

Chapter 3. Perception, Appearance, and Reality

We never ... originally and really perceive a throng of sensations, e.g., tones and noises, in the appearance of things...; rather, we hear the storm whistling in the chimney, we hear the three-engine aeroplane, we hear the Mercedes in immediate distinction from the Volkswagen. Much closer to us than any sensations are the things themselves. We hear the door slam in the house, and never hear acoustic sensations or mere sounds. Martin Heidegger (1977: 156)

Throughout the foregoing chapters, we have witnessed a battle between our commonsense claims to knowledge and the impending threat of skepticism. The principal territory for which this battle has been waged is our ordinary perceptual beliefs about the objects in our immediate, present environment. I have defended my claims to know (for instance) that there is a coffee cup before me, that it is presently raining outside, and that I have hands. Thus far, I have given a philosophical defence of a conditional conclusion: that *if I perceive that I have hands, then I know that I have hands*. Moreover, the antecedent of this conditional *is true*; I do actually perceive that I have hands. (I don't need a philosophical defence of this fact; it *just is* a fact. Nor do I need to have the higher-order knowledge that *I perceive that I have hands* for the antecedent to be true and for the consequent to follow; however I do happen to also think that I have this higher-order knowledge). It follows that I *know* that I have hands. None of the skeptical arguments that we have faced so far have been able to resist this conclusion.

Yet, we are not entirely in the clear. There is still one more challenge that needs to be faced, which comes not from epistemology, but from the metaphysics of perception. There is an old view of perception which is such that, if it were true, it would destroy the progress we've made against skepticism.

1 Direct realism and sense data

Consider first the naive picture of perception that I've been assuming so far. Hitherto, I've assumed that it is possible for me to *perceive that it is raining outside*, or *that there's a coffee cup before me*, or *that I have hands*. More generally, I've assumed that it is possible for me to perceive any external object that is currently unobstructed in my field of vision. I can perceive these objects—the rain, the coffee cup, my hands—*directly*, as it were. As long as I perceive them, I am aware of their existence; I do not need to *infer* their existence from any other perceptions of other objects. This view is commonly known as *direct*

realism. It is so-called because it holds that *external objects are real* and *we can directly perceive them*. This is the picture of perception that I had tacitly presupposed when I grappled with skepticism in previous chapters.

But what if this picture of perception is wrong? Then I will need to reconsider my anti-skeptical positions. And perhaps I will need to do so, because there have been several philosophers throughout history who have advanced an alternative account of perception. This view is known as the sense data theory. (It was traditionally endorsed by most empiricists, including Hume, Russell, and Ayer.) Our task for this chapter is to interrogate it and see if we can resist it.

The basic contention of the sense data theorist is that *external objects never* lay unobstructed to our view. We *never* perceive external objects directly, in principle. That is because, according to their view, we are only ever directly aware of *mental images* that are fleeting and mind-dependent. When I attempt to direct my gaze upon my coffee cup, what appears before my mind is not the coffee cup (the thing in itself), but rather a sensory image that exists within my mind. This sensory image may *resemble* the coffee cup (more or less), and be *caused* by the real coffee cup, but it is not the same thing as the coffee cup itself. It is a subjective entity, whereas the real coffee cup is an external object that resides outside of my mind. Philosophers throughout the ages have variously called these entities—the internal representatives of external objects—“mental images”, “appearances”, “ideas”, “impressions” and “sense data.”

What exactly are these “sense data”? Well according to these philosophers, the sense data have three defining traits. First, they are the things that we are *directly aware* of in perception. That is to say that we do not need to infer their existence from our perceptions of anything else. We see them and know of them by our perception, unmediated by any cognitive efforts. Secondly, the sense data exist entirely “within our minds”—their very existence depends on our perceiving them. When we perceive them, they exist; and the moment we look away, they fleet out of existence. Finally, it is claimed that the properties of sense data are entirely transparent to our view. The properties that they *appear to have* are the properties that they *actually have*. So for example, if it appears to me that I’m perceiving a green leaf, then I must be experiencing a sense datum that is actually green and leaf-shaped. This last property is supposed to impart the sense data with very special epistemological properties. It means that if I attend to my sense data alone, I should be able to access their properties in a way that’s entirely transparent. If I experience a sense datum that is green and leaf-shaped, then I should be able to immediately *know* that I have a green leaf-shaped sense datum, and I couldn’t be mistaken about the matter.

In the opening pages of this manuscript, I spoke vaguely of a ‘one-world picture’ and a ‘two-world picture’ of reality and our place within it. We can now see direct realism and the sense data theory as archetypes of these opposing ways of thinking. For according to the sense data theory, the world that I (directly) experience is the world of my own private sense data. When I have a perception, my cognitive interaction is with the elements of my own mind. It is

as if my mind has a little theatre inside it and I am the audience.¹ The role of my sense organs is to deliver images to the theatre screen, for me to then view. But the *only* things that I directly view are the images on the internal theatre screen (the sense data). I do not have direct access to the world outside of my mind. There may be an objective world out there, behind the screen. But if I am to ever obtain knowledge of it, I must infer this knowledge from what is displayed on the screen of my internal theatre.

