

IST 605: Human Information Processing

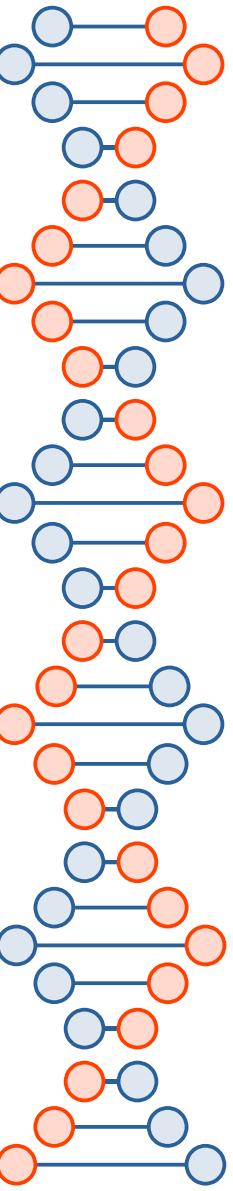
Perception

Denis L. Nkweteyim

02/12/25

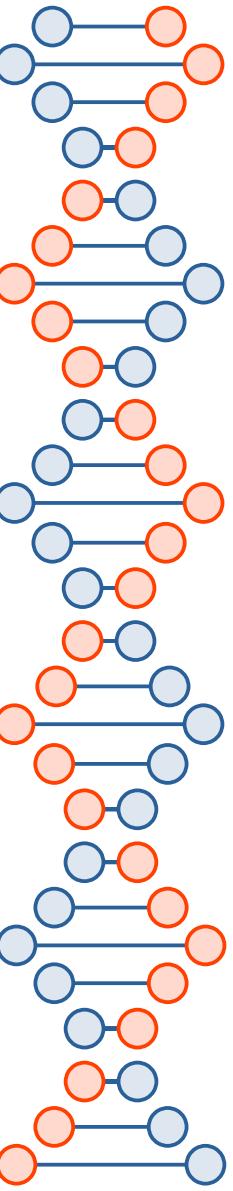
Denis L. Nkweteyim@UB - IST605 (HIP - Perception)

1



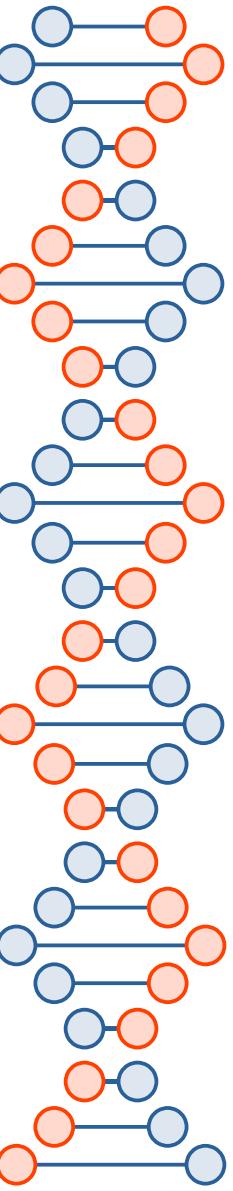
Outline

- The Problem of Perception
- Visual Perception
 - Low-Level Vision
 - Localization
- High-Level Vision
 - Feature Detection Theories
 - Structural Theories
 - Template Matching and Alignment
- Levels and the Integration of Information in Perceptual Context Effects
 - The Word Superiority Effect

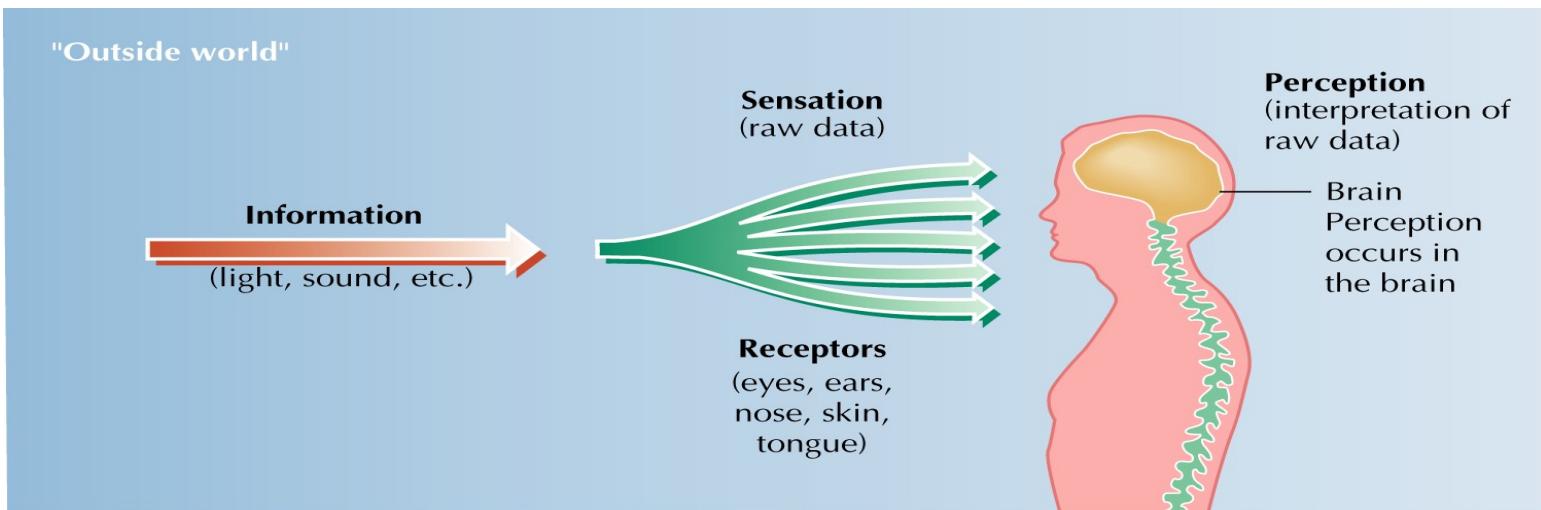


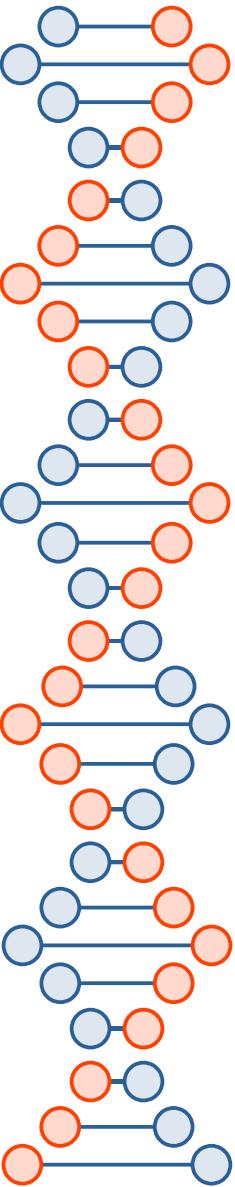
Perception

- Begins with sensation
 - Sensation is the process of receiving, converting, and transmitting information from the outside world.
 - Sensory organs contain receptors that transduce sensory energy into nerve impulses that are carried to the brain



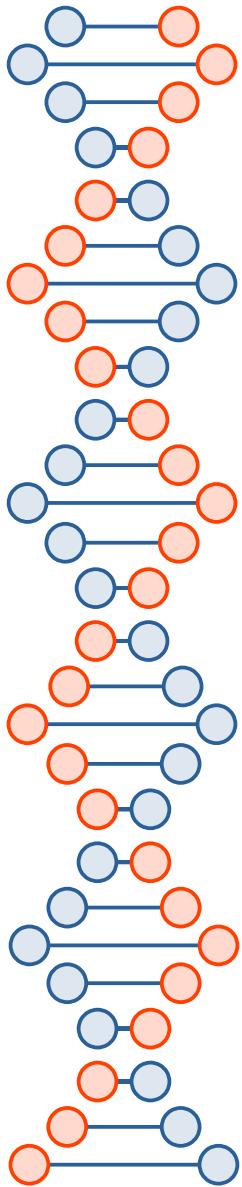
Perception





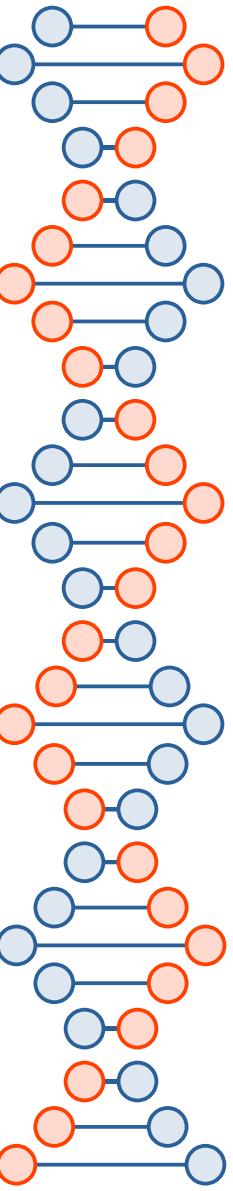
Understanding Sensation: (1) Processing

- Three Types of Processing
 - Transduction
 - The conversion of sensory stimuli into neural impulses that are sent on to the brain
 - i.e., sensory receptors convert a physical energy into nerve impulses
 - Sensory Reduction
 - Filters and analyzes incoming sensations before sending on to the brain
 - Coding
 - Converts particular sensory input into a specific sensation sent to differing parts of the brain



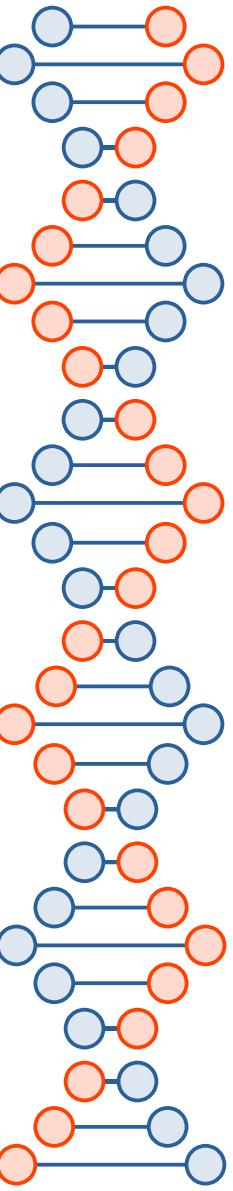
Understanding Sensation: (2) Sensory Thresholds

- Thresholds informed by Psychophysics
 - Subfield of psychology devoted to the study of physical stimuli and their interaction with sensory systems
 - Psychophysical tasks have been extensively used to draw conclusions on how information is processed by the visual and other sensory systems
- Testing limits and changes of sensory systems
 - Absolute Threshold
 - Smallest amount of a stimulus we can detect
 - Difference Threshold
 - Minimal difference needed to detect a stimulus change
 - Also called the just noticeable difference (JND)



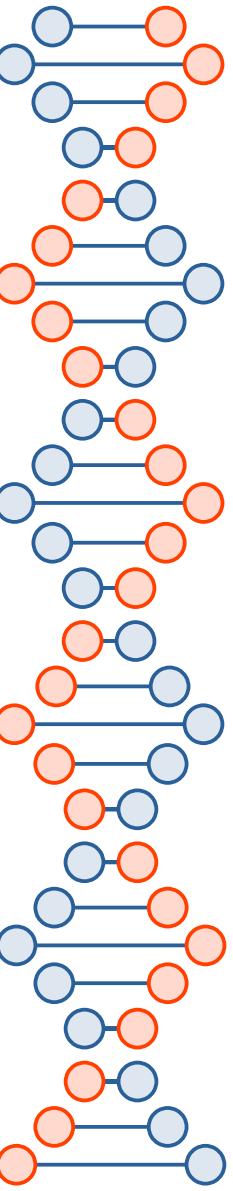
Perception

- Perceptual processes include
 - Selection: choosing which of many stimuli will be processed
 - Organization: involves collecting the information into some pattern
 - Interpretation: involves understanding the pattern
- Perceptions can be in error
 - For vision, leads to illusions
 - i.e., visual stimuli that are misinterpreted



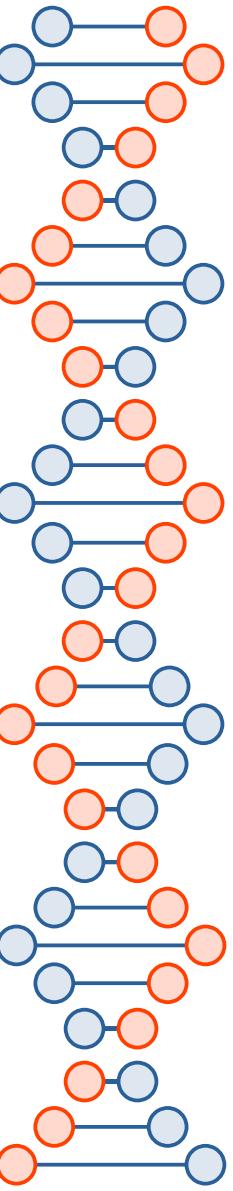
The Problem of Perception

- We perceive the world around us
 - Eyes – sight
 - Ears – sound
 - Feeling – of objects, warmth of surfaces, position of our limbs, etc.
- Various sense organs involved, yet we end up with a coherent world
- Yet, perception is very hard for a computer



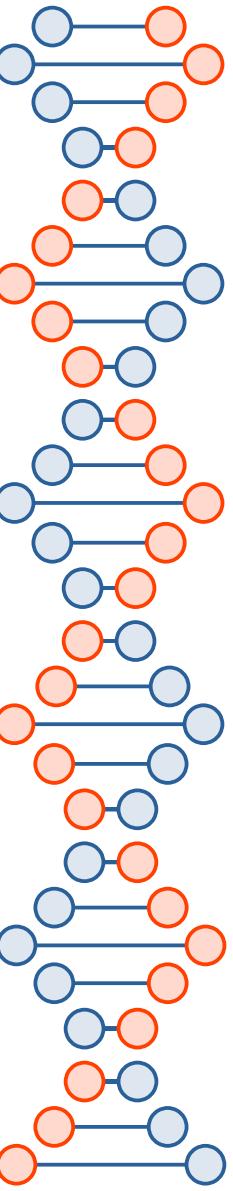
Why is Perception Difficult?

