

# 12

## Beyond Appearances: The Content of Sensation and Perception

*Jesse J. Prinz*

There seems to be a large gulf between percepts and concepts. In particular, concepts seem to be capable of representing things that percepts cannot. We can conceive of things that would be impossible to perceive. (The converse may also seem true, but I will leave that to one side.) In one respect, this is trivially right. We can conceive of things that we cannot encounter, such as unicorns. We cannot literally perceive unicorns, even if we occasionally “see” them in our dreams and hallucinations. To avoid triviality, I want to focus on things that we can actually encounter. We perceive poodles, perfumes, pinpricks, and pounding drums. These are concrete things; they are closely wedded to appearances. But we also encounter things that are abstract. We encounter uncles and instances of injustice. These things have no characteristic looks. Percepts, it is said, cannot represent abstract things. Call this claim the Imperceptibility Thesis. I think the Imperceptibility Thesis is false. Perception is not restricted to the concrete. We can perceive abstract entities.

This may sound like an obvious claim. We often use perceptual terms widely to say things such as: “I perceive a lack of agreement” or “I see where you are going with that argument.” But, by most accounts, these uses of perceptual terms are either metaphorical or, at any rate, different from the use of perceptual terms in cases that more directly involve the sense modalities: “I perceive distant rumbling”; “I see a red light over there.” The abstract cases are interpreted as involving the sense modalities, if only indirectly. The presumption is that we must first pick up something with our senses and then judge that there is, say, a lack of agreement. Moreover, the abstract cases are presumed to require a level of mental representation that is not perceptual in format. I want to deny all of this. Perceiving abstracta can be just like perceiving concreta.

Those willing to abandon the Imperceptibility Thesis might dig in their heels elsewhere. If there is no semantic gulf between percepts and concepts, there

I am deeply indebted to two anonymous referees and, especially, Tamar Gendler. This paper benefited tremendously from their detailed comments and excellent advice.

might be a gulf between *sensations* and concepts. As I will use the term, a sensation is a state in a perceptual stream that is phenomenally conscious. On the face of it, sensations are even more restricted than percepts. We can perceive concrete entities, because they are closely wedded to appearances, but, intuitively, we can sense only appearances themselves. We can sense poodly appearances, but sensing a poodle, where that means having a phenomenal experience of the property of being a poodle, seems impossible. There are no poodle qualia. Sense-data theorists traditionally made claims like this, and, as we will see below, an alluring argument can be marshaled in support. If we cannot sense concrete entities, sensing abstract entities is entirely out of the question. Sensations, it is said, cannot represent anything abstract. Call this claim the Insensibility Thesis. I think the thesis is false.

Most people agree that we can conceive of abstract things, but many deny that we can perceive and sense abstract things. Or, more accurately, they assume that percepts and sensations cannot have abstract contents. If we can be said to perceive and sense abstract things, it is only indirectly by combining percepts or sensations with concepts. On this approach, percepts and concepts do not represent abstract things, but perceivings and sensings can represent abstract things when combined with concepts, which are presumed to be something other than percepts and sensations. In other words, some authors allow that we can perceive something or sense something *as* falling under an abstract concept, where that means we produce a complex mental state that has both a percept or a sensation and a concept. Importantly, the semantic heavy lifting is done by the concept. Percepts and sensations do not themselves represent abstracta in these cases. Those who think that we can sense and perceive abstracta indirectly still hold the Imperceptibility and Insensibility Thesis. The point might be made by saying that such authors think we cannot perceive or sense abstracta directly. Percepts and sensations cannot represent abstracta themselves.

In this chapter, I will present evidence against the Imperceptibility Thesis and the Insensibility Thesis. That evidence is partially empirical. We are led astray by pretheoretical intuitions about the semantic properties of the representations available to our senses. Research on the nature of concepts and their interactions with sensory processes reveals a different picture. We can, in principle—if not always in practice—sense and perceive just about anything that we can conceive.

## 1 WHAT CAN WE PERCEIVE?

### 1.1 Seeing Things

To navigate through these issues, we need to clarify some terminology. First of all, we need a working definition of perception. I will treat “perceiving” as a success term. Uses in the case of hallucinations and dreams are parasitic on

successful seeing. On the success interpretation, an organism O *perceives* something X only if:

- (i) X impinges on O's sensory transducers;
- (ii) O forms a corresponding perceptual representation as a result;
- (iii) that representation is matched against stored representations that represent X.

This requires a few more definitions. A *sensory transducer* is a psychophysical mechanism that converts physical magnitudes into mental states or mental representations. Something *impinges* on a sensory transducer on an occasion if its instantiation causes transduction in that transducer on that occasion (where causation can be cashed out in counterfactual terms). A *perceptual representation* is a representation in a dedicated input system. A *dedicated input system* is an information-processing system that takes inputs from sensory transducers (and possibly elsewhere), and forms representations that have the function of representing inputs from those transducers. Dedicated input systems typically use proprietary codes; they use representations that have some syntactic and semantic properties not found in other input systems. As should be clear this definition applies only to uses of the word "perception" that involve these senses. Intellectual uses ("Descartes was very perceptive") don't count. The third condition in my definition refers to matching. Here, some caveats are necessary. Matching does not necessarily involve comparing two token representations. As I am using the term, matching can also occur when an input *triggers* a stored representation. The key idea is that perceiving involves recognition.

Some might take issue with the recognition requirement. Dretske (1969) argues that we can perceive something without recognizing it. He calls this non-epistemic perceiving. We can say, "I must have seen you at the market, but I didn't recognize you." On my definition, this is not a case of perceiving, though I would allow that it is a case of seeing. I have the conceptual intuition that seeing can be non-epistemic, but perceiving can't. Nothing will ride on this. Dretske's notion of non-epistemic perceiving can be captured by dropping condition (iii).

That does not mean that my definition is equivalent to what Dretske calls epistemic perceiving. I aim for something that is stronger than Dretske's non-epistemic perception, but weaker than his notion of epistemic perception. I think we need something in between. There are two reasons for this diagnosis.

First, Dretske may blur an important distinction that can be found in the pages of Wittgenstein (1953). Wittgenstein contrasts ordinary cases of seeing with "seeing-as." The latter phrase is colloquially reserved for cases where we need to apply a special skill or interpretive act. We can see a cloud as a warthog, but it sounds odd to say I see this fork *as* a fork. When I see the fork, I just recognize it. One might call this recognitional seeing. Recognitional seeing is not non-epistemic. A tarantula can see forks in a non-epistemic sense, but a tarantula cannot see a fork in the way we can, because tarantulas cannot (I presume)

recognize forks. Recognitional seeing requires the capacity to represent forks as such. But recognitional seeing differs from seeing-as in at least three respects. First, seeing-as involves two stages. Initially an object is recognized, and then it is interpreted in a new way. A duck is re-construed as a rabbit. Second, recognition need not involve the application of a concept. Concepts are representations that can be actively tokened by an organism (Prinz, 2002). In recognition, we often use representations that can only be tokened passively. I can recognize certain things that I cannot bring readily to mind in imagination or reflection. Recognition outstrips conceptualization. Seeing-as usually doesn't. To re-construe a perceived object, we typically deploy concepts (though perhaps not always—swapping figure and grounding, for example, can be done without concepts). Third, recognitional seeing is always factive, and seeing-as is often not. When we see a cloud as a warthog, there is no avian swine. Dretske does not use “epistemic seeing” to subsume non-factive cases, but neither does he distinguish epistemic seeing from seeing as. I want to make it explicit that perceiving, in the sense outlined above, is not equivalent to seeing-as.

