Visual Space: Depth

- Depth perception is based on multiple sources of information
- Oculomotor cues (accommodation and convergence)
- Pictorial cues (e.g. linear perspective, occlusion)
- Movement cues (e.g. motion parallax)
- Binocular disparity (stereo vision)

Oculomotor Cues

The occulomotor cues are accommodation and convergence.

Accommodation - As an object gets near the observer, the lens must change shape (bulge) to keep it in focus on the retina. If you can sense this change in shape, you could use it as a depth cue.

Convergence - As an object gets closer, the two eyes must converge (point inward) to maintain the image of the object on the fovea. If you can sense this, you can use it as a depth cue.

These two cues only work for objects close to the observer (inside of ten feet). Observers can use these cues, but the other three sources of depth information predominate.

Pictorial cues

A photograph is a two dimensional representation. The depth cues in a photograph are the pictorial depth cues.

This photo contains many of the pictorial depth cues.

The Pictorial Cues

1) Occlusion
2) Relative Height
3) Relative Size
4) Familiar Size
5) Atmospheric Perspective
6) Linear Perspective
7) Texture Gradient
8) Shadows

In experiments designed to examine the ability of humans to use these cues, we vary one cue and hold all other sources of depth information constant. The observer must view the scene with one eye to eliminate binocular vision. Because these cues work with one eye, they are also known as monocular depth cues.

Occlusion & Relative Height

When an object hides or partly obscures your view of another object, then the hidden object is farther away.

Occlusion is a relative depth cue.

Items higher in the visual field, but below the horizon are generally further away. Objects above the horizon are further away when they are lower in the visual field. Put another way, the closer an object is to the horizon, the further away it is.

Relative height is a relative depth cue.

Relative and Familiar Size

If two objects are of equal size, then if one is further away, it will occupy a smaller area on the retina. All other things being equal, a larger retinal image makes something appear closer.

We know the actual sizes of some objects. People are generally 5 to 6 feet tall, trees are taller than that. A quarter is larger than a dime.

If the images on the retina of a quarter and a dime are the same size, this implies that the dime is closer.

Both relative and familiar size are relative depth cues.


LIBERTY

Atmospheric Perspective

Objects at a distance appear less sharp and slightly more bluish. This is atmospheric perspective.

When an object that is far away reflects light to an observer, the light must travel through a lot of air (filled with dust, water vapor, etc.). Some of the light is diffracted, so that distant objects look less sharp. This is more likely to happen to short wavelength light (blue) than long wavelength light (red), giving the distant objects a bluish cast.

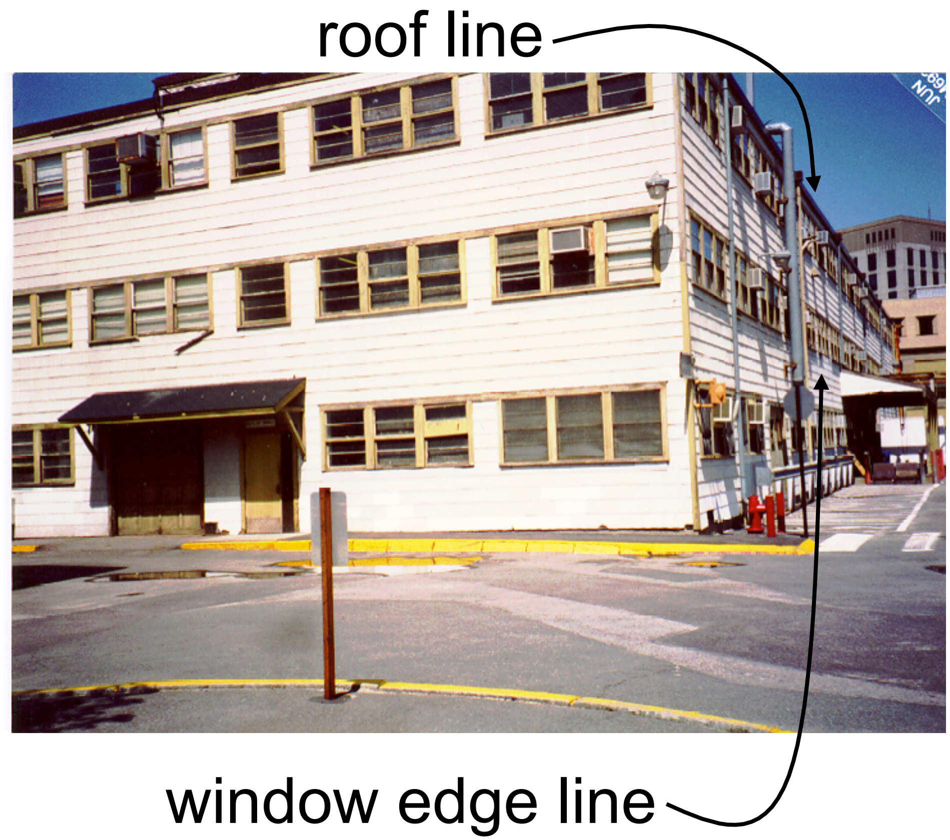
Atmospheric perspective is a relative depth cue. It also differs from location to location (depending on particles and water vapor in the air).

Linear Perspective

Lines that are parallel in the world appear to converge.

The greater the distance, the greater the convergence.

Linear perspective is a relative depth cue.



In this photo, the roof and window edge lines converge.

Texture Gradient

Items that are equally spaced in the real world will appear to be closer together if they are further away.

In the building picture, the windows on the side of the building are closer together (in the visual image on the retina) when they are further away.

The texture gradient is a relative depth cue.

Shadows

The position of shadows, relative to the object casting the shadow, can disambiguate the relative positions of the objects. See example in Figure 10.7 in text.

Motion Base Depth Cues

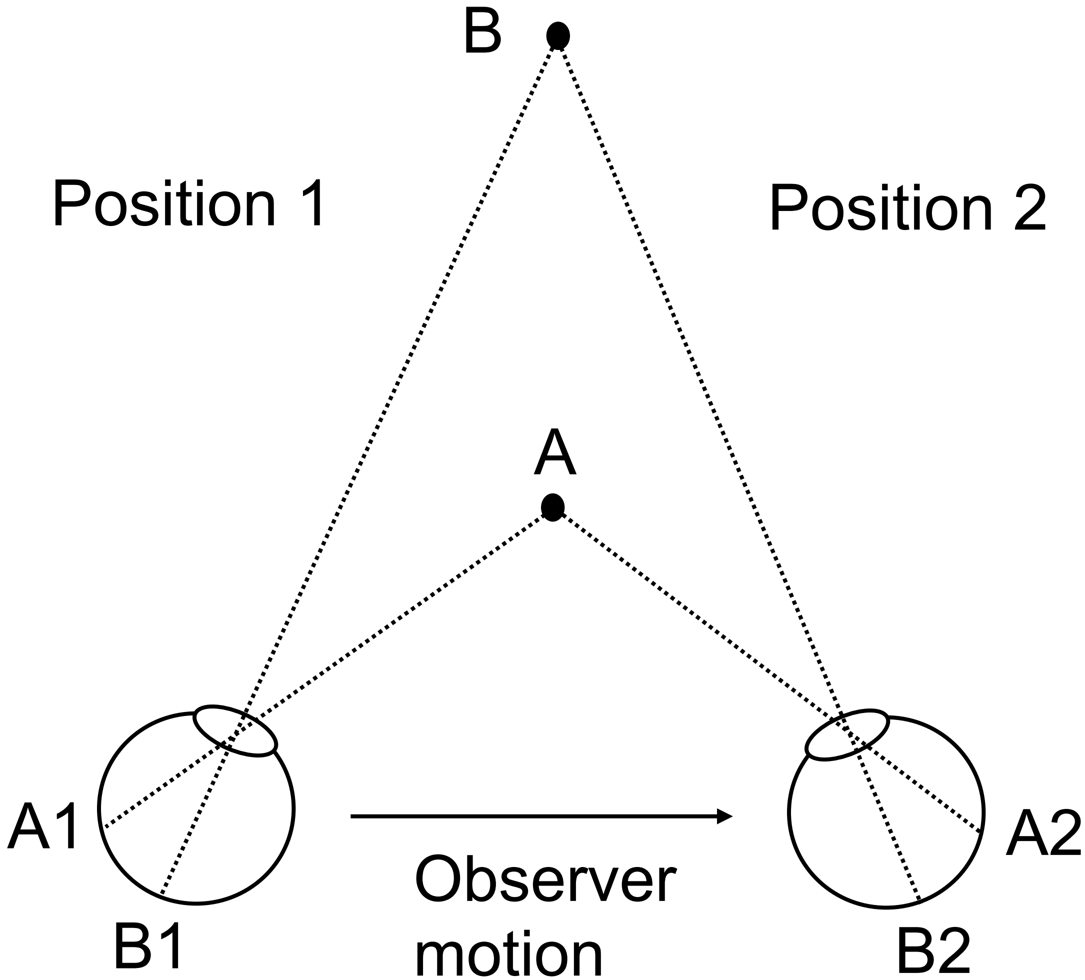
There are 2 motion based depth cues. These operate when you are moving. These are monocular depth cues.

1) Motion parallax - as we walk or move, nearby objects appear to move rapidly past us. Far objects appear to be stationary or move more slowly.

2) Accretion and deletion - when the observer moves, the surfaces of objects change in their relative positions. A nearer surface will obscure a farther object, but as the observer moves, the part of the far object that is obscured will change and previously obscured areas will come into view.

The motion based depth cues are relative depth cues.

