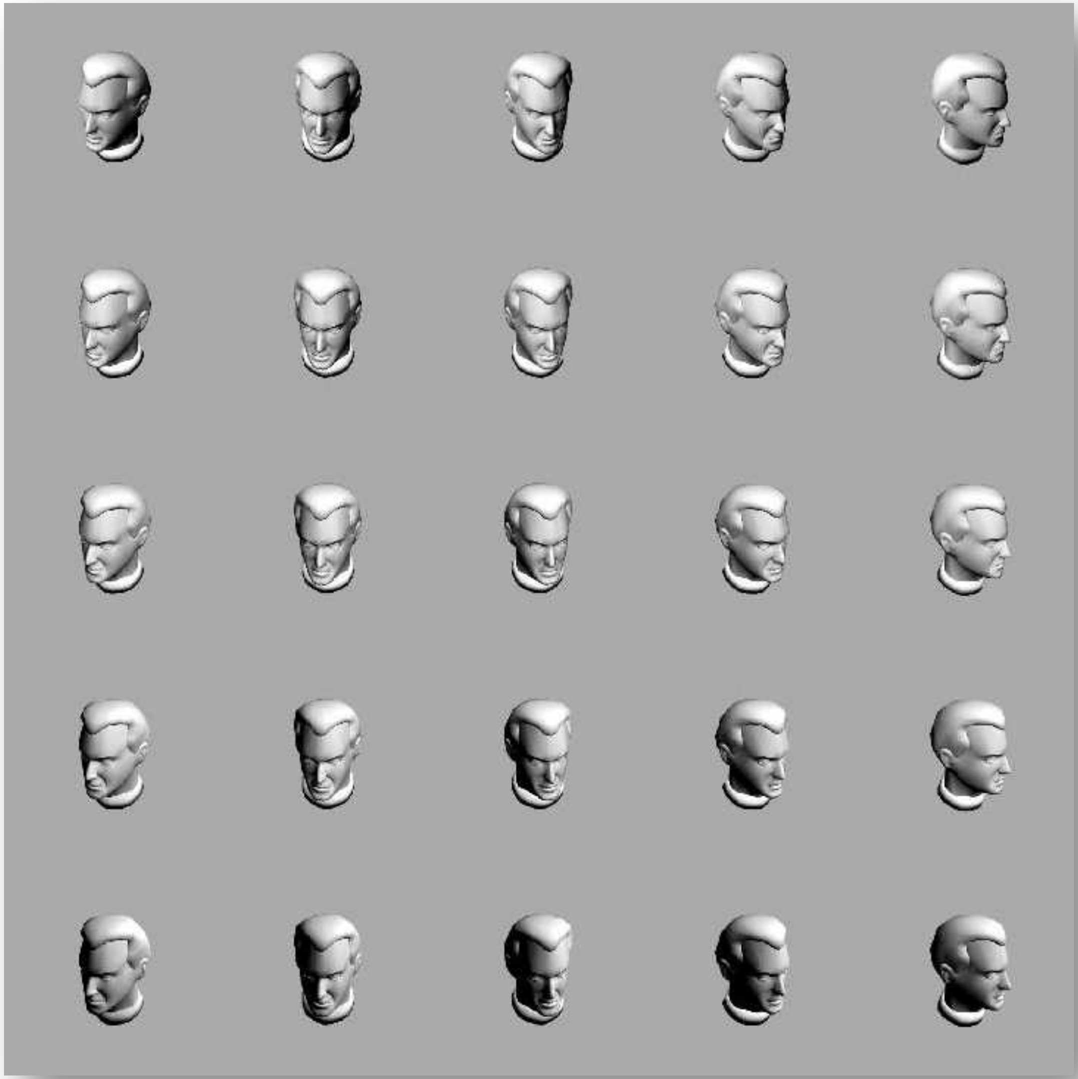
x

bd db bd bd bd db db bd db db bd

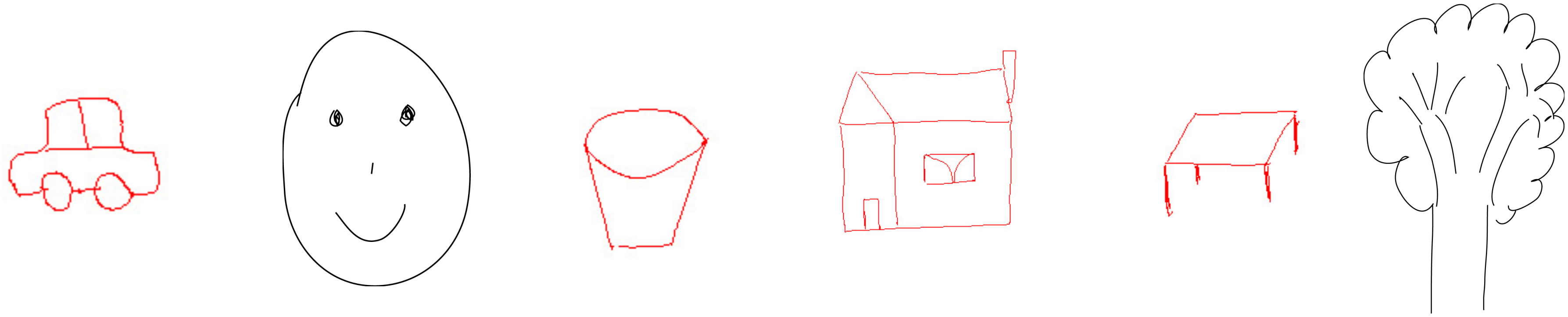
Other transformations require example-based training



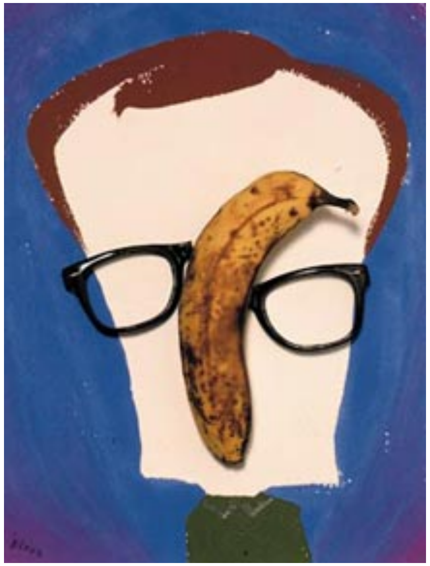
Tolerance to viewpoint and illumination changes