There is a second contrast that I want to draw between recognitional seeing and epistemic seeing in Dretske's sense. Dretske (1981) argues that epistemic seeing involves a process of digitalization. Information that is nested at one stage of sensory processing becomes un-nested. In the fork example, this might work out as follows: at some stage in visual processing, representations carry the information that a fork is present, by carrying information about other features, which are natural signs for the property of being a fork. In particular, we represent forkhood via forky appearances. At this stage of processing, one cannot perceive a fork as such. One cannot have forkhood as the object of perception. To perceive forks epistemically, we need to digitalize. We need to extract forkhood from forky appearances through a further stage of processing that abstracts away from the appearances and represents forkhood. A digitalized representation of forkhood does not represent forkhood in a way that depends on carrying information about forky appearances. For example, it might be a representation that is triggered by an open-ended range of forky appearances. If a representation is triggered by an open-ended range, it cannot represent any one of those appearances because it will not carry the information that any one of those appearances has been instantiated. Such a representation represents forkhood as such.

I reject this account of epistemic seeing. It entails that we can epistemically see forks only if we have representations that abstract away from forky appearances. This strikes me as terribly implausible. I think we can visually recognize forks, and hence epistemically see them, by means of fork images: representations that encode features of forky appearances. Dretske does not offer a good way of accommodating this possibility. I will suggest a way below. My proposal is compatible with Dretske's approach to intentionality, but it differs from the digitalization theory of epistemic seeing. That is why I prefer to talk about recognitional seeing (which is, in my terminology, just visually perceiving).

The first condition in my definition of perception makes it clear why we can perceive only those things that we encounter. In order for something to impinge on the senses, it needs to be present to do the impinging. But what exactly does presence require? As a first pass, we can say something is present if it is here now. This needs to be qualified. Consider very distant objects, such as stars. What could “here and now” mean such that stars count as being here now? It is tempting to say that something counts as “being here now” just in case it is accessible to our senses here now. This, of course, is circular if we are hoping to define accessibility in terms of presence. We can do a little better, however, if we go back to the idea of impingement. If impingement is a causal notion, then we can say that “being here” means being instantiated within sufficient spatio-temporal proximity to have causal impact on sensory transducers. I will eventually argue that this allows for the perception of abstract things.

To make sense of this claim, we will need a characterization of abstractness. Abstract is the opposite of concrete, or at least the opposing pole on a spectrum. These terms are tricky, however, because they are used differently by philosophers and psychologists. In saying that we can perceive anything that we can conceive, I must argue that we can perceive abstracta in both the psychological and philosophical sense, assuming such things (a) exist and (b) are conceivable. I will sometimes use the shorthand “philosophically abstract” and “psychologically abstract” to capture the contrast. In much of the discussion, I won’t bother with modifiers because context can disambiguate or because both kinds of abstractness are relevant.

In psychology, the concrete/abstract contrast is sometimes defined with reference to perceptibility; concrete things are perceivable and abstract things are not. Such definitions clearly won’t do for this inquiry, because I think abstract things are perceivable. But I think we can capture the basic idea behind the psychologists’ definition. Abstract and concrete can be *indirectly* defined in terms of perceptibility.

It all begins with appearances. Appearances must be defined relative to the senses. They are powers that external things have to cause representations in our dedicated input systems. Two things have the same appearance (from some vantage point) if they cause perceptual representations that are intrinsically type-identical (from that vantage point). Two things have similar appearances if the intrinsic properties of the perceptual representations they cause are similar. By “intrinsic properties,” I am thinking of the kinds of properties that we would mention in describing perceptual representations at a psychological- or personal-level of description—properties such as shapes or colors. By analogy, think of the visual information captured on the celluloid of a film (perhaps a 3D film). When we watch a film, we assign meaning to the images that pass before our eyes, but the film itself just captures shapes, colors, and movements. Appearances are like that. For simplicity, I will be assuming here that perceptual representations are like mental pictures. Abstractness and concreteness can be defined relative to

appearances. A thing is concrete to the extent that it would appear alike across different encounters when it impinges on the same sensory transducers under the same viewing conditions. A thing is abstract to the extent that it appears different across encounters. This characterization works for objects, events, and properties (though I will not talk about events). An object is concrete to the extent that it appears alike across encounters with it. A property is concrete to the extent that it appears alike across encounters with its instantiations. Properties have their appearances via their instantiations. The property of being a flounder is relatively concrete; the property of being a fish is less so; and the property of being a sea creature is even less concrete than that.

Here I must step back from the main thread of discussion to pursue a metaphysical aside. I want to clarify how the psychological abstract/concrete distinction relates to the homophonic metaphysical distinction, and I want to show that the Imperceptability Thesis seems plausible whether or not one believes in universals. To philosophically trained ears, talk of concrete properties sounds odd. Philosophers agree that concrete entities can exist in space and time, but they are divided as to whether properties can exist in space and time. Trope theorists say they can, and universalists typically say they cannot. Talk of “concrete properties” is fine for trope theorists (though they often use the term “abstract particulars”). For trope theorists, all properties are concrete in the philosophical sense. Notice, however, that this does not make all properties concrete in the psychological sense. For there is no guarantee, on trope theory, that two tokens of the “same” property will appear alike. Talking about two tokens of the same property is shorthand for talking about two tropes that bear a certain equivalence relation. Trope theorists talk about “exactly resembling tropes,” but “resemblance” here must not be understood in terms of appearance. Trope resemblance is usually taken to be a primitive relation, not analyzable in perceptual terms.

Trope theorists are often skeptical about the existence of philosophical abstracta. For them, there is no special puzzle about perceiving properties, because all properties are instantiated instances. But that does not mean trope theorists deny the Imperceptability Thesis. Trope theory does not entail that we can perceive things that are abstract in the psychological sense. On trope theory, ordinary property terms refer to equivalence classes of tropes, rather than individual tropes. Our concepts can represent these equivalence classes, but what about our perceptions? There is nothing in trope theory to guarantee that we can perceive equivalence classes of tropes. Equivalence classes of tropes are often psychologically abstract (their members have different appearances). It is far from obvious whether such entities are perceivable. Thus, trope theorists may be strongly inclined to accept the Imperceptability Thesis.

Now consider universalists. Unlike trope theorists, they deny the existence of metaphysically concrete properties, because they think properties are abstracta in the philosophical sense. But they can accept concrete properties in a

psychological sense. When I use the phrase “concrete properties” this can be translated into metaphysically benign talk by interpreting it as “properties whose *instantiations* are psychologically concrete.” Universalism seems to entail that we can perceive instantiations of properties, not the properties themselves. If properties are philosophically abstract, then they do not exist in space and time. If that is the case, they don’t seem to be in a position to impinge on our senses. Thus, like trope theorists, universalists may be strongly inclined to accept the Imperceptibility Thesis. The Imperceptibility Thesis does not hinge on a particular theory of properties.

That ends my metaphysical aside. And now back to our main question: what is perceivable? The most intuitive answer is that we can perceive appearances and we can perceive objects that are philosophically concrete, and when we do so, we do it by recognizing psychologically concrete properties. As properties get less concrete, they become less perceivable. Psychologically abstract properties, which are poorly correlated with appearances, cannot be perceived. Philosophical abstracta cannot be perceived either, because they do not come into physical contact with the senses. This, I think, captures folk intuitions about perceptibility. If questions about what we perceive were best answered by folk intuitions, the Imperceptibility Thesis would be confirmed. We can undoubtedly *conceive* of properties that are very abstract in both philosophical and psychological senses, but such properties cannot be perceived. If we encounter an event that instantiates injustice or a sentence that has the property of being true, we cannot perceive these properties. Injustice and truth can be encountered, but, intuition proclaims, they are imperceptible. This is the intuition I want to reject.