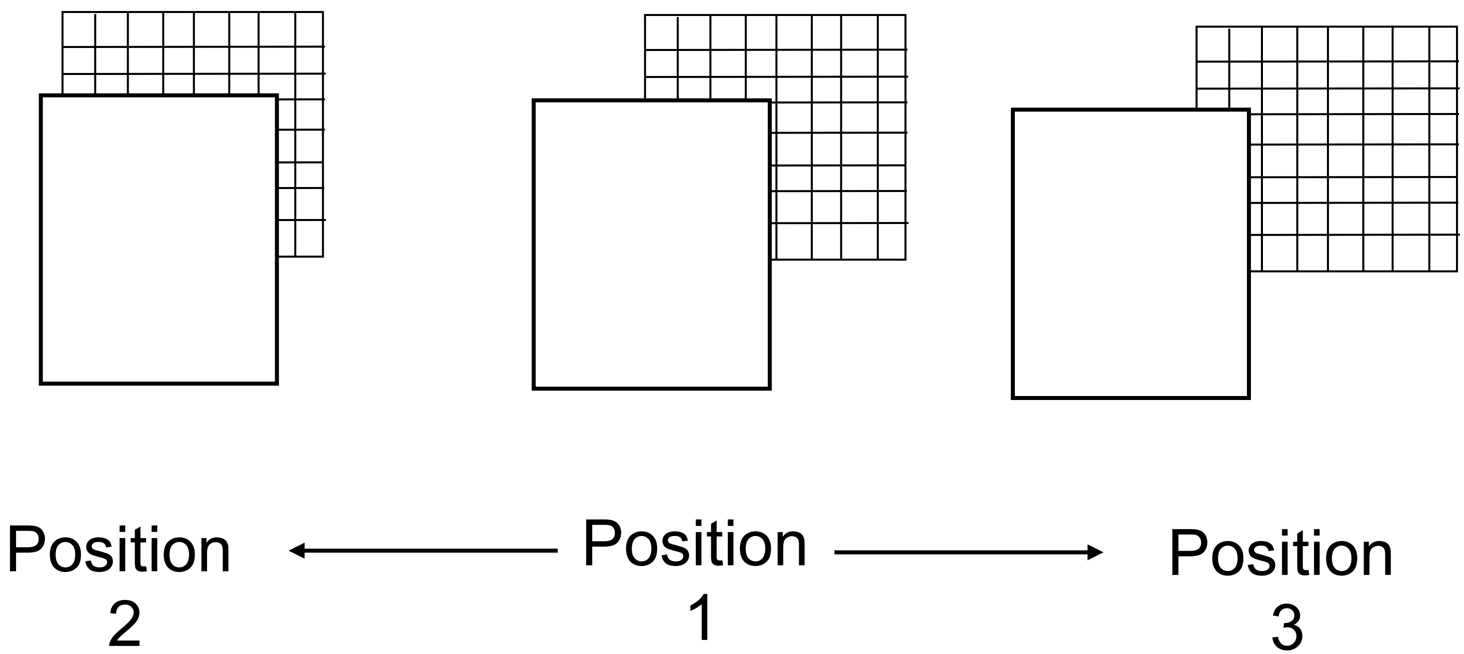
Motion parallax



As the observer moves from position 1 to 2, the image of A (closer) moves further on the retina than the image of B (farther).

Accretion and Deletion

For the objects in Position 1, observer movement left (1 to 2) produces deletion. Observer movement right (1 to 3) produces accretion.



Monocular Depth Cues Summary

	Close	Intermed	Far
Occlusion	X	X	X
Relative size	X	X	X
Accommodation & convergence	X		
Motion	X	X	
Relative Height		X	X
Atmospheric Perspective			X

Binocular Disparity

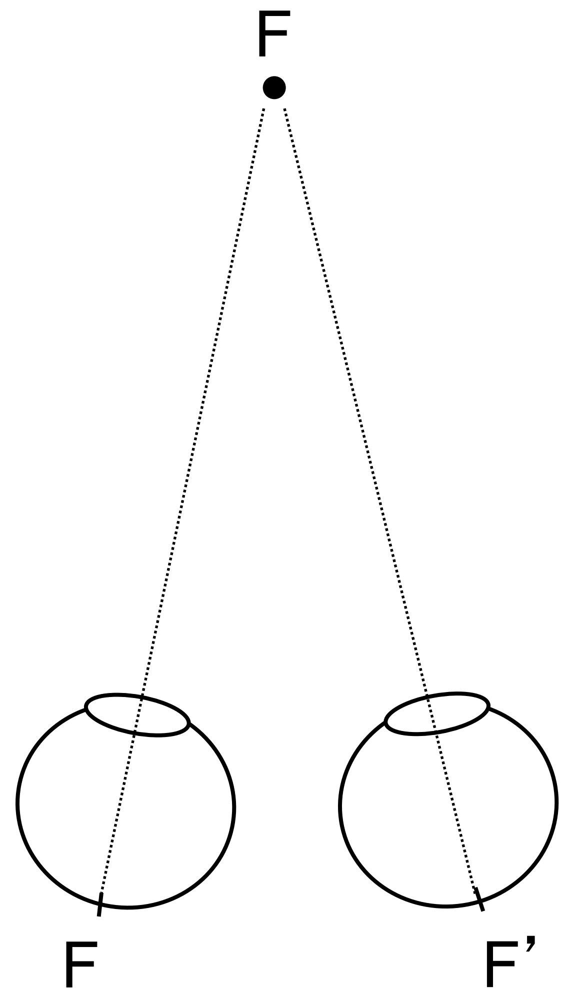
Humans have two eyes. Because they are a few inches apart, the retinal image of an object on one eye may be slightly different than the retinal image of the same object on the other eye. This is the depth cue known as binocular (retinal) disparity. The brain compares these two images as part of depth perception.

The impression of depth created by binocular disparity is called stereopsis. It can be used to create an illusion of depth in a picture using special viewing devices or glasses.

Corresponding Retinal Points

When you look at an object, its image falls on the fovea on each retina. The foveas are corresponding points on the two retinas.

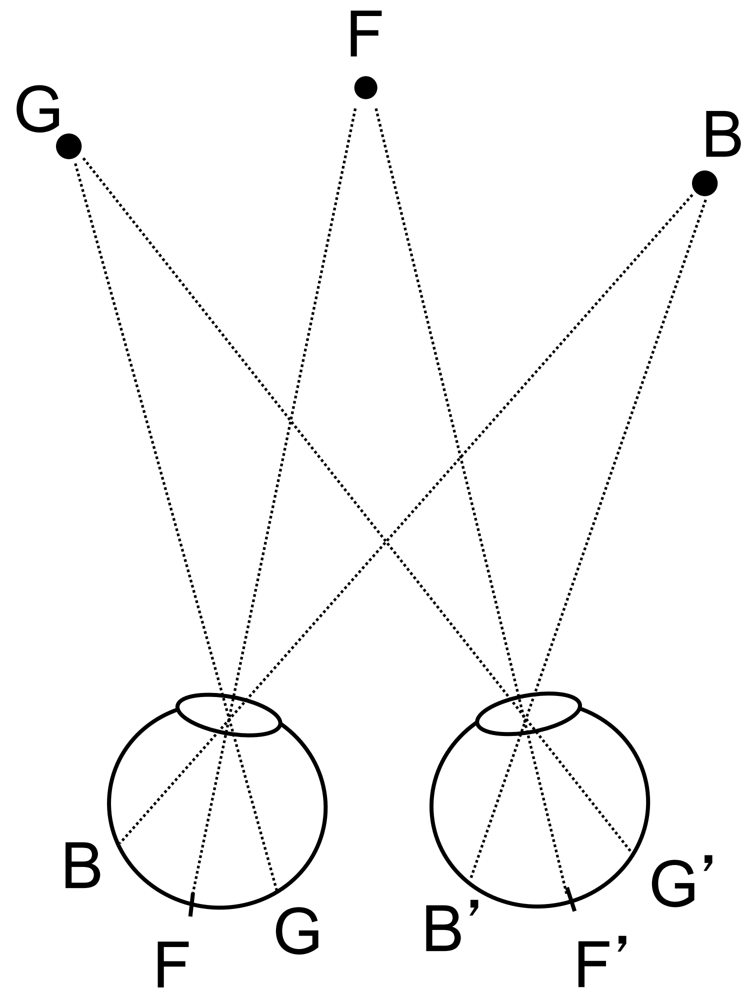
Corresponding points on the two retinas are points that would align if we were to superimpose the two retinas.



Corresponding Points - 2

In addition to the image on the fovea on each retina, any other objects at the same distance from the observer as the object being fixated will also have their retinal images fall on corresponding points.

Here, the images of objects B, F, and G all fall on corresponding retinal points.



Corresponding points - 3

All of the points in space at the same distance from the observer as their point of fixation fall onto corresponding points on the two retinas. An imaginary arc that passes through all of these points is called the horoptor.

