Chapter 1

The Domain of Visual Science

Contents

1.1 What is vision science? .................................................. 3
1.2 Three kinds of questions, three kinds of mapping rules .......... 6
1.3 Design questions ....................................................... 16
1.4 The mind/brain problem ............................................. 17
1.5 Linking propositions .................................................. 18
1.6 A more perceptual perspective ..................................... 23
1.7 An interdisciplinary field ............................................ 25
1.8 Summary: The domain of vision science ....................... 27

1.1 What is vision science?

Vision science is the study of vision, the visual system, and the relations between the two.

When vision scientists study vision, we study what and how well people see. The scientific goal is to describe and quantify our sensory and perceptual capacities – our capacities to respond to physical stimuli by using our eyes. What is the dimmest light we can detect? How fine are the finest details we can resolve? What is our color vision, or our stereovision, or our perception of motion like? How accurate are our perceptions of the sizes, shapes, and locations of objects? How well can we recognize objects? What can we see, and what can we not see?

When we study the visual system, we study the neural machinery – the optics, photochemistry, anatomy, and physiology of the eye and the parts of our brains that serve vision. How fine an image is made by the optics of the eye? How is light absorbed? Through what processing stages does incoming information pass? How is information recoded – what computations are performed – as we go from one stage of neural processing to the next? To what physical stimuli or aspects of the visual scene are neurons at different levels of the visual system tuned to respond?

What about the relations between the two? For many vision scientists, studying vision and the visual system separately is not the ultimate goal. Rather, the ultimate goal is to explain why we see as we do, on the basis of the properties of the neural machinery that makes seeing possible. Why can we detect some lights and not others; resolve some spatial details and not others; see objects
accurately under some conditions but not others? Vision scientists want to understand how the computations carried out by the visual system both enable and limit our visual capacities, and how they leave their marks on our visual perception. We will call these attempted explanations causal stories.

1.1.1 Entities and mapping rules

Let us now put these questions in a slightly broader context. As shown in Figure 1.1, vision science is concerned with three kinds of entities, and three kinds of mapping rules.

Let us start with the three kinds of entities. The first kind is physical objects, or physical stimuli – the physical objects and light sources that send light to our eyes. The second is physiological states – the states of the many varieties of neurons in the visual system, occurring in response to the physical objects and light sources that lie in front of us. And the third is perceptual or phenomenal states – conscious states that usually correspond remarkably well to the objects and other stimuli in the physical world. Typically, we will accept appropriate behavioral reports as a stand in for perceptual states. These three kinds of entities are shown by the boxes in Figure 1.1.

Between each pair of entities there is a set of mapping rules. We are looking for rules of correspondence of the form, entities $x_1, x_2$, in the physical domain occur in conjunction with entities $y_1, y_2$, in the neural domain, and with entities $z_1, z_2$, in the perceptual domain. The fundamental goal of vision science is to determine the mapping rules between each pair of entities, by whatever techniques are required, and however simple or complex these mapping rules might be. The three types of mapping rules are shown by the three arrows in Figure 1.1.

The three types of mappings are studied by three very different kinds of techniques. We study Type 1 mappings – mappings between physical stimuli and perceptual states – by means of the discipline of psychophysics, using sophisticated behavioral techniques to ask human subjects what they see when they view particular stimuli. We study Type 2 mappings – mappings between physical stimuli and neural states – with the techniques of visual neuroscience, such as presenting particular stimuli and recording the activities of particular neurons at various levels of the visual system.

What about Type 3 mappings – mappings between neural states and perceptual states? Suffice it to say that at this point the techniques used for exploring Type 3 mappings are much harder to define. At the same time, for many vision scientists, these are the heart of the matter, because as stated earlier, the ultimate goal of many vision scientists is to explain why we see as we do, on the basis of the properties of the neural machinery that makes seeing possible.

The world as perceived is a strikingly accurate representation of the physical world, allowing us both to perceive objects and to carry out appropriate motor activity with respect to them. But the existence of psychophysics notwithstanding, there are no direct causal mappings between physical and perceptual entities. The perceptual representations we have of the physical world are created by passing through the other two legs of the triangle: first through physical/physiological and then through physiological/perceptual mappings.

In the next few sections of this chapter, we will expand on each of these types of mappings. Then, in later sections of the chapter, we will expand at length on Type 3 mappings – mappings between neural states and perceptual states – because of their philosophical complexity and because of the fascination they carry for DT, the author of this book.
1.1. WHAT IS VISION SCIENCE?

The domain of vision science is comprised of three types of entities and three types of mapping rules. The entities are physical stimuli, physiological (neural) states, and perceptual (phenomenal) states. The mapping rules are: Type 1, from physical stimuli to perceptual states; Type 2, from physical stimuli to neural states; and Type 3, from neural states to perceptual states.

Figure 1.1: The domain of vision science. Three types of entities and three types of mapping rules make up the domain of vision science. The entities are physical stimuli, physiological (neural) states, and perceptual (phenomenal) states. The mapping rules are: Type 1, from physical stimuli to perceptual states; Type 2, from physical stimuli to neural states; and Type 3, from neural states to perceptual states.
1.1.2 An example: Grating acuity

Let’s take a concrete example of a physical/perceptual mapping. Figure 1.2 shows seven sets of regular black and white stripes, called square-wave gratings, and one homogeneous gray field. The stripes in each grating are half as wide as the stripes in the next coarser grating. At normal reading distance, you can probably see the spatial variation of light level across the all of the gratings (e.g., A-G). Now view the gratings from a 3 meter distance: what is the finest grating that you can see?

Somewhere between gratings E and G, your perception of black and white stripes probably fades perceptually into a uniform gray, and cannot be distinguished, or discriminated, from the homogeneous field H. Your grating acuity is defined as the finest stripes that you can just barely perceive as stripes, or discriminate from the homogeneous field on the basis of their spatial pattern. Grating acuity is a measure of the limit of detail you can resolve in space; it defines the limit of the spatial resolution capacity of your visual system.

We now elaborate on the three sets of entities and the three sets of mapping rules that make up the domain of vision science, using grating acuity as our example.