- Ambiguities in the world
- Two categories of ambiguity
 - Same objects seen from different vantage points and under different lighting conditions
 - Images we see are a result of light rays that have bounced off objects in 3-D space onto the retina, which is 2-D
 - An unlimited number of 3-D objects can result in the same 2-D representation on the retina, which the brain has to process
- We are not aware of these ambiguities, and amazingly, our perceptual system is capable of resolving them before we experience the visual world



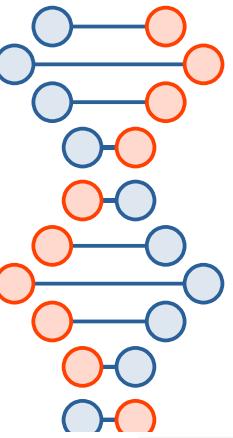
Signal Detection Theory

- Most of the time, sensation and perception are not completely clear cut
 - True for humans as well as machines
- Signal detection theory provides
 - A means to measure the ability to differentiate between information-bearing patterns (called stimulus in living organisms, signal in machines) and noise (i.e., random patterns that distract from the information)
 - The theory explains how we make decisions about stimuli against a background of uncertainty, i.e., in ambiguous situations
 - The goal of the decision-maker is to tease out the decision signal from background noise



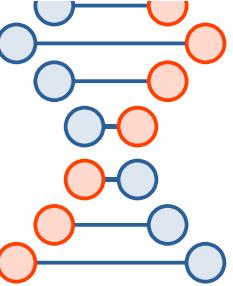
Signal Detection Theory

- Our decision about whether a signal is present or absent is a function of the strength of the stimulus, the strength of the noise, past experience, and expectations
- Key concepts
 - Signal: The stimulus that is being detected (e.g., a sound, a visual image)
 - Noise: Distracting information or random background interference that obscures the signal
 - Sensitivity: An individual's ability to discriminate between the signal and noise. This is independent of their decision-making criterion.
 - Criterion
 - Point at which signal is strong enough for us to say the stimulus is present
 - Criterion can be biased, i.e., shifted by many things outside of pure sensory processes
 - e.g., a person might be cautious, tending to say "no" more often, or liberal, tending to say "yes" more often

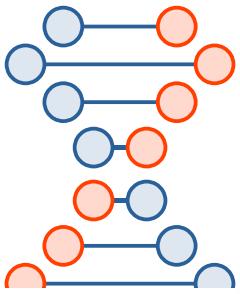


Signal Detection Theory

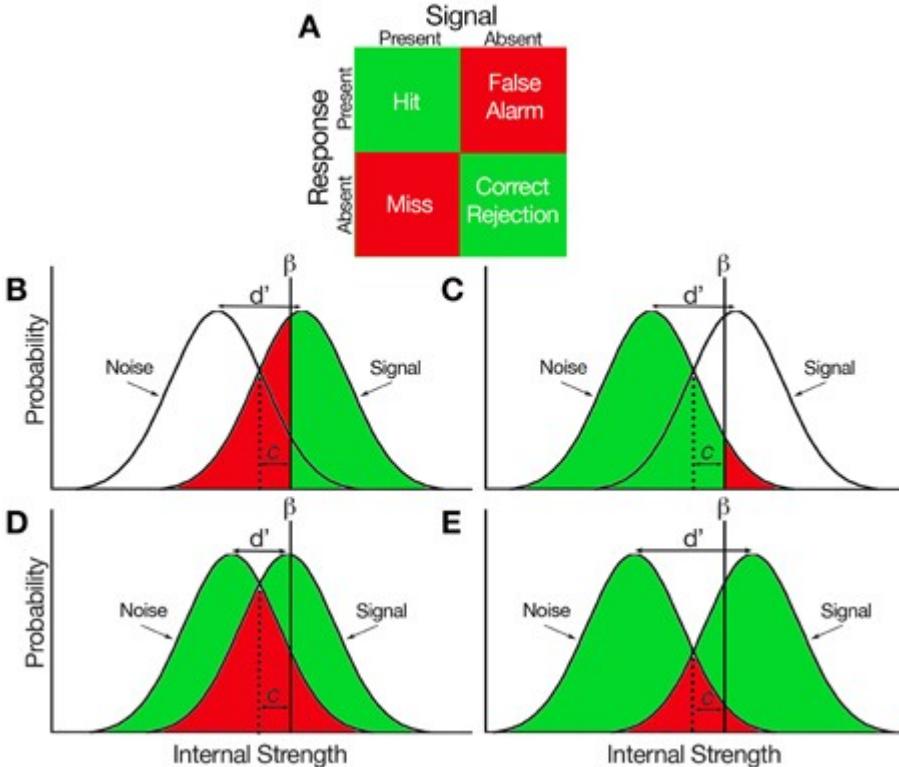
		Signal	
		Present	Absent
Response	Yes	Hit	False Alarm
	No	Miss	Correct Rejection



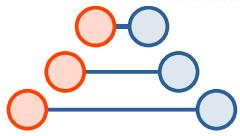
- Possible outcomes
 - Hit
 - The signal was present, and the person responded "yes".
 - Miss
 - The signal was present, but the person responded "no".
 - False Alarm
 - The signal was absent, but the person responded "yes".
 - Correct Rejection
 - The signal was absent, and the person responded "no"

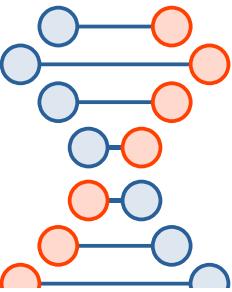


Signal Detection Theory – Illustration

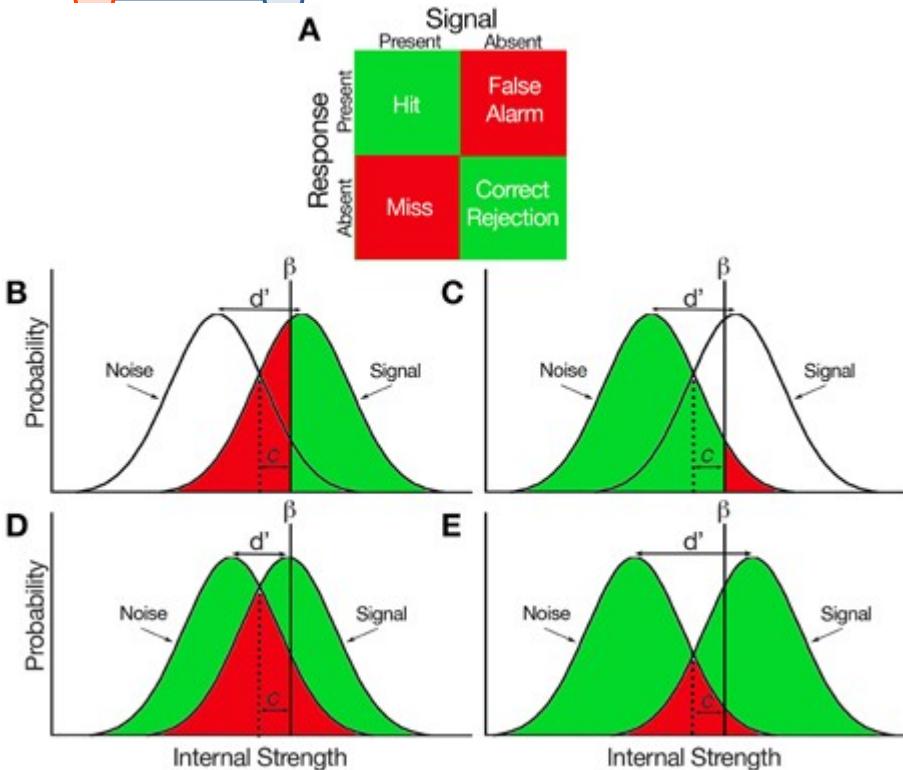


- Criterion (β)
 - A signal will be reported present when the internal signal is stronger than β and absent when the internal signal is weaker than β
- A hit
 - The probability that the subject reports the signal present when it is (Figure 1B, green)
- False alarm
 - The probability that the subject reports the signal present when it is absent (Figure 1C, red)
- A miss
 - The probability that the subject reports the signal absent when it is present (Figure 1B, red)
- Correct rejection
 - The probability that the subject reports the signal absent when it is absent (Figure 1C, green)

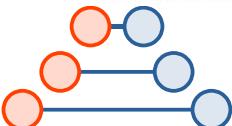


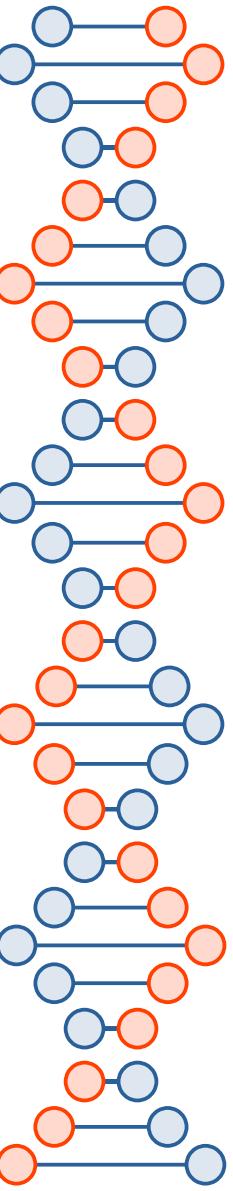


Signal Detection Theory – Illustration



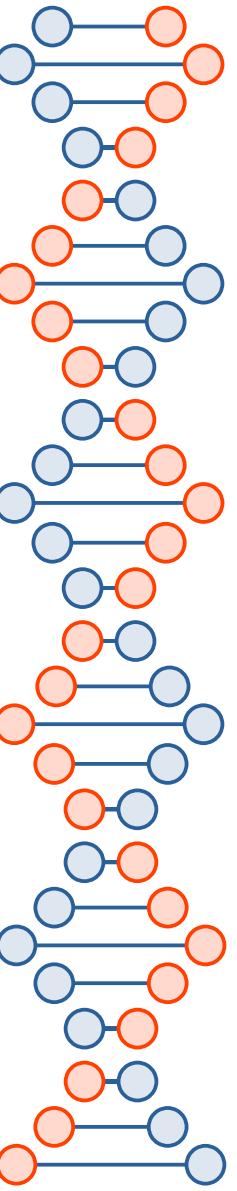
- Green indicates correct decision, red indicates incorrect decision
- (B) Proportions of hits and misses represented under the signal distribution
 - β reflects the subject criterion, c reflects bias, and d' reflects sensitivity which represents the difference in position between the two distributions
- (C) Proportions of false alarms and correct rejections represented under the noise distributions
- (D) A condition which hypothetically reflects low subject sensitivity
 - When the distributions are closer together (i.e., d' is smaller), the difference between the proportion of hits and false alarms is lower
- (E) A condition which reflects high subject sensitivity
 - When the distributions are farther apart (i.e., d' is larger), the difference between the proportion of hits and false alarms is higher





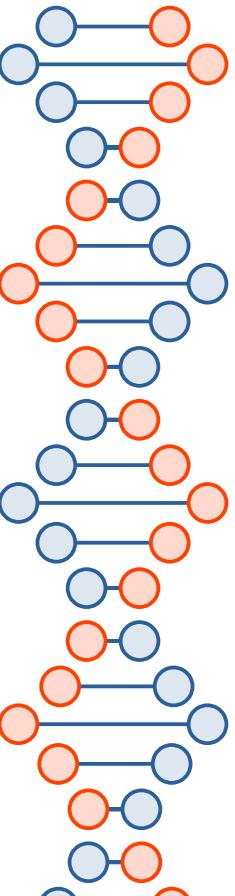
Signal Detection Theory

- Some application areas
 - Psychology
 - Sensory perception: Explaining how people hear a faint sound in a noisy room
 - Medicine and diagnostics
 - Medical imaging: Evaluating the accuracy of radiologists who must distinguish between a true signal (e.g., a tumour on an X-ray) and noise (e.g., surrounding tissue).
 - Engineering and technology
 - Telecommunications: SDT is used to analyze the performance of communication systems, helping to determine how well a signal can be detected amidst background noise
 - User experience (UX) research
 - SDT helps designers understand how users interact with products by analyzing how often users correctly notice important information (hits) versus how often they miss it (misses).



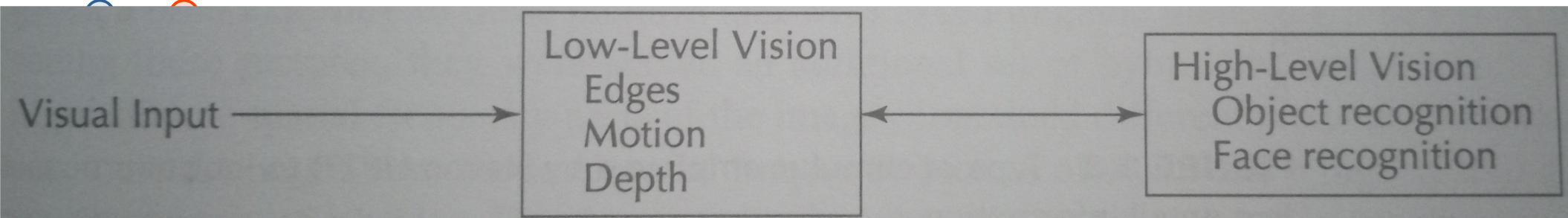
Outline

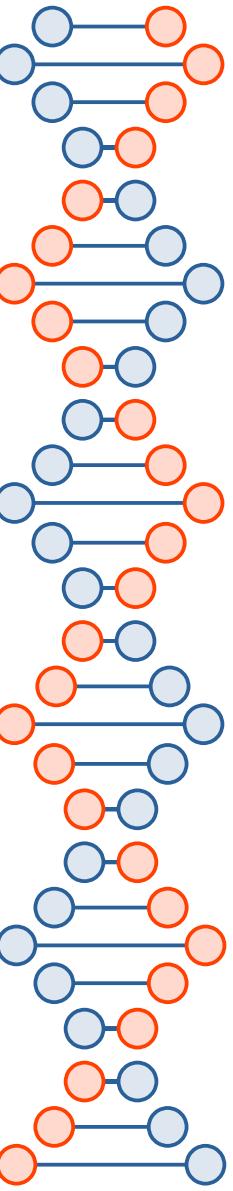
- The Problem of Perception
- Visual Perception
 - Low-Level Vision
 - Localization
- High-Level Vision
 - Feature Detection Theories
 - Structural Theories
 - Template Matching and Alignment
- Levels and the Integration of Information in Perceptual Context Effects
 - The Word Superiority Effect



Visual Perception

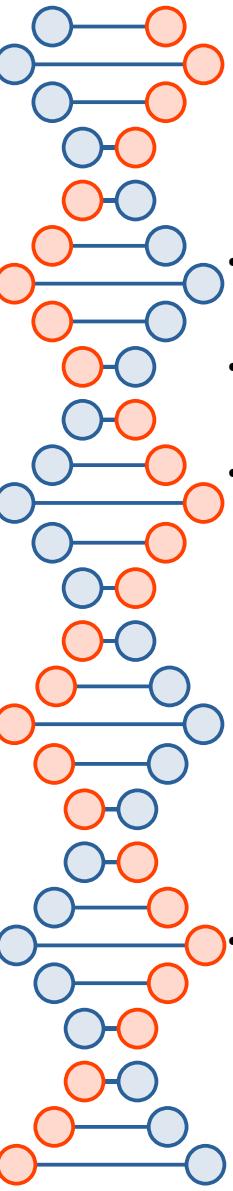
- An active process in which different levels of analysis interact to determine what we perceive and understand
- Low-level vision
 - Extracting preliminary information from the pattern of light reaching the retina
- High-level vision
 - Perception of larger scale elements of the world





Low-level vision

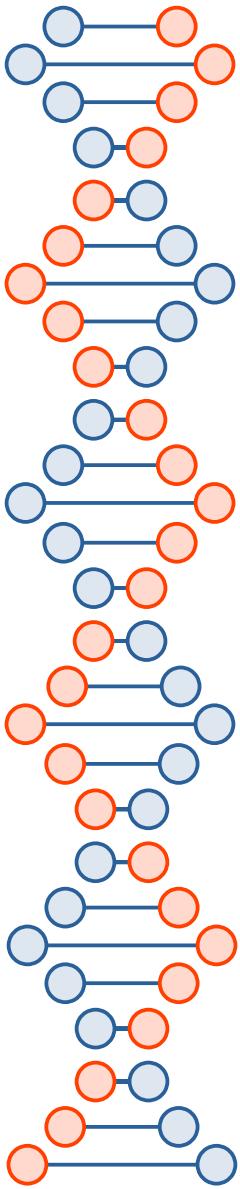
- Forming internal representations of the 3-D visual world, and extracting useful information to help people interact with their environment
 - Since there is an unlimited number of possible representations on the retina, the visual system seeks out information that will help the individual achieve their goals by extracting features, determining location of the objects, and how the objects are moving



