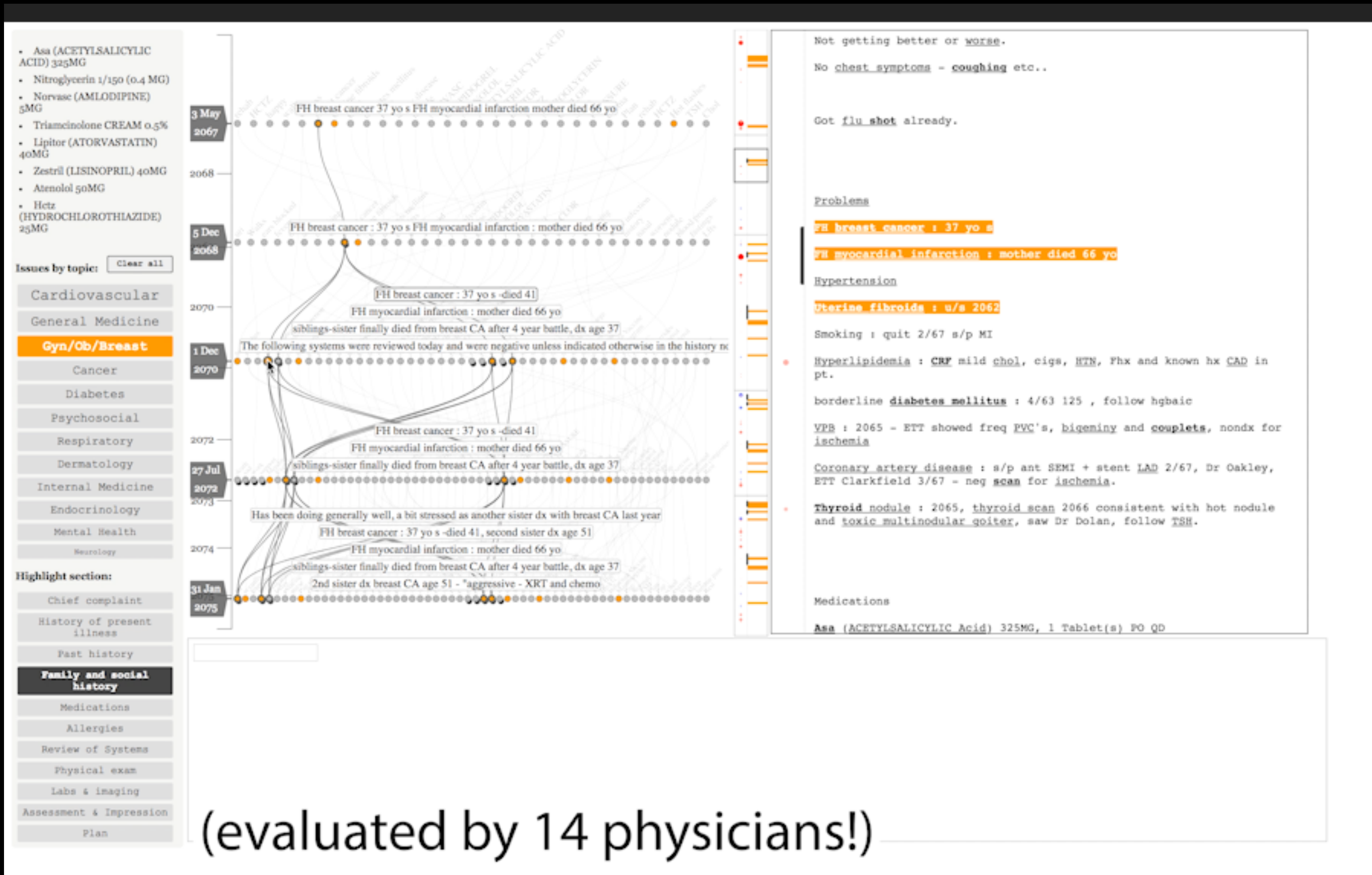
CSC2537 / STA2555 - INFORMATION VISUALIZATION

VISUAL PERCEPTION / DATA MODELS

Fanny CHEVALIER

UNIVERSITY OF TORONTO

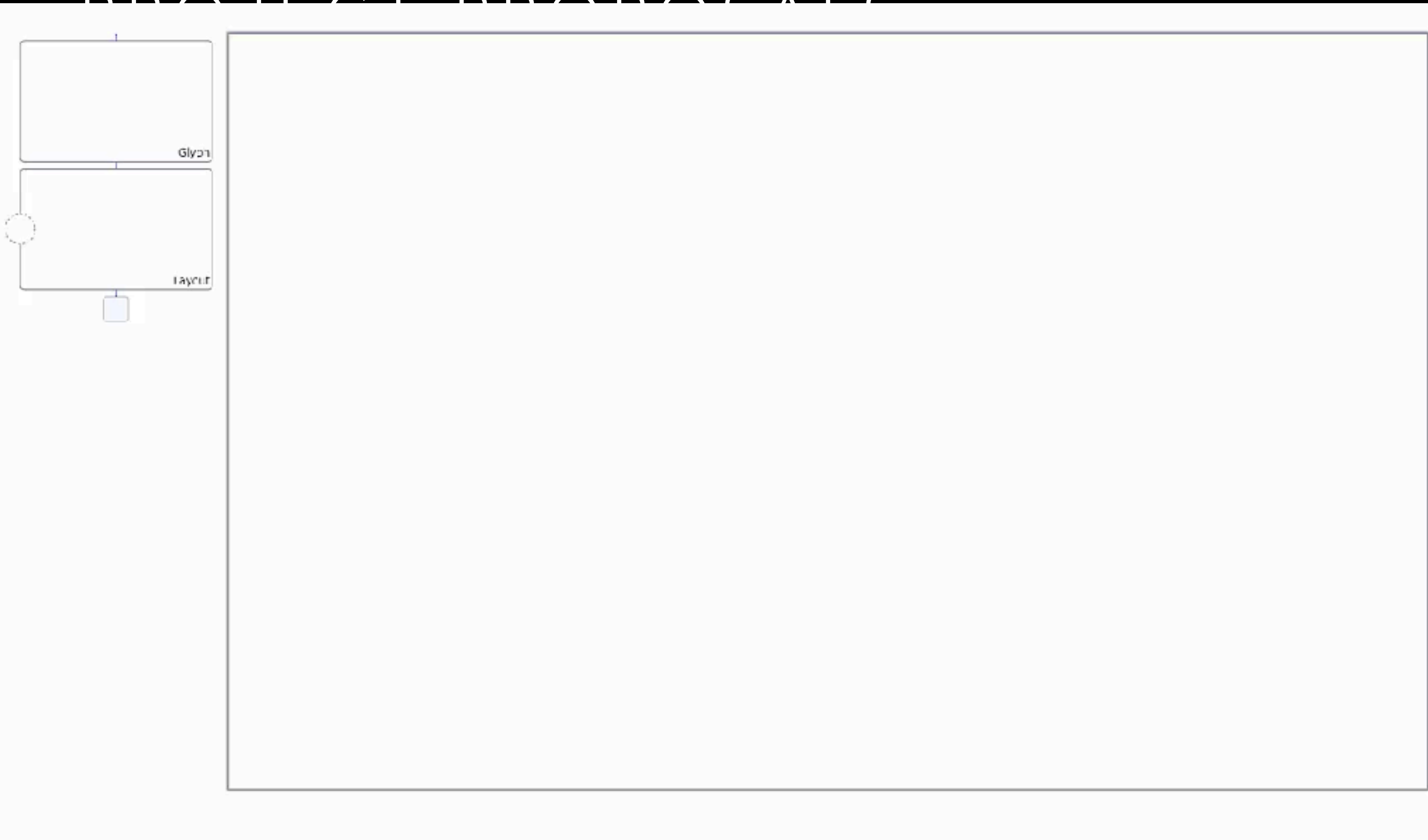
PROJECT PROPOSALS



(evaluated by 14 physicians!)

ACM CHI 2018 — Best paper honourable mention!

PROJECT PROPOSALS



ACM CHI 2018 — Best paper honourable mention!

PROJECT PROPOSALS

Visualization of mutation dynamics from serial sequencing and their clinical relevance in cancer genomics

TaeHyung Kim

Background. Recent breakthroughs in high-throughput technologies as well as their substantially declining of sequencing cost have enabled us to conduct disease studies across entire genome. It has become standard procedure to perform sequencing to detect somatic variants and study the cancer biology. It turned out that most types of tumours are results of multi-step process of mutation acquisitions. Heterogeneous mixture of genetically distinct subclones can be inferred from sequencing as well as it provides deeper understanding on treatment response and resistance especially when serial samples are available. Thus, there is an increase in the number of studies collecting multiple tumour samples from a single case whether it is a multi-regional and/or longitudinal. As an example, our recent study showed patterns of mutation clusters are associated with clinical events (Figure 1) (Kim et al, 2015). To extend this finding, we performed serial sequencing on 100 CML (chronic myeloid leukemia) and 95 MDS (myelodysplastic syndrom) patients before and after treatment to assess the impact or association of mutational dynamics and its clinical relevance.

…

PROJECT PROPOSALS

Visualization of mutation dynamics from serial sequencing and their clinical relevance in cancer genomics

TaeHyung Kim

Motivation, related work, and challenges Currently available methods are typically manual and labor-intensive. In particular, it only visualizes a single case or a sampling time-point, rather than mining for the general pattern in longitudinal data. For example, they focus on visualizing these mutational data using 2D heatmap where each row contains gene name that carries mutations and each column contains the patient ID. Cells are sorted either by the mutation frequency or the number of mutations per patient. In addition, visualization of clonal evolution in a single case is also performed manually in most cases. Fishplot, developed by Miller et al visualizes the mutation pattern from serial sequencing, but it requires specific format of the data (Miller et al, 2016). It only visualizes the clonal evolution in a single case, thus manual inspection after visualizing each case is required (Figure 2). In such case, only one or two representative cases are presented. Alternatively, an extra dimension to a heatmap (i.e. 3D heatmap where the extra dimension is the mutation status at 2nd sampling time) has been attempted, but it does not fully describe the value of serial sequencing. The 2D heatmap with the same format per each sampling time point has same issues. As such, there is a lack of tool, that leverages time-series information and visualizes mutation dynamics and its clinical relevance simultaneously.

…

TaeHyung Kim, CSC2524 Final Report

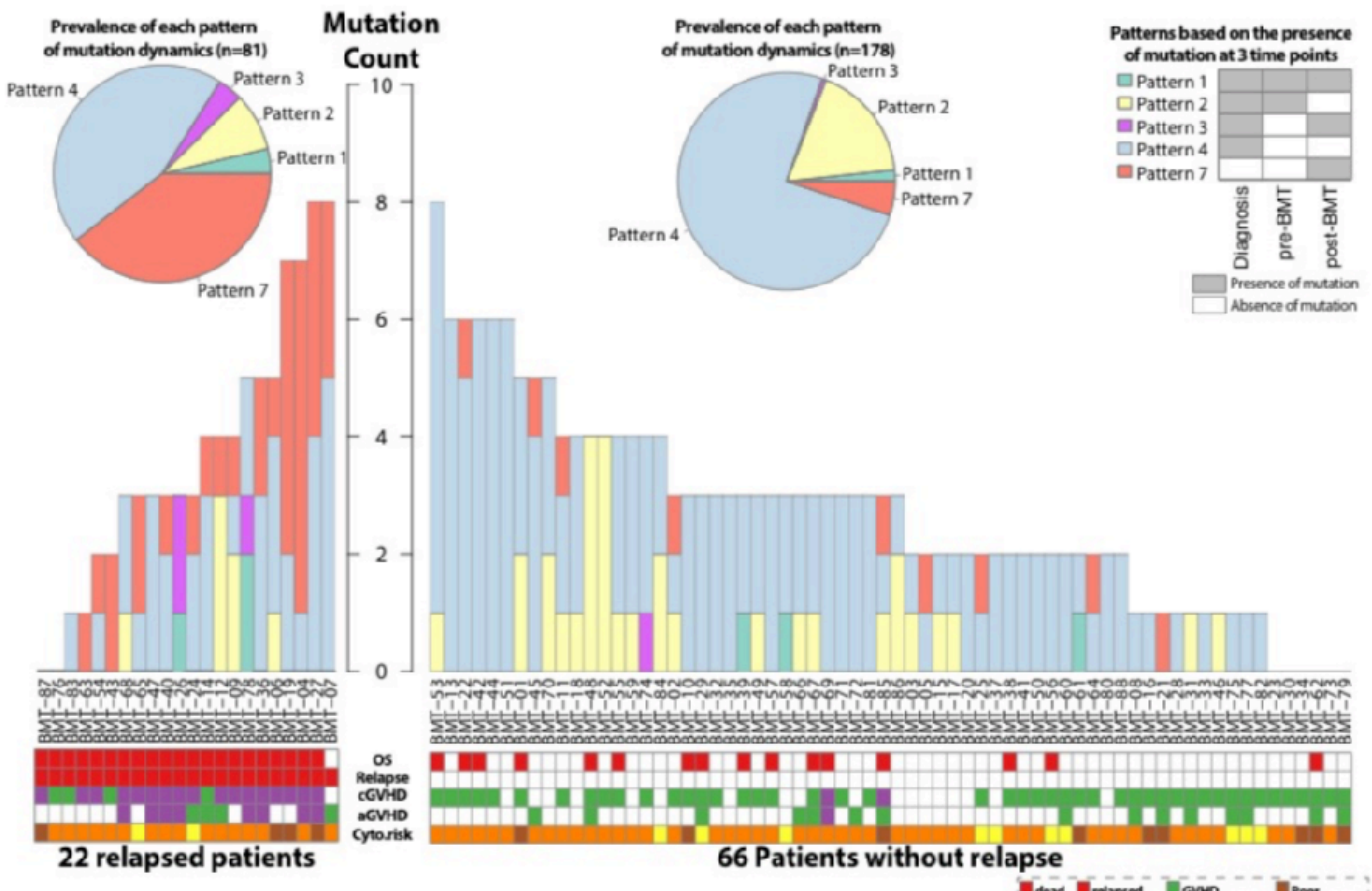
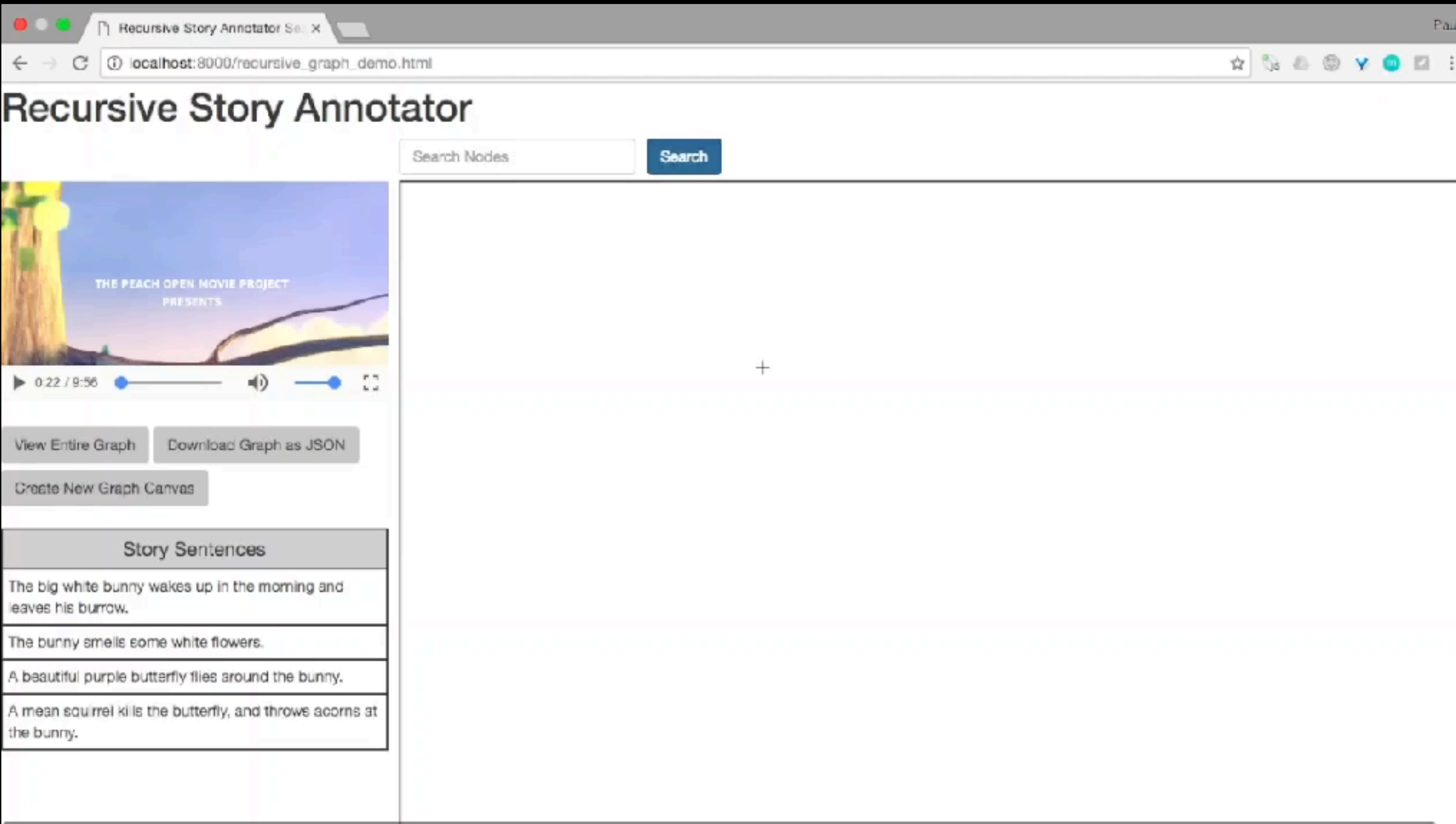


Figure 7

PROJECT PROPOSALS



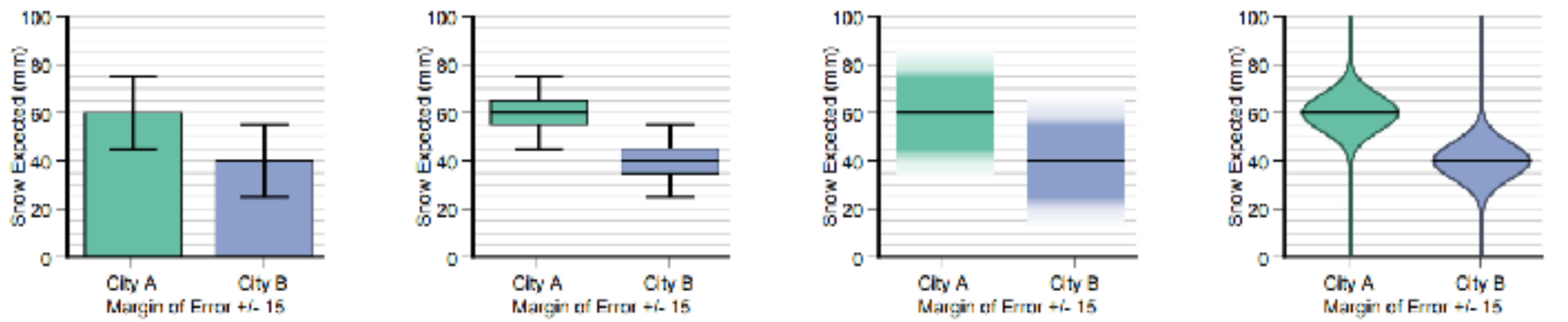
PROSPECTIVE PROJECTS

- Nutritional facts / compare products
- Visualization of music pieces (see TSO)
- Climate change
- Sexual harassement in academia https://docs.google.com/spreadsheets/d/1S9KShDLvU7C-KkgEevYTHXr3F6InTenrBsS9yk-8C5M/edit#gid=1530077352
- Data journalism
- Understanding visualization
- Explorable Explanations
- Explainable AI
- Sketching & Animation

UNDERSTANDING VISUALIZATION

Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error

Michael Correll *Student Member, IEEE*, and Michael Gleicher *Member, IEEE*



(a) **Bar chart with error bars**: the height of the bars encodes the sample mean, and the whiskers encode a 95% t-confidence interval.

(b) **Modified box plot**: The whiskers are the 95% t-confidence interval, the box is a 50% t-confidence interval.

(c) **Gradient plot**: the transparency of the colored region corresponds to the cumulative density function of a t-distribution.

(d) **Violin plot**: the width of the colored region corresponds to the probability density function of a t-distribution.

Fig. 1. Four encodings for mean and error evaluated in this work. Each prioritizes a different aspect of mean and uncertainty, and results in different patterns of judgment and comprehension for tasks requiring statistical inferences.

Abstract— When making an inference or comparison with uncertain, noisy, or incomplete data, measurement error and confidence intervals can be as important for judgment as the actual mean values of different groups. These often misunderstood statistical quantities are frequently represented by bar charts with error bars. This paper investigates drawbacks with this standard encoding, and considers a set of alternatives designed to more effectively communicate the implications of mean and error data to a general audience, drawing from lessons learned from the use of visual statistics in the information visualization community. We present a series of crowd-sourced experiments that confirm that the encoding of mean and error significantly changes how viewers make decisions about uncertain data. Careful consideration of design tradeoffs in the visual presentation of data results in human reasoning that is more consistently aligned with statistical inferences. We suggest the use of gradient plots (which use transparency to encode uncertainty) and violin plots (which use width) as better alternatives for inferential tasks than bar charts with error bars.

Index Terms—Visual statistics, information visualization, crowd-sourcing, empirical evaluation

UNDERSTANDING VISUALIZATION

Blinded with Science or Informed by Charts? A Replication Study

Pierre Dragicevic and Yvonne Jansen

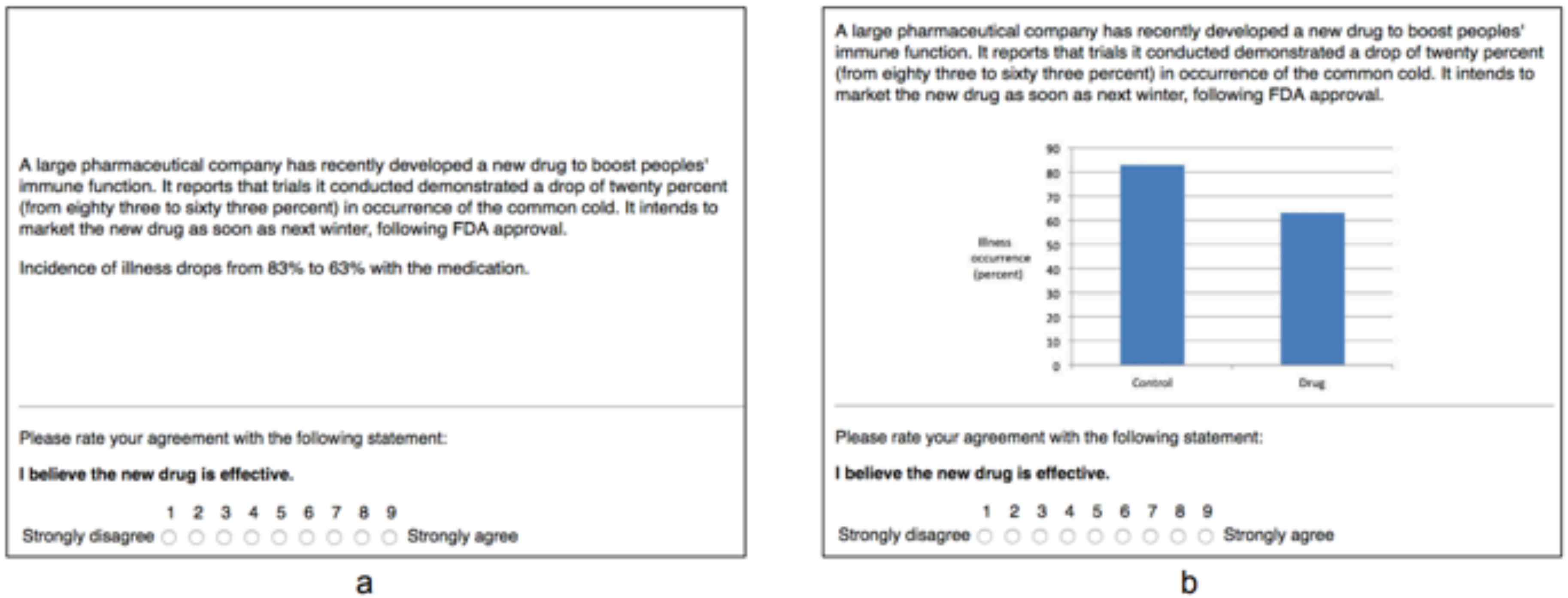


Fig. 1. First page of our second experiment, replicating experiment 2 from Tal and Wansink [49]. (a) no-chart condition, with an extra sentence repeating the two quantities with numerals; (b) chart condition: the extra sentence is replaced with a bar chart.

Abstract —We provide a reappraisal of Tal and Wansink's study "Blinded with Science", where seemingly trivial charts were shown to increase belief in drug ef cacy, presumably because charts are associated with science. Through a series of four replications conducted on two crowdsourcing platforms, we investigate an alternative explanation, namely, that the charts allowed participants to better assess the drug's ef cacy. Considered together, our experiments suggest that the chart seems to have indeed promoted understanding, although the effect is likely very small. Meanwhile, we were unable to replicate the original study's ndings, as text with chart appeared to be no more persuasive – and sometimes less persuasive – than text alone. This suggests that the effect may not be as robust as claimed and may need speci c conditions to be reproduced. Regardless, within our experimental settings and considering our study as a whole (N = 623), the chart's contribution to understanding was clearly larger than its contribution to persuasion.

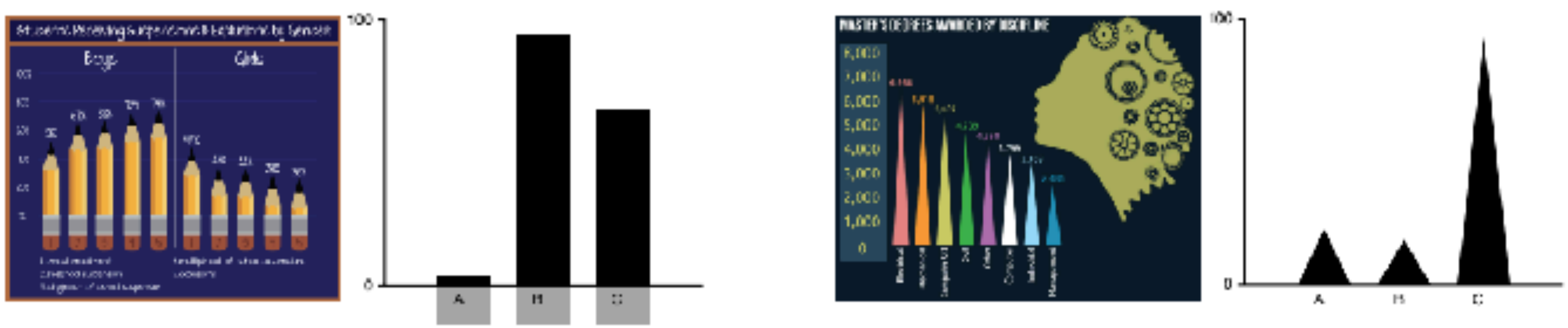
Index Terms —Replication study, persuasion, charts, data comprehension, methodology.

UNDERSTANDING VISUALIZATION

An Evaluation of the Impact of Visual Embellishments in Bar Charts

Drew Skau[1,4], Lane Harrison[2], and Robert Kosara[3,4]

[1]Visually Inc. [2]Tufts University [3]Tableau Research [4]UNC Charlotte



(a) Bars extend below zero line. (b) Triangle bar chart.

Figure 1: Two examples of embellished charts and abstracted versions of the embellishments.

Abstract

As data visualization becomes further intertwined with the field of graphic design and information graphics, small graphical alterations are made to many common chart formats. Despite the growing prevalence of these embellishments, their effects on communication of the charts' data is unknown. From an overview of the design space, we have outlined some of the common embellishments that are made to bar charts. We have studied the effects of these chart embellishments on the communication of the charts' data through a series of user studies on Amazon's Mechanical Turk platform. The results of these studies lead to a better understanding of how each chart type is perceived, and help provide guiding principles for the graphic design of charts.

UNDERSTANDING VISUALIZATION

Using Concrete Scales: A Practical Framework for Effective Visual Depiction of Complex Measures

Fanny Chevalier, Romain Vuillemot, and Guia Gali

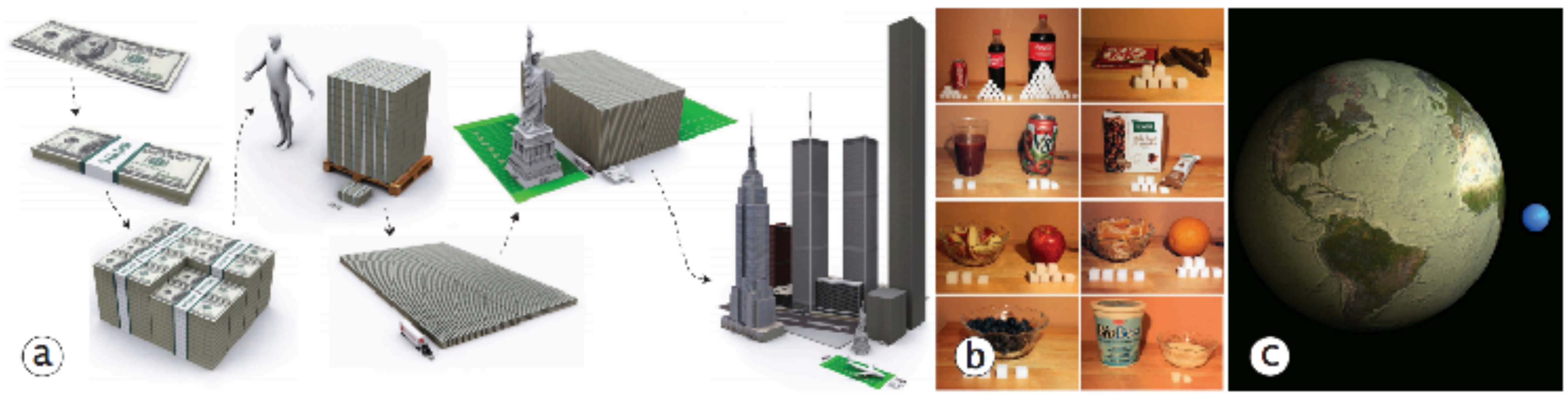


Fig. 1. Illustrates popular representations of complex measures: (a) *US Debt* (Oto Godfrey, Demonocracy.info, 2011) explains the gravity of a 115 trillion dollar debt by progressively stacking 100 dollar bills next to familiar objects like an average-sized human, sports fields, or iconic New York city buildings [15] (b) *Sugar stacks* (adapted from SugarStacks.com) compares caloric counts contained in various foods and drinks using sugar cubes [32] and (c) *How much water is on Earth?* (Jack Cook, Woods Hole Oceanographic Institution and Howard Perlman, USGS, 2010) shows the volume of oceans and rivers as spheres whose sizes can be compared to that of Earth [38].