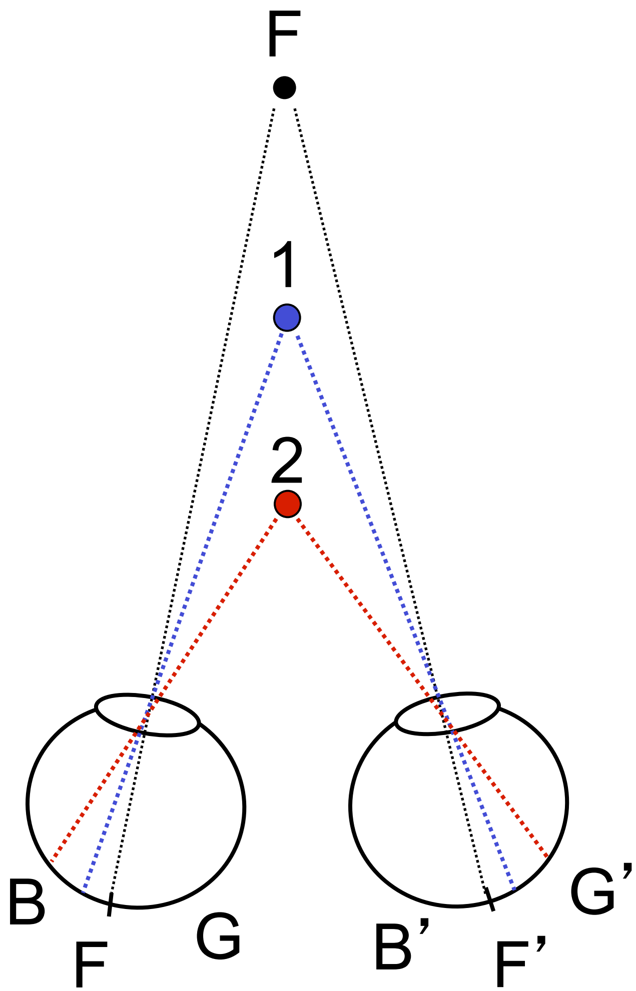
Think of this arc as part of a circle whose center in the observer. All points on this arc are at an equal distance from the observer.

Objects that are not on the horoptor project onto noncorresponding (disparate) retinal points. It is this disparity that is the binocular depth cue.

Retinal Disparity

Here, the observer is focused on object F. The images of objects 1 and 2 do not fall on corresponding retinal points. Points B and B' and G and G' are corresponding retinal points. The image of object 2 actually falls on points B and G'. The distance between where the image falls and corresponding points on the two retinas is retinal disparity.

It is a relative depth cue.



Retinal Disparity - 2

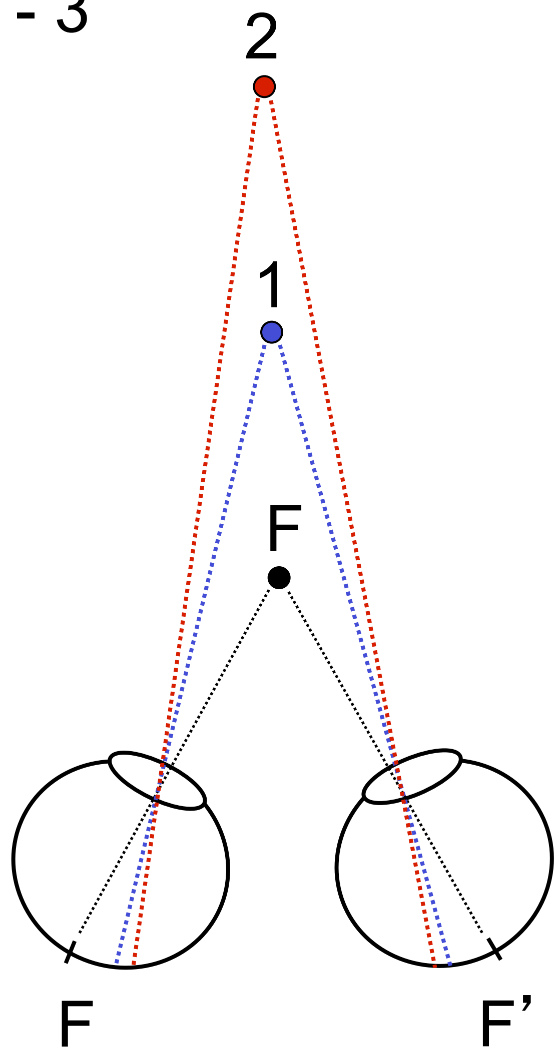
When an object is <u>closer</u> to the observer than the horoptor, its image on the retina is displaced outward, relative to corresponding points. The farther the object is from the horoptor, the greater the disparity. For objects closer than the horoptor, this is termed crossed disparity.

If the brain can compute the disparity, it can be used as a relative depth cue. The greater the crossed disparity, the greater the distance from the horoptor - and closer to the observer - the object.

Retinal Disparity - 3

If an object is farther from the observer than the horoptor, its image on the retina is displaced inward (toward the nose). This is the uncrossed retinal disparity cue.

The greater the distance from the horoptor - and the observer - the greater the uncrossed disparity.



Stereo Vision

Binocular disparity is the cue to stereo vision. If the brain can compute binocular disparity for corresponding parts of the retinal image, it can be used to judge relative depth.

Stereoscopes use pictures taken by a special camera which has two lenses and takes two pictures. When viewed so that the left picture is presented to the left eye and the right picture to the right eye, it creates the illusion of a 3-dimensional picture. The brain is using the disparity in the two pictures.

However, the pictures also have pictorial (monocular) depth cues. How would we determine if the brain can use binocular disparity without these other cues?

Stereo Vision - 2

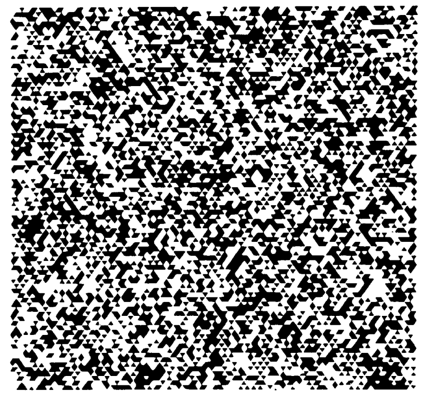
What we need is a picture with no monocular objects or depth cues so that if the “object” is visible, it is because of the observer’s ability to use binocular disparity.

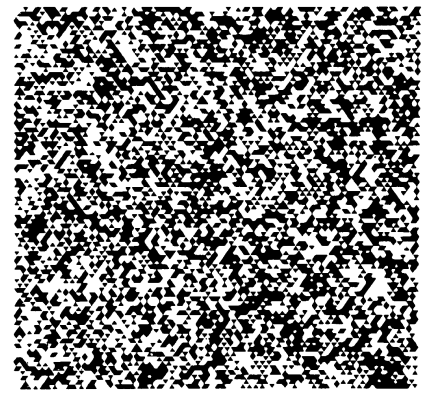
Julez (1971) created random dot stereograms. A grid is filled with black and white pixels, arranged randomly. This is presented to the left eye. The same grid has a region of pixels (say a square) displaced to the right by a small amount and the area left behind is filled randomly. This grid is presented to the right eye.

If the brain can establish the retinal disparity of the pixels in the two images, the observer should “see” a square in front of the background.

Random Dot Stereogram

A square has been filled randomly with black and white pixels. The left and right squares are identical except that some of the pixels in the right square have been displaced to their right (relative to their position in the left square).

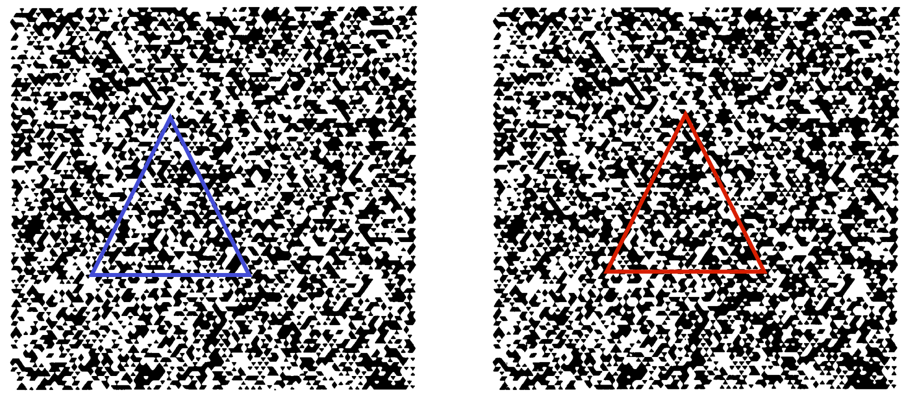




When viewed with a stereo viewer, an object should appear in front of the background.

Random Dot Stereogram - 2

Here, one of Julez' random dot stereograms is shown along with the region of pixels that has been moved. Since the disparity cue is an outward displacement, observers report a triangle in front of the background.



In the random dot grid on the left, the triangular region of pixels, one the left, is shifted to the right. The space left behind is filled. This results in the random dot grid on the right.

Random Dot Stereograms - 3

Since an observer can see an object based on only binocular disparity -

1) Binocular disparity can be used separately from all other cues to depth.

2) Binocular disparity is extracted before objects are created in perception. That is, binocular disparity is extracted then used to group together basic visual features to form objects (based on common disparity).

Depth Cues Summary - 2

1) Different cues work at different depths

2) For any particular depth, multiple cues are available in normal viewing conditions

3) For any particular depth, no one cue is essential

4) Under certain viewing conditions, some cues may not be available and depth perception may be limited

Physiology of Binocular Disparity

Cells in visual cortex are binocular. All have a "preferred" disparity. That is, all will give a maximal response to their preferred stimulus at a particular retinal disparity.

A particular complex cell, which responds to a 30 degree angle bar moving left to right may respond optimally to binocular input with a disparity of 30 minutes (one-half of a degree) of visual angle. This same cell will not respond to monocular input from one eye.

Cells in V1, V2, the dorsal pathway and the ventral pathway show sensitivity to binocular disparity.

Selective Rearing Studies

A second approach to examining the physiology of vision has been to rear animals (cats or monkeys) in controlled visual environments. Then, both behavioral testing and single cell recordings are done.