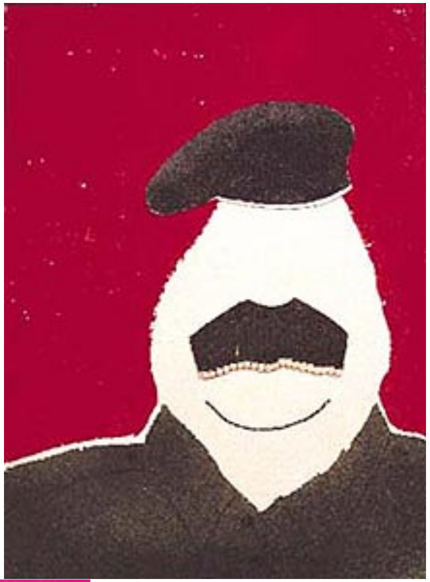


Recognition from minimal features



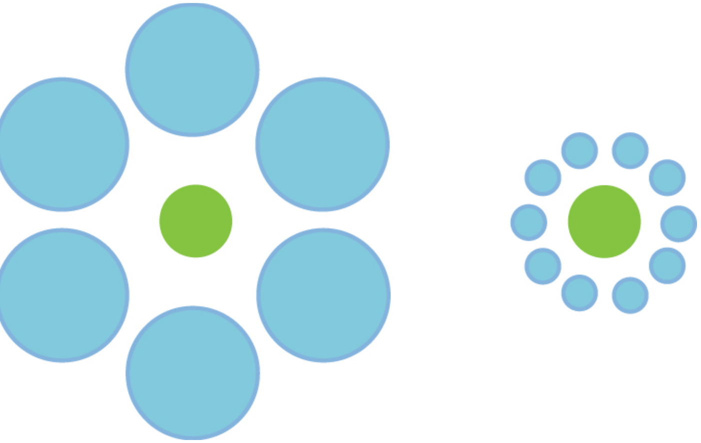
Recognition of caricatures





Images:
Hanoch Piven

Beyond pixels – Context matters



Visual recognition depends on experience

Recognition of images flashed for ~100 ms (demo)

Visual recognition can be extremely fast

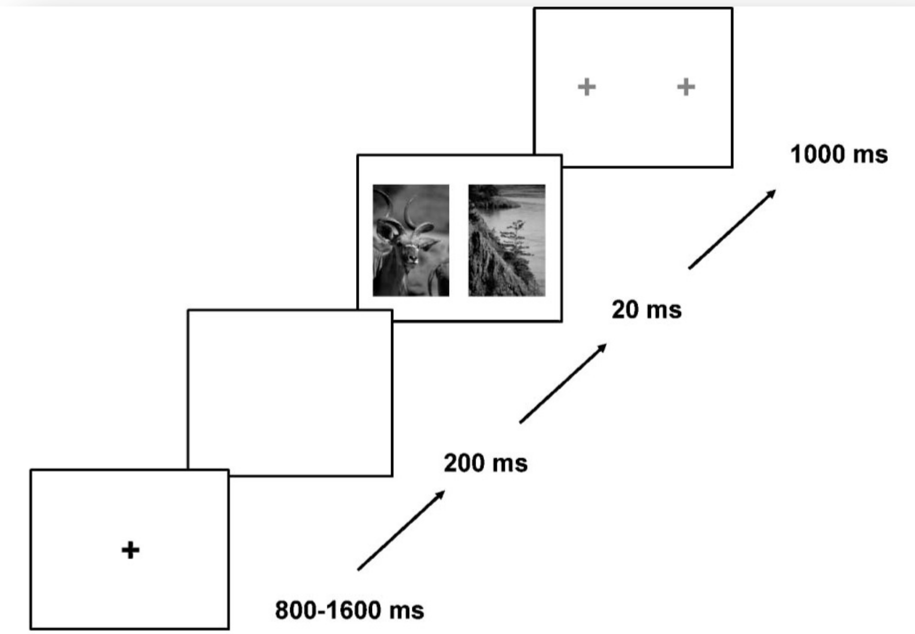


Fig. 1. Choice saccade task. After a pseudo-random fixation period, a blank screen (gap period) for 200 ms preceded the simultaneous presentation of two natural scenes in the left and right hemifields (20 ms). The images were followed by two grey fixation crosses indicating the saccade landing positions.

Table 1
Summary of behavioural results. Participant numbers correspond to those in Fig. 4

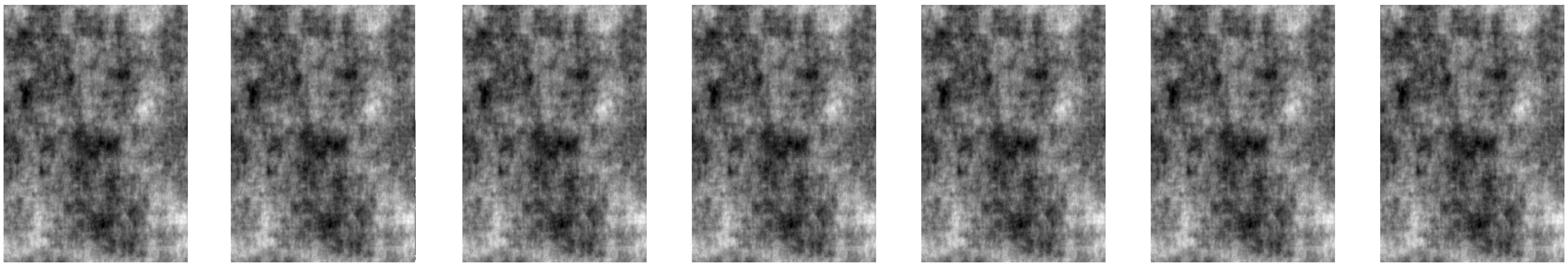
Subject	*N*	Accuracy (%)	Median RT (ms)	Min RT (ms)	Mean start (ms)
1	682	96.3	227	130	143
2	774	93.3	200	130	136
3	726	81.8	201	130	129
4	563	80.1	191	120	126
5	672	86.6	159	130	133
6	675	86.1	224	150	143
7	574	90.2	204	140	129
8	653	94.0	213	150	147
9	694	96.7	251	180	200
10	534	89.7	236	180	124
11	739	90.0	253	190	205
12	652	96.6	276	200	235
13	703	95.0	238	160	173
14	769	98.7	301	230	251
15	529	77.1	233	160	235
All	8998	90.1	228	120	140

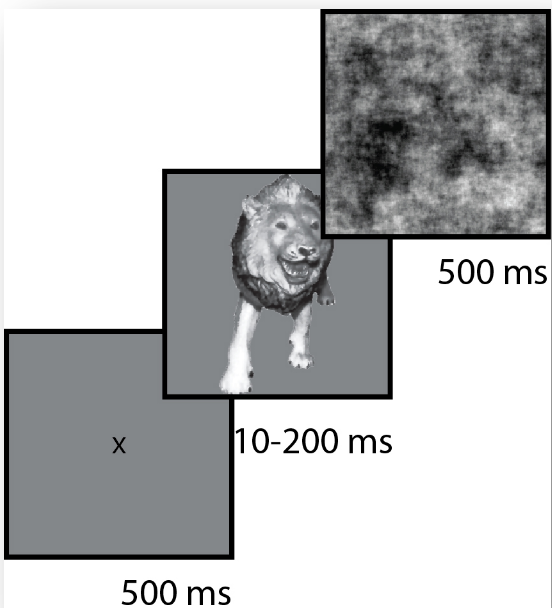
The second column of this table indicates the total number of trials per participant (see Section 2 for details). Columns 3–5 give the mean accuracy, median and minimum reaction time values for each participant shown in Figs. 3B and C. The last column indicates the onset latency of the mean eye trace for each participant (see Fig. 5).