1.2 Three kinds of questions, three kinds of mapping rules

1.2.1 Type 1 – From physics to perception

What are Type 1, or physical/perceptual, mappings? Vision scientists want to know how and how well people see – to measure and quantify human sensory and perceptual capacities. To find out, we bring people (usually called subjects or observers) into the laboratory, and use well-controlled physical stimuli and sophisticated behavioral, or psychophysical, techniques to measure their visual capacities. The results of such experiments yield objective descriptions of the facts about visual acuity, color vision, distance perception, object recognition, and so on. We will look in detail at psychophysical techniques in Chapter 2 and 3.

The most immediately interesting fact about grating acuity is that it is so readily definable. There is a range of coarse gratings that you can resolve, an abrupt transition, and then a set of finer gratings that you can’t resolve. Notice the first of many mismatches of properties between the physical and the perceived. The physical variation is continuous – there is nothing in the stimulus continuum that suggests a basis for any break of perceptual properties – but the change of perception from seeing to not seeing is abrupt.

At the perceptual level, grating acuity has several additional interesting properties. The finest grating you can see on the page varies with the viewing distance, the light level, and the part of your field of view in which the grating is presented. Try these experiments. First, prop up the book across the room, perhaps 20 feet away. From this distance, you will probably be able to resolve only the one or two coarsest gratings. Now walk toward the book. Every time you cut the distance of the book in half, you should be able to resolve one more grating.

Second, turn down the lights in the room in progressive stages, or prop the book up in the light of a window in the evening, when the outdoor light is steadily decreasing. As the light level decreases, your grating acuity will decrease, and you will need to move closer to resolve gratings that were easily resolvable at your original distance in full daylight. And third, instead of looking directly at the gratings, look above them by various amounts while still trying to resolve the stripes. The greater the eccentricity of the grating – the greater the displacement from your center of vision – the lower will be your grating acuity.
1.2. THREE KINDS OF QUESTIONS, THREE KINDS OF MAPPING RULES

A. 0.5 cy/cm
B. 1 cy/cm
C. 2 cy/cm
D. 4 cy/cm
E. 8 cy/cm
F. 16 cy/cm
G. 32 cy/cm
H. gray field

Figure 1.2: Seven square wave gratings and a blank (homogeneous) field. The gratings are labeled A, B, C, D, E, F, G. The gray field, H, approximately matches the gratings in average light intensity. If you have normal vision, at 3 meters distance you will probably be able to resolve gratings A-E, but not gratings F-G. The finest grating you can discriminate from the blank field, H, defines your grating acuity. Unfortunately, imprecise printing often makes irregularities in the finer grating that make them more visible than intended.
CHAPTER 1. THE DOMAIN OF VISUAL SCIENCE

By picking out the finest grating you can see under a variety of conditions, you have just been a subject in an (informal) psychophysical experiment. You have made a series of measurements of the mappings from physical to perceptual entities. You have encountered the perceptual phenomenon of grating acuity, and you have learned about several important parameters – distance, light level, and eccentricity – that influence it.

1.2.2 System Properties: Bumblebees can fly!

Grating acuity and its variations with distance, light, and eccentricity can be called system properties of vision. Such system properties are interesting in and of themselves. But they become more interesting when we realize that system properties provide us with logically compelling information about some of the physiological properties of the visual system, without a single physiological experiment having been done. Oddly, we are arguing that perceptual results imply physiological conclusions. A fancier way of saying this is, system properties place important constraints on models of the visual system. The constraints depend on whether you (the subject) resolve the grating or not.

If you resolve the grating, information physically present in the stimuli must have been retained.

When you discriminate a grating from the homogeneous field, it follows that information that the two stimuli differ is retained from the physical stimulus, all the way through every one of the series of anatomical/physiological stages that make up your visual system. It is retained through every link of a causal chain, right up to your conscious perception, and right out through whatever motor system you use to tell the experimenter you resolve the grating. The question then becomes: How is the information carried, or coded, at each anatomical stage?

If you don’t resolve the grating, information physically present in the stimuli must have been lost.

When a grating is present but you don’t discriminate it from the homogeneous field, it follows that information that the two stimuli differ must be lost somewhere, at one or more stages of processing in your visual system. Like information retention, information loss implies a physiological conclusion – there exists a stage or stages of visual processing at which the information is lost. The question then becomes: at what anatomical stage (or anatomical locus) is the information lost? Questions of this kind are sometimes called locus questions.

Because they specify the limits of actual visual function, system properties like spatial resolution exert a great deal of power over mathematical and physiological models. A classic joke based on the importance of system properties concerns some early aerodynamic engineers who tried to make a mathematical model of the bumblebee, to find out how it flies. They tried and tried, but the model bumblebee fell out of the air every time. Eventually, the engineers concluded that bumblebees can’t fly! But in fact, bumblebees can fly, so immediately we know that something about the model had to be wrong. Similarly, if a vision scientist makes a model of the visual system, and that model predicts that human beings can’t resolve gratings as fine as the ones you could resolve in Figure 1.1, we know immediately that something is wrong with the model.
In more abstract terms: the system properties of vision are “black box” measurements made on a highly complex, multi-stage anatomical/physiological system, and rarely reveal the details of the machinery inside the box. But even in the absence of knowledge about the inside of the box, system properties put fundamental constraints on models of how it works. If the person can do X, then the underlying visual system must be such as to allow X to occur. Any model that claims that the person can’t do X must be wrong. DT refers to arguments of this form as bumblebees can fly arguments. On the other hand, if the person can’t do X, then information is lost, and there must exist a locus (or loci) of information loss.

In addition to such logically compelling implications, system properties often play a less formal but very important theoretical role in vision science. That is, system properties can encourage speculation and theory-building concerning information processing within the visual system. Such speculation, and the computational models it generates, can provide the motivation for visual neuroscientists to go and look for particular kinds of elements or processing circuits within the eye and visual system. We will see many examples of this pattern of argumentation, from psychophysics to physiology, as we go along.

1.2.3 Type 2 – From physics to physiology

In this section we provide a brief preliminary encounter with Type 2 mappings. We begin with a simplified cartoon of the anatomy of the eye and the early parts of the visual system.

