Inner Sense

Vincent Picciuto and Peter Carruthers

This chapter considers whether any of the inner-sense mechanisms that have been postulated to detect and represent some of our own mental states should qualify as sense modalities. We first review and reject the four standard views of the senses, and then propose a set of properties that would be possessed by a prototypical sensory system. Thereafter we consider how closely the existing models of inner sense conform to the prototype. Some resemble a prototypical sense to a high degree, some much less so.

1. Introduction

A number of theorists have proposed the existence of an *inner* sense modality. According to some of them, the faculty of inner sense both represents certain mental states and explains how they are phenomenally conscious (Armstrong, 1968; 1984; Lycan, 1996). These forms of theory purport to explain how it is that perceptual states acquire a dimension of phenomenology or "feel". It is held that they acquire such properties by being detected and represented through the operations of an inner sense. Other proponents of inner sense, in contrast, are somewhat less ambitious, and make no attempt to solve the "hard problem" of consciousness. Rather, they appeal to a faculty of inner sense to explain how it is that we have privileged and authoritative access to certain of our mental states, including both perceptual states and propositional attitudes (Nichols and Stich, 2003; Goldman, 2006).[1] Here inner sense is deployed to explain how we have a certain sort of knowledge of ourselves, not to explain

[1] Although Nichols and Stich (2003) defend the existence of at least two introspective monitoring mechanisms (for identifying perceptions and attitudes respectively), they do not themselves use the language of "inner sense" to describe their view. However, since the mechanisms they describe are functionally equivalent to what might plausibly be taken to be forms of inner sense, we propose to consider their views alongside inner-sense accounts that are explicitly formulated as such.

what makes mental states phenomenally conscious. It is widely believed among philosophers that people have access to their own experiences and thoughts that is both privileged (not available to others) and authoritative (unable to be challenged by others). Inner-sense theories provide one candidate explanation of these supposed facts.[2]

All of the above authors appeal to an internal faculty or mechanism that is receptive to one's own mental states—either perceptual states, or attitudinal states, or both. Moreover, the internal faculty is generally regarded as an inner *sense* modality. The notion of "sense" here is not metaphorical. For example, Goldman (2006) presents a perception-like account of inner sense, including an internal receptor–transducer system that is sensitive to neurophysiological properties; and Lycan (1996), too, regards inner sense as a genuine sense modality. The main goal of this chapter is to examine the extent to which faculties of inner sense of the sort proposed by such authors can legitimately be described as *sensory* in character.

There is one initial point of clarification. "Inner-sense theory" should, more precisely, be called "higher-order-sense theory". This is because we already have senses that operate within the body, such as interoception and proprioception, that are not intended to fall under the scope of inner sense. On the contrary, these are first-order senses on a par with vision and hearing, differing only in that their purpose is to detect properties of the body rather than properties of the external world (Hill, 2004). According to the sort of inner-sense theory that is the topic of this chapter, these internally-directed senses, too, will need to have their outputs scanned and represented to produce *higher-order* contents so that those outputs can become phenomenally conscious or be attributed to oneself. In contrast, inner sense is supposed to detect and represent some of the subject's own *mental* states, not mere internal states of the body. If we were picking our terminology afresh, we would use the term "higher-order sense." But since the use of "inner sense" is now firmly established in the literature, we will reluctantly employ it.

There are roughly four standard proposals for how to individuate the various sense modalities. To be clear, these proposals are aimed at answering the individuation question:

[2] See Carruthers (2011) for a view that denies that we have privileged and authoritative access to our own thoughts, while allowing that we have such access to some of our own perceptual experiences.

"What is it about each individual sense modality that makes it a *distinct* sense from the others?" The four standard views are direct descendants of four criteria set forth by Grice (1962). There are, though, several other significant questions about the senses. The question we take up is: "What is it about any given mental mechanism that constitutes it as a *sense* modality at all, and distinguishes it from non-sensory modes of cognition such as inference?"[3]

Our primary focus is on this second question as it pertains to the faculty of inner sense. Our topic is whether inner sense is really a sense. While the four standard proposals are accounts of sense *individuation*, we think they each suggest a corollary answer to the second question, about what constitutes a system as a sense modality at all. In section 2 we begin by reviewing these proposals with the aim of extracting a set of properties or general constraints that can constitute a prototypical sense modality. Our overall goal in this section is to articulate a theory of what it is that distinguishes sensory systems from other mental mechanisms, rather than merely to analyze our concept of a sense. We present this theory in prototype-format to facilitate judgments of degree when we come to consider whether inner sense is really a sense. Then in section 3 we discuss a range of inner-sense accounts, considering to what extent they conform to or deviate from the prototype. In section 4 we conclude.

It should be emphasized that our aim is not to issue in a categorical judgment that any of the proposed characterizations of inner sense do or do not qualify it as a sense. We do not offer the proposed criteria as necessary and sufficient conditions, but as a general model to help guide one's thinking about inner sense (or any new candidate sense, for that matter). Nor do we think the senses must individuate sharply. Rather, we think the more interesting project is to examine the ways in which inner sense is like a prototypical sense and the ways in which it is not. Our answers will all be expressed in terms of *degree*.[4]

---

[3] Few authors have taken up this question. Some notable exceptions are Heil (1983), Shoemaker (1994), and Keeley (2002). Grice (1962) acknowledges the question but swiftly proceeds to focus on the individuation issue.