For the direct realist, on the other hand, perception is not all like that. For them, if I perceive a green leaf, then my perceptual state involves a cognitive relation that I bear *directly to this leaf*. Now, what does it mean to bear a cognitive relation *directly* to an external object? Well, for one thing, it means that *if this very leaf* were not there (if it were replaced by a hologram or I were having a hallucination) then I could not be having this perceptual state. Instead, I would be in a different state: I would be perceiving the hologram, or I would be hallucinating. In other words, my perceptual state essentially involves the leaf. In general, the direct realist thinks of my perceptual states as essentially involving the objects of my external environment. I am pictured, in their view, as one being amongst others in the world, and just as I can bump into other material objects, I can also relate to them by perceiving them.

Upon hearing of this dispute between direct realism and the sense data theory, many students of philosophy jump to the conclusion that this debate is best left to the scientists who study the physiology of sense perception. They feel that there must be something wrong with trying to discover how perception works by *a priori* philosophical reflection. However, when we actually listen to what the two views are saying, it isn't obvious how anything we could learn in the lab could be used to settle this dispute, without begging crucial questions towards the other side. The scientist can give us more details about the causal pathway between an object, our sense organs, and the region of our brain that is responsible for the cognitive uptake (in terms of photons, retinal stimulation, and neural activity). But how are these details supposed to settle the question of *what is it that we directly see, without the aid of inference?*

Perhaps the student is assuming that *the thing that we directly perceive* must be the last link in the causal chain before my cognitive uptake. Since the final nodes of the causal chain take place entirely inside my head (they are retinal imaging or neural firing), the student might infer that I never perceive external objects, and so they conclude that the sense data theory is true. But despite its initial temptation, this line of thinking represents an egregious misunderstanding of the dispute. It is a misunderstanding for two reasons. First of all, it is simply not true that I can perceive the penultimate events of the sensory causal process.² The final events of the sensory process are a matter of my neurons firing in my brain, and I have never in my life seen my own neurons fire. This mistake allows me to make a clarification. The dispute between the direct

¹The imagery of the 'Cartesian theatre' comes from Daniel Dennett, as a satire of this kind of theory.

²That is, unless I'm in the rare circumstance of viewing a live image of my brain—which I have never done.

realist and the sense data theorist is properly understood as a dispute over the *content* of perception (that is, what our perceptual states are *about*). It is not a dispute over the *vehicles* by which that content is delivered to the mind. Now, as for the student's second mistake, they have naively assumed that the penultimate events of the sensory process must be something like sense data. But remember, sense data are defined as having very peculiar properties. In particular, sense data are defined so that there is no difference between *how a sense datum appears to be* and *how the sense datum really is*. The sense data are supposed to be *epistemically transparent* to us. But what purely scientific discovery could support the existence of such a thing?

I conclude that the controversy over sense data is properly understood as a philosophical one. There is no simple way for empirical science to decide it, unaided by philosophical analysis.³ Here is an analogy to crystalize this claim. Picture two astronomers who are taking turns looking through a telescope that is pointed towards Saturn. One astronomer exclaims that they can see the rings of Saturn through the telescope. The other astronomer responds, "No you surely cannot. What you *really* see is *the images of Saturn's rings* that are projected onto the telescope lens." The first astronomer rejoins, "well sure, *if* I concentrate on the lens, I can concentrate on the image. But when I look *through* the telescope, I see the rings of Saturn." Which one of them is right? And more to the point, what is the proper way to determine who is right? One thing should be obvious. Their dispute cannot be settled by calling upon the telescope technician and asking them about the fine details about the internal structure of the telescope. Their dispute is clearly philosophical.

The debate between the sense data theorist and the direct realist is similar. The first philosopher claims that we can only ever *really* see the final stages (the mental images) of our sensory processes. The opposing philosopher insists that we can *use* our perceptual instruments to *perceive* external objects. Now, we cannot simply adjudicate their dispute by calling upon the scientist to supply more details of the sensory process. Their dispute concerns the representational content of perception, not the mechanical details.

2 Perception and knowledge

Properly speaking, the controversy over direct realism and sense data is principally concerned with the *nature* of perception. The subject thus belongs to metaphysics, and not epistemology. Yet, there is a clear relevance of this topic to epistemology. Our claims to knowledge about ordinary external things may very well be at stake

If direct realism were true, then I can (directly) perceive the ordinary objects that surround me in my external environment. I can perceive that I have hands, I can perceive that there is a green leaf before me, and so on. I do not need

³I do not want to rule out the possibility that there are relevant discoveries in cognitive science. My only point is that we cannot naively assume that science can solve this dispute and philosophy is irrelevant.

to infer their existence on the basis of any other beliefs of mine. In that case, these beliefs of mine are fit to be noninferential, foundational knowledge, which are known on the basis of perception. At least, that is what I have argued in previous chapters.