Low-level vision – Edge Detection

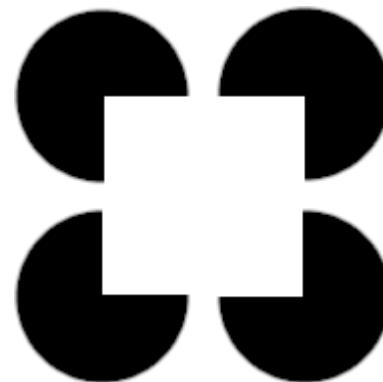
- Object boundary
 - Coarse edges of an object
- Parts of an object
 - The fine details of the object
- How does the visual system separate the boundaries from the fine details of an object?
 - The large letter H can be viewed as a figure with thick lines
 - The Ss that make up the figure are smaller, and have narrower lines
 - Edges like the ones in this picture are detected on the basis of low spatial frequency information
 - In contrast, edges like the ones that make up the narrower Ss are detected on the basis of high spatial frequency
 - Low spatial frequencies in edge detection correspond to gradual changes and coarse details, such as the overall shape and structure of an object
 - High spatial frequencies represent abrupt changes and fine details, like sharp edges and textures, which are crucial for identifying the precise contours of an image
- There is some evidence suggesting that processing of scenes begins with global (coarse) information and gradually incorporates local (i.e., fine detail) information
 - Other factors like size also play a part

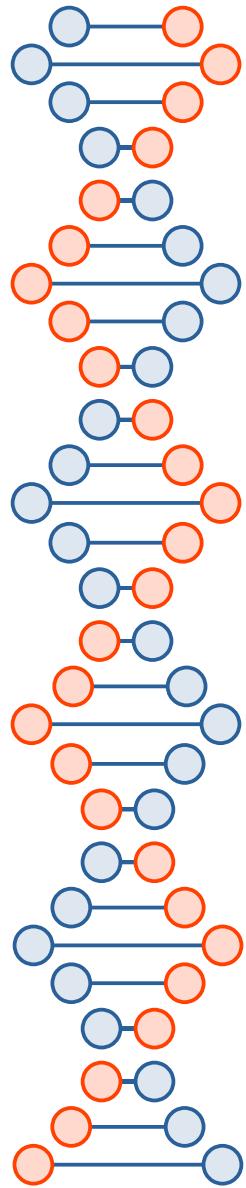
S S S S
S S S S S S S S
S S S S
S S S S



Low-level vision – Edge Detection

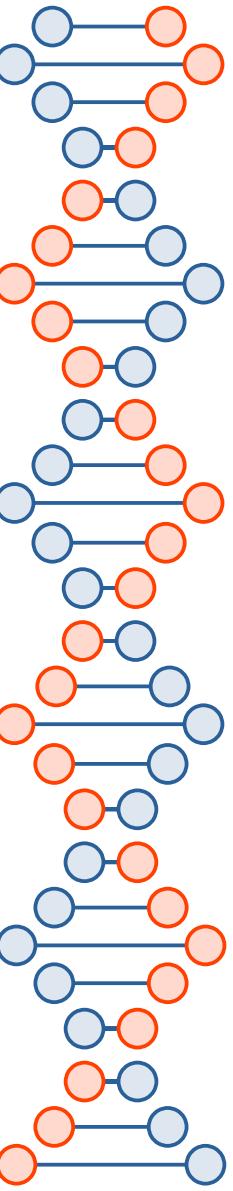
- Edge detection through differences in brightness between regions
 - Visual system enhances differences between dark and light regions in order to make edges stand out
 - In this image, not only is the square not there on top of the contours, the illusory square appears whiter than the surrounding white background





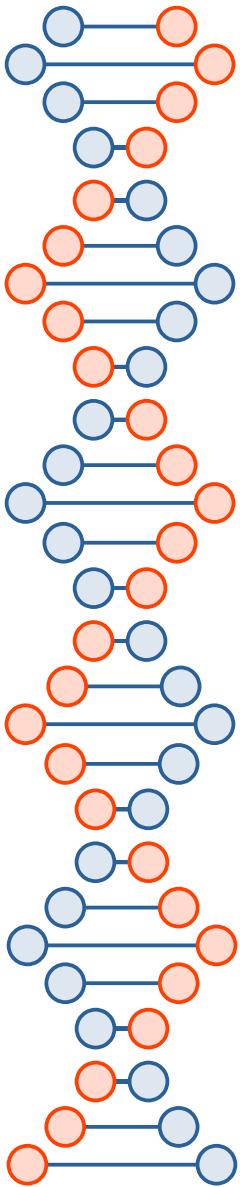
Low-level vision – Localization

- How do we
 - Locate an object in our environment?
 - Determine the position and movement of objects in the world?
- Steps
 - Recognize the object by segregating its visual information about each object from that of other objects (**segregation**)
 - Determine how far away the object is (**distance**)
 - If object is moving, get a sense of direction and speed



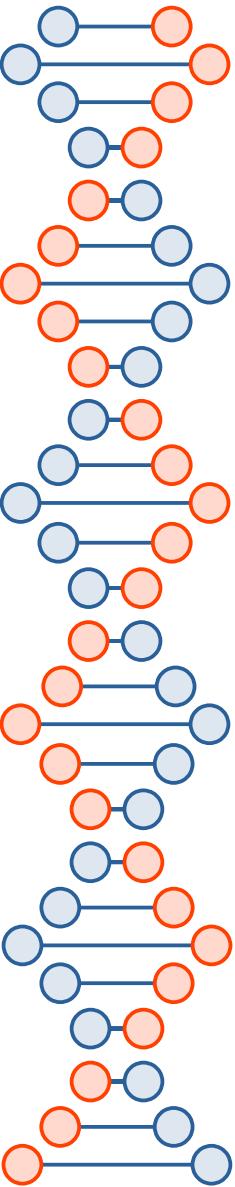
Segregation

- How does the perceptual system organize objects?
- Gestaltist psychologists proposed Gestalt Laws of Organization
 - Principles that describe how people perceive visual elements and organize them into unified wholes rather than separate parts
- Key Gestalt laws include
 - Similarity (grouping like items), proximity (grouping nearby items), closure (filling in gaps to see complete shapes), continuity (perceiving continuous, smooth lines), figure-ground (separating a foreground object from its background), and simplicity (perceiving the simplest possible arrangement)



Gestaltist Laws of Organization

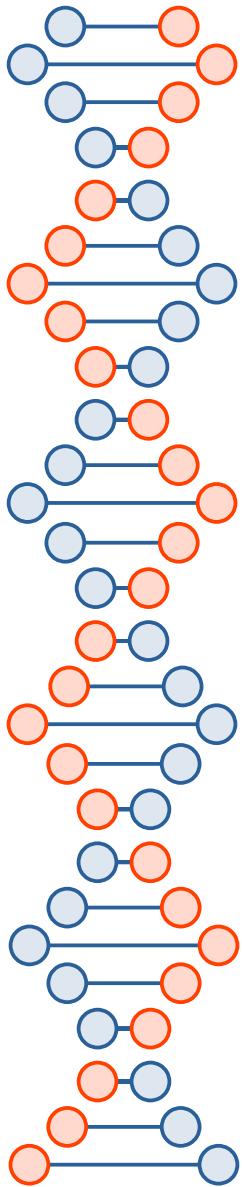
- Distinguishing between figure and ground
 - Blurriness
 - Objects in the foreground tend to be crisp and distinct while those in the background are blurry or hazy
 - Contrast:
 - High contrast between objects can lead to the perception of figure and ground
 - Size
 - Images that appear to be larger will be perceived as closer and part of the figure while those that are smaller will seem further away and part of the background
 - Separation
 - An object isolated from everything else in a visual scene is more likely to be seen as a figure and not background



Gestaltist Laws of Organization - Contrast

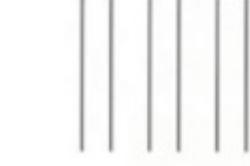
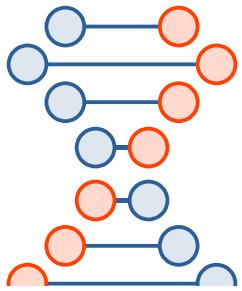
- High contrast between objects can lead to the perception of figure and ground
 - Example
 - Faces or vases illusion, also known as the Rubin vase
 - Depending on whether you see the black or the white as the figure, you may see either two faces in profile or a vase in the center





Gestaltist Laws of Organization

- Similarity: Objects that share similar characteristics, such as shape, color, or size, are perceived as being related or grouped together.
- Proximity: Objects that are close to one another are perceived as a group.
- Closure: The mind tends to complete incomplete shapes or figures by filling in missing information.
- Continuity: The brain prefers to see smooth, continuous patterns rather than disrupted or broken ones.
- Figure-Ground: This law describes the tendency to distinguish a figure from its background, separating elements into a clear foreground and background.
- Prägnanz (or Law of Simplicity): Also called the law of good figure, it states that people will perceive ambiguous or complex images in the simplest possible way.



Similarity

- The tendency to place items that look similar into a group

Proximity

- The tendency to place objects that are physically close to each other in a group

Closure

- The tendency to look at the whole by filling in gaps in a perceptual field



Continuity



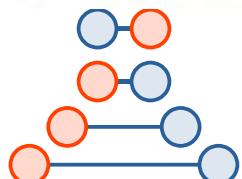
Connectedness

Continuity

- The tendency to perceive that movement of an object continues once it appears to move in a particular direction

Connectedness

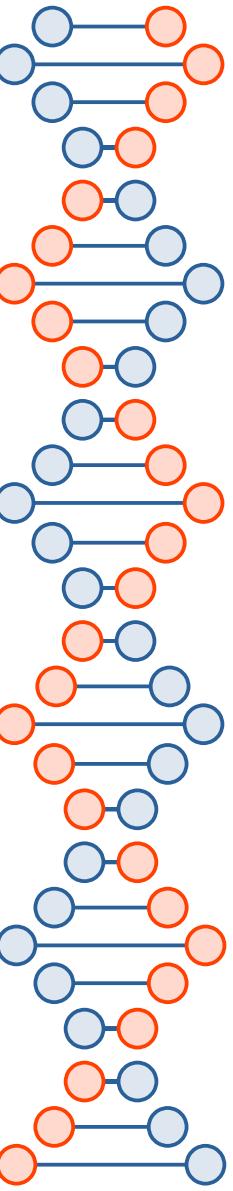
- Elements that are connected to each other by colour, lines, frames, or other means are perceived as more related and grouped than elements with no connection



02/12/25

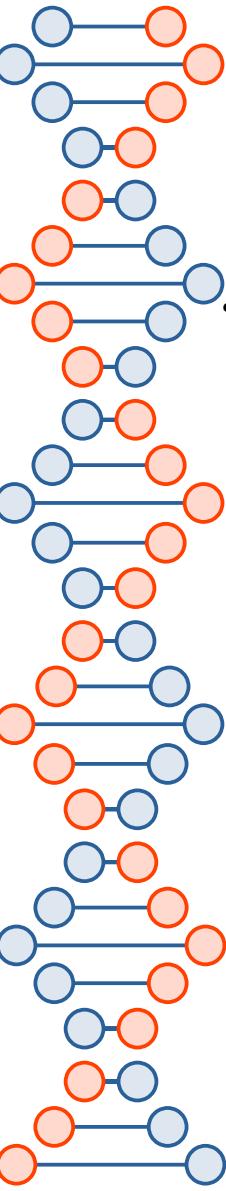
Denis L. Nkweteyim@UB - IST605 (HIP - Perception)

26



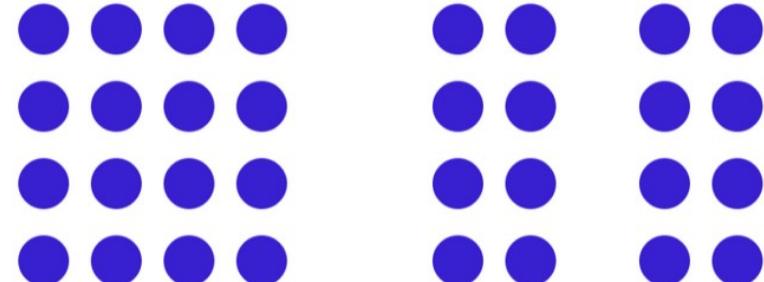
Gestalt Grouping Principles

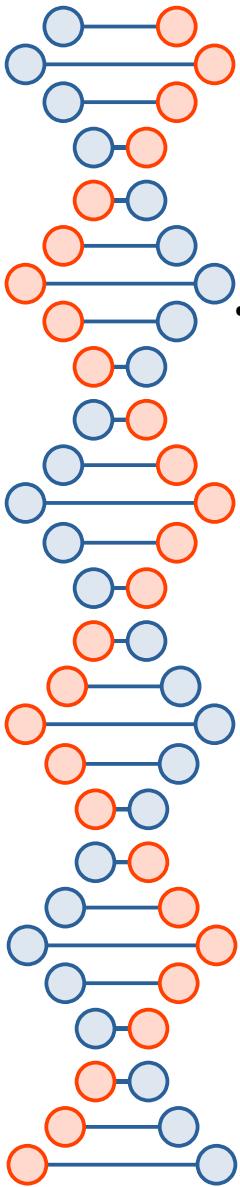
- Widely implemented in various phases of design
 - User Experience Design
 - User Interface Design
 - Interaction Design
 - Etc.