[4] In fact our view is a hybrid that combines several of the criteria endorsed by the standard four views, in the manner of Keeley (2002) and Gray (2005), but with additional constraints. However, since we are not attempting to propose necessary and sufficient conditions and since we do not think there is always a sharp answer to the question whether some mental mechanism constitutes a sense, our view is more in line with those of Heil (1983) and Macpherson (2011).

We should also stress that we will not be claiming that any of the candidate mechanisms of inner sense actually exist.[5] Our question is hypothetical: if the mechanisms that have been proposed by inner-sense accounts of self-awareness exist, then to what extent are they properly characterized as *sense* modalities?

2 The properties of a prototypical sense

As many have argued, each of the four standard views on its own fails to provide necessary and sufficient conditions for individuating the senses. Similarly, each of the corollaries of the four standard views fails to provide necessary and sufficient conditions for a mental mechanism to count as a sense modality. However, we think that each standard view suggests a relevant property that forms at least part of the characterization of a prototypical sense.

Proponents of the four standard views attempt to individuate among the different senses by appeal to the relevant *sense organs*, *proximal stimulus*, *proper objects/representational features*, or *phenomenal character*.[6] Keeley (2002) and Gray (2005) argue that roughly these criteria must be combined to form jointly necessary and sufficient conditions. We discuss each criterion separately before combining some with others to provide a model of a prototypical sensory system.

2.1 Sense Organs

According to the sense-organ criterion, the senses should be individuated on the basis of the parts of the body that constitute the receptor systems for each putative sense, together with the brain regions that process information emanating from those receptor systems. For example, the cochlea and relevant parts of the brain that are used to hear are distinct from the eyes/retina and relevant parts of the brain that are used to see.

That specific sense organs play a crucial role in each of the senses is intuitive. Presented as

[5] One of us denies that they do (Carruthers, 2011), while the other is neutral on the question.

[6] Different authors use different terminology to label these criteria. Our chosen labels have been gleaned from existing literature. It is unclear that anyone defends the sense-organ criterion as such. But many discuss it before moving on to integrate it with other factors.

an individuating condition, though, the sense-organ view encounters a fundamental difficulty. For there is no one physiological mechanism that seems necessary for a sense organ to constitute any given kind of sense modality. Bee eyes are very different from typical human eyes, consisting of multiple lenses each of which is directed at a distinct region of the bee's visual field. Yet it is surely clear that bees can see. If so, then the organ is insufficient to individuate the visual sense modality. In concluding that a creature has anatomical features that qualify as eyes, we must plainly be relying on some further criterion. In particular, we need some way of isolating the function of the relevant anatomical features.

While the sense-organ criterion fails as an individuation condition for the senses, it does form a very plausible component in our idea of a prototypical sense modality, and one might think that some or other organ is a necessary condition for a mental mechanism to qualify as a sense. But one plainly needs to build more into the idea of a sense organ than mere physical mechanism. For *all* cognitive systems are presumably realized in physical mechanisms of some sort. We need to know more about the function of the putative sense organ's relevant anatomical features. What seems crucial is that a sense organ should be charged with receiving, transposing, and generating representations from some set of physical stimuli. One possible way of spelling out the function of a candidate sense organ's relevant anatomical features is to appeal to the kind of energy (or range of a kind energy) to which they are sensitive. This leads us to consider the proximal stimulus criterion.

2.2 Proximal Stimulus

Proximal-stimulus accounts claim that what distinguishes one sense from another is the specific kind of physical stimulus or energy to which the putative sense organ is receptive (Heil, 1983, 2011). That is, to be an eye is to be receptive to light waves. To be an ear is to be receptive to pressure waves. The proximal stimulus criterion is a useful addition to the sense-organ criterion. However, proximal stimuli cannot always be so neatly carved up. The eyes of bees detect electromagnetic radiation, but they detect a different range than do human eyes (the ultraviolet range as well as portions of the "visible" range). Moreover, pit vipers have organs just below their regular eyes that are capable of detecting heat (electromagnetic radiation in

the infra-red range; Gray, 2005). As a result, we can set the proximal-stimulus criterion a dilemma. On the one hand, it can be claimed that what individuates the sense of sight is a mechanism that is sensitive to a specific range of electromagnetic radiation (the "visible" range). But in that case we will be compelled to deny that bees have a sense of sight, since they can detect forms of radiation that we cannot. Yet if it is claimed, on the other hand, that what individuates sight is sensitivity to *some* (unspecified) range of electromagnetic radiation, then we will have no option but to claim that pit vipers see heat. Indeed, we will be forced to conclude that they possess, not one, but two visual systems.

Our task is to delineate the properties of a prototypical sense, however, not to provide individuating conditions for the senses. And for this purpose a combination of the sense-organ criterion with the proximal-stimulus criterion provides a very plausible component of the prototype. One factor that inclines us to judge that pigeons and trout possess a magnetic sense, for example, enabling them to navigate via the Earth's magnetic field, is the discovery that they possess distinctive magnetism-sensitive structures in their heads that serve as the organ of magnetic sense (Wallcott et al., 1979; Walker et al., 1997). Likewise, it is partly the discovery of heat-sensitive pits underneath the pit vipers' eyes that inclines us to think that it has a heat-based sense modality.