Figure 1.3 and Figure 1.4 show an overview of the early parts of the human visual system. As shown in Figure 1.3A, the optics of the eye create an optical image of the visual scene, called the retinal image, at the back of the eyeball. The major optical elements that form the image – the cornea, the iris (and the hole in its center, the pupil), and the lens – are shown in Figure 1.3B. The retinal image is focused on the retina, a thin sheet of neural tissue that lines the back of the eyeball. The fovea is a small, central region of the retina, specialized for high acuity. When you looked at each acuity grating in Figure 1.2, you turned your eyes to place your fovea under the retinal image of that grating.

Figure 1.4A shows a cartoon of a small piece of the retina, and some of the types of neurons it contains. The photoreceptors at the back surface of the retina catch the incoming light, and initiate a set of neural signals. Several other types of cells within the retina process these signals before they arrive at the retinal ganglion cells. The output processes (axons) of the ganglion cells make up the optic nerve. As shown in Figure 1.4A, the optic nerve passes out through the retina and eyeball, and leaves the eye.⁴

As shown in Figure 1.4B, the optic nerve projects to a way station – the lateral geniculate nucleus (LGN), deep within the brain – before signals are sent on to the primary visual cortex, a cortical region located at the very back of the brain. From there, signals project outward and forward to many higher levels of cortical processing (see Figures 17.5 and 17.4 in Chapter 17 for an intimidating preview).

¹Notice that, oddly, the retina is “in backwards” – the photoreceptors lie at the back (or outer) surface of the retina. In consequence, the light must pass through all of the other neurons in the retina before getting to the photoreceptors for absorption; and the optic nerve must pass through the retina to get out of the eye, creating a blind spot in your visual field. The probable reason for having the retina in backwards is that (as we will see) the photoreceptors are highly active metabolically, and need to be near a blood supply; and a blood supply that traversed across the front of the retina would itself get in the way of the retinal image.
Figure 1.3: Overview of the eye and its optics. A. The optics of the eye form an image of the physical world on the back of the eyeball. B. This sketch shows a horizontal section through the right eye, labeling the major optical elements of the eye that work together to form the retinal image: the cornea, pupil, and lens. It also shows the retina, a thin sheet of neural tissue that lines the inside of the eyeball; the fovea, a central region of the retina specialized for high acuity; and the optic nerve, the nerve bundle that leaves the eye and carries visual signals toward the brain. (Modified from Cornsweet, 1970, Fig. 3.11, p.40)
Figure 1.4: Overview of the visual system. A. A small section of the retina (schematic), showing its layers of neurons. The photoreceptors capture the light, and pass on neural signals (via several other types of neurons) to the ganglion cells. The output processes (axons) of the ganglion cells travel across the inside surface of the retina to join in a bundle, and plunge through the retina to form the optic nerve. B. A sketch of the pathway from the eye to the visual cortex. The eye is at the lower left. The optic nerve and the optic tract carry the signals sent by the ganglion cells to the lateral geniculate nucleus (LGN). Axons from the LGN project to the primary visual cortex and from there to other cortical areas (not shown) for further processing (B modified from Kandel, Schwartz, and Jessell, 2000, Fig. 27.4, p. 527).
When we ask about Type XR or physical/physiological mappings, we are trying to trace the features of each physical stimulus through the visual system. We want to understand and quantify the characteristics of the eye’s optics, and the means whereby light is transformed into physiological signals. We want to know the anatomy and physiology of each stage of processing in the early visual system and the cortex, and the computations – information losses, retentions, and recodings – that take place at each anatomical locus within this multistage information processing system.

1.2.4 Type 3 – From physiology to perception

But what about Type 3 mappings? From a vision scientist’s perspective, the answers to Type 2 questions are made more interesting because the visual anatomy and physiology form the physical substrate of visual perception – these are the structures and processes that make seeing possible. But – here is a Type 3 question – which anatomical structures and which physiological computations are most critical to producing any particular system property of perception? And how, exactly, do they produce it?

In order to give an example of exploring a Type 3 question in depth, we now ask: What limits grating acuity? At what stage or stages of visual processing is the spatial resolution limit, revealed by the system property of grating acuity, imposed on the incoming sensory signal? At which anatomical locus is the information lost? From Figure 1.3 and Figure 1.4 you can already intuit that the limit could be imposed at any of several stages.

In this section, we will raise four specific possibilities. Each is an example of a Type 3 mapping. We emphasize that the treatment of these possibilities is qualitative and intuitive at this stage – we just want you to end up believing that the answer is not obvious, and there are several very different and plausible candidate explanations. In later chapters, we will take up each of these possibilities in much more detail. In the meantime, just imagine you are looking out at Figure 1.2 through the neurons in your visual system, and trying to use the incoming spatial pattern of light or neural activity to tell the difference between a grating and a homogeneous field. Which level or levels of the processing system is the one that limits your spatial resolution, and what features of processing impose that limit?

1. The optics of the eye?

The black and white gratings in Figure 1.2 are patterns of light that exist in the physical world. When you look at a grating, the optics of the eye make an image of that grating on your retina. The process is straightforward, as shown in Figure 1.5A. Rays of light coming from a particular point on the grating leave that point and travel in straight lines in all directions. An image is formed because the optical system captures a subset of those rays, and (ideally) bends each ray just enough so that all of the rays that start at a point on the object are reunited at a point in the image. Neighboring points on the object are represented at neighboring points in the image, with the result that an image of the grating is formed at the back of the eye.

But real optical systems are not perfect. As shown in Figure 1.5B, in reality the rays from a single point on the object do not converge perfectly to a single point in the image. They are slightly spread out in the image, forming a small irregular blob. Technically, the optical image of a point of light is called a point spread function because it describes how much the rays from a single point in the object are spread over the image. Rays from neighboring points form neighboring point spread...
1.2. THREE KINDS OF QUESTIONS, THREE KINDS OF MAPPING RULES

Figure 1.5: Idealized and realistic optical point spread functions. A. A point source and an idealized point image, drawn on the premise that all rays that leave the point source and enter the eye are perfectly bent by the optics, so as to end up at a single point on the retina. B. A more realistic, imperfect image – an extended “blob” of light – caused by optical imperfections in the eye. Some of the rays originating from the point source are bent too much or too little, so that they arrive near but not exactly on the idealized image of the point source. The distribution of light in the image of a point source is called a point spread function.
functions, and these imperfect blobs of light will soften the boundaries of the stripes of the grating in the retinal image.