Gestalt Principles: Examples in Design

- Law of Similarity
 - Similar objects on the interface are often considered as a part of the same group, by the human eye
 - Achieved by using similar shapes, sizes, or colours
 - Enables the user to figure out the functionality behind the various interface elements.
- Law of Proximity
 - Objects placed close on an interface are perceived to be related to those set far apart
 - Can even override the similarity created by using the same colours or sizes
 - Sometimes placing the objects far apart creates a sense of negative space, which is useful to distinguish the objects from the rest





Gestalt Principles: Examples in Design

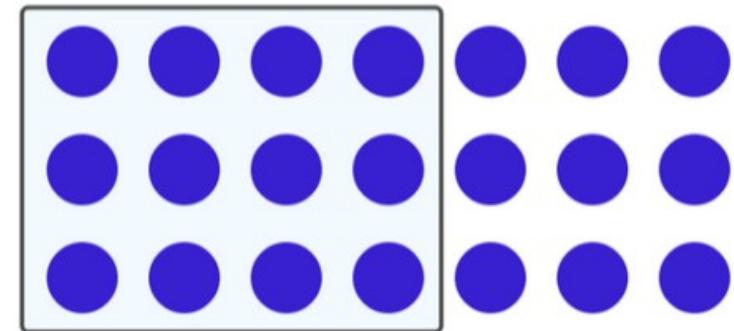
- Law of Figure Ground

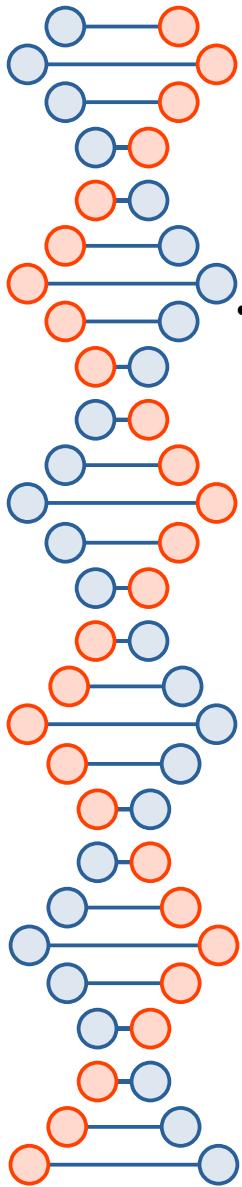
- We divide any visual into figure and ground
 - In this example, the image of the pro app logo is in the foreground and has the focus
 - The ground is the background of the visual
 - This law is widely used in creating logos for brands



- Law of Common-Region

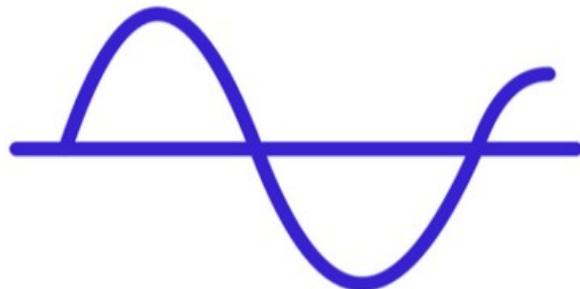
- Similar to the law of proximity
- Objects that are placed in the same enclosed area or space, are perceived to be related, e.g., creating cards on interfaces depicting unique user profiles



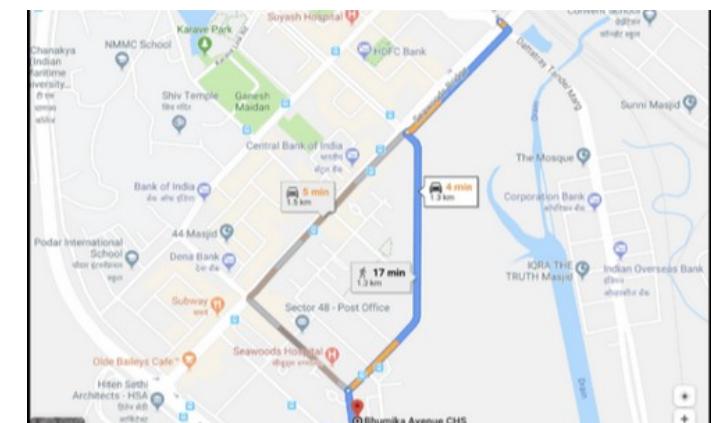


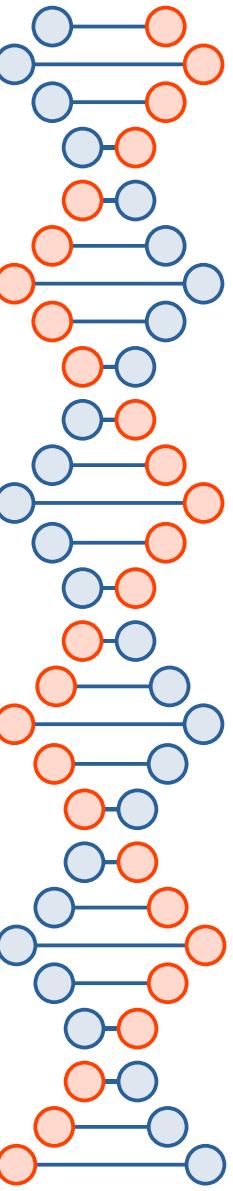
Gestalt Principles: Examples in Design

- Law of Continuity
 - The eye tends to follow the objects placed on a straight or curved line more smoothly than at broken lines or angles
 - They are even perceived as connected and help in the smooth navigation of a website or an application



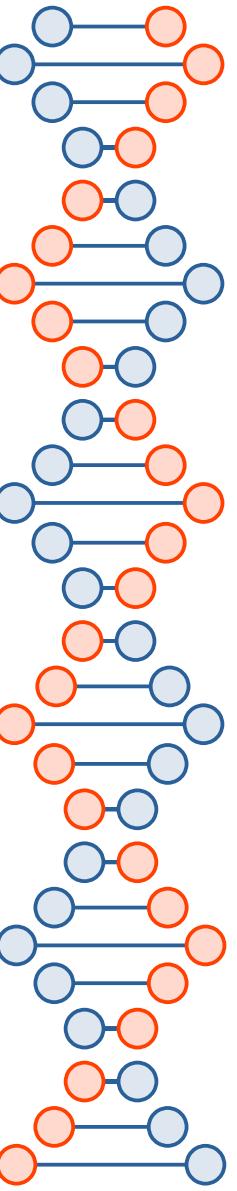
- Example: Law of Continuity – Google Maps
 - The eye follows the straight line and makes it easy to navigate the path





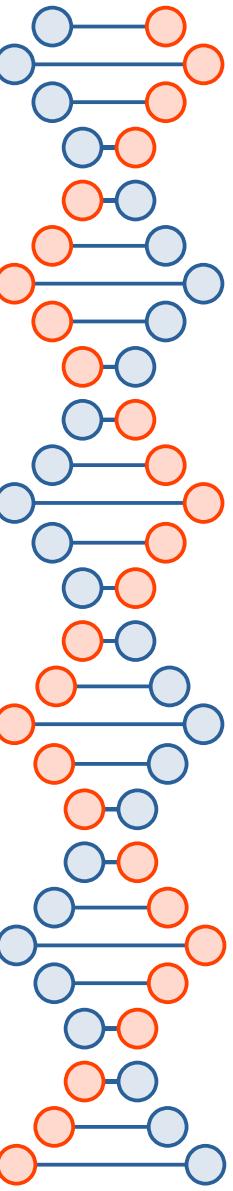
Low-level vision – Distance

- Retinal image gives information on where things are relatively to each other, vertically and horizontally
- But how is distance (depth) determined?
- Visual system uses several cues
 - Monocular cues
 - Texture gradients
 - Binocular cues



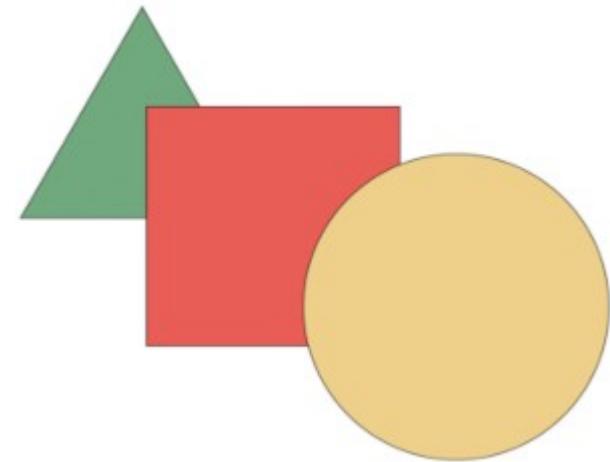
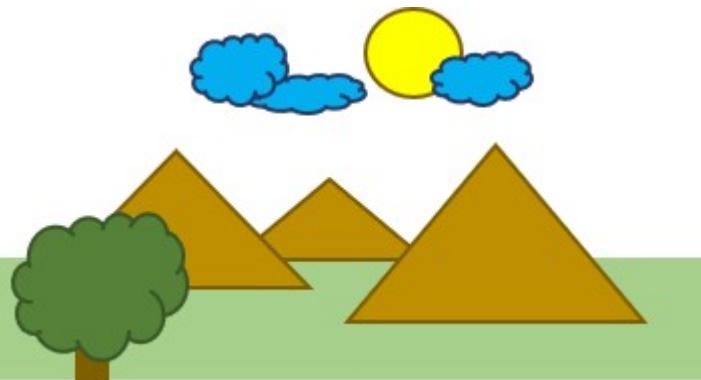
Monocular Cues

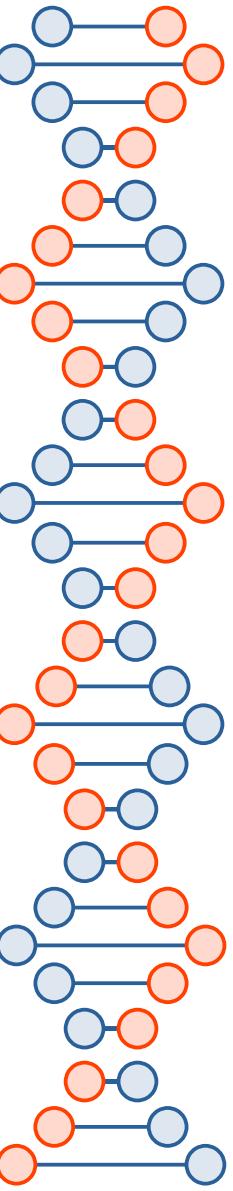
- Involve information coming from only one eye
 - Interposition
 - Linear perspective
 - Relative size
 - Shadows
 - Size constancy
 - Texture gradients
 - Motion parallax



Monocular Cues

- Interposition (occlusion)
 - When one object interrupts the contours of a second object, the first object is perceived as being in front of the second one

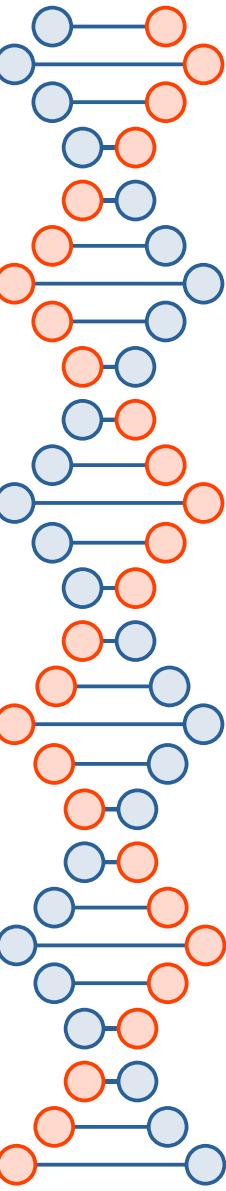




Monocular Cues

- Linear Perspective
 - When straight lines gradually converge, they are perceived as parallel, with the converging points seen as more distant





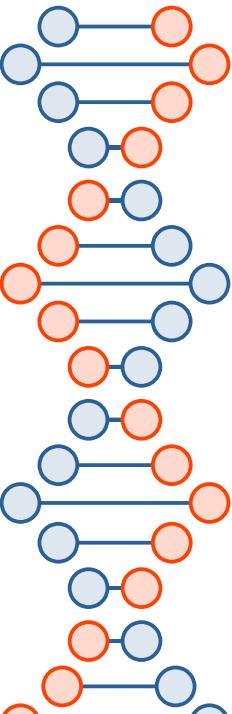
Monocular Cues

- Relative Size
 - If we assume that two objects are similar in size, we perceive the one that casts the smaller retinal image as farther away
 - Alternatively
 - The larger the object appears, the closer the object is to the viewer



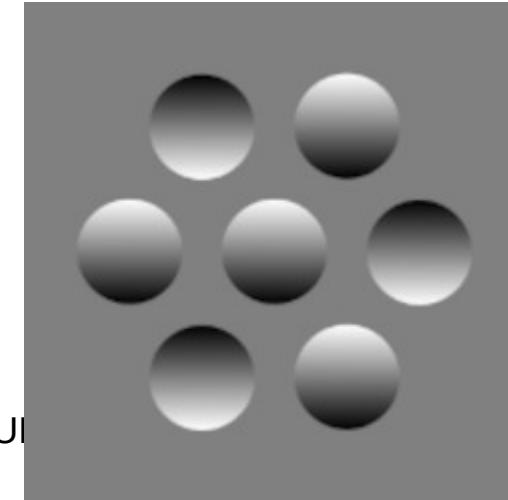
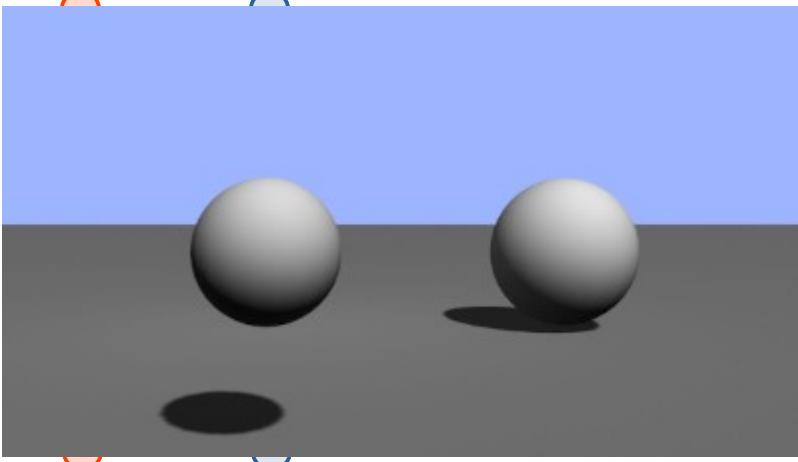
02/12/25



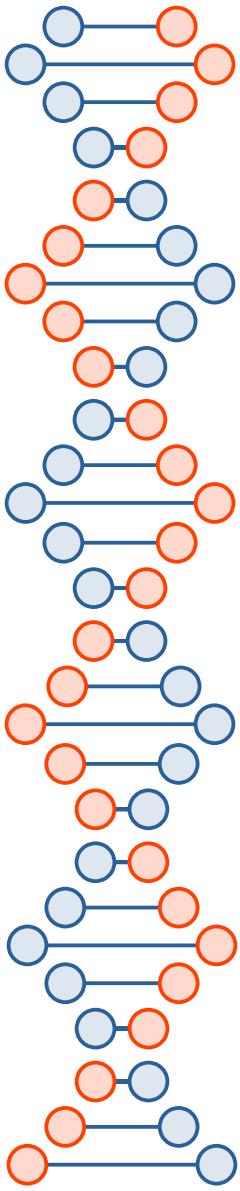


Monocular Cues

- Shadow
 - Visual system can use information from cast shadows to determine the apparent position of an object and its shape
 - Left picture
 - Both spheres occupy the same relative positions, but the one on the left is perceived to be floating and the one on the right resting on the surface
 - Right picture
 - Spheres have the same dimension, but shading is used to make some of them look more concave than others

