

To appear in *Perception and its Modalities*, Ed. D. Stokes, M. Matthen, and S. Biggs, New York: Oxford University Press (forthcoming, 2014)

Sorting the Senses

Stephen Biggs, Mohan Matthen, and Dustin Stokes

This volume is about the many ways we perceive. Contributors explore the nature of the individual senses, how and what they tell us about the world, and how they interrelate. They often represent competing views: for instance, some argue that perception uses the senses in concert, while others are content, at least for present purposes, to leave unchallenged the traditional assumption that we perceive through discrete senses. And the methods deployed sometimes differ from one essay to the next: some draw upon the sciences and engineering, while others rely upon conceptual analysis. Despite these differences, all the contributors agree that traditional theorizing about the senses is hampered by a neglect of the senses other than vision (based on the facile assumption that what goes for vision goes just as well for the other senses), and by the misconception that vision itself is a passive receptacle for an image thrown by a lens. In addition, many of the contributors believe that it is unduly restrictive to think of perception as a collation of content provided by individual modalities; they think that to understand perception properly, we need to build into our accounts the idea that the senses work together. The ambition of the volume is to begin to develop better paradigms for understanding the senses, and thereby to move toward a better understanding of perception.

The push for change begins with recognizing the shortcomings of tradition. We introduce the volume, accordingly, by sketching some of those shortcomings, and situating each contribution among the sketches. While we intend for the discussion to be informative in its own right, our sketches are offered to set the stage for the more detailed discussions that are offered by the essays that comprise the volume.

Some shortcomings of traditional theorizing are easier for the philosopher to spot than others. For some are evident when one carefully describes the phenomenology of sense-

perception. Others, however, are unearthed only when one considers philosophical issues in the light of results from the cognitive sciences.

As an example of a shortcoming that emerges from phenomenological description, consider this question: do we have auditory experience of sounds merely as coming from various directions, or do we hear some sounds as behind, farther away than, or partially occluded by others? Tradition tacitly assumes the former, on the grounds that sound impinging on the ear carries no mark of distance.¹ Reflection on our own auditory experience, however, makes clear that tradition is wrong: we often *do* hear some sounds as behind, farther away than, or partially occluded by others. With respect to distance, think of the sound of a train: it can sound not only as if it is approaching or receding, but also as if it is quite close by or distant. As for occlusion, think of the tap running in your kitchen sink, partially obscuring some of the conversation in the next room. Sounds are heard, then, as located in three-dimensional space.

As an example of a shortcoming that emerges only given empirical investigation, consider the nature of flavour. To the taster, it seems that flavour is a single composite quality of food, detected primarily by the tongue. The truth is more complex (Auvray and Spence 2008). Work by cognitive scientists has established that there are multiple sources of flavour perception, including not only taste receptors on the tongue (which account for the simplest taste components—salt, sweet, sour, bitter, and perhaps a couple more), but also olfactory receptors in the nasal tract (which react very differently when odorants emanating from the mouth flow over them in the retronasal, or opposite-to-sniffing, direction), and the trigeminal nerve, which is responsible for sensation in the face and for the motor control of biting and chewing. (It is the stimulation of the trigeminal nerve in the face that causes mint to have its characteristic coolness and chillies their characteristic heat—yet, the coolness and the heat are sensed *in the foods*, not in the face.) As well, of course, there is the complex experience of eating or drinking, of which flavour is only one part, with texture and mouth-feel as distinct but well-integrated components. Flavour,

¹ This is largely only tacit, since audition hasn't been discussed very much until recently—but now see Pasnau (1999), Nudds (2001), O'Callaghan (2007), Nudds and O'Callaghan (2009); Bulot and Egré (2010). The conclusion, one should note, is transferred from Berkeley's discussion of the impossibility of seeing distance—which in turn is inferred from the fact that the retinal image is two-dimensional.

though experienced as if it were delivered in a single perceptual act of sampling, is actually a composite of outputs from all these sources.

Since recognizing the shortcomings of traditional theorizing often requires considering philosophical issues in light of results from the cognitive sciences, theorizing about the senses is inherently interdisciplinary: philosophers who ignore the sciences will not encounter many of the results that constitute problems for traditional theorizing, and scientists who ignore philosophy will not appreciate many of the problems that could drive significant research. For example, it is philosophically significant that flavour is experienced as a property of food, not as events in the body occasioned by food. And this makes good sense from an ecological point of view: flavour helps us identify foods and their characteristics, and is not used as a marker merely of the *experience* of tasting various foods. But how does this happen? How is it, for example, that the irritation of the trigeminal nerve caused by chillies is felt to belong to the food in one's mouth? (Actually, it isn't just a few philosophical problems that are of scientific interest: one could go further and argue that without philosophy, scientists wouldn't have *any* problems to solve.)

We turn now to another glaring deficiency of the traditional view: it supposes that perceptual states reflect *just* the information available at the sensory receptors. In the case of vision, this information is supposed to be contained in the two retinal images; the nature of these images is supposed to be given in turn by geometrical optics and the theory of lenses, as investigated in a tradition that includes such landmark thinkers as Euclid, Ibn Haitham, and Johannes Kepler. Thus, philosophers from Berkeley down to A. J. Ayer (1956) supposed that visual awareness must be of a fusion of two planar distributions of coloured pixels. Since these images lack depth information—at least beyond the short distances in which there is detectable parallax displacement—any impression of depth or of object boundaries must come from learned associations. Thus, Berkeley held that there is no direct visual awareness of depth or distance (even at short distances, according to him) and that the appearance of depth must be added by the learned association of tactual ideas of distance with the inward turn of the eyes to fuse the two retinal images, the parallax displacement of these two images, and with the “confused” character of visual ideas that come from afar.

Assumptions of this sort were strongly challenged by the Gestalt psychologists of the early mid-twentieth century (Kohler 1930, Koffka 1935). Though a series of cleverly constructed perceptual illusions, they showed that depth, size, and object unity are essentially and inextricably contained in visual percepts. They then speculated about principles of scene analysis, which would explain how the cues available in the retinal images account for the richness of visual percepts. Although the resulting Gestalt principles have been somewhat eroded by subsequent research, this work made a permanent contribution: it is now universally accepted that the perceptual system *interprets* sensory data in accordance with principles that enrich the information available to the sensory receptors. The image of the world that we receive through the senses incorporates the results of this interpretative process, which is largely beyond the control of the perceiver.