Intuitively, what consequences would such optical imperfections have for grating acuity? A coarse grating will be represented faithfully, still recognizable and resolvable, but with slightly fuzzy edges. But we can intut that when the stripes in the image of the grating are about equal to the width of the point spread function, the blobs from neighboring stripes will begin to overlap, so the stripes will become less resolvable. As the stripes become even finer, the overlapping point spread functions could eventually produce a homogeneous wash of light, and we wouldn’t be able to tell the striped field from the homogeneous field in Figure 1.2.

Now, remember the argument you already know: bumblebees can fly. If you can resolve grating D (say) in Figure 1.2, you know immediately that the optics of your eye must be of at least sufficient quality to make a perceptible image of grating D on your retina. However, if you can’t resolve grating E (say), you only know that information from grating E is lost somewhere within your visual system. The optical imperfections argument suggests intuitively that the optics could be the level that imposes the limit on your grating acuity, and prevents you from resolving grating E. To find out, we’d have to find a way to measure the actual quality of the optics of the human eye, and develop a quantitative theory of the effects of optical quality on vision. We will return to this task in Chapter 4 and 5.

2. Photoreceptor spacing?

Within the eye, at the back of the retina, lies a layer of tightly packed, highly specialized neurons called photoreceptors. The retinal image falls on the matrix of photoreceptors, and each photoreceptor captures the light from its particular region of the two-dimensional retinal image. However, the photoreceptor sums the light it catches over its whole extended region, and doesn’t keep track of where each bit of light came from within that region. So the continuous retinal image is sampled piece-by-piece; that is, the photoreceptors perform a discrete sampling of the retinal image.

What consequences could discrete sampling have for grating acuity? As shown in Figure 1.6, each of the stripes in the optical image of a coarse grating covers many photoreceptors, so coarse gratings will yield variations of outputs across the matrix (or mosaic) of photoreceptors as a whole. By analyzing the spatial pattern of those outputs, we could readily tell that the input differed from that created by a homogeneous field. But a grating so fine that it puts more than one stripe on each photoreceptor is in danger of destruction, because it may yield no regular variation of output across the matrix of photoreceptors. In other words, a fine enough grating may produce the same spatial pattern of photoreceptor signals as does the homogeneous field of light, and thereby not be discriminable from it. At this point we might guess that the finest grating that gets through a discrete sampling matrix will have stripes just wide enough to put one stripe on each photoreceptor.

Again, bumblebees can fly. We already know that the limit imposed by discrete sampling at the photoreceptor mosaic can’t be worse than the behaviorally measured acuity. But if the optics don’t impose the resolution limit, the receptor matrix might. To find out, we’d need to know the sizes of the photoreceptors and their spacing, and make a quantitative model of the effects of discrete sampling. We return to this task in Chapter 5.
1.2. THREE KINDS OF QUESTIONS, THREE KINDS OF MAPPING RULES

Figure 1.6: Discrete sampling by the photoreceptor mosaic. A shows a schematic of the photoreceptor mosaic, viewed face on. Each circle is a photoreceptor. Each photoreceptor catches light from a small but spatially extended region of the retinal image, and sums the light over this region. B, C, D and E show images of four different square wave gratings, from coarse to fine. Intuitively, it seems likely that the grating in B will make a signal that varies systematically across the matrix of photoreceptors, but the finer gratings, especially the very fine grating schematized in E, might make only a homogeneous signal across the matrix. If so, the grating represented in E would not be discriminable from the blank field represented in F, and information about the spatial structure of the grating would be lost.
3. Neural convergence within the retina?

The concept of neural convergence has already been illustrated in Figure 1.4A: there are many more photoreceptors than ganglion cells. In fact, across the retina as a whole there are about 100 million photoreceptors but only about 1 million ganglion cells. That is, on average, over the retina as a whole, there is a 100:1 spatial convergence of neural signals. Intuitively, it is easy to imagine that unless special provisions are made, spatial resolution could be compromised here.

Bumblebees can fly. Even despite this average 100:1 convergence of photoreceptors onto ganglion cells, you already know that the ganglion cell layer, like all of the other layers of the visual system, must pass on information allowing us to resolve the finest gratings we do resolve. But if the optics and the photoreceptor spacing don’t limit grating acuity, maybe neural convergence in the retina does. We’ll look at this question again in Chapter 13.

4. Later levels of the visual system?

Beyond the retina, at the cortical level, there is a long series of anatomical stages and physiological recodings of visual information. In looking for the limits of grating acuity, our question about each level would be the same. How does that level manage to preserve and pass on information about the finest grating we can resolve? How is information about the grating carried (or coded) at this level? And if the resolution limit is not imposed before this level, might it be imposed here, and if so, by what computational process? The second half of the book deals with these later levels of the visual system. [Before you go on, why not lay a bet as to which level imposes the resolution limit, and give the best justification you can at this stage for your answer.]

1.2.5 Causal stories and locus questions

We now want to introduce two more terms that DT finds useful: *causal stories* and *locus questions*. *Causal stories* are proposed explanations of perceptual events on the basis of neural events. In other words, they are theories linking vision (perception) and the visual system (physiology). For example, the attribution of the grating acuity limit to the optics of the eye, or to the properties of the retinal mosaic, or to a combination of both, are all causal stories. Causal stories can be speculative, or they can be argued on the basis of quantitative theory.

*Locus questions* are a particular kind of causal story: Where within the neural information processing system is information lost, or importantly recoded, in such a way as to bring about the correspondence between physical stimuli and perception? The four options for the locus of information loss in grating acuity represent four possible answers to a locus question. The usefulness of these concepts will become more meaningful through examples encountered throughout the book.

1.3 Design questions

Finally, a fourth type of question – design questions – is worthy of mention. Design questions are *why* questions. These questions are concerned with why human vision and the human visual system take the form they take. For example, why is our grating acuity as good as it is, and why is it not better? Design constraints are imposed from many sources: the laws of physics, the physiological properties of neurons, and the effectiveness of various visual coding schemes for various purposes. In addition, the design of the visual system is shaped by the competing evolutionary pressures
that combined to shape the organism as a whole. Many competing pressures have acted upon the
design of the visual system, and the current features of the visual system are doubtless historical
compromises among these pressures. The answers to design questions are usually speculative, but
often instructive and interesting as well.