But now let's see how our situation would change if the sense data theory were true. Under this way of thinking of perception, I can no longer directly perceive the external objects around me. The only objects that I immediately perceive are my own private sense data. It is likely fair to say that, on this view, our perception-based foundational knowledge will entirely consist of beliefs about our sense data. So I can know (without inference) that I am currently experiencing a green leaf-shaped mental image. But that is not the same thing as knowing that there is a real leaf that exists before me. According to the sense data theory, I cannot have non-inferential, perception-based knowledge of that. It's as if we are bound to our own subjective cage, where the inward-facing boundaries are all mind-dependent appearances. (We only ever see the images on the screen of our internal theatre.) So then, how could I know that there is a real leaf that exists before me?

If the sense data theory were true, then my only hope for attaining knowledge of the external world is to *infer it* from the traces it may supposedly leave in my sense data. We see the images displayed on our internal screen, and we use these as *clues* as to what exists behind the screen. We develop *hypotheses* and *theories* that postulate 'real' objects that exist with a measure of permanence—we conjecture that there's a real leaf that won't cease to exist the moment I look away. We take the patterns of our sense data to *confirm* such hypotheses. We *hope* that the evidence in our sense data is strong enough to support our theories about the real world. Sometimes it is claimed that the real world hypothesis is supported by *inference to the best explanation*.⁴ The thought here is that the existence of real, permanent objects (like the leaf) is the simplest and most fruitful way to explain all of the patterns that I observe in my sense data.

Whatever the initial promise of this proposal, there are several well-known difficulties with the sense data theorist's strategy to overcome skepticism. Since my aim isn't to defend the sense data theory, I will only briefly mention a few. For one, the theory predicts that we never observe the external objects. So as David Hume pointed out, we have no grounds for supposing that the external objects *resemble* our sense data in any respect—it is impossible for us to compare them and check for resemblance. Secondly, it is unclear whether the inference-to-the-best-explanation strategy can result in genuine *knowledge*. After all, this strategy never rules out the alternative explanations for our sense data, such as the hypothesis that we are brains-in-vats that are stimulated with false impressions. Lastly, it might even be doubtful that the real-world hypothesis is the *simplest* explanation for the patterns we observe in our sense data. After all, the real-world hypothesis requires a universe that is finely-tuned for life and about 14 billion years of unlikely events that culminate in our

⁴Bertrand Russell proposes this anti-skeptical strategy in the second chapter of *The Problems of Philosophy*

evolution. The 19th century physicist Ludwig Boltzmann may have suggested a simpler hypothesis: perhaps by sheer random chance, a mass of atoms in empty space happened to momentarily coalesce into a fully functioning brain intact with memories and sense data.⁵ Perhaps this ‘Boltzmann brain’ is me. Obviously, this Boltzmann brain scenario is extremely improbable. But given all of the random chance events that need to occur for humans to evolve in a fine-tuned universe, it may be even *less* improbable than the real-world hypothesis.

I am not going to pursue any of these skeptical problems for the sense data theorist any further. My point is only to impress you with the epistemic problems that are peculiar to the sense data theory. Given all of this, it might very well turn out that the direct realist has the unique advantage to vindicating our ordinary claims to knowledge. It could very well be that this knowledge is possible on direct realism and impossible on the sense data theory. If that’s our situation, then we have all the more reason to examine the arguments for sense data theory.

Some philosophers might see in this situation the grounds for a ‘transcendental’ argument for direct realism. A transcendental argument is one that is based on the preconditions for knowledge. A philosopher who would run such an argument would first argue that direct realism is a precondition for knowledge about external objects. They would then claim that we manifestly possess such knowledge, and so direct realism must be true. As for myself, I do think that these philosophers have a point. But I will not be content to rest my case on this transcendental style of reasoning. Instead, I think that we also need to critically examine the arguments for the sense data theory, and develop our own counterarguments.

3 The arguments for sense data

It is now time to face the most pressing question of this chapter: *why would anyone subscribe to the sense data theory?* By far the oldest and most influential reason stems from the phenomena of perceptual illusion.

A perceptual illusion occurs whenever we fix our gaze upon an external object, and the object appears to have properties that it does not in fact have.⁶ Let’s say that I fix my gaze upon a straight wooden stick. If I half-submerge this stick in water, it will appear as bent. Or if the wind blows it out of my hand into the distance, it will appear to shrink. Or if I stare at the stick while I use my finger to nudge my eyeball from the side, it will appear to me as if there are two sticks in my field of vision, rather than one.