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. Vision Res, 46(11), 1762-1776.

Backward masking

10 ms 20 ms 30 ms 40 ms 50 ms 100 ms 200 ms





Doubles?



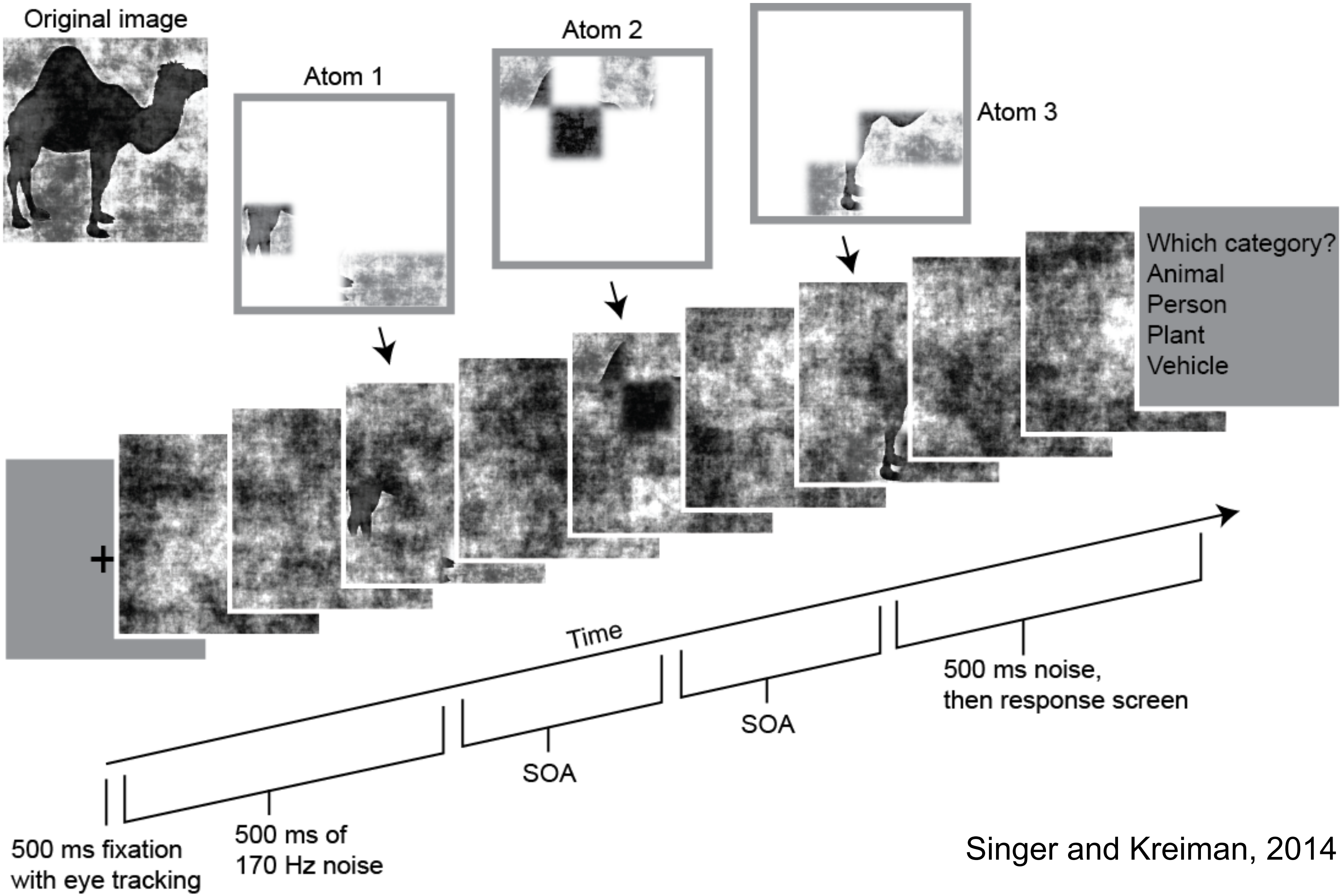




http://www.francoisbrunelle.com/

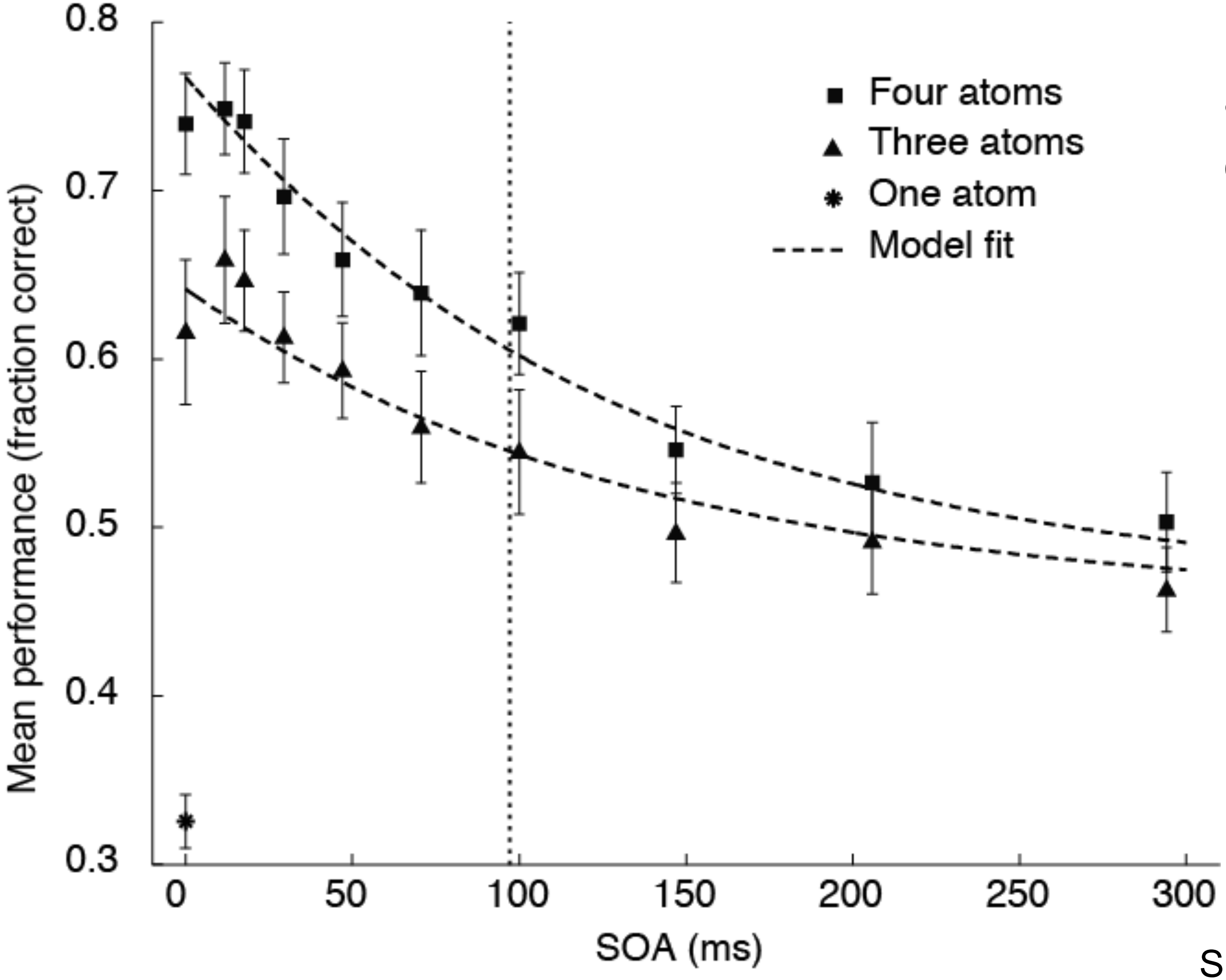
