Chapter 3: Perception

Overview

Covering the field of perception in one chapter necessarily means narrowing our focus. There is good reason to limit ourselves to vision: vision is the most important sense to human beings and it is the most researched and best understood of the senses. Think about the distinction between sensation (stimulus detection) and **perception** (the *interpretation* of sensory information). It is difficult to conceptualize how much we unconsciously add to our sensations to create perceptions. To understand the interaction between sensation and perception, it is important to first understand the physiology of the visual system. Chapter 3 begins by reviewing the structures of the eye, and then moves on to the pathway the visual information travels to be processed and interpreted within the context of the environment. Information pertaining to "what" an object is, such as colour and shape, travels from the visual cortex to the temporal lobe via a pathway known as the ventral or "what" pathway. Information pertaining to where an object is, or its movement in space, travels from the visual cortex up to the parietal lobe via the dorsal or "where" pathway. The challenge for the visual system is to impose order on the very rich visual stimuli that we are constantly confronted with.

Reading is a good example of the use of context in (**top-down**) processing. Generally, when you are reading a paragraph, you are not paying much, if any, attention to the individual marks that make up each of the letters. Not only are you not attending carefully to those details but you barely even perceive them. Words become so familiar to us that we recognize them by their overall shapes instead of having to read the individual letters. Researchers studying eye movements during reading have discovered that the eye is likely to skip over short, familiar words. We tend to use high-level expectations to assume the identity of words. A good demonstration of this (Reicher, 1969) is discussed in the text. Reicher's participants were to indicate whether they had seen a "k" or a "d" in a string of letters. They were faster and more accurate to identify that it was the letter "k" if the letter string was a real word like "kite" than if it was not (e.g., "kprt"). If participants were assessing the individual components of the letter strings (i.e., **bottom-up** processing), then it shouldn't matter whether it was a real word or not; both "kite" and "kprt" include four letters, one of which is a "k." But it *does* matter: people use top-down processing, even when the item is presented very quickly.

The discussion of **change blindness** provides a chance to examine an issue of visual perception that has very "real-world" implications. Students are fascinated by people's inability to detect changes. There are many demonstrative videos available (portions of well-known movies, for example) that make clear the potential for change blindness. The University of Illinois website (described below) includes a video of the "asking for directions" paradigm and is a compelling illustration of a limitation in detecting perceptual change.

Feature integration theory (**FIT**) proposes that we recognize things by first extracting **features** pre-attentively. Discussion of this theory is a nice precursor to attention issues that will come up in the next chapter.

The **Gestalt approach** to perception proposes that an examination of small components of a perceptual image does not add up to the whole. Elements that are similar and that are close to each other tend to be grouped together. The visual system tends to provide closure to figures that are almost enclosed. We are biased to perceive good continuation in figures; lines that could objectively be seen as discontinuous are subjectively continuous. This tendency toward visual organization applies as well to movement—items that move together in a similar way tend to be grouped together.

It is useful to consider the actual **transformation** between the outside world and our inner representations. One way of conceiving of the transformation is to assume that we construct a representation based on the external environment. J.J. Gibson proposed that, rather than creating internal representations, information can be *directly apprehended* from the environment.

Visual perception can easily be taken for granted; it is not something we typically think about very much. Models of **pattern recognition** (**template matching**, **feature detection**, **recognition by components**) demonstrate with their imperfections the complexity and sophistication of human perception. But perhaps the best way of gaining an appreciation for how much visual processing is used on the world is by learning about individuals who are incapable of such processing. A person with **visual agnosia** can "see" visual stimuli in the sense that he can describe basic forms and shapes but he is unable to recognize objects. Prosopagnosia involves a lack of recognition for faces in particular. The form of a human face is incredibly important to human beings—even newborn babies prefer to look at faces over other stimuli. The existence of prosopagnosia demonstrates that face processing seems to involve a distinct form of visual perception.

Learning Objectives

In this chapter students will:

- Describe the basic physiology of visual perception.
- Outline basic facts about perception learned from the visual deficits of patients with brain damage.
- Describe how objects are consciously perceived and recognized.
- Illustrate the importance of context and observer knowledge in theories of perception.
- Examine the relation between perception and action.
- Explore the nature of multimodal perceptions.

Key Concepts with Illustrative Examples

Akinetopsia (motion blindness) (see page 54)

A person who suffers from akinetopsia is unable to perceive the motion of objects. Click on the link to watch a brief video detailing the experience of someone with akinestopsia. https://www.youtube.com/watch?v=B47Js1MtT4w

bottom-up vs. top-down influences (see page 55)

Bottom-up influences on perceptual experience arise from basic components of an object. Top-down influences arise from high-level knowledge, expectations, and goals.