⁵Boltzmann did not actually believe that he was a ‘Boltzmann brain.’ The point of the thought experiment is to compare the probability of a Boltzmann brain randomly forming with the probability of human beings evolving in a fine-tuned universe. Both of them are extremely unlikely events. But given all of the random chance events that are required for humans to evolve, the latter may even be less likely than the former.

⁶Do not confuse an *illusion* with a *hallucination*. An illusion occurs when a real object appears to have illusory properties. A hallucination occurs when no object is present for one to perceive, but a hallucinatory experience makes it appear as if there is one.

A number of philosophers have taken these considerations to be conclusive proof for the sense data view. Here is what David Hume has to say on the matter:

The [stick], which we see, seems to diminish, as we remove farther from it: but the real [stick], which exists independent of us, suffers no alteration: it was, therefore, nothing but its image, which was present to the mind. These are the obvious dictates of reason; and no man, who reflects, ever doubted, that the existences, which we consider, when we say, this house and that tree, are nothing but perceptions in the mind, and fleeting copies or representations of other existences, which remain uniform and independent. *Enquiry* XII.1 page 118.⁷

With all of Hume's confidence, you would expect to find an argument that's beyond refutable. But what is the argument exactly? From what Hume has provided us, we seem to have the following premises and inference.

1. The thing which we directly see appears to diminish.
2. But the real external object (the stick itself) does not diminish.
3. Therefore we do not directly see the external object; what we directly see is something else (its 'mental image'; a.k.a. sense data).

If this argument were valid, then we could presumably extrapolate its conclusion to all other cases of perception. Firstly, if this sort of argument worked, then the things we see in illusory cases of perception must have all of the properties of sense data. In particular, what we see must *actually have* all of the properties that it *appears to have*. Why? Because if it didn't, then we could run the argument again to show that we're *actually (directly) seeing something else*. Secondly, it would be preposterous to claim that we only ever have sense data during illusory perceptions; after all, this phenomena is pervasive. Whatever stuff we see in illusory experiences must be the same kind of stuff that we always see. So if this argument were valid, then we have established that we only ever see sense data, *full stop*.

But alas this argument *isn't* valid. The first premise says that the object of perception *appears* to diminish. The second premise says that external object *does not actually* diminish. It does not follow that our object of perception is not an external object.⁸ Why? Because it is entirely consistent for one and the same object to both *appear to have a property* while *not actually having that property*. So Hume has in no way established his conclusion.

The same thing could be said about the occasion where we dip the straight stick into the water and it appears to be bent. The sense data theorist would proffer this argument.

⁷Hume's original example is a table, but there's no harm in changing his example to be consistent with mine.

⁸The logical form of the argument is: (1) The x is apparently-F; (2) The y is not-F; (3) Therefore the x ≠ the y. But clearly this is invalid.

- 4 The object that one directly perceives appears bent.
- 5 The real stick is not bent; it is straight.
- 6 Therefore one does not directly perceive the stick.

And again, this is invalid for precisely the same reason. There is no conflict between one and the same physical stick being both *apparently* bent and *actually* straight. So for all this argument has shown, we may perceive the real stick.

Maybe we should be more charitable to the sense data theorists. Maybe we should revise their argument into something that at least meets the standard of formal validity. To do so, we would have to swap the first premise with something that concerns the actual properties of the objects of perception. In that case, their argument would go like this.

- 7 The object that one directly perceives is (*actually*) bent.
- 8 The real stick is not bent; it is straight.
- 9 Therefore one does not directly perceive the stick.

Now at least we have an argument whose conclusion follows logically from the premises. But does it fare any better as an argument against direct realism? I'm pleased to say that it doesn't. This time around the argument has a problem with its first premise: there is no reason to think that it is true. At least with the previous argument, we knew that the first premise was true; we can happily grant that we perceive an object that *appears* bent. But the new argument trades in that uncontroversial premise with a controversial one—that the object we perceive *is* bent. But why think that we perceive any object that *really is* bent, when we are looking at a straight stick that appears bent?

The direct realist certainly will not grant this premise. On their assessment of the situation, what we perceive is a straight stick that appears to be bent. But that doesn't imply that we see anything that actually is bent. The only apparent reason to suppose that we see a bent object is if one assumes that we see an image of the stick rather than the stick itself. But if that's right, then the revised version of his argument is blatantly question-begging.

For their next stratagem, the sense data theorist might try to change the topic to a different kind of illusion. They'll ask us to fixate our gaze upon the stick, and nudge our right eyeball from the side, making it appear as if there are two sticks rather than one. Following this experiment, they will put forward this argument.

- 10 There are two objects of your perception.
- 11 But there is only one physical stick.
- 12 So at least one of the entities that you perceive is not an external object; it's a mental image.