In a cat raised in a normal environment, there are binocular disparity cells at birth and subsequently.

In a cat raised with monocular experience, at adulthood there are no binocular cells and the animal is stereo-blind.

Selective Rearing - 2

For the cat, as little as three days of monocular rearing during the the interval between 4 weeks and 16 weeks age can alter the ocular dominance and binocular response of cells. Extended monocular rearing after 16 weeks has virtually no effect.

Thus, for the cat, there seems to be a critical period during which environmental exposure can "tune" or alter the response properties of visual cortex cells.

Human Selective Rearing

In humans, we do not do selective rearing studies. However, visual disorders do occur "naturally" and their effects have been studied.

Two natural effects are amblyopia and strabismus. Amblyopia is a loss of acuity in one eye. Strabismus is a muscle imbalance between the eyes that prevents proper coordination of the two eyes. Both can lead to a "selective rearing" in that the infant/child will not be exposed to a clear, coordinated visual world on the two retinas.

If this condition is uncorrected for the first 3-4 years of life, the child will (likely) be stereo-blind.

Depth Summary

As a final note, consider nature's design of the visual systems of different animals.

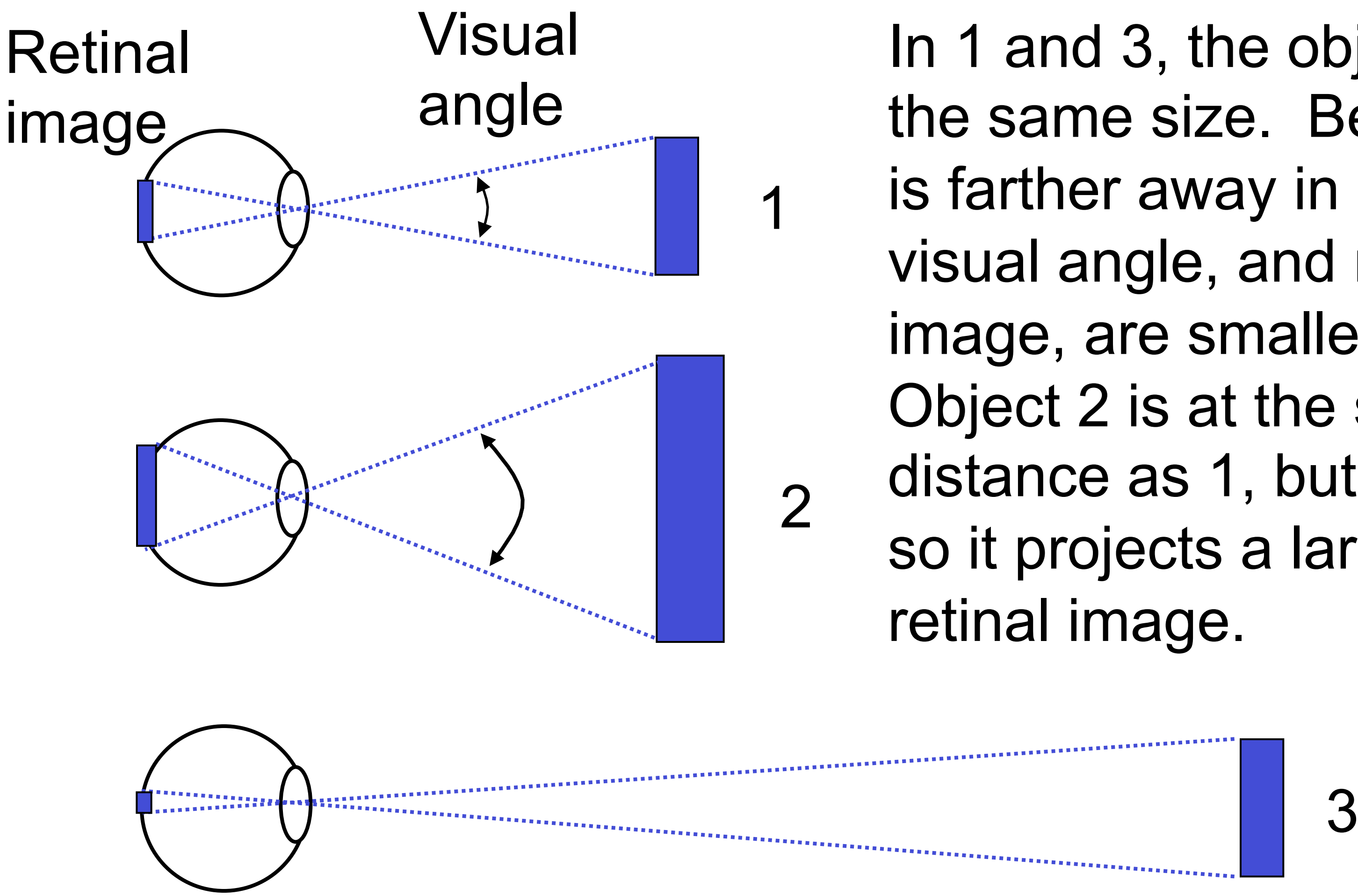
1) To use binocular disparity, the animal needs two eyes with overlapping fields of view. Animals with frontal eyes (cats, humans) have good stereo vision.

2) To get a wide field of view, place the two eyes on opposite sides of the head. This can give a 360 degree field of view, but with virtually no overlap between the two eyes. Animals with lateral eyes (rabbits, pigeons) have little (pigeon) or no (rabbit) stereo vision. Their depth perception relies on the other cues.

Size and Illusions

- Size perception involves the retinal image (retinal size) and perceived distance.
- Many illusions of size involve the misperception of depth
- These illusions do not occur often in the natural environment because of the multiple perceptual cues to depth.

Retinal Size and Distance



In 1 and 3, the object is the same size. Because it is farther away in 3, the visual angle, and retinal image, are smaller. Object 2 is at the same distance as 1, but is larger so it projects a larger retinal image.

Size Constancy

Size constancy occurs when we correctly perceive an object's physical size in spite of changes in the size of the image on the retina.

Size constancy could be the result of "knowing" the object and of the sizes of different objects on the retina. That is, it is an effect of familiar size and relative size.

Size constancy could be the result of combining retinal size and distance. This is called size-distance scaling.

How would we investigate these?

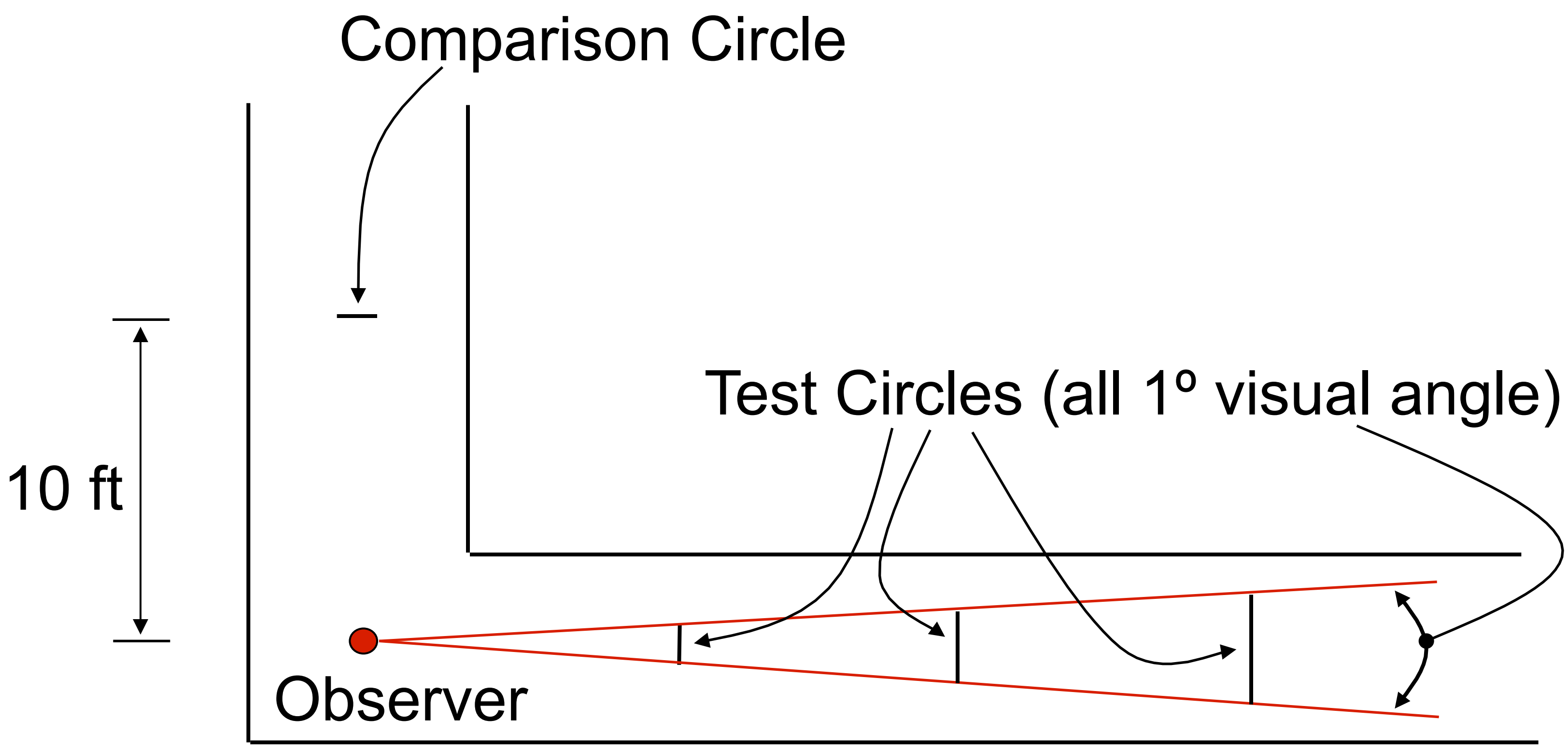
Size Constancy and Depth

In the early 1940s, Holway and Boring set up an experiment to reveal what information observers used in size perception.