Abstract—From financial statistics to nutritional values, we are frequently exposed to quantitative information expressed in measures of either *extreme magnitudes* or *unfamiliar units*, or both. A common practice used to comprehend such complex measures is to relate, re-express, and compare them through visual depictions using magnitudes and units that are easier to grasp. Through this practice, we create a new graphic composition that we refer to as a *concrete scale*. To the best of our knowledge, there are no design guidelines that exist for concrete scales despite their common use in communication, educational, and decision-making settings. We attempt to fill this void by introducing a novel framework that would serve as a practical guide for their analysis and design. Informed by a thorough analysis of graphic compositions involving complex measures and an extensive literature review of scale cognition mechanisms, our framework outlines the design space of various *measure relations*—specifically relations involving the re-expression of complex measures to more familiar concepts—and their visual representations as graphic compositions.

Index Terms—Concrete scale, scale cognition, visual comparison, graphic composition, visual notation

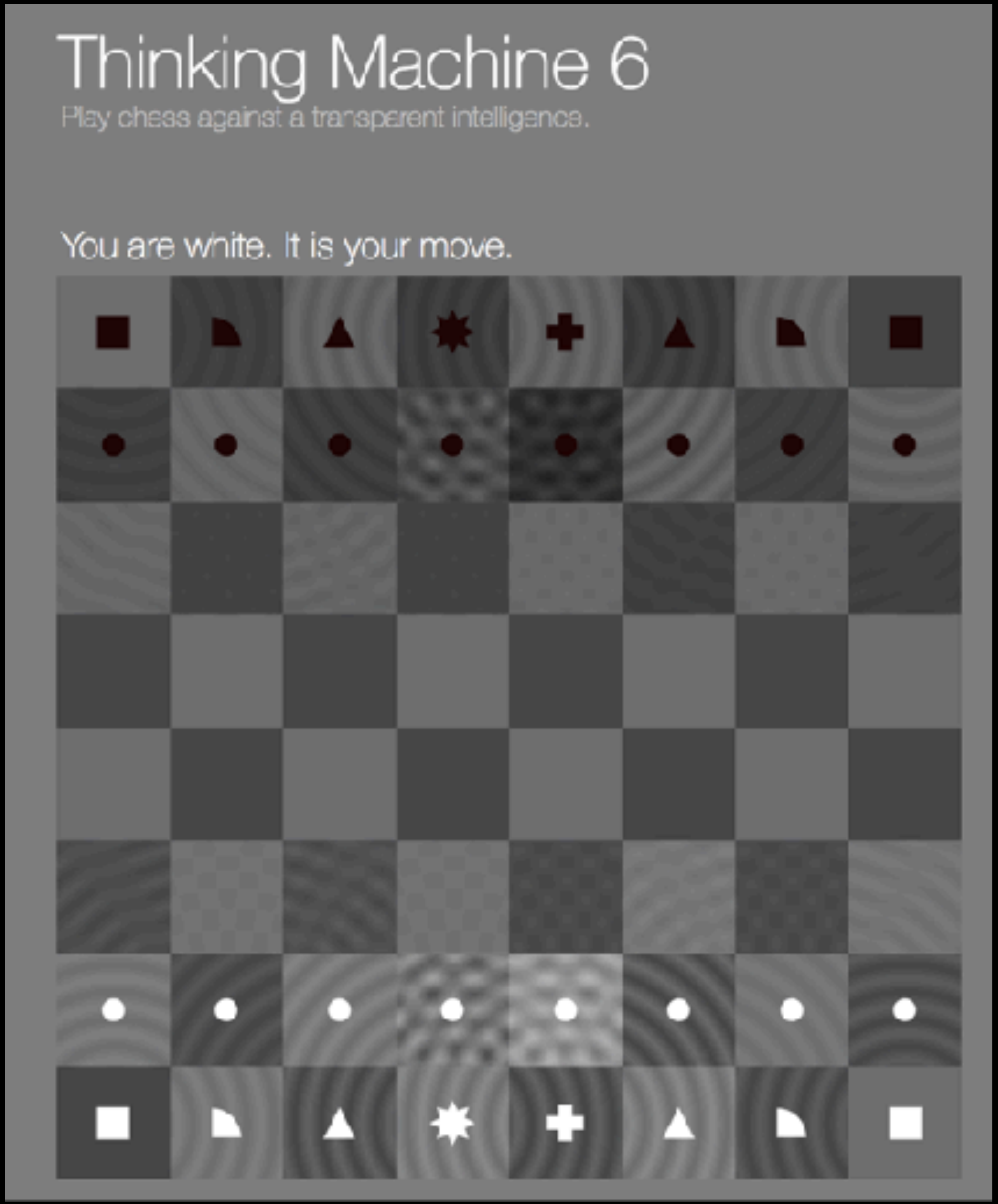
EXPLORABLE EXPLANATIONS

Lion cubs play-fight to learn hunting skills. Rats play to learn social & emotional skills. Monkeys play to learn cognitive skills, to practice problem-solving and creativity.

And yet, in the last century, we humans have convinced ourselves that play is useless, and learning is *supposed* to be boring. Gosh, no wonder we're all so miserable.

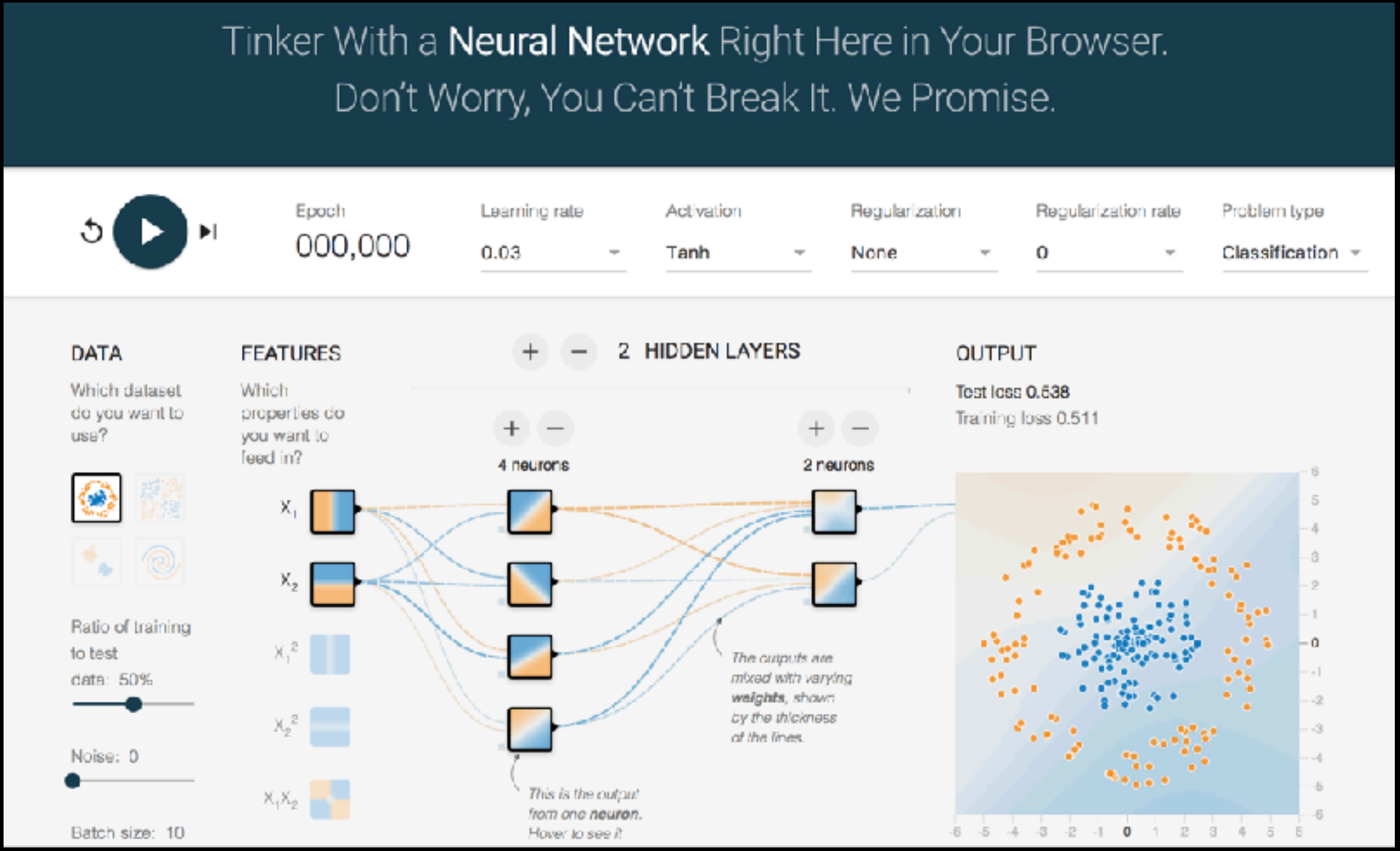
http://explorabl.es/

MAKING PROCESSES VISIBLE



http://www.bewitched.com/chess/

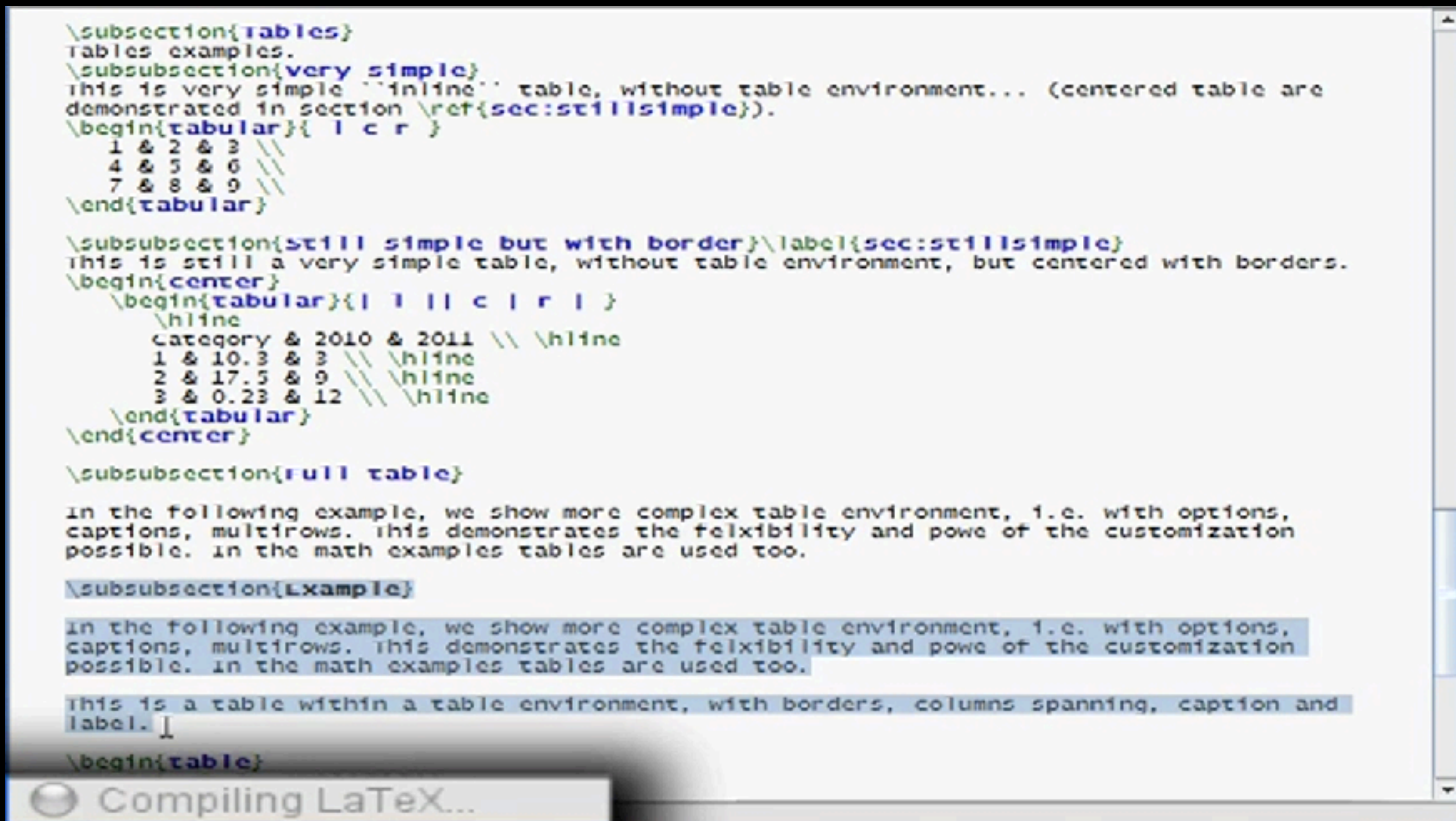
TENSOR FLOW PLAYGROUND



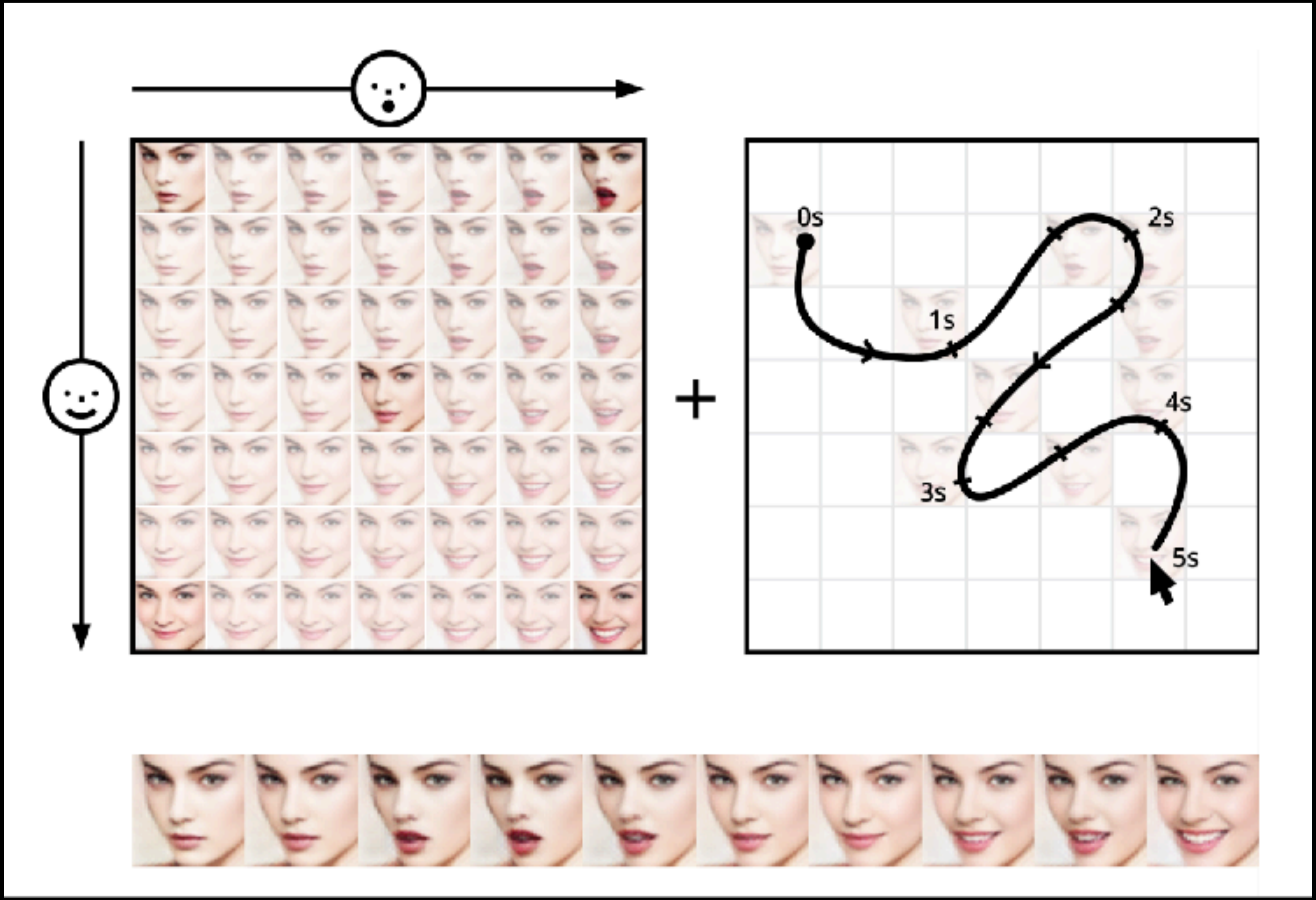
http://playground.tensorflow.org/

more cool stuff at: http://hint.fm/

ANIMATED TRANSITIONS (OF INTERNET SEARCHES?)



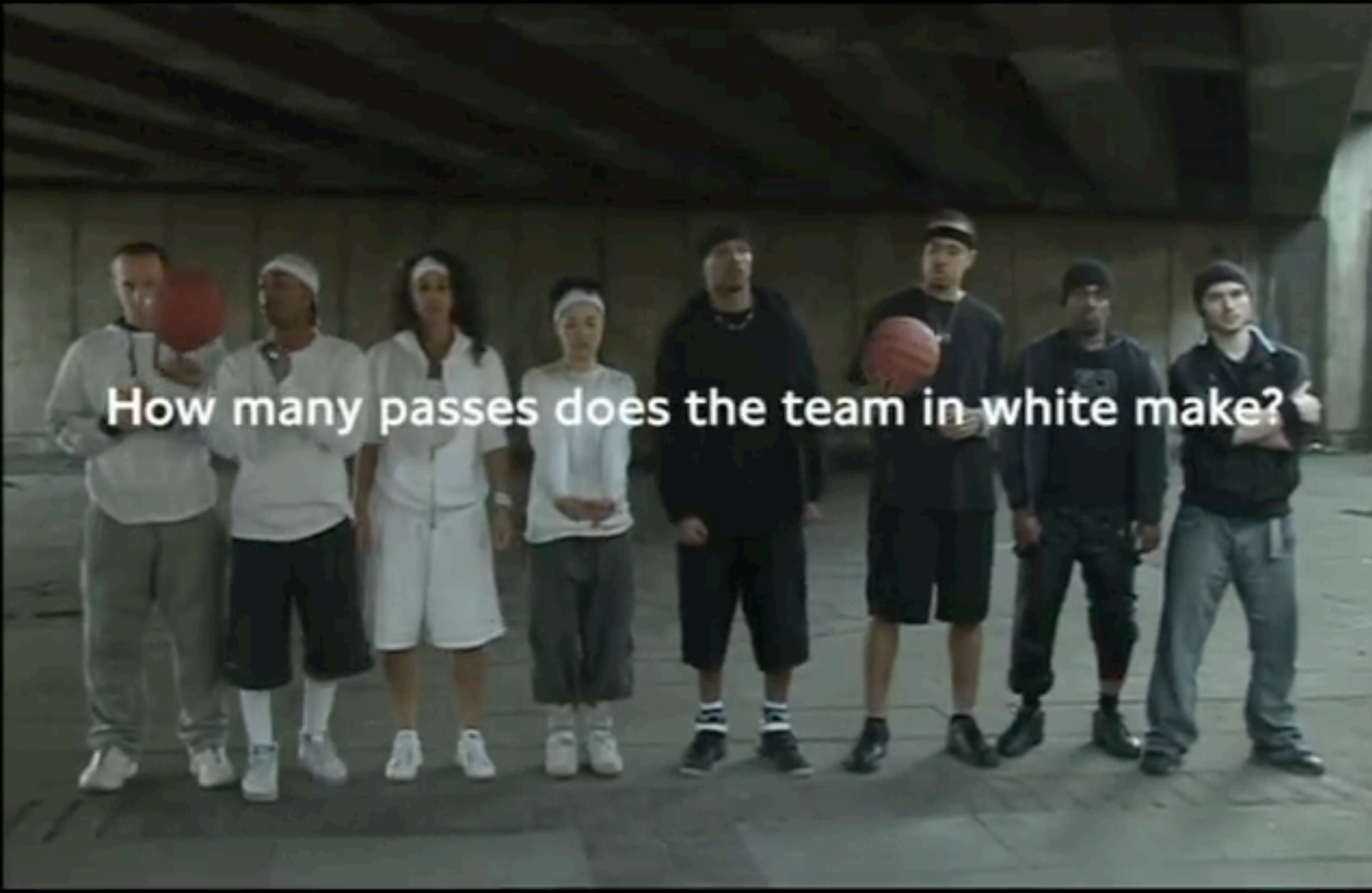
(SKETCH-BASED) TOOL FOR AUTHORING ANIMATION

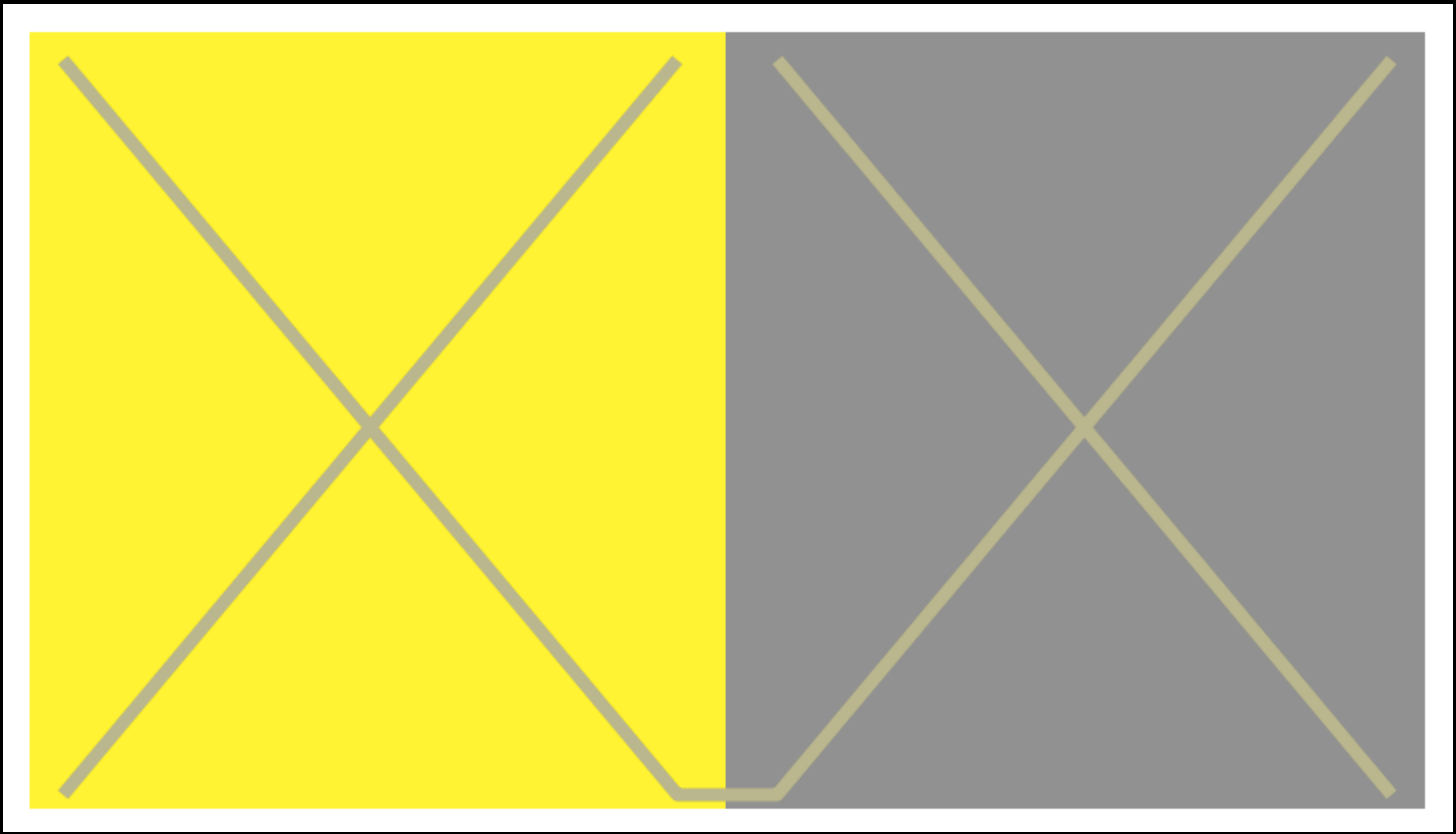


Generating Animations by Sketching in Conceptual Space Tom White*, Ian Loh*

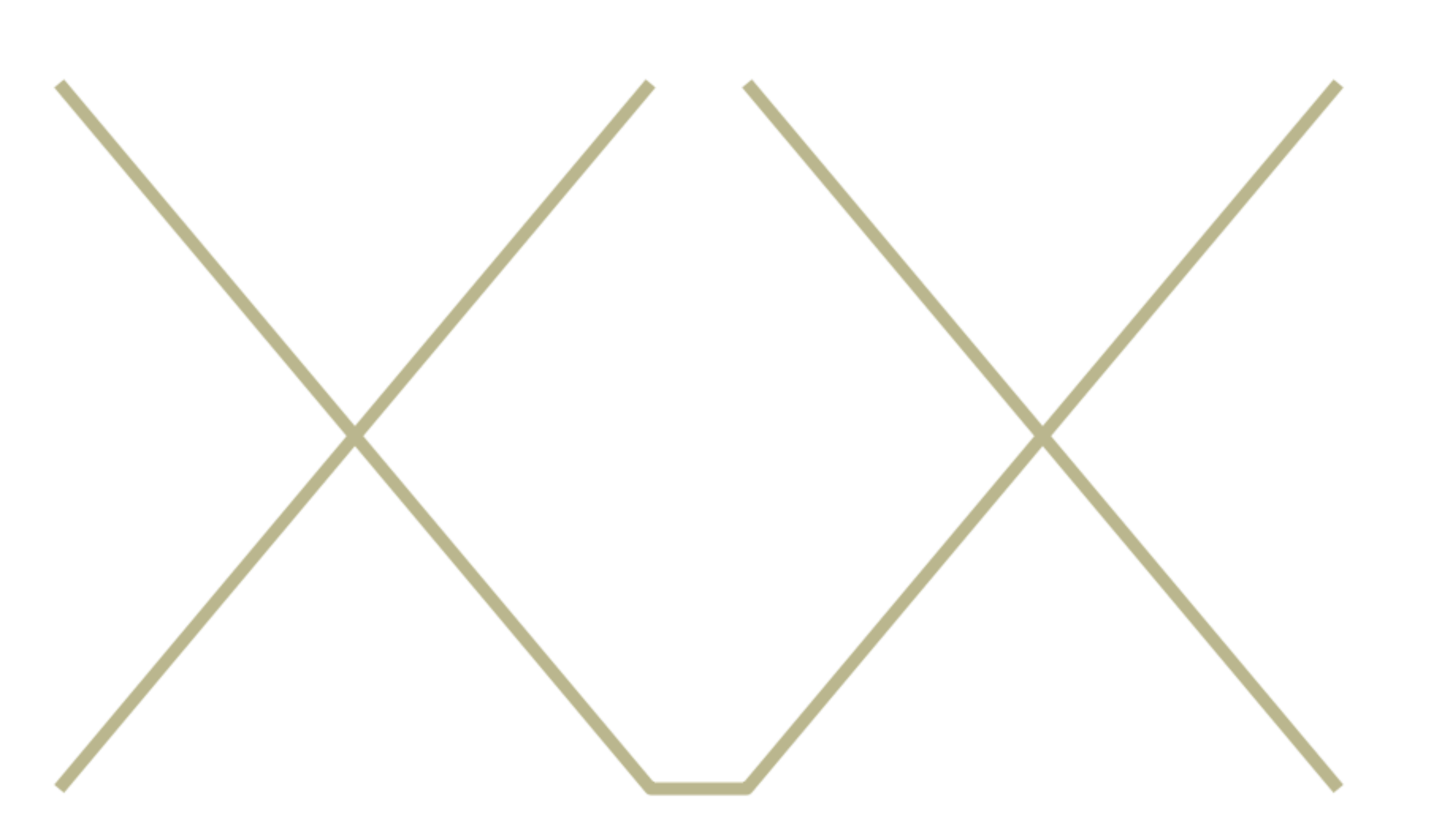
"I ONLY BELIEVE IN WHAT I SEE WITH MY OWN EYES"

BRAIN BUGS

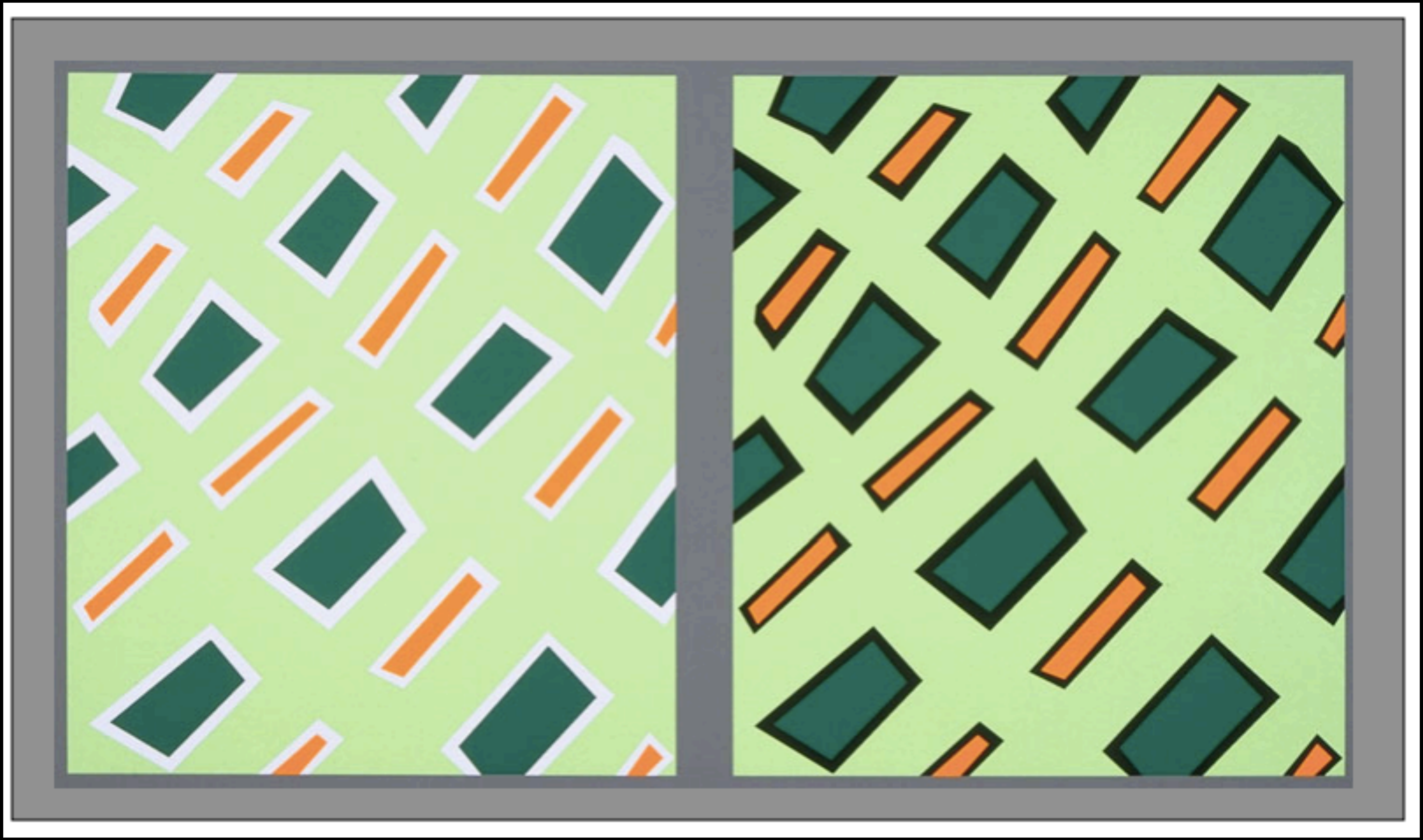

How many passes does the team in white make?