13 If at least one of the entities you perceive is a mental image, then both are.

14 Therefore the two entities you perceive are mental images.

At first this argument seems a bit more compelling than the last. When I nudge my eye and cause myself to see double, there is a loose way of speaking in which I could say that “I’m seeing two things.” But would it be *literally* true that I’m seeing two things? Or to rephrase the question, are there really two things that I am perceiving?

Obviously the sense data theorist would think so, *because they think that I’m perceiving two sense datums*. But what should we say if we’re not already convinced of sense data theory? Is it possible to resist the first premise?

I do believe that it is. The direct realist can deny that the first premise is true without sacrificing any plausibility of their position. Instead of admitting that we perceive two objects, they can insist that we perceive *only one stick*. The illusion consists *not* in seeing *two distinct objects*, but in seeing *one object* (the stick) *that appears to be in two distinct spatial locations*. As long as this is a plausible way to describe your perceptual state, the direct realist should not fear this argument.

So we see, once again, that this new version of the argument from illusion begs the question against the direct realist. The sense data theorist will accept the first premise, but *only because they already accept sense data theory*. Whereas the direct realist will deny the first premise and offer their alternative way to describe the illusion. So how should we respond to this situation? Should we become more partial towards the sense data theory? I don’t think so. If we have not yet found any other good reason to believe in the sense data theory, this particular argument should not convince us.

So much for the argument from illusion. At this point, you may be wondering whether the sense data view has any other arguments in its favour. And in fact it does have another argument to offer, so we are not yet out in the clear. This argument is a bit more recent; it stems from the phenomena of *hallucination*. Before we get to the argument, let’s pause to get clear on the distinction between *illusion* and *hallucination*. As we’ve been using these terms, an *illusion* occurs when we fix our gaze upon a *real* object, and then from our perspective, the object appears to have properties which it does not in fact have. A *hallucination*, on the other hand, occurs when there is no real object present before our eyes, but we nonetheless undergo a hallucinatory experience that makes it seem to us as if there is one. (This can happen in dreams or when we take hallucinogenic drugs.)

Now, according to this new argument, it is entirely possible for a subject to have hallucinatory experiences that perfectly resemble the experiences that I am having now. For the sake of illustration, we can imagine an unembodied brain that is kept alive in a vat full of life-sustaining fluid. Picture all of its nerve endings as hooked up to a supercomputer, and imagine that the computer feeds these nerves with electrical impulses that resemble sensory stimulation. If

the technology is advanced enough, it might be possible to serve the brain with hallucinatory experiences that appear to the brain just like how my own experiences appear to me.

We have already seen how these thought experiments are dear to epistemology, but this time our focus is on the nature of perception. In this context, the sense data theorist could attempt to argue for their view by appealing to the brain-in-the-vat's perceptual states. The brain in the vat, they will point out, is not perceiving any physical objects. That is because all of their experiences are hallucinatory. But still, they will insist, the BIV must be perceiving *something*, since it is having some sort of experiences. But since it isn't perceiving anything physical, it must be perceiving something internal, *namely sense data*. Lastly, the sense data theorist will argue that the BIV and myself must be *perceiving the same things*, since our experiences are qualitatively indistinguishable. If all of this is right, then they can conclude that I am perceiving sense data. Their argument has followed all of these steps.

- 15 My hallucinating BIV-counterpart is perceiving something.
- 16 But if my BIV-counterpart is perceiving anything, it is sense data.
- 17 I perceive the same things as my BIV-counterpart perceives.
- 18 Therefore I am perceiving sense data.

Now we have, laid bare before us, the direct realist's final challenger. Should this be the test that finally defeats their view?

For one last time, the direct realist will deny one of their opponent's premises, and do so without forfeiting the plausibility of their view. To see how, notice how the sense data theorist has assumed, in their first premise, that a hallucinatory state still involves *some* object of perception. On their conception, the hallucinating subject is still perceiving *something*. But must that be true? I think that we can plausibly challenge this way of thinking about hallucination. The direct realist will say, instead, that the hallucinating subject is not *really* perceiving anything at all. There are no objects that my BIV-counterpart can perceive. It might *seem* to my BIV-counterpart that it perceives something—e.g. it will seem to them that they perceive that they have hands, or that there is a coffee cup before them—but they don't *actually* perceive anything. That's just what it is to have a hallucinatory experience: it seems to them that they see things that they do not in fact see.

I do not believe that there is anything exceedingly implausible about this response from the direct realist. It is quite natural to say that a person who hallucinates is not really perceiving anything. But this response can sound *very* implausible if it gets confused for something else. So just to be clear: the direct realist is *not* claiming that my BIV counterpart isn't having any qualitative experiences as the result of their sensory stimulation. Surely the BIV, or any other victim of hallucination, may still enjoy the full range of *qualitative experiences* and *sensations*. The direct realist isn't denying *that*.