It is still an open question, nonetheless, exactly how the senses extract perceptual content from receptor information. Three of the essays in this volume explore new models. Andy Clark (chapter 1) discusses “predictive coding”, a model on which the perceptual system makes predictions about what it will encounter and compares these predictions with incoming information, adjusting as necessary when discrepancies are found. Mohan Matthen (chapter 2) introduces a notion of active perception, which is a temporally extended process of deliberate probing of the environment, using bodily movement and interaction with objects of investigation. And Nicholas Shea (chapter 3) investigates the interaction between learned associations and incoming data—he asks to what extent interpretation is shaped by past experience.

Another serious shortcoming of traditional theorizing about the senses is its strong commitment to the idea that all perception is modality specific. The basic proposition underlying this commitment is that any sensory idea belongs to exactly one modality, which implies that there is no single sensory experience that is, say, audio-visual. Of course, visual and auditory experiences sometimes are coordinated, as when we see and hear a person speak. Tradition maintains, nonetheless, that the auditory experience is phenomenally distinct from the visual experience and hence should be treated as a separate entity. Relatedly, it implies that insofar as this and other perceptual experiences

seem to the subject to be integrated wholes, rather than mere aggregates of individual experiences, it is not because of the intrinsic nature of the experiences themselves, but rather, because of some further feature of the psychology in which they are embedded.

These assumptions are at play in Kant's discussion of the unity of apperception. Kant argues that a subject treats her visual and auditory experiences as belonging to herself and as being jointly informative about the world, and that this "treating as" is required for achieving the unity among the senses that one finds in perceptual experience. He would have proceeded differently had he taken perceptual experience to be a unitary product of information provided by the different sense modalities, rather than separate experiences that need then to be blended or added to one another. Although the view that perception is modality specific is a view about *experience*, it intimates that the senses themselves are discrete, rather than integrated in virtue of their intrinsic natures. And it assumes, further, that (for example) the act of hearing and seeing a person speak must be post-perceptual. As we shall see in a moment, this is contestable.

One question that immediately arises from the modality specific perception view is this: what happens when a single quality is expressed in more than one modality? This puzzle expresses itself, for example, in Molyneux's Question:

Suppose a man born blind, and now adult, and taught by his touch to distinguish between a cube and a sphere of the same metal, and nighly of the same bigness, so as to tell, when he felt one and the other, which is the cube, which is the sphere. Suppose then the cube and the sphere placed on a table, and the blind man made to see: Quaere, Whether by his sight, before he touched them, he could now distinguish and tell which is the globe, which the cube? To which the acute and judicious proposer answers: 'Not.' For though he has obtained the experience of how a globe, and how a cube, affects his touch; yet he has not yet attained the experience, that what affects his touch so or so, must affect his sight so or so... (Locke, *Essay Concerning Human Understanding* II, IX, 8)

If the visual idea of a globe is phenomenally distinct from the tactual idea, how is the man who is newly sighted supposed to identify the unfamiliar visual idea with the tactual one? The French philosopher, Denis Diderot, gave this problem an additional twist. The blind man takes the measure of a three-dimensional object by a temporally extended series of tactual impressions that unfolds as he feels it: how could such a temporal sequence of

impressions be equivalent to the single idea that the sighted have of the same three-dimensional object?²

Though Molyneux’s Question continues to be an open research problem (Held et al 2011), several experiments challenge the modality-specificity of perception by identifying percepts naturally taken to be *multimodal*, e.g., audio-visual. Consider the well-known “McGurk effect”. In one demonstration (McGurk and MacDonald, 1976), subjects see a video of a person saying “ga”, alongside a synced audio track of a person saying “ba”. They report *hearing* “da” (or a relevantly similar phoneme, such as “ta”). Plausibly, this is because their auditory experiences result from a sub-personal reconciliation of the visual and auditory stimuli—‘da’ is intermediate between ‘ga’ and ‘ba’ from the point of view of speech production. (Think of where in the mouth the closure occurs for these syllables.) Since subjects’ *auditory* representations of speech result, to a significant degree, from the *visual* stimulus, i.e., from the hearer’s visual perception of how the speaker produces speech, one is tempted to describe this as a case where *two* senses essentially contribute to *one* percept (i.e., the subject’s “da-experience”), and thus, a case where a single experience is audio-visual.

Consider next the “rubber hand illusion” (Botvinick and Cohen 1998). Here, a subject looks at a rubber-cast of a hand, while her own hand is hidden from her view. The rubber hand and her hidden hand are stroked in unison with a brush. Subjects report *feeling* the brush stroking the rubber hand. Thus, the feel of the brush on the subject’s hand—obviously, she cannot feel the inanimate rubber hand being stroked—is location-shifted by vision. In other words, *visual* input affects subjects’ representations of *tactile* stimuli. Again, one is tempted to describe subjects’ experiences here as multimodal.

An advocate of the traditional view may respond that, although the input to these perceptions is multimodal, they are nonetheless not cases of multimodal *experience*. Though they are caused by stimuli associated with (what are ordinarily taken to be) distinct sense organs, they are experiences characteristic of a single modality—audition in the case of the McGurk effect and touch in the case of the rubber hand illusion. Here, the respondent assumes that which sensory kind an experience belongs to depends entirely on

² See Diderot’s *Letter on the Blind*, chapter 3 of Morgan 1977.

its phenomenal character, i.e., what it is like to have it. It seems, phenomenally, to McGurk subjects that they hear a “da”; the “phenomenal character” of their perception is as if it is auditory. Thus, defenders of the traditional view take it to *be* auditory. The McGurk effect is simply an auditory illusion, they insist, a case where subjects *hear* ‘da’ when ‘ba’ is spoken. The genesis of the illusion has no bearing on its modality.