The fundamental Design question about grating acuity is: What factor or combination of factors
necessitates that visual resolution have the limits that it has? Is it that the optics can’t be any
better? Or can’t the photoreceptors be any smaller? Or can’t there be any more ganglion cells?
Or that is there a constraint imposed on some later level of the system? Or might it be that no one
level is to blame for the spatial resolution limit, but rather that the limit is imposed by conflicting
design necessities, and many levels of the system conspire to impose this limit in a more complex
way? What would have to be changed in order to improve our acuity by a factor of two, and what
would be the cost? [Again, before you go on, write down your guesses about the answer to these
Design questions.]

1.4 The mind/brain problem

We now return to Type Y mappings – mappings between physiological and perceptual entities.
To begin, we must turn briefly to the philosophy of mind. Vision science can be particularly
perplexing from a philosophical perspective, because it seems as though with Type Y questions,
vision scientists hope to explain mental events (visual perception) on the basis of physiological
events (neural activity). This hope brings us to close encounters with the mind/body or mind/brain
problem.

For centuries philosophers have argued about the nature of the relationship between mind and
brain. In particular they have argued about whether mind and brain are a single physical entity
(a position called materialism), a single mental entity (a position called idealism), or two separate
entities (a position called dualism); and if two entities, whether one of the two holds a causal
priority over the other. Many variants of each of these positions have been formulated, and the
debate continues to fascinate philosophical audiences across the centuries. For recent treatments,
see Chalmers (1996) and Metzinger (2000).

Most vision scientists are probably most comfortable with a materialist perspective. That is,
most of us probably believe, implicitly if not explicitly, that between mind and brain, the brain
is the primary causal agent. Moreover, perceptual events become less mysterious when they are
viewed simply as high-level properties of the brain. To support this view, an analogy can be made
between the properties of chemicals and chemical compounds, and the properties of brains and
conscious states. Just as water can be viewed as a high-level property of hydrogen and oxygen, so a
conscious perception can be viewed as a high-level property of a complex neural network. [Learning
this argument scratched a huge and persistent itch for materialist DT. Does it do the same for other
vision scientists, such as you?]

Taking the argument further, some philosophers make use of the concept of emergent properties.
An emergent property of X can be defined as a high-level property of X that cannot be predicted,
either in practice or in principle, from the characteristics of the lower-level elements from which X
is made. Given this definition, some philosophers argue that consciousness is an emergent property
of complex neural networks. But the concept of emergence is itself controversial, and many vision
scientists and philosophers would sharply disagree with its usefulness in the mind/brain debate.
1.4.1 Finessing the mind/brain problem: Mapping rules

Given the inevitable continuation of these debates, it seems to DT that rather than developing a science that depends upon a single view of the mind/brain problem, vision science would be wise to finesse it. That is, we should try to find a formulation of the questions of vision science that will be robust, and survive across many or all of the different philosophical stances on the mind/brain problem.

In 1870, Ewald Hering (of whom you will hear more later), lecturing at the Imperial Academy of Sciences in Vienna, laid out the classic finesse:

“If then, the student of neurophysiology takes his stand between the physicist and the psychologist, and if the first of these rightly makes the unbroken causative continuity of all material processes an axiom of his system of investigation, the prudent psychologist, on the other hand, will investigate the laws of conscious life according to the inductive method, and will hence, as much as the physicist, make the existence of fixed laws his initial assumption… it only remains for him to make one more assumption, viz., that this mutual interdependence between the mental and the material is itself also dependent on law, and he has discovered the bond by which the science of matter and the science of consciousness are united into a single whole...”

This, then, by no means implies that the two variables above mentioned – matter and consciousness – stand in the relation of cause and effect...to one another. For on this subject we know nothing. The materialist regards consciousness as a product or result of matter, while the idealist holds matter to be the result of consciousness, and a third maintains that mind and matter are identical; with all this the physiologist, as such, has nothing whatever to do; his sole concern is with the fact that matter and consciousness are functions one of the other.”

In other words, Hering argues that we can set aside the philosophical problem, and get on with finding the lawful relationships (or mapping rules) that he assumes to hold between neural and perceptual states. We will adopt this perspective for the purposes of this book. [Does Hering’s declaration provide a satisfactory finesse of the mind/brain problem? How would vision science be different if different vision scientists took different stances in regard to the mind/brain problem?]

1.5 Linking propositions

We now turn to another of the major philosophical themes of this book: the topic of linking propositions. Let’s take the next logical step beyond Hering’s assertion that lawful relationships – mapping rules – exist between visual perception and visual neurophysiology. Can anything more be said about the properties of these mapping rules? The question isn’t just, do neural states map to perceptual states? It’s which neural states map to which perceptual states?

1.5.1 Mueller’s axioms of psychophysical correspondence

Interestingly, an elaborate set of mapping rules was explicitly formulated right at the beginning of the discipline of psychophysics. The 19th century scientists who founded the discipline were motivated not just by an interest in sensations and perceptions, but also by a desire to use perceptual observations as a tool for drawing inferences about the workings of the brain. They argued that perceptual (mental) events and brain (material) events were of two different kinds, described by language from two different realms of discourse. Therefore, if conclusions about brain events were
to be drawn from facts about perceptual events, some kind of special linking statements would be needed. Their attempts to specify the necessary arguments were concisely formulated by G.E. Mueller in 1896, and are known as *Mueller’s axioms of psychophysical correspondence*.

Mueller’s first three axioms have been translated as follows (Boring, 1942, p. 89):

“1. *The ground of every state of consciousness is a material process, a psychophysical process so-called, to whose occurrence the presence of the conscious state is joined.*

2. To an equality, similarity, or difference in the constitution of sensations...there corresponds an equality, similarity, or difference in the constitution of the psychophysical process, and conversely. Moreover, to a greater or lesser similarity of sensations, there also corresponds respectively a greater or lesser similarity of the psychophysical process, and conversely.

3. If the changes through which a sensation passes have the same direction, or the differences which exist between series of sensations are of like direction, then the changes through which the psychophysical process passes, or the differences of the given psychophysical process, have like direction. Moreover, if a sensation is variable in n directions, then the psychophysical process lying at the basis of it must also be variable in n directions, and conversely.”

