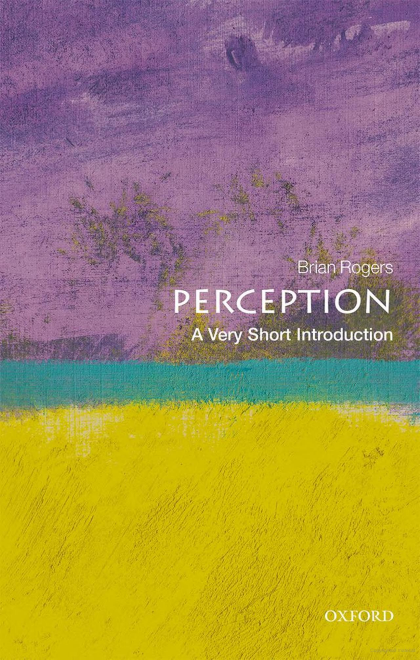


Brian Rogers

# PERCEPTION

## A Very Short Introduction

OXFORD

Brian Rogers

# PERCEPTION

## A Very Short Introduction

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries



First edition published in 2017

Impression: 1



Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2017942397

ISBN 978–0–19–879100–3

Printed in Great Britain by
Ashford Colour Press Ltd, Gosport, Hampshire

# Contents

# Acknowledgements

I shall always be very grateful to the late Richard Gregory for introducing me to the delights of visual perception while I was a first-year undergraduate student and for his inspiration and friendship over the following years. My graduate supervisor and subsequent close friend, Stuart Anstis, showed me how to do (and how not to do) experiments as well as being a source of constant encouragement. The late Ian Howard showed me how to write more clearly and concisely and we shared an interest in building experimental apparatus to investigate perception in more real-world situations. I am also grateful to the many other colleagues and friends who have influenced the way I think about perception including Alan Cowey, Jan Koenderink, Christopher Tyler, Barbara Gillam, Alan Gilchrist, Bill Epstein, and my post-docs and students.

# List of illustrations

# Chapter 1
# What is perception?

## Definitions

What is perception? The word itself can be used in two different ways. First, it can refer to our *experience* of seeing, hearing, touching, tasting, and smelling objects and individuals in the surrounding world. As human observers, we are able to describe our subjective experiences—I see the sky as blue, I see my desk as having a particular size and shape, and I feel the texture of my sweater. Seeing is the particular word used to describe our visual experiences in contrast to our auditory, tactile, olfactory, and gustatory experiences. However, the word perception can also be used to refer to the *processes* that allow us to extract information from the patterns of energy that impinge on our sense organs. Thinking about perception as a set of processes has the advantage that it includes situations where there is no subjective experience—such as the control of our balance using visual and proprioceptive (felt position) information. In these cases, the information reaching our senses guides and controls our behaviour but there is no accompanying subjective experience. For the remainder of the book, I will use the word perception in this more general sense to refer to the use of sensory information whether or not it creates a subjective experience.

Some writers, such as the 19th-century scientist Hermann von Helmholtz, have suggested that there is a distinction between *sensations* and *perceptions*. He considered sensations to be the result of early or 'low-level' sensory processes in contrast to our perceptions that are thought to be a consequence of 'higher-level' processes that give sensations their meaning. However, this distinction raises awkward questions: first, is it possible to experience raw sensations that lack meaning? And second, how are we able to access the outputs of these low-level processes? A further assumption in the perception literature is that the processes responsible for our perceptions are not at all obvious or straightforward. For example, on the first page of Richard Gregory's excellent book *Eye and Brain*, he writes:

> We are so familiar with seeing, that it takes a leap of imagination to realise that there are problems to be solved. But consider it. We are given tiny distorted images in the eyes, and we see separate objects in surrounding space. From the patterns of stimulation on the retinas we perceive the world of objects, and this is nothing short of a miracle.