They gave observers control over the size of a “comparison” circle 10 feet from the observer. The observer adjusted the size of the comparison circle to match it to that of a test circle. The distance of the test circle from the observer was varied (10 feet to 120 feet).

The visual angle of the test circle was held constant at 1 degree.

Experimental Set-up

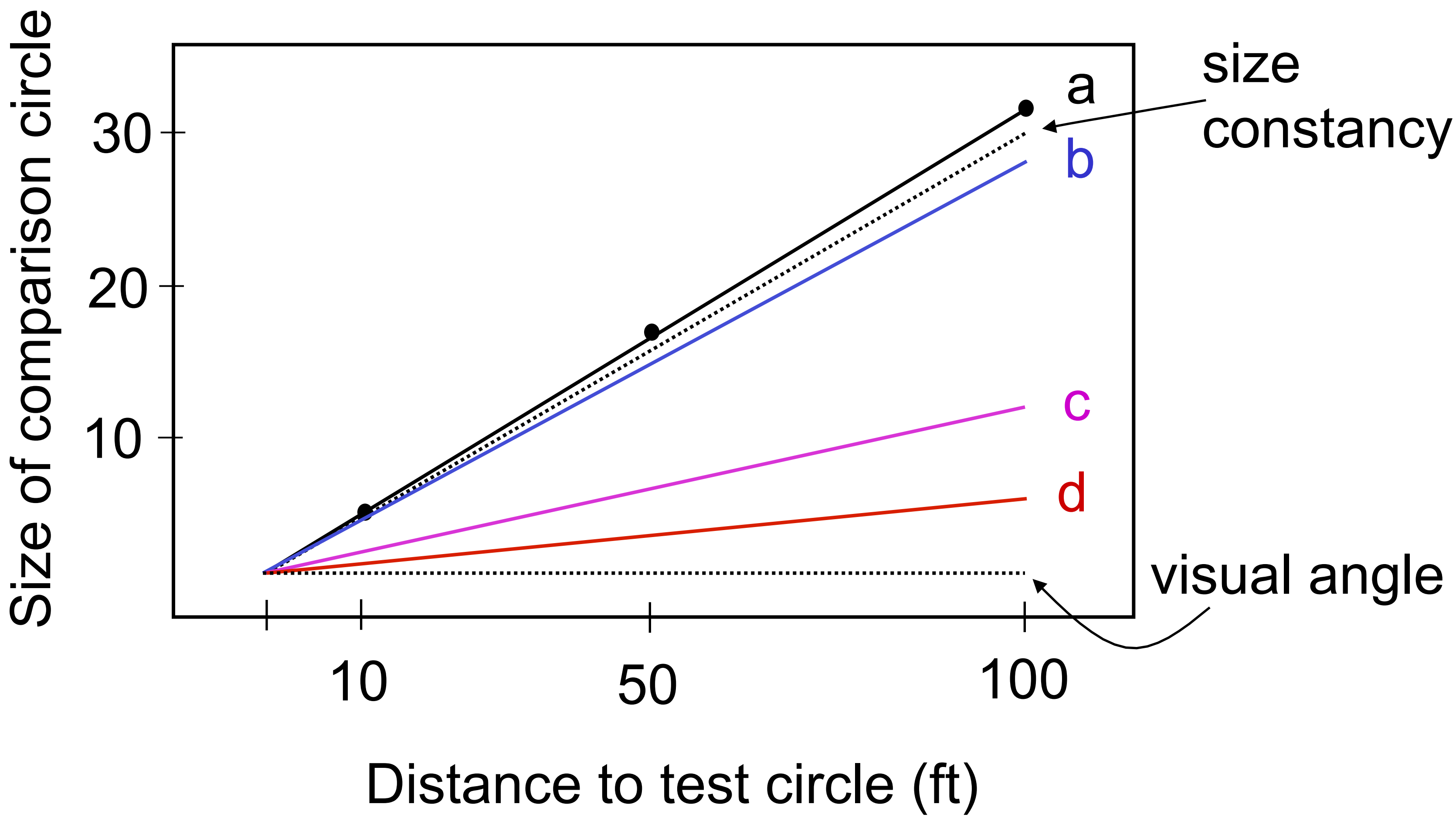


Experiment - 2

A key element to this experiment is that all of the test circles are sized so that their retinal images are the same size. All circles have the same visual angle.

When the observer was allowed to use both eyes under normal lighting, they matched the comparison circle to the physical size of the test circles very accurately. This is size constancy. The line labeled a in the accompanying graph shows these results.

Experiment Results



Experiment Results - 2

When the observer used only one eye to observe the test circle, the results shown by line b were found. The estimates of test circle size decreased slightly. This indicates that when binocular disparity was removed, a bit less size constancy was obtained.

When the observer had to look at the test circles monocularly through a peep hole, line c represents the results. Even less size constancy.

When the hallway was draped in black crepe to eliminate pictorial depth cues, the results from line d were found. Here, size constancy is almost gone since the data are similar to retinal size (visual angle).

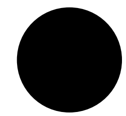
Experiment - Conclusions

Size constancy is the result of size-distance scaling. When depth (distance) cues are eliminated, size constancy breaks down.

Size-Distance Scaling is the idea that perceivers combine retinal size and distance to perceive physical size: $S = k(R \times D)$. Perceived size S is the combination of retinal size (R) and perceived distance (D).

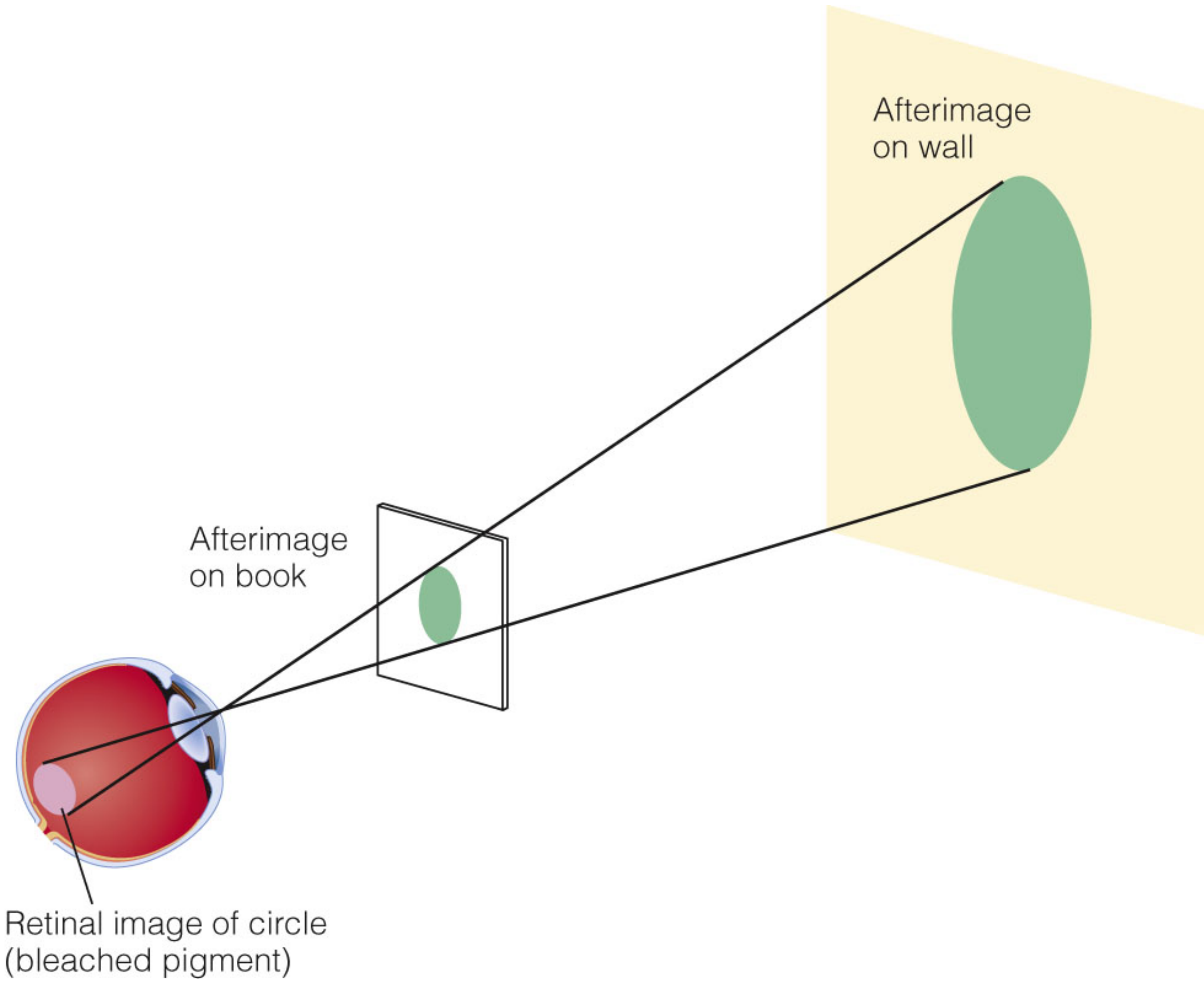
A demo of size-distance scaling

Look at the small dark circle, at a comfortable reading distance, for 60 seconds. Then, look at the blank white area below it. A small, close afterimage will appear.