This discussion reveals that a lot can depend on whether we should give greater weight to the source of the information or to the phenomenal character of the resulting sense-impression. One might predict that ordinary ways of thinking would favour the second alternative, since the first turns on a determination made by a cleverly designed experiment to which one would ordinarily have no access. Participating in such an experiment might persuade us that the sources of information used by speech perception are unexpectedly diverse, but it does not easily persuade us that we don’t simply *hear* ‘da’ or ‘ba’ as the case may be. This brings out another problem for traditional theorizing: its criteria for *individuating* the senses face serious challenges, an issue we consider a little later.

Returning to multimodal perception, one might ask: What exactly is it that is multimodally perceived? In typical cases, there is no obvious answer. Consider a baseball player catching a ball (an example borrowed from Casey O’Callaghan). She sees the ball come into her glove, hears the sound of the ball making contact with the glove, feels the ball against her palm; she is also aware of her own bodily actions of reaching for and grabbing the ball. At first glance, this seems like a perfect example of a single occurrence being apprehended by many sense-modalities, each one an independent source of information, the whole being a sum of these independent contributions. Is there a single event or object here that is perceived by all modalities? One could argue that there is not. Vision engages the ball and its motion; audition picks up how the ball hitting the gloved hand disturbs the air; touch is concerned with how this collision affects the skin and muscles of the hand; bodily awareness monitors how the hand moves to grab the ball. These are four distinct events, constituting a single causal nexus, but arguably detected independently by different senses.

This is the sort of question that Casey O’Callaghan (chapter 5) explores. He claims that multimodal perception is, in part, the unification of these distinct events into a coherent environmental whole. He concedes, for the sake of argument, that one could

apportion these experiences to different modalities, but he questions whether the sum of single-modality experiences exhausts the total informational content of perceptual experience, given that the latter includes the connectedness and integration and integration of the distinct event mentioned. Matthew Nudds (chapter 6) takes a somewhat different tack: he explores how auditory and visual information can both be integrated in a single persistent material object, as in various “ventriloquism” effects, where seen objects are identified as sound *sources*. (Think of the ventriloquist’s dummy, which is misidentified as the sound-source because of its moving mouth.) Nudds’s approach fits well with Matthen’s line on active perception (chapter 2), which identifies interactive multimodal perception as a method of arriving at this kind of source-identification. In other words, while O’Callaghan identifies coherent nexuses of events as the target of multimodal perception, Nudds and Matthen take the complementary view that multimodal perception targets the material objects that are the *sources* of nexuses of non-visual sense-objects, such as sounds and smells.

However this may be, Charles Spence and Tim Bayne (chapter 4) present evidence that indicates a significant limitation on multimodality: human subjects, they suggest, can never be perceptually aware of two modalities at the same time. Of course, the system circulates among the modalities extremely fast and this is what accounts for the appearance, and possibly the functionality, of multimodality. Spence and Bayne maintain, however, that in any given very brief interval of time, one is aware only of ideas of one modality. (It is an open question how this works for flavour: is it impossible for us to be aware of taste and trigeminal stimulation simultaneously?) This calls into question the idea that there can be genuinely multimodal awareness: at the very least, there cannot be *simultaneous* multimodal awareness, if Spence and Bayne are correct. For in their view, multimodal awareness is a temporally extended collocation.

At the outset, we noted another deficiency of traditional theorizing about the senses: it all too often presupposes that theorizing about perception in general can proceed by studying only vision, and then taking claims about vision to generalize, with at most minor modifications, to non-visual senses. Matthew Nudds and Casey O’Callaghan (2009) emphasize this point,

Philosophical discussions of sensible and secondary qualities have focused upon color and color experience, while debates about perceptual content primarily concern the content of visual experience. (*ibid.*, 1)

While it might seem obvious in the case of vision that perceptual experience is *transparent* [i.e., that one is conscious of external objects, not of the experience itself *eds.*] or that space is required for objectivity, gustatory and olfactory experiences might tell otherwise (*ibid.*, 2)

The focus on vision is problematic at least because, in its thrall, philosophical theorizing has often suggested that all of the senses relate to knowledge and action in exactly the same way, i.e., as vision does. Various scientific results suggest, however, that the senses differ in these regards. To correct the error, and to provide a much richer and more pluralistic account, we have contributors who provide analyses of touch (Fulkerson), audition (Kulvicki), olfaction (Batty), and flavour (Spence, Smith, and Auvray).

What can we learn from non-visual perception? Consider, first, O’Callaghan’s question above. What are the objects of the non-visual senses? Interest in this question started with an influential paper by Robert Pasnau (1999). It is often taken for granted that audition picks up sounds, and that sounds are disturbances of the air. But, Pasnau argued, this cannot be right. For, disturbances of the air are spatially pervasive: when the starter’s gun is fired, the air throughout the stadium is disturbed. Yet, the sound of the gun is heard as localized; it is heard as coming from the gun. So what is sound? What, in other words, is audition’s target when we hear something? It is an event, says Pasnau—in this case, the firing of the gun. The disturbance of the air is simply the medium through which news of this event is conveyed to the ears of the competitors and spectators.

We noted earlier how the Gestalt psychologists had drawn attention to the objects of vision—material objects, not proximal stimuli. Pasnau’s argument shows how thinking beyond proximal stimuli can render the senses heterogeneous. Vision may focus on persistent material objects; audition, however, is concerned (according to him) with events. What about the other modalities? A number of essays in this volume address the issue. Matthew Fulkerson (chapter 7) distinguishes touch from the modalities that deliver information about the subject’s body by first noting that tactual properties are “bound” to external objects (as distinct from pains, itches, and the like) and then considering the role of active haptic exploration in detecting and binding these external properties. John Kulvicki (chapter 8) argues (against Pasnau) that audition detects the sonorous properties

of objects, rather than the events that these properties result in. Clare Batty (chapter 9) argues that olfaction is recognition-based rather than identification-based, i.e., it is concerned with the *kind* of substance it is engaging with, rather than with its location or boundaries. Charles Spence, Malika Auvray, and Barry Smith (chapter 10) consider the multi-input character of flavour, and argue that taste is *not* a discrete modality with input into flavour perception, but is rather an integrated part of the flavour system.