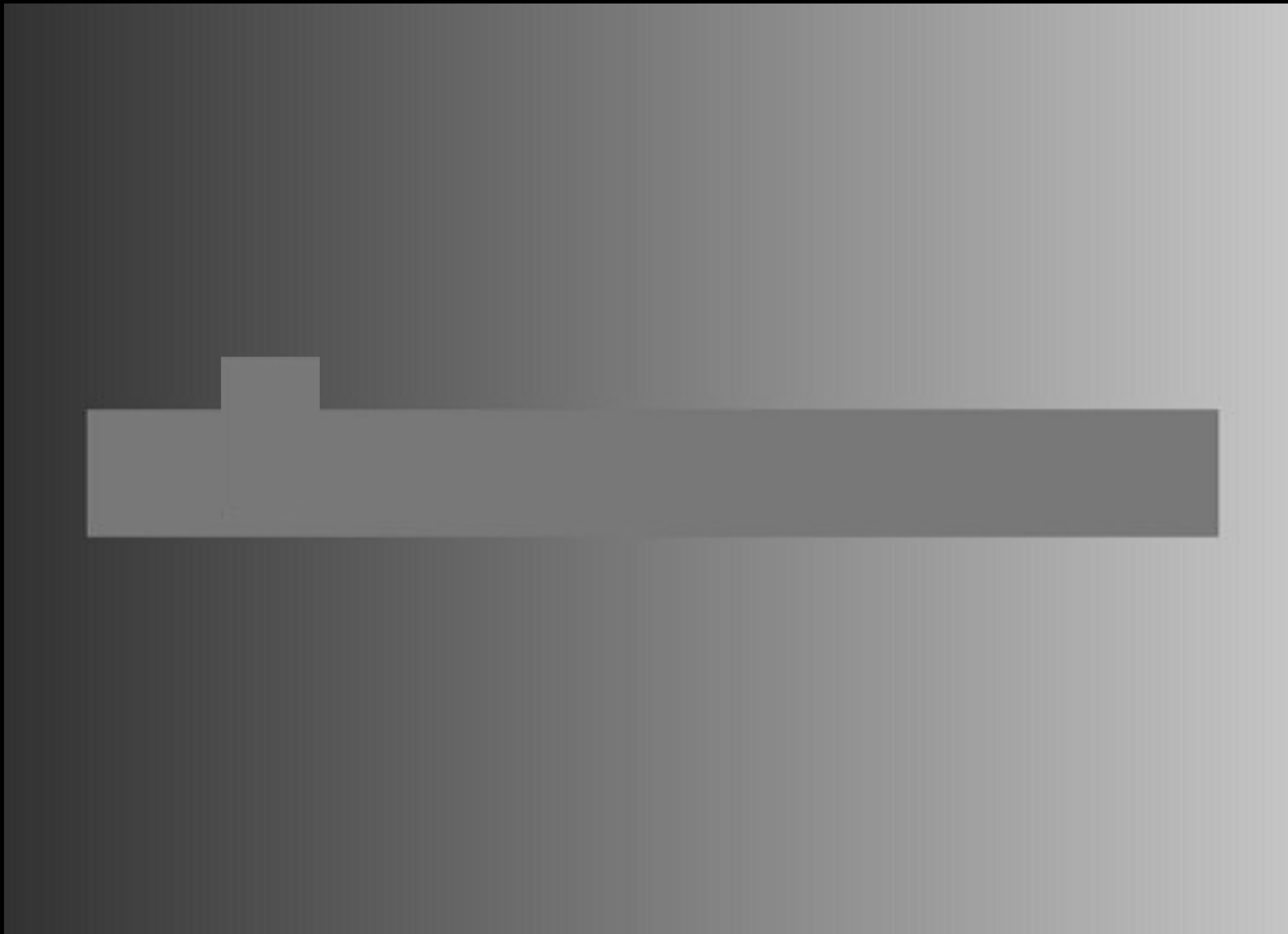


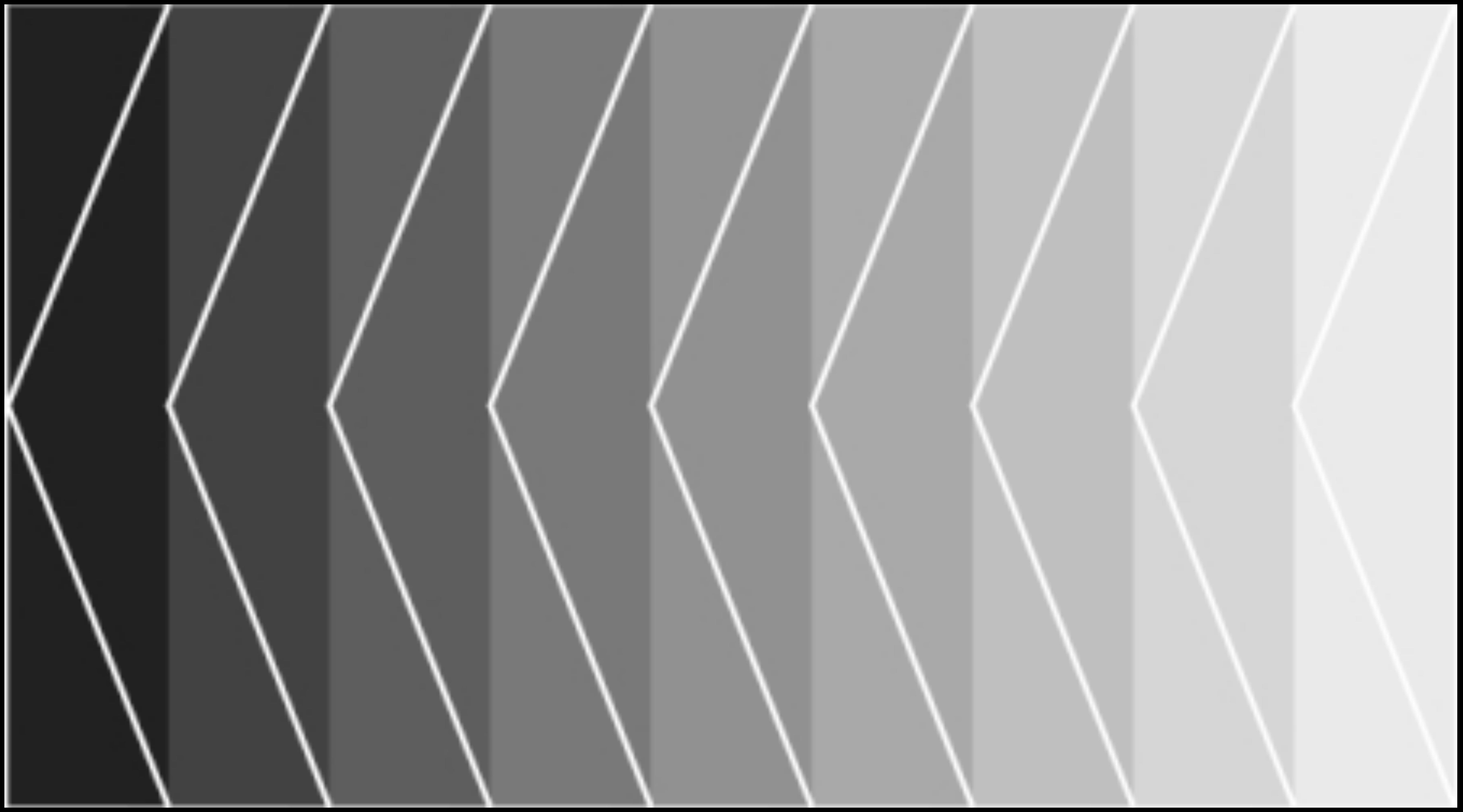
SIMULTANEOUS CONTRAST

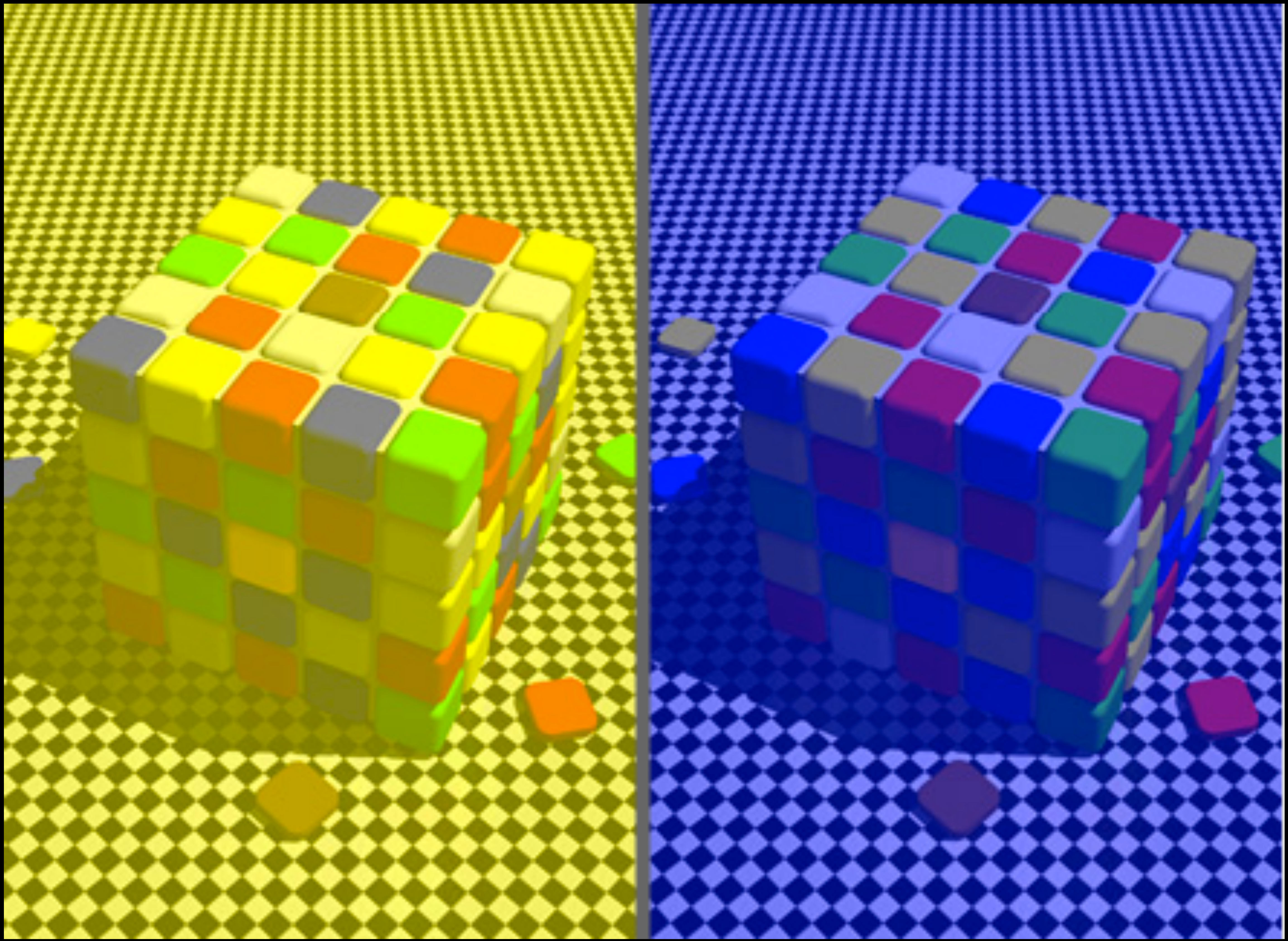


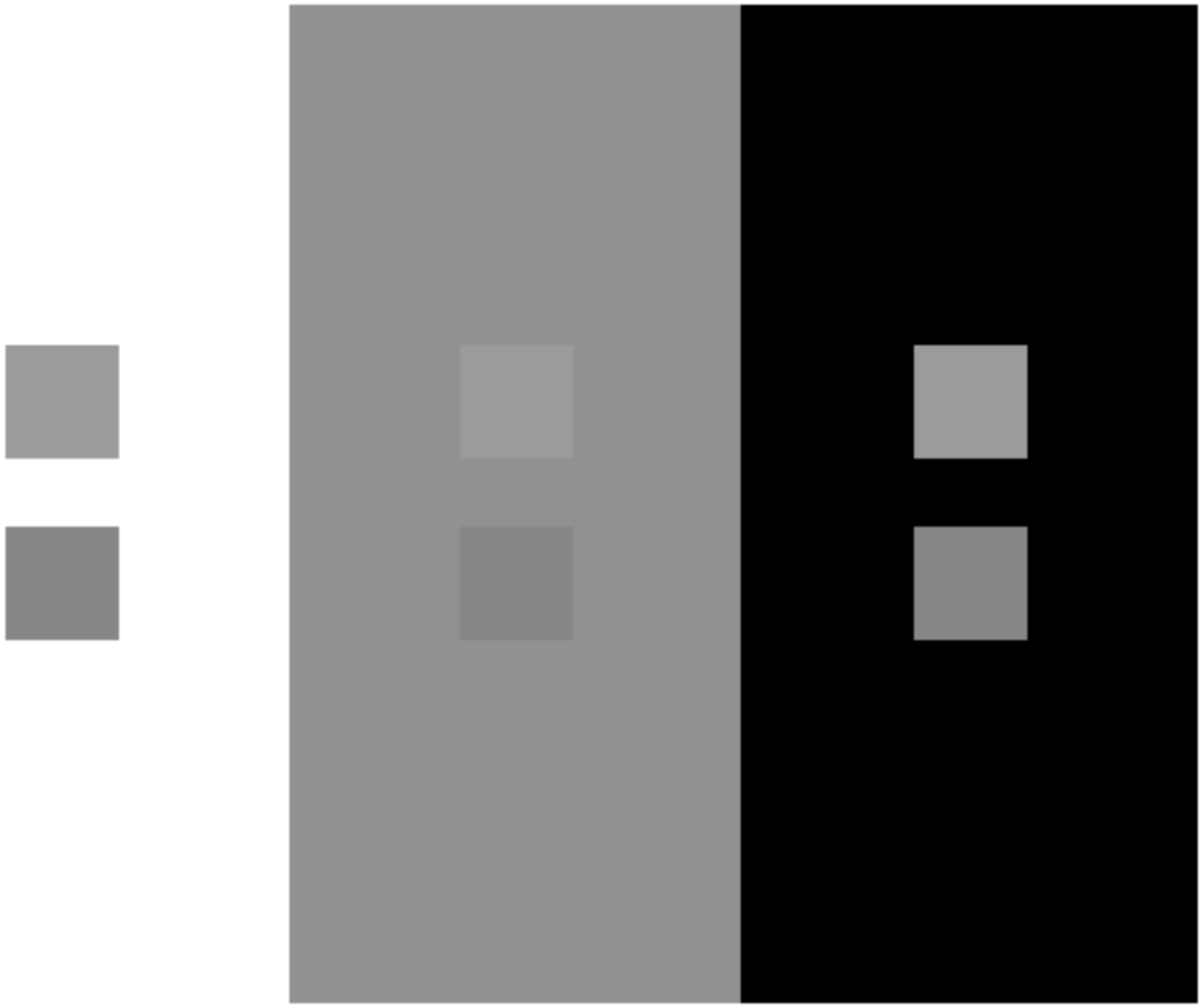
BEZOLD EFFECT



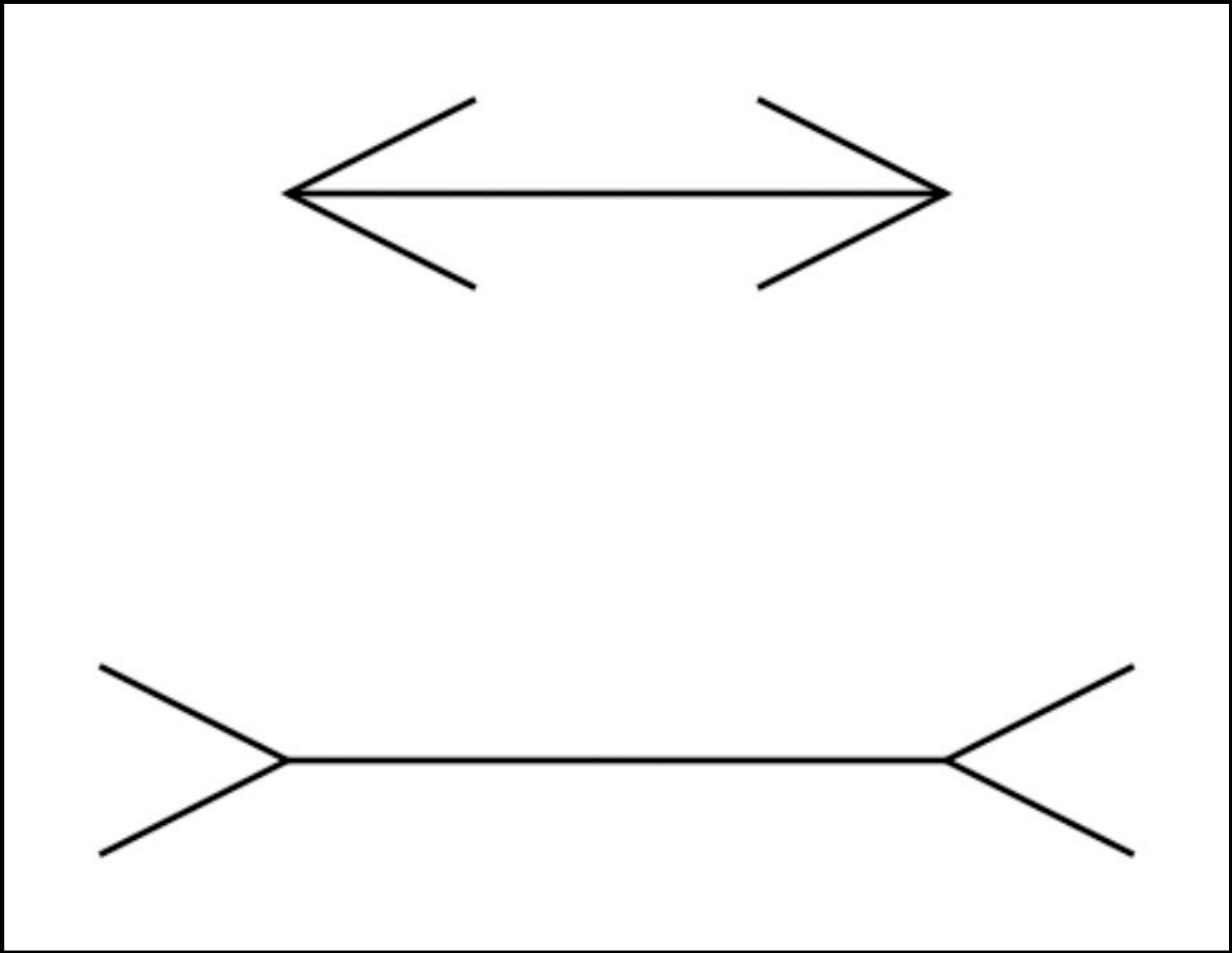




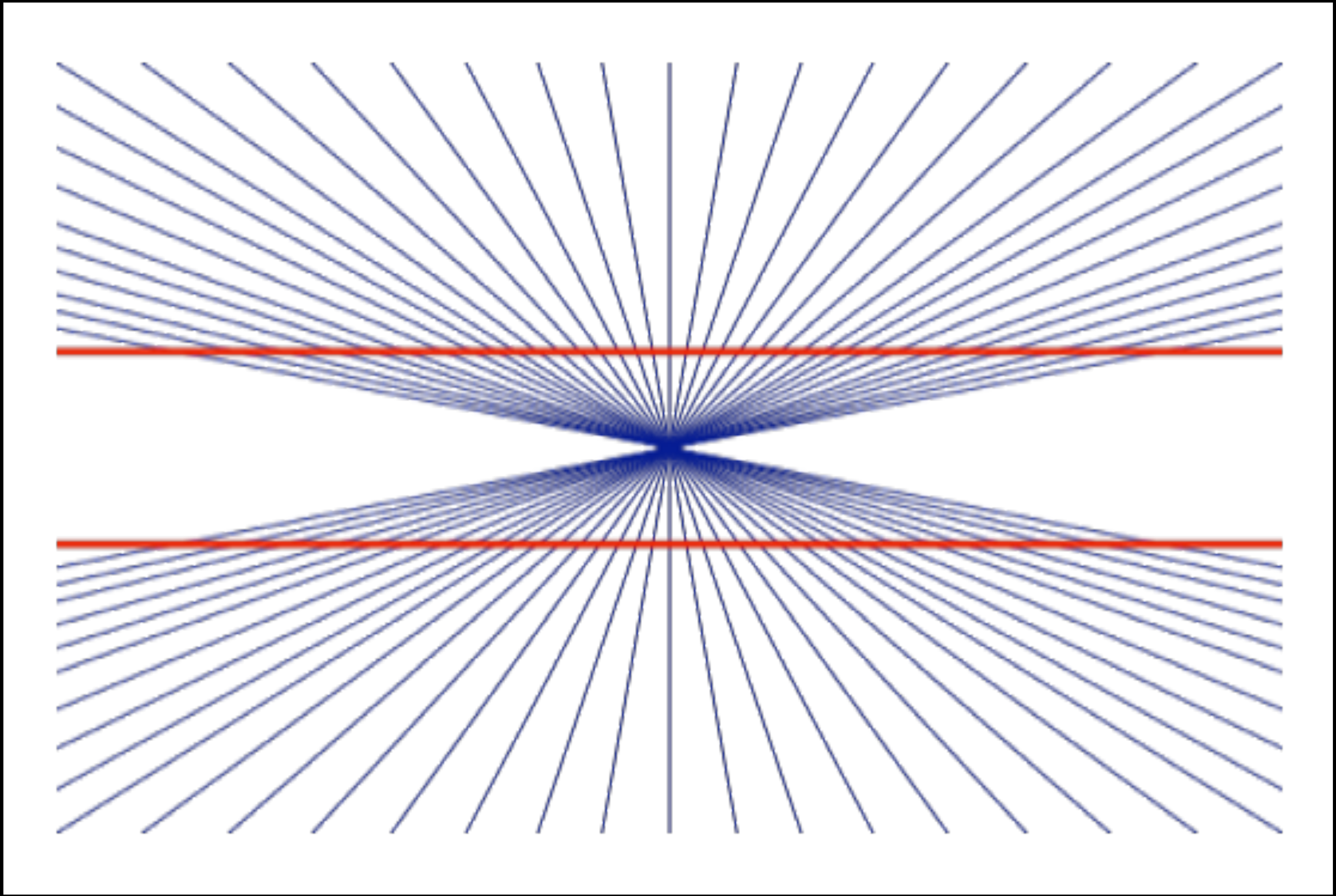


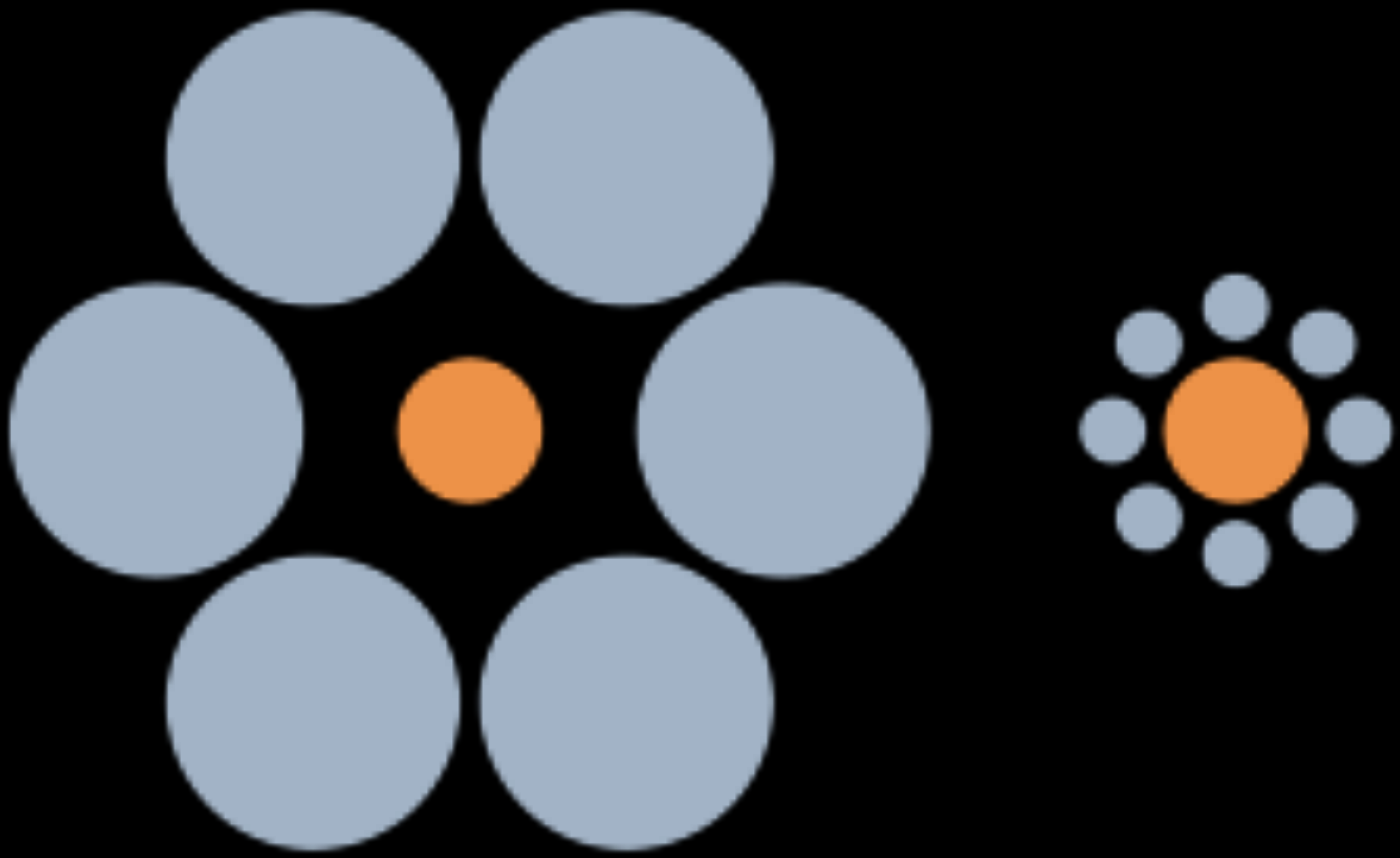


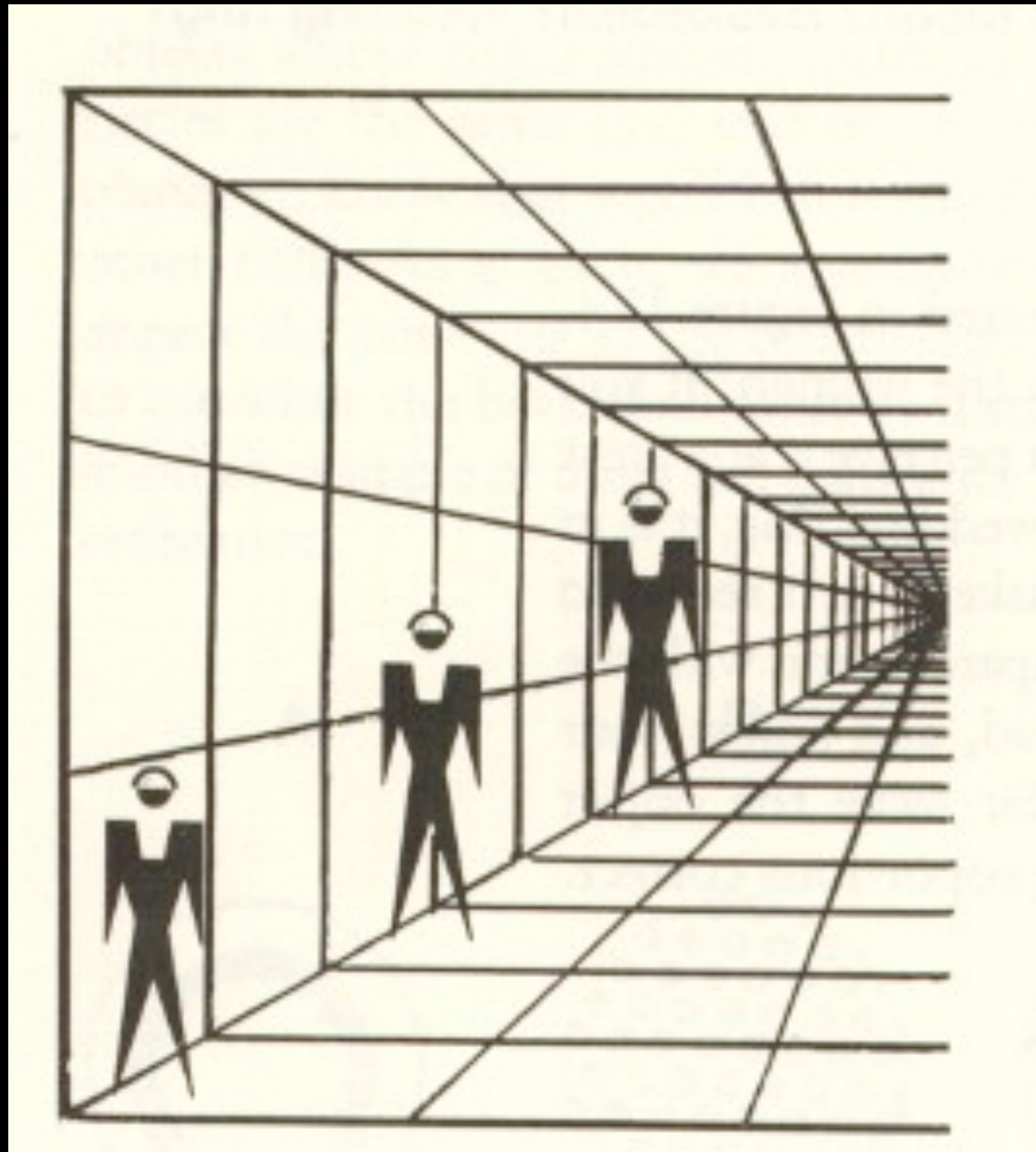
From Fairchild, *Color Appearance Models*

