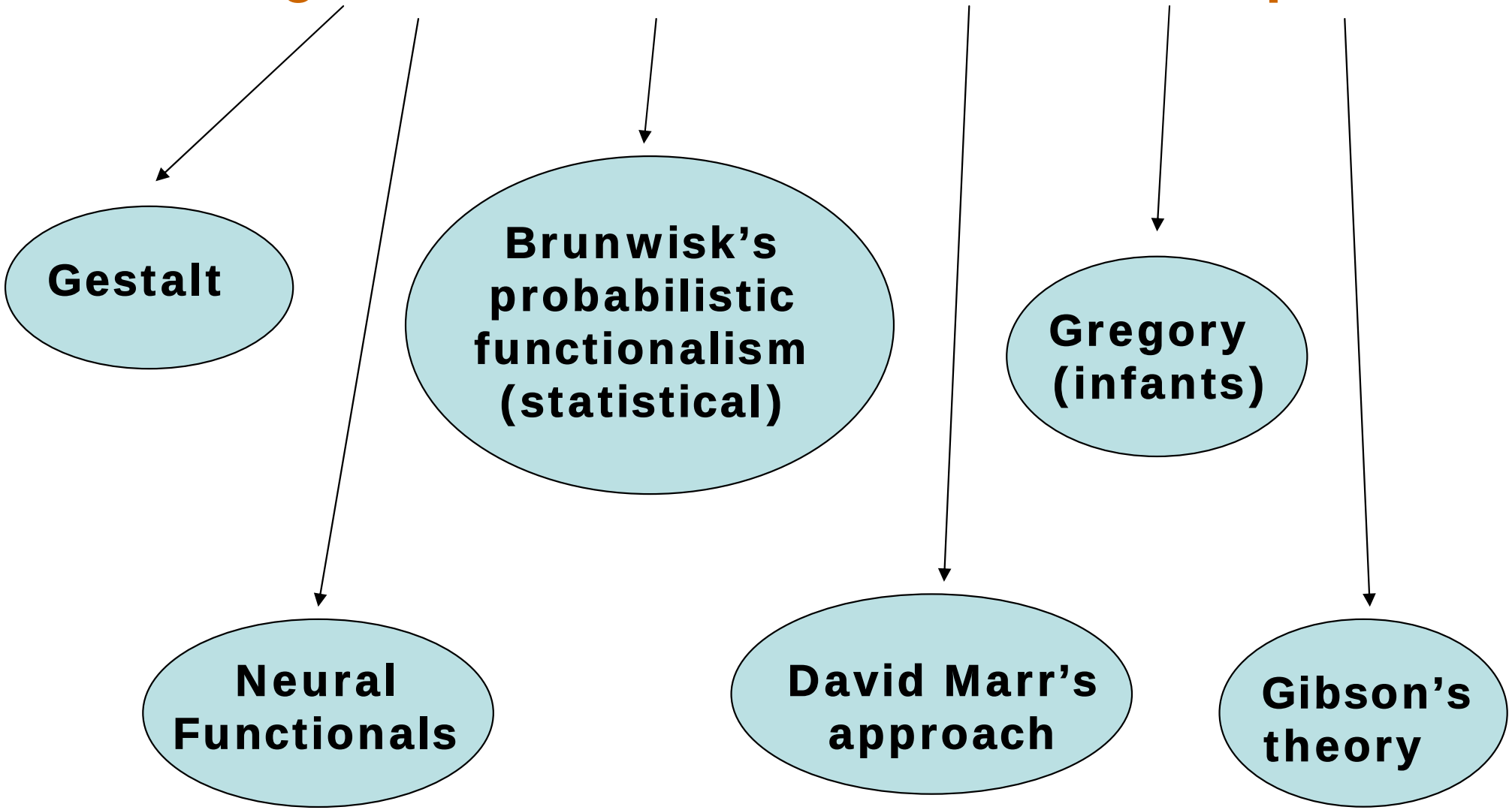
INTRODUCTION

Here, we will *try* to understand how attempts were made to describe theories of visual perception.

- **development of backgrounds**
- **underlying assumptions**
- **strengths and weaknesses**
- **current status and applicability**

What is VISUAL PERCEPTION??

Categories of theories in Visual Perception



In sciences, a theory is a logically self-consistent model or framework describing the behaviour of a certain natural or social phenomenon (thus either originating from observable facts or supported by observable facts).

In contrast, a hypothesis is a statement which has not been tested yet.

Theories are formulated, developed and evaluated according to the scientific method.

A theory is an established paradigm that explains all or much of the data we have and offers valid predictions that can be tested.

In science, a theory can never be proven true, because we can never assume we know all there is to know.

Instead, theories remain standing until they are disproven, at which point they are thrown out altogether or modified to fit the additional data.

