

SEEING WHAT YOU HEAR: CROSS-MODAL ILLUSIONS AND PERCEPTION

Philosophical Issues, forthcoming

Casey O'Callaghan

Bates College

October 28, 2007

1. Introduction

Vision dominates philosophical and empirical thinking about perception and perceptual experience. It furnishes the puzzles any philosophical theory of perception must solve and delivers the evidence and intuitions by which we measure such theories. Color vision, spectrum inversion, the waterfall illusion, blindsight, change blindness, and inattentional blindness all have driven philosophical theorizing about perception in recent years. The terminology deployed in discussing perception frequently is explicitly visual: *appearance, image, scene, perspective, observe*.

Increasingly, philosophers and cognitive scientists hope to discover what there is to learn about perception from modalities other than vision. Nothing guarantees that an adequate theory of vision extends neatly to audition, touch, olfaction, or gustation. Martin (1992) focuses upon touch to express skepticism whether any theory of perception generalizes to all the sense modalities. Lycan (2000) intimates that smells and olfaction hold important lessons concerning perceptual representation.¹ Theories of sounds and audition with surprising revisionist consequences recently have been proposed.²

Exploring modalities other than vision is good philosophical methodology because it

¹ See also Batty 2007.

² See, e.g., Casati and Dokic 1994, 2005; Pasnau 1999, 2000; Nudds 2001; and O'Callaghan 2007.

reveals where our vision-based understanding of perception succeeds and where it fails. Sometimes other modalities confirm what we learn from vision, but sometimes the lessons conflict. This expanded perspective furnishes challenging new puzzles to drive perceptual theorizing. Considering non-visual modalities thus is valuable in developing a comprehensive, general understanding of perception.³

But it does not go far enough. Developing a theory of audition, olfaction, taste, or touch, even when it does not merely extrapolate from vision, but faces up to perceptual phenomena drawn from the relevant sense modality, risks succumbing to a more insidious form of visuocentric thinking. In short, it enlists the methodological assumption that we can understand the perceptual modalities in isolation from each other.

I do not mean that considering each modality in its own right is not a fruitful strategy or that comparing and contrasting sense modalities is not an important way to learn about perception. For instance, recognizing that the sounds we hear, unlike the material objects we see, do not appear to be wholly present at a given moment is critical to understanding audition.⁴ Contrasting olfaction's spatial characteristics with vision's challenges, *prima facie*, the dependence of objective experience upon space.⁵ Recognizing that vision presents objects arrayed at a distance from your body while touch requires contact elucidates two varieties of perceptual acquaintance with an extended object.

I mean that an adequate, complete understanding of perception requires comprehending the ways in which what goes on with one sense modality *impacts* what goes on with another. Theorizing about perception is not just a matter of assembling

³ One nice example of this is Matthen 2005, ch. 13, sections III and IV, 282-92.

⁴ See O'Callaghan 2007, especially ch. 2.

⁵ See Smith 2002, ch. 5.

independently viable stories about vision, audition, olfaction, and the rest. Considering the relationships and interactions among perceptual modalities sheds light on what is most striking about perception: its capacity to furnish a sense of awareness of a world of things and happenings independent from oneself.

In this paper, I present a puzzle about audition that stems from a noteworthy difference between vision and audition. The puzzle is important because it emerges from attempting to understand hearing in a way that abstracts from other sensory modalities. The puzzle resists solution on such an understanding. I suggest, however, that grasping the import of a number of surprising *cross-modal* perceptual illusions helps to resolve this puzzle. Appreciating these illusions, moreover, tells against a widespread and otherwise intuitive conception of the role of the senses in perception and perceptual experience. According to this way of accounting, your perceptual experience of the world comprises characteristic experiences that are specific to each of the different sense modalities -- your overall perceptual experience is an assemblage of modality-specific components. The cross-modal illusions tell against this widespread conception because explaining them requires recognizing an aspect of perceptual experience that cuts across the boundaries of the sense modalities. In philosophical terms, it requires recognizing both a component of experiential content and an aspect of perceptual phenomenology that are shared by distinct perceptual modalities. Perceptual experience thus cannot be understood exclusively in modality-specific terms. What's more, this multi-modal dimension to perceptual experience makes a critical contribution to our sense of awareness of a world of things, happenings, and features that are distinct from ourselves and our experiences.