Try reading the following sentence:
For epxeirecned redares, ltetres dno't hvae to be in the rhigt palecs.

Bottom-up influences indicate that all is not as it should be: the letters are scrambled. Top-down influences, though, allow us to decipher the sentence anyway.

Gestalt psychology (see page 62)

Gestalt psychology proposes that component parts do not determine perception of a whole object.

jumbled word effect (see page 66)

If you have ever tried to proof read your own term papers, you probably know that you will very often miss very obvious typographical errors. In part, this is due to the jumbled word effect. This is the ability to read words in sentences even when some of the letters are mixed up. The jumbled word effect occurs, in part, because you "see" the words that you expect to see. Once you have written the paper the first time, you expect to see the words that you intended to write and sometimes miss what you actually typed.

moon illusion (see page 62)

The moon illusion refers to the tendency for the moon to appear different in size depending on whether it is near the horizon or high in the sky. There are several explanations as to why this occurs, some of which are shown in the following video.
https://www.youtube.com/watch?v=49RztN4Bqu0

optic ataxia (see page 72)

Optic ataxia is a condition characterized by a deficit in the ability to successfully reach for objects. This is especially true when objects are presented in the periphery of vision. Persons with optic ataxia do not lose the ability to identify objects supporting the existence of two pathways for the processing of visual information—the ventral pathway for processing and identifying objects, and the dorsal stream for processing spatial information and movement. Optic ataxia is one symptom of another disorder known as Balint's syndrome which is caused by damage to the parieto-occipital area (part of the dorsal stream).

pattern recognition (see page 56)

Pattern recognition, a term borrowed from the field of computer science, refers to the ability to recognize something for what it is. Humans are incredibly good at certain types of pattern recognition. We can, for example, easily recognize spoken words. To appreciate just how good we actually are at this, consider the long road of speech recognition software development. Siri's not bad but she’s lost with certain accents: http://www.dailymail.co.uk/news/article-2053684/Scottish-iPhone-users-lost-translation--Siri-understand-accent.html

template-matching theory (see page 60)

Template-matching theory is the hypothesis that the process of pattern recognition relies on the use of templates or prototypes. You have probably seen a movie or TV show where police are searching a data base to find a match for a fingerprint found at a crime scene. The computer takes the image of the fingerprint and then compares it to all the finger prints in its data base. If there is a print in the system that has similar lines to the image, it flags it as a match. The finger print system relies on a system very much like template-matching theory.

visual agnosia (see page 52)

Individuals with visual agnosia are unable to visually recognize objects. Oliver Sacks, in his book *The Man Who Mistook His Wife for a Hat*, presented the case of “Dr. P.” When presented with a glove, he described it as “A continuous surface . . . it appears to have five outpouchings . . . a container of some sort.” He is able to describe the globe very precisely, but he has no idea what it is. Note that *vision* is preserved; *visual perception* is not.

word superiority effect (see page 67)

The word superiority effect refers to the finding that it is easier to identify a letter (e.g., p) if it appears in a word (e.g., warp) than if it appears alone. First described by Cattell (1886) and later extended by Reicher (1969), we can generalize the finding to words in that if you are asked to read a random list of unassociated words (e.g., face, tool, drink), it would take you longer than if you were to read words that were either in a sentence, or are related (e.g., face, eyes, mouth).

Discussion and Debate Ideas

1. This chapter largely ignores most of the human senses. The coverage is a pretty straightforward reflection of research in the field of perception: vision is where it’s at. Have students discuss why there is less interest in the other senses. If rats conducted perceptual research, it would be all about smell. But humans are, more than anything, visual creatures.

2. Encourage students to *really* think about how they recognize objects. Ask them which, if any, of the theories of pattern recognition best intuitively captures the phenomenological experience? Have them discuss whether it is useful to ask such a question, or is pattern recognition performed at such a low-level that we don’t have access to insight on it anyway?

3. Put together a slideshow of visual illusions (a good resource can be found here: http://dragon.uml.edu/psych/illusion.html) as a demonstration of how our subjective experience doesn’t always match up with the objective characteristics of a scene. Our perceptual tendencies have roots in evolutionary adaptations. Have students speculate about specific evolutionary explanations of various illusions.

We have a tendency to overestimate vertical distances relative to horizontal distances and so the grey line looks longer than the black line. Why does this make sense from an evolutionary perspective? Vertical distances pose a distinct cost of falling (one that's not an issue in horizontal distances). Perhaps, we have evolved to see vertical as longer as a way of helping to guard against falling.

4. Have a class discussion about the challenges facing someone with visual agnosia. How well could someone with visual agnosia function in everyday life? What would be some particularly serious hurdles? What strategies might she adopt to get by? How could you design methods of rehabilitation?