First, one qualification is required. Folk psychology allows us to say, “He could see the injustice of her actions,” or “I can see that you are right.” This might be taken as evidence for widespread folk acceptance of the thesis that I will be defending. I don’t think that is the right interpretation. In the introduction, I suggested that such locutions exploit a metaphorical or extended use of perceptual vocabulary. Notice that it sounds a little weird to say, “I *literally* saw the injustice in her actions,” or “I could see the injustice, despite the fleck of dirt in my eye.” Blind people can see abstract things, in this extended sense. Tiresius could see that Oedipus was doomed. There is no entailment from “I see P” to “I see.” In the cases such as “seeing injustice,” the folk will resist the inference emphatically. After all, injustice doesn’t look like anything. I will part company with the folk. I think we can see abstract things in a literal way—i.e., via the senses.

## 1.2 Abstract Images

The Imperceptibility Thesis says that we cannot see any of the abstract properties that we encounter. Its denial is the thesis that we can perceive some abstract things. I want to go even farther and argue that virtually no abstract property is,

in principle, imperceptible. In making this case, I will focus on visual perception, although it will be easy to extrapolate from my examples how abstract perception could arise in other modalities.

To make the case for seeing abstracta, I will begin with a general theory of how perceptual representations represent. By default, we should assume that perceptual representations represent in the same way that other kinds of mental representations represent. If we have a good theory of representation, then the question about perceptual content becomes an empirical question. We need not worry about mongering intuitions or relying on introspection.

This is not the place to *defend* a theory of mental representation, but fortunately many defenses can be found in the literature. The kind of theory that I favor derives from Dretske's (1986) teleological/informational intentionality. I have defended a version of Dretske's approach against standard objections elsewhere, and I have applied it to concepts and to emotions (Prinz, 2000b; 2002; 2004). Roughly, the idea is that mental representations represent that which they have the function of detecting. Put a bit differently, a mental representation represents that which it is set up to be set off by. Two conditions are implied here. First, there is an informational condition: the representational content of a representation must be something that is capable of setting off that representation. If a mental representation *M* represents a content *C*, then instances of *C* have the power to cause tokens of *M* when they are encountered. This causal link must be reliable. In his initial formulations, Dretske (1981) captured the idea of reliability by supposing that, within certain boundary conditions, the probability of *C* given a tokening of *M* is equal to 1. This is too strong a requirement, and it saddles Dretske with the extra burden of specifying what those boundary conditions are. One can get by with the simpler claim that *C*s have the power to cause *M*s.

The second condition in Dretske's psychosemantics is teleological: a representation must be *set up* for the purpose of being set off by that which represents. Talk about purposes can be cashed out historically. Roughly, the idea is that a representation type *M* came to exist by virtue of being set off by an instance or instances of its content. *M*s would not exist were it not for prior encounters with *C*s. Those encounters can occur in the learning history of a particular representing organism or (with innate representations) in that organism's ancestral past. Dretske's first condition ensures that my *water* concept will refer to a substance that is typically clear and tasteless, for those are the substances that set off *water* tokens. Dretske's second condition selects out the specific clear, tasteless, liquids, that played a role in the acquisition of my *water* concept, namely H<sub>2</sub>O.

I think Dretske's theory of intentionality must be supplemented with one more condition. Content is further constrained by what I have called semantic markers, following Putnam (Putnam, 1975; Prinz, 2002). As I use the term, a semantic marker is not a kind of representation, as the term may imply, but



a way of using a representation. Different patterns of use determine what general ontological category a representation represents. For example, semantic markers determine whether a representation represents an individual or a natural kind. If there were no semantic markers, it might be indeterminate whether my water concept referred to H<sub>2</sub>O in general or to the specific sample of H<sub>2</sub>O that was ostended when I first learned the concept. Semantic markers are also used to determine whether a representation represents distal or proximal stimuli, and they may play a role in coping with certain forms of Quinean indeterminacy. Patterns of use may include inferential roles that distinguish between interpretations that would be equally viable if content were determined entirely by stimulus detection and causal history. I mention this add-on to Dretske's theory for completeness. It will not play much of a role in the discussion below.

The Dretskean approach to intentionality is controversial, but it is widely regarded as one of the best theories under contention. Other approaches to intentionality are compatible with the conclusions that I will draw about perception, but I will assume that Dretske's theory is essentially correct. Given that assumption, we can return to the question of perceptual representation. The question about what perceptual representations represent can now be traded in for the question of what those representations are set up to be set off by. This is an empirical question. It requires investigating, for any given perceptual representation, what sets it off and how it came into being. As an example, take a population of cells in a particular region of primary visual cortex. We can measure what sets these off by determining what kinds of stimuli are present when they fire. Perhaps they are responsive to lines at a particular angle of orientation. To determine how these cells came to be responsive to lines of that kind, we must engage in some historical guesswork. We can surmise that they came to have their current response profile in virtue of events in the organism's lifespan or in virtue of events in the organisms evolutionary past (or, most plausibly, some combination of these). If we conclude that these cells came to detect angles by virtue of having been set off by angles by the past, then we conclude that they are angle representations. Other cells in the visual system are responsive to colors, directions of motion, depth, and so on.

Now consider a more complex case. There are cells in higher visual areas that seem to be responsive to specific kinds of objects, rather than colors or shapes. For example, it is known that monkeys have cells in the fusiform gyrus that are especially responsive to faces (e.g., Desimone, 1991; for humans, see Kanwisher et al., 1997). It is not unreasonable to think that they are set up for this purpose, given the importance of faces for social creatures. It is also known that the visual stream is highly plastic. An encounter with a novel object can cause cells in the visual system to become responsive to that object in the future. For example, if you show a monkey a piece of wire bent in a peculiar pattern, cells in the inferotemporal portion of its visual stream will become detectors for that very pattern (Gauthier and Logothetis, 2000). It has been surmised that small

populations of inferotemporal cells become detectors for individual objects, such as particular faces (Gross and Sargent, 1992). Kreiman et al. (2000) have found individual cells in human medial temporal cortex that respond to multiple images of the same person. They found, for example, a cell in one subject that was responsive to pictures of Bill Clinton. It is safe to assume that this cell was set up for the purpose of Clinton detection. It came to play that role through encounters with Clinton (or images of Clinton). (Things may be more complicated for cells that respond to Saddam Hussein, because he allegedly had a group of look-alikes standing in for him in public appearances.)

A cell that is responsive to Clinton's face may be interpreted as representing Clinton's face. Clinton's face is not merely an appearance. It varies across angles, expressions, and viewing conditions. Clinton's face is a concrete object, though; it is closely correlated with appearances. Note, moreover, that Clinton's face is highly correlated with Clinton, the person. A cell that reliably detects Clinton's face is, thereby, a Clinton detector. Cells that are involved in detecting *particular* faces may attain this function in virtue of the fact that they are good tools for detecting particular persons. The idea of being a good tool might be captured by downstream effects. Millikan (1989) explains the teleological contribution to content by appeal to "consumers," the systems that make use of representations. If used to coordinate responses towards a person, and not just a face, the cell that responds to Clinton's face may have the function of detecting Clinton. Such downstream effects may be one way of cashing out the idea of semantic markers. The Clinton face cell may, in virtue of its use, be marked as a person-representation. If so, this cell is a Clinton representation, and not just a Clinton-face representation.

Clinton, of course, is a relatively concrete object, so the existence of Clinton representations in the visual stream would not establish the possibility of seeing things that are abstract. But the Clinton example will help provide a strategy. Let's shift to a more abstract domain. Consider numerosity. When we perceive groups of objects we can perceptually determine whether there are, say, three objects or four. We share this ability with infants, nonhuman mammals, and birds (Dehaene, 1997). Groups of three vary significantly in their appearances, but the visual system is capable of tabulating quantity across a wide range of variation. Let us suppose, plausibly, that there are cells in the visual system that have the function of firing when we encounter groups of three. I think such cells represent the property of threeness or, at least, the property of being a three-item group. They are invariant across a wide range of concrete inputs, none of which are correlated well with their firing. Threeness has the highest degree of correlation. Moreover, these cells are used to keep track of quantity. They contribute to simple arithmetic, and, when we see a group of three strawberries, say, we know to reach and grasp three times.