2. *A puzzle about audition*

According to an intuitive line of thought, sounds and their audible qualities are *proper* intentional objects of audition. Sounds are inaccessible to other sense modalities. Furthermore, sounds and their audible qualities are *immediate* objects of auditory perceptual experience, in the sense that your awareness as of a sound need not occur by or in virtue of your awareness as of something else. For instance, your awareness as of a sound does not occur in virtue of your awareness as of private auditory sense data because the latter do not seem to be objects of your awareness at all. Finally, sounds and audible qualities appear to *exhaust* audition's immediate objects. Whatever else you hear, you hear it by or in virtue of hearing a sound. Hearing a bell, it is natural to think, requires hearing a sound.

Audition thus differs from vision. Vision appears to have among its immediate objects ordinary material things like tables and chairs. But sounds are not ordinary objects. A sound is unlike a table or a chair. Sounds are not heard to have crisp spatial boundaries, to be solid, or to be wholly present before you at any given moment. A sound occurs or takes place over time, and patterns of change through time are essential to the identities of many common sounds, such as spoken words or police sirens. We perceptually individuate sounds according to their causal sources, but sounds do not seem simply to be properties or qualities of ordinary objects. Sounds do not seem audibly bound to ordinary objects in the way that colors and shapes visibly appear to qualify their bearers. Sounds are audibly distinct from ordinary material things.

Nonetheless, auditory experience is *object-* and *event-*involving. You might hear a bell, a train, or a muffler, and you might hear a collision, a conversation, or a glass breaking. However, you seem to hear such things by or in virtue of hearing their sounds, since your auditory awareness as of an ordinary object or event depends upon your awareness as of a sound. You would not have heard the train or the conversation had you not heard a sound. Audition affords awareness of the muffler or the glass breaking in virtue of its capacity to reveal the sounds of the muffler and the breaking.

One might argue that we merely learn about ordinary objects and events on the basis of audition without strictly speaking hearing them. Forming beliefs about your environment on the strength of auditory experience might result from extra-perceptual judgments, associations, inferences, or otherwise cognitive processes. The advantage is that this explains audition-based learning about things that are not themselves sounds without the burden of explaining how we *hear* non-sounds. Three considerations, however, tell against such an account. First, reflection seems to reveal that we simply hear things like dogs and ice cream trucks, and it is the perceptual seemings that are in question. Some nevertheless worry that we are not particularly adept at reporting how things perceptually seem; perhaps beliefs masquerade as auditory experiences. A strategy to quiet this worry appeals to our skill at detecting phenomenological contrasts.⁶ Suppose you are listening to the sound of a vacuum cleaner running in the next room. Now suppose we replace the vacuum with a very good recording of a vacuum running, and that we tell you this. Though you neither judge nor believe that there is a vacuum running next door, your auditory experience remains intact. Auditory phenomenology is unaffected. You need not believe or be disposed to believe that there is a vacuum running

⁶ Cf. Siegel 2006*a, b*, 2008; O'Callaghan 2007, ch. 3.

for your experience to seem to be as of a vacuum cleaner. Finally, the patterns of action and reaction afforded by audition support a perceptual understanding of our awareness of ordinary things and happenings. We quickly orient and direct visual attention toward the apparent source of a sound, and we reflexively duck upon hearing something rapidly approaching. This would make little sense if we heard only a sound to approach.⁷

The puzzle, then, is how audition, whose immediate but proper objects are sounds, could furnish awareness *as of* an ordinary object or happening. How could hearing a sound be a means to hearing a glass breaking or to hearing a muffler? Part of the issue is that a sound seems like such a different sort of thing from a commonplace material object or occurrence. How could awareness *as of* a sound ground awareness *as of* something extra-acoustic? It is helpful to contrast this with other varieties of indirect perceptual awareness. Perhaps there is a sense in which we visually experience objects by or in seeing their features such as color and shape. Sounds do not in this sense auditorily seem like features of ordinary material things. The natural candidates, pitch, timbre, and loudness, however, are qualities of sounds. Consider, next, seeing an object by seeing its facing surface. The latter is a mereological part of the former, and we recognize it as such. Sounds are not obviously mereological parts of ordinary things. Consider, third, seeing a face by seeing its televised image. Faces are among the things we see, and the capacity to recognize a face suffices for the capacity to recognize a pictorial rendering of a face.⁸ The two notably share attributes such as colors and perspectival characteristics

⁷ One response is that we do not hear non-sounds, but perceive things we see to *produce* or *generate* sounds. We never hear non-sounds, since the experience of the production of sounds by visual objects is essentially bimodal. This view is articulated and defended by Nudds (2001). The problem, as auditorily-guided action shows, is that the phenomenology of generatedness or production occurs even when the source is unseen.