These discussions of the non-visual senses raise several interesting questions. For example, if one can sense flavours on the basis of input to multiple kinds of receptors, one wonders what else can be sensed. Picciuto and Carruthers (chapter 11) ask a broadly related question: are the so-called “inner” senses genuine sense modalities? Humans are capable of higher-order representations of first order perceptual states and cognitive states. And some have argued that calling this faculty a ‘sense’ is not mere metaphor. Picciuto and Carruthers proceed by first identifying a sense modality prototype (which we’ll discuss later) and then analyzing various inner sense theories accordingly. They conclude that in this light, at least some of the inner sense mechanisms posited by theorists *are* senses rightly so called.

One might suggest that, in the face of the traditional overemphasis on vision, we should, as an antidote, just stop thinking about sight and seeing for a while, to allow work on the other senses to catch up. But this is not a viable program. For, the traditional conception of vision is itself an unrealistically distorted idealization. For example, vision is wrongly thought to be transparently available to consciousness, and innocent of any systematic distortions that would reduce its trustworthiness. And these characteristics are assumed to be just as true of the other modalities as they are of vision. A more sensitive treatment of vision might lead us to question the homogeneity of such assumptions across the modalities—it could be that the senses are not homogeneous with regard to either outward directedness (or “transparency”) or trustworthiness. And, once we start thinking about interactions with other modalities, we might begin to think differently about how vision itself operates.

Howard Hughes, Robert Fendrich, and Sarah Streeter (chapter 12) take up the question of how the traditional conception of vision misleads. There are cases of vision

that are too ephemeral to count as reasons for knowledge; there are even cases of vision that are completely unconscious. Conversely, there are many cases where damage to parts of the brain result in an inability to identify objects of various kinds (such as faces or letters), even though visual acuity is maintained. Finally, there is the recruitment of visual areas of the brain for the processing of input from other senses, raising questions about the plasticity of visual processing.

Visual substitution technology helps problematize this last point, i.e., the plasticity of visual processing. Starting with Paul Bach-y-Rita at the University of Wisconsin Medical School, a series of devices was developed that converted camera input into the stimulation of a tactile vibrator array or sound sequence. With vibro-tactile substitutes for the retina, subjects using these “sensory substitution devices” begin to be able to sense distal objects as projected to the tactile array, and to report sensations of depth and perspective. Exploring the implications of sensory substitution, Ophelia Deroy and Malika Auvray (chapter 13) wonder what kind of awareness is provided by these devices: is it like touch or like sight? Their answer is that it is like neither but is rather a third sort of thing, which has some characteristics of both. Mazviita Chirimuuta and Mark Paterson (chapter 17) explore a different aspect of this question: to what degree does sensory substitution suggest that the spatial framework of the senses is transferable across the senses. On this point, their treatment should be set alongside Matthen’s account of space (chapter 2). Matthen argues that since active perception is multimodal, it requires cross-modal spatial location. This either requires that one of the modalities is privileged (which Stokes and Biggs assert (chapter 14) and Chirimuuta and Paterson, in effect, deny), or that space is “pre-modal.” The latter seems to have been Kant’s view, which Matthen defends.

It is arguable that the philosophical way of thinking about vision is itself fundamentally mistaken. We think of colour—perhaps the most fundamental visual property—as a similarity structure. This kind of thinking harks back to Aristotle, who thought of the colours as arrayed between just two poles, black and white. It’s difficult to see how Aristotle could have thought of (for example) brown and green as one “more white” and the other “more black”, but this illustrates the unreliability of this kind of intuition. By now, of course, we think of colours as a *three*-dimensional similarity space, with all the colours arrayed between white and black, red and green, and blue and yellow.

C. L. Hardin discusses the inadequacy of this schema. He recounts the experiments that create reddish-green experiences in (some) observers. He also discusses the possibility that colour vision treats the “unique hues” as distinguished points in colour-similarity space, not merely null points for one or other of the hue dimensions.

Erik Myin, Ed Cooke, and Karim Zahidi (chapter 16) explore an issue central to cross-modal influences, asking how changing the functional properties of a sense would affect associated experiences. They ask, for example, whether the phenomenal character of auditory experience would become like the phenomenal character of olfactory experience if audition took on some of the functional roles actually played by olfaction. Rather than taking a firm stand on this question, however, they answer that even those who think that functional properties determine phenomenal character can accept that, say, vibro-tactile sensory substitution does not generate experiences with a visual phenomenal character—roughly, because the functional properties associated with touch that are essential to vibro-tactile sensory substitution may play the greatest role in determining the phenomenal character of vibro-tactile experiences. While all contributors agree that vision should not be taken as the paradigm sense, Dustin Stokes and Stephen Biggs (chapter 14) consider what makes vision special. They argue that, alone among the sensory modalities, the visual possesses rich spatiality, and consequently, the visual dominates the spatial sense of the other senses. They suggest that this visual dominance makes vision especially important both psychologically and epistemically, which partly explains (but does not excuse) the traditional focus on vision.

A closely related shortcoming of traditional theorizing about the senses is its insufficient exploration of criteria for individuating the senses (see MacPherson 2011, Matthen forthcoming). A criterion for individuating the senses is a criterion for identifying the kinds of senses there are, and to which kind of sense any individual sensing belongs. Philosophers have proposed a number of candidates, many of which have been operative in relevant sciences. Several contributors suggest that these familiar criteria should be abandoned, in favour of either a preferred alternative or a view on which the senses work seamlessly together—see, especially, chapters under the heading “Relating the Modalities”. We now consider the most familiar four criteria, which distinguish senses by

one of receptive organ, proper sensible, representational character, or phenomenal character (cf., Grice 1962).