None of the theories in VP (Visual Perception) are great;

w.r.t. say, Einstein, Newton, Charles, Coulomb, Maxwell, Fourier etc.

Other examples of *non-great* theories:

Mandelion genetics, Marx's theory of historical process, Keyne's economics, Young-helmholtz theory of colour vision, etc.

Well, old theories are abandoned and new theories developed / altered

Among all theories in VP, Marr's work has led to a lot of good ideas and high-quality research. We may term it a good utilitarian or working theory.

Vision is a sense which enables us to view physical stimuli. For stimuli to reach our brain and be interpreted into a meaningful picture, our body must complete a complicated process.
This process is broken into two parts:

Visual Sensation and Visual Perception.

Visual Sensation, which occurs so we can detect the physical stimuli, begins with the reception of information. Light (electromagnetic energy) enters the eye and is focused on the retina.

This electromagnetic energy is then transduced by rods and cones (light sensitive cells) into electrochemical energy enabling it to be transmitted to the brain.

The electrochemical energy is then transmitted along the neurons of the optic nerve to the Occipital Lobe (Visual Cortex of brain).

Visual Perception, which occurs so we can make sense of the stimuli that has been sensed, begins with selection.

This process involves sorting and isolating features of a stimulus which are kept separate on the neural pathway until reaching the brain.

Feature detector cells then respond to specific features of the stimuli and match them up with ideas from the existing mental set in the process of organization.

Once organized into a meaningful whole, interpretation of stimuli takes place thereby giving meaning to the visual stimuli that is being experienced.

Sensation may be regarded as a **physiological** process and **perception** as a **psychological** process.

However, the difference between sensation and perception is not always clear. We all have the capability of sensing the same stimuli but our interpretation depends on the mental processes involved.

Prior experiences, memory, emotional state, knowledge, expectation and psychological characteristics of an individual all contribute to the way stimuli will be perceived.

The difference between visual sensation and perception can be found in an example of two children viewing a dog. One child approaches the dog to pat it while the other crosses the road to avoid it.

Both children sensed the same stimuli, but due to different mental processes, perceived the stimuli differently.

The child who avoided the dog may have been bitten in the past and the child who approached it may never have been bit before.

Visual perception principles are rules that we unconsciously apply to incoming physical stimuli to shape them into a meaningful whole for easier interpretation.

There are three main categories of perception principles: Perceptual constancies, depth cues and gestalt principles.

Perceptual constancies are seen in action when an individual perceives an object as maintaining the same features even when the image on the retina changes.

An example of this would be found in *shape constancy*. As we move around, the angles of objects which we perceive change shape on our retina but by automatically applying the principle of shape constancy, we realise the stimuli does not physically change shape despite the images.

An example of shape constancy can be found in a spinning coin. As the coin rotates, the image cast upon our retina changes from a full circle to an eclipse, down to a slender rectangle, then back through these shapes again. However, at each stage of the spinning we still perceive it to be a circular coin.

Depth cues are either provided by the environment or our own body and help us to determine where things are located.

They can be classed as monocular (requiring one eye) or binocular (requiring two eyes), primary (within body) or secondary (take place in environment).

An example of a secondary depth cue is texture gradient. Texture gradient provides a depth cue because we perceive diminishing detail to mean there is a greater distance between us and the stimuli.

The principle of texture gradient can be seen in a field of purple flowers. We can see the petals in detail on the flowers closest to us, but as we looked down the field, the petals blur into one purple 'sheet', therefore telling us those flowers which we cannot clearly distinguish are further away than those which we can.

Gestalt principles allow us to make sense of what we are perceiving by finishing incomplete, or rationalising incoherent, stimuli.

One of the most widely used gestalt principles is closure. This is the process whereby we mentally complete an incomplete stimuli by adding the contour of the image.

Many examples of where this process is used can be seen in the advertising industry. 'Telstra' have recently changed their logo to include incomplete stimuli. The 'T' in their logo simply has a horizontal bar above a circle which has a vertical strip missing from it. Although the horizontal or vertical bars do not meet, we perceive a 'T'. We are hereby using closure to complete the incomplete stimuli.

Although perceptual principles help us to make sense of ever-changing stimuli in an ever-changing world, psychological factors greatly determine how we eventually perceive the stimuli. All psychological factors are interrelated, however the factors influencing perception will differ from person to person.

One factor which affects the way stimuli can be perceived is the perceptual set. Previous experience or previous information that makes up the perceptual set can falsely lead us to expect something from a stimuli, therefore creating an error in our perception.

An example that is both relevant to this factor and myself as a student can be found at school. I have observed two classes anticipating the same test. One class have been told by their teacher that the test is hard, while the other class have been told it is easy. After sitting the test, the class who expected it to be hard exit the room complaining of their difficulties and the class who expected it to be easy, exit confidently.