⁸ See Carroll 1985. This need only apply to human perceivers.

that depend upon the relative spatial arrangement of features and viewing angle. The image is a likeness of the face. Sounds and objects or collisions do not share features in this straightforward way. The sound is not a likeness of the thing. Consider, finally, that sometimes we claim to see fire in virtue of seeing smoke. This, however, does not constitute visual awareness as of a fire unless you can attend to the fire itself. Hearing a sound does frequently afford the opportunity to attend and demonstratively refer to its source.⁹ The puzzle is how audition affords awareness as of the source, so it may be heard to make the sound.

The puzzle, then, concerns what experientially grounds your awareness as of a sound source -- an ordinary object or happening -- where that awareness depends upon your awareness as of a sound. Notice that in order to generate the puzzle, you need not be auditorily aware of the specific kinds to which extra-acoustic things belong. You need not, for instance, be aware of something as a muffler or a backfire. Rather, you need only enjoy auditory awareness as of something like a *source*, an *object*, or a *happening* beyond a sound itself. How do we *auditorily* experience such particulars, given that sounds are, in the first instance, what we hear and that to which we ascribe audible qualities? The puzzle concerns how awareness as of a sound affords auditory perceptual experience as of an extra-acoustic object or happening.¹⁰

⁹ Cf. Campbell 2002, 63-4. Some (e.g., Snowdon 1992) suggest that the capacity to demonstratively refer or attend suffices for direct perceptual awareness. This might be thought to conflict with the claim that auditory awareness of sound sources requires awareness as of a sound. The epistemic notion of directness, however, might come apart from the metaphysical or experiential notion of immediacy. Bermúdez (2000) argues that one might directly perceive some particular while not immediately perceptually aware of it. Thus, one might directly perceive a sound source despite only mediated perceptual experience of that source.

¹⁰ The puzzle, it is worth noting, is not a metaphysical puzzle. It is not a puzzle concerning how the relationship between sources and sounds affords perception of sources. Rather, it concerns,

3. *The composite snapshot conception*

The puzzle concerning audition, I want to suggest, has its origins in a conception of the role of the senses in perception and perceptual experience that is underwritten by visuo-centrism in perceptual theorizing. The picture stems from thinking of the senses as providing discrete modes or channels of awareness. The idea, in brief, is that perceptual experience comes in distinctive varieties corresponding to different *modalities*, and that these distinct modalities of awareness collectively make up one's overall perceptual experience. Intuitively, seeing differs from hearing, hearing differs from smelling, smelling differs from touching, and touching differs from tasting. Perceiving nevertheless just is a matter of seeing, hearing, smelling, touching, and tasting.¹¹ The relevant kind of difference is one that is accessible to the subject of the experience -- it is phenomenologically apparent. Just by introspecting our perceptual experience, we recognize both that it consists in experiences associated with different modalities and how such experiences differ from each other.¹²

The intuitive picture has three central commitments. First, one's total perceptual experience comprises visual, auditory, tactile, etc., experiences. Experiences corresponding to the different sense modalities are *constitutive* of one's overall perceptual experience. Second, one's total perceptual experience is *exhausted* by one's visual, auditory, tactile, etc., experiences. There is no more to what we perceptually experience

according to one way of putting it, the resources of auditory representation. How does representing a sound ground one's capacity to represent an ordinary object or event?

¹¹ This particular list need not exhaust the experiential modalities. It serves as a placeholder for a complete accounting of whatever modalities of awareness exist.

¹² Tye (2003, 2007) and Nudds (2001) also critically examine this intuitive conception. As will become clear, my characterization of the intuitive view and my reasons for challenging it differ from those Tye and Nudds offer.

than what we visually, auditorily, tactilely, etc., experience. Third, visual, auditory, tactile, etc., experiences each have their own *distinctive* and recognizable phenomenological character. There is something distinctive about what it is like, for example, to undergo a visual experience that differs from what it is like to undergo an auditory experience or a tactile experience. Each modality-specific kind of perceptual experience has its own unique character that could not be shared by any other perceptual modality. For instance, no auditory experience could share the phenomenology of a given visual experience. Two visual experiences, however, might be phenomenologically alike. This implies, for example, that for any object you can perceive through a given modality, there is a phenomenologically unique or distinctive way of experiencing it in that modality that cannot be shared by any other modality in which you might experience that object.¹³