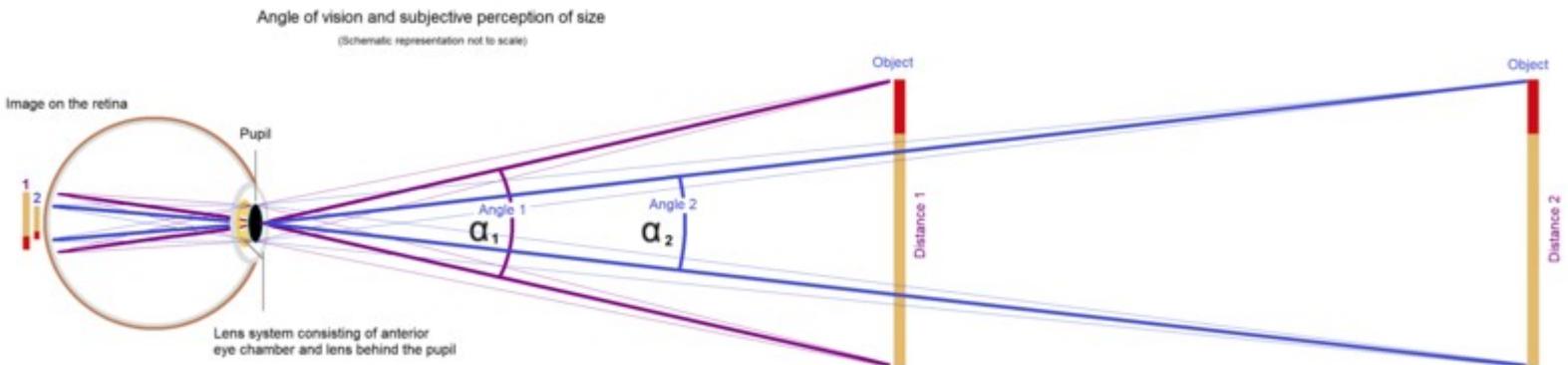
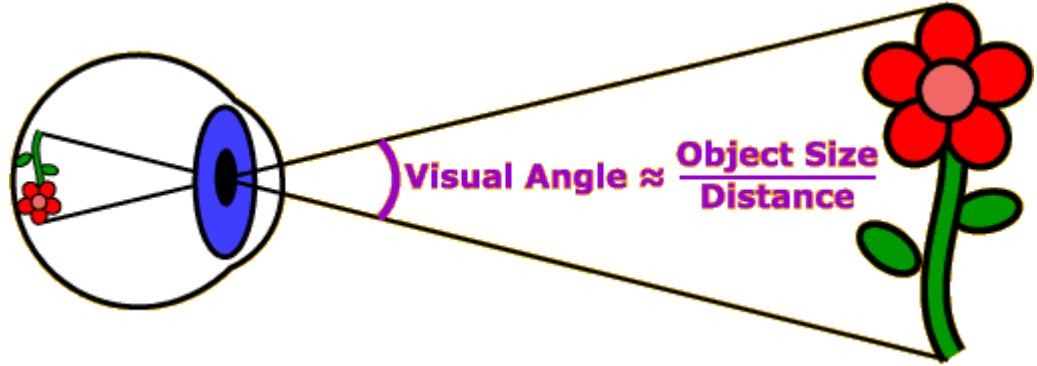


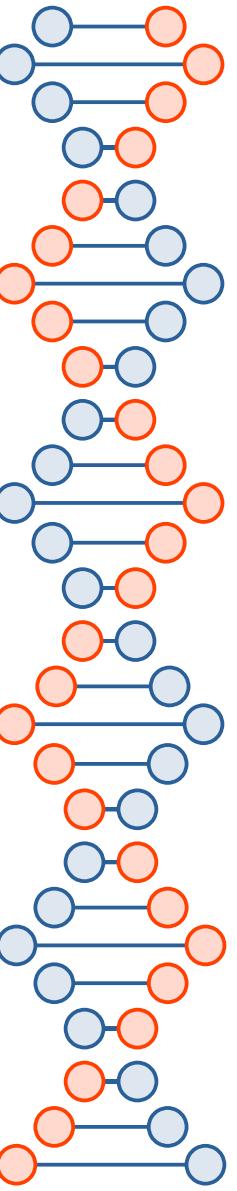
Monocular Cues



Size of an object on retina
depends on the visual angle

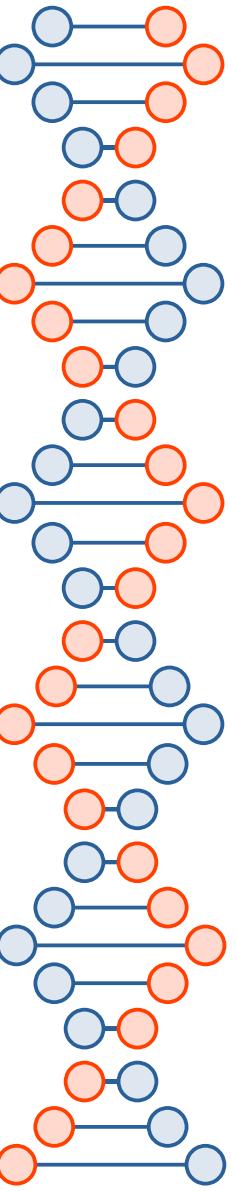
Same object has smaller
visual angle the farther away
it is from the eye





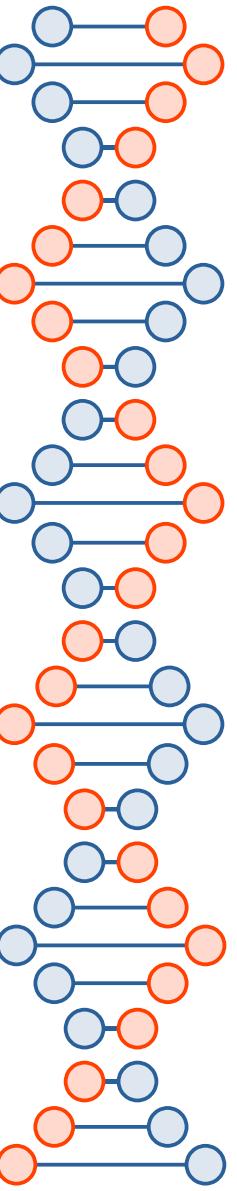
Monocular Cues

- Size constancy
- Shape constancy



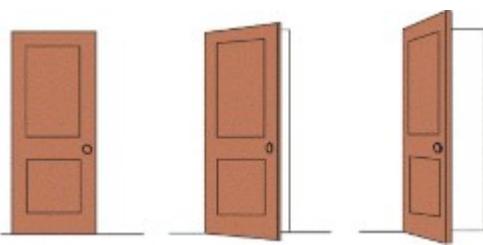
Size Constancy

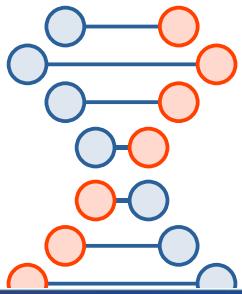
- An object observed from different distances is perceived to have a constant size (in spite of varying visual angle and image size on retina)
- Put in other words
 - As an object moves further or closer, its actual size stays the same
- Hence
 - As an object appears to become larger we realize it is **getting closer**, not bigger.
 - As an object appears to become smaller we realize it is **moving farther away**, not getting smaller



Shape Constancy

- A familiar object's shape perceived as constant when observed from various angles, positions and orientations

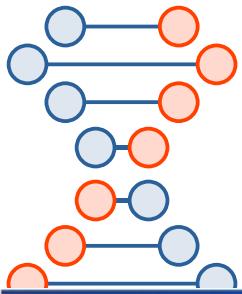




Size Distance Relationship Example 1

- Which of the horizontal bars is longer?

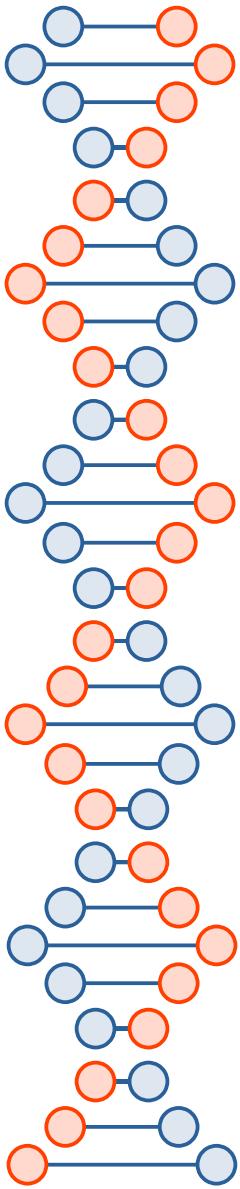




Size Distance Relationship Example 1

- The are actually the same size



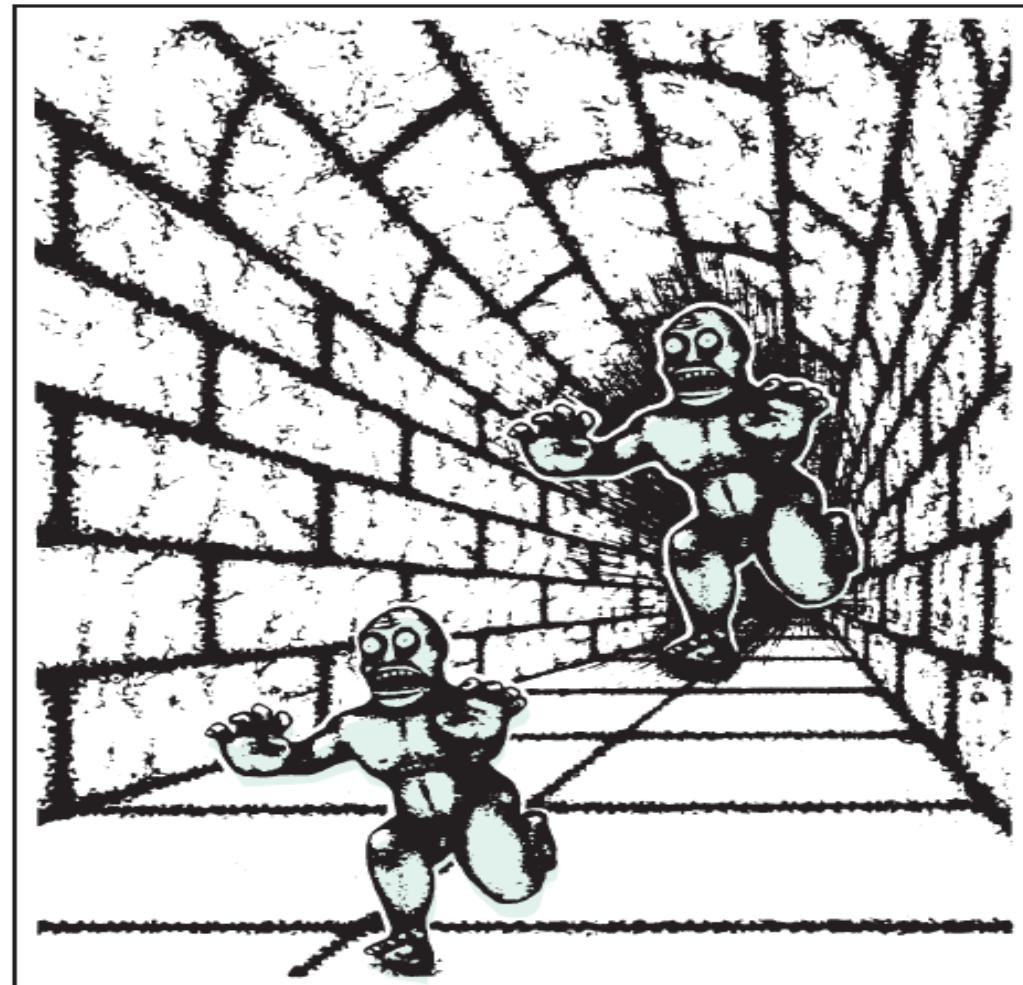


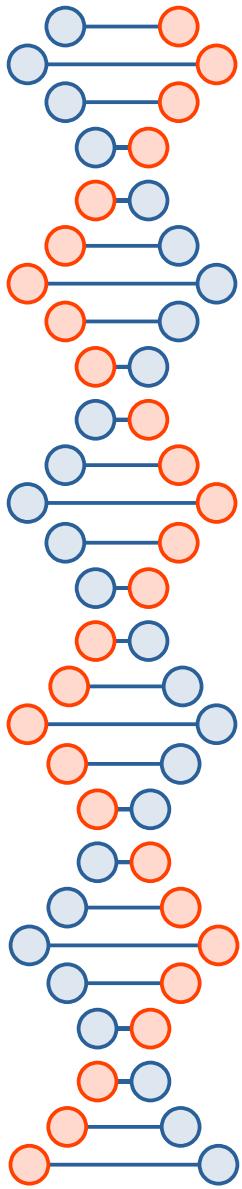
Size Distance Relationship Example 2

- Which of the beasts is bigger?

02/12/25

Denis L.



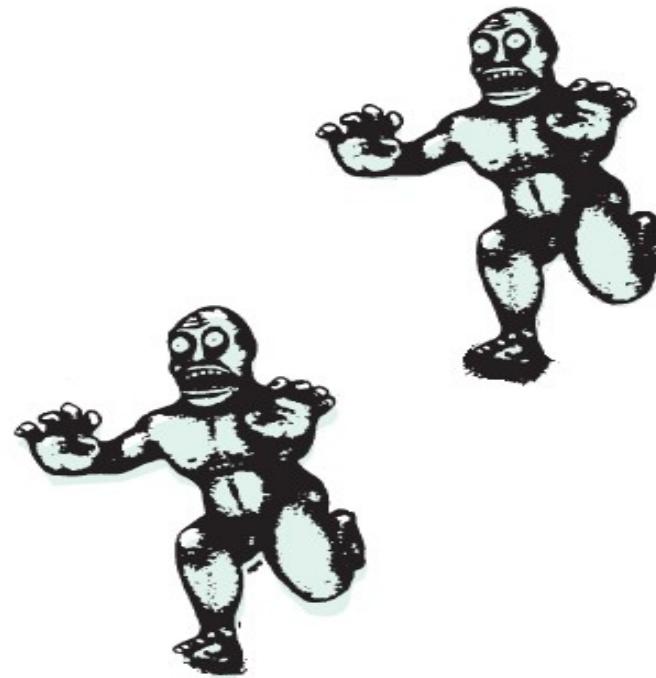


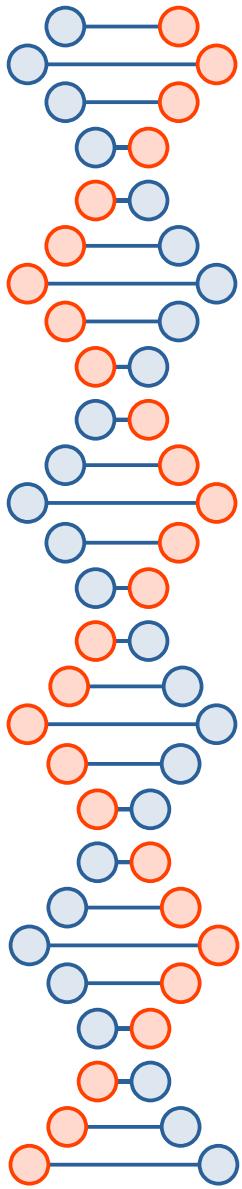
Size Distance Relationship Example 1

- Same size?