Francois Brunelle

Is information integrated over time?



Singer and Kreiman, 2014

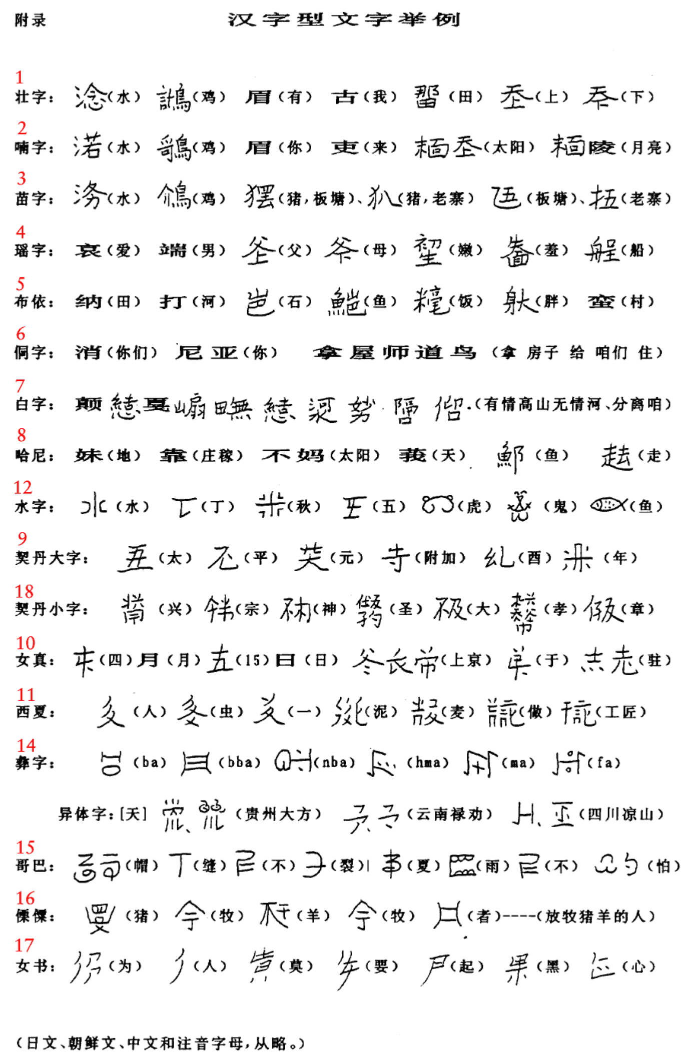
Rapid decay in recognition of asynchronously presented object parts



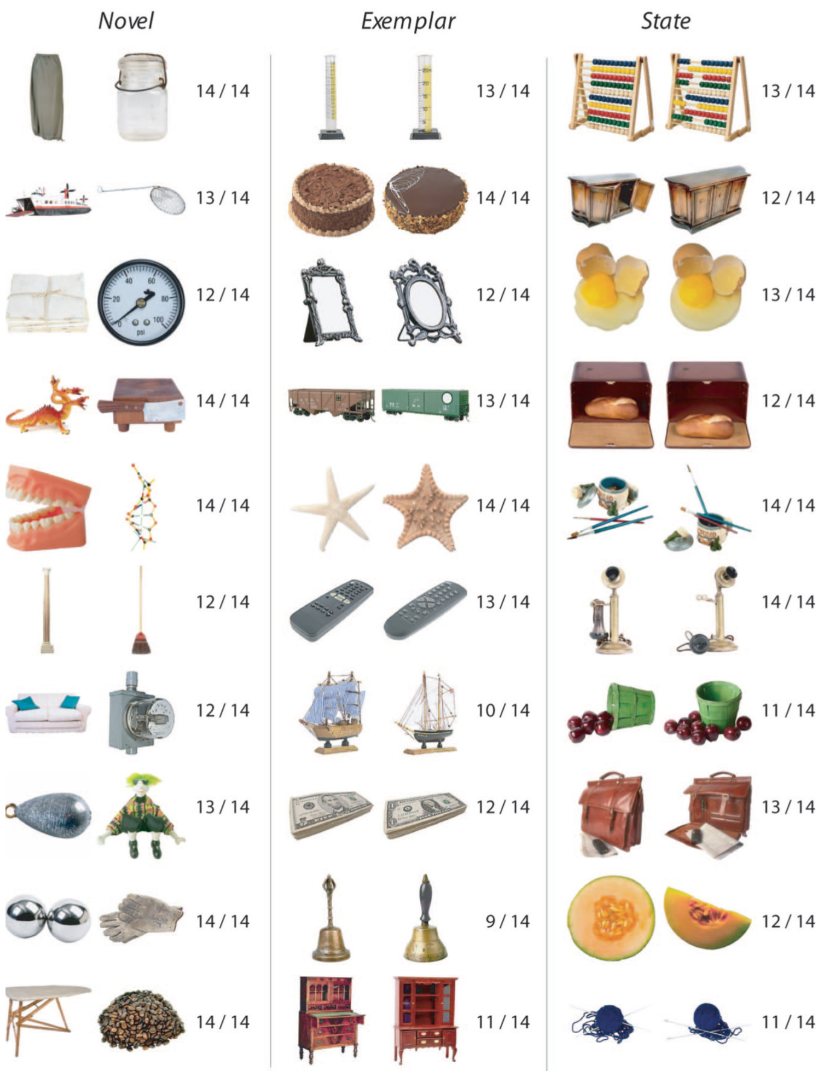
Brief asynchronies disrupt object recognition