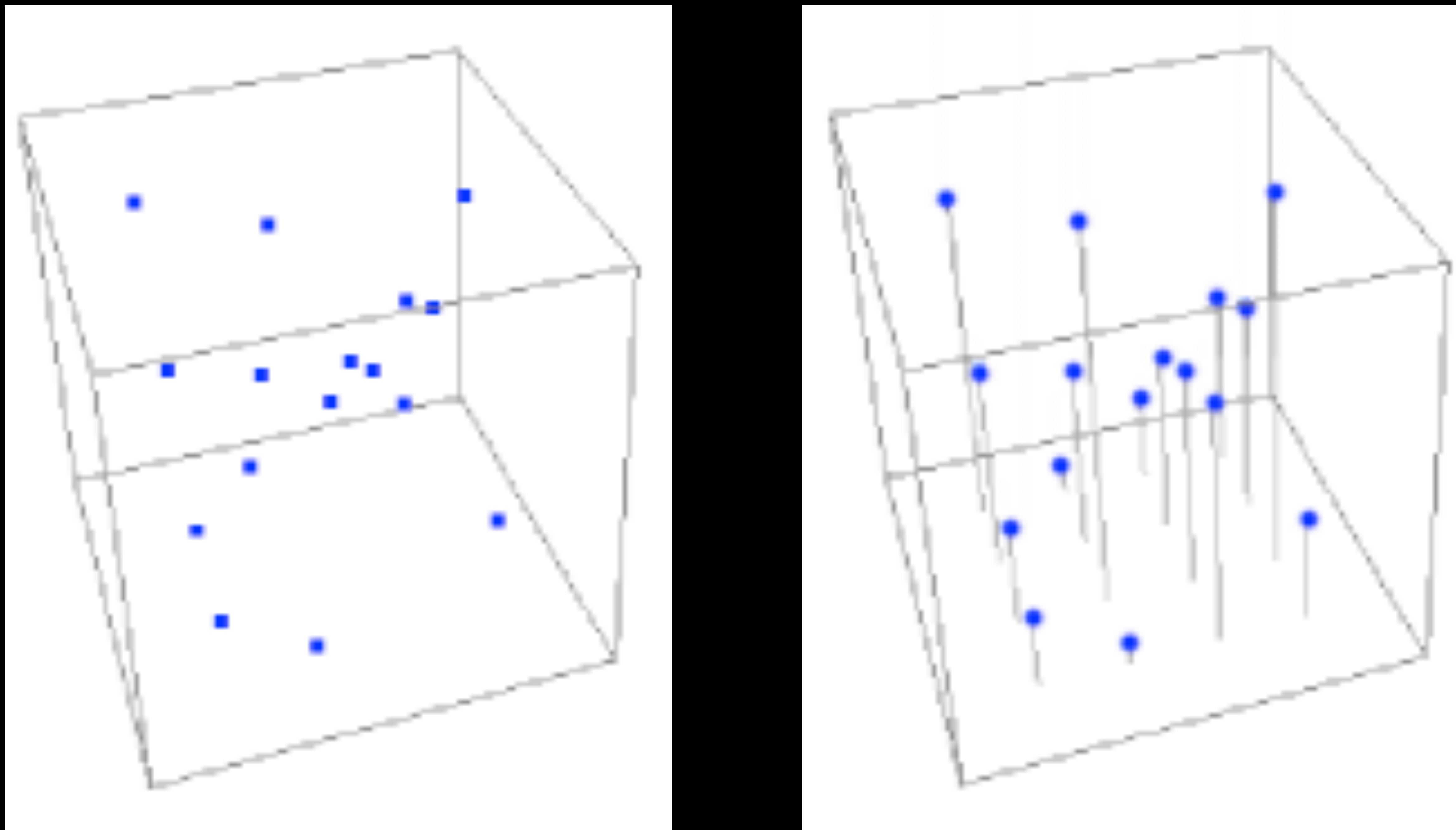


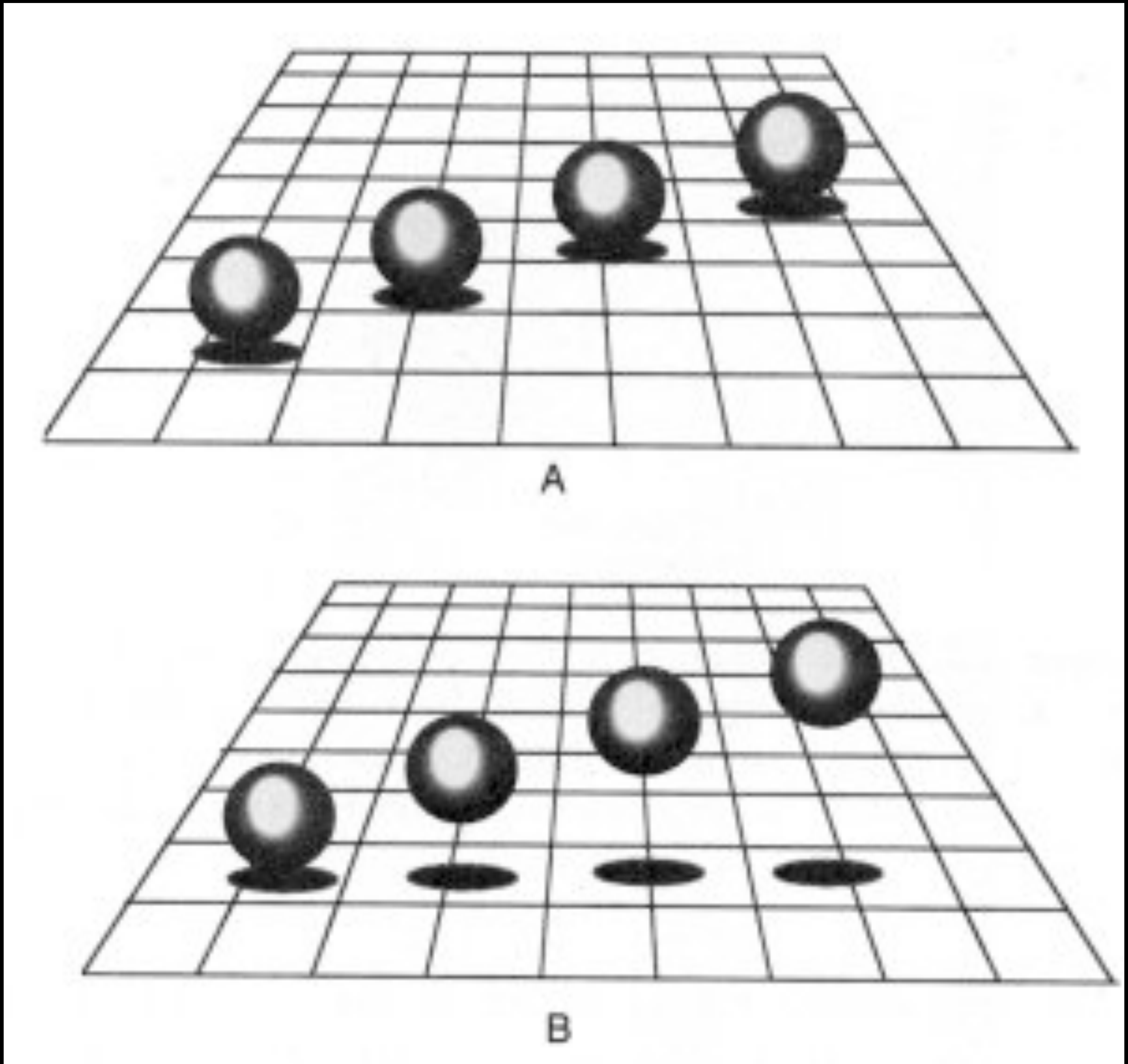
Brain Bugs





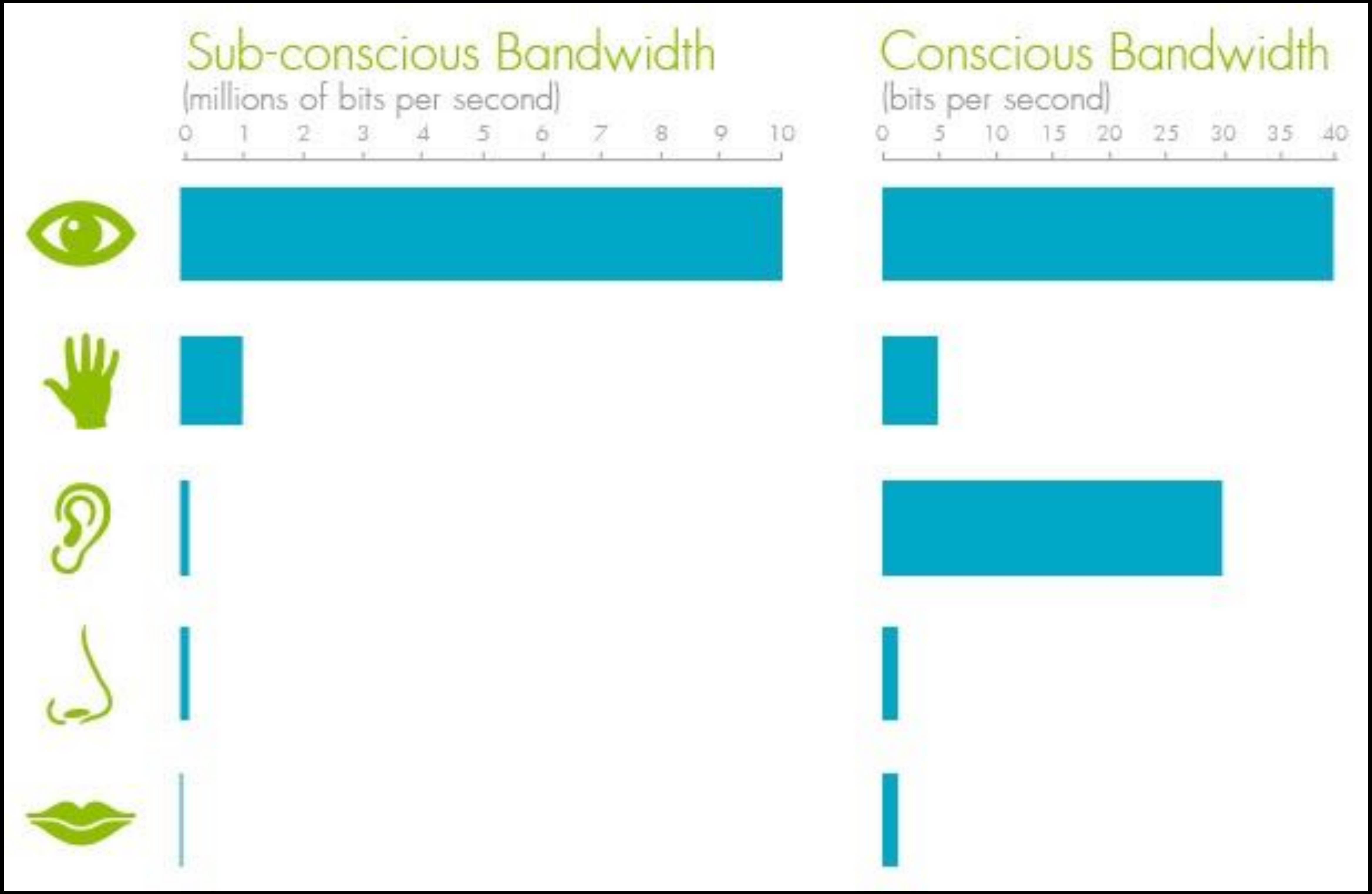





A
B

VISUAL PERCEPTION & COGNITION

WHY RELYING ON VISUALS?



[Source: http://www.creativebloq.com/design/science-behind-data-visualisation-8135496]

VISUAL PERCEPTION & COGNITION

KNOWING HOW WE PERCEIVE … TO BETTER REPRESENT

`

Temporal perception

- reaction time: 200ms to initiate a conscious observation
- stimuli <100ms apart are not perceived

Visual acuity

- Lines can be detected from 0.5″
- Distance between two lines from 30″ to 1′

Implications for Information Visualisation

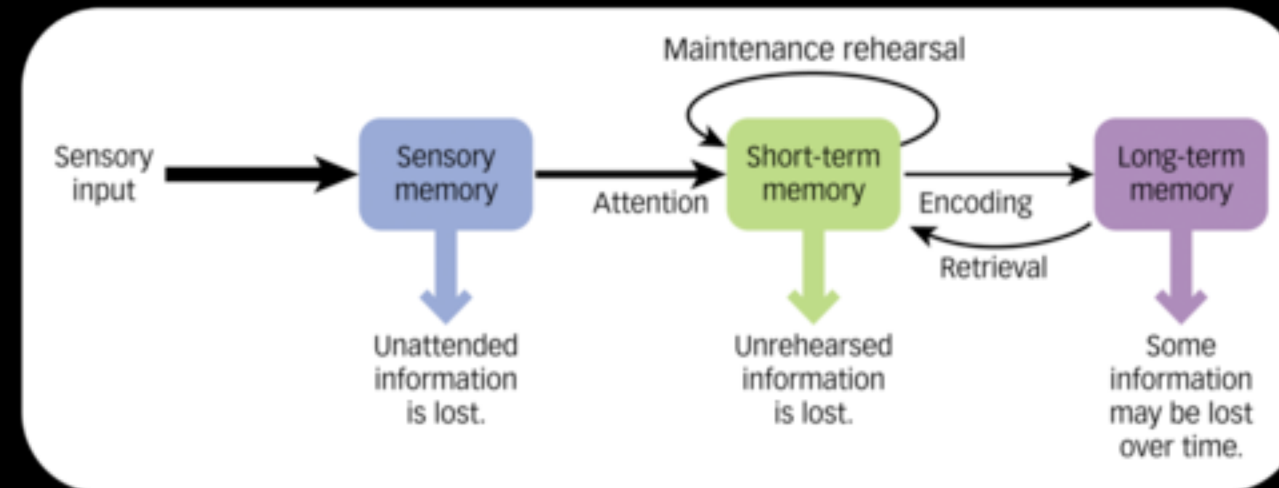
- Animations should have cycles > 1/10° seconds
- Large datasets: guarantee that the data displayed remains above limits

arcminute (′) = 1/60 of one degree. arcsecond (″) : 1/60 of one arcminute

COGNITION & VISION: COGNITIVE LOAD

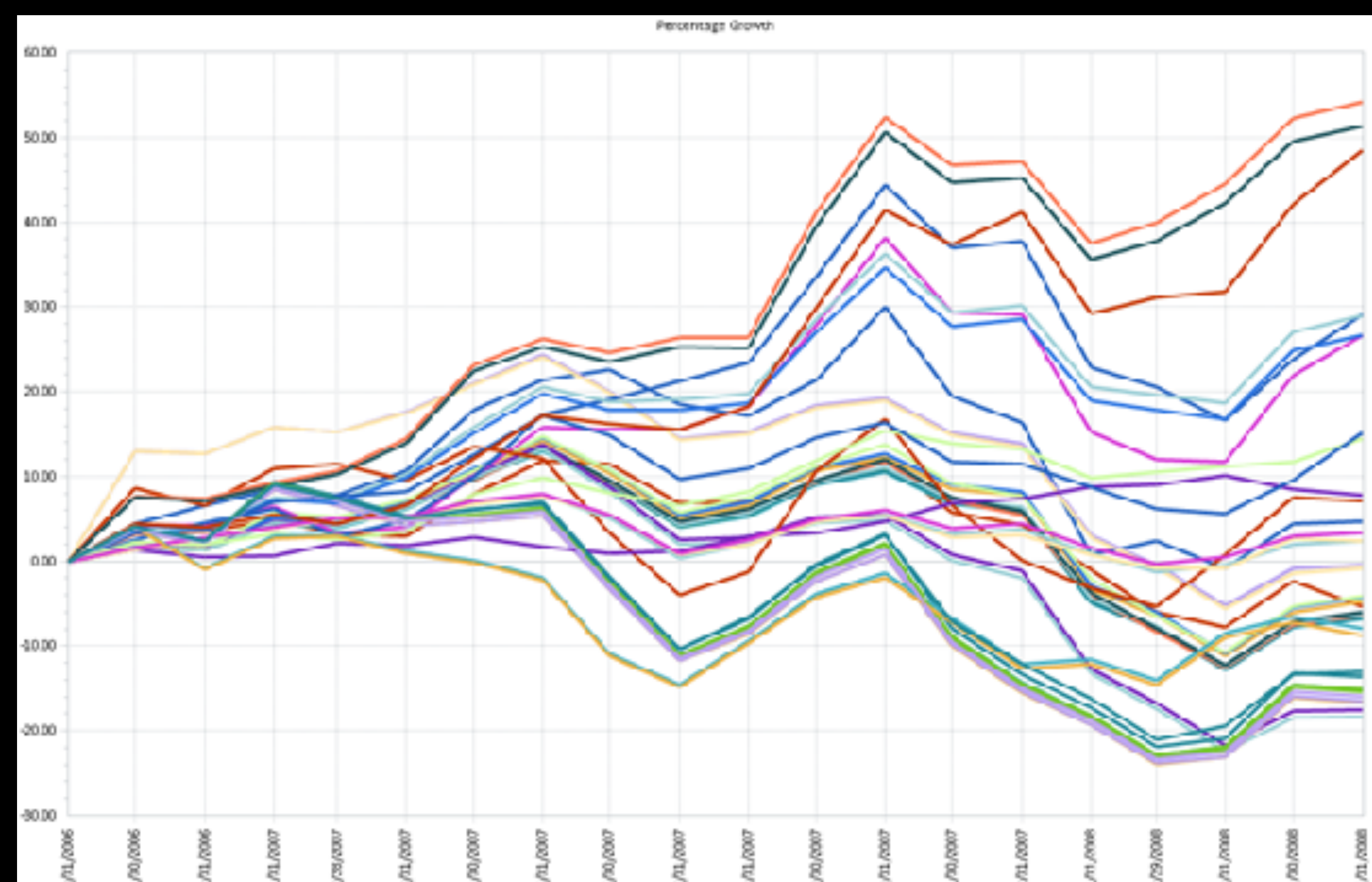
Problems with cognitive load

- **short-term memory** = working memory
- **memory span** is limited : [Miller 1956]
  7 ± 2 *independent* memorable items
- Critical for visualization of large datasets



Implications for Infovis

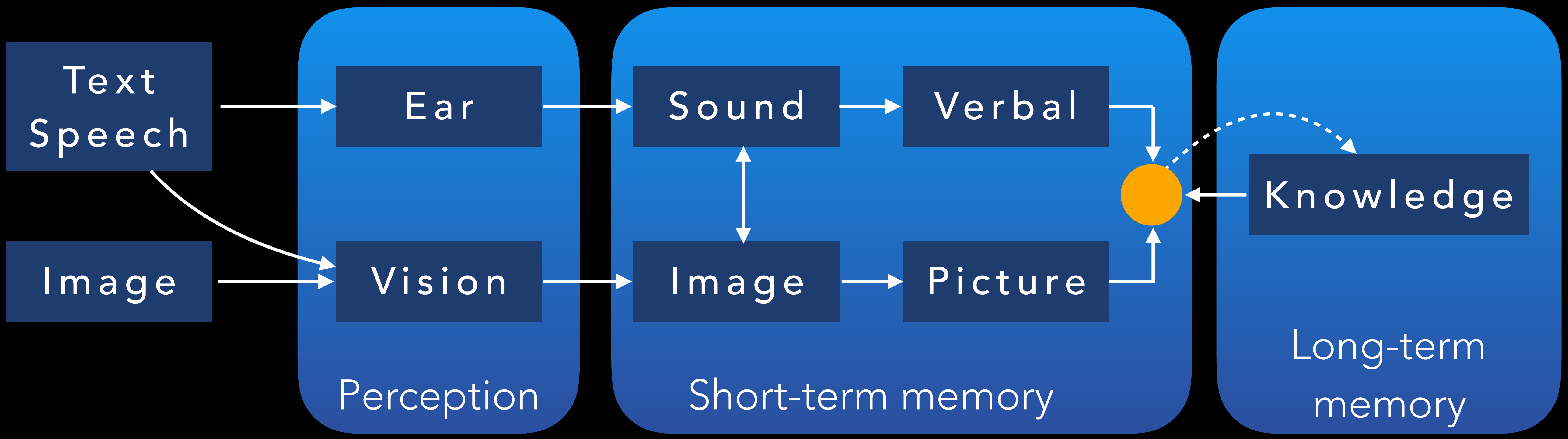
- Group, aggregate in ***chunks*** (analogy: phone numbers)
- Never require to compare more than 3 independent elements at a time



COGNITION & VISION: COGNITIVE LOAD

Cognitive load : workarounds

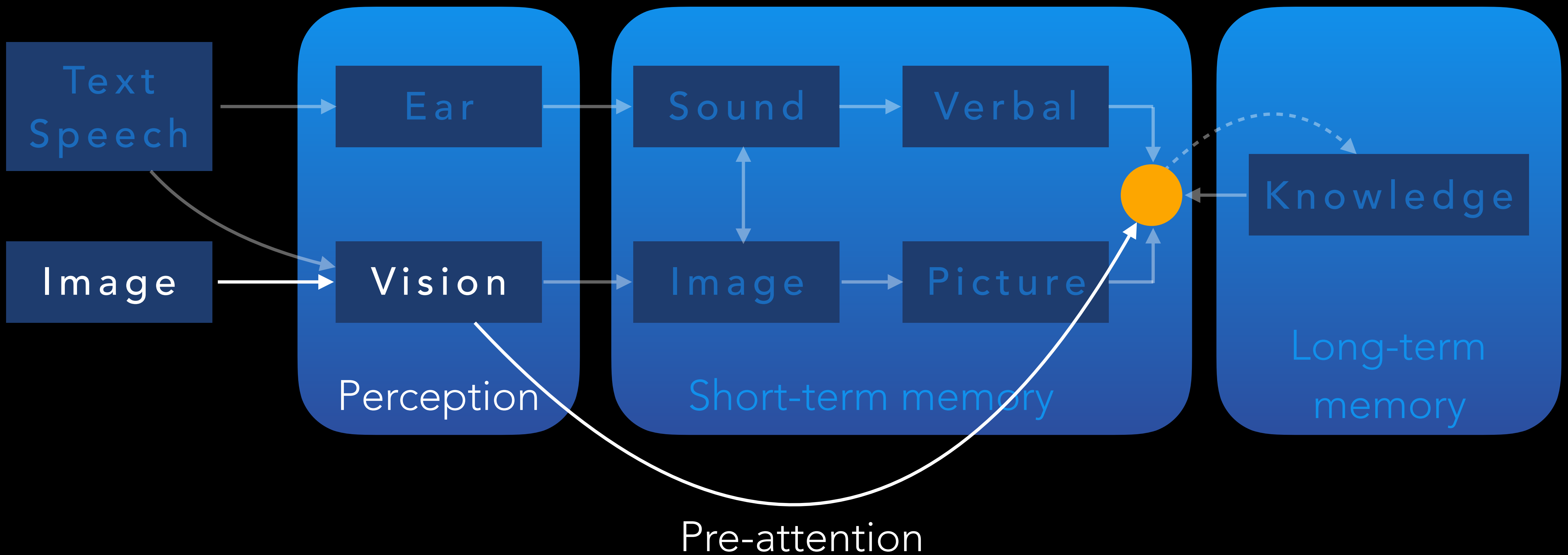
- **Multimodality** to limit cognitive overload
  Multimodality —> different cognitive pathways (i.e. visuospatial sketch pad, phonological loop, episodic buffer) [Baddeley A., Wilson B.A., 2002]



- **Pre-attentive perception**

PRE-ATTENTIVE PERCEPTION [Treisman & Gormican, 1988]

- Some visual features are processed pre-attentively, e.g. without focusing attention
- Low-level (unconscious) cognitive processes
- Reduced reaction time: <200ms (eyes movement > 200ms)
- Witness of our evolutionary story



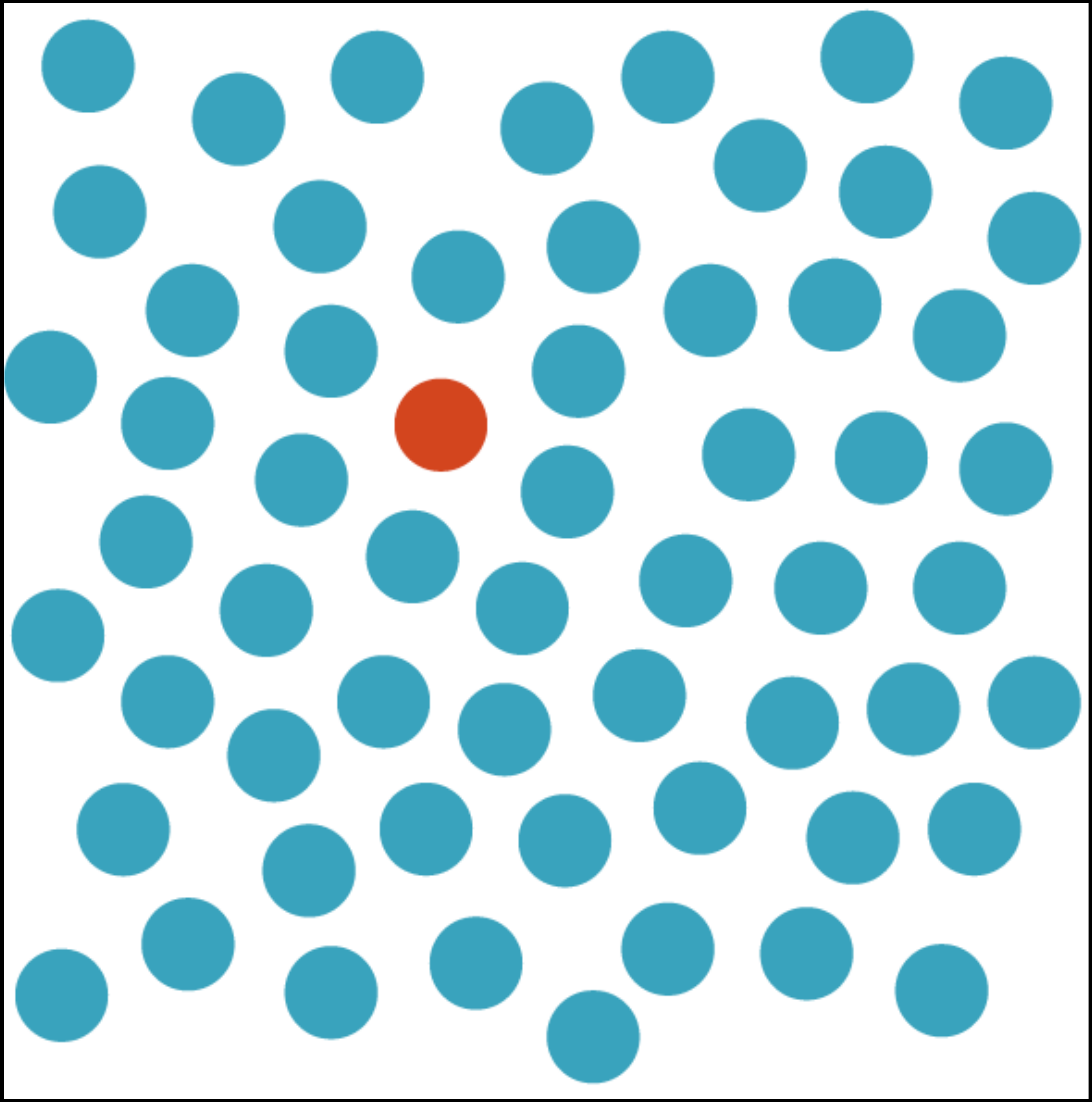
PRE-ATTENTIVE PERCEPTION [Treisman & Gormican, 1988]

- Some visual features are processed pre-attentively, e.g. without focusing attention

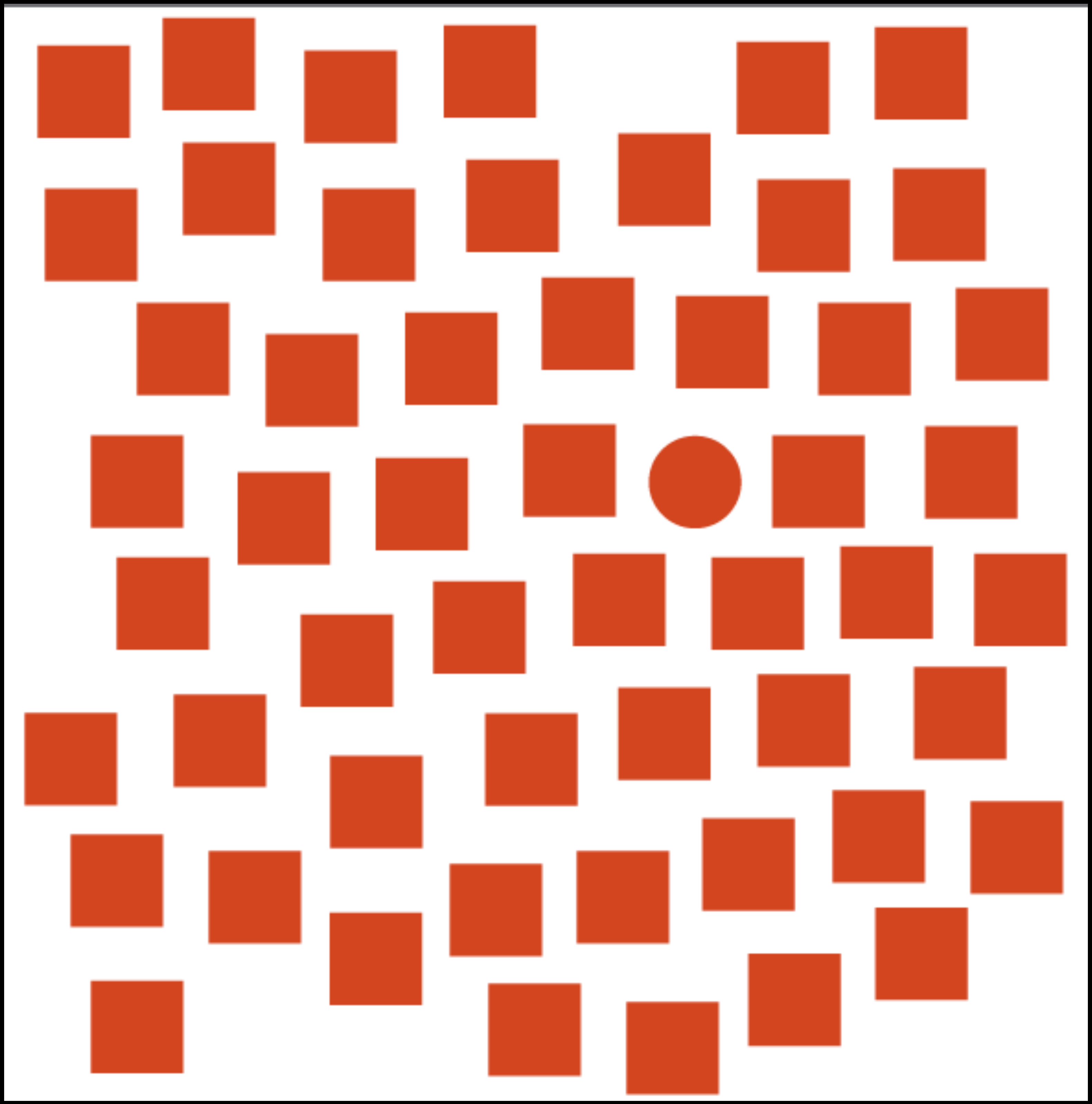
Implications for Information Visualisation

- No cognitive load
- Direct processing : what must be perceived immediately

Is there a red circle?