In Chapter 2, I am going to take issue with this characterization but for the moment we can see how this view is consistent with the fact that our perceptual systems can be fooled—we sometimes perceive the world incorrectly and experience illusions. One of the best-known examples is the Ames Room, named after its American inventor Adelbert Ames (Figure 1). Unlike most of the rooms we encounter, the Ames Room is not rectangular in shape but instead is trapezoidal with the far left corner much further away from the observer. However, when the room is viewed through a peephole using just one eye, it is seen as a normal rectangular room. Moreover, the two individuals standing in the two corners appear to be of very different sizes. There is a discrepancy between the reality—a trapezoidal-shaped room—and what we see—a normal, rectangular room—and hence our perception is labelled as illusory or non-veridical.

**1. The Ames Room is trapezoidal in shape but from the particular peephole from which this photo was taken, the room is seen as a normal, rectangular room.**



More recently, a similar theme of the difficulties faced by our perceptual systems has been highlighted by those working in the field of artificial intelligence and computational vision in their attempts to model our perceptual processes. In the visual modality, this is referred to as the problem of *inverse optics*. Conventional optics provides us with a description of how the patterns of light from the surrounding 3-D world reach the eye and create a 2-D image across the array of receptors we call the retina. In contrast, inverse optics—how we go from the 2-D retinal image to the 3-D world out there—is seen as problematic because there are many different 3-D scenes in the world that could have created exactly the same 2-D image (Figure 2). As a result, it is not obvious which of the possible real-world scenes was responsible for creating the image. Note that inverse optics is not just a problem for our perception of the 3-D world—similar problems exist in the perception of colour, motion, and form, as we shall see later.

**2. The ambiguity of the retinal image illustrated in Magritte's painting *The Human Condition*. The real-world 3-D scene and the 2-D painting on the canvas create the same pattern of light at the eye.**

## Approaches to understanding perception

A theme that is common to both the traditional approaches to perception and the more recent computational research is that there is insufficient information in the patterns of light reaching the eyes (and similarly for the other senses) to explain the richness

**3. A photo of a hollow (concave) mask of Beethoven but we perceive it to be a normal (convex) face.**



and the lack of ambiguity in what we perceive. For many, the solution to this problem lies in our past experience and our knowledge of the world. Experience tells us that rooms are typically rectangular, rather than trapezoidal, and that faces are typically convex rather than concave. Hence, when we are faced with a trapezoidal room (Figure 1) or a hollow face (Figure 3), we choose the more likely interpretation based on our past experience. For the computational researchers, the solution to the ambiguity of the 2-D image is expressed in terms of the *constraints* that are incorporated into the processing mechanisms.

The alternative view, due largely to the work of the American psychologist James Gibson, is that there is actually a wealth of information in the patterns of energy reaching our senses and that the task for any perceptual system is to extract that information. As a consequence, there is no need to invoke ideas of inference or hypothesis testing. The difference between these two radically different conceptions of our perceptual processes is a theme that will recur throughout the book and you may find it remarkable that there is still no real consensus as to which of these approaches to the subject is the more appropriate. In writing this book, my intention is to de-mystify the process of perceiving by showing that there is, in reality, plenty of information to specify not just the low-level aspects of our perception—such as the lightness, colour, motion, depth, and form of objects in the surrounding world—but also their affordances (the significance of the information) for us as perceivers.

# Chapter 2
# Perceptual theories—direct, indirect, and computational

## Background

The idea that the information reaching our senses is insufficient to account for the richness of our perceptions can be traced back to the writings of Hermann von Helmholtz in the 19th century. He thought of perception as a process of making 'unconscious inferences' about what is out there in the world. More recently, Richard Gregory has argued that our perceptions are like hypotheses in science—we make guesses about what is there in the surrounding world that are based on information from our knowledge and past experience. A variant on the same theme has been suggested by the American psychologist, Irvin Rock, who thought of our perceptions as being the result of 'intelligent, thought-like processes'. All three proposals are referred to as *indirect* or *constructivist* theories of perception and they share the idea that we need to construct or reconstruct our perceptions from limited sensory information. This view of perception sounds very plausible because we, as humans, are able to make *cognitive* inferences and hypotheses that go beyond the available evidence. Moreover, if our perceptions are only guesses or inferences about the external world, we should expect that we will sometimes perceive the world incorrectly and suffer from illusions, and this seems to be the case.