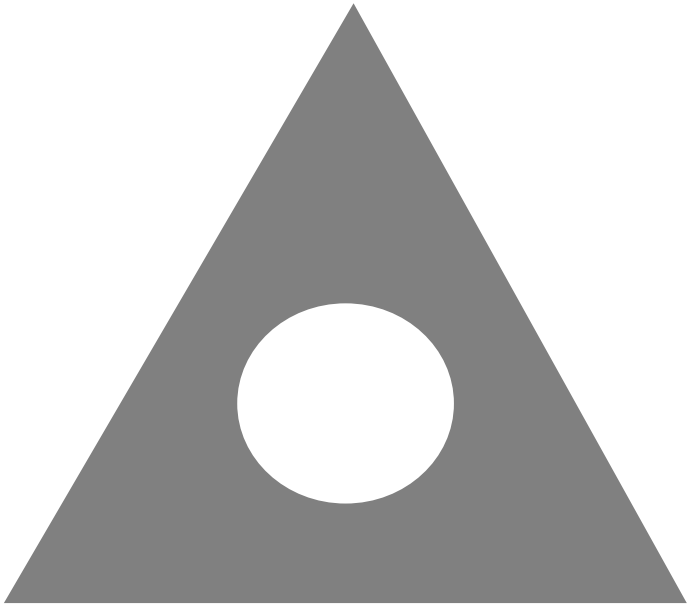
Another psychological factor which can often influence the perception of incoming information is/are(?) the psychological determinants.

This reasoning follows the idea that information we are currently in need of or interested in determines our attention and can affect our interpretation of sometimes ambiguous stimuli.

Gestalt's Theory

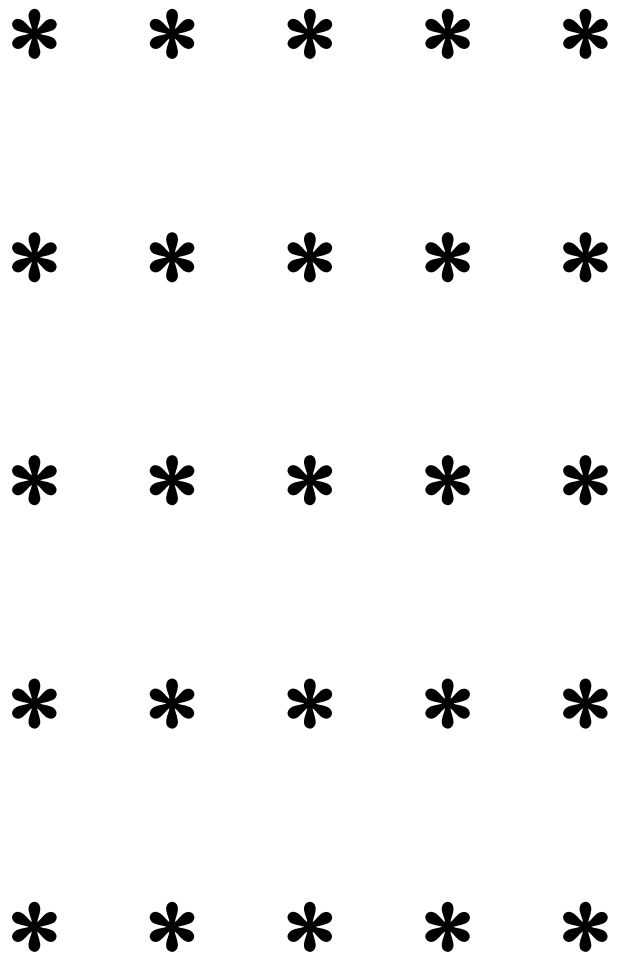
Perception as a dynamic, organized process:

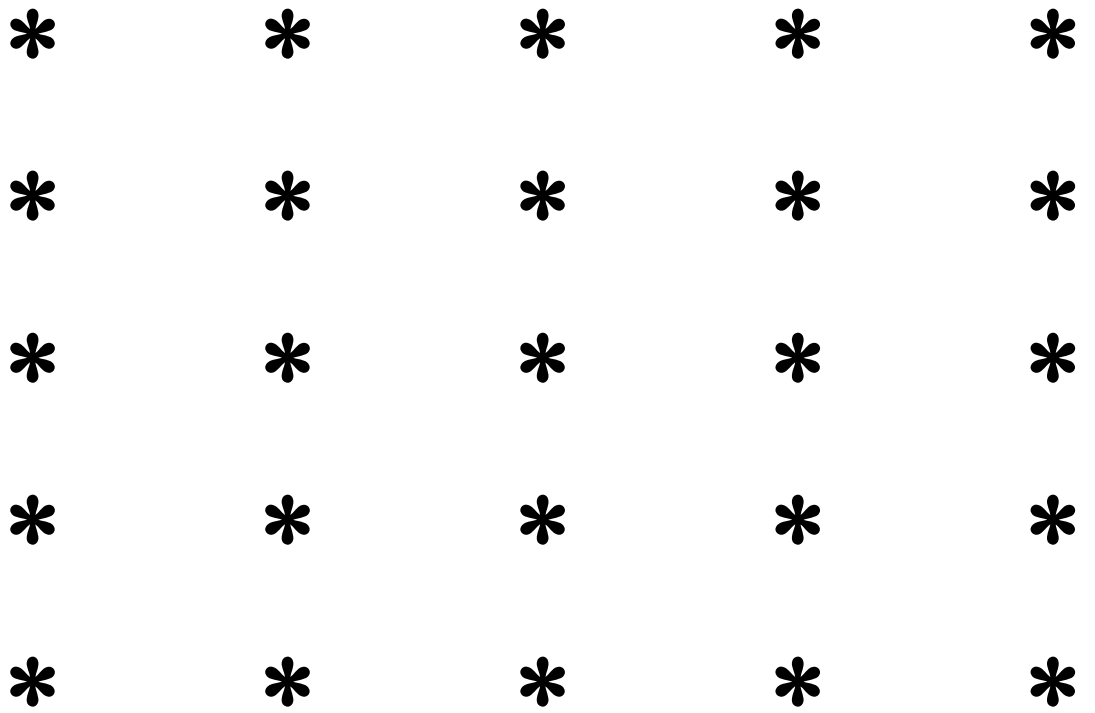
Figure-background distinction ?



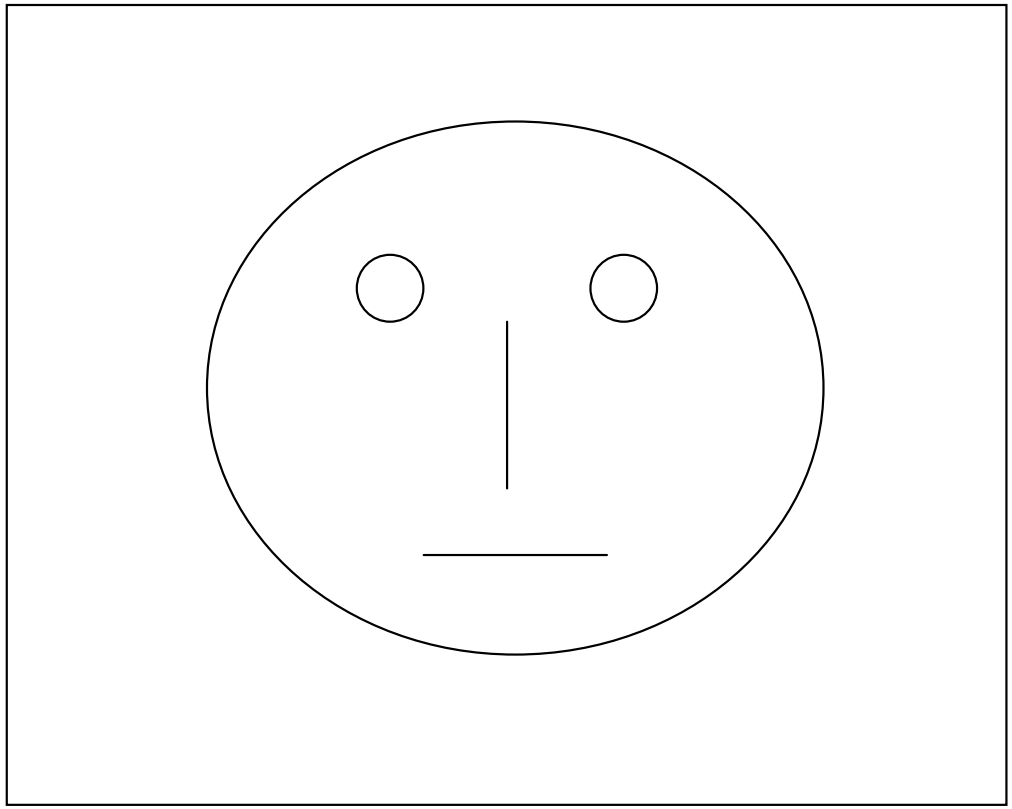
Know about the face-vase illusion ?