Of course, this won't help with larger numbers. To represent exact quantities much larger than three, Deheane makes the obvious suggestion: we count. More

specifically, Deheane says we use representations of number-words in our language systems. These representations are perceptual. They are representations of the sounds or symbols used in linguistic communication. To determine that there are, say, fifty-seven toothpicks on the table, we just shift attention from pick to pick, assigning each a label: one, two, three . . . By the time we get to the fifty-seventh toothpick, we form a visual image of it along with an auditory image of the word “fifty-seven.” This is a perceptual state, and it is one that is set up to be set off by sets of fifty-seven items. We can thereby perceive sets with arbitrarily high cardinalities (though performance limitations will prevent us from going too far). By this method, we can perceive the difference between a chiliagon and a 999-sided figure. We need only count the sides!

The Dretskean approach to semantics can explain how we represent things that are abstract in the psychologists’ sense. A representation can refer to things that do not appear alike if it is reliably caused by those things, despite the differences in their appearance. Specific populations of numerosity neurons represent threeness, because those neurons are best correlated with threeness, and not any particular group of three things.

This story about how we refer to psychologically abstract properties only delays the question about philosophical abstracta. That question is even more pressing. The philosopher will object that, though threeness may be correlated with cells firing, it cannot be causally responsible for their firing, because threeness doesn’t exist in space and time. Only concrete instantiations of threeness can have causal efficacy. Thus, our cells can represent instantiations of threeness, but not threeness itself. One might reply by stipulating that representation does not require causation, but only correlation and function. But this won’t help. If threeness exists always, then it is correlated with every mental representation that ever gets tokened. And, even if representation doesn’t require causation, perception, as I have defined it, does. Fortunately, there is a solution.

Notice, first, that the problem of referring to properties has an analogue in the case of referring to concrete individuals. Strictly speaking, our mental representations of individuals (such as Clinton) are not correlated with the *existence* of Clinton. Clinton always exists (or at least he exists until he dies), so his existence is correlated with all my representation tokens. To get around this problem, we define reference in terms of *encounters*. My Clinton representation is tokened when I encounter him (or representations of him, in which case reference is one step removed). Above, I said that we encounter something if it becomes causally accessible to us. Now, it would be crazy to deny that properties are causally efficacious. For many philosophers, causal efficacy is a necessary condition for realism. If properties exist, they must pull their weight. One way to explain the causal efficacy of properties is by appeal to counterfactuals. *Very* roughly, we can say that a property P is causally efficacious in item O’s bringing about effect E, if E would not have occurred if O had not instantiated P. Properties need instantiations to have an impact, but, once instantiated, they deserve causal

credit. If we buy into this picture, there is no problem saying that we encounter properties. We come into causal contact with them (via their instantiations). So reference to properties can be characterized in the same way as reference to concrete individuals. (Fodor (1990) even argues that we only refer to properties, because he thinks that reference depends on laws, and laws are defined in terms of properties.) I conclude that perceptual representations can refer to abstract properties. We need to be able to say that properties are causally efficacious, and once we do that (by whatever theory of causation turns out to be correct), then we have the resources to attribute abstract contents to perceptual states.

The preceding story explains how a universalist can allow reference to abstract properties. The problem of abstract reference facing trope theory is different. I have been talking about abstractness in the philosophical sense. Trope theorists deny that such properties exist, so they need not worry about how something that isn't in space and time can have a causal impact on the world. Properties are tropes and tropes are present. But trope theorists face a related problem. They must explain how a concept could refer to an equivalence class of exactly resembling tropes. The members of an equivalence class are rarely compresent. If representation depends on detection relations, how can a representation refer to an equivalence class?

The question can be addressed by dissecting the notion of reliable causation. A representation that is caused by a particular trope will also be caused by exactly resembling tropes. Thus, there is a distributive sense in which such a representation is caused by the equivalence class of tropes, i.e., it is caused by the members of that class. Compare: pinpricks cause pain. Furthermore, a representation that is caused by all (or most) tropes in a class, is also better correlated with the class than with any individual member of the class. The class itself has no causal efficacy above and beyond the efficacy of its members, but, semantically, it is the class, and not the members that have the kind of causal correlation. When we talk about reliable causal relations between two things, we mean a relation that is causal and highly correlated. The relation between equivalence classes of tropes and tokenings of a mental representation satisfies both of these conditions. The class is causally related to the tokenings via its parts, and it is correlated with tokenings, when taken as a whole. The causation is distributive and the correlation is collective. If we interpret the Dretskean account so as to allow this, reference to trope classes will be secured.

I conclude that mental representations in our perceptual systems can refer to both philosophical and psychological abstracta. They can refer to properties whose instances don't look alike, and they can refer to those properties (and not just their instances) even if they are universals or equivalence classes of tropes. If defensible, this is a useful result. But, for my thesis, something stronger is required. I want to show that we can perceive *any* conceivable abstract thing that we can encounter. To defend this claim, it will help to consider another example of abstract perception.

Suppose you are teaching a class and you want to find out how many of your students are philosophy majors. Being a philosophy major is an abstract property. Philosophy majors do not look alike. So you tell your students to raise their hands if they are philosophy majors. I would contend that a visual representation of raised hand under these circumstances quite literally represents being a philosophy major. When we ask people to signal something through hand-raising, we effectively assign the perception of those hands a particular informational function. For a brief interval, the perception caused by raised hands reliably detects the property of being a philosophy major, and those perceptions are selected to serve that purpose. We can think of the perceptions as having a “passing function” assigned temporarily for a particular purpose.

This is a surprising result. I am not claiming merely that we can infer that there are philosophy majors by looking at raised hands. I am claiming that the perceptual state caused by raised hands represents the property of being a philosophy major. It represents that property, since that’s what it is set up to be set off by. Folk intuitions allow only that we can perceive properties like this in an indirect way. It is hardly controversial to say that we can “perceive” abstract properties by first perceiving something else that indicates that the abstract property has been instantiated, and then representing the abstract property in a subsequent mental episode. On this model, seeing the property of being a philosophy major is very different than seeing red. I want to claim that there is no difference. Seeing red is a matter of having a visual state that represents red. Seeing the property of being a philosophy major is a matter of having a visual state that represents that property. By assigning transient meaning to our visual states, we can literally and directly see the property of being a philosophy major.

This case underscores the fact that percepts can come to represent abstract properties in at least two ways. The first way is illustrated by the numerosity example discussed above. In cases like that, perceptual systems *store* representations that detect abstract properties. We see threeness by matching a visual experience against a representation of threeness that is built into the visual system. Here, the “matching” that takes place between occurrent percept and stored percept does not require calling up a representation from somewhere outside of the visual system. A visual encounter with three objects triggers a representation of threeness by directly activating it in the visual stream. The hand-raising case works differently. There, we assign a meaning to perceptual states by matching it to an image that is called up from elsewhere. We call up an image of raised arms, and that image is *transiently* linked to our beliefs about what raised arms indicate in this context. Both kinds of examples present a challenge to the Imperceptibility Thesis. They show that we can represent abstract properties in perception. A fan of the Imperceptibility Thesis might concede the point, but deny its generality. The case of numerosity perception may be very unusual, and the hand-raising case is a bit contrived. We still need a more general argument for the perceptibility of abstract things.

I think such an argument is available and, indeed, it falls out of the semantic theory that I just endorsed. If Dretskean semantic theories are right, then everything that we can represent is represented using Dretskean detectors or combinations thereof. Let's consider abstract concepts, such as *injustice* or *truth*. If these things are to represent in a Dretskean manner then they must be either reliable detectors for those abstract properties, or combinations of reliable detectors. Detection is a relationship between mind and world. In order for an item in the mind to become a detector for something out there, the thing out there must cause that item to occur. That causal relation must be mediated by the senses. The senses are the only avenue by which things out there can cause states in here. So, if Dretske is right about reference, then referring concepts must be, at the very least, built up from representations that get correlated with the world through perceptual representations. Those representations may be extremely complex. They may be highly variable. They may include representations of words, sentences, or other public symbols. But, no matter what, they must be perceptual. Every referring concept, no matter how abstract, must be linked to a collection of perceptual representations that play an indispensable role in establishing reference. This is going to be the key point in debunking the Imperceptibility Thesis.