5. Have students look around the classroom for real-world examples of the Gestalt laws of perceptual organization. Are there any that seem to appear more often than others?

6. What advantage would there be to processing facial information in the fusiform face area, and information about the body in a completely separate area (extrastriate body area)?

7. Persons with autism have difficulty recognizing faces and appear to have fewer cells in the fusiform face area. Discuss this finding within the context of how this affects the evolutionary view of the FFA. Do people with autism have difficulty with facial recognition because they have fewer cells or do they have fewer cells because they have problems recognizing faces? Discuss the reciprocal interaction between the sensory system and the environment.

8. Tucker and Ellis (1998, 2001, 2004) found that participants respond faster to images that afforded a particular type of action (e.g., using the right and vs. using the left hand). Have students generate a list of objects that are typically designed for right-handed persons. Some examples can be found at the following site: http://lefthanded.cdf-design.com/against_us/. How might this affect the coordination between the perceptual system and action system of people who are left-handed?

9. How does context affect facial recognition? Have students think of examples where they saw a person out of context and couldn't remember from where they knew the person.

Further Reading, Media Suggestions, and Teaching Aids

1. **Faceblind: www.faceblind.org**

   This website is a collection of information on prosopagnosia put together by researchers at Dartmouth College, Harvard University, and University College London. It provides a comprehensive description of the disorder, a list of research questions being investigated, and tests for assessing your own ability to recognize faces.

2. **The McGurk Effect: http://www.youtube.com/watch?v=aFPtc8BVdJk**

   Play the video once, as is, and have students guess the sound. Many will report hearing "da." Have them close their eyes. They should now hear "ba." Turn off the sound. Now, the mouth movements should look like "ga." This provides a fascinating demonstration of how we do not perceive sounds in isolation. In fact, we integrate both the sound and the visual information.

3. **Optic Flow and the Moving Room: https://www.youtube.com/watch?v=F4xenIulg_8**

   This video demonstrates the integration of visual information with other sensory information. The video shows how a sense of balance is developed in infants based on the visual information in the environment. When the researcher moves the wall of a child who has been walking for only one month, the child is unable to maintain their balance.

4. **Cool Optical Illusions**
   https://www.youtube.com/watch?v=hrhGTR54E5k&ebc=ANyPxKpE7LacwQxscvApSipljyW3rENAFzU9qt4tq3mTr0rv8P40IrgY4Gu_FPSS18gJAyDis0zJOJqb0sucEvsmHg7LmaA

   This video demonstrates some novel optical illusions. View the above video and have students explain each illusion within the context of the chapter. Some can be explained in terms of Gestalt principles, while others are a bit more complex.

5. **Goldstein, E. B. 1981. The Ecology of J. J. Gibson's Perception. *Leonardo*, 14, 191-195**.

   This article summarizes J. J. Gibson's approach to the interaction between perception and action as written in his three books *Perception in the Visual World; The Senses Considered as Perceptual Systems;* and *The Ecological Approach to Visual Perception*. Topics covered include the influence of texture gradients on movement, the invariant nature of perception as we move through our environment, and perception as an independent process.

Homework or Study Questions

1. **With reference to specific examples, what are some strengths and weaknesses of each object recognition theory?**

   Template-matching involves storing mental copies of environmental stimuli in memory and comparing objects to those templates. For example, the letter "T/t" has two lines that meet at right angles—if you see a stimulus that fits that template, then it is a "t." A template matching explanation makes sense for very simple stimuli. The problem, of course, is that there are all kinds of variations on even simple items, like letters. It seems inefficient (and unlikely) that we would build an inventory of templates that would work to classify "t's" of different fonts and varieties of handwriting. We must be able to judge items as "similar enough" to the template. But how similar is that? The details of template-matching theory are murky.

A feature detection explanation might describe that the letter "T" has a horizontal line on top of a vertical line, unlike the letter "L" which has the horizontal line underneath and to the right of the vertical line. While less rigid than template-matching, feature detection must involve some sort of list of defining features, which can get complicated for three-dimensional objects. This is especially true for things that can move, change form to some extent, and look very different from different angles. As well, it could be argued that a feature detection explanation would have difficulty explaining this:

UDOQPG	MNIXWV
GOUPQD	NEXIWM
DOZQUG	IXNWME
COQPGD	VIZWMN

If you were to compare the two displays and try to find the "Z" in each, you'd probably find that it is much easier to find it in the display on the left than on the right because the left display has letters that have a lot of curves, whereas the display on the right has letters than are angular. It's as if we are able to "shut off" certain feature detectors when looking for the "Z" in the display on the left. A feature detection would have difficulty explaining how we look for the *absence* of features.