Mueller’s axioms have several important properties. First, notice that Mueller called these linking statements *axioms* — statements that could not be proved, but that had to be assumed to be true if the discipline was to be pursued. With these axioms in place, one could use perceptual facts to deduce some aspects of the workings of the brain. Today we would be more likely to call these statements premises or assumptions. DT argues that they are a special class of assumptions, and they enter into all claims about physiological/perceptual mappings in vision science. Moreover, they have a huge impact on the science we choose to do, and they govern the kinds of arguments that we entertain.

Second, notice that the first axiom differs fundamentally from the second and third. Mueller’s first axiom states the very general premise that all perceptual processes arise from material processes, with the material process taking causal priority. However, it says nothing further about the forms these neural/perceptual correspondences might take. So far, the possibility is left open that any material process, or brain state, could give rise to any mental process, or perceptual state. In the mapping of physiological states to perceptual states, chaos could reign. Remember, however, that Hering earlier rejected this option, preferring to assume that “this mutual interdependence between the mental and the material is itself also dependent on law”. Since the rest of the universe is lawful, it makes sense to DT to assume that mappings between neural and perceptual states are lawful too.

The second and third axioms, in contrast, are specific lawful relationships that might be assumed to hold between perceptual and neural states. Mueller recommends the assumptions that identical perceptual states imply identical neural states, and vice versa; similar perceptual states imply similar neural states, and vice versa; and so on through a long list.

Statements like these are sometimes called *isomorphisms* — (assumed) *similarities of form* — between neural and perceptual states. More specifically, DT calls the mapping rules in Mueller’s second and third axioms *relational isomorphisms*. Notice that in a relational isomorphism, percep-

---

2 Notice that as used by Mueller (and his translator) the term *psychophysical process* was used to mean a special variant of a physiological process — a physiological process to which a conscious state is joined. But in more modern writings, the term *psychophysics* always refers solely to methods for describing and quantifying sensations and perceptions (Chapters 2 and 3). Other terms, such as *neural correlate of consciousness*, are used to refer to Mueller’s so-called psychophysical process.
tual states are compared to perceptual states, neural states to neural states; and the isomorphism is that the same *relationship* that holds between perceptual states is assumed to hold between neural states (cf. Coombs, Dawes, and Tversky, 1970). The propositions are that identical perceptual states imply identical neural states, and vice versa; similar perceptual states imply similar neural states, and vice versa; and so on.

The discussion of the possible limits on grating acuity (above) illustrates the use of relational isomorphisms. If you review each of the four cases, you will find in each case the assumption that due to a particular spatial processing imperfection within the visual system, as the grating gets finer and finer, the spatial distributions of signals produced by the grating and the homogeneous field become more and more similar, and eventually become *identical*. The linking proposition that enters each argument is that an identity of neural processes creates an identity of perceptions; and vice versa, the identity of the two perceptions implies that the neural identity has been reached.

In general, relational isomorphisms are not logical necessities. Two identical perceptions could in principle arise from different brain states if the mappings from brain states to perceptual states were chaotic, or if they were many-to-one instead of one-to-one (Teller and Pugh, 1993). Other rules of relational isomorphism, such as those arising from similarity, are also easily challenged. The relational isomorphisms that enter into our beliefs about physiological/perceptual mappings are premises, not logical necessities. But they are certainly convenient!

### 1.5.2 Updating Mueller’s first axiom: The Universal linking proposition

There has been surprisingly little work on the axioms of mental/material correspondence since 1896. Brindley (1960) treated the topic briefly, proposing the name *linking hypotheses* as a name for these rules of correspondence. Since they are rarely actually hypotheses, DT (Teller, 1980) suggested the name *linking propositions*. DT defines a linking proposition as a *claim that a particular mapping occurs, or a particular mapping principle applies, between neural and perceptual states*. She argues that linking propositions lie at the heart of vision science. Any causal story – any attempt to explain perceptual events on the basis of neural events, or vice versa – will necessarily include a linking proposition. Moreover, linking propositions are often implicit rather than explicit, and part of the fun of vision science is ferreting them out and examining them.

In thinking again about Mueller’s first axiom, it makes sense to DT to reformulate it for modern times. The first axiom can be called the **Universal linking proposition**: all perceptual states and processes are implemented in neural states and processes. Most vision scientists would doubtless endorse this premise, because the alternative is to argue that perceptual states can exist without being accompanied by neural states. Notice, however, that the exact words that one uses to phrase the Universal linking proposition would vary with one’s position on the mind/brain problem. Shall we say, perceptual states and processes are *implemented* in neural states and processes; or *arise* from them, or are *enabled* by them, or *emerge* from them, or that neural states and processes *cause* perceptual states and processes; or vice versa? [As you begin to read the vision literature, watch for these variations of meaning.]

### 1.5.3 Linking propositions and relational isomorphisms

Mueller’s remaining axioms can be called the *Propositions of Relational Isomorphism*. The premise is that particular relational isomorphisms exist between perceptual states and their neural imple-
1.5. LINKING PROPOSITIONS

Each variations on the relational isomorphisms can be considered a particular linking propositions.

One might ask how linking propositions are related to causal stories. Causal stories are specific whereas linking propositions are abstract. For example, as discussed in the next chapter, the identity proposition is: identical neural states imply identical perceptual states. An example of a causal story using this linking proposition is: identical retinal images of a grating imply identical grating acuity. Linking propositions are templates that can be used to build a great variety of causal stories.

Beyond Mueller’s axioms, there are many other linking propositions. We briefly articulate two more – analogies and computational linking propositions – in the next two sections, before returning to a more general discussion.

1.5.4 Analogies and nothing mucks it up

The term isomorphism has also been used in a second way in the context of linking propositions (cf. Pessoa, Thompson, and Noe, 1998). DT calls this second category of isomorphisms analogies. These are isomorphisms that bridge between perceptual and physiological domains, by making an analogy between some aspect(s) of perceptual and neural states. For example, think about your perception of a set of broad black and white stripes, such as those in Figure 1.2A. In speculating on the neural state that underlies this perceptual state, you might assume (or show) that there is a region of the retina across which the firing rates of neurons take on a similar pattern – a set of neurons firing more slowly, say, to provide the neural correlate of a black bar, and a similar set of neurons, displaced by the equivalent of one bar width, firing more rapidly to provide the neural correlate of a white bar. The plots of whiteness/blackness against space, and firing rate against space, would look very similar. In other words, there is a visual similarity, or analogy, between the perceptually and neurally defined patterns.