Consider some examples. Suppose you want to know whether a particular statement is true. Someone says, "It is snowing outside." You want to see if she is being honest with you, so you have a look at the window, and you see snow. When you see that snow, it confirms that it is snowing outside, and it also confirms that the sentence "It is snowing outside" is true. You can be said to *perceive* the truth of that sentence insofar as you used your concept of truth to establish a way of recognizing truth through perception.

Here's another example: it is possible to perceive an uncle, but, intuitively, it is not possible to perceive the property of being an uncle. That property seems too abstract. But this intuition is a bit hasty. If you go to a family gathering and want to determine whether there are any uncles present, you can simply search for people you know to be uncles. If you recognize your own uncle nestled on the sofa while searching for uncles, your perception of him qualifies as a perception of unclehood. It is not just a perception of an instance of unclehood, but, arguably, a perception of unclehood itself since you have assigned perceptions of uncles the transient meaning of unclehood. The perceptual representation caused by seeing your uncle has the function of detecting that abstract property on this occasion.

Now consider moral concepts. I think we often apply concepts such as *good* and *bad* by paying attention to our emotions. If you experience guilt after doing something, your perception of the guilt represents that you have done something bad. It represents that moral fact, because guilt is set up to be set off by bad behavior (Prinz, forthcoming). This kind of account does reasonably well for thin moral concepts, in Williams's (1985) sense, but it is less plausible for more

sophisticated moral concepts, such as *injustice*. It is less plausible that any particular emotion reliably indicates the presence of injustice. Indignation may be designed to register injustice, but to know that you are indignant and not merely angry, you need to know something about the factors that triggered that emotional response. Let's consider how one might recognize injustice without relying on telltale emotions. Instances of injustice are events of various kinds. Suppose you read in the paper that a federal court has upheld a decision to allow a corporation to maintain discriminatory hiring practices. These are just words on a page, but as you read them you engage in a series of inferences, mediated in large part by language. "That's unjust!" you conclude. The concept you are expressing manages to refer to the property of injustice, whatever that is, precisely because it is causally responsive to situations of this kind. There are events in the world, such as bad court decision, that are unjust. These events are often encountered indirectly through linguistic symbols. Through linguistic inference, we can derive the conclusion that these events each involve an authority-sanctioned unequal distribution of costs or benefits to equally deserving individuals or groups. But a token instance of the concept of injustice needs to be comprised of that complex description, and the recognition of injustice needs to involve a complex linguistic inference. Recognition requires only that we have some mental representation, *of any kind*, that has been set up to be set off by inequitable distributions. Suppose that we infer, through linguistic inference, that a particular event would be unjust if it were ever to occur. When that event subsequently does occur, we can represent it as unjust without going through all of the inferential steps. Imagine that you have been following the court battle between a corporation and those against whom the corporation has discriminated, and you have come to the conclusion that the jury would be acting unjustly if they were to favor the corporation. Imagine, further, that you are in the courtroom when the judge reports the jury's decision. When you hear the words "not guilty," your representation of those words in this context carries the information that an injustice has occurred. Here, I would contend, you are literally perceiving injustice. You are perceiving injustice because you have assigned a passing function to the words "not guilty." Those words are, at this moment, a tool for detecting injustice. Just as we can perceive a philosophy major by seeing a hand, we can perceive injustice by hearing a couple of words.

These examples can be used to illustrate the problem with Dretske's (1981) theory of digitalization. In that earlier work, Dretske argued that you can epistemically represent a property in perception only by forming a representation that carries information about that property without carrying information about other properties in which that property is nested. In other words, if A carries information about B (but not conversely), a representation of A is not thereby a representation of B. To represent B, one needs a representation that abstracts away from A. One can represent forkhood only by abstracting away from forky appearances. The examples that I have been offering do not depend on an abstraction process of this

kind. Consider uncles again. When you use a mental image of your uncle to determine whether there are any uncles in the room, you do not abstract away from his appearance. Your uncle image continues to carry information about your uncle's appearance. However, for that moment, it serves as a representation of the property of being an uncle by virtue of the function it has been assigned. Dretske's (1986) teleological theory of content can be played against his earlier proposal by showing how we can bypass the need for digitalization.

The present proposal bypasses the need for digitalization in a second way. Dretske (1981) implies that digitalization is a *stage* in information processing: we first represent appearances using one representation, and then go beyond appearances using another. There is one representation of uncle-like appearances, and then an abstract representation registering the presence of uncles. I think these two stages can be collapsed. The representation of unclehood is token identical to the representation of uncle appearances. We perceive the abstract property *by means of* a perceptual representation of a concrete particular.

This alternative to the digitalization story has an important implication. There is a sense in which everyone might agree that we can perceive abstract entities. Everyone might agree that we can use appearances and concrete objects as evidence to confirm that an abstract property has been instantiated. I can see that Barbara is a philosophy major because she raised her hand. This kind of talk is commonplace, and, if interpreted in a certain way, it is uncontroversial. I see that Barbara is a philosophy major by virtue of the fact that she raised her hand. I see abstractions by virtue of their concrete signs. The use of the phrase "by virtue of" renders the "seeing" indirect. I don't literally see that Barbara is a philosophy major, if that means I directly experience that through vision. Rather, I see something else (her hand) which allows me to conclude that she is a philosophy major. This modest suggestion is not the one that I am defending. I am not suggesting that we *indirectly* perceive abstract things *by virtue of* directly perceiving concrete things. Rather, I am suggesting that we *directly* perceive abstract things *by means of* directly perceiving concrete things. The perception of abstract things is direct, because it does not require that we perceive something else in advance. And it is perception "by means of" the concrete things, because our images of concrete things *constitute* our perceptions of abstract things; they are not merely instrumental causes of those perceptions. Thus, we literally perceive unclehood and the property of being a philosophy major. This is a controversial claim, but I think it is true.

The controversial claim also generalizes. Every concept that refers by reliable detection can be triggered by a perceptually mediated encounter with the world. The perceptual representations that cause a concept to be tokened on such an encounter *could* be set up as passing representations of the property designated by that concept. I would actually take this one step further. I believe that concepts *just are* stored records of perceptual states. Once we adopt the view that concepts refer by reliable detection and recognize that detection is perceptually



mediated, we must conclude that every concept for which this story is true is associated with a collection of perceptual representations that play this mediating role. Those representations can be said to constitute a mental file. Elsewhere, I have argued that we should define concepts in terms of mental files of this kind (Prinz, 2002). On any given occasion when a concept is tokened, it will be mentally represented using a representation drawn up from a particular file, i.e., the file that establishes a reliable relation with the property represented by that concept. Those representations will vary from occasion to occasion as context demands. Since these files comprise nothing but stored perceptual representations, every token of a concept will itself be a stored perceptual representation. Concepts that do not refer by reliable detection are built up from those that do. If this theory is right, then all concepts are stored records of percepts or combinations of stored percepts (where “percept” is shorthand for perceptual representation). This is not an original proposal. It is a redressing of the theory of concepts defended by Locke and Hume.

This is not the place for a defense of concept empiricism. I want only to trace out an important implication. If empiricism is true, then all our referring concepts, no matter how abstract, are built up from stored percepts. For any concept built up from stored percepts, there should be some condition under which those percepts could be matched with percepts caused by an encounter with objects or events in the world. When such a match takes place, we can be said to perceive the property represented by the percepts used to achieve the match. Thus, if empiricism is true, we should be able, in principle, to perceive the properties represented by any referring concept, no matter how abstract, provided those properties can be encountered. This means that the Imperceptibility Theory is deeply mistaken. Seeing is not restricted to the concrete.