Aristotle distinguished the eyes, ears, skin, nose, and tongue as the five sense organs. He also observed that each sense organ is sensitive to the world in virtue of a different medium through which information about the object of perception is transmitted. Influenced by these and other similar empirical observations, many have thought that the senses have fundamentally different modes of sensitivity and different objects of perception. This suggests two criteria for individuating the senses: the *receptive organ criterion* holds that for each receptive organ there is exactly one sense; the *proper sensible criterion*, which offers at least a necessary condition for being a sense, holds that each sense uniquely detects at least one property (e.g., colour for vision). These criteria seem to converge, yielding the familiar five senses: the eyes, which alone detect colour, are distinct from the ears, which alone detect sound, and so on. Accordingly, Aristotle's approach was thought to be undisputable until at least the middle of the 19th century. At this point, scientists came to realize that the senses do not end at the receptive organs. The input provided by these organs awaits further processing in the brain. And this processing need not be as distinct and separate as the sense organs themselves. The discovery of sensory data-processing thus weakens the case for modality-specificity.

Putting the point about sensory processing aside, a little reflection on our physiology problematizes each criterion. The sense organs that Aristotle recognizes rely on multiple kinds of receptors, each of which is sensitive to different properties: the eyes contain rods and (three kinds of) cones; touch relies on stretch, pain, and pressure receptors; olfaction relies on a variety of chemical receptors; and so on. The processing pathways associated with these receptors, moreover, are kept distinct for at least some short distance in early sensory processing in the brain. It is far from obvious, then, that Aristotle's criteria yield the familiar five senses—why vision, rather than a “rod-sense” and a “cone-sense”? More broadly, it is far from obvious how to apply either criterion. Is the joint presence of rods and cones in the eye an adventitious by-product of the use of a lens to focus light? Or is there a deeper reason why these cells should be thought of as contributing to a single modality? What about the many kinds of touch receptor?

An alternative proposal resembles the proper sensible criterion, holding that senses are distinguished by the properties they represent, which makes for clear distinctions since each sense uniquely represents at least one property. This *representational criterion* requires an account of the conditions in which a sense represents a property. The most common naturalistic accounts hold (in one way or another, and many qualifications aside) that a mental state represents just those properties that *cause* it. So, for example, vision represents colour, shape, and size since these cause visual representations.

Given any such causal account of representation, however, the representational criterion struggles to account for the McGurk effect and rubber-hand illusion. Recall that in the former subjects' *auditory* representations of speech result, to a significant extent, from the *visual* stimulus, while in the latter their *proprioceptive* and *tactile* representations result, to a significant extent, from the *visual* stimulus. Given any causal account of representation, this suggests that, contrary to the representational criterion, audition and touch can represent proper sensibles of vision.

An advocate of the representational criterion may respond with an alternative account of representation. One alternative holds that which properties are represented in any given experience is determined entirely by the phenomenal character of that experience. On this interpretation, the representational criterion collapses into the last of the four familiar criteria, the *phenomenal character* criterion, which holds that there is exactly one sense for each kind of sense experience. Vision is distinct from audition, then, because the phenomenal character of visual experience (i.e., what it's like to see) differs from the phenomenal character of auditory experience (i.e., what it's like to hear), and so on.

The phenomenal character criterion faces a few obvious challenges. First, the evidential value of phenomenological claims is disputed. Second, the criterion is no help in the attempt to individuate non-human senses. It is no help, for example, when determining whether bat echolocation is hearing, vision, or something else entirely, because we cannot know on the basis of phenomenology what it's like to echolocate in the way bats do. This is a serious concern since an account of the kinds of senses there are should be a full-fledged account, not limited to the kinds of human senses there are—unless only humans have senses. Third, as we illustrate next, phenomenology provides

little help in individuating human senses in controversial cases. This is a serious shortcoming because controversial cases are precisely where guidance is needed most.

Sensory substitution phenomena present further challenges to each traditional criterion. Take, for example, Tactile Visual Sensory Substitution (TVSS), in which subjects detect shape and perspective distally on the basis of tactile stimulation projected from a camera. The representational and proper sensible criteria imply that TVSS is a form of *colourless vision*. The receptive organ criterion implies that TVSS allows one to detect shape and perspective *distally* through *touch*. Both implications are problematic. The phenomenal character criterion provides little guidance, in part because those using such devices struggle to classify their experiences into any sensory kind, as Deroy and Auvray and (chapter 13) report. Thinking about sensory substitution systems even makes some wonder whether any single, suitable criterion could motivate an answer to the question, “Am I seeing or feeling?”

These kinds of concerns have inspired an intriguing suggestion: rather than treating the familiar individuation criteria as competing, one should treat them as dimensions of a single criterion, which identifies senses by their position in a multi-dimensional state space (Macpherson 2011). Senses may cluster in this space in ways suggestive of the traditional categories, but they may not. In any case, since this *multi-dimensional criterion* is not committed to just one feature that makes a sense a sense, it can flex with new theory and data.

Although the multi-dimensional criterion is interesting, it faces its own challenges. Developing the view requires overcoming technical problems with characterizing the state space—e.g., specifying how the dimensions relate, quantifying them, etc. Some implications of the view, moreover, may be unpalatable; for example, many may reject the implication that there is a distinct sense for each (occupied) position in the state space. The multi-dimensional criterion, moreover, may retain some of the problems from each traditional criterion; for example, taking each criterion on board does nothing to tell us whether rods and cones are distinct receptive organs or parts of a single receptive organ, and thus, whether they constitute multiple dimensions of the state space or only one. In chapter 18, Fiona Macpherson takes these challenges on, thus giving further substance to her new start.

Of course, there are other ways to react to difficulties with individuating the senses.³ The most radical reaction rejects individuation altogether. On the traditional view, “sensory images” would faithfully reflect information separately gathered by the eyes, ears, etc., each of which would have its own proper sensible(s). But it is questionable whether this model is ultimately defensible. More broadly, once one acknowledges that sense experiences are highly processed entities, the prime motivation for sharply distinguishing kinds of senses may be lost. There are, of course, various ways to counter this concern. The crucial question is whether counter-considerations are any better than *ad hoc* ways to save an outmoded distinction.