Rather, they are denying that there are any *objects* that are perceived by the BIV and other victims of hallucination—not sense data or anything else. To see the point of their objection, it might be helpful to reflect on the fact that *perception* is not just any experiential state. Perception is special, because it has representational content. To have a genuinely *perceptual* state, one must have an experience *that also represents some object as being some way*. One can have experiences that don't represent anything, but such experiences would not be perception.

Bearing this mind, the direct realist's understanding of hallucination may even be *more* plausible than the sense data theorist's. Consider again my BIV-counterpart that has an hallucinatory experience of having hands. The direct realist is content to say that the BIV undergoes these hallucinatory experiences, but it doesn't have any relevant perceptual state that represents a *real* object. If it *were* to represent any real object, you would think that it would represent the BIV's hands. But since the BIV doesn't have hands, there cannot be any such state.⁹ However, the sense data theorist (on the other hand) will insist that the BIV does have a perceptual state that represents an actually-existing object. But the BIV's state cannot be representing the objects that it would appear to be representing—it isn't representing the BIV's hands (because those don't exist). So the sense data theorist is forced to say that the BIV's perceptual state represents something else: a surrogate object of perception—a fleeting sense datum. But rather than invoke this ad hoc substitute object of perception, isn't it simpler to accept that there is no object of perception during hallucination?

4 The proof that there is no sense data

So far we have not seen any conclusive reasons for the sense data theory—ones that would persuade a direct realist. Is that enough to reject it in favour of direct realism? It might be. Perhaps at this stage a direct realist could claim that their view deserves a sort of 'default' status, since it meshes with common sense. In that case, if there are no good arguments against direct realism, then we should accept it as true by default. Or perhaps the direct realist could appeal to the transcendental argument that I alluded to earlier. They could argue that knowledge about the external world is impossible without direct realism, and since we clearly have some knowledge of this kind, we should conclude that direct realism is true. In my opinion, both of these argument strategies have some merit to them. But I will not rest my case on either of them here. For I believe that we can do even better. I believe that we can prove that the sense data theory is false.

⁹It is a logical truth about perception that one can only perceive something if it exists. If pink elephants don't exist, but you appear to see a pink elephant during an LSD trip, we would not say that you *actually* perceived a pink elephant. Instead we would say that it *appeared to you* as if there were a pink elephant. You *thought* you perceived a pink elephant; but you didn't. Similarly, if S perceives that P, then it follows that P is true. If I *really* perceive that I have hands, then I have hands. For if I did not have hands, then I couldn't *really* perceive that I do.

To see this proof, remember how sense data are defined. They are supposed to have three defining properties: (I) they are the objects that we are directly aware of in perception, (II) their existence depends on our perception, and (III) their properties are supposed to be entirely transparent to us—the properties that they *appear to have* are the properties that they *actually have*. I think that I can show that no object of perception can meet condition (III).

Here is a thought experiment. (This thought experiment and subsequent argument is modelled after Timothy Williamson’s *anti-luminosity argument*.¹⁰) Imagine that you spend your entire morning staring at a computer screen. At $t_1 = 8:00$ AM the screen displays a solid shade of green. You continue to stare, and from moment to moment you do not notice any difference in the computer screen’s colour. However, it does gradually change, but at a pace that is too slow for any human to recognize. By $t_n = 12:00$ PM the screen is distinctly blue.

Now suppose that the sense data theory is true. Since the screen *appears* to you to be green at t_1 , you must have a green coloured sense datum as the object of your perception. By the time t_n comes around, you no longer have a green coloured sense datum; you have a blue coloured one instead.

But here’s the catch. For any two adjacent seconds, t_i and t_{i+1} , the sense datum you perceive at t_{i+1} *will appear to you to have exactly the same colour as the sense datum from t_i* . But according to the sense data theory, if the t_{i+1} sense datum *appears* to have the same colour as the t_i sense datum, then they *do* have the same colour. It follows that the t_2 sense datum must have precisely the same shade of green as the t_1 sense datum; and the t_3 sense datum must have precisely the same shade of green as the t_2 sense datum; and so on and so forth. So we now have, at our disposal, all of the premises we need for a *reductio* argument against the sense data view. Let $t_0 = 8:00$ AM, $t_1 = t_0 + 1$ second, $t_2 = t_0 + 2$ seconds, \dots , $t_i = t_0 + i$ seconds, and $t_n = 12:00$ PM.

If the sense data theory were true, then...

P₀ At t_0 you have a green coloured sense datum (with a particular, precise shade).

P₁ At t_1 your sense datum has precisely the same colour as it does at t_0 .

P₂ At t_2 your sense datum has precisely the same colour as it does at t_1 .

...

P_n At t_n your sense datum has precisely the same colour as it does at t_{n-1} .