02/12/25

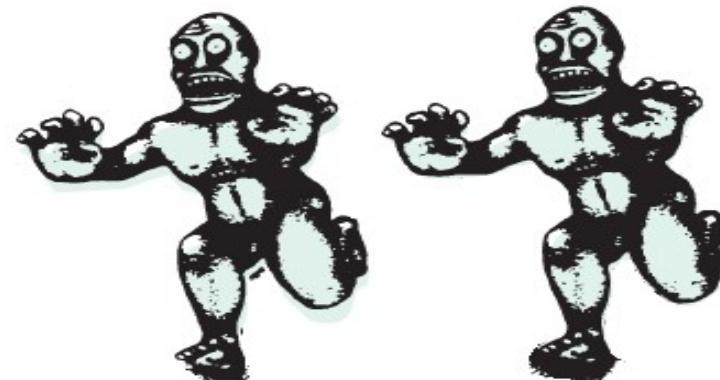
Denis





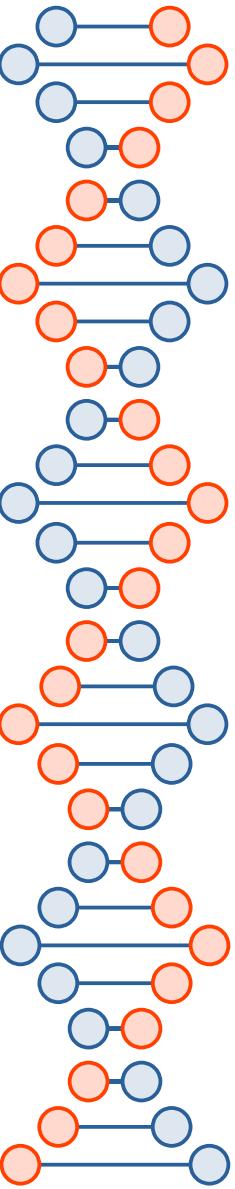
Size Distance Relationship Example 1

- Yes, same size?



02/12/25

Denis L

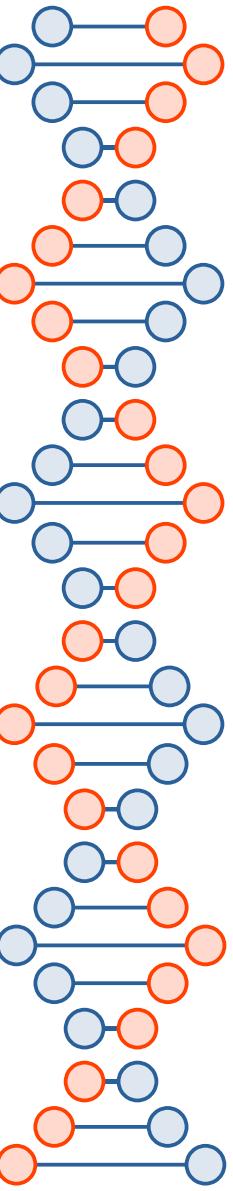


Monocular Cues

- Texture gradients
 - Method of determining depth by noting that distant objects have a smoother texture than nearby objects



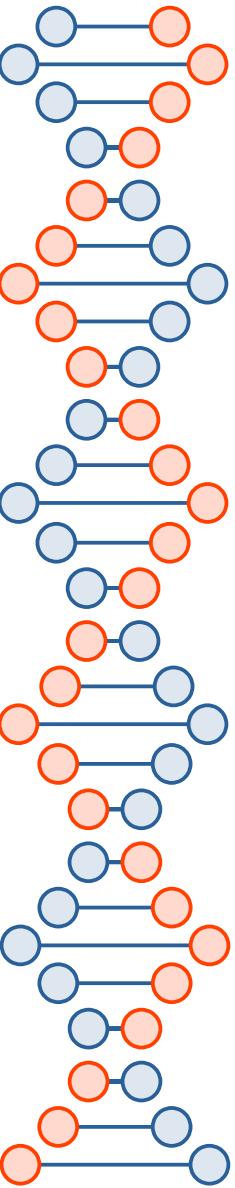
(perception)



Monocular Cues

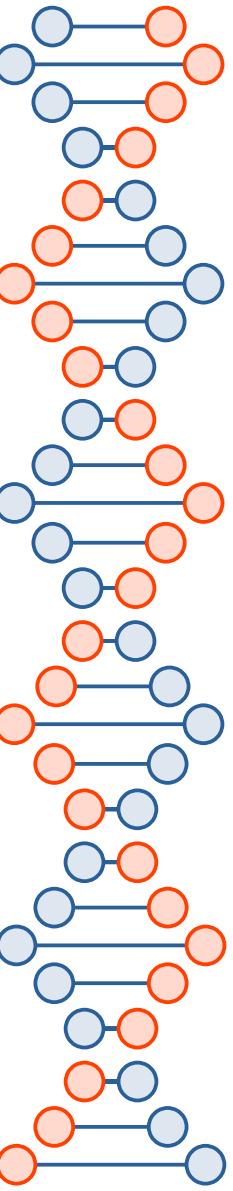
- Relative Clarity
 - Method of determining depth by noting that distant objects are less clear than nearby objects
 - Tends to work outdoors





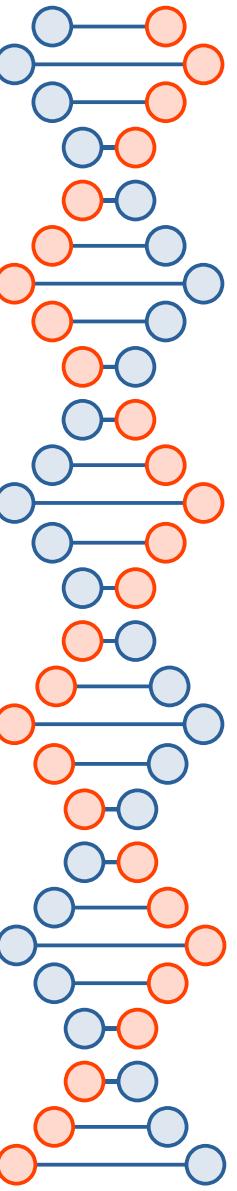
Motion Parallax

- So far, we have seen distance perception in situations in which both the observer and the observed scene are stationary
- In real life, we move about, and motion provides an important source of information about the spatial arrangement of objects around us
- As head is moved from left to right, images on the retina move from right to left
- Motion parallax, also known as relative motion, provides an effective cue to depth



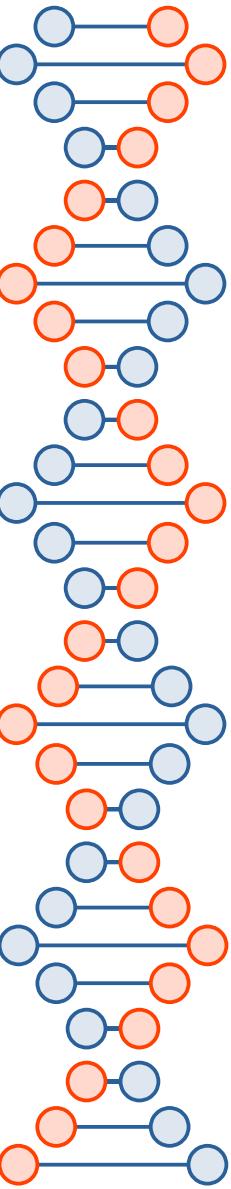
Motion Parallax

- A person who is moving can determine depth by focusing on a distant object
 - Nearby objects appear to move faster than far away objects
 - Examples
 - As you drive, nearby trees appear to pass by quickly, but the mountains move more slowly
 - The moon does not appear to move at all as you drive
 - Objects closer than the object of focus will appear to move in the opposite direction
 - Objects further away than the object of focus will appear to move in the same direction as the subject is moving



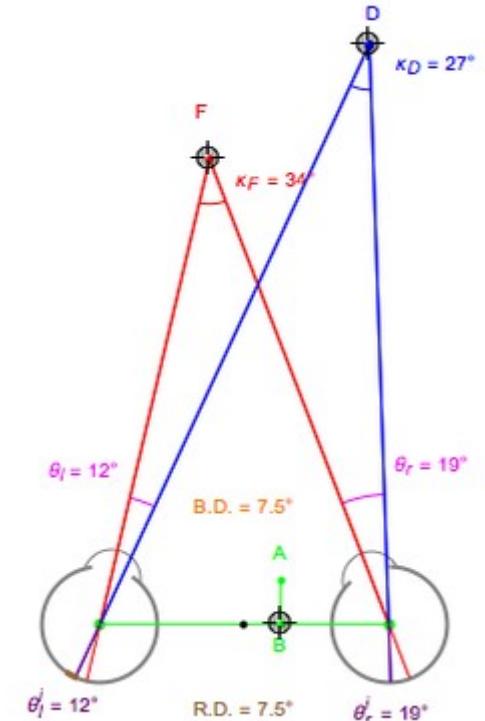
Binocular Depth Cues

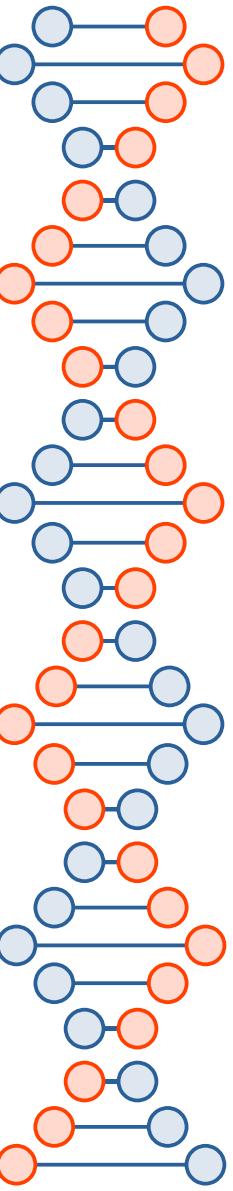
- Depth cues that require the use of both eyes
- Stereopsis
 - The process of interpreting information from both fields of vision (i.e., from binocular vision)
 - Major contributor to depth perception
- Binocular vision happens because each eye receives a different image because they are in slightly different positions on one's head (left and right eyes)



Binocular Depth Cues

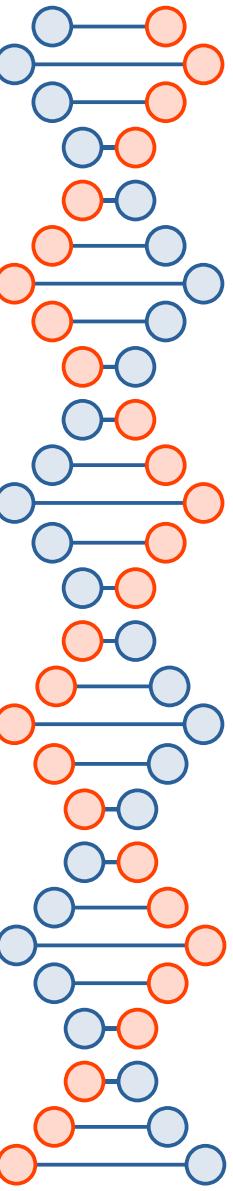
- Example
 - Points D and F are at different distances from the observer
 - They present slightly different retinal images
 - The angles (distance) between the two images on left eye (θ_l) is different (disparate) from the distance on the left eye (θ_r)





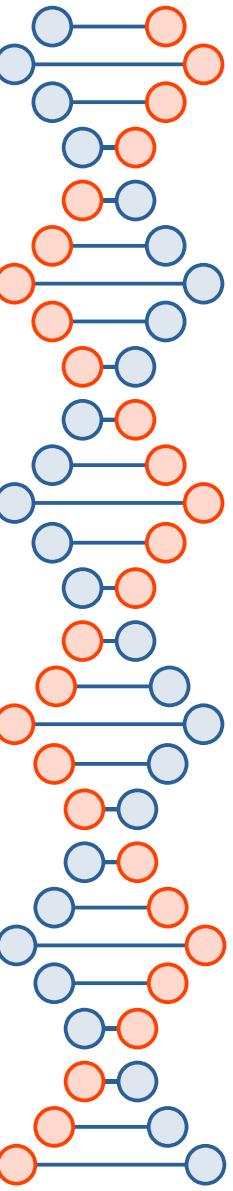
Binocular Depth Cues

- Retinal Disparity
 - A binocular depth cue for depth perception
 - Results from slightly different images produced by the retina of the left and the retina of the right eye
 - The brain merges these two different viewpoints into a single, three-dimensional image
 - Is most effective when the item is quite close to the person
 - As objects move further away, the retinal disparity decreases, and as they get closer, the disparity increases, a difference our brain uses to create a sense of depth



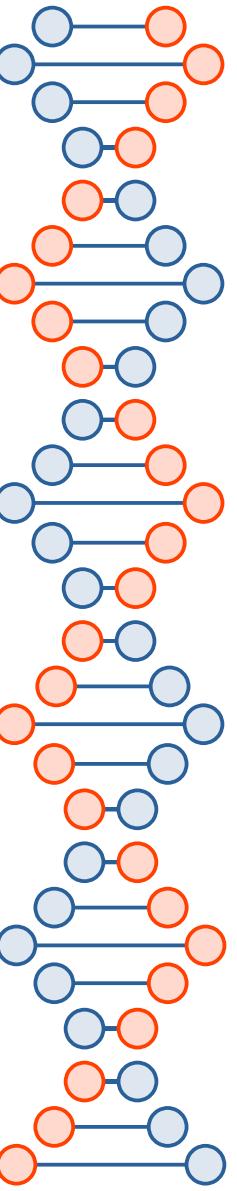
How retinal disparity works

- Different angles
 - The two eyes are separated by a small distance, so they capture the same scene from slightly different angles
- Brain integration
 - The brain receives these two distinct images and merges them into one coherent 3D perception
- Depth judgment
 - The amount of difference, or "disparity," between the two images helps the brain determine the distance of an object
 - Objects that are closer have a larger disparity, while more distant objects have a smaller one



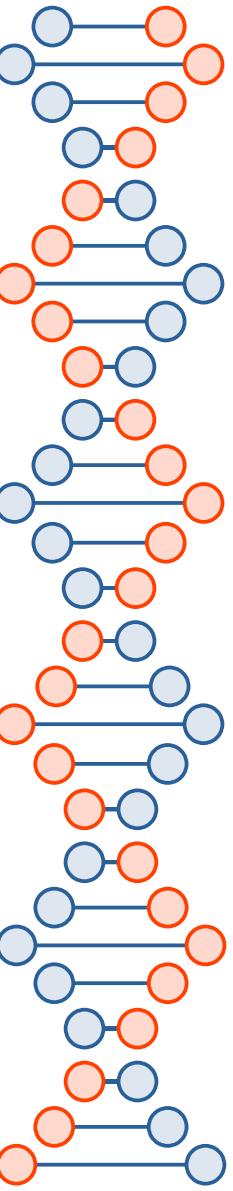
Outline

- The Problem of Perception
- Visual Perception
 - Low-Level Vision
 - Localization
- **High-Level Vision**
 - Feature Detection Theories
 - Structural Theories
 - Template Matching and Alignment
- Levels and the Integration of Information in Perceptual Context Effects
 - The Word Superiority Effect



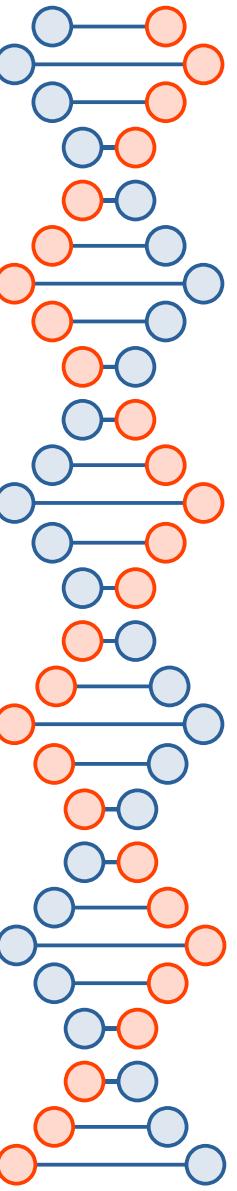
High-level Vision

- Refers to the later stages of visual processing
 - Focus is on understanding the meaning and identity of objects and scenes, rather than just analyzing low-level features like edges and colors
 - Involves interpreting visual input to recognize objects, understand their shapes and properties, and recognize the overall structure of the environment