¹³ *Distinctiveness* is tricky to spell out. It might involve the rather weak claim that there always exists a *global difference* between the phenomenology of experiences in distinct modalities. For example, a complete description of the phenomenology of one's auditory experience differs from that of one's visual experience at a time. This implies, for experiences in different modalities, that there always exists some phenomenological feature that one has and the other lacks. Though the overall phenomenology of experiences in different modalities must somehow differ, this understanding of distinctiveness is compatible with a great deal of phenomenological similarity. This characterization of distinctiveness is a common defense by those who hold that phenomenological character is a subset of representational content against the objection that one could represent some object, such as a Frisbee, visually or by touch. The objection is that the pure representationalist cannot capture the phenomenological difference between such experiences. The representationalist responds that there is plenty else besides being a Frisbee-representing experience that differs between the two experiences, such as that one represents redness and the other represents coolness-to-the-touch. The arguments I develop in this paper are mostly compatible with this very weak reading of distinctiveness. However, all but a few reject this response and advocate something stronger. Some appeal to modes of presentation or modalities of representation to capture the phenomenological difference between experiencing a Frisbee in vision and in touch (cf. Chalmers 2004, 2006). Some appeal to non-representational qualitative differences. Correspondingly, a stronger understanding of distinctiveness holds that any two experiences in distinct modalities differ phenomenologically. So, for instance, seeing a dog and hearing a dog differ phenomenologically in ways that do not require appeal to the complete phenomenology of one's visual and auditory experiences. Applied to all aspects of an experience, this implies the strongest reading of distinctiveness, according to which experiences in distinct

According to this understanding, one's total perceptual experience at a time is an assemblage or composite of modality-specific experiences. Perceptual experience comprises discrete, modality-specific components or 'snapshots'. Each such modality-specific experience has its own recognizable and distinctive character. You would not, for instance, mistake an auditory experience for a visual experience. Call this traditional conception of overall perceptual experience the *composite snapshot conception*.¹⁴ According to this picture, each sense modality delivers a discrete snapshot of the world from its unique perspective, and the aggregate of these snapshots -- a composite snapshot -- constitutes and exhausts one's total perceptual experience.

It is fair to say that the composite snapshot conception is the traditional empiricist view of experience. Consider David Lewis's characterization of the 'color mosaic' conception of visual experience:

Those in the traditions of British empiricism and introspectionist psychology hold that the content of visual experience is a sensuously given mosaic of color spots, together with a mass of interpretative judgments injected by the subject.¹⁵

Quality mosaics, involving pitches or smells or tastes, for instance, constitute the sensory fields belonging to each modality. Unsurprisingly, traditional empiricists struggled to secure access to the mind-independent world given this quality-mosaic conception of experience. Where each modality affords awareness only of qualities accessible uniquely to a given sense modality, no obvious mark of the objective or extra-mental exists.

modalities share no phenomenological features. This paper targets all but the weakest, global version of distinctiveness.

¹⁴ Alva Noë's (2004) discussion of what he calls the 'snapshot conception' of visual experience is unconnected with, though it served as inspiration for, this way of thinking of overall perceptual experience.

¹⁵ Lewis 1966, 357.

The composite snapshot conception need not rule out 'common sensibles' accessible to more than one modality. For example, objects and shape properties might be experienced through vision and through touch. This conception does assume, however, that such experiences are modality-specific and distinctive, and that they depend constitutively and exclusively upon awareness of sensibles accessible uniquely some one modality. The experience of seeing an object and its shape consists in experiencing a patterned field of color. Tactile experience of an object and its shape is feeling the texture and resistance a surface offers. There is, therefore, a distinctively visual way of experiencing objects and shapes, which can be captured and characterized exclusively in visual terms, that differs from the tactile experience of objects and shapes.

The traditional conception stems from thinking of the senses as distinct systems or channels of awareness. The sense modalities are understood to involve separate processes, and to work in isolation from each other until some relatively late stage. Each modality as a result delivers experiences with a distinctive qualitative character that could not be replicated by another modality. Each furnishes only an experiential ingredient for one's total perceptual experience.

This conception animates the long history of resistance to answering affirmatively Molyneux's question whether someone without the relevant sort of background experience could visually identify a cube formerly only felt. How, without conjoint experience to associate them, could experiences in distinct modalities seem commensurate? Russell's striking claim that visual, tactile, and auditory space are distinct from each other, and from the space of science, is a symptom of this conception.