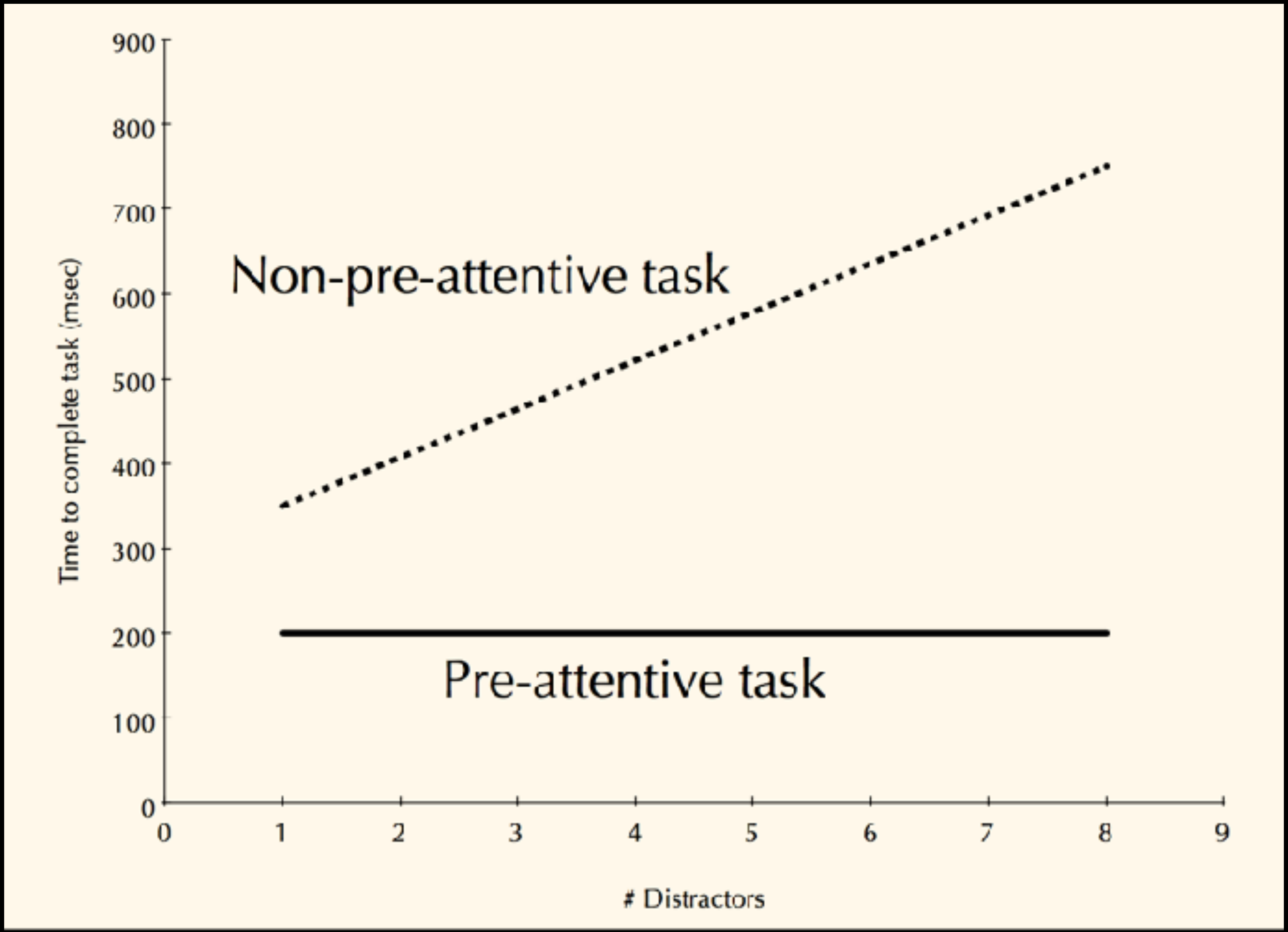


Is there a red circle?



PRE-ATTENTIVE PERCEPTION

Takes the same amount of time, regardless of the number of distractors



[Slide inspired by J. Wirz & S. Bedrid course material]

Find the 3's

1424164963575984759217659684748917284 82

2859588198294509685048506958476121240 44

0746748989851714959691245676596080208 60

6083654164964575906439804792485769607 81

2859607999187128452681014959691245677 81

8742416496457576596081495969124567012 85

9607991649645751278799187128452984969 12

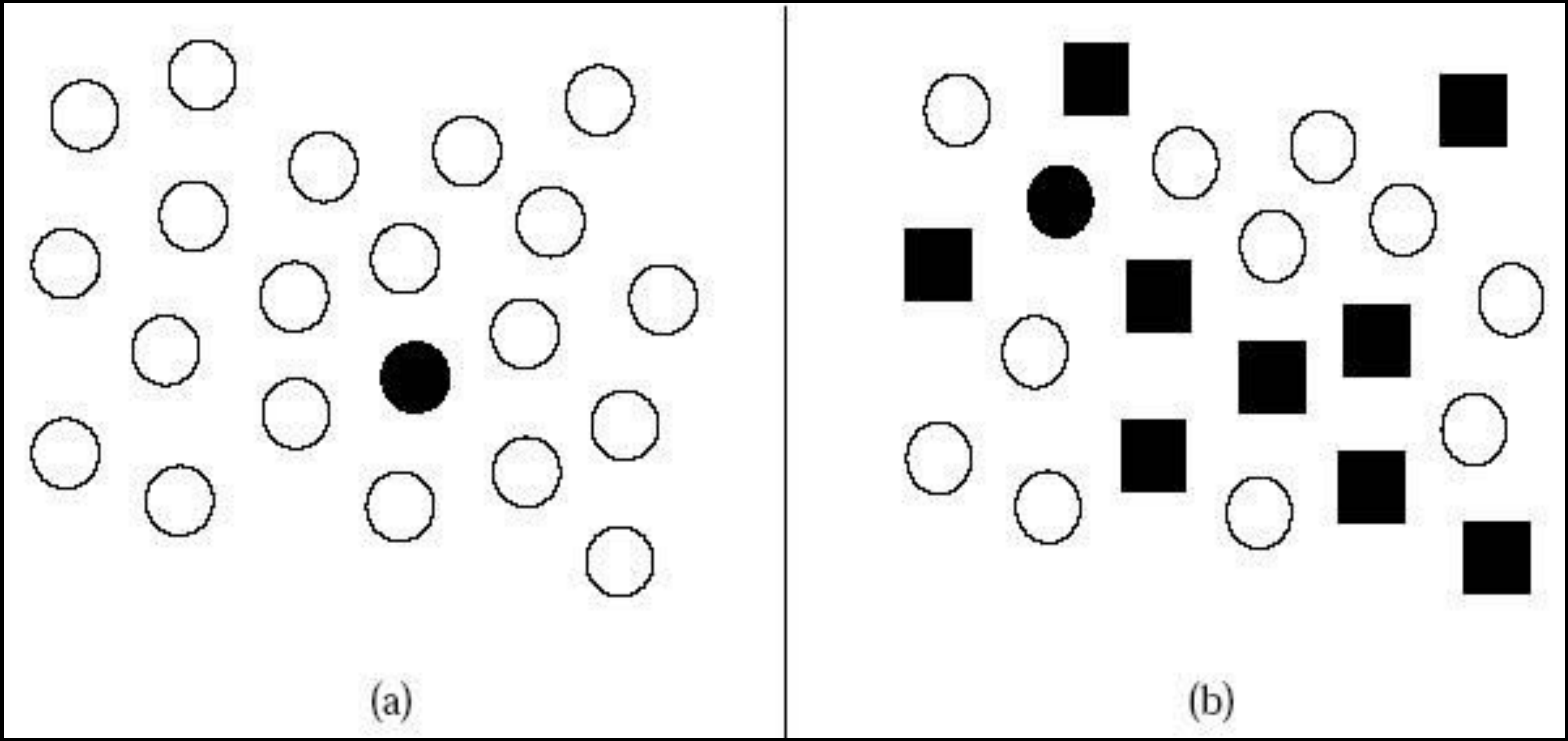
2235916496457595881982509635765960805 96

Find the 3's

1424164963575984759217659684748917284 82
2859588198294509685048506958476121240 44
0746748989851714959691245676596080208 60
6083654164964575906439804792485769607 81
2859607999187128452681014959691245677 81
8742416496457576596081495969124567012 85
9607991649645751278799187128452984969 12
2235916496457595881982509635765960805 96

PRE-ATTENTIVE PERCEPTION

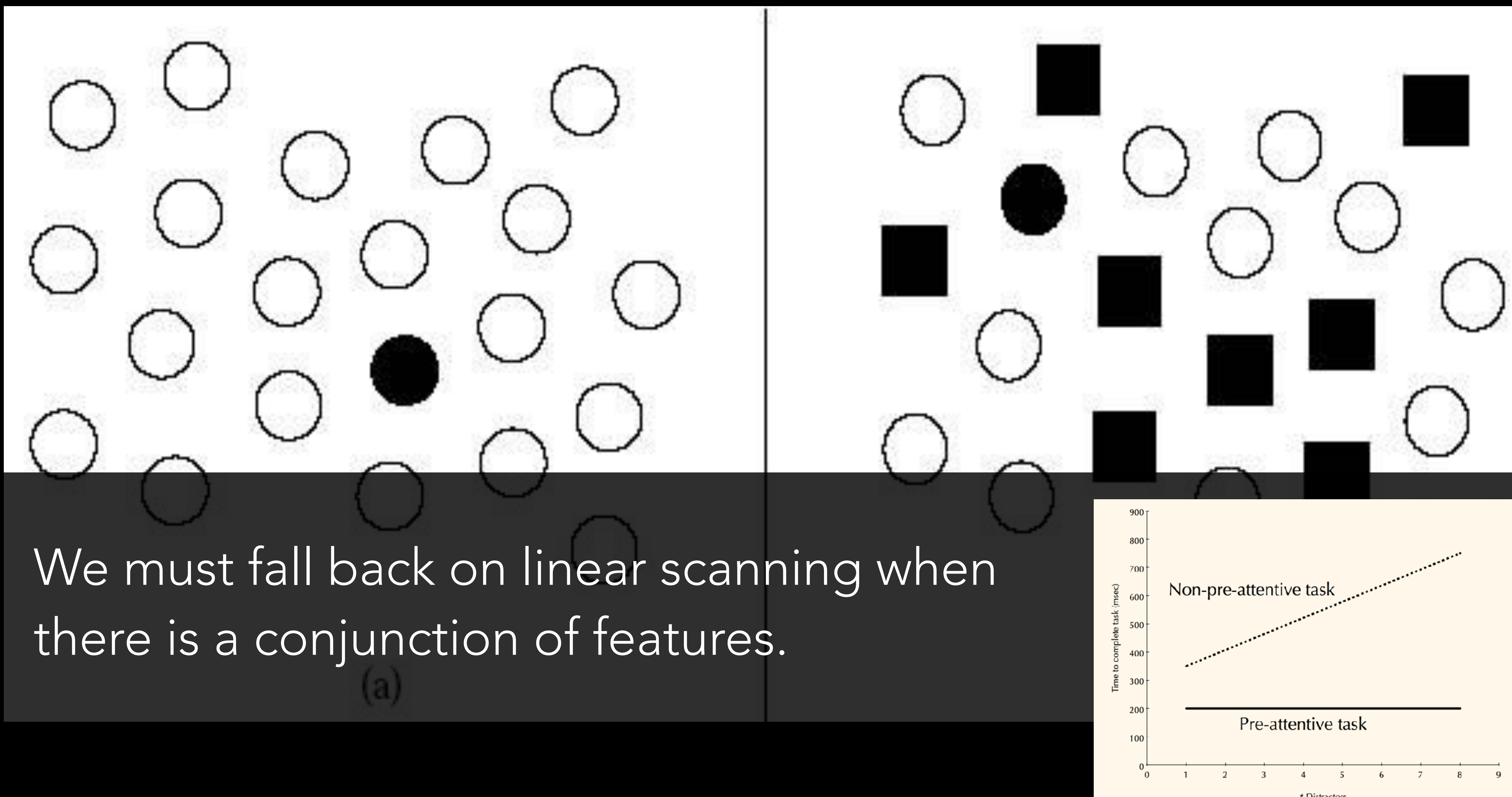
Only works when the distractors differ from one feature:



[Healey et al., 1996]

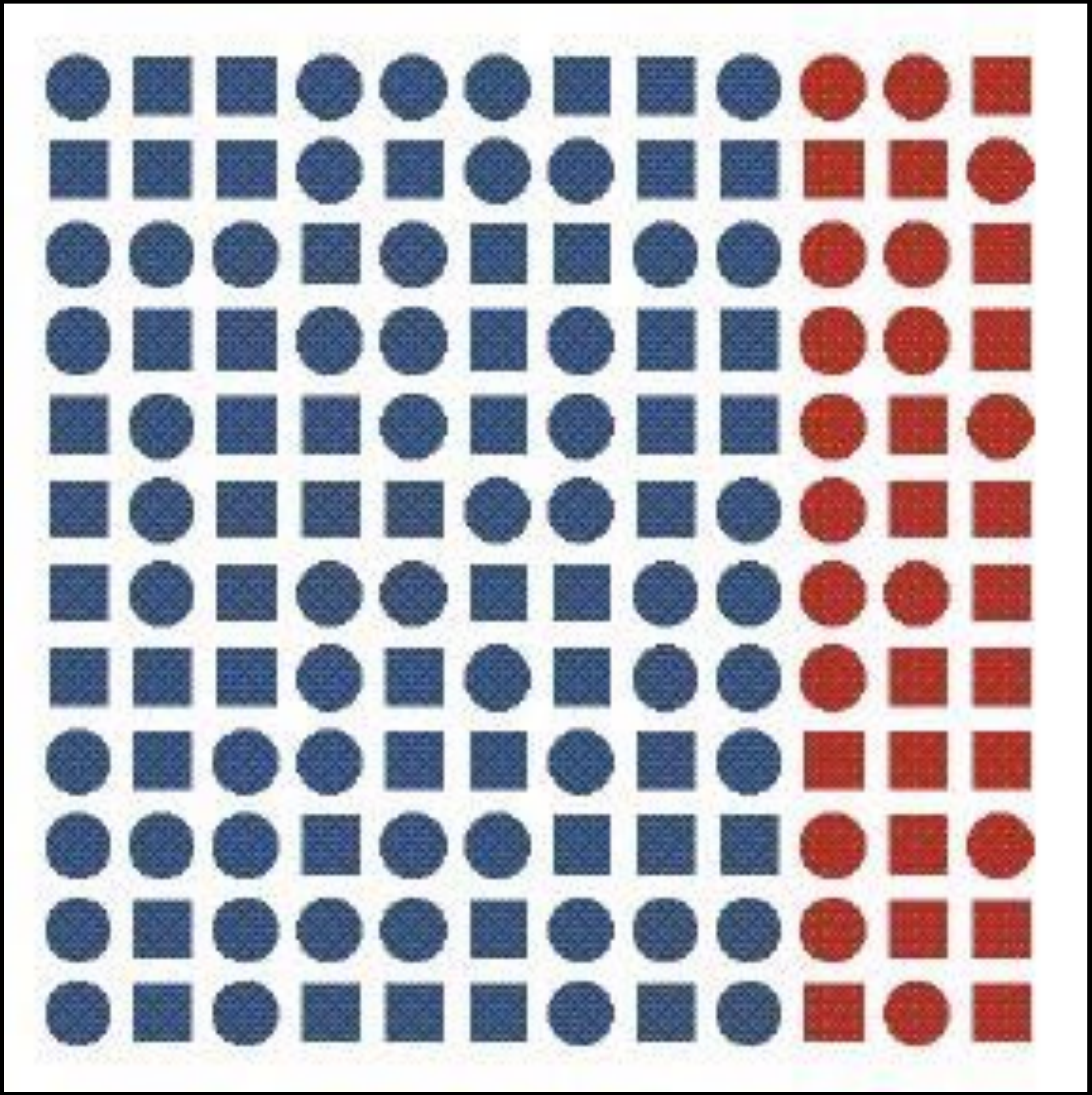
PRE-ATTENTIVE PERCEPTION

Only works when the distractors differ from one feature:

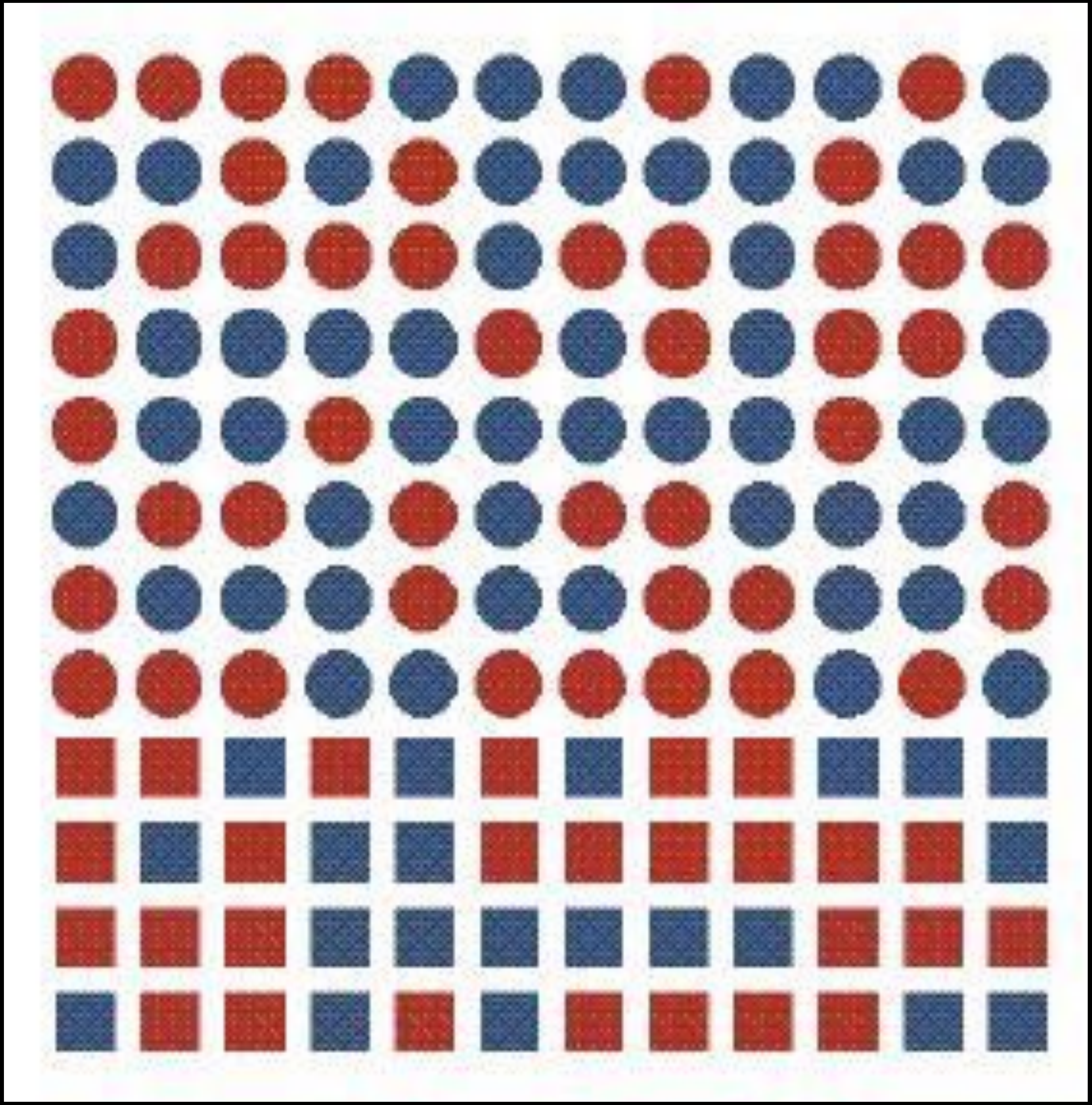


We must fall back on linear scanning when there is a conjunction of features.

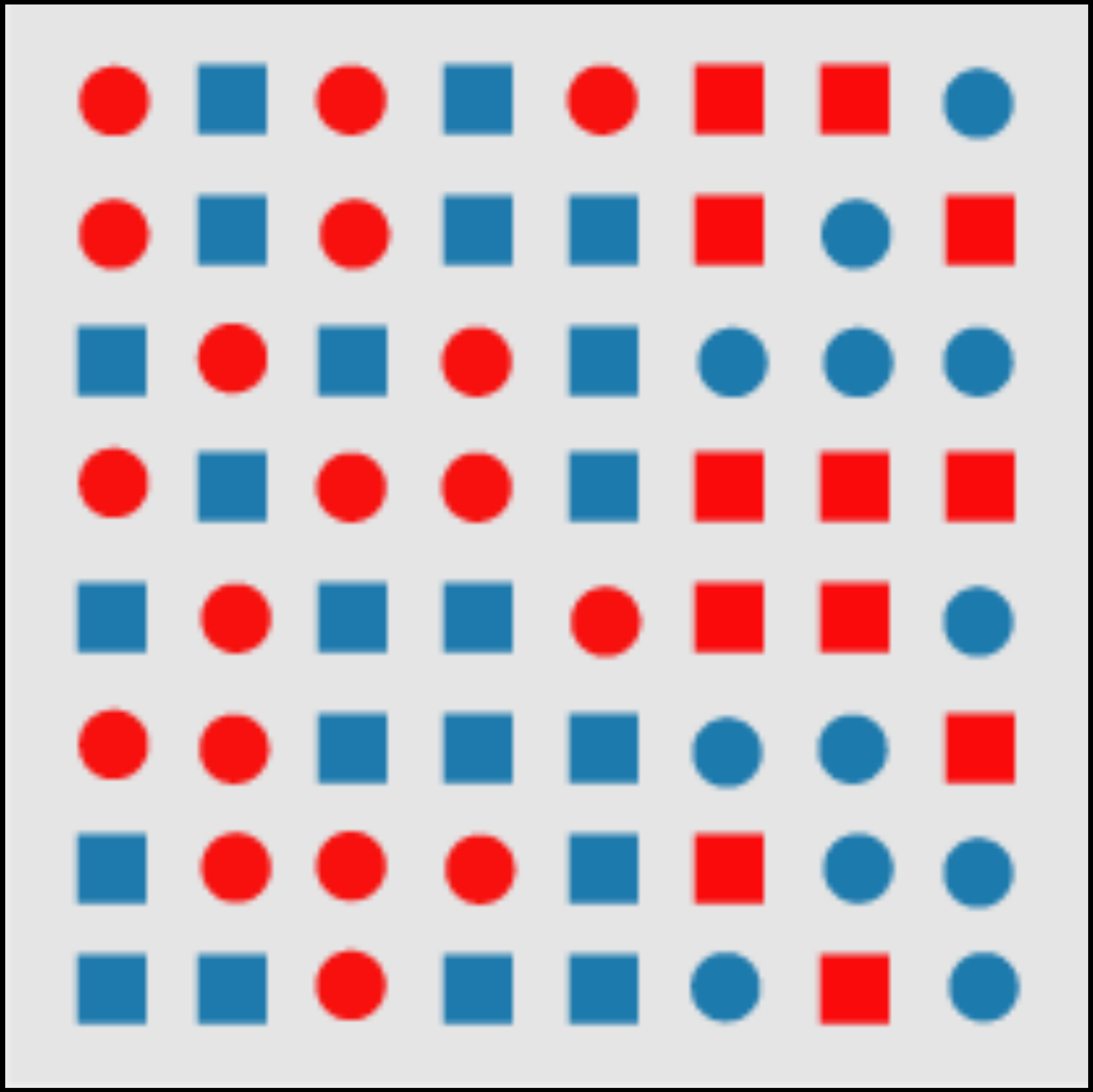
Is there a boundary?



Is there a boundary?

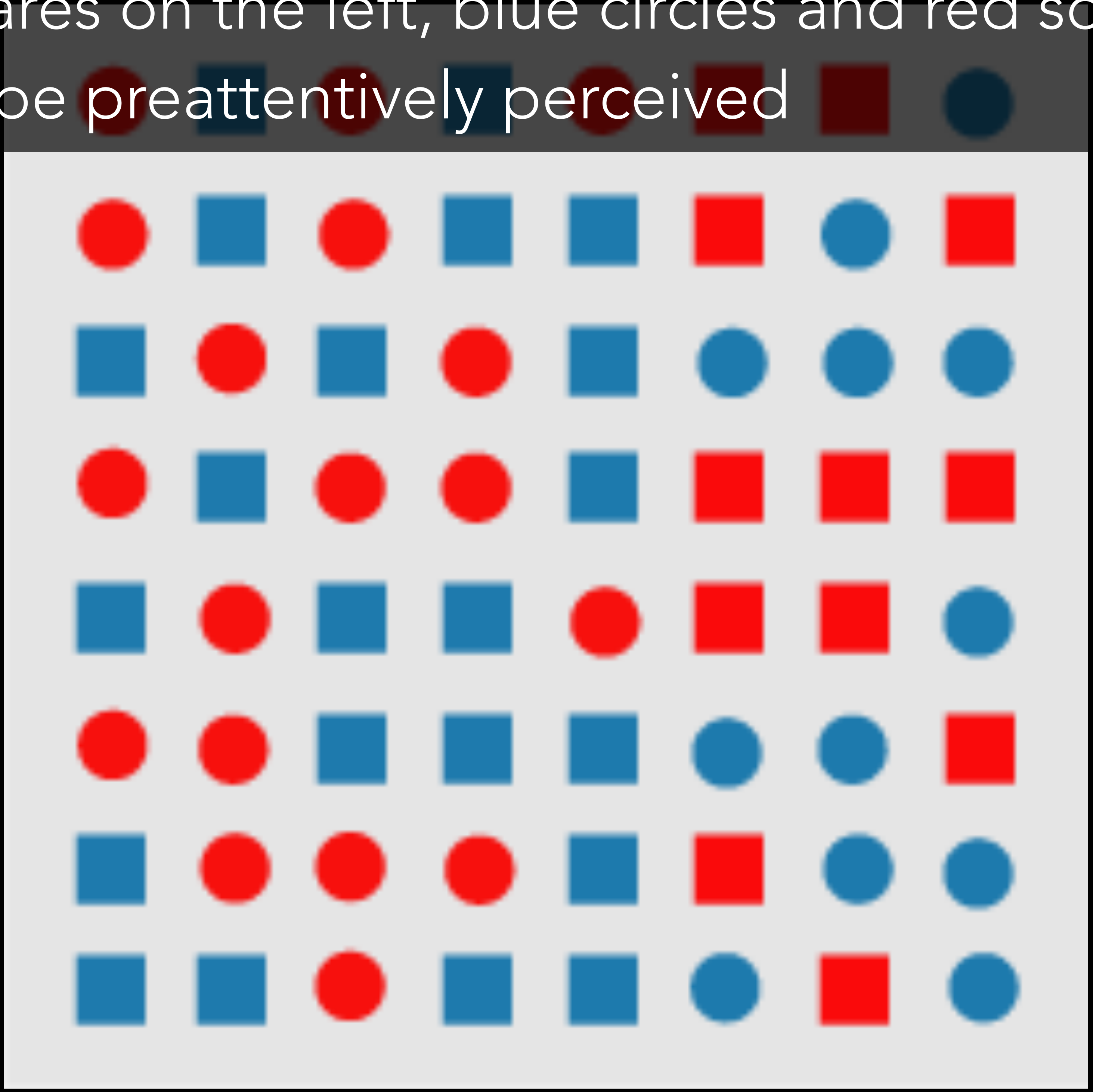


Is there a boundary? (hint: YES!)