An interesting feature of analogies as linking propositions is that they often enter vision science as nothing more than two pictures that look alike. But typically, many pieces need to be added to the argument to make a compelling theory or explanation of the perception on the basis of the neural activity. In fact, DT argues that statements that she calls “Nothing Mucks it Up” provisos must be involved, implicitly if not explicitly. These assumptions are of the form that nothing within the visual system, between the neural pattern and the perception, interferes with the control of the particular neural pattern over the perception. In general, the earlier in the visual system the neurons on which the analogy is based occur, the more complicated and tenuous the required “Nothing Mucks it Up” provisos would seem likely to be.

1.5.5 Computational linking propositions

Another interesting set of linking premises might be called computational linking propositions. That is, what computations would you be willing to assume can take place within the mapping between neural and perceptual states? Must this mapping always be simple or 1:1, or might computations be parts of the mapping rules? And, how fancy are the complications we will allow?

Here’s an example. Gustav Fechner, one of the very earliest psychophysicists, was struck with the idea that it might be possible to discover a mathematical formula that would specify the mapping of physical intensity to perception. Applying Fechner’s argument to the case of brightness, suppose that brightness grew, not linearly with physical intensity, but with the logarithm of physical
intensity. (This is approximately true in some cases, but not in others. But suppose it were true.) The question is, must the logarithmic transformation take place within the physiological system, so that the mapping from neural state to perceptual state is always linear; or might the growth of neural signals with intensity be linear throughout the visual system, and the logarithmic transformation come about in the mapping from neural to perceptual states? Vision scientists might well differ in their premises here.

1.5.6 From the sublime to the ridiculous

To DT, an interesting property of linking propositions is that they are often implicit. But once a linking proposition is made explicit, there is often a surprisingly good consensus among most vision scientists about its acceptability as a premise. At the one extreme lie some relational linking propositions, like Mueller’s axiom of identity. Most vision scientists would probably regard this proposition as clearly true perhaps even analytically true or tautological (cf. Brindley, 1960). We will return to elaborate it further in Chapter 2.

At the other extreme, there are some candidate linking propositions that we would doubtless all reject. For example, we would probably be uncomfortable with the assumption that the neural code for seeing a three-dimensional object must be (literally) a neural circuit with the same three-dimensional shape as the object, somewhere within the brain. We would doubtless deny that a neuron that signals redness would have to be literally red (but then, what would a redness neuron have to be like?). And we would think it silly to argue that when we perceive a dance we must do so with dancing neurons, as shown fancifully in Figure 1.7.

These examples are chosen because they define the ends of a continuum from high acceptability to silliness. But the credibilities of other kinds of linking propositions, such as analogies and computational linking propositions, fall between the two extremes, and are worth some thought.
The moral of the story is this. Whenever a causal story is proposed, a linking proposition lies within it. Is it a linking proposition that most vision scientists would readily accept, or readily discard? Would there be a consensus, or might different vision scientists differ in the kinds of linking propositions they are willing to incorporate as premises into their causal stories?

1.6 A more perceptual perspective

Historically, vision has been studied from two different perspectives. The first is a sensory perspective, in which vision scientists tend to emphasize the simpler aspects of vision, and (in general) attempt to account for them on the basis of the codings and recodings of information that take place within the early processing stages of the visual system. The second is a perceptual perspective, in which we tend to emphasize the more complex aspects of perception, and (in general) attempt to account for them on the basis of the more complex and higher level aspects of visual processing.

So far in this chapter we have been taking a classically sensory approach to vision. But let’s switch briefly to a more perceptual approach. From the perceptual perspective the interesting parts of vision science lie not in the sensory details such as grating acuity, but rather in the complex system properties of perception. And the most fundamental phenomena are not illustrated by drawings on the pages of a book, but by looking at the world around you. [Look at the world around you!]

As you look around, you see a scene that contains three-dimensional objects of particular sizes, shapes and colors in particular three-dimensional locations. These objects may move, but their essential characteristics of size, shape and color tend to remain constant across viewing conditions (that is, we as perceivers have good size, shape, and color constancy). You are often able to recognize objects across time and across contexts. The perception scientist wants to describe and quantify these more complex system properties, and to understand the properties of the (presumably high level) neural processes that make these high level features of perception possible. We will argue later, for example, that incoming sensory information must be combined with stored information and processed with complex computational algorithms before it can provide the neural basis of high level visual perception.

Until relatively recently, physiological study of high level visual processing was still in its infancy. Since little relevant information was available, many perception scientists were not much interested in knowing the details of information processing within the visual system. In fact, some have denied the value of understanding anatomy and physiology for understanding perception.

But this situation is changing. By now, a great deal of information about the anatomy and physiology of cortical processing – Type 2 mappings – is well established, as we will show. Moreover, in the past few years a number of vision scientists have undertaken studies of single neurons, searching for the neural basis of particular perceptual processes and phenomena. As this knowledge comes in, perception scientists are beginning to invent causal stories about perceptual events and processes based on neural events and processes. Moreover, recent, more global analyses arising from neural imaging techniques such as functional magnetic resonance imaging (fMRI), have also drawn the interest of perception scientists toward neuroanatomy and neurophysiology, and toward causal stories relating perception to neurophysiological activity.
CHAPTER 1. THE DOMAIN OF VISUAL SCIENCE

Figure 1.8: Illusory contours. The continuous contours at the left and the illusory contours at the right both give rise to similar perceptions of a triangle. Why?

1.6.1 Perception and linking propositions

Let us develop a couple of examples of linking propositions that might be involved in causal stories in high level perception. First, take the case of the triangles illustrated in Figure 1.8. This figure illustrates the phenomenon of illusory contours. The perception of a set of three borders can arise from at least two very different stimuli: three physical borders formed from solid lines; or a set of “pacmen” at three corner locations. And in both cases, similar triangles can be perceived. Suppose we were to decide to search for the neural cause of this odd perceptual similarity. Where would we start, and why?