It is plain that the mere presence of a physical mechanism sensitive to some range of physical energy or set of physical properties does not constitute a prototypical sense, however. For there are a great many detection systems in the body that we might hesitate to categorize as sense modalities. Many of them play a role in bodily homeostasis, but never give rise to beliefs, and nor do they guide behavior (except indirectly). For example, there are physical structures that detect blood pressure and others that measure heart-rate, as well as numerous other receptors in the internal organs of the body (Vaitl, 1996). But people are generally at chance in attempting to judge the rate of their own heart-beat (Brener and Jones, 1974; Pennebaker, 1982). An important component of our idea of a sensory system is that the structures in question should deliver representations that guide the animal's intentional behavior, as we now discuss.

2.3 Proper Objects / Representational Content

According to this account, the senses are individuated by appeal to the kinds of objects and/or properties to which each putative sense is receptive and which its outputs represent (such as color in the case of vision or pitch in the case of sound). Some have objected, however, that there are far too many types of object that can be perceived by a given sense (Sorabji, 1971). Others have argued that the proper objects criterion fails because a *conjunction* of properties is generally given in experience, which will include properties supposedly distinctive of more than one sense (Grice, 1962). Similarly, Nudds (2004) claims that the trouble with the proper objects criterion is that it only allows us to individuate a sense when it functions in isolation. We will not pause to evaluate these criticisms here. For when our task is not to individuate the different senses, but rather to describe a prototypical sense modality, then we no longer need to be concerned with specific kinds of representational feature.

It is surely plausible, however, that any full-blown sense modality would produce representations of properties of the environment (or of the subject's own body), and that these should have a role in guiding the subject's behavior. (As we noted in section 2.2, this will mean that many detector-mechanisms within the body fail to qualify as senses to the fullest extent.) While these representations might (or might not) be distinctive enough for the individuation project, it will be necessary to include other components in a description of a prototypical sense, as we will see shortly. It will also be important to ask whether there are further restrictions on the kinds of representation in question. This issue will be taken up in section 2.6.

Some theorists attempting to provide individuation-conditions for the senses have claimed that representational content itself does not adequately characterize *what it is like* to undergo an experience produced by a given putative sense modality. This motivates some to appeal to the phenomenal character criterion.[7]

---

[7] Grice (1962) called this feature the "special introspectible character" of experiences. We take Grice here to be referring to what many philosophers now call phenomenal consciousness or, for better or worse "what it's likeness." Since we wish to remain neutral about whether all introspected states are states with phenomenal character (states that *must* be phenomenally conscious), we call this criterion the phenomenal character criterion. Nudds (2004) refers to the criterion as the "experience" criterion. However, this is too general and ignores the crucial distinction between conscious and unconscious experiences.

2.4 Phenomenal Character

According to this criterion, each sense has its own distinctive kind of phenomenal character, and it is these that individuate the senses (Grice, 1962; Leon, 1988; Martin, 1992; Lopes, 2000). It is like one thing to see a cube, but it is like quite another thing to feel a cube in one's hand. While there are current defenders of the phenomenal character account, it begs an important question when made part of a theory of what it is to have a sense *at all*. This is because higher-order theorists of consciousness will want to distinguish the senses themselves from what gives their output phenomenal character (Lycan, 1996; Carruthers, 2000; Rosenthal, 2005). Moreover, if such higher-order theories are correct, then it is far from clear that phenomenally conscious experience will be widespread in the animal kingdom. Indeed, it may be that the relevant forms of higher-order representation are unique to humans, or perhaps restricted to primates (Carruthers, 2000). If so, then possession of phenomenal character will fail to be a prototypical property of sense modalities in general. For these are arguably possessed by almost all forms of creature (Carruthers, 2006).

It is an open question, then, whether phenomenal consciousness is a prototypical property of a sense modality. The answer will depend on the correct form for an account of consciousness (whether first-order or higher-order), together with facts about the cognitive powers of nonhuman animals. Since for present purposes we do not wish to rely on the truth of a higher-order account, we will include possession of phenomenal character in the discussions to come, noting that its inclusion is controversial.

2.5 A Combined Properties Account

Our examination of the four standard views of sense-individuation has led us to an initial sketch of a prototypical sense modality. The latter should consist of a physical organ that is sensitive to some range of proximal stimuli, producing representations (perhaps phenomenally conscious representations) that serve to guide the organism's intentional behavior. In the present section we examine the account provided by Keeley (2002), who is one of the few theorists to focus on the question of the conditions under which a mental mechanism counts as sensory. This will

enable us not only to confirm, but to elaborate and extend, our account of the prototypical sense. Keeley's goal, however, is to provide a set of individually necessary and jointly sufficient conditions for a system to count as a sense. Our own goal is weaker. It is to describe the components of a prototypical sense. Our view is that there are likely to be numerous systems in the natural world that count as senses to some degree (of which mechanisms of inner sense might constitute good examples), and that the interesting question is *the extent to which* these systems approximate to a prototypically sensory one.

Keeley's account combines some of the criteria discussed above with others. He eliminates phenomenal character as a necessary condition for a sense modality. But he does so for different reasons than those discussed in section 2.4. (He is concerned with troubles posed by "qualia", whereas we think that these can be explained naturalistically; Carruthers, 2000; Picciuto, 2011.) But in positive mode, he presents the following four conditions: (i) Physics, (ii) Neurobiology, (iii) Behavior, and (iv) Dedication. We will discuss them in turn.

Keeley's physics condition is roughly the proximal stimulus criterion considered in section 2.2. It specifies the external physical stimulus to which a putative sense is sensitive, thereby fixing the "space of possible modalities."[8] Gray (2005) points out a problem for the physics condition, however. This is that, as the pit viper case shows, there might be good reason to suppose that some distinct senses will be sensitive, not to a distinctive kind of energy (e.g. electromagnetic radiation), but to a distinct *range* within a kind of energy (e.g. infrared radiation). Gray therefore argues for a modification to the physics condition. We agree, and will thus assume that this condition should be stated in terms of *ranges* of physical energy of a distinctive kind, or *sets* of related physical properties (think of taste or smell).