Still, distinctions between, say, visual and auditory stimuli, visual and auditory receptive organs, visual and auditory experience, etc. may be inescapable. Imagine trying to explain the McGurk effect without them! One might be tempted, therefore, to adopt a deflationary stance on individuation, taking the senses to be *pragmatic*, rather than *ontological*, categories. On this way of thinking, an individuation criterion that is appropriate in one context may be inappropriate in another; perhaps, for example, one criterion is appropriate for discussing cross-modal priming and another for discussing attention-cuing. The goal of individuation, then, would not be discovering or legislating which senses there are, but something else, perhaps finding an individuation criterion that achieves the best fit with other theoretical commitments in a given context.⁴ Roberto Casati, Jérôme Dokic, and François Le Corre (chapter 19) consider these issues while exploring relations between the project of individuating the senses on the basis of folk psychology and the project of individuating the senses on the basis of empirical science.

The difficulties of individuating the senses arise largely because theorists treat the senses as entirely distinct conduits of information. The critic resists this traditional view by maintaining that perception is essentially *multisensory*.⁵ She notes, first, that the senses

³ Matthen (forthcoming) reacts by distinguishing sensory from perceptual modalities. The latter are individuated by the kinds of activity we undertake when we actively probe our surroundings by means of perception. Matthen’s claim is that the traditional modalities are individuated by mutually reinforcing sets of perceptual activities.

⁴ See Nudds 2004 for similar suggestions, and Macpherson 2011 for further discussion.

⁵ Those who agree that perception is multisensory can dispute both what grounds its being multisensory, and where its multisensory nature can be detected. Some might hold, for example, that while the phenomenal characters of sense experiences are discrete, the neural bases for “information pick-up” are continuous, or vice versa.

interact in ways that traditional theorizing would not predict, in ways that (arguably) they could not interact if they were largely discrete conduits of information. She explains this by observing that the senses inform us about a single reality beyond our minds. Given that the information garnered from different sense organs and different groups of receptors are about the same external world, one would expect these information channels sometimes to reinforce and sometimes to act as checks on one another. Thus, a well-designed perceptual system would allow cross-communication between processing streams.

Interactions within the complex flavour system illustrate the point. As noted earlier, flavour is the product of the taste receptors in the tongue, retronasal olfaction, and trigeminal stimulation—receptors that belong to distinct senses on most criteria. Although these sources can be activated independently of one another or in sub-groups, in ordinary flavour perception each influences the others, which, as noted, leads Spence, Auvray, and Smith (chapter 10) to treat them as a single perceptual system. Other senses, moreover, can influence these sources themselves: the apparent taste of tasteless dyed water seems to depend on the colour of the dye (except, of course, when subjects are blindfolded (see Clydesdale 1984). Similar points can be made for the other senses. Just as ordinary flavour perception uses multiple kinds of inputs, vision uses rods and cones, touch uses stretch, pain, and pressure receptors, and so on—at least touch, moreover, is strongly influenced by the exploratory activity of the perceiver. Just as vision can influence taste, moreover, it influences audition in the McGurk effect, proprioception in the rubber hand illusion, and touch in a variety of cases—indeed, even visual *imagination* can influence non-visual perception (see Stokes and Biggs, chapter 15). This all intimates that the perceptual system uses any available information, integrating it at various stages of processing, with little concern for receptive organ or proper sensible.⁶

⁶ Potential neural bases for these kinds of inter-sense interactions have been identified. For example, a series of studies find that when a *tactile task* concerns object identification (or shape, size, or orientation discrimination), the *visual cortex* is regularly active. Further, this activation seems to be functionally operative; when the visual cortex is disrupted with transcranial magnetic stimulation (TMS), performance on the tactile discrimination tasks diminishes significantly. Remarkably, this visual effect on touch is evident even when visual stimuli are removed. Researchers infer that visual *imagery* is invoked (sometimes tacitly, sometimes strategically) to aid in the tactile task (Sathian et al 1997; Zangaladze et al 1999; Zhang et al 2004; and see Djordjevic et al 2004 for similar research on visual-olfactory interaction). Similarly, although the lateral occipital complex (LOC) (a region in the ventral visual pathway in the

One might wonder why perception integrates information across modalities so freely given that integration can mislead us, as vision “misleads” taste in the dye-experiment, audition in the McGurk effect, and proprioception in the rubber-hand illusion. The reason may be that integration, Stokes and Biggs suggest, is ordinarily beneficial. Subjects who see a briefly flashed picture (say, of a dog) are much more likely to identify the depicted object correctly when also presented with a semantically congruent sound (say, barking) than when presented with a semantically incongruent sound (say, whistling), or no sound at all (cf., Chen and Spence 2010). This may reflect a kind of “priming” that takes advantage of environmental regularities: one dog-indicating stimulus (barking) readies the perceptual system to identify other stimuli as emanating from a dog, and this aids both in interpreting ambiguous stimuli and in locating the source of the barking.

Interactions among the senses also inform us of predicative relations. In the ventriloquist effect, the ventriloquist produces sounds with minimal lip and mouth action, but moves a dummy’s mouth, with the effect that the sound is narrowly and irresistibly located in the dummy. Why is this? Not because vision picks up auditory information, but rather because it sometimes provides information that helps resolve auditory information in conditions that are not conducive to accurate resolution. Thus, social interactions demand that we know *who* is speaking and, ideally, his or her *intentions* vis-à-vis conveying a message, and this task is improved with the influence of vision. Giving visual information a role in the interpretation of the voice stream improves perceivers’ ability to gather semantic information. Indeed, a similar explanation might be given of the McGurk effect. A few contributors consider how multi-level interactions among the senses might be beneficial.

Whatever the value of the interactions, a multisensory view holds that each sense is influenced by others. Some go further down the path of integration, claiming that each sense can be influenced by learned states. Recent work in cognitive psychology may

human brain) is traditionally considered to be responsible for processing visual information, given substantial research suggestive of task-specific and performance-enhancing activation of the LOC during various touch tasks, some researchers now suggest that this area is multisensory, responsible for processing object-level or macrospatial information, which can then be used for either visual or touch tasks. See Amedi et al., 2002; Amedi et al., 2001; James et al., 2002; and Zhang et al 2004 for discussion of opposing interpretations of these data.

support this view. Studies find, for example, that pictured objects are perceived as being their typical colour even when the picture is adjusted to achromatic grey: subjects see an achromatic grey picture of a banana as yellow and an achromatic grey Smurf cartoon character as blue (Delk and Filenbaum 1965). This suggests that one's culture may influence what one senses. Reinforcing this suggestion, German subjects see an achromatic grey UHU glue tube as being its typical colour (yellow), but do not see an achromatic grey Ferrari icon as being its typical colour (also yellow), presumably because the former is more familiar to them. (See Hansen et. al 2006; Olkkonen et al 2008; Witzel et al 2011 for these studies.)