To begin with, space as we see it is not the same as space as we get it by the sense

of touch; it is only by experience in infancy that we learn how to touch things we see, or how to get a sight of things which we feel touching us. But the space of science is neutral as between touch and sight; thus it cannot be either the space of touch or the space of sight.¹⁶

If awareness of space consists in awareness involving features unique to a given sense modality, it is unsurprising that the question arises whether space experienced through one modality is identical with space experienced through another.

The puzzle about audition set out earlier traces to the composite snapshot conception. How could audition confer awareness as of something extra-acoustic? Audible qualities belong to sounds, and sounds do not appear to comprise extra-acoustic individuals. The question concerns how we can enjoy auditory experiences as of objects and happenings beyond the proper objects of audition, and how we hear them as the sorts of things we might touch and see, if audition furnishes its own distinctive variety of perceptual experience that can be captured entirely in audition-specific terms. The problem is, given that sounds are so unlike ordinary objects, happenings, and their features, nothing specific to audition seems capable of grounding such awareness.

The traditional story I have been discussing is false in crucial respects and incomplete in others. Appreciating an important class of perceptual effects that has gone unrecognized or underappreciated by philosophers provides good reasons to believe the composite snapshot conception of experience is incorrect. Perceptual experience is not exhausted by an assemblage of discrete, distinctive, modality-specific ingredients. The composite snapshot conception should be abandoned. The cross-modal perceptual illusions I will discuss, however, do not have merely negative implications. They provide

¹⁶ Russell 1912, 29.

the resources for a solution to the puzzle about audition I have described and illuminate perception in perhaps its most significant respect. Coming to terms with cross-modal perception teaches what we could not otherwise have learned with attention restricted to vision, or to any other individual modality.

4. Cross-modal illusions

The cases I have in mind are ones in which what you sense with one modality affects what you experience in another.¹⁷ The well-known ventriloquist effect, for example, involves an illusory experience of the location of a sound that is produced by the sound's apparent visible source. Visible movements of a puppet's mouth affect where you hear the voice to come from. The effect, however, is not limited to the perception of speech. Research has detailed the ways in which even a minimal visual stimulus impacts perceived auditory location.¹⁸ Seeing a circular disk affects where subjects auditorily experience a beep to come from. The effect is neither inferential nor cognitive, but results from cross-modal perceptual interactions.¹⁹

Visual capture of proprioceptive location and tactile size are further examples in which vision impacts spatial experience in another modality. For instance, altering the visible location of your hand by placing it under a prism impacts the proprioceptive experience of your hand's orientation. Presenting you with a photograph of a hand or a

¹⁷ I should be quite clear that I am speaking now of sense modalities as perceptual systems or processes, and not simply as ways of experiencing or kinds of experience. When I say that one sense modality impacts experience in another, I mean that processes associated with one sense modality impact experience associated with another. I am not suggesting, for instance, that one experience impacts another experience.

¹⁸ See Howard and Templeton 1966.

¹⁹ See, e.g., Bertelson 1999; Vroomen et al. 2001.

rubber hand likewise impacts the proprioceptive experience of your hand's location.²⁰ In addition, seeing an object that is larger than one placed in your palm affects the apparent size of the object you are holding.²¹

Cross-modal interactions, however, are not limited to vision's impact upon the experience of space through other sense modalities. The fascinating McGurk effect, for instance, involves a qualitative auditory illusion stemming from the perception of speech.²² To evoke this strikingly robust illusion, subjects are shown video of a speaker articulating the velar /ga/ sound, which is pronounced with the back of the tongue on the soft palate. At the same time, the audio of the bilabial /ba/ sound, pronounced with the two lips together, is presented through speakers. The vivid experience as of hearing the sound of the alveolar /da/, which is pronounced with the tip of the tongue on the palate behind the teeth, results. Simply averting your gaze from the speaker's lips results in a marked change to the phoneme you seem to hear. The gesture that produces the audible /da/ phoneme is in one sense a midpoint between those which produce the /ba/ and /da/ phonemes. The McGurk effect occurs when conflicting auditory and visual information about speech is reconciled into a kind of 'average' or parsimonious percept.

Each of the preceding illusions is compatible with the following explanatory principle: *vision wins*. When, in these cases, information from vision conflicts with what you would expect to experience through another sense modality, vision exerts its influence and alters experience in the other modality. Perhaps straightforward visual dominance suffices to explain the cross-modal illusions. In that case, vision's dominance in perception might vindicate visuo-centric theorizing about perception and perceptual

²⁰ See Hay et al. 1965; Pick et al. 1969.