The neurobiology condition is a more detailed account of the sense-organ criterion. According to Keeley, a legitimate sensory organ must have three characteristics. It must physiologically respond to a naturally occurring range of physical stimulation; it must be wired up properly to the central nervous system; and it must include an "end organ," from which the informational signals to the central nervous system initiate (2002, 14). So while the organ itself is not sufficient to constitute the existence of a sense, some kind of organ of this general sort is

[8] "External" should be taken to mean external to the sense in question, not external to the body or brain.

at least necessary.

Keeley's third criterion is "behavior." An organ that detects a specific kind of proximal stimulus must enable the organism to discriminate behaviorially between stimuli that differ only in terms of a particular physical energy type. The third criterion is supposed to address a problem that arises when considering only the first two criteria. This is the problem of vestigial sense organs. Roughly, the problem is that a sense organ sensitive to a range of physical stimuli might meet conditions (i) and (ii) but still not suffice to constitute a sense, because one may have such an organ but never make use of the information it processes. An implicit condition on a sense, then, is that the organism actually makes use of the information it generates. We agree; for as we noted in section 2.2, there are numerous detector mechanisms within the body that one might be reluctant to characterize as senses. This is because, while the information they deliver gets used by some system or other, that information is not available to guide the intentional behavior of the organism. A prototypical sense, then, should involve an action-guiding component.

We need yet another condition, according to Keeley (2002), because a sensory organ can fulfill the first three conditions without enabling the organism to *perceive* those stimuli. The last condition, "dedication," is supposed to provide the final component. Dedication is "the evolutionary or developmental importance of the putative sense to an organism. For example, we ought not attribute an electrical modality to an individual unless electrical properties of the world are part of the normal environment of that individual and to which the organism is attuned" (17). An eye might "detect" mechanical energy in that it will respond to pressure, say, but it is "receptive" only to electromagnetic stimulation, because it has evolved specifically to enable the organism to discriminate that sort of information.

We are happy to accept the dedication condition as providing one component of a prototypical sense modality. However, Keeley's four conditions are not complete. A prototypical sense should also issue in fine-grained (or nonconceptual) representations with mind-to-world direction of fit, as we explain section 2.6.[9]

[9] While we use the language of nonconceptual content throughout, this is for convenience only. For everyone allows that sensory representations are distinctively different from those employed in thought. Even McDowell

2.6 Nonconceptual World-Directed Representations

While we noted in section 2.4 that it is highly controversial to include the phenomenal character criterion in an account of a prototypical sense, there is a related, less contentious, suggestion. For one distinctive property of phenomenally conscious states is that they have a special sort of fine-grained, nonconceptual, content. But presumably the perceptual states of nonhuman animals possess such content, even if they aren't phenomenally conscious. Moreover, these contents have mind-to-world direction of fit. (If there is a mismatch, it is the perceptions that are in error, not the world.) One might wonder, then, whether a prototypical sense modality would produce, not just *any* sort of representation that can guide the subject's behavior, but specifically *nonconceptual* forms of representation with mind-to-world direction of fit.

Suppose that by the year 2050 the "science" of parapsychology has advanced to the point that prescience is recognized to be a real phenomenon. There are some individuals, it is by then discovered, who can foretell the future in quite reliable ways. And suppose that there is some structure that is discovered in the heads of these individuals that proves sensitive to future-occurring events (perhaps as a result of quantum entanglement or some-such). But the outputs of the system are just beliefs. A prescient individual will just find herself believing that there will be a train-crash in Baltimore the next day, for example. She doesn't in any respect "see" the train crash, or experience any kind of nonconceptual representation of it (whether consciously or unconsciously). She just forms a belief. The lack of nonconceptual content in such a case would make us less inclined, at least, to describe the prescience-mechanism as a sense modality. Paradigmatic examples of existing sense modalities suggest that a prototypical sense should produce nonconceptual representations among its outputs. Perhaps the prescience-mechanism is like a prototypical sense in many other ways, but in this one important way it is not.

As for mind-to-world direction of fit, consider the structures in the nasal cavity that detect

---

(1994), for example, who denies the existence of nonconceptual content, allows that perception deals in a special class of fine-grained indexical concepts. If necessary, our discussion could be couched in such terms.

and respond to pheromones. It is presently unclear exactly what role these play in human cognition, but in other animals they modulate sexual attraction and other forms of affective behavior (Dulac, 2000). Suppose it is discovered, then, that the role of the pheromone-detection system is to make any opposite-gendered person to whom one is attending at the time seem to some degree sexually attractive or repulsive. In short, suppose that pheromones issue directly in feelings of desire. This, too, would surely disincline us to regard the pheromone system as a sense modality, despite the fact that it contains an organ that is sensitive to a specific sort of physical stimulus, and despite the fact that it produces nonconceptual representations of attractiveness as output. For even though, at the input level, there are representations with mind-to-world direction of fit, the representations that the system produces as output have the wrong direction of fit. They don't represent the world as being a certain way. Rather, they issue in feelings of desire or repulsion. Prototypical sensory systems are for representing the world, not for changing it (or not directly, anyway).