Taken from https://www.csc.ncsu.edu/faculty/healey/PP/index.html

a boundary defined by a conjunction of features (here red circles and blue squares on the left, blue circles and red squares on the right) cannot be preattentively perceived

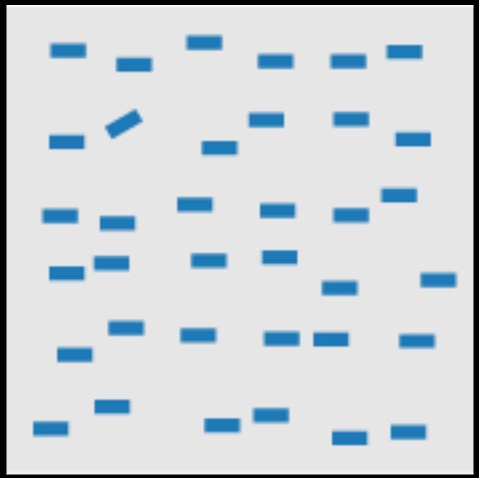


Taken from https://www.csc.ncsu.edu/faculty/healey/PP/index.html

(SOME) PRE-ATTENTIVE VISUAL FEATURES

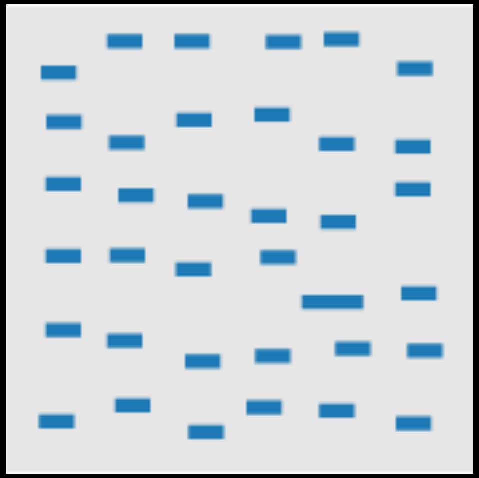
orientation

[Julész & Bergen 83]
[Sagi & Julész 85]
[Wolfe et al. 92]
[Weigle et al. 2000]



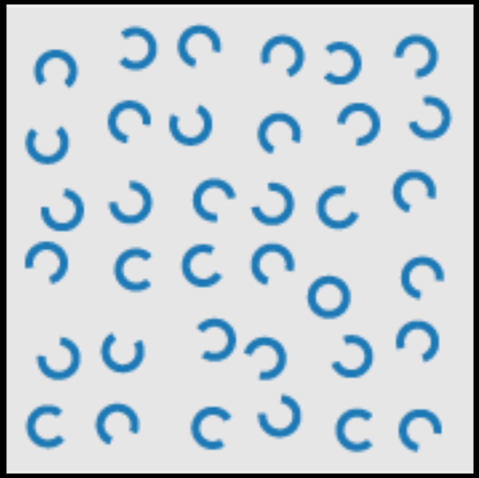
length, width

[Sagi & Julész 85]
[Treisman & Gormican 88]



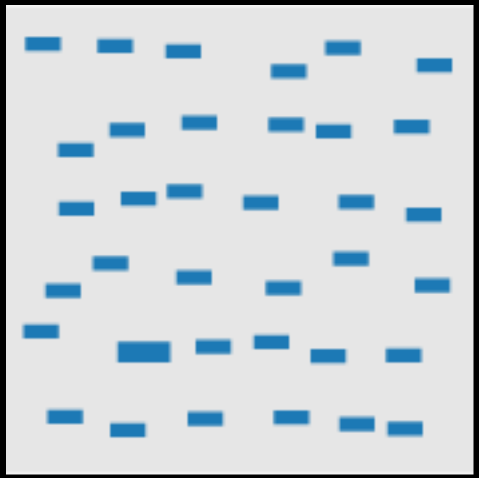
closure

[Julész & Bergen 83]



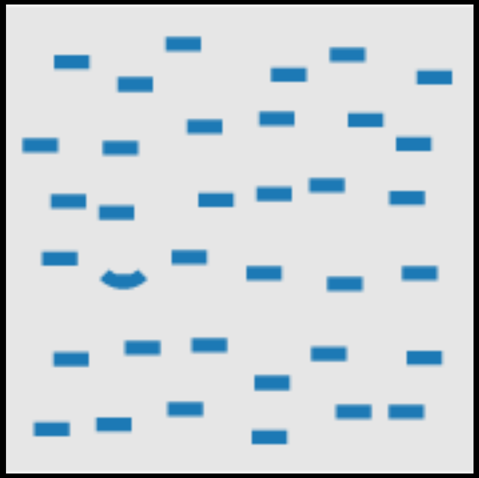
size

[Treisman & Gelade 80]
[Healey & Enns 98]
[Healey & Enns 99]



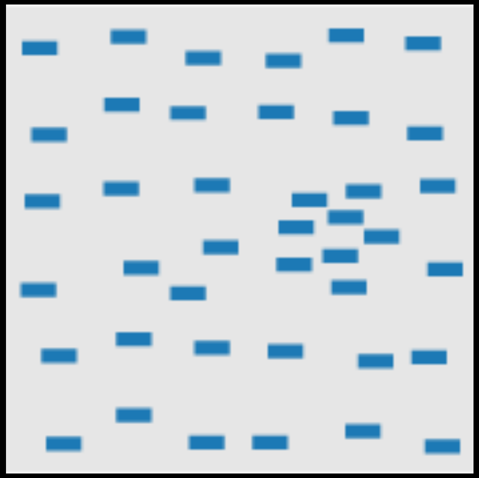
curvature

[Treisman & Gormican 88]



density, contrast

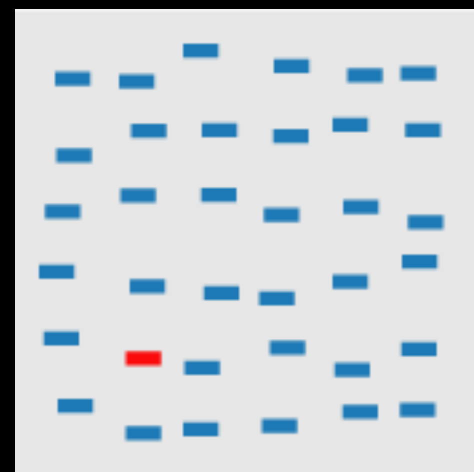
[Healey & Enns 98]
[Healey & Enns 99]



(SOME) PRE-ATTENTIVE VISUAL FEATURES

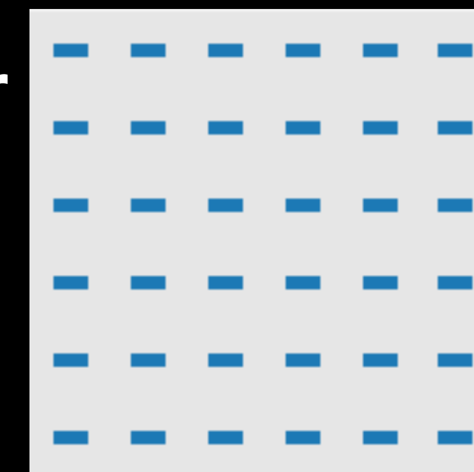
color, hue

Nagy & Sanchez 90; Nagy et al. 90; D'Zmura 91; Kawai et al. 95; Bauer et al. 96; Healey 96; Bauer et al. 98; Healey & Enns 99



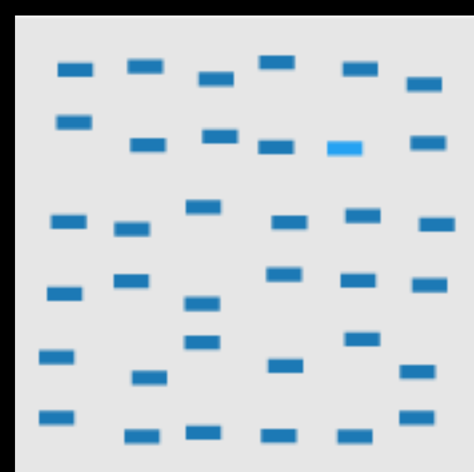
flicker

[Gebb et a. 55; Mowbray & Gebhard 55; Brown 65; [Julész 71]
[Huber & Healey 2005]



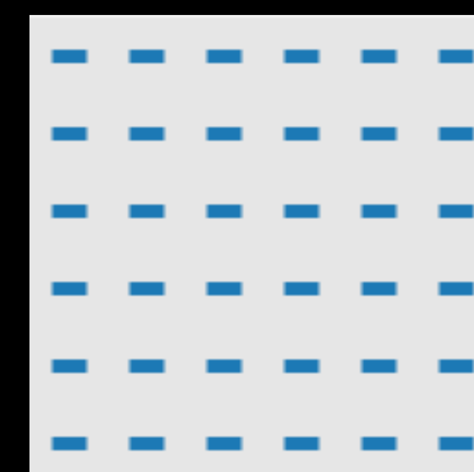
intensity

[Beck et al. 83]
[Treisman & Gormican 88]
[Wolfe & Franzel 88]



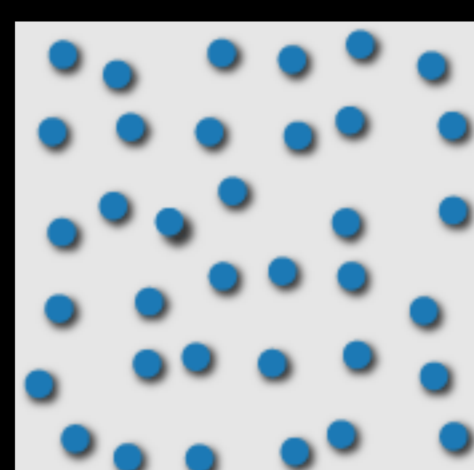
direction of motion

[Nakayama & Silverman 86; Driver & McLeod 92; Huber & Healey 2005]



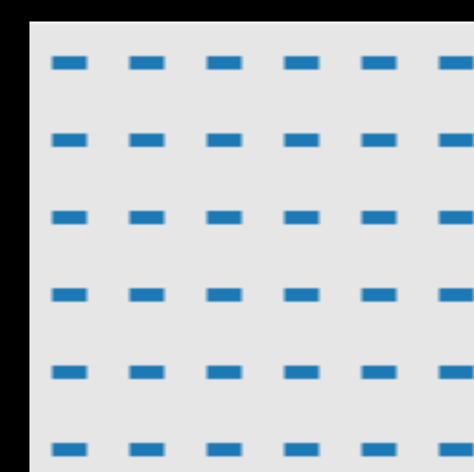
3D depth cues

[Enns 90b; Nakayama & Silverman 86]

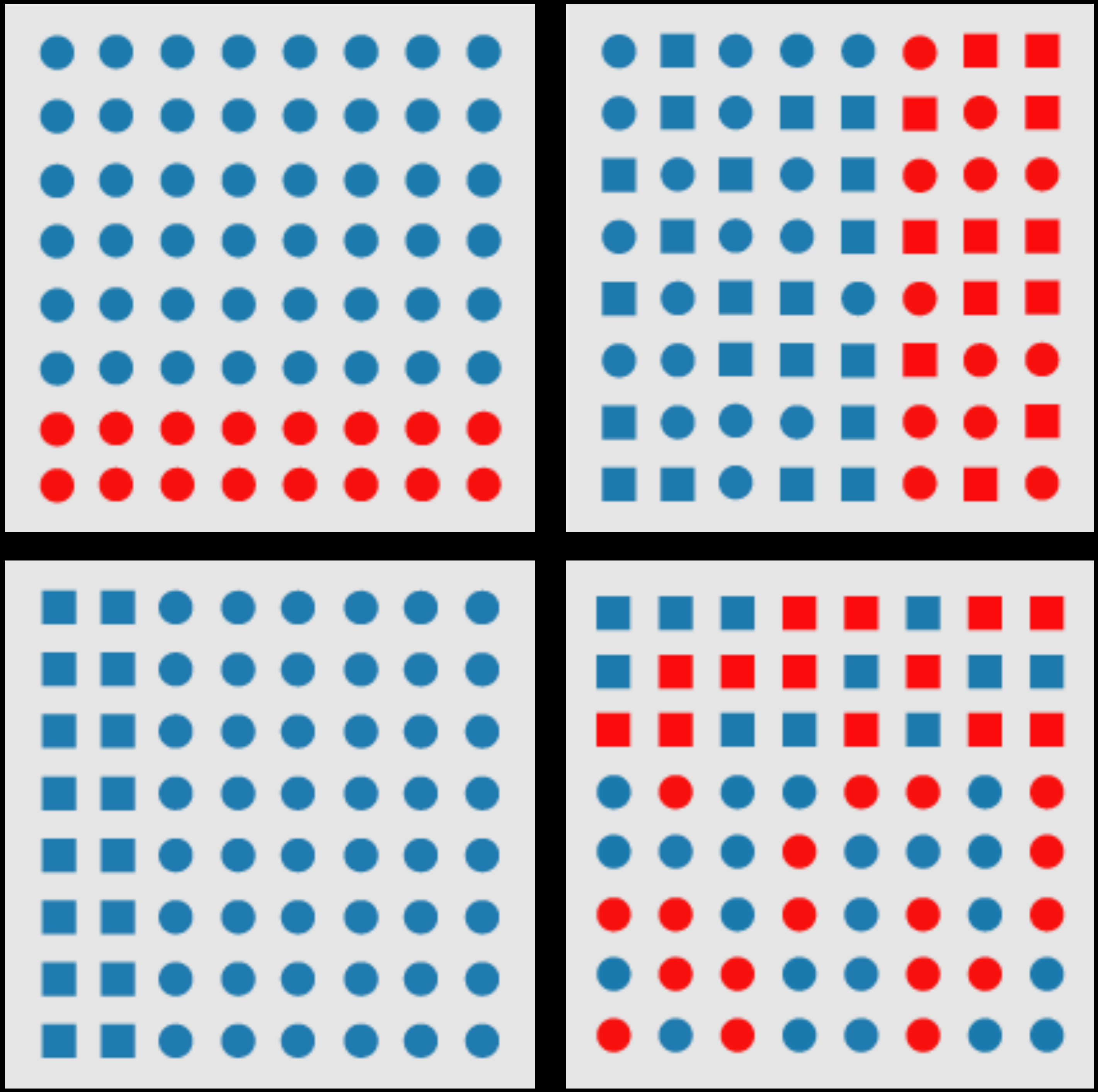


velocity of motion

[Tynan & Sekuler 82; Nakayama & Silverman 86; Driver & McLeod 92; Hohnsbein & Mateeff 98; Huber & Healey 2005]



Note that these various features are not created equal!



We seem to have a strong bias towards color perception over shape perception, etc…

Taken from https://www.csc.ncsu.edu/faculty/healey/PP/index.html

What does all of this mean?

[Slide inspired by J. Wirz & S. Bedrid course material]

1. Certain tasks that depend on pre-attentive features can sometimes be done "for free" by our brains:

Target detection *Boundary detection*
Region tracking *Counting (estimation)*

2. The more of our story we can tell using pre-attentive features, the faster and better our viewer will "get it".

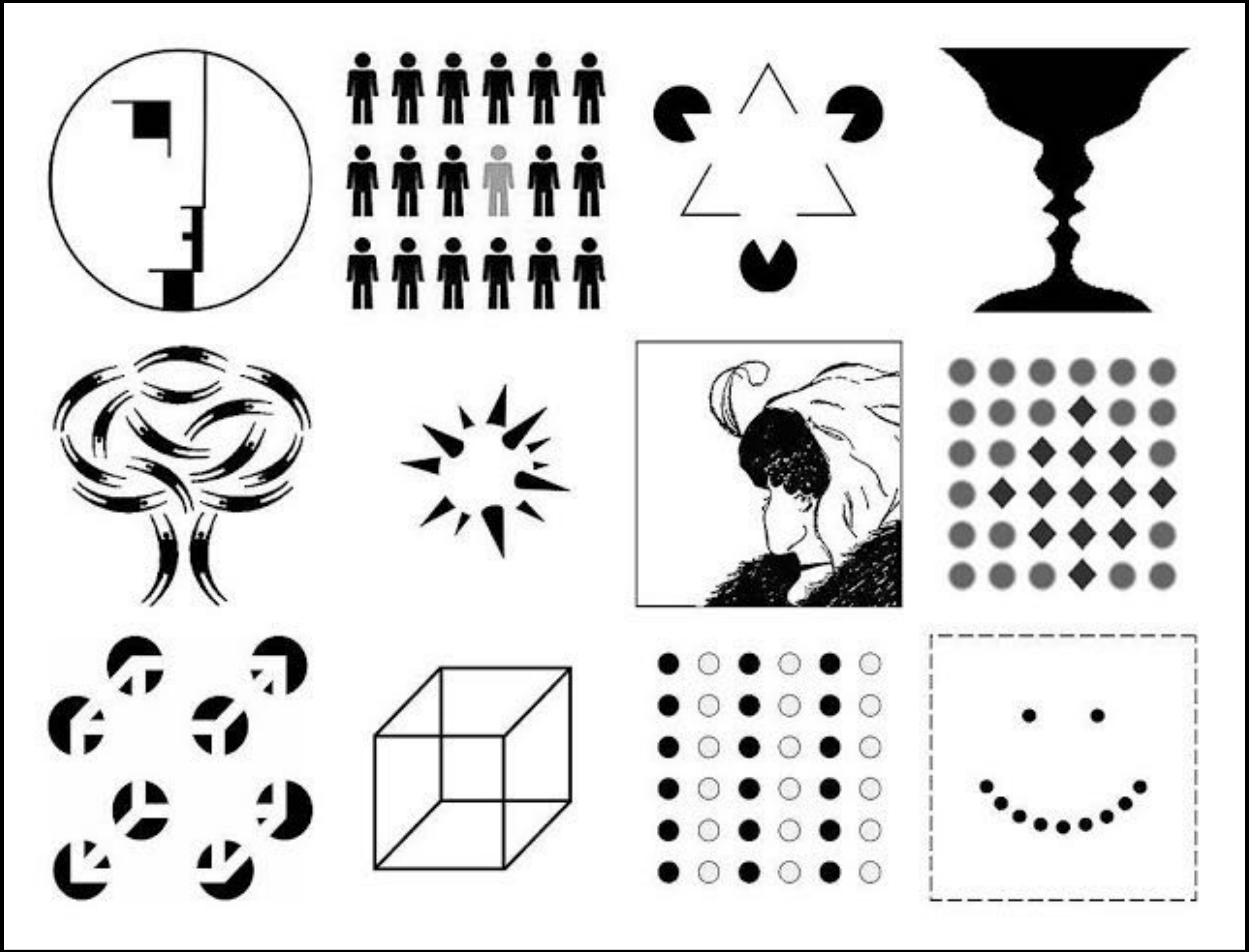
[Slide inspired by J. Wirz & S. Bedrid course material]

3. We can easily mess up our viewer's ability to interpret our visualisation by "triggering" pre-attentive perception inappropriately!

Many of the things that make a bad visualisation "bad" can be traced back to problems relative to pre-attentive processing.

[Slide inspired by J. Wirz & S. Bedrid course material]

GESTALT PSYCHOLOGY



DEFINITION

The Gestalt psychology is a **theory of perception** that is often summed up by:

"The whole is other than the sum of the parts"

— Kurt Koffka (1922)

THE BASIC IDEA:

Our brains operate less on individual points, lines, etc…

… but rather on **higher-level constructs** …

… which is what our perceptual systems are **optimised for.**

[Slide inspired by J. Wirz & S. Bedrid course material]

The Gestalt psychology notably describes the **perception of forms** by the visual system.
It relies on four **principles**:

- Emergence
- Reification
- Multistability
- Invariance

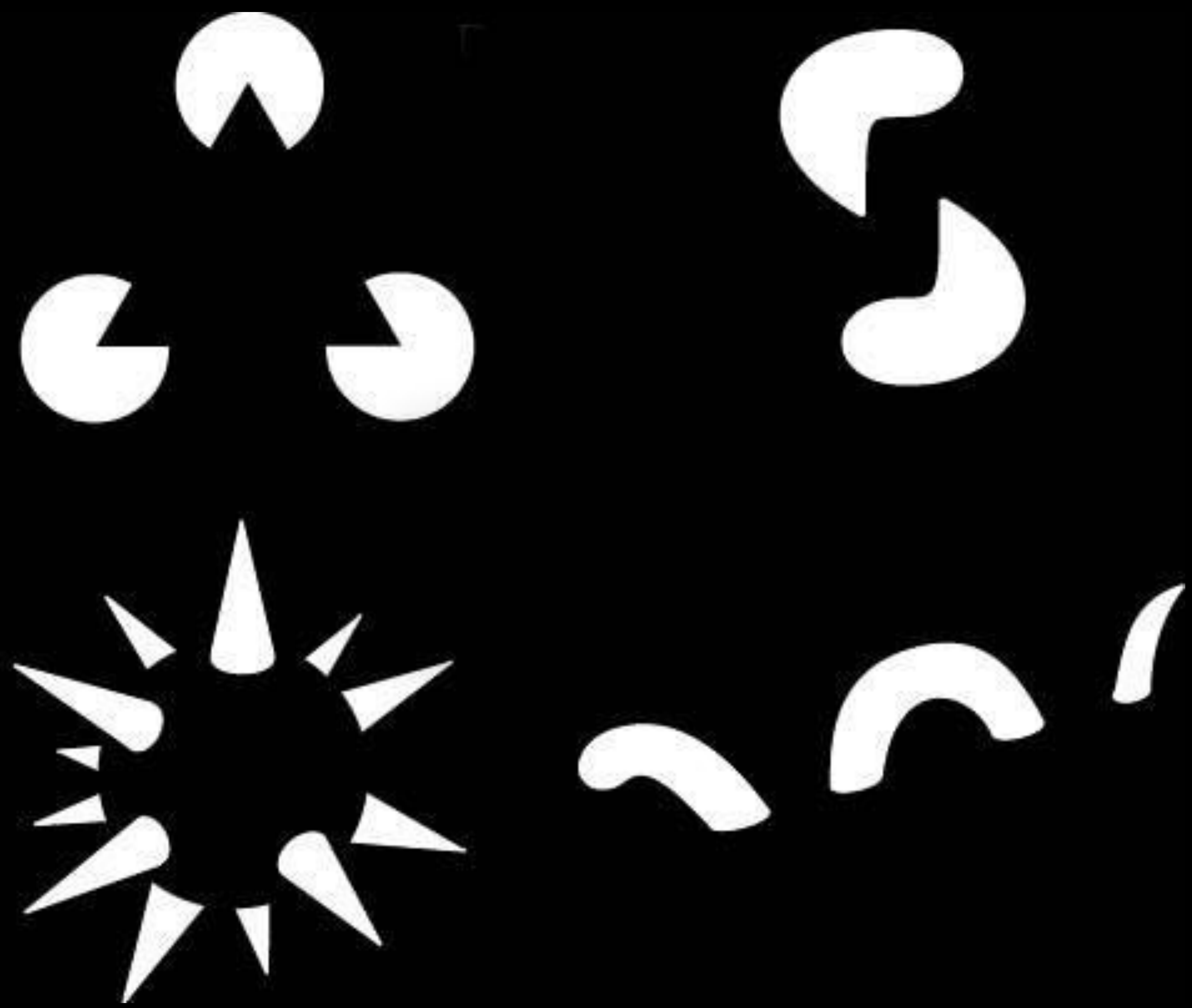
It also describes our visual perceptions by a set of **laws**.

EMERGENCE

The **global perception** can **not** be explained by the **sum of its parts**.

EMERGENCE

REIFICATION



REIFICATION

The **perception** contains **more spatial information than the stimulus** on which it is based: **part of the perception is generated**.

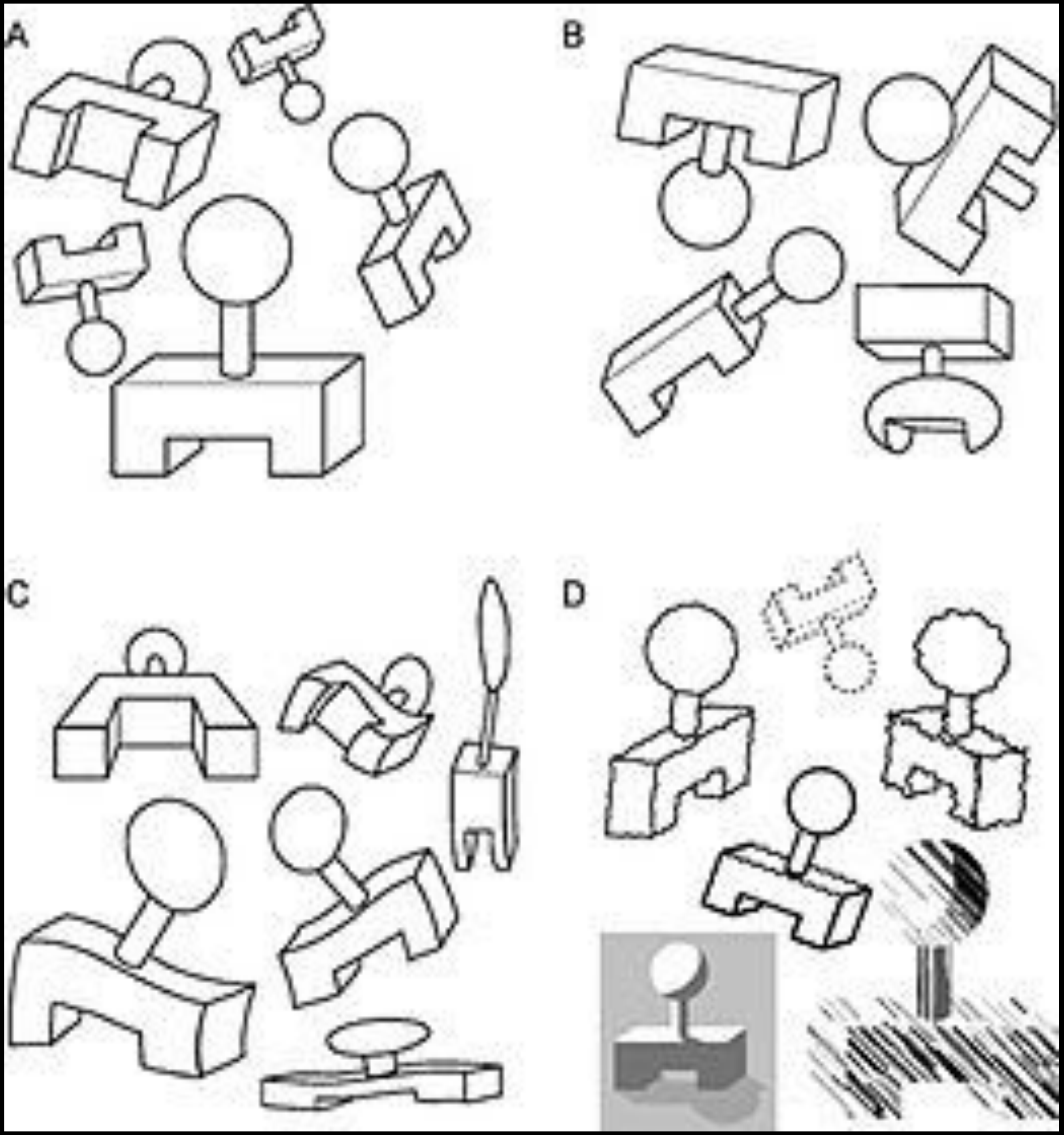
MULTISTABILITY

"My wife and my mother-in-law." (1915)

MULTISTABILITY

Ambiguous stimuli can generate **different perceptions** but they **can not coexist** simultaneously.

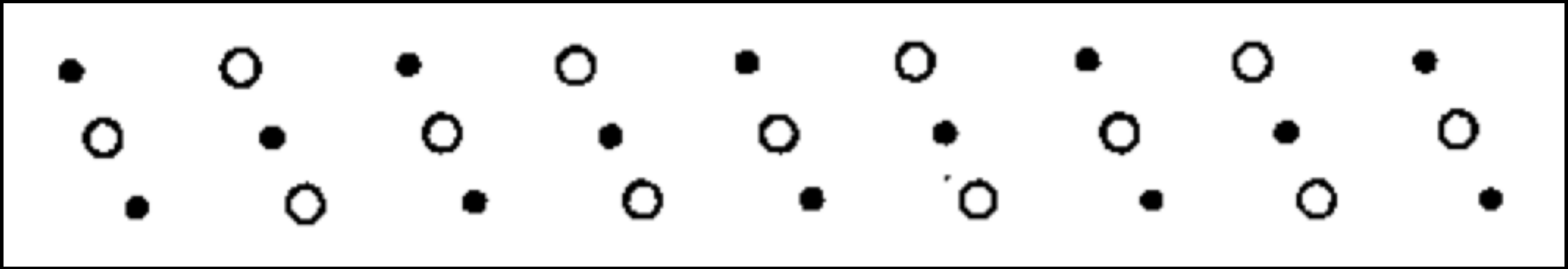
INVARIANCE



INVARIANCE

Objects are **recognized independently of various variations**, such as geometrical transformations, lighting, etc.