One doesn't need to be an empiricist to accept this conclusion. The crucial premise is that abstract concepts can be associated with perceptual representations, and, indeed, that they must be in order to secure reference to properties in the world. Concepts themselves do not need to be identified with percepts; they need only be associated with percepts. If the non-empiricist accepts this, then there is only one more premise needed to secure abstract perceiving. If a percept that is associated with an abstract concept is applied in perceptual recognition, it imparts the meaning of the concept with which it is associated. In other words, when we recognize things by using stored percepts that we have previously associated with a concept (permanently or in passing), those percepts represent whatever the associated concept represents. If representation is determined by detection and function, this assumption is very plausible. Percepts that are associated with concepts have the function of detecting whatever the associated concept detects. Anyone tempted by this approach to psychosemantics should be comfortable with the idea of abstract perception.

This story is not just Humean; it is Kantian. Kant tells us that concepts without percepts are empty. He also tells us that concepts are rules for

constructing perceptual representations, and that, in imagination, we bring concepts to bear on perceptual episodes (a theory of imagination that he shared with Descartes). Any theory that allows abstract concepts to bear on perception via the deployment of perceptual representations has the resources to debunk the Imperceptibility Thesis. Empiricist theories make this possibility especially obvious, but empiricists are not alone.

## 2 WHAT CAN WE SENSE?

### 2.1 The Realm of the Senses

The claim that we can *perceive* abstract properties does not entail that we can *sense* abstract properties. Perceiving and sensing are not the same thing. Sensing occurs when something that impinges on our sense organs causes a sensation. I will define a “sensation” as a representation in a dedicated input system that is consciously experienced. By “conscious,” I mean phenomenally conscious or having qualitative character.

It is sometimes presumed that sensations are not representational. Traditional sense-data theories, for example, propose a level of perceptual processing that is qualitative, but uninterpreted. Perception is the process of assigning meaning to sense-data. I think this picture is implausible. If representation amounts to functional detection, then representation goes all the way down. The earliest stages in sensory processing are involved in detecting information. Sensations carry information, and they have the function of doing so. But what sort of information do they carry? What do our sensations represent?

Intuitively, sensations are representationally impoverished in comparison with perceptions. It seems perfectly natural to say that we represent concrete objects in perception. It is not very controversial to say that we perceive tigers, tables, and typewriters—as such. It seems less plausible to say that we sense these things. If we restrict sensation to those aspects of perception that have phenomenal character, then it is natural to conclude that we cannot represent anything that makes no phenomenal difference. Sensing a real tiger and a Hollywood prop tiger can be qualitatively alike. The content of those sensations must, therefore, be alike as well. And, conversely, if two sensations are phenomenally different, they must have different content. A sitting tiger looks different from a prancing tiger, and that difference in looks is a difference in our sensations. Such sensations must differ in content. Using the terminology from above, it seems that sensations represent appearances. They represent colors shapes, textures, and sounds from particular vantage points.

This is an argument for the Insensibility Thesis. If sensations represent appearances, then they cannot represent anything abstract. Indeed, they cannot represent some things that are not maximally concrete. Things are concrete to

the extent that they appear alike. Tigers appear alike, but not maximally so. Only appearances themselves have this guarantee. The appearance of something from a particular vantage point is the power it has to cause perceptual representations in a viewer from that vantage point. Two appearances are type-identical if they are powers that could cause identical perceptual representations. The argument that we can perceive only appearances goes as follows:

- P1. Sensations represent only things that make a phenomenal difference.
- P2. Appearances are the only things that make a phenomenal difference.
- C. Thus, sensations represent only appearances.

One might try to challenge this argument by rejecting the second premise. One might argue that something other than appearances contributes to phenomenology (see Siegel, this volume, ch. 14). I don't find this plausible. I suspect that most of our perceptual representations are picture-like (or like sound recordings or textures, etc.). In vision, that means perceptual representations capture things like shapes and colors. To say that phenomenology goes beyond appearances is like saying that phenomenology includes features that could not be captured by a camera or a tape recorder. Is there any reason to think that this is the case?

Philosophers who believe that phenomenology outstrips appearances typically appeal to introspection. They point to examples in which we are invited to recognize that the same appearances can be phenomenally different. Block (1995) gives the example of hearing words. If a monolingual English speaker hears the German word *Hund* it is just a sound. If a German speaker hears it, the experience is phenomenally different. Does this show that phenomenology goes beyond appearances? I think not. The problem is, if we grant the two listeners have different phenomenal experiences, we can explain those differences in terms of further appearances. The German speaker may form a visual image of a dog, or an auditory image of an associated word; she may have an emotional feeling of familiarity or recognition; she may form spontaneous motor images of dog-related behaviors. Hearing familiar words triggers a cascade of responses, documented through a thousand lexical priming studies, which may impact phenomenal experience.

This response reveals a serious danger of appealing to introspection when arguing for differences in phenomenology. Introspection may do a reasonably good job of telling us when two phenomenal experiences differ (especially if one can experience them consecutively, which is difficult in Block's example), but introspection is less accurate when it comes to telling us what such differences consist of. This fact is precisely what undermined introspectionism as a methodology in psychology. Wilhem Wundt and his followers insisted that all experiences were constituted by images. Dissenters, such as Oswald Külpe and Robert Woodworth, claimed that we could have imageless thoughts. Roughly, the debate was about whether phenomenal experiences included features that

were not specific to any sensory modality. Both sides used the same experimental methods: they asked subjects (often serving as subjects themselves) to do some mental task and then report on whether any imagery had been used. For example, they asked subjects to form the intention to move, and report what it felt like. Unfortunately, the different camps got correlatively different results (see Woodworth, 1906). Defenders of imageless thought got subjects to claim that they were not experiencing any images when they formed the intention to move, and opponents of imageless thought got subjects to explain every aspect of their experience imagistically. The intention to move might be reported as a tensing of the muscles or as a kinesthetic image of a movement. With no way to adjudicate between these sides, the methodology had to be abandoned. The moral is that intuitions about what kinds of representations are used in phenomenal experiences are unreliable. Those who use introspective evidence to support the claim that phenomenology outstrips appearances are offering too little.

In the debate about imageless thought, I think we have independent reason to favor Wundt. In particular, I think we have reason to think that all the mental representations used in thought are modality specific. There are no amodal symbols, no common code, and no language of thought (Prinz, 2002; Barsalou, 1999). If this is right, then no thought, conscious or unconscious, is imageless. Some of the evidence for this claim comes from the fact that we find modality-specific representations throughout the brain. There is no center of thinking or cognitive engine that hovers above the input systems. So-called higher cognitive structures often contain modality-specific subregions, and any regions that have claim to being polymodal typically play an auxiliary role in thinking; they reactivate modality-specific regions of the brain or coordinate transfer of information between the senses. In addition, when we look at brain activity during cognitive tasks, we see modality-specific activation. For example, when we ask people to form the intention to move, areas of motor cortex (unsurprisingly!) are especially active. If thought were imageless, we might expect to see loci of activation in a modality-neutral region, which could be regarded as the central executive that makes decisions about behavior before issuing motor commands.

These remarks about modality specificity also bear on the question of whether phenomenology outstrips appearances. They offer a way of explaining the intuition that favors that hypothesis. When you look at a table, it doesn't seem that you are merely experiencing colors and shapes. You seem to be experiencing tablehood. Phenomenologically, it seems like a table experience. I have been trying to suggest that such intuitions are unreliable, but I also think they can be explained. If empiricism is right, then our concept of tables comprises stored images of tables. So, if we were to represent the property of tablehood in thought, we would use an image. The images available to us in sensation may seem to represent objects (rather than mere appearances) by virtue of being very much like the representations used to represent objects. When we form

*judgments* about our phenomenal states, we assume that they represent objects, but the phenomenal states themselves may represent nothing but appearances. This undercuts any attempt to refute P2 in the argument for the Insensibility Thesis. Empiricism can help explain away the persistent belief that phenomenology outstrips appearances.