2.7 The Prototypical Sensory System

Pulling together all of the ideas discussed above, we suggest that a prototypical sense modality will: (1) be sensitive to some range of physical energy or set of related physical properties, (2) include a detector mechanism that transduces that energy or those properties into informational signals sent to the central nervous system where (3) they are used to guide the intentional behavior of the organism (perhaps issuing in phenomenally conscious sensations). In addition, a prototypical sense will (4) have as its evolutionary function the detection and representation of the physical energy or properties in question, and (5) will issue in nonconceptual representations with mind-to-world direction of fit. While a full account of a prototypical sensory system would no doubt need to include some specification of the comparative importance of each component, a simple listing of the components will be sufficient for our purposes here.

3 To what extent is inner sense a sense?

In the present section we consider the extent to which different models of the inner-sense

faculty match the account of the prototypical sense modality outlined in section 2.

Inner-sense theory was first proposed by Locke (1690), but the view has been defended more recently as a theory of phenomenal consciousness by Armstrong (1968, 1984) and Lycan (1996), and as a theory of self-knowledge by Nichols and Stich (2003) and Goldman (2006). Recall that while both types of theory are inner-sense theories, they have quite different explanatory goals. Armstrong and Lycan use it to construct a theory of phenomenal consciousness, whereas Nichols and Stich and Goldman use it to construct accounts of self-knowledge. This difference is not relevant to our aims in this chapter (except insofar as it impacts the shape of the theories in question). For we are concerned with the ways in which inner sense resembles a prototypical sense modality (and the extent to which it does so), irrespective of distinct theoretical applications.

Each of the three most recent accounts has somewhat different implications for our question. We will discuss Lycan in section 3.1, Nichols and Stich in section 3.2, and Goldman in section 3.3.

3.1 Scanned Sensory-Output Models (Lycan)

According to Lycan (1996), humans (and perhaps some nonhuman animals) not only have sense-organs that scan the environment or body to produce fine-grained representations that can serve to ground thoughts and action-planning, but they also have *inner* senses, charged with scanning the outputs of the first-order senses to produce equally fine-grained but higher-order representations of those outputs (allegedly rendering the latter phenomenally conscious).

On this account, inner sense is a perception-like faculty, and is presumed to include a receptor–transducer system of some sort. Plainly, too, its outputs can guide intentional behavior (enabling people to make reports about their conscious experiences, for example). Moreover, those outputs have fine-grained nonconceptual contents with mind-to-world direction of fit (although in this case "the world" comprises the outputs of the first-order senses that are targeted by inner sense). In addition, since inner sense is presumed to have the *function* of representing the outputs of our first-order senses, it appears that on this account

inner sense resembles a prototypical sensory system to some quite high degree.

Note, however, that on this account the mechanism of inner-sense does not issue in outputs that are themselves phenomenally conscious. According to Lycan (1996), it is the outputs of our first-order senses that become phenomenally conscious when represented by inner sense, whereas the latter representations (the *outputs* of inner sense) are *not* phenomenally conscious. For this would require that they, in turn, were detected and represented by some sort of third-order inner-sense mechanism. If the account were intended merely as a theory of self-knowledge, however, rather than an account of phenomenal consciousness, then there would be nothing to prevent one from claiming that the outputs of the sensory monitoring mechanisms are themselves phenomenally conscious (except that there seems little introspective support for such a view).

One might worry, moreover, that the proposed receptor–transducer system would be too fragmentary and distributed to qualify as a sensory organ. For the various first-order sensory systems are realized in quite different areas of the brain. So it might seem that multiple receptor systems would need to be involved, rather than just one. But it is a mistake to think that this would constitute any sort of problem for the view that inner sense is genuinely a sense. For our tactile sense, too, comprises a great many receptors of a number of different kinds distributed over the surface of the body. So it can't be a requirement on a sensory organ that it should comprise a unitary localized structure. Moreover, although neither Armstrong nor Lycan develop their views in this way, it might be said that the receptor mechanism of inner sense has to wait on the "global broadcast" of attended sensory representations (Baars, 1988, 1997). Since global broadcasting results in sensory representations being made widely available to other systems within the brain irrespective of sensory modality, this would mean that the inner-sense receptor mechanism could be a single local structure after all.

More significantly, one might worry that there is nothing physically distinctive about the properties that an inner-sense organ would be designed to detect. For one might question whether the neural properties of our perceptual states differ from those that the realizing mechanisms for *any* cognitive process would need to be sensitive to (where we wouldn't be tempted to talk of a sense). Ultimately all cognitive processes need to be realized in neural

ones, and so all cognitive mechanisms will need to be sensitive to relevant properties of other such processes when taken as input. And one might expect that the mechanism involved in inner sense would pick up on the very same set of properties.

Consider, for example, the mechanism in humans and other animals that estimates the numerosity of a set from a perceptual representation of it. Since this mechanism is physically realized in the brain, it must at some level of description be responding to physical properties of the neural signals that code for the presence of the set in question. If these are the same kinds of properties that the inner-sense mechanism is sensitive to, then we face a dilemma. Either the inner-sense mechanism lacks a sensory organ (because there is nothing distinctive about the physical properties detected), or sensory organs will turn out to be rampant in the brain, existing at every physical interface between one cognitive system and another.