Racial stereotypes may have a similar effect on visual perception. In online matching tasks, subjects tend to judge greyscale faces with typical Caucasian features as lighter, and greyscale faces with typical Afro-American features as darker, in spite of the target faces being identical in luminosity. This effect can even be produced for racially ambiguous faces simply by labelling them 'WHITE' or 'BLACK' (Levin and Banaji 2006). Culturally specific associations have even stronger effects on other senses. For example, a wealth of recent research suggests that beliefs regularly, powerfully influence flavour perception—e.g., beliefs about the price of the wine, the brand of ketchup, and where one consumes a meal affect both the capacity for successful flavour distinctions and the pleasure reported.

These results suggest that beliefs and culturally acquired concepts directly affect visual experience. However, the mechanism of these influences is unclear. Is it really *belief* that influences perception (Farkas and Bitter 2012)? Suppose that subjects are *told* that they are about to be presented with achromatic banana images. Would the resultant belief modify their tendency to experience bananas as yellow, or would the bias to yellow persist despite a contrary belief? Might the effect be due to association and conditioning rather than belief? It is unclear, therefore, how much this discussion overlaps with the spirit of the New Look movement in psychology, led by Jerome Bruner and colleagues in the mid-20th century. The movement held that perception is continuous with cognition (beliefs, desires, values) in that how we perceive the world via the senses is always informed by one's "mental set". The New Look movement produced many results, similar to those discussed above. These results were interpreted as showing that cognition

profoundly influences perception—i.e., that cognition “penetrates” perception—but as indicated, the jury is still out on the extent of the phenomenon.⁷

One might worry that any influence of cognition on perception would be epistemically pernicious. How could the senses work quickly enough, deliver objective reports, or even update beliefs at all if they were influenced by background beliefs (cf., Fodor 1984), let alone desires? Although these concerns are serious, merely raising them does not decide the issue. In the first place, these kinds of influences need not be unconstrained: a dog’s bark might predispose a perceiver to interpret visual stimuli as dog-like, but it does not follow that she will be perceptually disposed to interpret scientific observations in a way that favours her preferred theory. Perhaps the function by which we combine input from online stimuli with input from memory is faster, more objective, and tracks truth better than a system that ignores input from memory, or allows memory input only at the level of judgment. Perhaps, then, by integrating incoming and remembered information the perceptual-cognitive system generates a “report” that optimally guides action. Both Clark’s discussion of predictive coding and Shea’s discussion of top-down versus bottom-up effects are especially relevant here (chapters 1 and 3, respectively).

The discussion thus far has touched on many questions. How are improvised stimuli built into rich perceptual experiences (see chapters 1-3)? Is perception always modality specific or can it be multimodal (see chapters 4-6)? What can we learn by investigating senses other than vision (see chapters 7-12)? What can we learn through a careful reappraisal of vision (see chapters 13-15)? How should we think about individuating the senses, and, more broadly, how do the senses interrelate (see chapters 16-19)?

These questions, and the many others that are addressed in this volume, invite a foundational question that largely remains in the background: what makes something a sense modality? This question has received almost no attention, perhaps because it is

⁷ For recent theorists who have argued that some of this data is best explained in terms of cognitive effects on perception (called ‘cognitive penetration’), see Macpherson 2012; Siegel 2011; Stokes 2012; Stokes and Bergeron (unpublished manuscript); Wu 2013. See also Deroy 2013 for worries about whether the alleged effect is a genuine *cognitive* one. See Stokes 2013 for an overview of the debate.

taken to be uncontroversial that a sense is a faculty by means of which a subject monitors her immediate environment in “real time”. This characterization, however plausible it may be, does not define the property *being a sense*, at least because it provides only a necessary condition. After all, the body monitors blood CO₂ levels and uses the information it gathers for homeostatic regulation in real time, e.g., to regulate respiration and heartbeat, but surely *that* is not a sensory process. Brian Keeley (2002) made an important start to sorting out this question: he suggests that the senses convey information from the outside world into the central nervous system in order to guide specific responses to a certain kind of ambient energy. On the other hand, it is quite plausible that there is a pain sense. What is the difference? As Picciuto and Carruthers (chapter 11) put it in their consideration of the “inner sense”: “What constitutes a system as a sense modality?”

There are two approaches to this question. Picciuto and Carruthers look for essential differences between those systems that count as sense modalities and those that do not. There are at least two crucial differences between the senses and mechanisms such as the blood CO₂ monitor. They “issue in nonconceptual representations with mind-to-world direction of fit.” And these representations “are used to guide the intentional behavior of the organism (perhaps issuing in phenomenally conscious sensations).”

A complementary approach, taken by Mohan Matthen (forthcoming), is to focus on the functions of the senses *as a group*. As noted earlier, the senses furnish us with *additive* complementary information: there is a barking (audition) dog (vision); here is a cool (touch), shiny (vision) ball. The blood CO₂ monitor does not complement other information gathering faculties in this way; that is, the evidence does not contribute to a nexus. More notably, the evidence of the other senses does not affect its control of bodily processes. Second, the senses are used to *learn* about the environment and *anticipate* what will happen in ways that non-sensory mechanisms are not. Specifically, whereas the senses feed into learning systems (such as association and conditioning) that initiate action on the basis of not only sensory input but also learned associations within and also across modalities. Picciuto and Carruthers focus on the representational states that issue from the senses; Matthen emphasizes that these representations work together across the modalities. This cooperation is, in the end, the rationale for multimodality: the senses are designed to work as a group.

One should now have a sense not only of the shortcomings of traditional theorizing about the senses, not only of how much room there is for a new, better understanding, but also of how challenging achieving a better understanding will be. Despite the challenges, we are optimistic that an improved understanding is possible. The papers commissioned for this volume are a testament to that optimism. We hope that they will pave the way to a new, better understanding of the senses, and in turn a new, better understanding of perception.