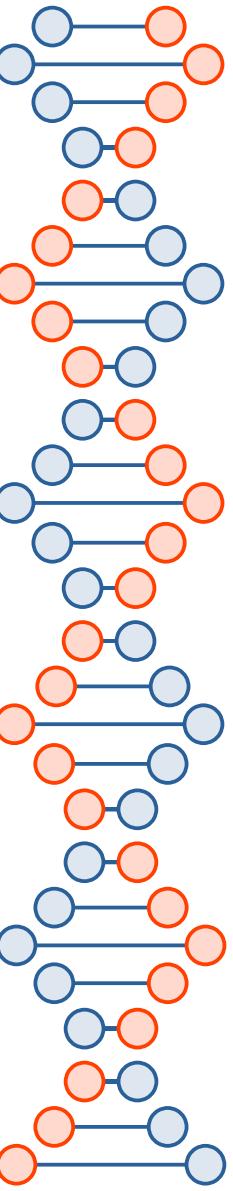
High-level Vision

- Key aspects of high-level vision
 - Object recognition: Identifying specific objects, such as "cup" or "face," from a visual scene
 - Scene understanding: Moving beyond single objects to interpret the overall scene, such as recognizing an event or the spatial relationships between objects.
 - Property estimation: Gauging physical properties of objects, like their shape, material, and position, not just the properties of the light on the retina.
 - Integration with cognition: Connecting visual information with other cognitive processes like short-term and long-term memory to give the input meaning.
- In computer vision
 - High-level vision concepts are central to artificial intelligence, with deep learning models like convolutional neural networks used to perform complex tasks like image recognition, object detection, and scene classification



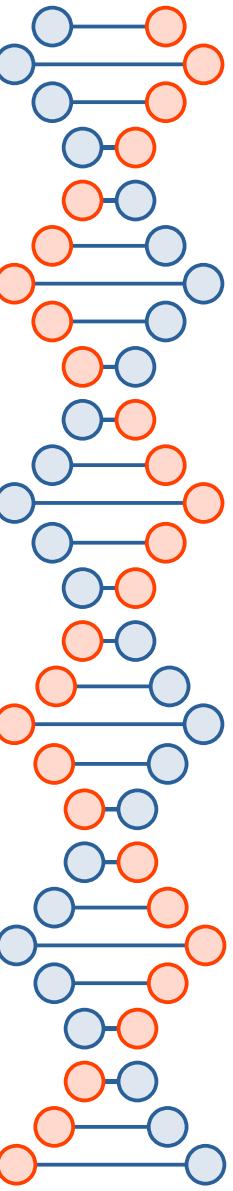
High-level Vision

- How do we identify (recognize) objects in the world based on visual properties like their shape
 - Feature detection theories
 - Structural theories
 - Template matching and alignment



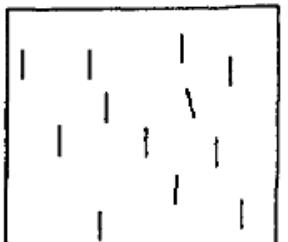
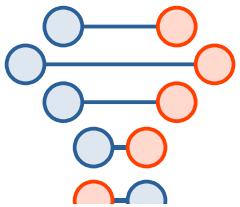
Feature Detection Theories

- Key idea of feature-based (i.e., featural) models
 - Objects are composed of separable, distinct parts (i.e., features)
 - A pencil is similar to a pen because both are about the same size and shape and both are used for writing
 - Also, features provide a vocabulary for constructing objects
 - Just as 26 letters of the English alphabet can be used to compose more than 100,000 words
 - A small set of visual features may be sufficient to describe a large number of objects
- According to featural models
 - Recognizing an object involves decomposing a complex stimulus into its features and then matching the features against the features in the representations of objects in long term memory
 - Just as each letter of the alphabet has unique features, objects may also have a unique set of features

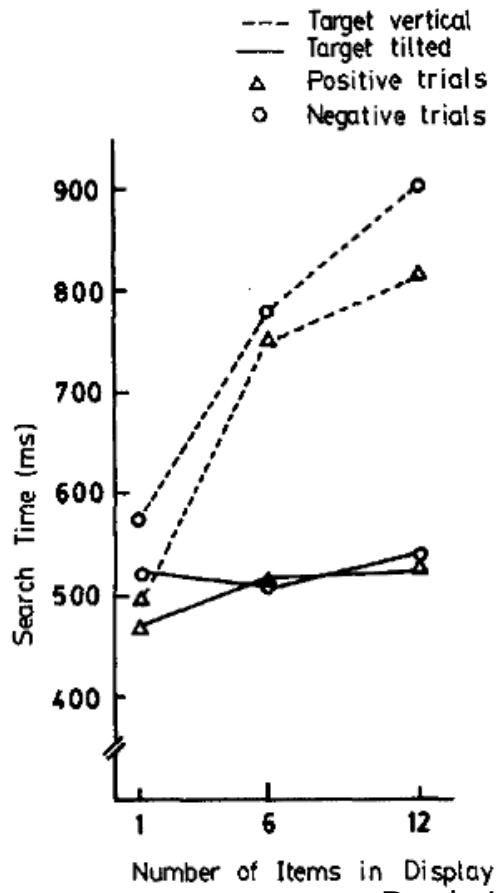


Feature Detection Theories

- Visual search task
 - The process of locating a target object among distractors
 - Commonly used in laboratory settings to study visual attention for simple objects, where participants search for a target defined by a known feature, like colour
 - Performance is typically measured by how quickly and accurately a person finds the target
 - Sample experiment
 - People either looked for a vertical line segment against a background of tilted line segments, or looked for a tilted line segment against a background of vertical line segments



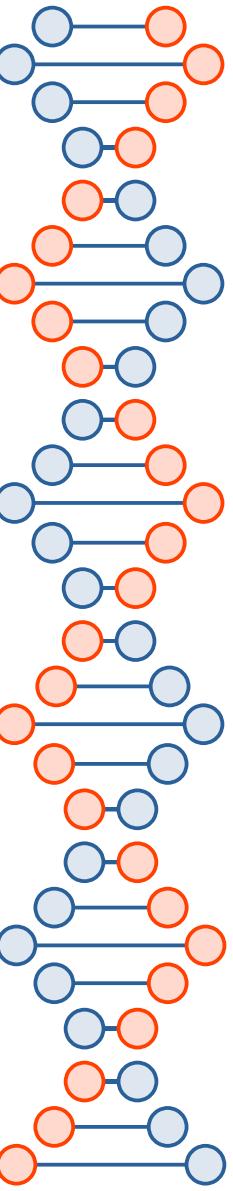
Feature Detection Theories



Any departure from the standard (tilted or vertical line) is coded as having an extra feature that the standard lacks

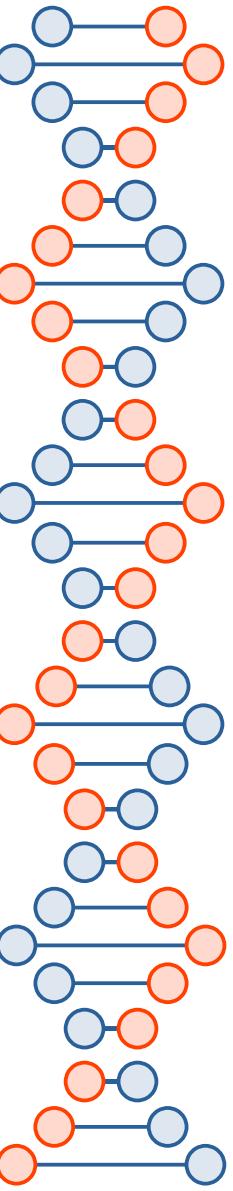
The extra feature is readily detected against background of tilted/vertical lines





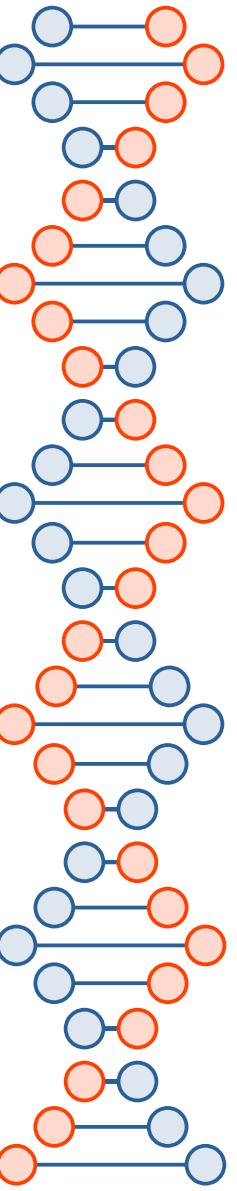
Feature Detection Theories

- Problem with feature detection models
 - Incomplete as theories of object recognition as they ignore spatial relationships among features
 - Examples
 - Describing T as a letter with a horizontal line (–) and a vertical line (|) does not distinguish between a T and a +
 - If letters are features of words, what distinguishes atom from moat, or moat from otma?



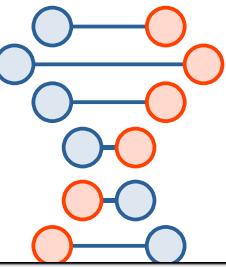
Structural Theories

- Structural description theories focus explicitly on spatial relationships
- Structural descriptions include
 - Features that describe parts of objects, and
 - Spatial relations that describe how the parts are connected
 - Example
 - The T and + example can be described along the lines of
 - T: t-junction(Line1, Line1)
 - +: cross(Line1, Line1, 90°)
 - Key advantage of structural description theories
 - Ability to recognize a large number of objects without the need to recognize a large number of features



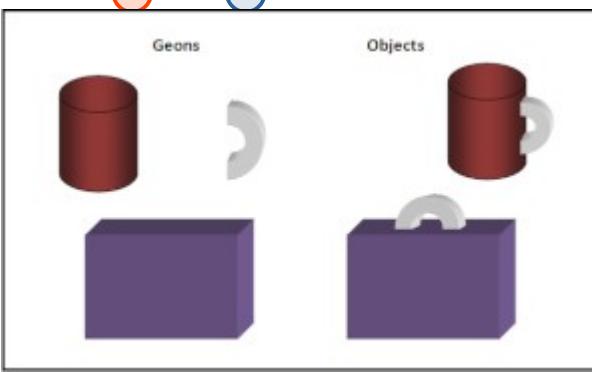
Structural Theories

- Biederman's Recognition-by-Components (RBC) theory of object recognition
- RBC theory
 - Humans are able to recognize objects by breaking them down into a basic set of simple, 3D geometric shapes called geons (geometric icons)
 - Geons
 - Simple 3-D shapes: cylinders, blocks, cones, wedges, etc.
 - Biederman suggested there are a total of 36 geons
 - Structural descriptions of complex objects are built from these components using spatial relations

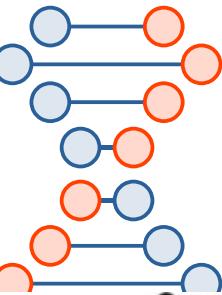


Structural Theories

- Key Concepts of the RBC Theory

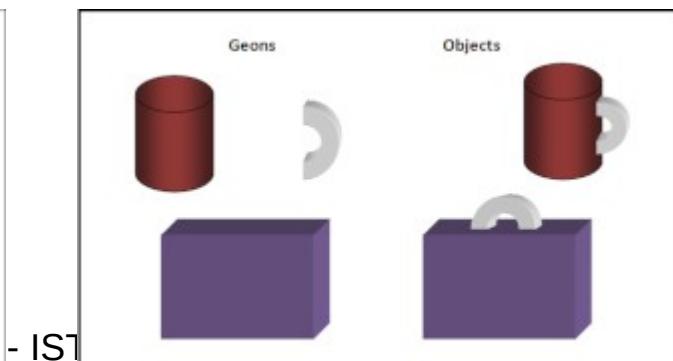
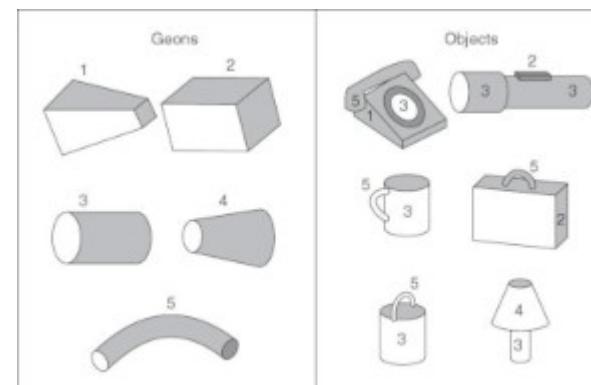
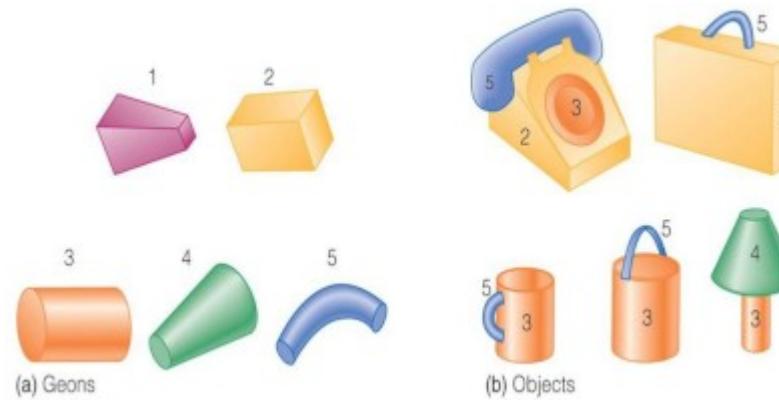
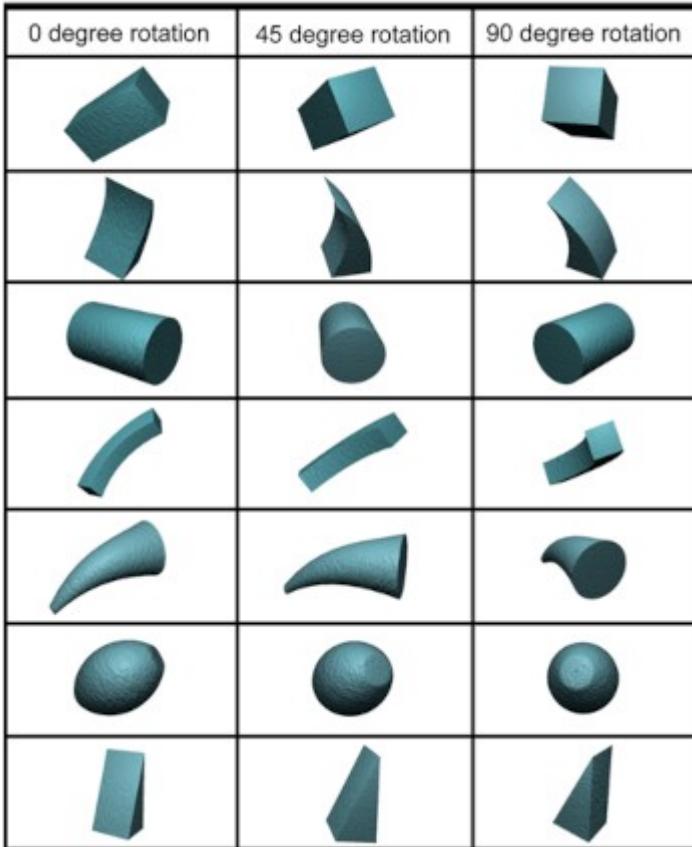