It might be possible to reply to this difficulty, in part by utilizing the idea of global broadcasting once again. For suppose this turns out to have a distinctive physical signature in the brain (perhaps involving synchronized neural oscillations in the high-frequency range 40-150 Hz; Rees et al., 2002). And suppose, too, that each kind of neural process that realizes the outputs of the various first-order sense modalities includes some modality-distinctive physical signature. Then one might envisage an inner-sense mechanism with a complex component structure. One component would be sensitive to those physical properties of a perceptual experience that realize its representational content. (This would be no different, in this respect, from any other cognitive mechanism.) But the others would be sensitive to more widespread neural oscillation frequencies in the brain, as well as to the physical properties distinctive of each sense. On the assumption that the latter are not detected as such by other mental mechanisms, then this would serve to distinguish the organ of inner sense from other cognitive systems after all.

It might also be possible to reply to the difficulty more directly. For perhaps one can draw distinctions among cognitive systems regarding the level at which it is appropriate to describe their inputs. While all mental mechanisms must be sensitive to neural properties at *some* level of description, in many cases it might be more appropriate to describe them as sensitive to contentful or computational–syntactic properties instead. If this is so for the vast majority of

cognitive systems, but *not* for inner-sense mechanisms, then this would vindicate the idea that the latter (and only the latter) contain physical transducers that respond to physical properties of mental states as such. However, it is unclear to us whether and how such a distinction might be justified. In any case deep questions are raised about how one should select the appropriate level of explanation in cognitive science.

We conclude that the model of inner sense proposed by Lycan (1996) resembles a prototypical sensory system to some quite high degree. It is an evolved system, it generates nonconceptual representations with mind-to-world direction of fit, and these can guide the subject's intentional behavior. It is possible, too (albeit much more problematic), that the receptor component of the mechanism might count as a sensory organ that detects a distinctive range of physical properties.

3.2 Belief-Output Models (Nichols and Stich)

Nichols and Stich (2003) argue that there are a number of distinct monitoring mechanisms at work in the human mind–brain. There is at least one such mechanism for monitoring one's own perceptual experiences, issuing in knowledge that one is seeing or hearing something, for example. And there are other such mechanisms for monitoring one's propositional attitude states, especially beliefs and desires. They think that the mechanisms for monitoring perceptual states will need to be quite differently structured from those that monitor propositional attitudes. We will discuss their account of perceptual self-knowledge first, before turning to the case of belief.

Nichols and Stich suggest that the perceptual monitoring mechanism would need to possess a complex internal structure. For it needs to receive nonconceptual representations of aspects of the environment or body as input while delivering fully-conceptual beliefs as output, such as, "I am seeing the color grey" or, "That looks like a rock." In our view they overstate the need for internal complexity. For many of the accounts of perception constructed by cognitive scientists suggest that the output of perceptual systems is already partly conceptual. While vision, in particular, issues in fine-grained nonconceptual representations of colors, textures, and shapes, conceptual representations are also bound into our visual percepts as a result of visual

processing (Kosslyn, 1994). Hence one sees something *as* a rock or *as* one's mother's face, for example. So the mechanism that constructs higher-order awareness that one is seeing a rock would just need to redeploy one of the concepts embedded in the perceptual state that it receives as input, while also determining that *seeing* is the appropriate modality (perhaps on the basis of cues like color).

Whether Nichols and Stich's perceptual monitoring mechanism is simple or complex, however, it is plain that it has few of the properties of a prototypical sense modality. On the plus side, it does produce representations with mind-to-world direction of fit that can guide the subject's behavior, and it might well have been designed to do so. But the representations in question are higher-order beliefs, not nonconceptual perception-like states. Moreover (as we will see in connection with Nichols and Stich's postulated belief-monitoring mechanism), there is no suggestion that the perceptual monitoring mechanisms contain anything like an *organ* designed to detect some form of physical energy or set of physical properties. On the contrary, the mechanisms are thought to be syntactic in nature (or quasi-syntactic, in the case of perceptual monitoring). Like other inferential mechanisms in the human mind–brain, they are held to respond to syntax-like properties rather than to brain states as such.

Nevertheless, despite the fact that the outputs of the mechanism are mere beliefs (and hence are in this respect not perception-like), they might on some views be phenomenally conscious. (Nichols and Stich themselves are silent on this issue.) This is one respect in which it matters that the account is designed just to explain our knowledge of our own perceptual states, and not to explain phenomenal consciousness. For this leaves a defender of inner sense free to endorse the views of Strawson (1994), Siewert (1998), Pitt (2004), and others, claiming that thoughts as well as experiences can be phenomenally conscious.[10] In particular, it can be claimed that the thoughts about one's own experiences that are a product of the inner-sense mechanism are sometimes phenomenally conscious.

Turning now to the postulated monitoring mechanisms for beliefs and desires, these are thought by Nichols and Stich (2003) to be simpler than the monitoring mechanism for perceptual states. Thus the belief-monitoring mechanism is designed to take any of one's

[10] For arguments to the contrary, see Tye and Wright (2011) and Veillet and Carruthers (2011).

current beliefs as input and to embed that representation into the content of a belief attribution. For example, taking as input the judgment that global warming is real, it embeds whatever syntax-like structure carries the content, *global warming is real* into a higher-order representation with the content, *I believe that global warming is real*. The desire-monitoring mechanism operates similarly. But in both cases it is plain that the mechanisms in question would have few of the properties of a prototypical sense modality.

While these mechanisms have the function of delivering representations with mind-to-world direction of fit that can guide one's intentional behavior, the representations are not at all perception-like (although they might, on some views, be phenomenally conscious), and the mechanism that generates these representations is not much like a sensory organ. While these inner-sense systems might have a complex internal structure, with attitude-detector components located in the varied regions of the mind–brain where beliefs and desires are produced, they are sensitive to syntactic properties of the input, not neural properties as such. Hence what is detected by these monitoring mechanisms won't be physically distinctive, but will be the same as the properties that are detected and deployed by any other inference mechanism. So inner sense, on this account, will only resemble a prototypical sensory system to some smallish degree.