GESTALT LAWS OF GROUPING



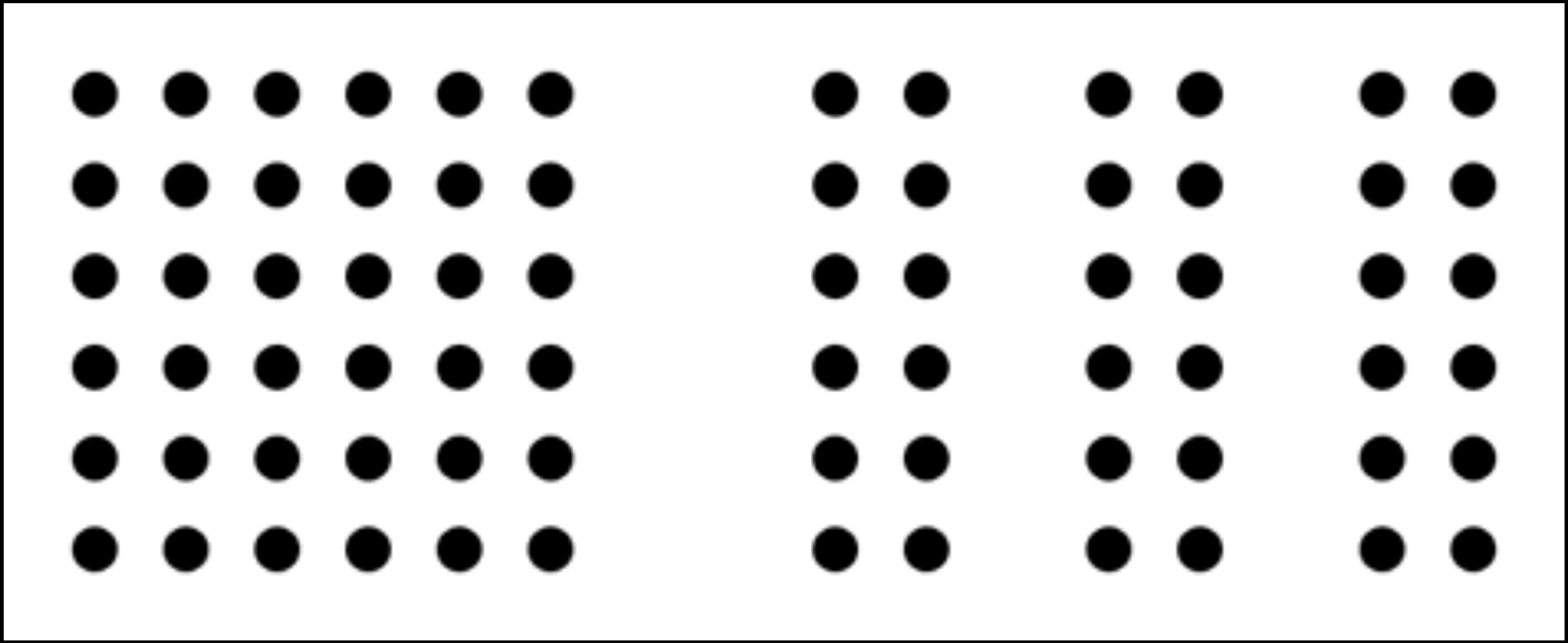
The **laws of grouping** state how **low-level perceptions** are **grouped** into higher-level objects.

Good Gestalt (Prägnanz)

We tend to order our experience in a manner that is regular, orderly, symmetric, and simple.

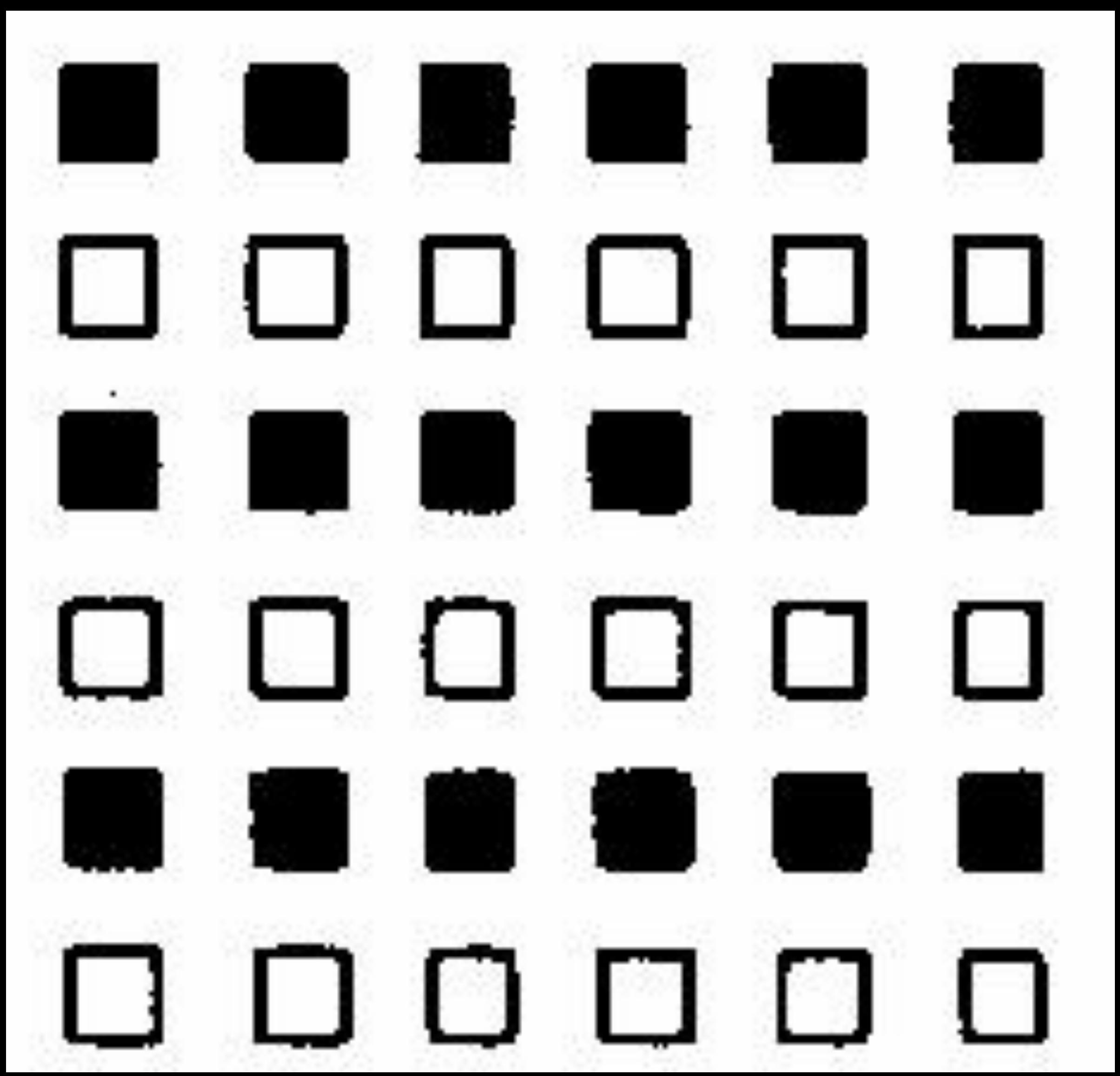
LAW OF PROXIMITY

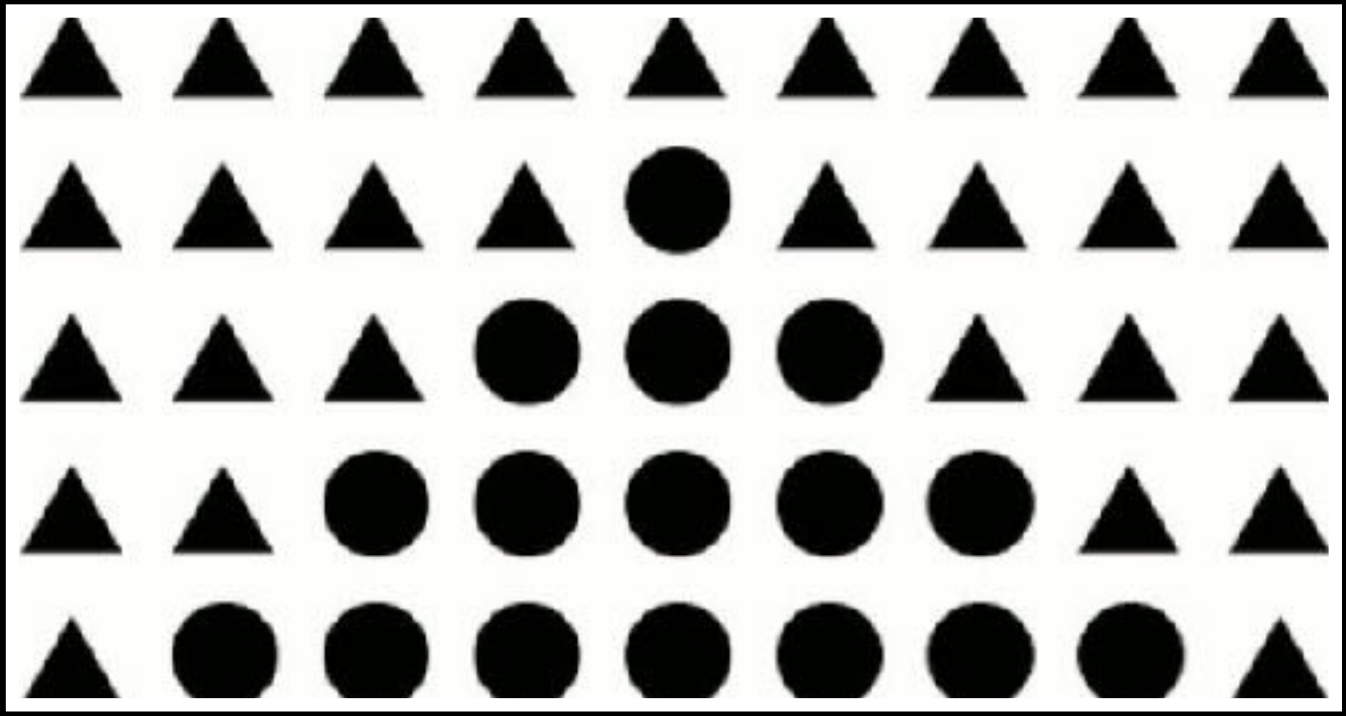
Objects that are close tend to be perceived as a group.



LAW OF SIMILARITY

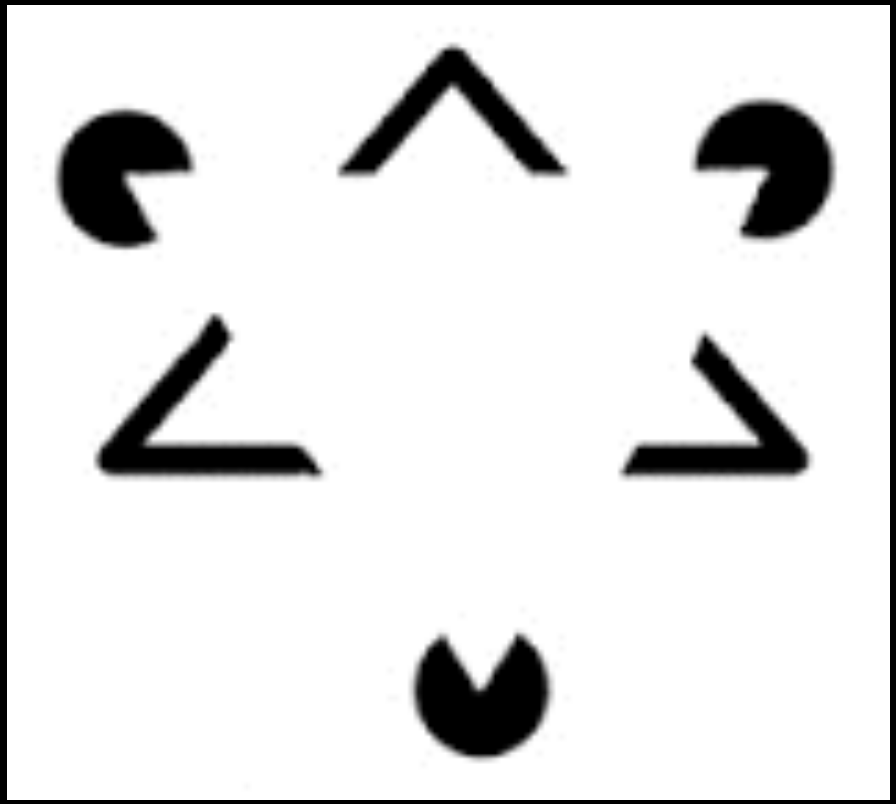
Objects that are similar (in shape, color, shading, etc.) tend to form a group.

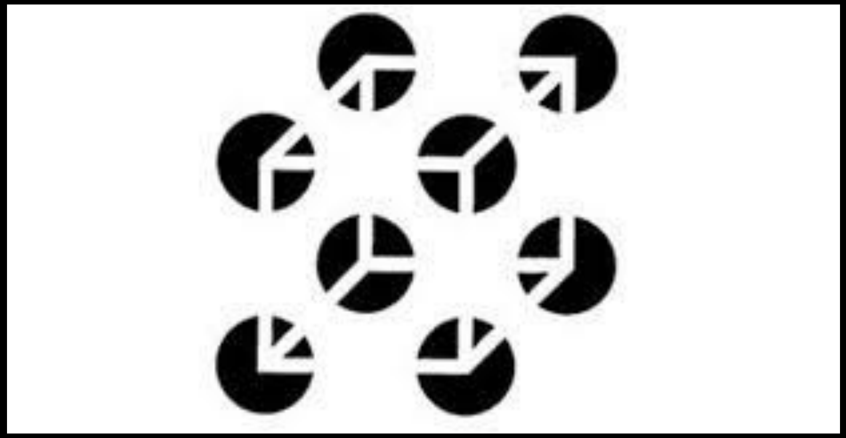
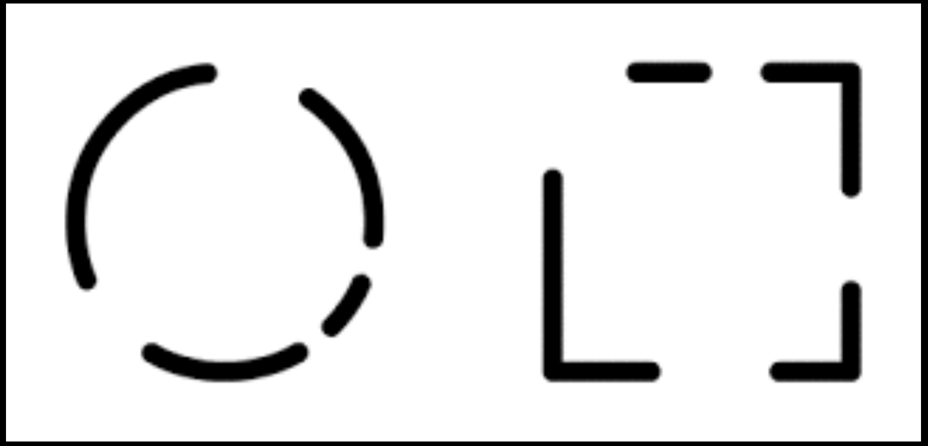




LAW OF CLOSURE

The perception fills gaps in stimuli.


LAW OF SYMMETRY

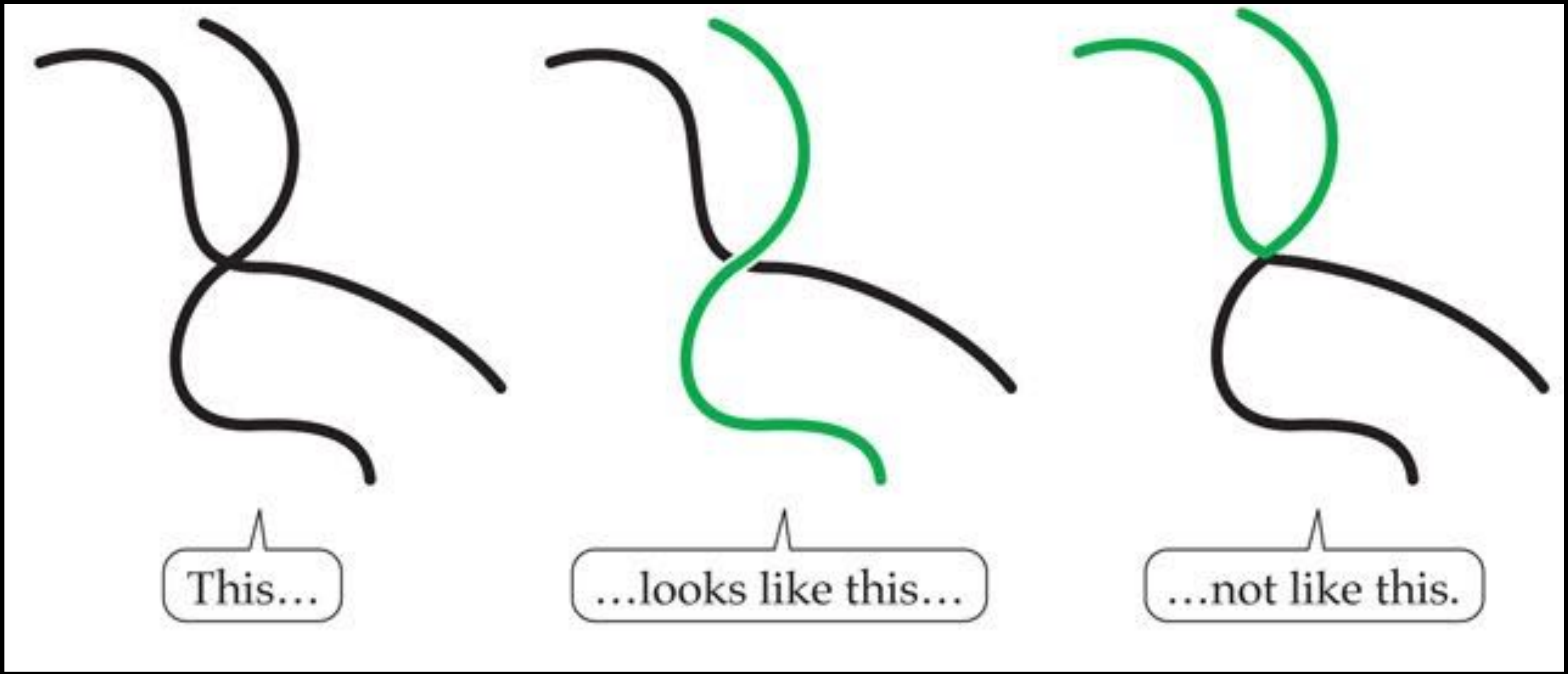
Objects with symmetric disposition tend to be perceived as forming a whole.

[ ]{ }[ ]

How many groups of elements are there?

LAW OF CONTINUITY

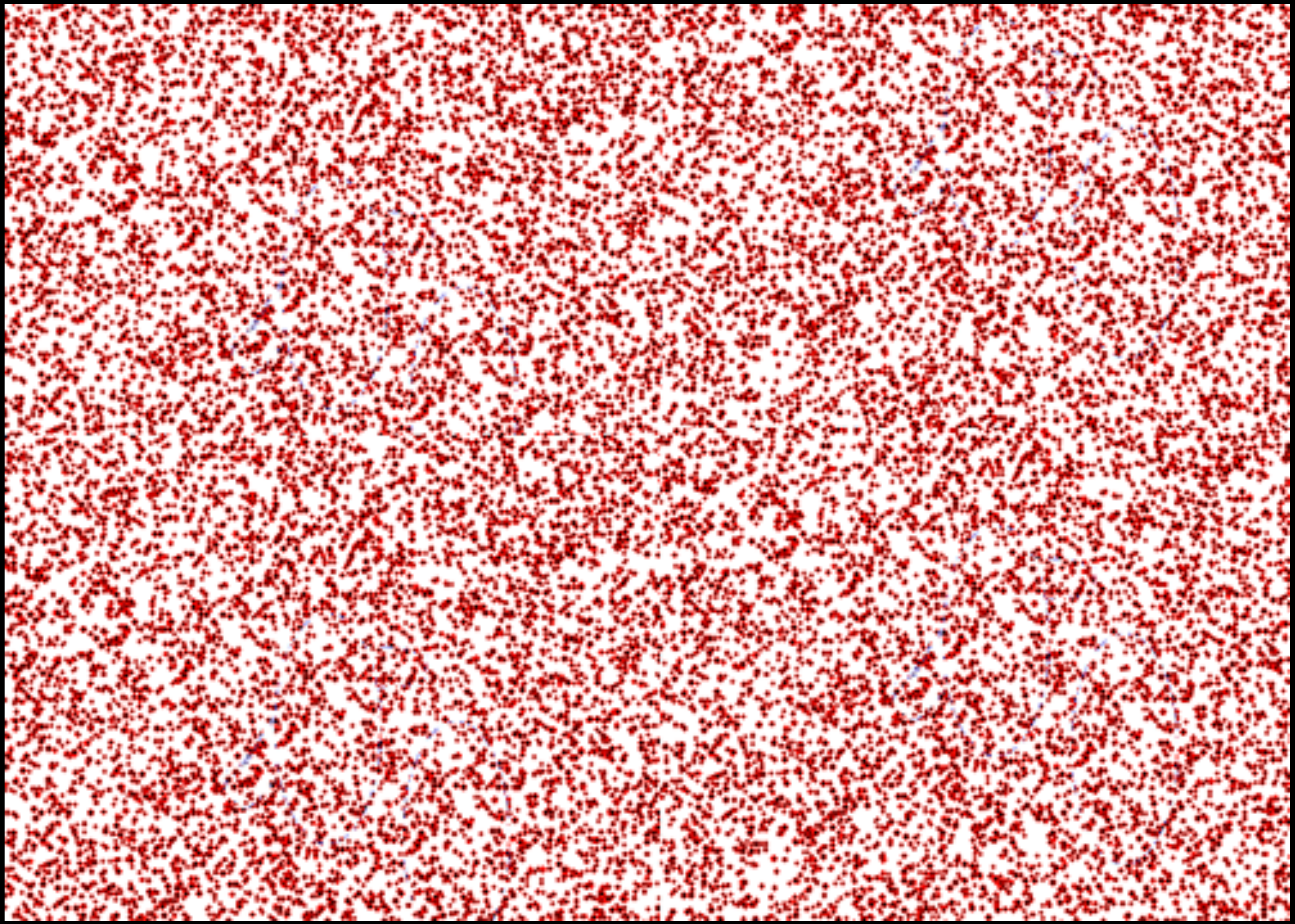
Ambiguous stimuli are perceived preferentially with the interpretation that is the most continuous.



[Source: Goldstein's Sensation and Perception]

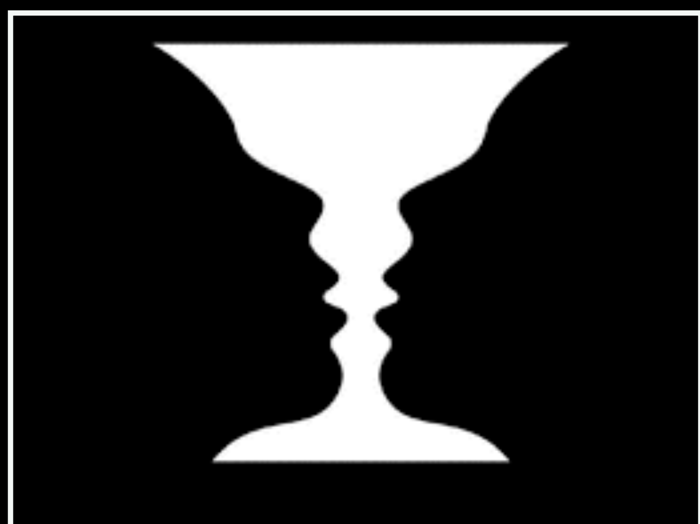
LAW OF COMMON FATE

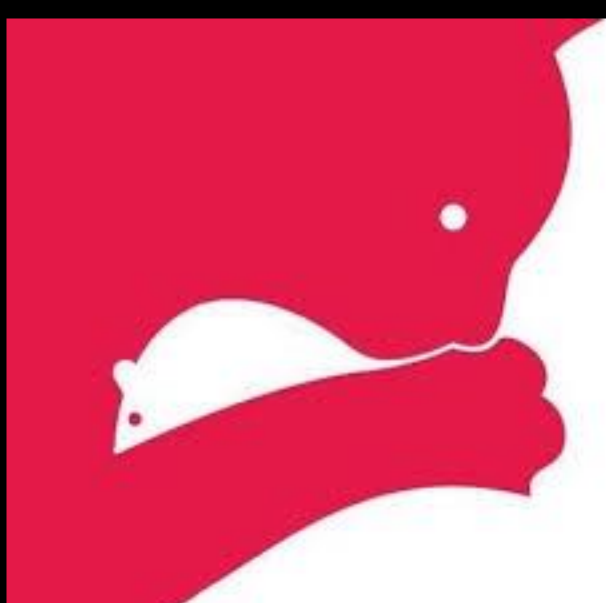
Objects evolving together are perceived as a group.



LAW OF FIGURE & GROUND

Elements are perceived as either a **figure** (element of focus) or **ground** (background on which the figure sits)



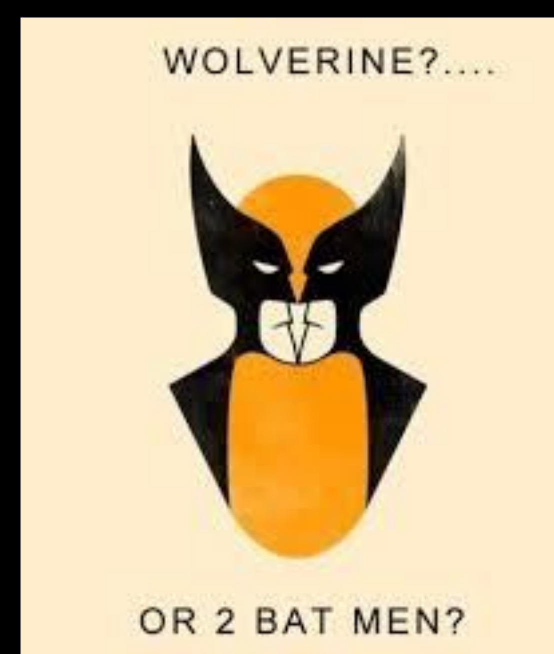
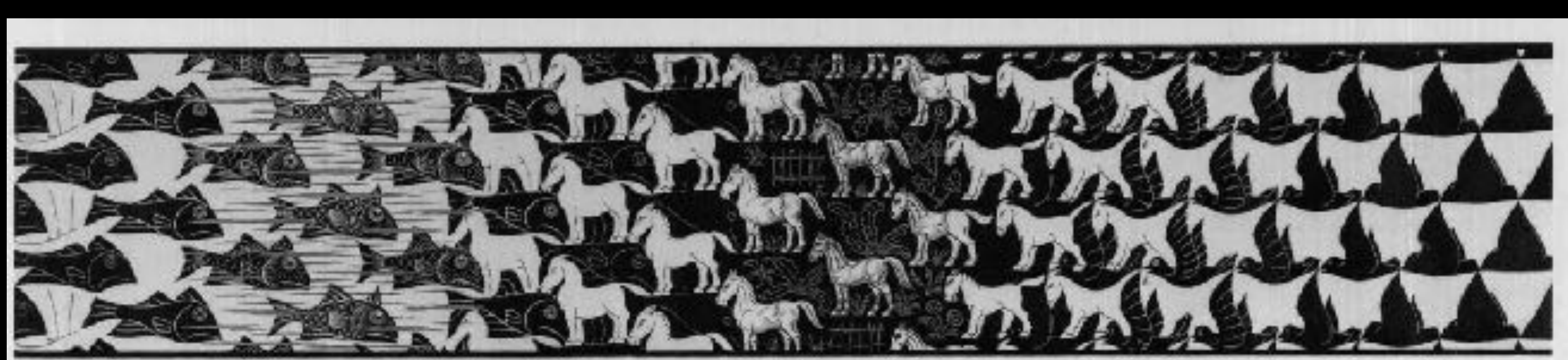