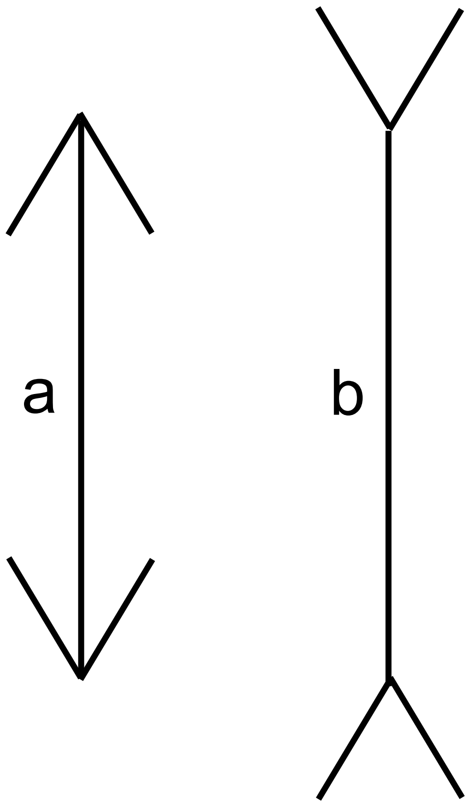


Repeat this, but after staring at the circle, look at a wall 10 feet or so away. A large, distant afterimage will appear.

How does size-distance scaling explain this illusion?


Afterimage
on wall
Afterimage
on book
Retinal image of circle
(bleached pigment)

Visual Illusions



Which line looks longer - a, b, or neither?

This is the Müller-Lyer illusion. Even though the line segments in the center of a and b are the same length, b looks longer.

An Explanation

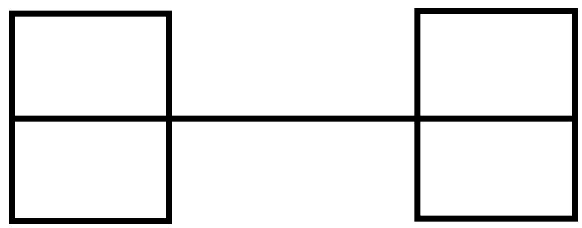
One possible explanation for the Müller-Lyer illusion is that the visual system miss-applies depth cues. Size-distance scaling then results in one line segment looking longer.

Using the previous figure, a is an edge projecting toward you (near) while b is an edge projecting away from you (far). Both lines project the same visual angle. To do this, the far edge must be larger. This explanation proposes that the Müller-Lyer is a result of a depth illusion influencing size-constancy.

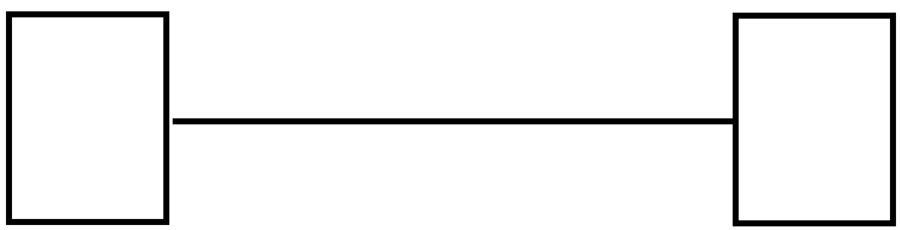
Müller-Lyer Part 2

How would we test this depth illusion explanation?

Try variations of the Muller-Lyer that do not have a depth illusion. This example has no obvious depth. Does a look longer or does b look longer?



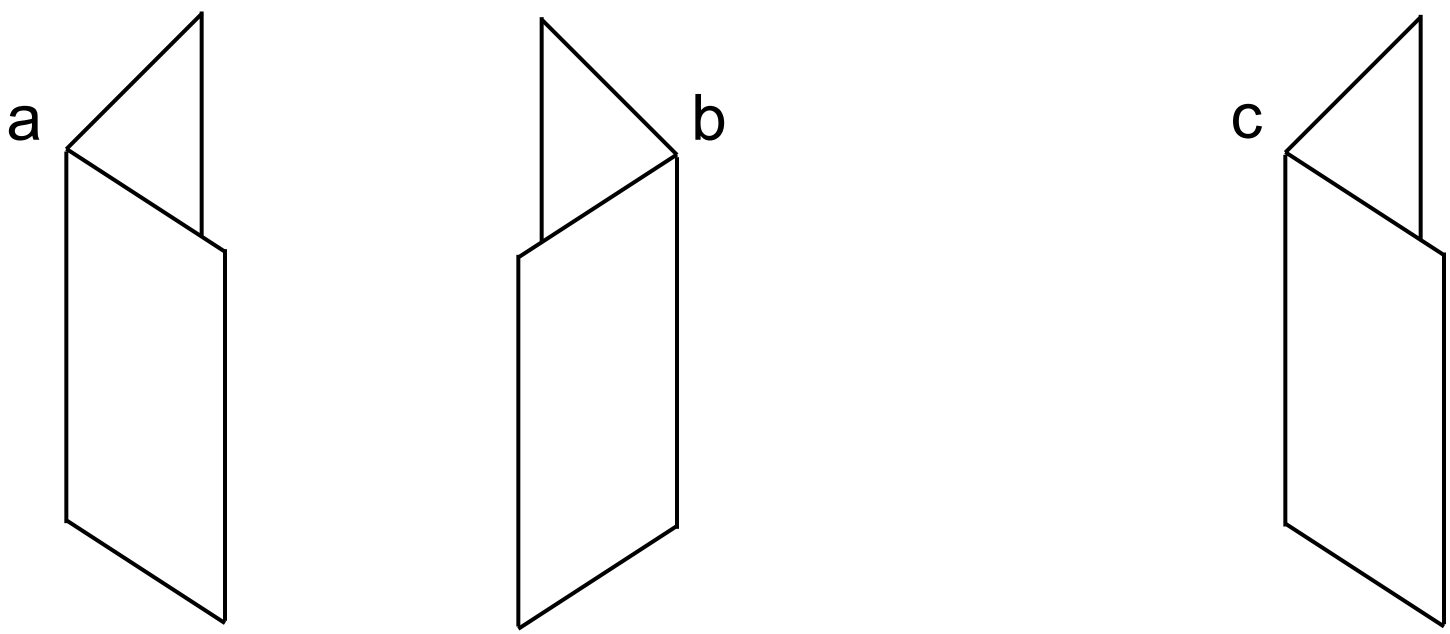
a



b

Müller-Lyer Part 3

In this variation, we manipulate depth to make points a, b, and c all at the same depth. Which distance is longer: a-b or b-c?



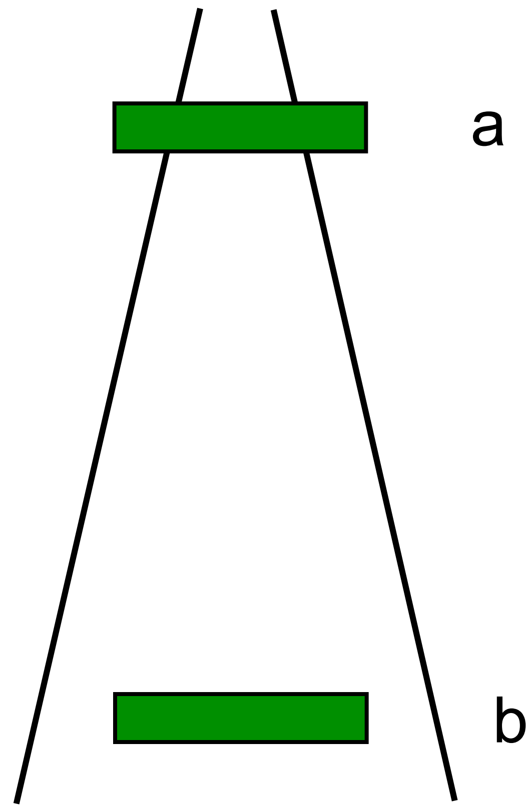
A Second Explanation

An alternative to depth is that these variations of the Müller-Lyer are all *length* based illusions. If the perception of line length in the Müller-Lyer is based on both actual line length and overall figure length, then the longer figures should appear to have longer lines.

This explanation works, but is based on a concept that has no separate verification - that length perception is made up of two components.

Additional Size Illusions

Does rectangle a look larger or does rectangle b look larger? Even in this very simple version, a looks larger. As pictorial depth cues are added, the size of the illusion grows. This is the Ponzo illusion.