- Geons: Primitives that serve as the fundamental building blocks of all objects we perceive
- Structural Descriptions: Object recognition occurs when the visual system identifies component geons and the qualitative spatial relationships between them
 - e.g., "a cylinder and a handle" for a mug, a cube and a handle for a suitcase
- Viewpoint-Invariance
 - The properties used to identify geons (e.g., straight vs. curved edges, parallelism, and symmetry) are view-invariant
 - We can recognize an object from almost any viewpoint or orientation because these fundamental properties remain the same in the 2D image projected onto the retina
- Efficiency
 - The modest number of geons, combined with a vast number of possible spatial arrangements, allows for the representation and recognition of virtually an unlimited number of complex objects
- Analogy to Language
 - Geons vs phonemes (basic units of sound in speech)
 - Just as a small set of phonemes makes up all words in a language, a small set of geons makes up all objects

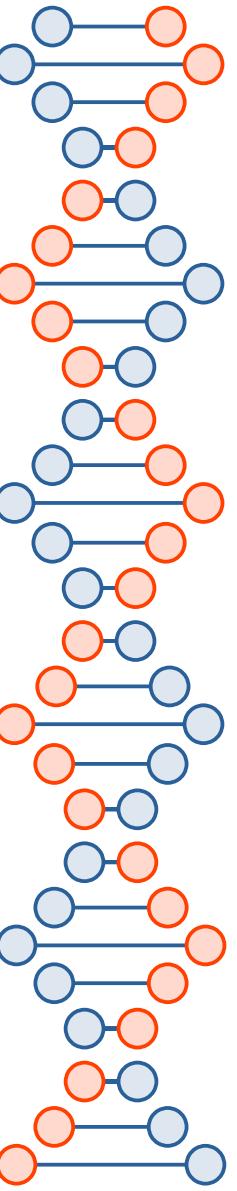


Structural Theories

- Biederman's recognition by components theory of object recognition

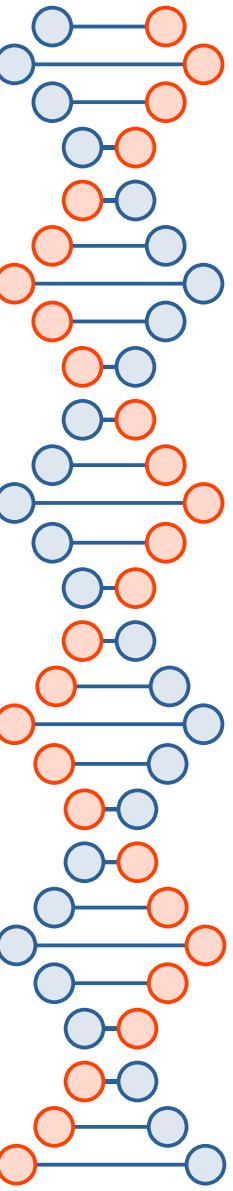


- IS1



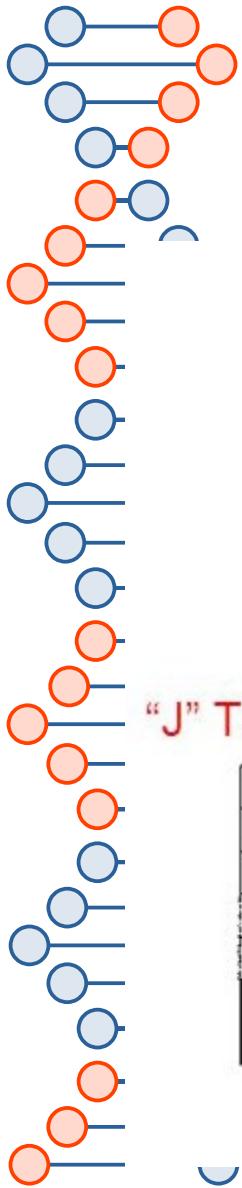
Structural Theories

- Biederman's recognition by components theory of object recognition
- Strong points
 - Explains how objects can be recognized using only a limited number of basic elements whose presence can be extracted from lower level elements of the visual image
 - Theory has been implemented in a computer model, which demonstrates that it is capable of forming the basis of an object recognition system
- Limitations
 - Decomposition of an image into geons may not be refined enough to distinguish between objects that have the same parts (e.g., wolf vs cat)
 - Some objects (e.g., loaf of bread) are difficult to decompose into parts that characterize the object

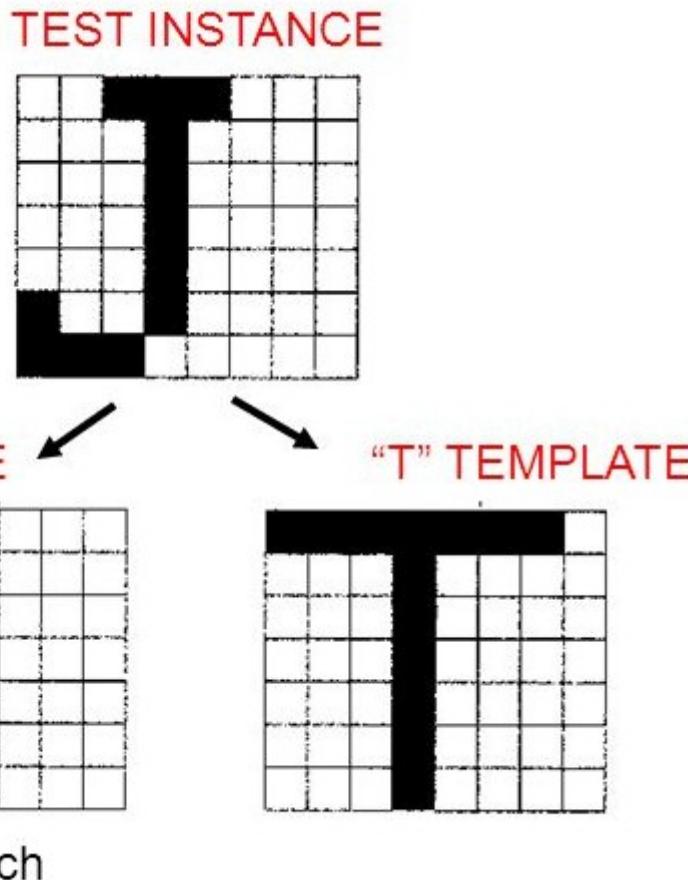


Template Matching and Alignment

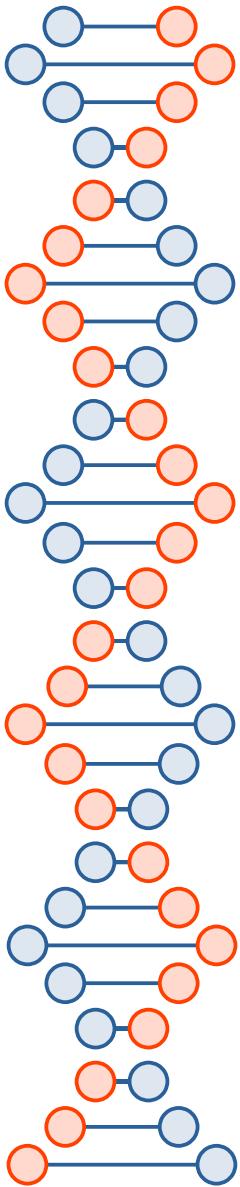
- Feature theory and structural description theory
 - Assume that people's representations of visual objects consist of lists of elements that describe the image
- Template matching approach
 - Assumes that people's representations of visual objects are 2-D arrays of picture elements (i.e., pixels)
 - A template is a copy of the image
 - Matching of a template involves finding corresponding elements between the current visual image and the template and determining how well they fit



Template Matching and Alignment

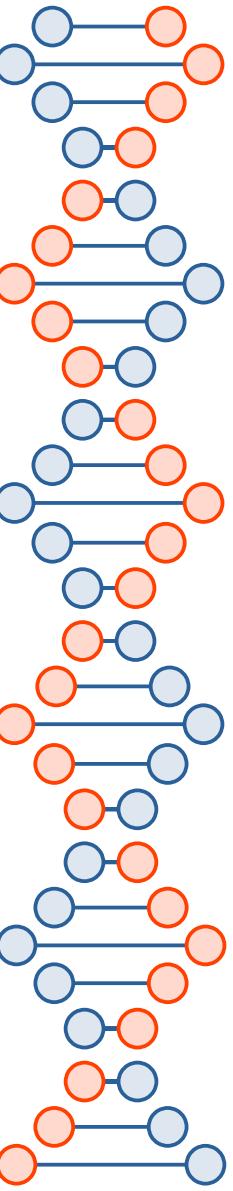


- If decisions are made based on number of matching cells, then the test instance would be classified as a J



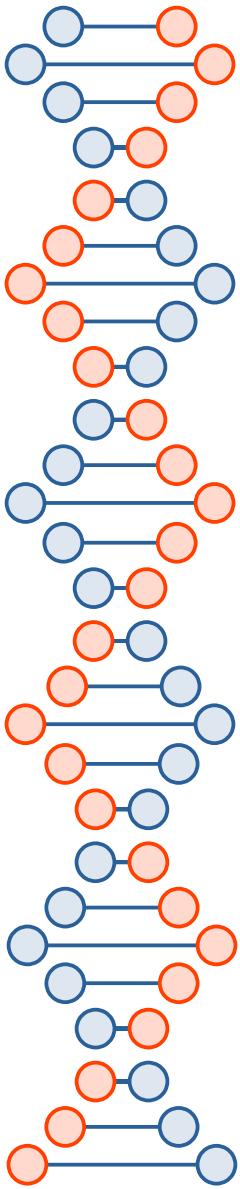
Template Matching and Alignment

- Approach seems too simplistic
 - Example: changing the size of the test instance or rotating it will cause it not to match
 - But humans don't face this trouble
- Proponents of the approach suggest as follows
 - 2 stages in object recognition
 - Use a set of transformations to align the object with the models or templated stored in memory
 - Search for the best match



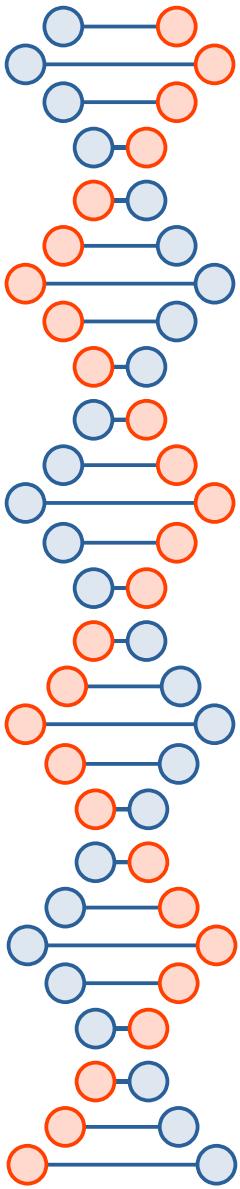
Outline

- The Problem of Perception
- Visual Perception
 - Low-Level Vision
 - Localization
- High-Level Vision
 - Feature Detection Theories
 - Structural Theories
 - Template Matching and Alignment
- **Levels and the Integration of Information in Perceptual Context Effects**



Levels and the Integration of Information in Perceptual Context Effects

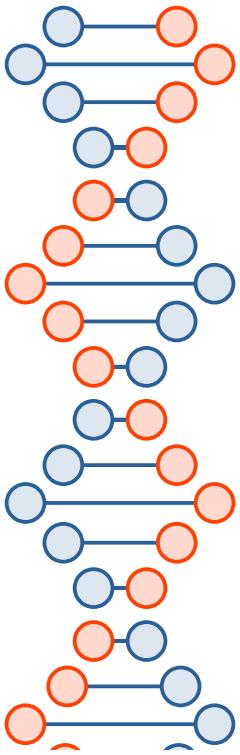
- Perception of information depends not only on lower level input, but also on surrounding context
- Bottom-up processing
 - Processing that begins with the sensory receptors (low level information) and works up to the brain's higher-level processing of the information
 - Relies on raw sensory data
- Top-down processing
 - Processing that uses prior knowledge, expectations, and context to interpret sensory information
 - Context influences perception



Levels and the Integration of Information in Perceptual Context Effects

- Because of context effect
 - Top picture
 - Middle image perceived as 13 reading top-to-bottom, and as B reading left to right
 - Middle picture
 - The second character in the first word is perceived as H, and in the second word as A, even though the characters are the same
 - Bottom picture
 - We can still perceive the letter E, even though much of it is blotted out





Top-down Processing Example: The Word Superiority Effect

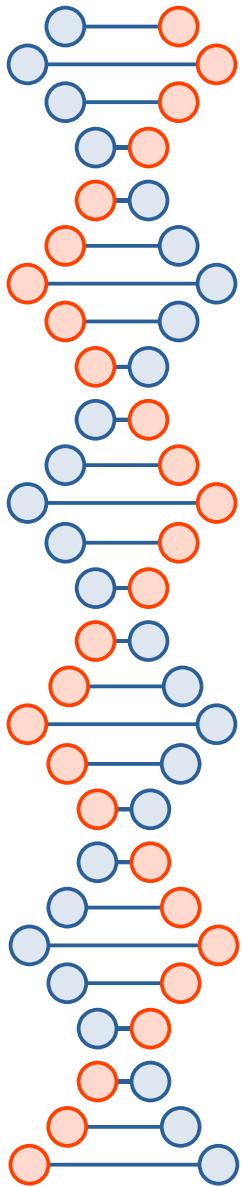
- Word Superiority Effect
 - The phenomenon that people have better recognition of letters presented within words as compared to isolated letters and to letters presented within non-word strings
 - Top down processing of the word, before the individual letters are processed
 - Meaningfulness of the word helps to identify its component letters
 - May also be due to a pronounceable difference where non-words are harder to pronounce than real words
 - Letters in "TREE" are identified more easily than letters in "TVXC"

CARD

CRAD

CRDA

DRCA



Top-down Processing Example: The Word Superiority Effect

- Similar effect seen in sentence structures
 - Background knowledge on what is being read is used to guide the reader and this is likely to allow us to predict the next likely word in the sentence
 - If text is unfamiliar, then it is harder to predict and therefore it will be read more slowly
 - The following sentence will likely be read slowly as the words are random and it is therefore impossible to predict the next word in the sentence:
 - **ORANGES DISHWASHER COMPUTER SUICIDE REVISION PARANOID CHICKEN BANANA**
 - Additionally, if an unexpected word appears in or at the end of the sentence, then reading is slowed whilst the sentence is processed
 - **I WOULD LIKE TO GO OUT AND DRINK A MICROWAVE**
 - As the word "microwave" is not an expected word, it slows the reading processing down and forces the reader to double-check as it is not a predicted word.