3.3 Mixed Models (Goldman)

Goldman's is perhaps the most well-developed account of inner sense. On this view, introspection is a meta-representational process involving "recognition, redeployment, and translation" (2006, 254). Like first-order sensory systems, Goldman claims that inner sense involves a "transduction" process, taking neural properties of mental states as input. In response to those inputs, the transducer produces a representation in a proprietary code. Goldman calls this the introspective code, or the "I-code." The I-code operates in the language of thought and encodes certain properties of a given mental state that correspond to various aspects of stored concepts. Goldman isn't fully confident of exactly which properties are encoded in any given I-code representation. However, he speculates that an I-code representation of the attitude-type *hope*, say, will include its general type (that it is an attitude

as opposed to a perception), a doxastic parameter with degrees of doubt or uncertainty as its values, together with a valence parameter with degrees of desire or aversion as its values. According to Goldman, these properties are then matched to the stored concept, *hope*, and thus the attitude is recognized.

Simultaneously with typing the kind of mental state in question, the inner-sense faculty also proceeds to self-attribute the content of the state, according to Goldman. It does this either by redeploying the content of the detected state into a meta-representation (in cases where the state in question is an attitude), or by "translating" it into conceptual form (where the state in question is a perceptual one). For example, it redeploys the content of a state of hoping that it will be sunny (namely the content, *It will be sunny*) into the content of a higher-order belief with the content, *I hope that it will be sunny*. In connection with a visual perception of a sunny day, in contrast, the nonconceptual content in question first needs to be translated into conceptual form with the content, *It is sunny*, before being embedded into a higher-order belief state with the content, *I see that it is sunny*. (Note that Goldman accepts the need for translation in addition to redeployment on the same grounds as Nichols and Stich, and hence this aspect of his view is open to a similar criticism.)

To what extent does inner sense, thus conceived, resemble the prototype of a sense modality? Consider Goldman's account of self-knowledge of attitude states first. One important point to stress (which distinguishes the account from that of Nichols and Stich) is that the output of the mechanism is a mixed conceptual–nonconceptual representation. For example, while the output in a given case conceptually represents the content of a detected belief state (while also conceptually representing that it is a belief that is detected), the output of the mechanism will additionally represent the degree of certainty with which the belief is held, as well as the degree of positive or negative valence that attaches to it. In this respect, at least, the inner-sense mechanism seems very much like a prototypical sensory system. For the latter, too, generates mixed conceptual–nonconceptual representations, as we noted in section 3.2. And in both cases the representations produced have a mind-to-world direction of fit, and in both cases the output can be phenomenally conscious.

While Goldman does not say so explicitly, he can presumably maintain that the inner-sense

mechanism has the evolutionary function of detecting and self-ascribing mental states. Moreover, he tries hard to make it seem that the mechanism in question would qualify as a sense *organ* of some sort, transducing neural properties into I-code representations. At this point, however, Goldman (2006) confronts essentially the same difficulty as Lycan (1996). For at some level of description *all* mental mechanisms need to be sensitive to neural properties, since all cognitive processes are realized in neural ones. Much may depend, then, on the detailed structure of the inner-sense mechanism, or on whether a distinction between cognitive mechanisms that detect syntactic properties and those that detect neural ones can be cashed out, as we noted in section 3.1.

Goldman's account of self-knowledge of perceptual states, in contrast, is in one way more like that of Nichols and Stich than it is like Lycan's. For there seems to be no scope for nonconceptual representations of doxastic strength in a higher-order representation of what one sees or hears. Rather, the inner-sense mechanism will produce a conceptual representation with the content, *I see that it is sunny* (say). However, if valence representations as well as conceptual representations can be embedded into the content of perceptual states, as many think, then the valence-parameter in the higher-order I-code representation will generally be assigned some value. In that case there will at least be a *dimension* of nonconceptual content in the output of the inner-sense mechanism as it operates on perceptual states. And so to this extent the output can be phenomenally conscious, even without needing to rely on the views of those who think that thoughts *per se* can be phenomenally conscious (Strawson, 1994; Siewert, 1998; Pitt, 2004).

It seems, then, that on Goldman's (2006) account the mechanism that monitors perceptual states resembles a prototypical sensory system less closely than does the mechanism postulated by Lycan (1996). In contrast, Goldman's account of the mechanism that monitors attitude states resembles a prototypical sense *more* closely than do the mechanisms postulated by Nichols and Stich (2003).

4 Conclusion

If the account of a prototypical sense modality outlined in section 2 is correct, then there are

likely to be a range of possible inner-sense mechanisms that conform more or less closely to the prototype. In section 3 we have found that existing theories of inner sense are distributed along just such a spectrum. Each of these theories is somewhat underspecified. This makes it difficult to judge, in some cases, whether the prototypical properties are present. But we can confidently conclude the following:

All of these postulated mechanisms are perception-like in the minimal respect of having as their function the creation of representations of a certain sort (higher-order representations), which have mind-to-world direction of fit, and which can guide one's intentional behavior.