There is, however, a very different conception of our perceptual processes that was first suggested by the American psychologist James Gibson. He argued that far from being insufficient, the information reaching our senses is actually rich and, as a consequence, there is no need for elaboration or construction by processes such as inference or hypothesis testing. In his view, perceptual information merely needs to be 'picked up'. Moreover, he regarded the ambiguous and illusory situations that are used by the indirect theorists to support their theoretical position to be a consequence of the impoverished stimulus situations that are typical of many traditional experiments on perception. He pointed out that in the majority of those experiments, observers have to remain stationary and are obliged to keep their eyes fixed on a particular point in an unchanging scene. He contrasted this with everyday life where we are constantly moving around and gazing in different directions. Under these circumstances, our movements through the world create a constantly changing pattern of light at the eye(s) that he called *optic flow*. Optic flow not only provides information about the 3-D structure and layout of the world (exterospecific information) but also information about the observer's movements with respect to the world (propriospecific information).

This alternative theory of perception, which is usually referred to as *direct* perception, not only differs from indirect theories in the assumption that there is a sufficiency of information but also in terms of its conception of what perception is for. Traditional indirect theories have assumed that the primary purpose of perception is to create our subjective experiences—our *qualia* as the philosophers describe them. In contrast, Gibson claimed that the primary role of perceptual processes was to guide action and he had very little to say about subjective experience.

> Perceiving is an achievement of the individual, not an appearance in the theatre of his consciousness.

In this respect, Gibson can be thought to have more in common with the computational researchers like David Marr since it is clear that there is no place, and indeed no need, for subjective experience in a machine vision system. The idea of direct perception can also be thought to be more consistent with what we know about the perceptual systems of animals lower down on the phylogenetic scale. Their behaviour appears to be controlled more directly by the available sensory information and there appears to be no need to invoke ideas of inference or hypothesis testing, either consciously or unconsciously.

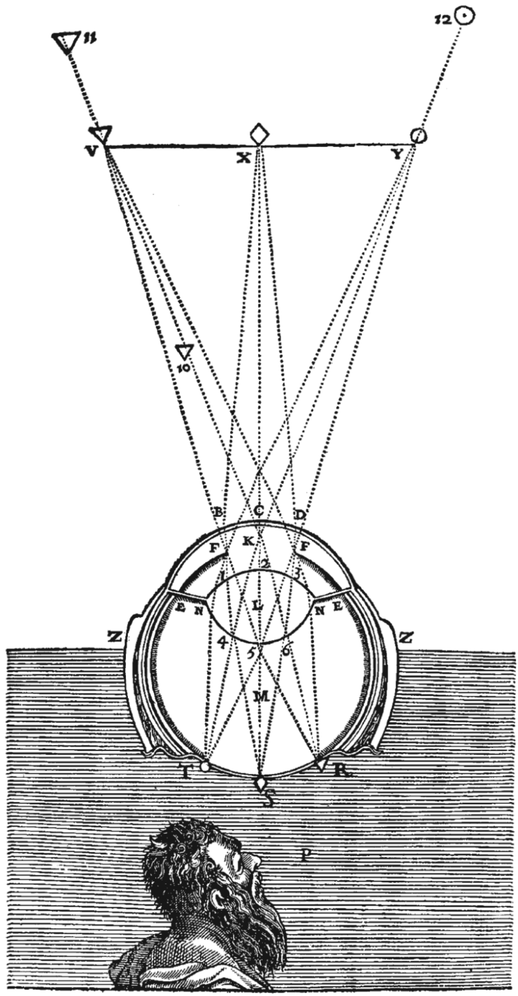
In his book *Vision*, David Marr argued that his view of perception had much in common with that of James Gibson. He wrote: 'In perception, perhaps the nearest anyone came to the level of computational theory was Gibson.' However, Marr went on to say that 'although some aspects of his thinking were along the right lines, he did not understand properly what information was, which led him to seriously underestimate the complexity of the information-processing problems involved in vision…'. Two questions arise. First, is it appropriate to characterize perception as an 'information processing task', and second, what does it mean to say that the information-processing problems of vision are 'complex'? Attempts to build machine vision systems to do basic perceptual tasks have proved to be extremely difficult but in what sense are the perceptual mechanisms of biological systems 'complex'? I will revisit the idea of complexity in Chapter 6.