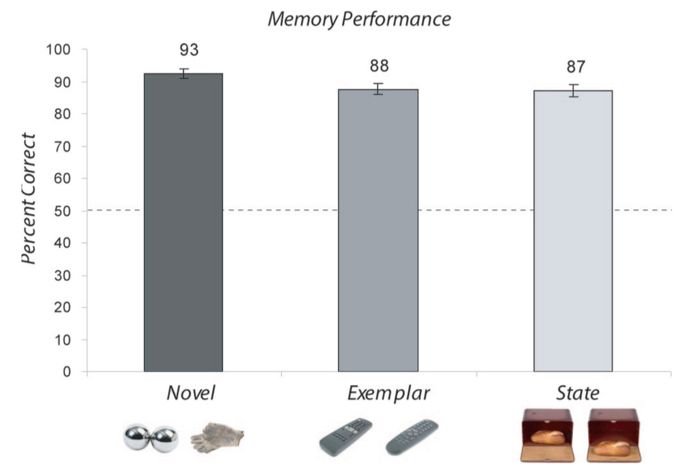
Singer and Kreiman, 2014

The visual system has a very large capacity



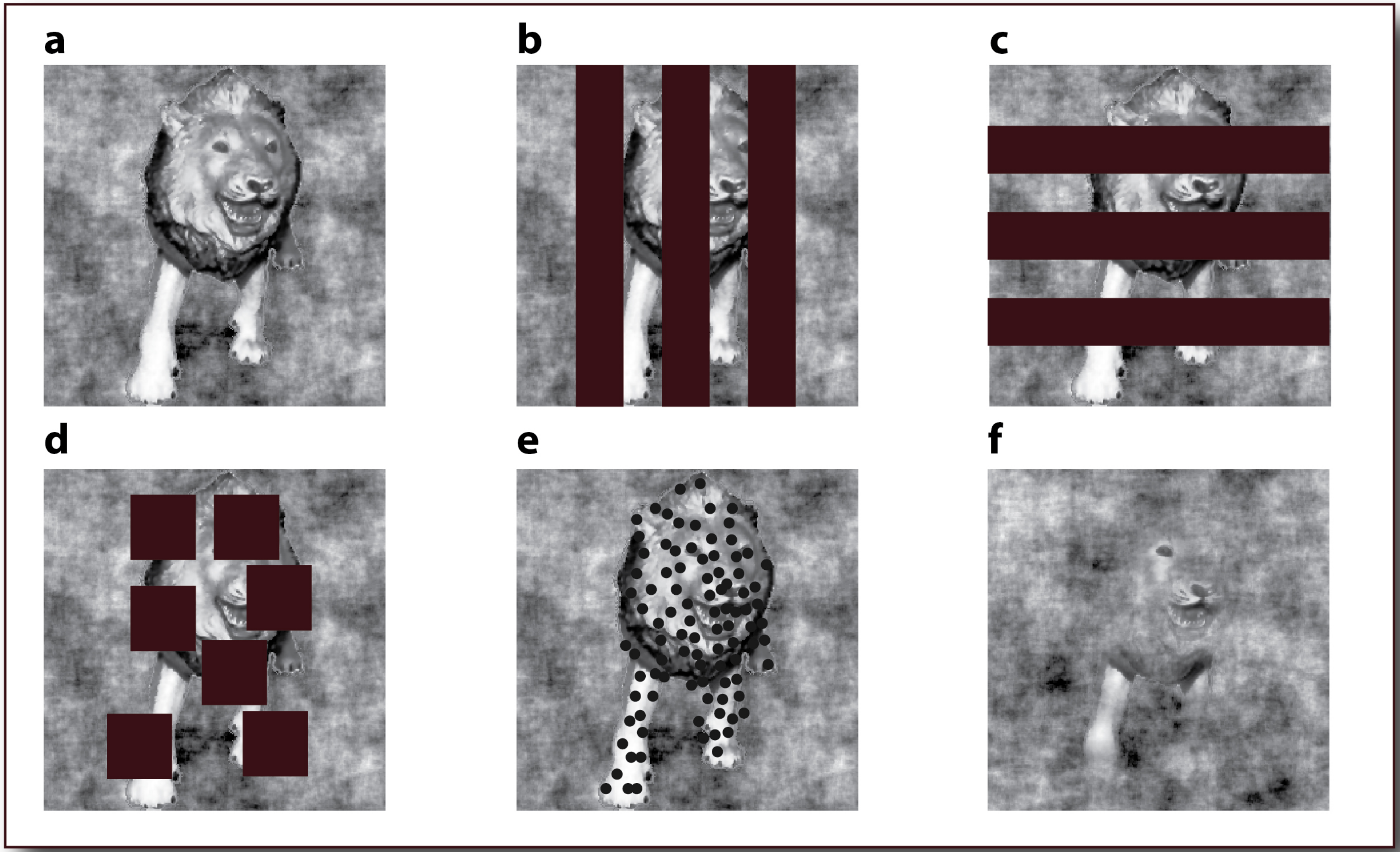
A massive recollection capacity



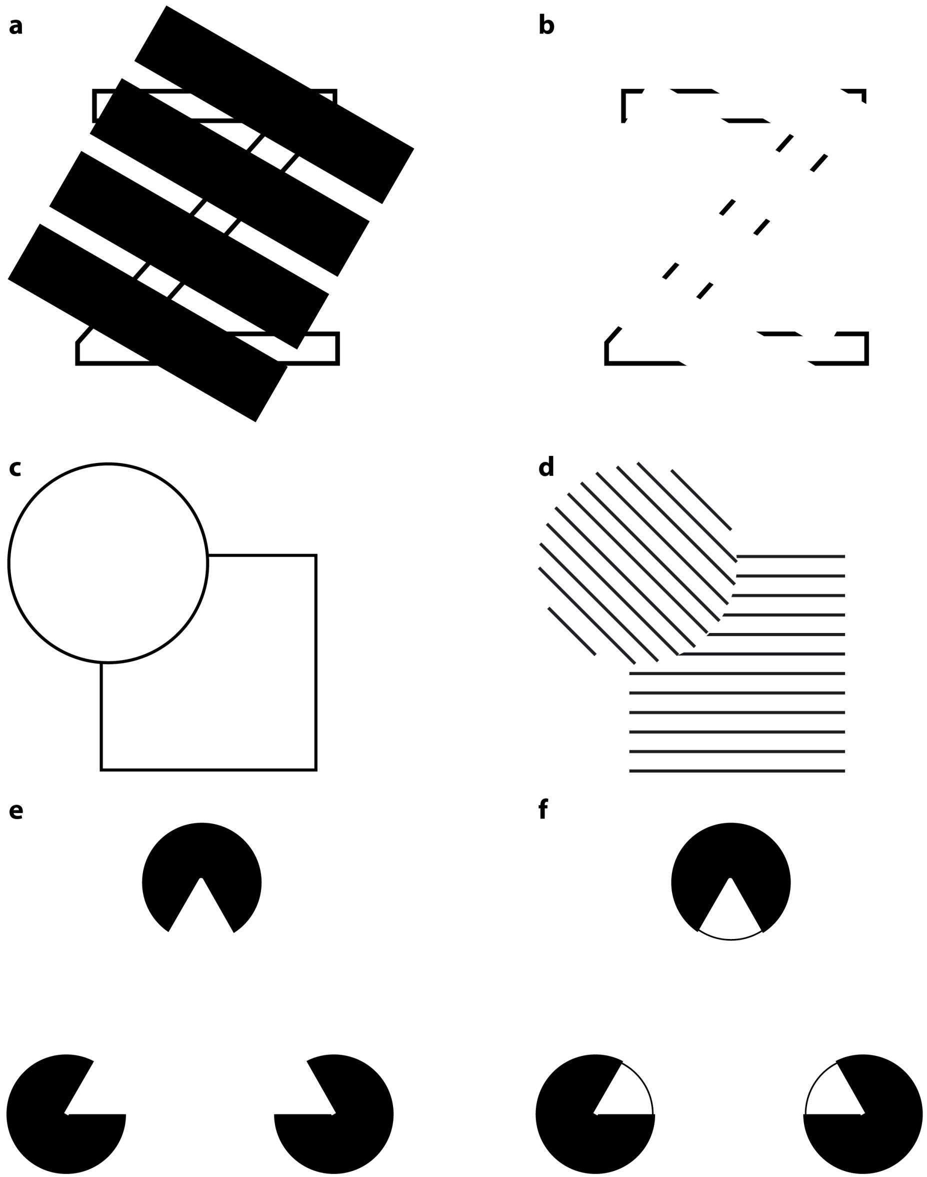


Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. Proc Natl Acad Sci U S A, 105(38), 14325-14329.

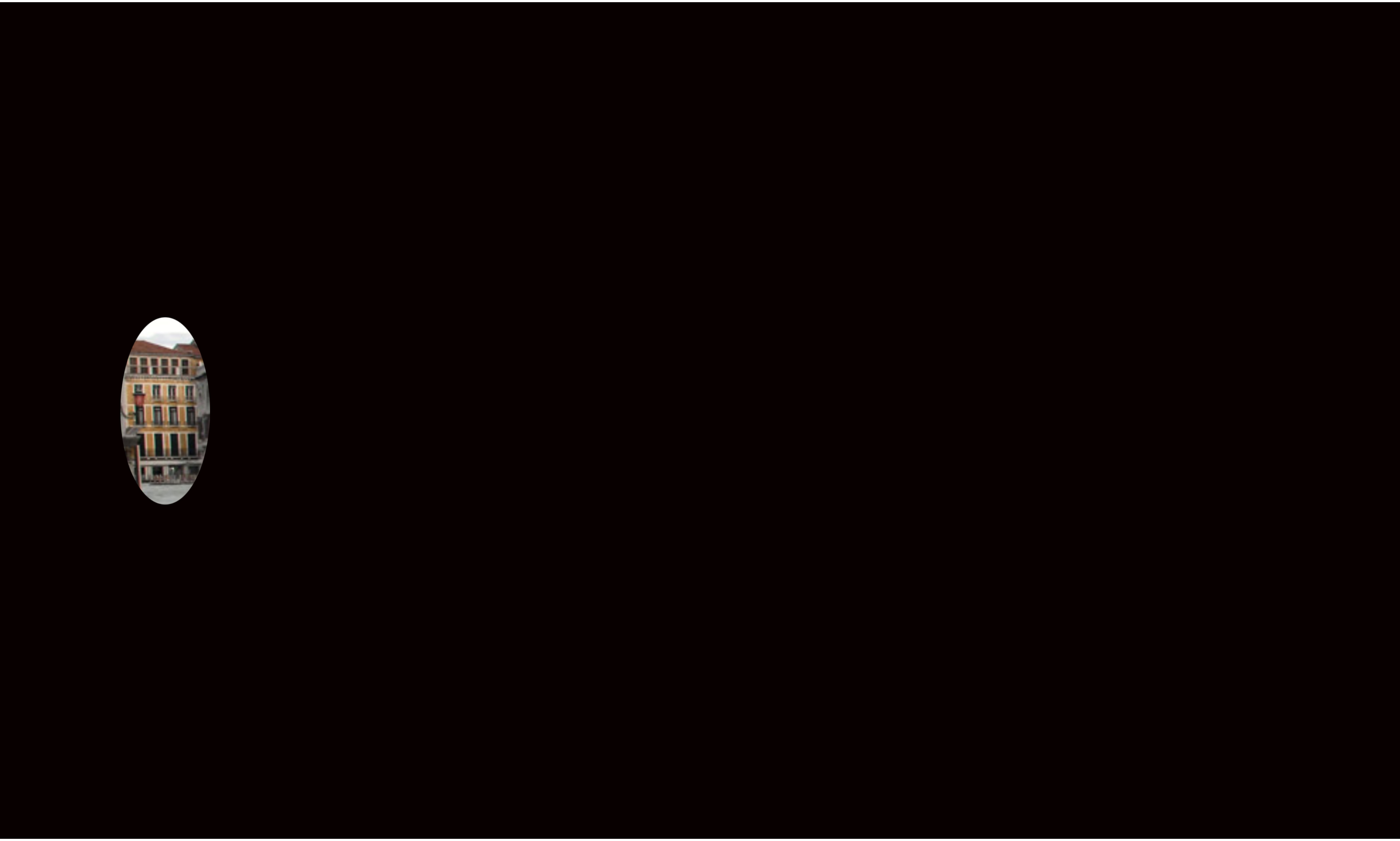
Pattern completion: Objects can be recognized from partial information