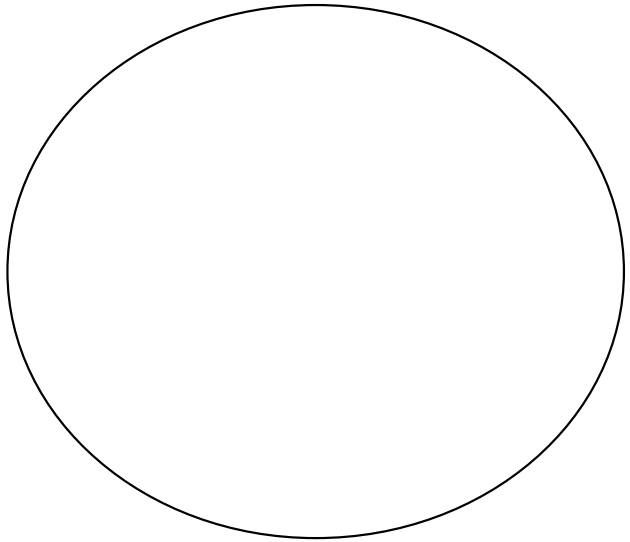
Laws of grouping:

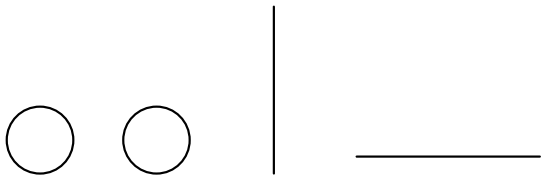




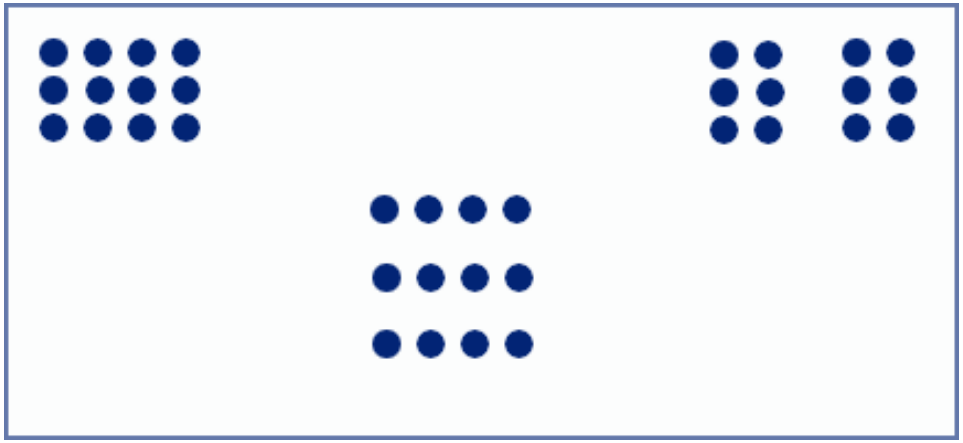
Wholes and parts:

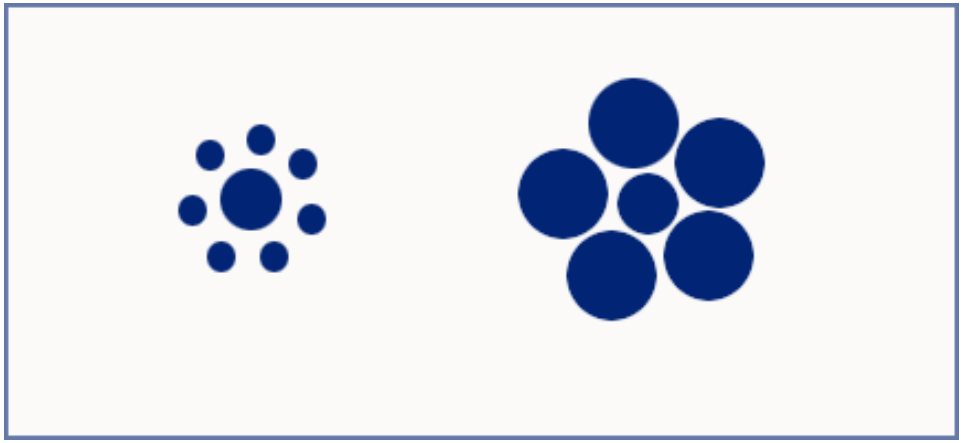






Also, which is the back-ground color ?