Does this mean that the Insensibility Thesis is right? I am inclined towards a negative answer. The culprit is P1, not P2.

## 2.2 Deep Sensations and the Duality of Content

I argued against the Imperceptibility Thesis by arguing that perception involves matching incoming percepts to stored percepts. Under certain conditions, the incoming percepts can take on the semantic content of the stored percepts. The stored percepts can represent abstract properties and, therefore, so can the incoming percepts. This argument hinges on the claim that perceiving involves recognizing. The matching process is a recognition process. Sensation does not require recognition. Indeed, I think that sensation takes place at a level of processing within our input systems that typically precedes the level at which recognition is achieved (Prinz, 2000a). So there is no trivial way to adapt the argument against the Imperceptibility Thesis to the Insensibility Thesis. Another strategy is needed.

My skepticism about the Insensibility Thesis stems from empirical work on perceptual systems in the brain. A little background will help. Perceptual systems are hierarchically organized. Low-level perceptual subsystems capture information about local features of a stimulus. In vision, for example, low-level subsystems encode information about small edges derived from discontinuities in light. Intermediate-level subsystems integrate local features together into more coherent representations. In vision, intermediate-level representations encode information about contours and surfaces. They achieve some degree of color constancy, superimpose illusory contours, and encode information about depth derived from binocular disparity. No level of processing in the visual stream corresponds more faithfully to what we report in conscious experience (Jackendoff, 1987; Prinz, 2000a). High-level perceptual subsystems use representations that are invariant across a range of perceiving conditions. In vision, high-level representations are often relatively invariant across orientations, scale changes, and even changes in handedness (whether a feature of a stimulus is on the right or the left). These invariant representations are ideally suited for object recognition. If perception were to be located within the visual stream, high-level areas (in inferotemporal cortex) would be the best bet. Sensation, on the other hand, is better identified with intermediate-level representations.

So far, this offers little hope for the opponent of Insensibility. If sensation precedes recognition, then it precedes the stage at which representations of

abstract properties are recruited for matching. There may, however, be a role for abstract representations in sensation. It is well known that pathways in our input systems travel both forwards, from our sensory receptors, and backwards, from higher perceptual areas. In the visual system, for example, there are massive back projections from inferior temporal cortex, associated with high-level vision, back into prestriate areas, associated with intermediate-level vision. The exact function of these back projections is unclear, but there are some reasonable hypotheses. One of the most plausible suggestions is that back-projections are used to form mental images (Kosslyn, 1994). In imagery, we reactivate perceptual systems using representations stored in memory. High-level representations, which can be relatively schematic, are used as instructions for reforming rich sensory arrays. The images that result inherit their meaning from the concepts or words that guide their formation. If you form an image of George Washington, it will represent Washington by virtue of having been drawn up from a mental file that has the function representing Washington. Mental images almost certainly re-engage intermediate-level perceptual areas, because images are often conscious and consciousness arises at the intermediate level.

Back-projections are clearly used during mental imagery tasks, but there is also physiological evidence that back-projections are active during conscious visual perception (e.g., Lamme and Roelfsema, 2000). One explanation is that imagery is used during perception to improve performance (Kosslyn, 1994). Information coming through our sensory receptors is often degraded. In vision, the objects around us are often occluded, poorly illuminated, insufficiently foveated, or moving too quickly to be adequately perceived. We take in enough information to make good guesses about object identity, but doing so often requires that we fill in information that has been lost. To do this, we use available information to call up less degraded perceptual representations from memory, which can be back-projected into earlier visual areas to enhance the signal. As Plato notes in *The Republic*, our input systems allow us to make guesses as to what we are perceiving, and those guesses then help us perceive. A good guess can also tell us where to saccade or focus attention, increasing the chances of picking up the details relevant for recognition. In sum, Kosslyn speculates that perception and imagery work in concert. Something like this may happen all the time. Even when an input is clear, high-level perceptual centers may send back efferent signals for other purposes. Back-projections result in a “re-entrant” processing loop, which may have computational advantages, just as back-projecting connectionist nets can outperform feed-forward nets (Edelman, 1993). Projecting images backwards can also facilitate binding. Treisman’s (1998) influential theory of attention postulates high-level feature maps, which help us associate shapes with colors in vision. When a visual stimulus is presented too quickly, we sometimes bind the colors to the wrong shapes. Longer exposure allows us to produce more stable high-level representation and project them back into earlier processing subsystems.

Bringing this together, perception may work in the following way. After transduction, a signal is propagated through a hierarchically organized sequence of subsystems, which begin by producing representations of local features and move on to representations that are more global and invariant. Sensations occur at the intermediate stage, between these two extremes. When a high-level representation is formed, it is matched against representations stored in memory. This process sometimes occurs before the perceptual system can fully discern a stimulus—a kind of perceptual guessing. The resulting high-level representation, whether a mere guess or not, is back-projected into earlier perceptual subsystems. This improves the signal, facilitates binding, and allows for selective, strategic processing.

After back-projection, sensations are no longer purely bottom-up. They are blends of incoming signals and mental images produced by centers further down the processing stream. The representations used downstream, which have been matched with representations stored in memory, can inherit semantic properties from those stored representations. This was one of the points that I tried to defend in the discussion of perception. Now I want to suggest that the semantic properties of high-level perceptual representations that have been matched against stored perceptual representations can be exported to earlier processing levels. Sensations can take on new meaning once they intermingle with representations coming down from on high.

Suppose you see a dog. Initially, your sensations will represent nothing but the appearance of that dog—an ephemeral array of colors and shapes. But, in a few dozen milliseconds, recognition is achieved, and the resulting high-level representations are projected backwards. The sensation may remain unchanged, phenomenologically, but it is now a blend of incoming signals and top-down signals. Just as a mental image of a dog would represent a dog, this blended conscious percept may represent a dog. It incorporates a representation, drawn from memory, that was set up to be set off by dogs.

This account can be extended to allow for the sensation of abstract properties. If the arguments of section 1 go through, then we can perceive just about any abstract property that we encounter. In perceiving abstract properties we relate incoming percepts to stored percepts that have been set up to detect abstract properties. Those very same stored percepts can be back-projected. I can form a mental image of injustice by projecting back an image of a scene stored in the mental file by which I track injustice perceptually. The image would presumably be a depiction of an event in which something unjust had occurred. It represents injustice insofar as it is drawn up from the injustice file. I can also back-project such an image in real time, while perceiving unjust events. The result will be sensations of injustice. In this way, we can perceive properties that are extremely abstract. There may be performance constraints on what can be back-projected, but I doubt that there are any limitations in principle. The contents of sensation are no more limited, in principle, than the contents of perception. The Insensibility Thesis is false.

At this point, one might be tempted to enter an objection. When I form an image of an unjust event, say a judge upholding an unjust verdict, it does not seem to be an image of injustice. Indeed, the whole idea of an image of injustice seems bizarre. Some people have the intuition that we can consciously experience concrete objects, in addition to appearances (e.g., one can have an image of a dog), but having an image of injustice is beyond the pale. Injustice just can't be depicted. Isn't this a *reductio* of the hypothesis that I have been defending?

I think this intuition is a *reductio*, but of another thesis. It is a *reductio* of premise 1 in the core argument for the Insensibility Thesis. According to that premise, sensations represent all and only things that make a phenomenal difference. The intuition that we cannot represent injustice in an image derives from the fact that an image of injustice would be indistinguishable from an image of a certain class of appearances. An image of injustice might be indistinguishable from an image of the sound of the phrase "not guilty" as it falls from a judge's lips. Since these images look alike, there is a strong inclination to say that they represent the same thing, namely, a mere appearance. We might *infer* that the appearance is *evidence* for injustice later on in processing, but that has no bearing on the content of the sensation itself.