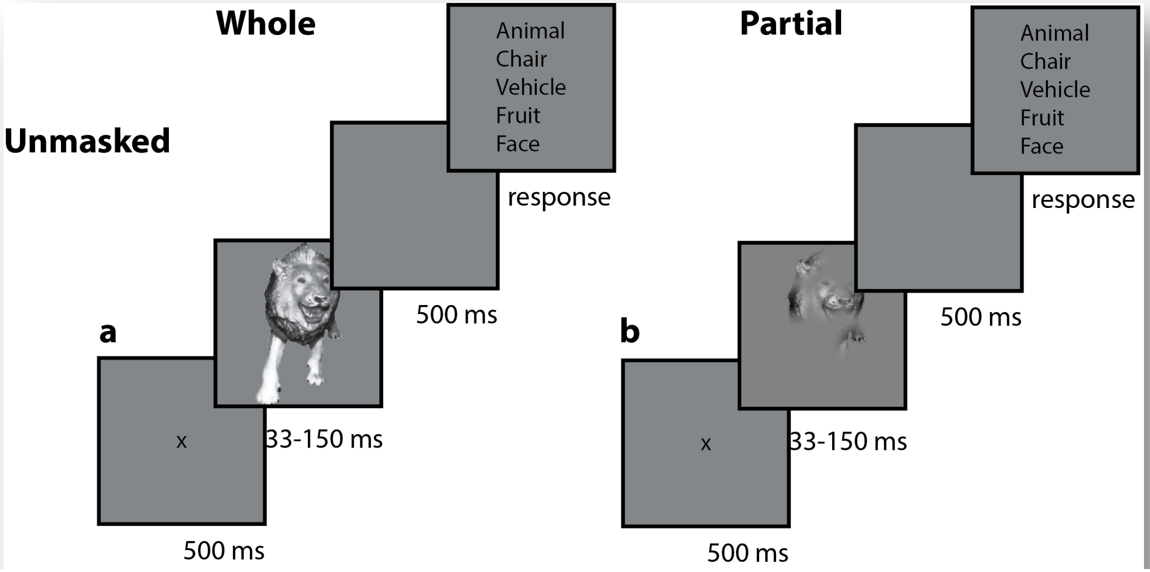
Amodal completion



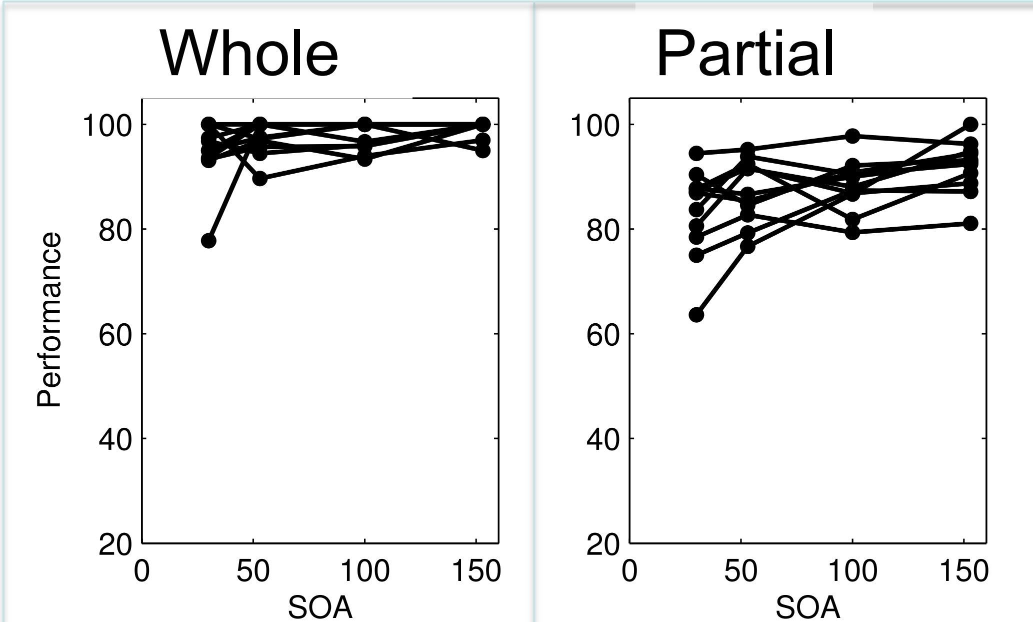
Object recognition from partial information