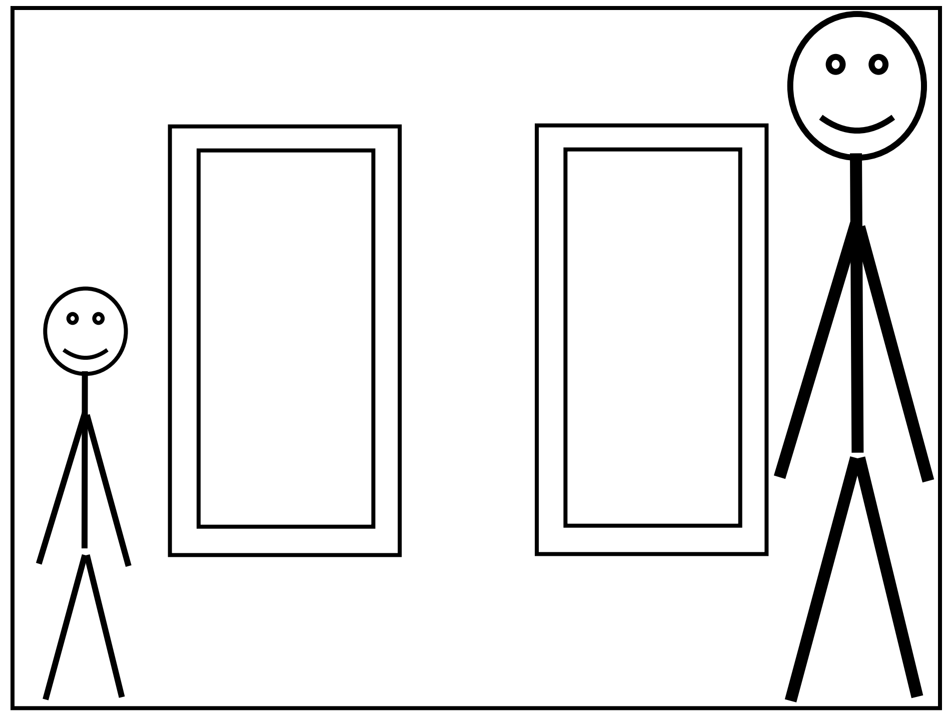
Ponzo - 2

Which of the two rectangles looks larger, the one on the left or the one on the right?

Size Illusions

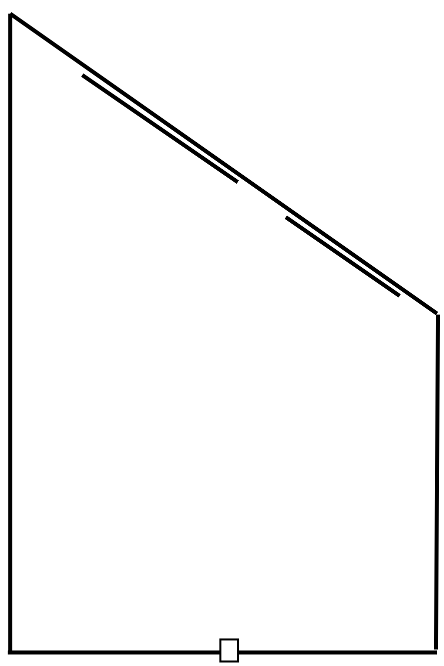
In addition to the Ponzo illusion, another size illusion based on depth miss-perception is the Ames Room.

The observer, looking through a peep-hole with one eye, sees the “image” at the right. The individual on the right looks much larger than the individual on the left.

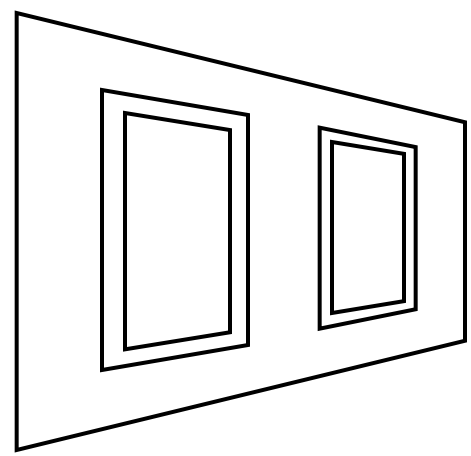


Ames Room

Actual shape of Ames room, viewed from above (left) and the shape of end wall (right). The room is built to precisely counter linear perspective.



Peep Hole



End Wall

Size Illusions & Depth

The Ponzo illusion and the Ames room clearly involve illusions of size that depend upon misperceiving depth.

The moon illusion, where the moon appears larger near the horizon than overhead, may also be partly based on depth.

Returning to the Müller-Lyer, are all of its variations based on the same underlying principles? Does depth play a role? How would we test this and separate it from the length perception explanation?

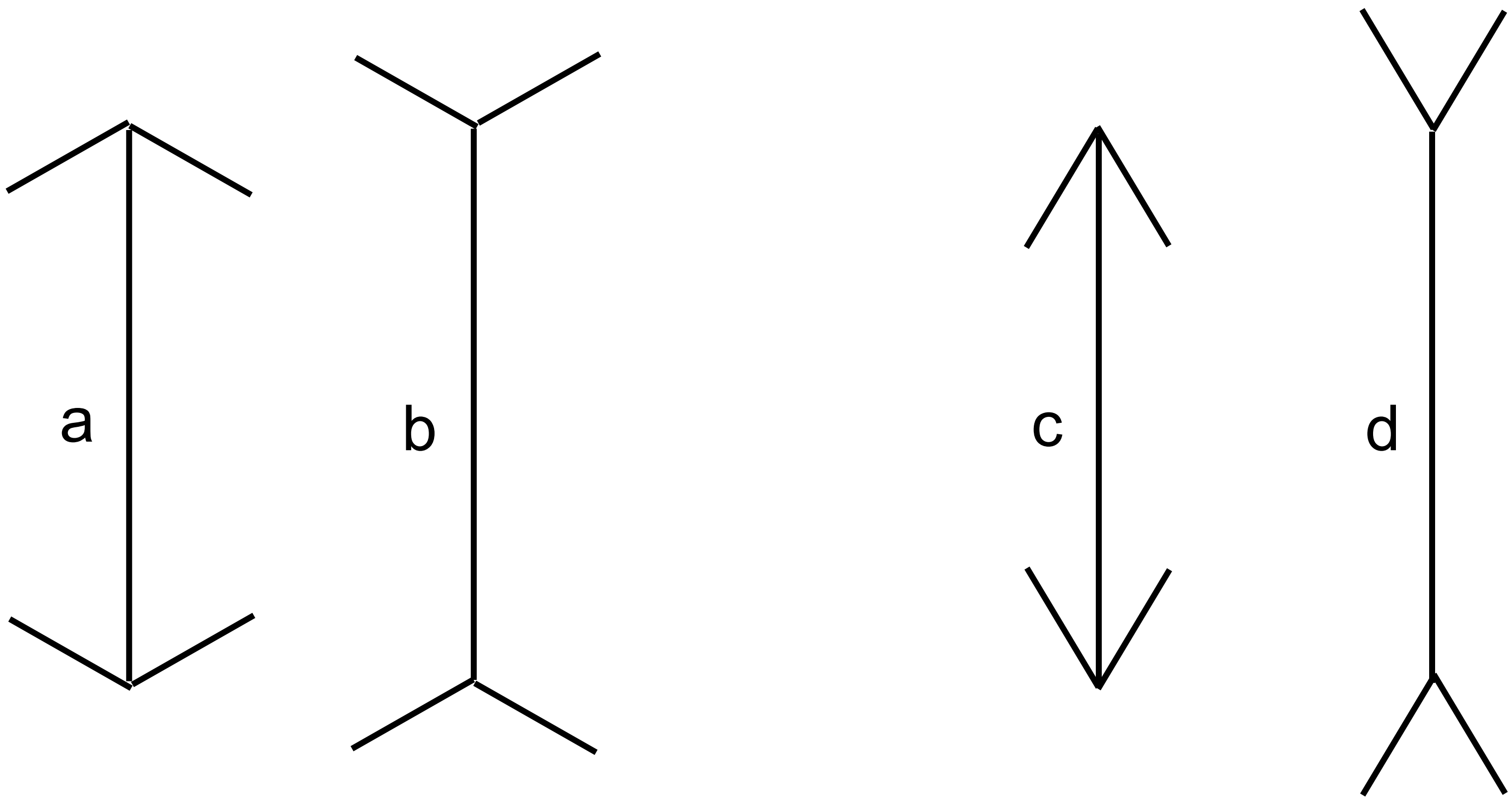
Müller-Lyer Test

Instructions:

Look at the two figures (a & b). Which appears longer? What percentage longer?

Look at the two figures (c & d). Which appears longer? What percentage longer?

Müller-Lyer Test - 2



The Ecological Perspective

In the ecological approach, as described by Gibson, the concern is for perception in the natural environment. For size and depth, these four principles are important.

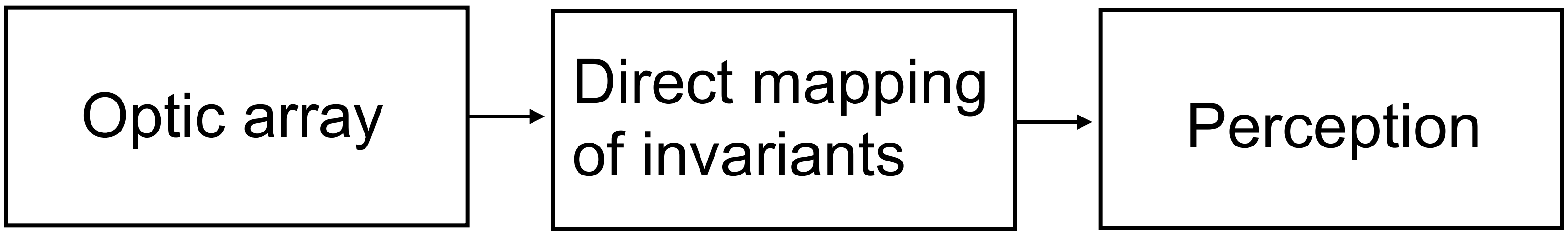
1) The stimulus is described in terms of information in the environment as it projects onto the retina - the optic array.

2) Important information for perception is created by movement of the observer.

3) Motion and the optic array lead to invariant information (cues that remain constant).

4) Invariant information leads directly to perception.