FIGURE & GROUND IN ART

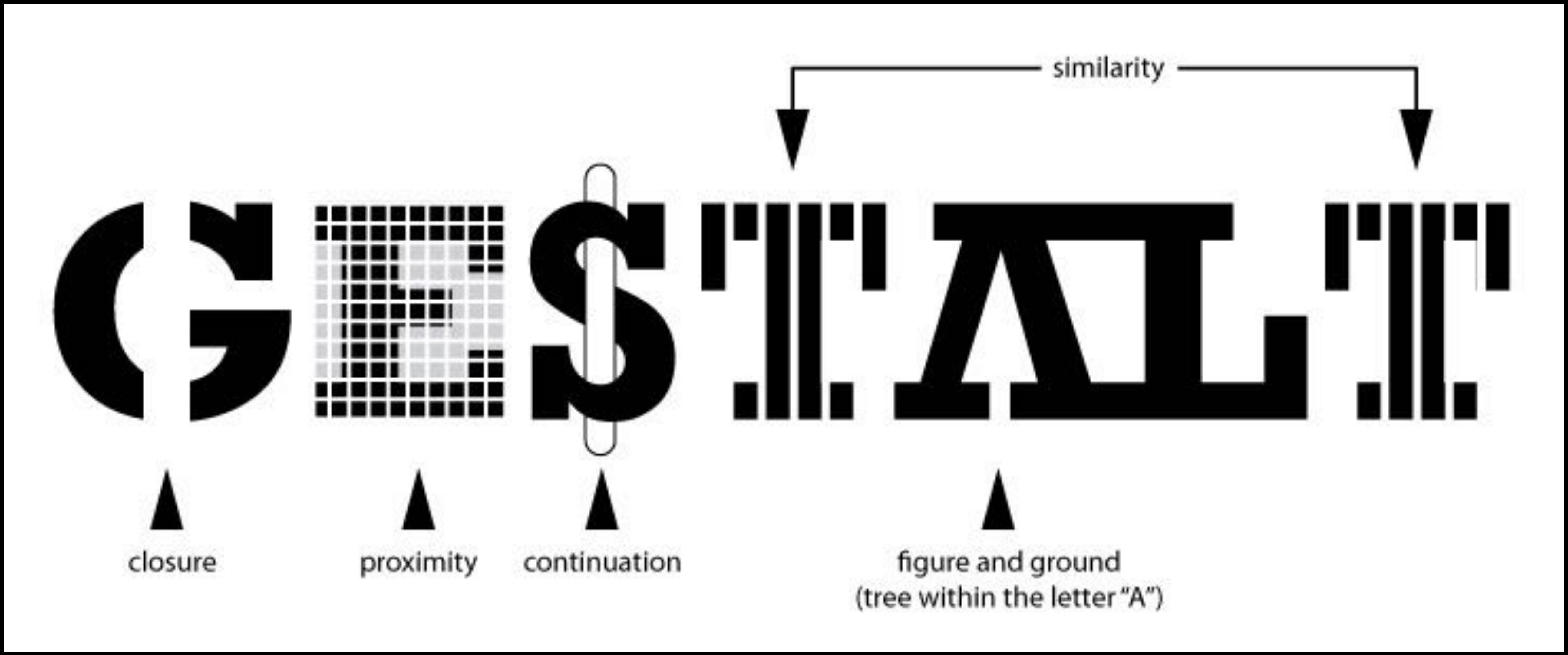
FIGURE & GROUND IN ART

FIGURE & GROUND IN DESIGN

FIGURE & GROUND: BEFORE GESTALT

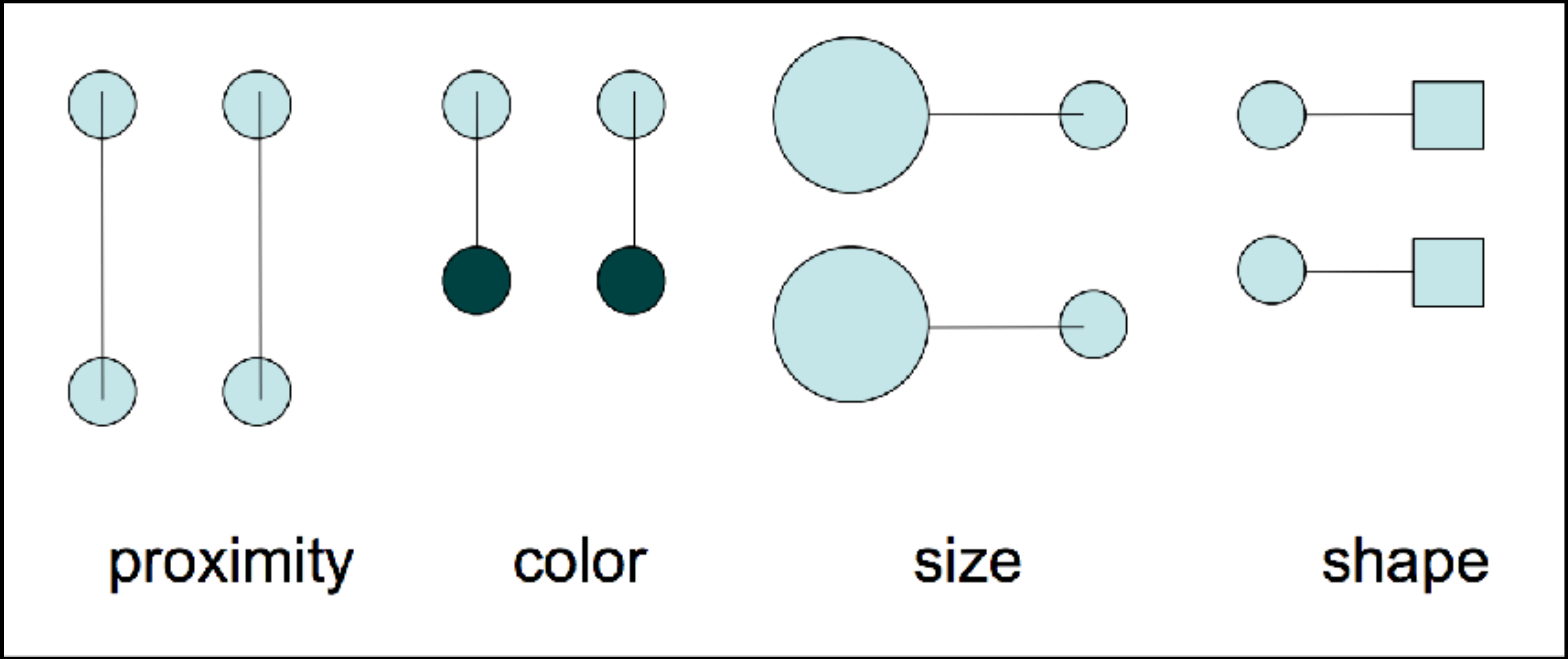


Escher’s Metamorphosis


similarity
closure
proximity
continuation
figure and ground
(tree within the letter "A")

MORE LAWS!
LAW OF CONNECTEDNESS

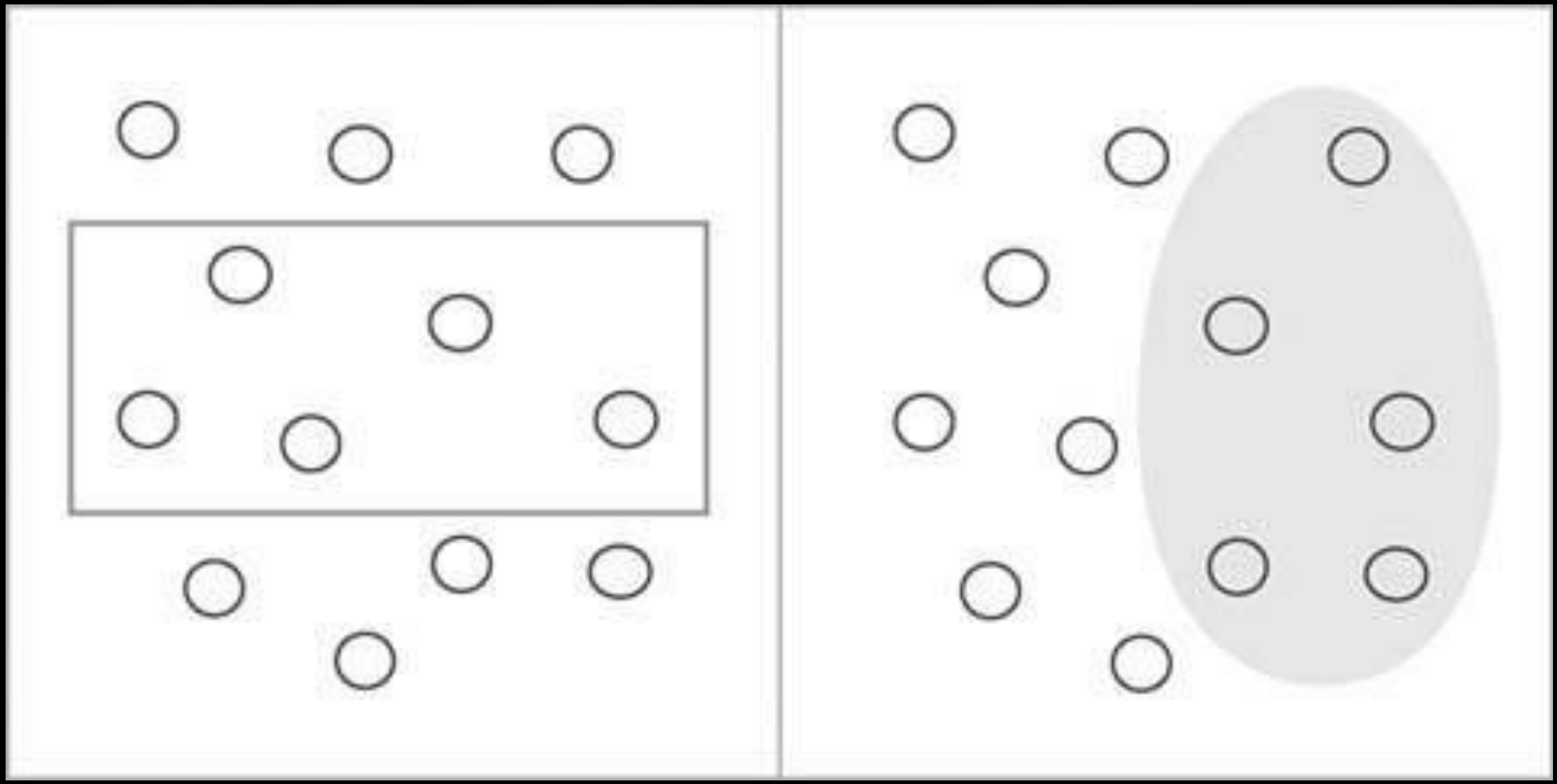
Things that are linked are perceived as belonging to the same group.



More powerful than proximity, color, size, shape…

MORE LAWS!
LAW OF ENCLOSURE

Objects that are enclosed are perceived as a group



Again, more powerful than proximity, color, size, shape…

IN SUMMARY

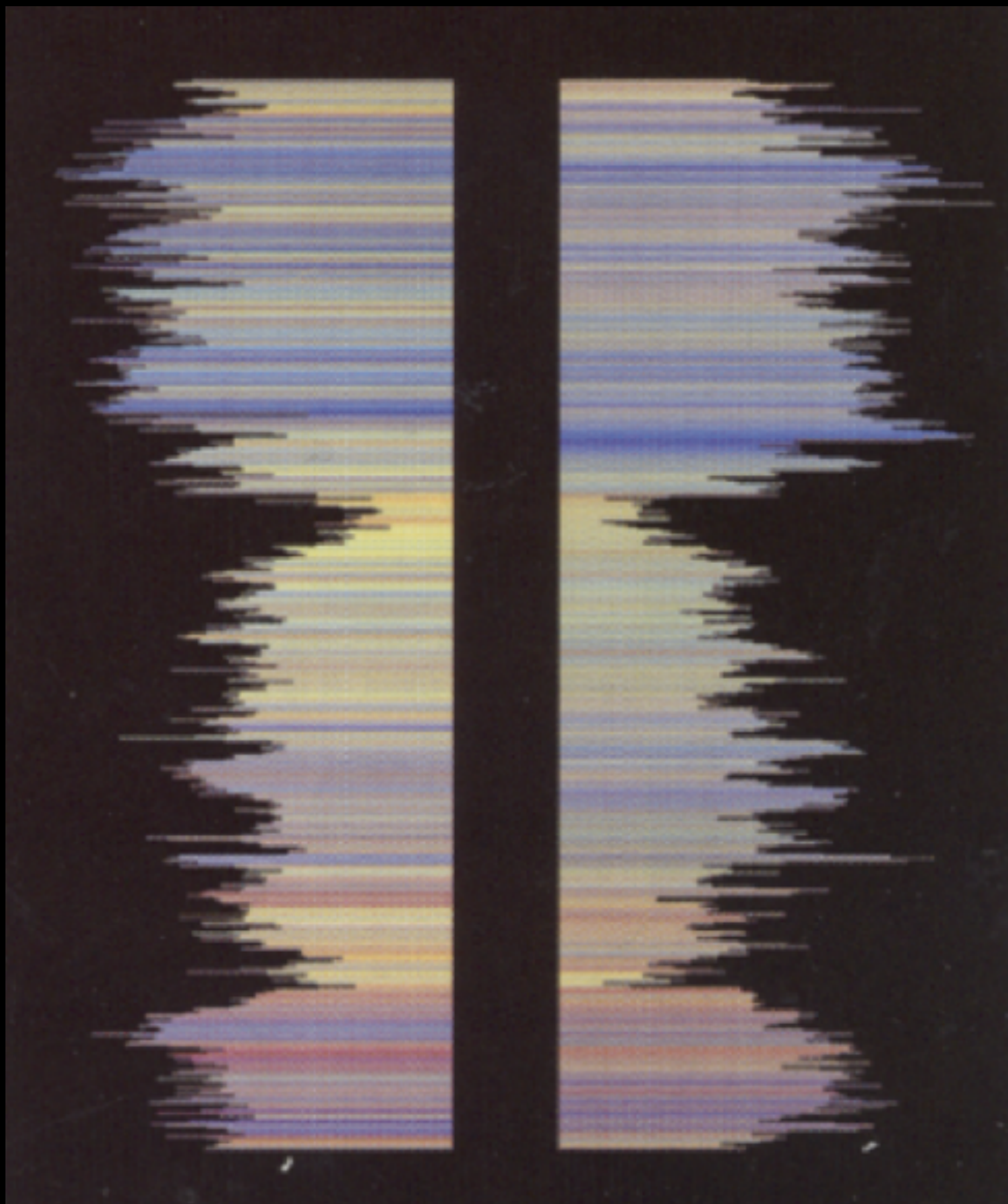
Our brains take lots of perceptual “shortcuts”…

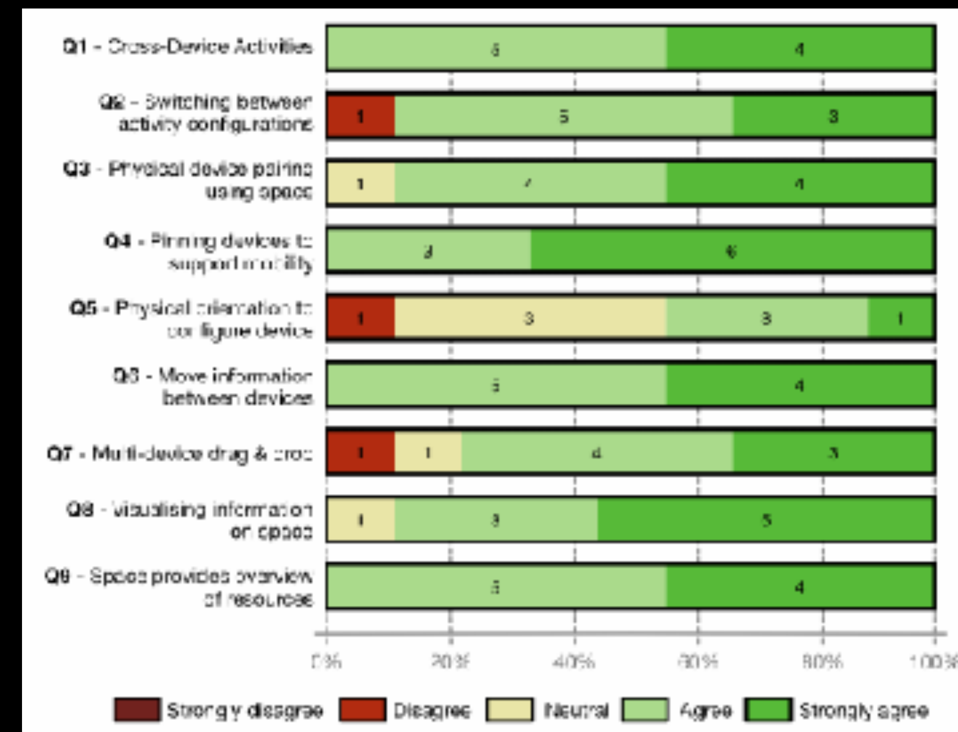
… which can either help or harm our visualizations!

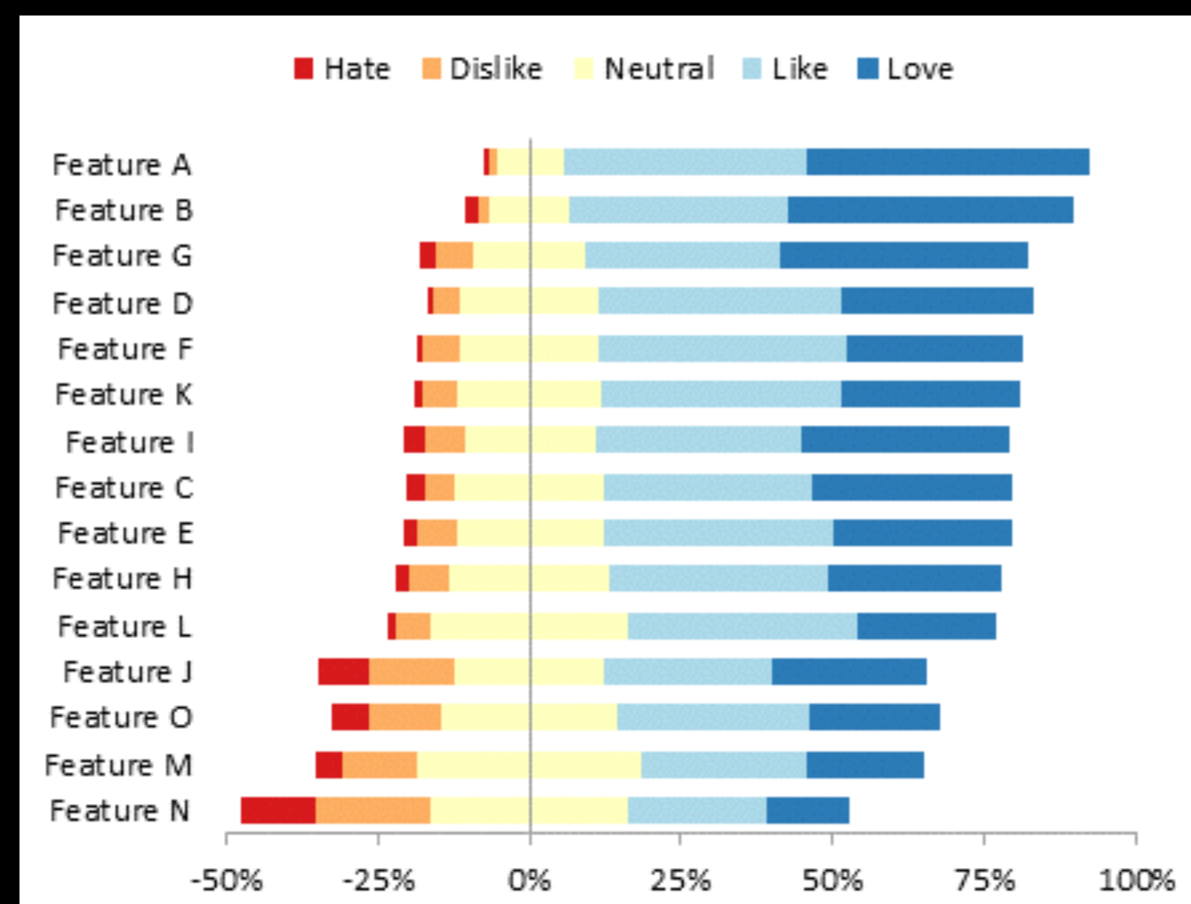
It is not enough to simply show something, we need to pay attention when and how it is shown.

A GOOD UNDERSTANDING OF PERCEPTUAL AND COGNITIVE PROCESSES IS CRITICAL!

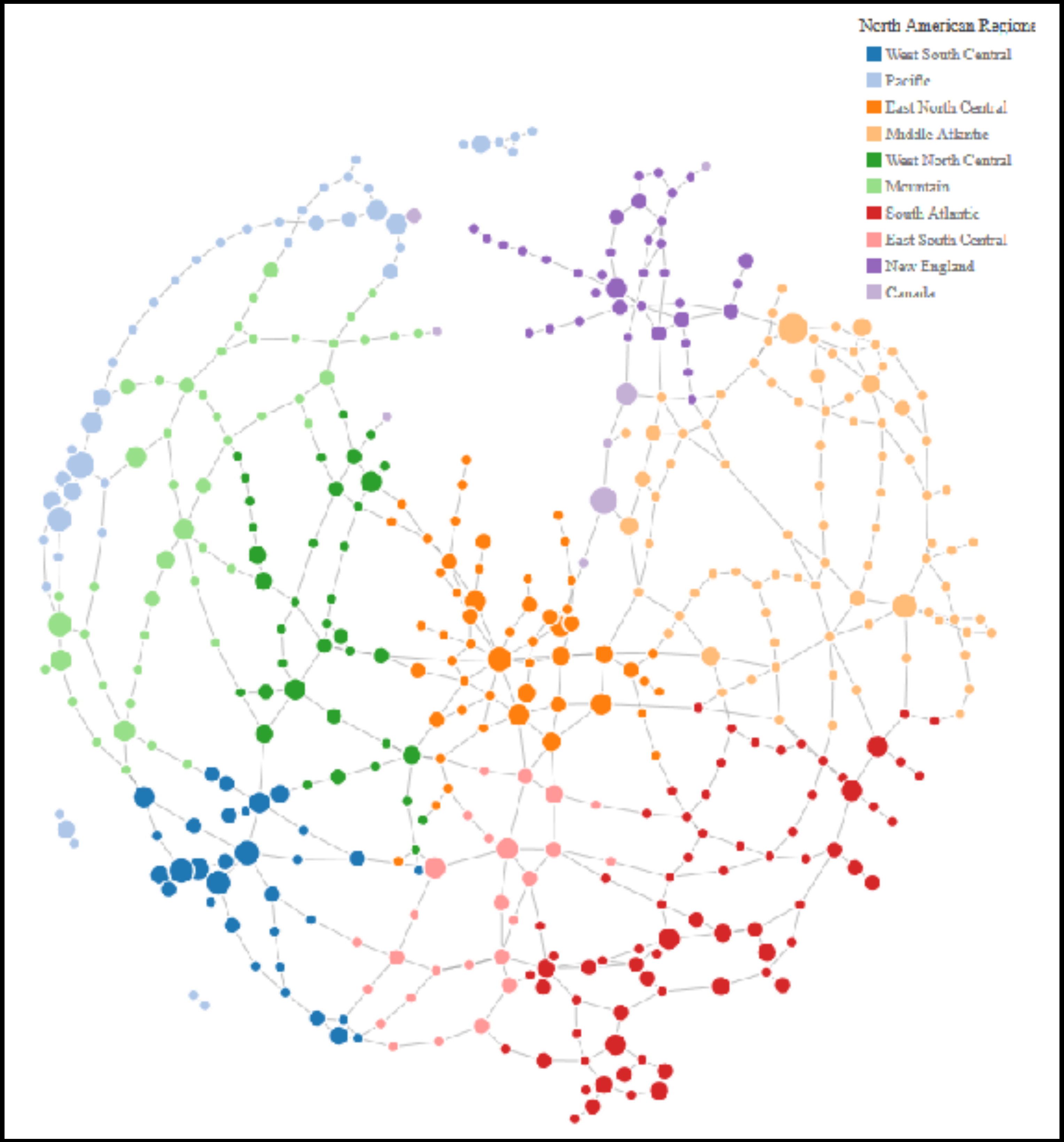
IMPLICATIONS FOR INFOVIS



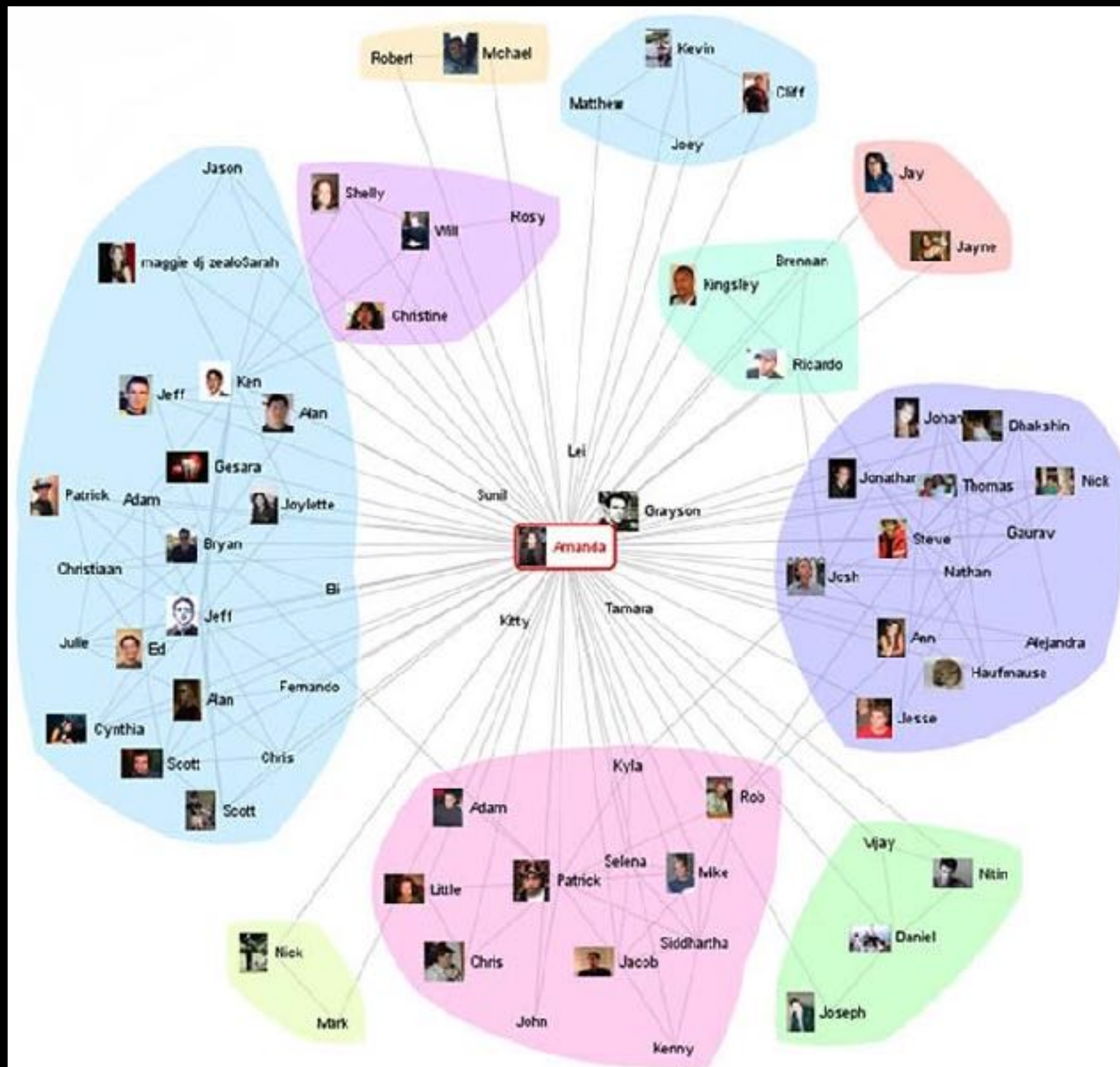




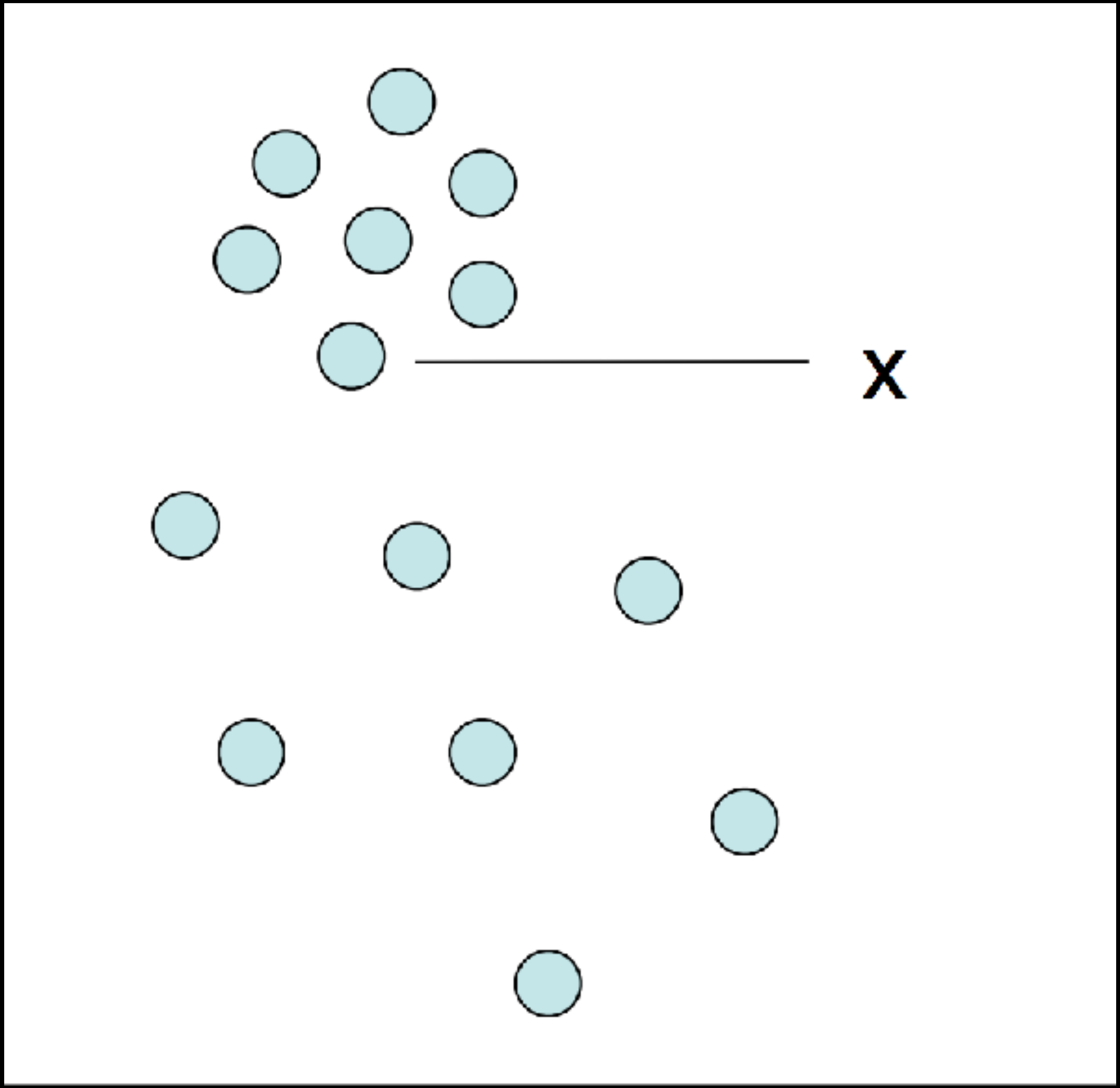
IMPLICATIONS FOR INFOVIS



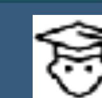
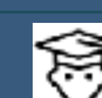
IMPLICATIONS FOR INFOVIS



IMPLICATIONS FOR INFOVIS



SCHEDULE

9 JAN.	WELCOME - INTRODUCTION / PROSPECTIVE PROJECTS
16 JAN.	VISUAL PERCEPTION & DATA MODELS
23 JAN.	GUEST SPEAKER : JUSTIN MATEJKA (9:30AM)
30 JAN.	EXPLORATORY DATA ANALYSIS + ***STUDENTS PAPER PRESENTATIONS***
6 FEB.	GRAPHS & NETWORKS + ***STUDENTS PAPER PRESENTATIONS***
13 FEB.	**PROJECT: MID-TERM REVIEW** 
20 FEB.	READING WEEK
27 FEB.	GUEST SPEAKER : ISABEL MEREILLES
6 MAR.	INTERACTION & ANIMATION + ***STUDENTS PAPER PRESENTATIONS***
13 MAR.	***STUDENTS PAPER PRESENTATIONS***
20 MAR.	***STUDENTS PAPER PRESENTATIONS***
27 MAR.	***STUDENTS PAPER PRESENTATIONS***
3 APR.	**PROJECT: FINAL PRESENTATIONS** + WRAP UP 

http://www.cs.toronto.edu/~csc2537h