## Indirect or constructivist theories of perception

Let us consider these three different conceptions of what it means to perceive—the indirect, the direct, and the computational—in more detail. The origins of indirect or constructivist theory can be traced back to the physicist, physiologist, and vision scientist Hermann von Helmholtz who emphasized the importance of experience in shaping our perceptual abilities. His ideas can be contrasted with the nativist views of his rival Ewald Hering who believed in the importance of innate mechanisms in perception. The two are best known for their differing ideas of how we perceive colour—Helmholtz arguing for a trichromatic theory and Hering arguing for the involvement of opponent processes (see Chapter 3). However, their disagreement about the role of learning versus innate mechanisms is best illustrated in their contrasting views of how *visual direction* (the direction of objects with respect to the eye) might be coded. Simple geometry shows how the light coming from different points in space with respect to the eye is focused on different regions of the retina (Figure 4). As a consequence, there is a systematic relationship between the position of a point in space (with respect to the eye) and the location of its image on the retina. Both Helmholtz and Hering assumed that different receptors in the retina have different and unique *local signs*, that is, different receptors are tagged or labelled with a particular visual direction with respect to the eye. They differed, however, in their views of how those retinal receptors acquire their local signs. Hering believed that local signs were innately given whereas Helmholtz argued that they needed to be learned (calibrated). Helmholtz suggested that the information used to calibrate the local signs comes from the magnitude and direction of the eye movements needed to bring the image of a particular peripheral point in space onto the fovea. As an aside, it is interesting to note that in spite of Helmholtz's strong empiricist views, his theory assumes that some things must be innate. In particular, (i) that eye movement signals are given (i.e. they do not need to be learned) and (ii) that the fovea of the retina has a special status in that its local sign is innately given. Helmholtz's suggestion that local signs need to be calibrated is, of course, consistent with the philosopher George Berkeley's idea that 'touch educates vision'.

Helmholtz's empiricist view of perception was accompanied by a belief that what we perceive goes beyond what is available to the senses. Insufficiency of information is the essence of constructivist theory. He argued that our perceptions are based on a process of




**4. The projection of light rays onto the retinal surface of the eye illustrated in René Descartes's *La Dioptrique*. There is a one-to-one relationship between the location of a point with respect to the eye (e.g. point 12) and the location of its image on the retina (T).**

'unconscious inference' whereby we seek meaningful explanations of the patterns of stimulation reaching our senses. In his *Handbook of Physiological Optics* he wrote:

> The sensations of the senses are tokens for consciousness, it being left to our intelligence to learn how to comprehend their meaning…

More recently, Richard Gregory put forward a similar idea in his proposal that we use induction to create 'perceptual hypotheses' from the data obtained from the senses. Gregory drew a parallel with the way in which we make hypotheses in science that go beyond the limited empirical data. In other words, both Helmholtz and Gregory share the assumption that perception is an inferential process with the need to make guesses (inferences or hypotheses) about what the sensory information means. In his book *The Logic of Perception*, Irvin Rock suggested that our perceptions are the result of 'intelligent thought-like processes'. All three indirect theories make similar assumptions: (i) perception is a process of construction rather than the result of some direct, mechanistic process; (ii) the information reaching our senses is insufficient to account for the richness of what we perceive; (iii) meaning is added rather than contained in the sense data.



These ideas seem eminently plausible because we, as humans, are clearly able to make inferences and hypotheses at a conscious, cognitive level that go beyond the available evidence. Rock contrasted his idea of indirect perception—in which there is no simple, one-to-one relationship between sense data and what we perceive—with Gibson's idea of direct perception, which Rock characterized as a 'stimulus theory'. In animal learning, we know that particular stimuli trigger particular responses and Rock argued that Gibson's direct theory implies that there is a direct link between stimuli and our perceptions (S → P). To refute this idea, Rock pointed out that there are many situations where a single stimulus can give rise to multiple percepts, using examples of ambiguous figures like the Necker cube and Rubin's vase