Recognition by components proposes that when we recognize objects, we are, in some way, breaking them down into their components (geons) and noting where the components join together. The pattern is then matched to information in the memory to recognize the object. This explanation allows for much more complexity in analysis. A problem with the recognition-by-components theory is that, like template matching and feature detection, it only emphasizes the perception of component pieces of a pattern. What happens when we look at a visually-rich picture like the Mona Lisa? We are quickly able to say what it is but we are only able to do so by focusing on the complex pattern as a whole rather than the individual components (the geons, that is).

2. Describe Selfridge's pandemonium model of pattern recognition.

The model consists of three levels:

- At the data level is the image itself, a pattern of features.
- At the next level are "cognitive demons." Cognitive demons are on the look-out for particular patterns and will shout when they find one they're looking for. The volume of the shout increases with greater similarity to the searched-for pattern. With multiple demons shouting, pandemonium ensues (hence the name).
- At the top level is the decision demon. The decision demon determines which pattern is recognized by judging which cognitive demon is shouting the loudest.

3. Come up with an instance of each of the Gestalt organizational principles.

Instances will vary greatly but may include photographs, videos, or descriptive examples of the following:

- Principle of experience

- Principle of proximity
- Principle of closed forms
- Principle of good contour
- Principle of similarity
- Principle of common movement

4. **Review the diagram of the eye (p. 53) and label each of the parts described in the book. Follow the pathway of light from the front of the eye to the back.**

- Cornea → pupil → iris → lens → retina

5. **Name and describe the two pathways responsible for processing specific components of visual information.**

The ventral "what" pathway is the pathway from the visual cortex to the inferior temporal lobe. It processes information that helps us identify an object including visual characteristics such as shape and colour. The dorsal "where" pathway projects from the visual cortex to the parietal lobe and is responsible for processing information about object movement of object location. The dorsal pathway allows us to act upon objects in our environment.

6. **Explain the empirical theory of colour vision. How does it explain individual differences in colour perception?**

The empirical theory of colour vision suggests that colour perception is influenced by prior experience with the way different illuminations affect colour. With this theory individual differences in colour perception occur due to a) prior experience with different sources of light and how they affect colour perception, and b) the inferences that the visual system make as a result.

7. **How are the McGurk effect and the visual prepotency effect related?**

The McGurk effect and the visual prepotency effect both demonstrate the importance we place on visual information in interpreting auditory information.

8. **Explain the intentional binding effect and provide an example.**

The intentional binding effect refers to the perception that events that take place after one has taken action occur sooner than they actually did. An example is if you are asked to press a button and then you hear a tone. If you are asked to estimate the time between when you pressed the button and when you heard the tone, you will likely estimate the time as being shorter than it actually was.

9. **In February 2015, a controversy arose over the colour of a dress. While some saw the dress as being blue and black, others saw it as white and gold. Provide an explanation for this phenomenon. What does this tell us about colour perception?**

One explanation for this is that the photo of the dresses was taken against slightly different backgrounds. The backgrounds give the visual system two different ideas as to the type of light illuminating the two dresses and the visual system adjusts the colour it perceives on the dresses. This suggests that colour perception depends on both the wavelengths of light as well as the colour information provided by the context of the background.

Suggestions for Research Paper Topics

1. Do all human beings use Gestaltist organizational principles similarly? See if you can find any evidence of cultural differences in perceptual organization.

2. Summarize a case study of a patient with optic ataxia.

3. Review J.J. Gibson's theory of ecological optics and attempt to apply it to another (e.g., auditory) mode of perception.

4. What hazards might someone who suffers from akinetopsia (motion blindness) face in their everyday life? What does this tell us about the adaptive nature of the visual system?

5. Why might the brain be organized in such a way that different types of visual information are processed in different areas of the brain? What are the advantages and disadvantages of this organization of function?

6. There is some controversy about the true function of the fusiform face area, as some researchers believe it processes more than just faces, but is involved in processing of familiar objects. Review the evidence supporting and contradicting both sides of the controversy.

7. Elderly persons are at an increased risk of injuries due to falling. How might this be explained according to J. J. Gibson's ecological optics?

8. Facial recognition is an important part of our social experience. How would prosopagnosia affect the social experience? How might this differ from someone who is blind? Why?

9. Watch the following video: **http://www.youtube.com/watch?v=aFPtc8BVdJk**. The McGurk effect (demonstrated in the video) can be eliminated if you have the person close their eyes, as they are no longer influenced by the visual information. What happens if the person were to move their eyes to the side of the visual image? See how far away from the central visual point on the image you have to move your eyes before you no longer experience the McGurk effect. How might you explain what you find?