REFERENCES

- Amedi, A., Jacobson, G., Hendler, T., Malach, R., & Zohary, E. (2002) “Convergence of visual and tactile shape processing in the human lateral occipital complex” *Cerebral Cortex*, 12: 1202-1212.
- Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001) “Visuo-haptic object-related activation in the ventral visual pathway.” *Nature Neuroscience*, 4: 324-330.
- Auvray, Malika and Spence, Charles (2008) “The Multisensory Perception of Flavour,” *Consciousness and Cognition* 17: 1016-1031.
- Ayer, A. J. (1956) *The Problem of Knowledge*, London: Macmillan.
- Bullot, Nicolas and Egré, Paul (2010) *Review of Philosophy and Psychology* vol 1, no 1.
- Chen Y.-C. and Spence C. (2010) “When hearing the bark helps to identify the dog: semantically-congruent sounds modulate the identification of masked pictures,” *Cognition* 114, 389–404
- Clydesdale, F. M. (1984) “The influence of colour on sensory perception and food choices.” In J. Walford (Ed.), *Developments in food colours*. London: Elsevier Applied Science: 75–112.
- Delk, J. L. and Filenbaum, S. “Differences in Perceived Color as a Function of Characteristic Color,” *American Journal of Psychology* 78: 290-293.
- Deroy, O. (2013) “Object-Sensitivity Versus Cognitive Penetrability of Perception.” *Philosophical Studies* 162 (1):87-107
- Djordjevic, J., R.J. Zatorre, M. Petrides, and M. Jones-Gotman (2004) “The Mind’s Nose Effects of Odor and Visual Imagery on Odor Detection.” *Psychological Science* 15(3): 143-8.
- Farkas, Katalin and Bitter, David (2012) “Perceptual Learning and the Cognitive Penetration of Perceptual Experience,” unpublished paper presented at the Workshop on Perceptual Learning, University of York (England), 20th March 2012.
- Fodor, J. (1984). Observation reconsidered. *Philosophy of Science* 51: 23-43.

- Hansen, T., M. Olkkonen, S. Walter & K.R. Gegenfurtner. 'Memory modulates color appearance.' *Nature Neuroscience* 9 (2006) 1367-8.
- Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011) "The newly sighted fail to match seen with felt." *Nature neuroscience*, 14: 551-553.
- James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). "Haptic study of three-dimensional objects activates extrastriate visual areas." *Neuropsychologia*, **40**, 1706-1714.
- Keeley, Brian L. (2002) "Making Sense of the Senses: Individuating Modalities in Humans and Other Animals," *Journal of Philosophy* 99: 5-28.
- Koffka, Kurt (1935) *Principles of Gestalt Psychology* London: Paul, Trench, Trubner
- Kohler, Wolfgang (1929) *Gestalt Psychology* New York: Liveright.
- Levin, D. and Banaji, M. (2006) "Distortions in the Perceived Lightness of Faces: The Role of Race Categories," *Journal of Experimental Psychology: General* 135: 501-12.
- Macpherson, Fiona (2011) Introduction to Macpherson (ed.) *The Senses: Classic and Contemporary Philosophical Perspectives: Classic and Contemporary Philosophical Perspectives*. Oxford University Press, USA.
- _____ (2012) "Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism." *Philosophy and Phenomenological Research* 84.1: 24-62.
- Matthen, Mohan (forthcoming) "Individuating the Senses," in M. Matthen (ed.) *Oxford Handbook of the Philosophy of Perception*.
- Morgan, Michael J. (1977) *Molyneux's Question: Vision, Touch, and the Philosophy of Perception* Cambridge: Cambridge University Press.
- Nudds, Matthew (2001) "Experiencing the Production of Sounds," *European Journal of Philosophy* 9: 210-229.
- Nudds, M. and O'Callaghan, C (2009) *Sounds and Perception: New Philosophical Essays* Oxford: Oxford University Press
- O'Callaghan, Casey (2007) *Sounds: A Philosophical Theory* Oxford: Oxford University Press.
- Olkkonen, M., T. Hansen, and K.R. Gegenfurtner. (2008) 'Colour appearance of familiar objects : effects of object shape, texture and illumination changes.' *Journal of Vision* 8: 1-16.
- Pasnau, Robert (1999) "What is Sound?" *The Philosophical Quarterly* 49: 309-324.
- Sathian, K., Andro Zangaladze, John Hoffman, & Scott Grafton (1997) "Feeling with the Mind's Eye." *Neuroreport* 8(18):3877-81.

- Siegel, S. (2011) "Cognitive Penetrability and Perceptual Justification." *Nous* 46: 201-22.
- Spence, Charles (2012) "Is it the plate or is it the food? Assessing the influence of the color (black or white) and shape of the plate on the perception of the food placed on it," *Food Quality and Preference* 24: 205-208
- Stokes, D. 'Perceiving and Desiring: A new look at the cognitive penetrability of experience.'
(2012) *Philosophical Studies* 158: 479-92.
- _____(2013) 'Cognitive penetrability of perception' *Philosophy Compass*
- Stokes, D. and V. Bergeron. (unpublished manuscript). 'Modular architectures and Informational encapsulation: A dilemma.'
- Zangaladze, A., Charles Epstein, Scott Grafton, & Krish Sathian (1999). Involvement of Visual Cortex in Tactile Discrimination of Orientation. *Nature* 401: 587–590.
- Zhang, M., V.D. Weisser, R. Stilla S.C. Prather, & Krish Sathian (2004). Multisensory Cortical Processing of Object Shape and its Relation to Mental Imagery. *Cognitive Affective Behavioral Neuroscience* (4): 251–259.
- Witzel, C., Valkova, H., Hansen, T., Gegenfurtner, K. 'Object knowledge modulates colour appearance' *i-Perception* 2 (2011): 13-49.
- Wu, W. (2013)'Visual Spatial Constancy and Modularity: Does Intention Penetrate Vision?' *Philosophical Studies* 165(2): 647-669.