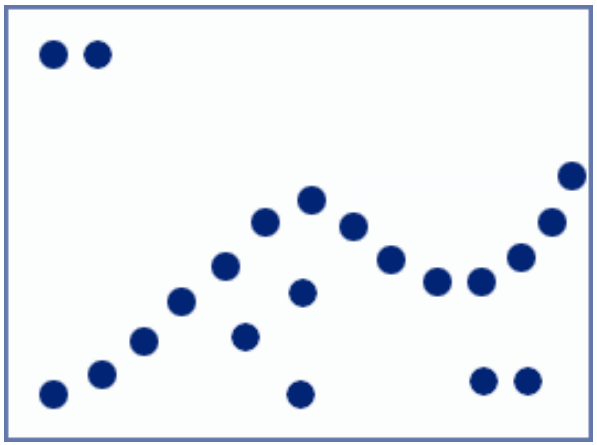




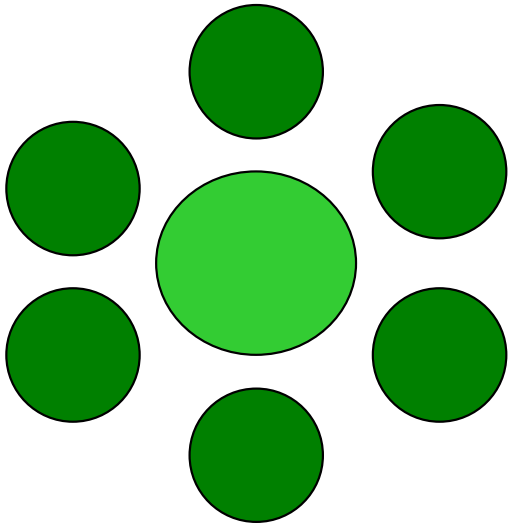
PROXIMITY

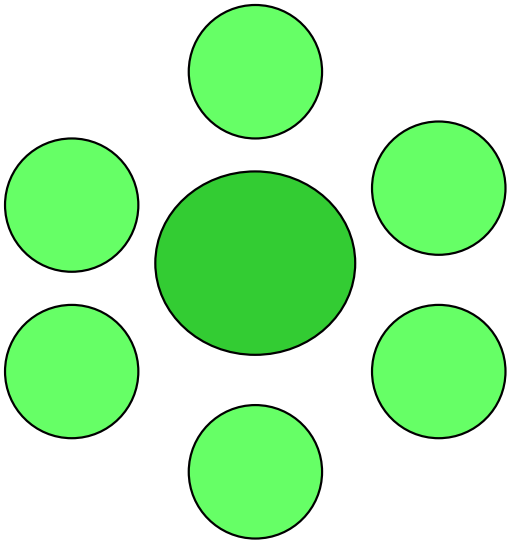
CLOSURE

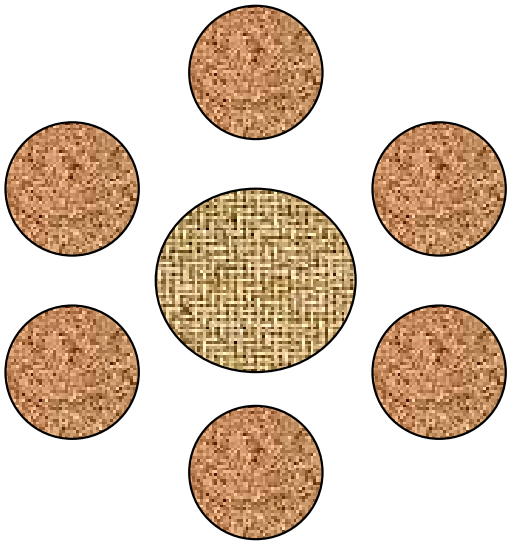
GOOD CONTINUITY



SIMILARITY







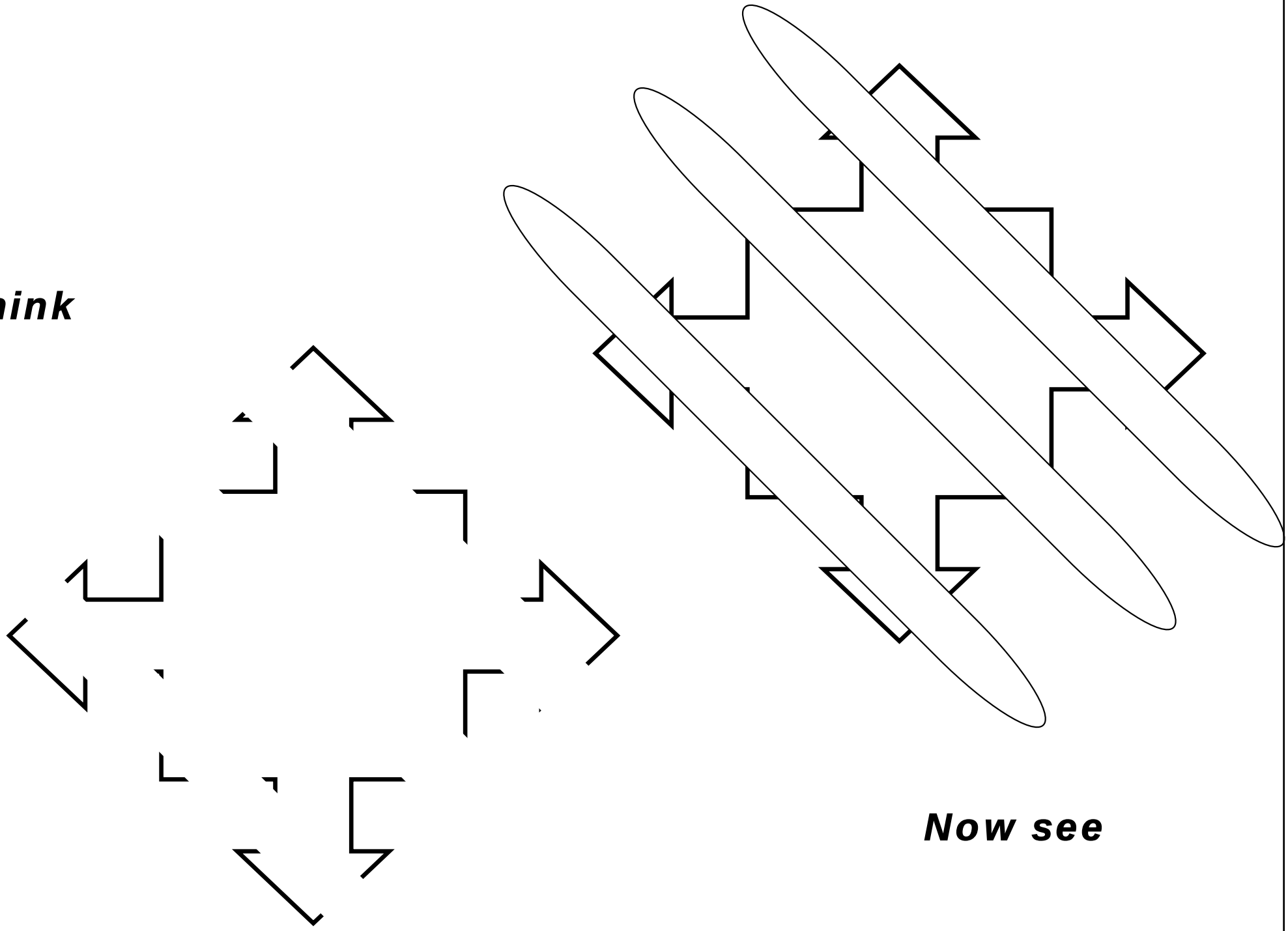
Infant perception:

- Habituation and discrimination
- 3D object vs. its photograph
- sensitive to orientation
- size constancy after birth