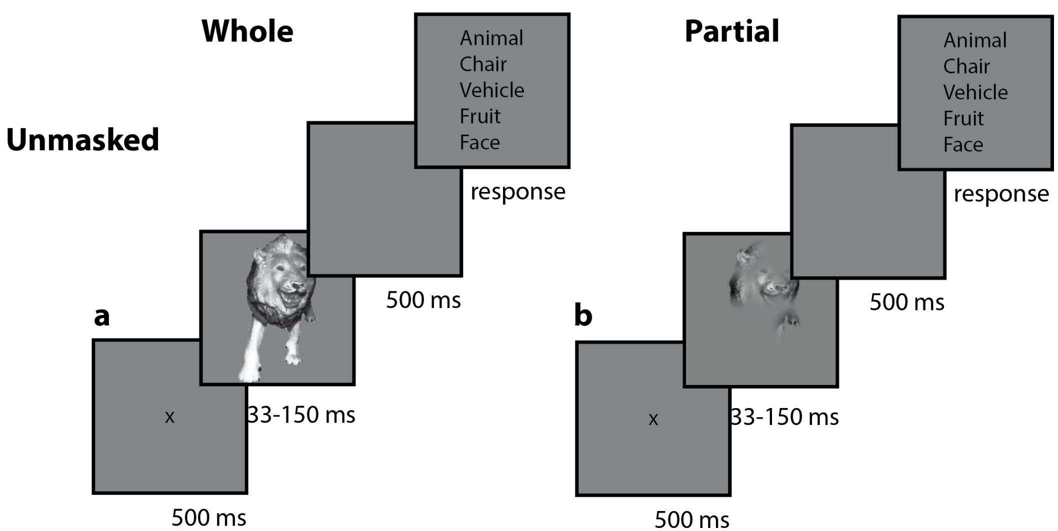


Object completion task

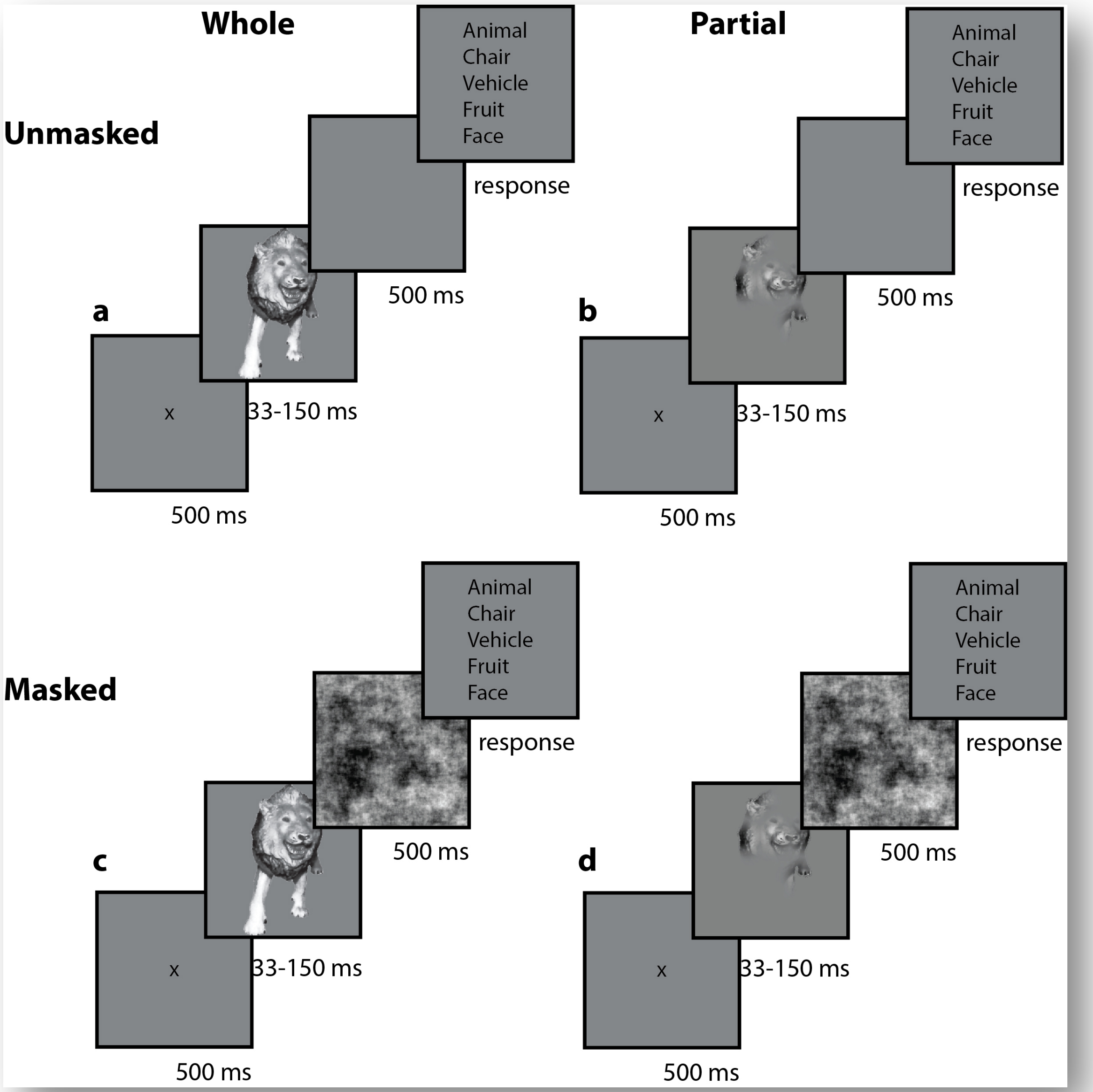


Object completion (unmasked condition)

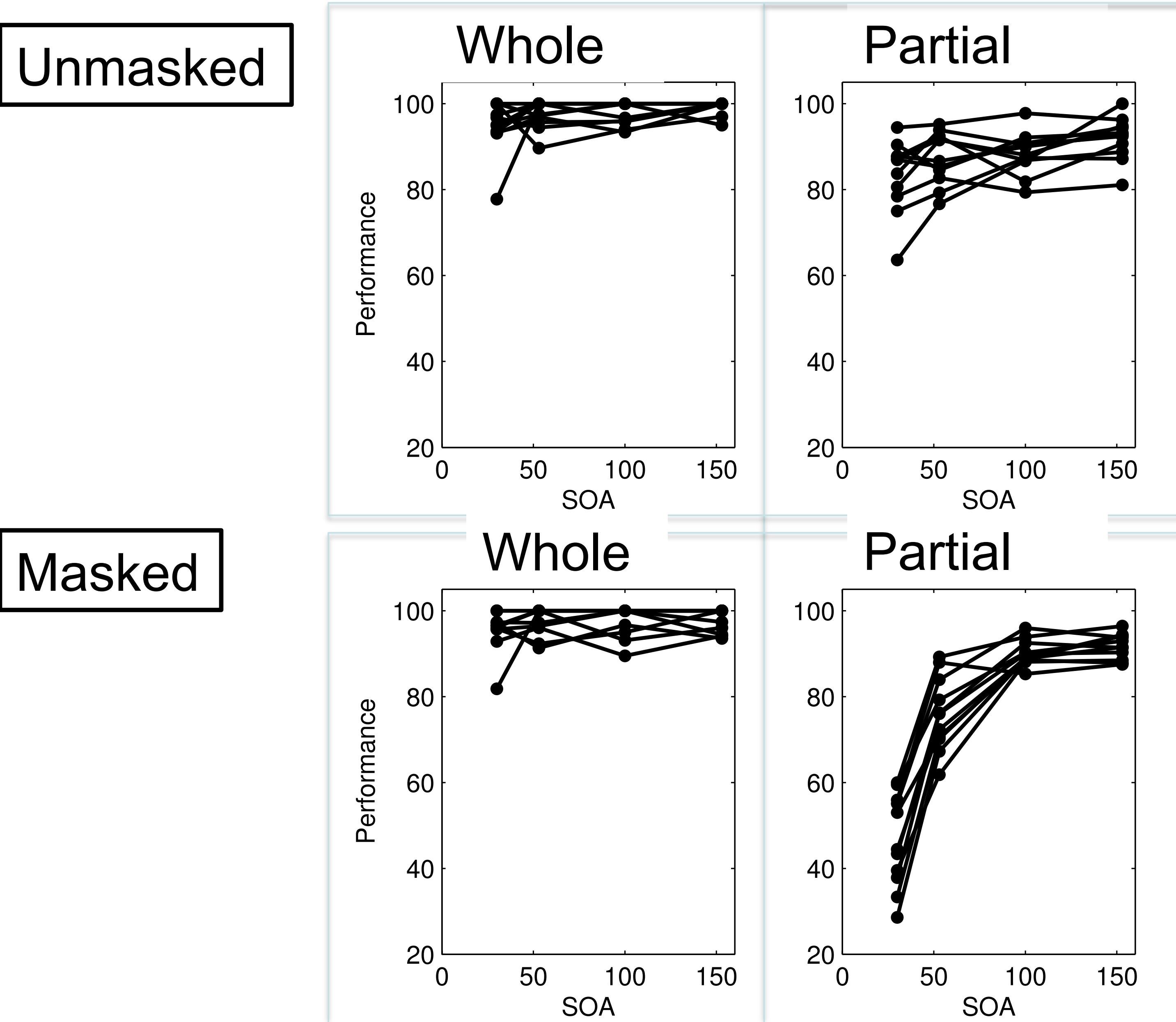




Object completion task (masking)



Object completion (unmasked condition)



Further reading

- Regan, D. Human Perception of Objects (2000). Sinauer Associates. Sunderland, Massachusets.
- Frisby, JP and Stone JV. Seeing (2010). MIT Press. Cambridge, Massachusetts.

<u>Original articles cited in class (see lecture notes for complete list)</u>

- Potter, MC (1969) Recognition memory for a rapid sequence of pictures. Journal of Experimental Psychology 81:10-15.
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. Vision Res, 46(11), 1762-1776.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. Proc Natl Acad Sci U S A, 105(38), 14325-14329
- Mooney CM. (1957). Age in the development of closure ability in children. Canadian Journal of Psychology 11: 219-226
- McKone et al, Frontiers in Psychology, 2013
- Singer and Kreiman (2014). Short temporal asynchrony disrupts visual object recognition. Journal of Vision 12:14.
- Tang, H., et al. (2014). "Spatiotemporal dynamics underlying object completion in human ventral visual cortex." Neuron **83**: 736-748.
- Tang, H., et al. (2014). "A role for recurrent processing in object completion: neurophysiological, psychophysical and computational evidence." CBMM Memo(9).