Moreover, the mechanisms postulated by Lycan (1996) and Goldman (2006) are also perception-like in that their outputs are either nonconceptual or (like other forms of perception) mixed conceptual–nonconceptual representations. In addition, if these mechanisms aren't intended to account for the phenomenally consciousness status of the states that they target, then it can be claimed that their outputs are phenomenally conscious. They therefore conform to the prototype of a sense modality to some significant degree—indeed, to the extent that it would not be misleading, at least, to describe inner sense as a sense.

It also remains possible that the internal structure of the systems postulated by Lycan and Goldman might include components that differ from those employed in other sorts of cognitive system, in such a way that they detect, as such, the neural properties that realize our mental states. If so, then these would qualify as fully-fledged sensory systems (or almost so, since on some views phenomenal qualities might be lacking in the output), and inner sense would turn out to be a genuine sense modality.

The mechanisms postulated by Nichols and Stich (2003), in contrast, fare much less well as putative sense modalities. While they produce representations with mind-to-world direction of fit that can guide one's intentional behavior, they only create conceptual representations, not nonconceptual ones. Moreover, because they respond to syntactic properties rather than neural ones, they fail to include anything resembling a sensory organ or transducer mechanism. So these mechanisms, if they exist, only remotely resemble a prototypical sensory system.

References

Armstrong, D., (1968). *A Materialist Theory of the Mind*. London: Routledge.

Armstrong, D., (1984). Consciousness and causality. In D. Armstrong and N. Malcolm (eds), *Consciousness and Causality*, Oxford: Blackwell.

Baars, B. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Baars, B. (1997). *In the Theatre of Consciousness*. Oxford: Oxford University Press.

Brener, J. and Jones, J. (1974). Interoceptive discrimination in intact humans: Detection of cardiac activity. *Physiology and Behavior*, 13, 763-767.

Carruthers, P. (2000). *Phenomenal Consciousness*. Cambridge: Cambridge University Press.

Carruthers, P. (2006). *The Architecture of the Mind*. Oxford: Oxford University Press.

Carruthers, P. (2011). *The Opacity of Mind*. Oxford: Oxford University Press.

Dulac, C. (2000). Sensory coding of pheromone signals in mammals. *Current Opinion in Neurobiology*, 10, 511-518.

Goldman, A. (2006). *Simulating Minds*. Oxford: Oxford University Press.

Gray, R. (2005). On the concept of a sense, *Synthese* 147, 461-75.

Grice, H.P. (1962). Some remarks about the senses. In R.J. Butler (ed.), *Analytical Philosophy*, Oxford: Blackwell.

Heil, J. (1983). *Perception and Cognition.* Berkeley, CA: University of California Press.

Heil, J. (2011). The senses. In F. Macpherson (ed.) 2011.

Hill, C. (2004). Ouch! An essay on pain. In R. Gennaro, (ed.), *Higher-Order Theories of Consciousness*, Philadelphia: John Benjamins.

Keeley, B. (2002). Making sense of the senses: Individuating modalities in humans and other animals, *Journal of Philosophy*, 94, 5-28

Kosslyn, S. (1994). *Image and Brain*. Cambridge, MA: MIT Press.

Leon, M. (1988). Characterising the senses, *Mind and Language,* 4, 243-270.

Locke, J. (1690). *An Essay Concerning Human Understanding*. (Many editions now available.)

Lopes, D. (2000). What is it like to see with your ears? The representational theory of mind, *Philosophy and Phenomenological Research"* 60, 439-453.

Lycan, W. (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.

Martin, M. (1992). Sight and touch. In T. Crane (ed.) *The Contents of Experience*, Cambridge: Cambridge University Press.

McDowell, J. (1994). *Mind and World*. Cambridge, MA: MIT Press.

Macpherson, F. (ed.) (2011). *The Senses.* Oxford: Oxford University Press.

Nichols, S. and Stich, S. (2003). *Mindreading,* Oxford: Clarendon.

Nudds, M. (2004). The significance of the senses, *Proceedings of the Aristotelian Society,* 102, 31-51

Pennebaker, J. (1982). *The Psychology of Physical Symptoms*. Düsseldorf: Springer-Verlag.

Picciuto, V. (2011). Addressing higher-order misrepresentation with quotational thought, *Journal of Consicousness Studies*, 18, 3-4, 109-36.

Pitt, D. (2004). The phenomenology of cognition. *Philosophy and Phenomenological Research*, 59, 1-36.

Rees, G., Kreiman, G., and Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3, 261-270.

Rosenthal, D. (2005). *Consciousness and Mind*. Oxford: Oxford University Press.

Siewert, C. (1998). *The Significance of Consciousness*. Princeton University Press.

Shoemaker, S. (1994). Self-knowledge and "Inner Sense". *Philosophy and Phenomenological Research*, 54, 249-314.

Sorabji, R. (1971). Aristotle on demarcating the five senses, *Philosophical Review,* 80, 55-79.

Strawson, G. (1994). *Mental Reality*. MIT Press.

Tye, M. and Wright, B. (2011). Is there a phenomenology of thought? In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*, Oxford: Oxford University Press.

Vaitl, D. (1996). Interoception. *Biological Psychology*, 42, 1-27.

Veillet, B. and Carruthers, P. (2011). The case against cognitive phenomenology. In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*, Oxford: Oxford University Press.

Walker, M., Diebel, C., Haugh, C., Pankhurst, P., Montgomery, J., and Green, C. (1997). Structure and function of the vertebrate magnetic sense. *Nature*, 390, 371-376.

Wallcott, C., Gould, J., and Kirschvink, J. (1979). Pigeons have magnets. *Science*, 205, 1027-1028.