¹⁰This argument is both a simplification and an application of Williamson’s argument. His argument is not targeted specifically against the sense data theorist, but rather a more general class of epistemologists who share some of the sense data theorist’s sensibilities. In an interview, I heard Williamson remark that his teachers at Oxford taught him this argument against the sense data theory, and he subsequently expanded it. I do not know who is the original source of this argument.

P_{n+1} Therefore (from P_0 to P_n) it follows that at t_n your sense datum is green.

P_{n+2} But at t_n your sense data is not green; it is blue (by assumption).

Each of the premises (P_1 - P_n) are consequences of the sense data theory's claim that the sense data have precisely the properties that they appear to have. Since we have reached a contradiction (P_{n+1} & P_{n+2}), it follows that the sense data theory must be false.

So what is the upshot of all of this? In summary, we have just found that the defining traits of sense data theory have landed us in a contradiction. The sense data theory entails that the objects of perception are *completely transparent to us*. But if that were true, then there couldn't even be the tiniest imperceptible change to our sense data. However, that contradicts the facts of gradual change. Change can occur so gradually that it doesn't even *seem* to change over short intervals of time. This shows that the original sense data theory cannot be true according to the definitions it gives us.

(In my experience, I know that some philosophers can be suspicious of this style of argument because of its superficial resemblance to the paradoxes of vagueness. But a proper understanding of the argument will reveal that it doesn't exploit the paradoxical aspects of vagueness. I will respond to this worry, but I will save the discussion for an appendix to this chapter.)

So now I've offered my main reason for rejecting the sense data theory. To quickly summarize, the sense data theory posits the existence of subjective mental images that exist within the mind and serve as the objects of perception, as opposed to objective external objects. A chief motivation for their claim is the phenomena of illusion, which reveals a gap between appearance and reality. But the sense data theorist goes one step further and reifies *appearance* and *reality* into two distinct kinds of objects: *the appearances* (the mental images, the sense data) and *the real objects* (e.g. external physical objects). Now, to keep with their motivations, they cannot allow for any gap between appearance and reality when it comes to the sense data themselves, and so they imbue the sense data with a spectacular epistemological feature: *how they seem to be is how they are*. But as my argument shows, this feature is their own undoing. There cannot be any entities (subjective or otherwise) that meet the requirement of *always being exactly how they appear to be*. The phenomena of gradual change shows this to be impossible, because we do not have the power to discern slight change over small time intervals. This observation undermines the sense data theory in two respects. Not only does it show that the sense data are defined to have impossible properties, but it also undercuts the argument from illusion. For it shows that, no matter what the objects of perception turn out to be, illusions will always be possible *in principle*. Hence, the phenomena of illusion provides no reason not to take the objects of perception to be real, external objects.

5 Conclusion

As it happens, the epistemological themes of this chapter generalize much further than I have led on. There has been a long history of epistemologists who exhibit similar tendencies to the sense data theorist. Impressed by the distinction between *appearance* and *reality*, they refuse to grant foundational status to our knowledge of external objects. They doubt that our cognitive reach can extend (directly) beyond the elements of our own private mental lives.

As a result, they retreat to epistemological foundations that are supposed to be on firmer ground. Typically this will be our knowledge about our own private mental states. These are the kinds of states that are alleged to be ‘epistemically transparent’ to us; so we can know about them safely and securely, and this knowledge comes prior to anything else. It’s as if our first-personal knowledge forms a kind of *cognitive sanctuary*, which serves as our retreat from skepticism. And then once we’ve retreated into the sanctuary of knowledge about our minds, we may attempt to re-ascend to gain knowledge of the external world. We are thus lured into an epistemological misadventure of trying to infer the external world from the traces left within our sanctuary.

In his book, *Knowledge and its Limits*, Timothy Williamson develops an argument to show that this entire ‘retreat to the sanctuary’ approach to epistemology is founded on a myth. His result is that there simply is no special realm of knowledge that enjoys epistemic transparency—whether it concerns our own private mental lives or anything else. There are no ‘transparent’ states that are such that, *whenever they obtain*, we are *always* in a position to know of them. The argument that Williamson uses for this conclusion is in broad outline very similar to the proof that I presented in the previous section; indeed, my proof is modelled off of his. Although his proof is a bit more complicated than mine because it has greater ambitions. Since my only target was the sense data theory, I was able to present something simpler. But if we are able to see the lesson as it applies to the sense data theory, we can gain the more general insight: there is no realm of epistemically transparent states that are uniquely apt to serve as foundational knowledge.