- preference to attractive faces
- motion perception – crawling and just walking

Other factors:

- Grouping by similarity – textures, color, shape similarity
- symmetry and information – target and mate selection

Think

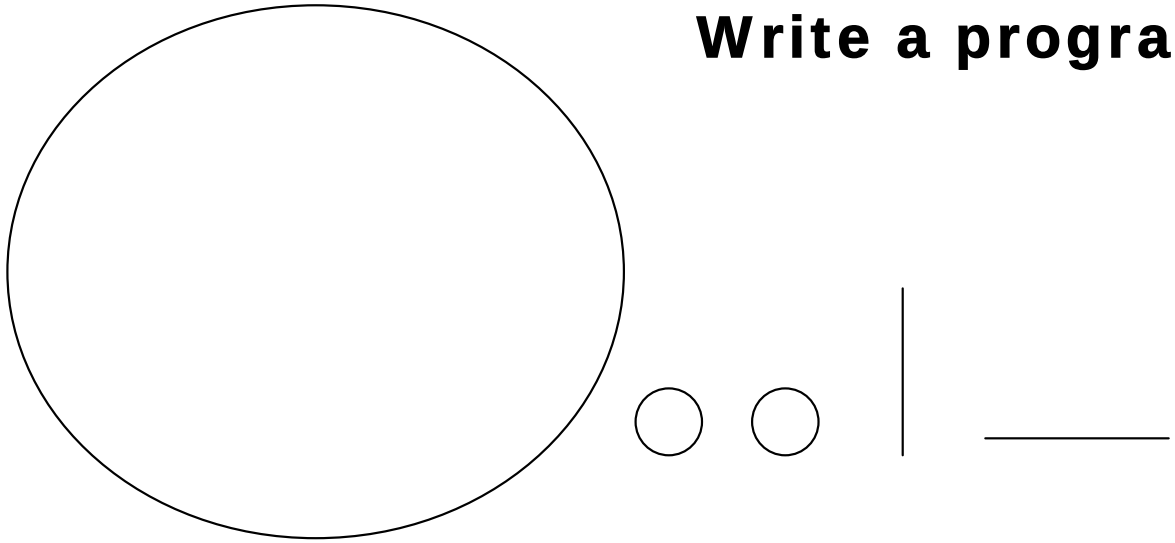


Now see

Write a program to generate the following sequences:

- **1212121212……..**
- **314159265262……..**
- **01834862738531……**

Write a program to generate:



Or the whole face:

Kolmogorov complexity defines the shortest program / algo, that computes / solves the above problems

Simplicity, minimum principle are now used by vision scientists to solve problems in vision, along with algo. Info. Theory (Shannon)

Brunwick's probabilistic functionalism

- **Difficult to describe using a well articulated theory**
- **Lots of experiments and conclusions**
- **Mainly conceptual but highly relevant**

Outline of Brunwick's work:

• Statistical nature of cues, many of them simultaneously arising out of the world (uncertain)

• Object and events are complex, and rarely occur in isolation

• Rapidly arriving at a decision or valid perception based on uncertain information

- **Inspite of this (uncertainty), the perceivers get their decision right (quickly and amazingly) under most non-trivial circumstances**

- **Else difficult to survive – how does this happen ?**

- **Brunwick – understand the ecology or environment, under which these cues arise; observe perception under complex life-like conditions and how it evolved**

Some terminologies and concepts:

- **Proximal variables (cues) are sizes and shapes of retinal images**

- **Distal variables are actual properties of things and events out in the world.**

- **The world is stable, but the retinal image is continuously changing (look around in this room, and visualize)**

Example of probabilistic nature of cues:

You are hungry (so much so, that eat or die), and you are searching for a fruit to eat, in a jungle *(no Tigers ☺)*

You know that edible fruits are darker, redder, softer and sweeter. All cues are not visual, and all are imperfect (you also know). e.g. some poisonous fruits are sweet too, soft fruits may be rotten, etc.

What would do you do in a situation of such uncertainty. Commit harakiri equivalent to a suicide ? No pun intended - it is a matter of life and death for you. Our ancestors (in geological time scale), must have faced this situation many times.

If they have failed on most occasions, then……..

We are all very smart intuitive statisticians, in real life (may not be in the MATHS paper). Very efficiently combine and weight the cues and be selective in choosing.

Heard of decision and feature fusion, somewhere ?

Brunwick (1956) – Perception is uncertainty geared. It aims for smallness of error at the expense of the highest frequency of precision.

Perception involves the evaluation of evidence from different sources, the estimation of relative probabilities and decisions about the attainment of goals. Thinking is similar, but no exactly so.

Thinking is a time consuming, deterministic (certainty geared) and discontinuous (sudden attainment after lengthy pauses).

You may think well (clever) but not a good perceiver.

Other work: Validity of cues, Lens model, size constancy, grouping and spatial proximity etc.

Brunwick was an experimental psychologist, with no background of statistics.

Brunwick's main concern was to explain the uncertain relationship between things in real world (3D distal stimuli) and structure of the visual image (2D proximal stimuli).

Effect of perspective distortion on shape and size; light and material properties dictating intensity, etc. make visual clues ambiguous.

How do we resolve the complexities and ambiguities of vision ? Tiger charging at us, or spear (or Shoaib's bouncer) heading towards your head. We need to make a quick decision, even at the expense of small errors. No time to think. Experience and evolution has helped us to survive.

Empirical studies (later developments in theory of vision) have shown that statistical patterns do correlate the distal and proximal cues.

Hypotheses of vision (2003):

Visual percepts are manifestations of the accumulated influence of visual experience with inherently ambiguous stimuli.

Therefore understanding what we see and why will depend on understanding the probabilistic relationship between stimuli and their sources that has shaped human visual physiology and its perceptual consequences.

We as species are here and survive because many billions of our ancestors have managed to resolve the complexities and ambiguities of vision.

So lets revisit: we have distal and proximal cues/stimuli. The distal stimuli arise from the real world. Proximal stimuli are patterns of light energy arriving at the eye and forming retinal images.

We have no option, but to use proximal stimuli to perceive the world. Does proximal stimuli represent the distal world very well ?

Some empirical studies have revealed that statistical patterns do correlate the distal and proximal cues.

Hence, the final comment of Brunwick:

Perceivers must behave like 'intuitive statisticians'.

- Theories of Visual Perception; by Ian E. Gordon; Psychology Press (Taylor and Francis Group), 2004
- Computational Neuroscience of Vision; by Edmund Rolls, Gustavo Deco; Oxford University Press, 2002
- Curvature Scale Space Representation: Theory, Applications and MPEG-7 Standardization; Farzin Mokhtarian and Miroslaw Bober; Kluwer Academic Publishers, 2003
- Multiple View Geometry in Computer Vision; R. Hartley and A. Zisserman; Cambridge University press, 2003.
- Machine Vision ; W. E. Snyder and H. QI; Cambridge University press, 2004.
- Introduction to Statistical pattern recognition; K. Fukunaga, 2nd Ed. Academic Press, New York, 1990.

End of Lectures on

Gestalt
and
Brunwick

more for the batch next year