I think this contention is wrong, but not entirely off-base. One wants a way of capturing what images have in common, qua images. Suppose I form an image of my dog, Fido. It might be indistinguishable from an image caused by seeing my neighbor's dog, Rover. I submit that these two images differ in content. But there is clearly a sense in which they are alike. That similarity might be captured by simply saying that they feel alike, but this leaves us with the thorny problem of how to individuate and compare conscious feelings. It seems we can characterize the similarity in the two images as a similarity in content. They seem to represent the same appearances.

This invites a proposal. Perhaps sensations have two different kinds of content. On the one hand, they represent superficial appearances. On the other, they represent the deeper properties that those appearances are used to detect. I recognize Fido by his appearance. An imagistic representation of Fido represents Fido *and his appearance*. A painting of Fido also represents Fido and his appearance. It is designed to represent both. I suspect that nothing represents Fido without also representing some perceivable attribute of Fido, including the name "Fido," which represents the word "Fido" (I would argue) as well as the dog.

Now go back to injustice. An image of injustice will represent whatever appearance we are using to detect injustice on a particular occasion. But that does not prevent it from representing injustice itself. Representations have dual content. This suggestion can actually be found in the pages of Locke's *Essay* (1690). He says that ideas (which are, arguably, conscious images, for Locke) represent both nominal and real contents. The real content of an idea is, often, some deep property or essence that can be very loosely correlated with



appearances. Nominal contents are appearances; they are the superficial features that we use to identify and classify things. The nominal content of our idea of gold comprises properties such as yellowness and hardness. The real content is gold, the substance, which has an essence that was (in Locke's time) unknown. I have argued elsewhere that Locke's dual-content thesis is right for concepts (Prinz, 2000b). I have also argued that it is true for emotions. An emotion represents an appearance of bodily perturbation, and it represents an organism-environment relation that bears on well-being (Prinz, 2004). Now I am suggesting that sensations can have dual contents as well.

The first premise in the argument for the Insensibility Thesis is partially right. Sensations do represent those things that make a phenomenal difference, and those things are appearances. But the premise is partially wrong. Sensations do not only represent appearances, they can also represent deeper properties, including properties that are highly abstract. The intuition that sensations cannot represent abstract properties is mistaken. How do we know it's mistaken? The leading theory of how mental representations represent entails that sensations represent abstract properties under certain circumstances. We have independent and solid reasons for believing that sensations go beyond appearance. Thus, the intuition that they don't is mistaken. This entails that we shouldn't trust our intuitions about sensory content. And, if we cannot trust our intuitions about sensory content, then the main support for the first premise falls away. The claim that sensations represent only appearances is based on the assumption that our intuitions about what mental representations represent are completely reliable. They are not. A good psychosemantic theory must accommodate intuitions about paradigm cases, but we should expect to give up other intuitions once a good theory is selected.

If I am right, then sensations can have content that makes no phenomenal difference. When we sense an abstract property that content does not alter the character of experience. We can experience something without it affecting our experience. This is not at all paradoxical when we consider that semantic content can, in general, be determined by factors external to the representations that have that content.

### 3 CONCLUSION

I have been arguing that we can sense and perceive abstract properties. So far, however, I have been a little bit vague about the limitations on abstract sensation and perception. If I am right, then every referring concept can be represented using perceptual representations, and these can, in principle, be used in perceptual matching or sensory back-projection. Thus, the range of things we can conceive does not significantly exceed the range of things that we could perceive. Perhaps we rarely deploy abstract concepts during perception. Perhaps limits on

memory or the size of the mental-imagery buffer or other performance factors place limits on the perceivable. Perhaps concepts that get their meaning through combination, rather than reliable detection, are too cumbersome to use in our input systems. I am agnostic about these questions. My point is that the alleged representational gulf between perception and conception may be much smaller than we usually realize.<sup>1</sup>

AQ: Please check the footnote cue "1" in the text. The text has been moved as unnumbered footnote in manuscript. Please provide footnote for cue 1.

## REFERENCES

- Barsalou, Lawrence W. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Block, Ned. (1995). On a Confusion about a Function of Consciousness. *Behavioral and Brain Sciences*, 18, 227–47.
- Dehaene, Stanislas. (1997). *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.
- Desimone, Robert. (1991). Face-Selective Cells in the Temporal Cortex of Monkeys. *Journal of Cognitive Neuroscience*, 3, 1–8.
- Dretske, Fred. (1969). *Seeing and Knowing*. Chicago: University of Chicago Press.
- Dretske, Fred. (1981). *Knowledge and the Flow of Information*. Cambridge, Mass.: MIT Press.
- Dretske, Fred. (1986). Misrepresentation. In R. Bogdan (ed.), *Belief: Form, Content, and Function*. Oxford: Oxford University Press.
- Edelman, Gerald M. (1993). Neural Darwinism: Selection and Reentrant Signaling in Higher Brain Function. *Neuron*, 10, 115–25.
- Fodor, Jerry A. (1990). A Theory of Content II. In *A Theory of Content and Other Essays*. Cambridge, Mass.: MIT Press.
- Gauthier, Isabel and Logothetis, Nikos (2000). Is Face Recognition Not So Unique After All? *Cognitive Neuropsychology*, 17, 125–42.
- Gross, Charles G. and Sargent, Justine (1992). Face Recognition. *Current Opinion in Neurobiology*, 2, 156–61.
- Jackendoff, Ray. (1987). *Consciousness and the Computational Mind*. Cambridge, Mass.: MIT Press.
- Kanwisher, Nancy, McDermott, Josh, and Chun, Marvin. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for the Perception of Faces. *Journal of Neuroscience*, 17, 4302–11.
- Kosslyn, Stephen M. (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, Mass.: MIT Press.
- Kreiman Gabriel, Koch Christof, and Fried Itzhak. (2000). Category-Specific Visual Responses of Single Neurons in the Human Medial Temporal Lobe. *Nature Neuroscience*, 3, 946–53.
- Lamme, Victor A. F. and Roelfsema, Pieter R. (2000). The Distinct Modes of Vision Offered by Feedforward and Recurrent Processing. *Trends in Neurosciences*, 23, 571–9.
- Locke, John. (1690/1975). *An Essay Concerning Human Understanding*; P. Niddich (ed.). Oxford: Clarendon Press.

- Millikan, Ruth. G. (1989). Biosemantics. *The Journal of Philosophy*, 86, 281–97.
- Prinz, Jesse J. (2000a). A Neurofunctional Theory of Visual Consciousness. *Consciousness and Cognition*, 9, 243–59.
- Prinz, Jesse J. (2000b). The Duality of Content. *Philosophical Studies*, 100, 1–34.
- Prinz, Jesse J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, Mass.: MIT Press.
- Prinz, Jesse J. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.
- Prinz, Jesse J. (forthcoming). *The Emotional Construction of Morals*. Oxford: Oxford University Press.
- Putnam, Hilary. (1975). The Meaning of “Meaning.” In K. Gunderson (ed.), *Language, Mind, and Knowledge* (pp. 131–93). Minneapolis: University of Minnesota Press.
- Treisman, Anne. (1998). Feature Binding, Attention and Object Perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 353, 1295–1306.
- Williams, Bernard. (1985). *Ethics and the Limits of Philosophy*. Cambridge, Mass.: Harvard University Press.
- Wittgenstein, Ludwig. (1953). *Philosophical Investigations*; G. E. M. Anscombe (trans.). Oxford: Blackwell.
- Woodworth, Robert S. (1906). Imageless Thought. *The Journal of Philosophy, Psychology, and Scientific Methods*, 3, 701–8.

AQ: Please  
update the  
reference Prinz,  
Jesse J.  
(forthcoming).