It is important to see that Williamson’s conclusion is not skeptical. In fact it is quite the opposite. For all that Williamson has said, we are still positioned to know a great deal about our own private mental lives. The point, instead, is that this knowledge does not enjoy a *unique* epistemic privilege. We can just as well know many things about our external environment, by using our perception. The former kind of knowledge is no more special than the latter kind of knowledge. So there is no good reason to treat the former as foundational knowledge to the exclusion of the latter. Not only was this tendency founded on a myth (the myth of epistemically transparent states), but it plays right into the skeptic’s hands. The correct approach for defeating skepticism, according to Williamson, is to refuse, from the start, to ever retreat into the sanctuary.

The conclusions of this chapter are not an epistemological panacea. I have argued that we can know (directly) of the *existence* of external objects through perception. So I can know that I have hands, and I can know that there is a

coffee cup before me. But there still remains a distinction between *the properties that they have* and *the properties that they appear to have*. Sorting out an object's *real* properties from its *apparent* properties can (at times) be no easy feat. But at least we have a foothold that the sense data view would never have given us.

Appendix

In this chapter I presented an argument that I claimed to be a proof against the sense data theory. But the style of argument that I used is bound to rouse suspicion within a certain group of jaded philosophers. They will be suspicious because the argument ostensibly appeals to *vagueness*, and we all know that vagueness is a hazard for paradoxes. Witness the classic sorites paradox. A single grain of sand does not make a heap of sand. Nor does two grains of sand. Nor three. And it seems that for any number of grains of sand, if we do not yet have a heap of sand, then we cannot make a heap by adding only one more grain. But if we follow this line to its logical conclusion, then we'll have to conclude that there are never any heaps of sand. How paradoxical! Clearly that's wrong! Now, what is the lesson of the sorites argument? Well according to many philosophers, we *know* that there are heaps of sand. We thereby know that there *must* be something wrong with the sorites argument and other arguments like it. What exactly they think has gone wrong is not the issue. The point is, that when we're dealing with vague properties like *heap* and *green*, we need to tread lightly.

Having heeded this lesson, some philosophers will respond to my proof with a speech like this. "Look, I concede that your argument is formally valid. But still, your argument appeals to vague properties (green and blue) and slow incremental change. And we know that arguments like this are liable to produce false conclusions. So even if I cannot identify a false premise or an invalid inference in your proof, we should still be suspicious of all arguments of this sort."

In response to this philosopher's speech, I would emphatically deny that my proof exploits the vagueness of *greenness* or *blueness* in any way that should cause concern. In fact, I can rephrase the argument using *perfectly precise* colour predicates and yet the same problem will still arise. Let's say that at t_1 the screen emits light that has a wavelength of *exactly* 540 nanometres (a middling wavelength for green colour) and at t_n it emits light of exact wavelength 420 nanometres (a middling wavelength for blue colour). We can then stipulate that the wavelength decreases from 540 nm to 420 nm at a constant rate, gradually enough so that the change over a second would be imperceptible to the human mind.

But all of this increased precision does not help the sense data theorist one bit. Because still they must say that we have a certain (precise) sense datum at t_1 , we lack that sense datum at t_n , and so our sensory states change over

time—slowly, gradually, and at a constant rate.¹¹ But despite the sensory states gradually changing, it does not *seem* to us to change over small intervals of time. It must change even when it doesn't seem to change. But this proves that our sensory states are not totally transparent to us. The way they *seem to be* is not always *how they are*. And so the sense data theory must be mistaken.

What this tweaking of the proof shows is that my argument doesn't essentially exploit the vagueness of colour properties. Rather, it hinges on the fact that the human mind does not have infinitely precise powers of discrimination. In particular, it does not have the powers to discern slow, gradual, constant change through short intervals of time. But if the sense data theory were to be true, then we would need to be able to discern this, because every slight change in our sense data would need to be apparent to us, and so we would need infinitely precise powers of discrimination. Since we do not possess these superpowers, the sense data theory must be false.

This last point bears repeating. The thrust of the argument against sense data is *not* that we cannot discern the vague boundary between *green* and *blue* because that boundary is indeterminate. Rather, the *real* thrust of the proof is that we cannot discern the *slight changes over short intervals* that must occur for gradual change to take place. And since we cannot discern this, these states cannot be entirely transparent to us.¹²

¹¹You can even imagine that we have a *perfectly precise* language for describing our sense data, so that each sense datum at each time interval has its own perfectly precise colour predicate that applies to it. With this perfectly precise description of our sense data, there will be *more colour predicates that express more sense-data colours* than we can humanly discern. And *that's* the real thrust of the argument.

¹²Even though I'm speaking of an argument whose sole target is the sense data theory, this last discussion echoes the defence of Williamson's anti-luminosity argument from *Knowledge and its Limits*. What I have said here can also be understood as dispelling a misunderstanding of his anti-luminosity argument. Contrary to what some philosophers believe, Williamson's anti-luminosity argument does not hinge on vagueness. It too hinges on our inability to discern the slight changes over short intervals that must occur for gradual change to exist.