Ecological Perspective - 2



This diagram shows the sequence of steps in Gibson's proposal. It is often referred to as "direct perception" since perception is based on invariant qualities extracted directly from the optic array.

The Optic Array

The optic array is the pattern of light that falls upon the retina. Gibson emphasized the rich source of information about the external world carried by the optic array. Gibson used projective geometry to describe how light from a 3-D world is reflected onto the 2-D surface of the retina. Based on this analysis, he then emphasized that perception involves the extraction of qualities that are invariant in the optic array. These qualities stay the same even when the observer's vantage point changes.

Observer Motion

A second key element to the Ecological Perspective is that the observer is usually moving relative to the visual world. Even if standing still or sitting, we are often moving our head or our eyes. While the image on the retina changes, some properties of the optic array remain constant - invariant. These are the critical aspects of the optic array on which perception is based.

Invariant Information

1) Texture gradient. Spacing of texture (pattern) in the optic array.

2) Flow pattern. Motion of regions across the retina.

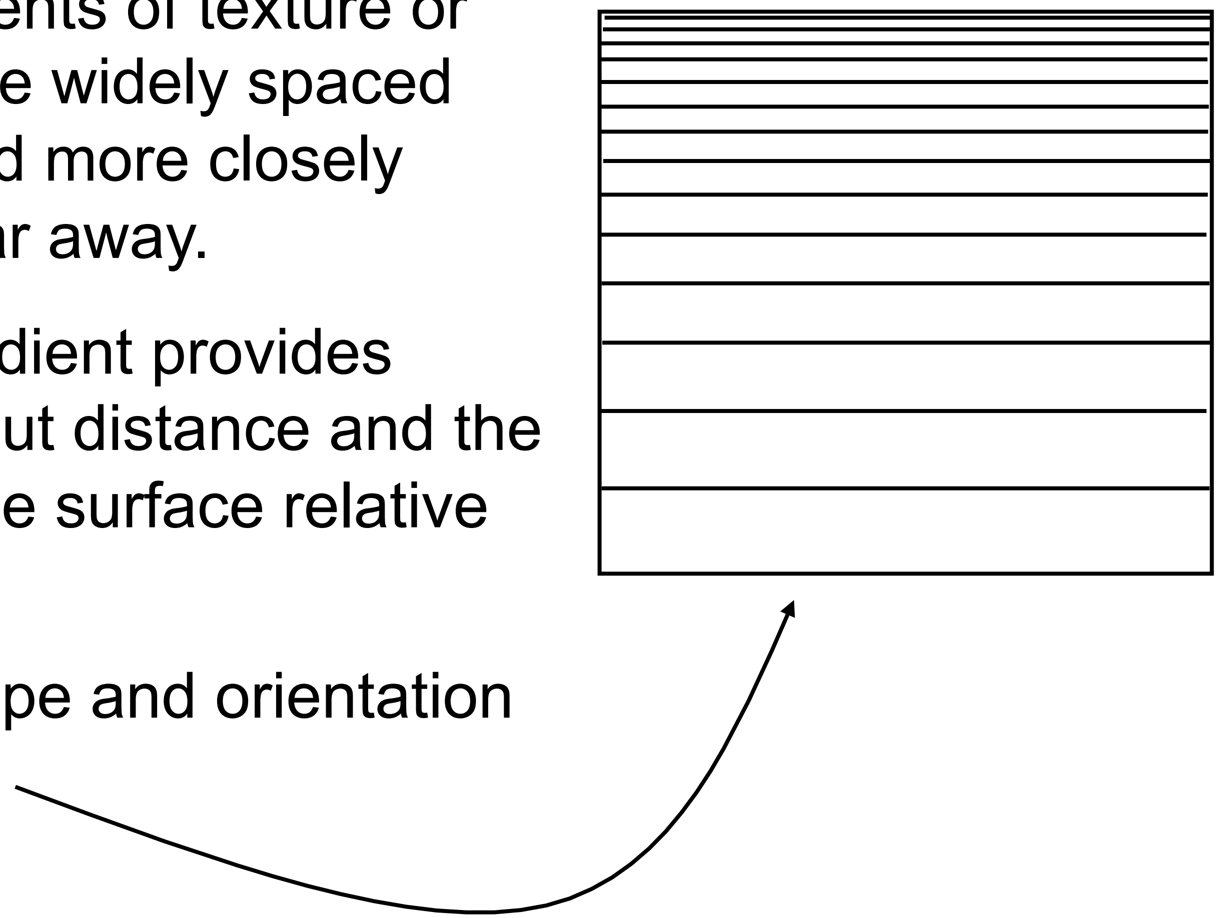
3) Horizon ratio. Two objects of the same size have the same proportion of their height above the horizon.

Texture Gradient

Individual elements of texture or pattern are more widely spaced when closer and more closely spaced when far away.

The texture gradient provides information about distance and the orientation of the surface relative to the observer.

What is the shape and orientation of this surface?



Texture Gradient - 2

Gibson emphasized that the texture gradient was invariant (as a depth cue) with motion of the observer and viewing angle. However, with change in viewing angle, the orientation of the surface to the observer does change. Thus, the texture gradient is a rich source of information.

Vasarely's painting Vega-Nor (in the Albright-Knox Art Gallery) shows how the spacing of elements can be used to create an illusion of shape and depth.

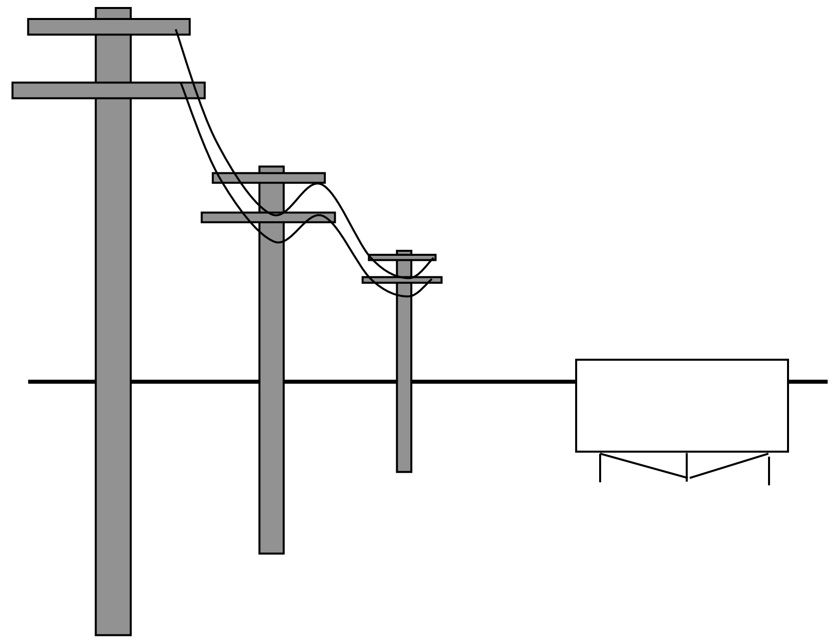
Flow Pattern

In discussing depth perception, relative motion and accretion and deletion at edges were described.

Both of these motion cues to depth emerge from the flow pattern (how the optic array changes as the observer moves). Here, Gibson emphasized the changes in the optic array for the entire field of view. This pattern of movement indicates the direction and speed of the movement of the observer as well as the depth relations among objects in the visual field.

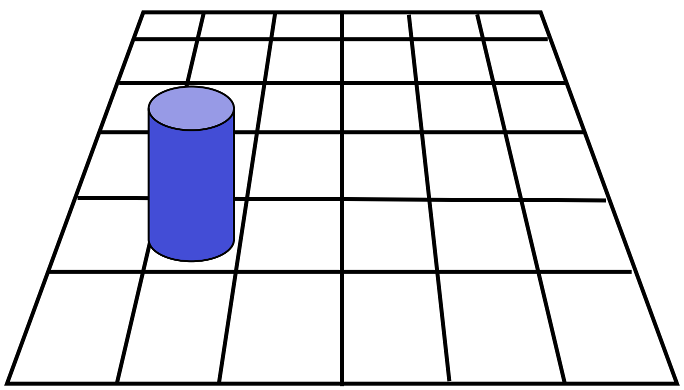
Horizon Ratio

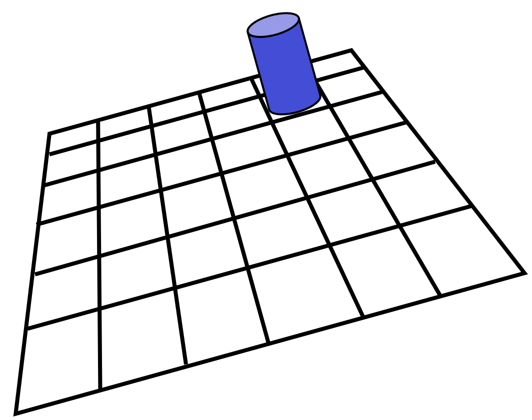
In this illustration, all of the telephone poles appear to be the same height. All have the same proportion of their height above and below the horizon. The billboard appears shorter than the telephone poles since less of its height is above the horizon.



Direct Perception

Do the two cylinders look to be the same size? In the ecological approach, we directly perceive them to be the same size because they cover equivalent amounts of the texture (background). Thus, size perception is direct and not based on size-distance scaling.





Ecological versus Information Processing Approaches

In size perception, the information processing approach emphasizes the extraction of information about retinal size and depth, then combining the two into a size percept.

The ecological approach emphasizes the direct perception of size from the invariants of the texture gradient and horizon ratio. No intervening computation is done.

The resolution to this difference in approach is still being investigated. The neural coding in perception is complex and not fully understood.