One linking proposition we could adopt would be a similarity (or identity) proposition: that within the visual system, there will exist neural elements that respond similarly (or maybe identically) to these two stimuli – the lines and the pacmen. Such a speculation will lead us to examine individual neurons at various levels of the visual system, and test them with both kinds of stimuli. The goal would be to try to find such neurons, and to locate the earliest level at which they occur. But the whole enterprise rests on a relational isomorphism: the premise that the two similar perceptual states – the two perceptions of a border – indicate the presence of two similar states of the same neurons arising from the two physical stimuli. [Is this an isomorphism you would endorse? Is it a logical necessity, a reasonable speculation, or just a silly argument?]

And as a complex example, we ask the same question about what DT calls the Ultimate Code: is there a form that the neural code may or must take, in order to give rise to the conscious perception of an object or a scene? Some of you have undoubtedly already adopted the premise that every time we see the same object (famously, our grandmother, as we will see), there is a neural stage at which the same individual high-level neuron or small set of neurons must be active in the same way (cf. Barlow, 1972). Others would reject this specific premise, but argue that some form of isomorphism must be involved in the mapping from neural to perceptual states, even though we cannot at present say what it will be. And yet others would argue that, much as we might prefer the universe to be otherwise, for complex perceptual phenomena there will turn out to be no consistent, definable isomorphism between perceptual and neural states.
1.7. AN INTERDISCIPLINARY FIELD

DT hastens to add that the problem of the Ultimate Code will not be solved in this book! And yet it provides a high-level example of the kinds of questions that initially attract many scientists and philosophers to the field of visual perception. It also illustrates the point that different implicit linking propositions lead to different physiological predictions and influence the choice of experiments a scientist undertakes.

1.7 An interdisciplinary field

It should be obvious by now that vision scientists cannot afford to respect the boundaries between classical scientific disciplines, much less be chauvinistic toward any of them. Instead, we first define the questions of interest, as we have done in this chapter. Then we look around to see what kinds of classical disciplines can provide us with the expertise we need to address our questions. Specialists of many different kinds are invited – we need all the help we can get! The disciplines that unite to form the field of vision science, and the expertise we need them for, include at least the following.

Psychophysics and perception:
To describe and quantify the system properties of vision.

Physics:
To describe the nature of light and the optical quality of the eye.
Photochemistry:
To describe the interaction of light with matter in the photoreceptors.

Neuroanatomy:
To describe the structure of the various parts of our visual systems.

Neurophysiology:
To describe the information processing characteristics of individual neurons and neural circuits at each stage of the visual system.

Cell biology:
To characterize the internal workings of the cells, and their mechanisms of communication.

Molecular genetics:
To describe the genetic control of the various parts of the visual system.

Optometry, ophthalmology:
To describe the disorders that can occur in vision and in the visual system, and use them to help us understand normal vision and the normal visual system. Also, to import the accumulating knowledge about vision and the visual system into medical practice, in order to help patients with vision problems.

Engineering:
To describe information processing systems; to provide conceptual and mathematical tools for describing complex systems and their properties.

Computer science:
To discover design principles and computational algorithms that might help us understand information processing within the human visual system.

Mathematics, statistics:
To provide tools for modeling the three kinds of mapping rules.

Philosophy, logic:
To provide logical analyses of our most basic scientific concepts.
Cognitive neuroscience:
To provide descriptions of the cognitive processes that affect perception, and models of the mechanisms that underlie it.

In fact, over the years specialists from all of these fields have been drawn into vision science. As a consequence, the field is conceptually very rich and sophisticated, and new ideas are always arriving from different sources. For DT, the excitement has lasted a lifetime. Everyone in the field comes from somewhere else; everyone has some areas of deep expertise, and some areas where he or she is an amateur.

1.8 Summary: The domain of vision science

In this chapter we defined vision science as the study of vision, the visual system, and the relations between the two. We argued that vision science spans three domains: physical stimuli, neurophysiological states, and perceptual states; and that vision scientists are interested in the mapping rules among these three domains.

The questions in which vision scientists are interested were illustrated by using the example of grating acuity. We defined grating acuity as a Type 1 phenomenon – a physical/perceptual mapping. We then provided a brief description of Type 2, or physical/neurophysiological mappings, mostly within the retina. Finally, we posed the Type 3 question: a locus question about neural/perceptual mapping. What stage or stages of visual processing limit grating acuity? We suggested four candidate answers. We will examine each of these possible answers in detail in subsequent chapters. The point is that these are the kinds of questions that vision scientists ask, and the ranges of answers that we find satisfying.

In pursuit of the essence of Type 3 questions, we then introduced the concept of a linking proposition: a claim that a particular mapping occurs, or a particular mapping principle applies, between neural and perceptual states. We updated Mueller’s first axiom to define the universal linking proposition – all perceptual states and processes are implemented in neural states and processes. We distinguished the universal linking proposition from the propositions of relational isomorphism, which postulate specific, relational mapping rules between sets of perceptual states and sets of neural states. We also briefly mentioned several other kinds of linking propositions, including analogies and computational propositions, to which we will return throughout this book.

Finally, DT has had three broad goals in writing this book. First, vision science is a complex interdisciplinary field, influenced by concepts from many kinds of science and by many kinds of scientists, and the beginning student is likely to have some difficulty with the parts that are the farthest from home. The first goal of this book is to provide a united, self-consistent set of tutorials, using simple examples, to make the science as a whole accessible. It is hoped that the tutorials in the various chapters of this book will slow down the moving train just enough for students with many different backgrounds to jump on.

Second, we have argued that vision science is a sophisticated discipline, spanning across physical, perceptual and physiological realms. To DT’s knowledge there is no deliberate, consistent exposition of the forms of argumentation common in vision science. The arguments and premises are often
implicit. It is hoped that making them explicit will demystify them, and thereby help students get on the train. And of all the implicit elements, linking propositions seem to be the most consistently hidden in the shadows, and therefore the most fun to bring out into the light.

A final goal of this book is to encourage the seamless integration of sensory and perceptual approaches to vision science. Perhaps some successful causal stories from early processing levels will provide forms of argumentation that will be useful in evaluating causal stories at higher processing levels.

In Chapter 2 and 3 we examine the methodological tools with which vision scientists study physical/perceptual (Type 1) mappings, and a sample of the results they have found. In Chapter 4 and 5, we examine the optics of the eye, and the marks that the optical system leaves on our perceptions. Then in Chapter 6 and 7 we examine the workings of photoreceptors, both individually and in sets, and begin to analyze the code for color vision.