²¹ Rock and Victor 1964.

²² McGurk and Macdonald 1976.

experience.

It is now clear that this is false. Recent discoveries reveal perceptual illusions that involve the impact of what you sense through another modality upon what you experience visually. For instance, Shams, Kamitani, and Shimojo (2000, 2002) have discovered a class of illusions in which audition affects vision. Shams and colleagues report that when subjects are shown video in which a circular black disk briefly appears or flashes one time, while presented with audio of two brief beeps, they report visually experiencing two flashes. The *sound-induced flash illusion* involves the modulation of visual experience by audition.

A single flash accompanied by multiple beeps is perceived as multiple flashes.

This phenomenon clearly demonstrates that sound can alter the visual percept qualitatively even when there is no ambiguity in the visual stimulus.²³

A number of features of this result are significant.²⁴ First, it does not result from inferential or otherwise cognitive processes, such as deploying a decision strategy for responding to ambiguous or conflicting experiences. Not only do phenomenology and response times tell against this, but cross-modal effects take place where we might expect vision and audition each to be clear and univocal. Second, generating the illusion does not require a semantic contribution, such as that learned through experience with a familiar bimodal context (such as speech perception or musical instruments). Generating the illusion does not require experience within any specific bimodal context; it occurs in naïve subjects in novel situations with simple stimuli. Third, this and other cross-modal effects are automatic and stem from interactions that take place at a relatively low level.

²³ Shams et al. 2002, 152.

²⁴ See Shams et al. 2000, 2002 for discussion.

Bertelson and de Gelder (2004), in the context of spatial cross-modal effects, claim that the relevant processes are pre-attentive: 'Cross-modal interaction reorganizes the auditory-visual spatial scene on which selective attention later operates.'²⁵ Watkins et al. (2006) report audition's impact on vision at some of the earliest stages of visual processing (V1) when the sound-induced flash illusion occurs. The sound-induced flash illusion is a phenomenological change to the character of visual experience that is produced by audition and that persists through changes to stimulus characteristics and setting. It results from cross-modal perceptual mechanisms by which audition impacts visual experience.

5. Explaining cross-modal illusions

Cross-modal illusions are surprising. That vision could be so misled by the mere presence of a sound or that seeing a rubber hand would sway proprioception is hard to believe even after it is demonstrated. Allegiance to an understanding of the senses as discrete, independent, encapsulated modes of perceptual awareness generates the surprise. Given our conception of the role of the sense modalities in perception and perceptual experience, we find their interaction and impact upon each other counterintuitive. According to this view of the senses, seeing is a different affair from hearing, functionally and experientially. Cross-modal interactions put considerable pressure on this way of understanding the role of the senses.

What then are the consequences of these results for thinking about perception and perceptual experience? It is useful to contrast the cross-modal cases I have been discussing with another familiar phenomenon in which stimulation affects experience

²⁵ Bertelson and de Gelder 2004, 165.

usually associated with another sense modality: *synaesthesia*. Those who suffer synaesthesia might systematically and persistently experience colors in response to sounds; they might experience shapes or textures in response to tastes; they might have auditory experiences in response to colors, shapes, patterns, or flavors. Some individuals enjoy color responses to graphemes and numerals.²⁶ Though many synaesthetic effects are thought to be merely associative, some clearly are sensory. Enhanced performance by synaesthetes in what otherwise would be serial search tasks suggests popout effects indicative of sensory phenomenology. For instance, synaesthetes are able quickly to discern a figure eight pattern formed by '5's embedded in a field of '2's.²⁷

Synaesthesia is relatively rare (LSD-induced synaesthesia rarely persists apart from occasional flashbacks). Estimates until recently placed the rate at roughly 1 person in 2,000. New work suggests the vastly higher rate of 1 in 22, but this includes forms of synaesthesia that do not involve qualitative sensory phenomenology.²⁸

Synaesthesia is a quirk of sensory processing. That is, synaesthetic effects result from a kind of processing error or a bit of sensory miswiring. Synaesthesia always involves some illusion. Synaesthetes do not literally perceive the sound of a color, the shape of a taste, or the color of a number since colors lack sounds, tastes lack shapes, and numbers lack colors. One cannot perceive through audition the color of a sounding object (in particular, a color it lacks) or through gustation the shape of tasted food. The experience always involves a measure of misperception. Though in some limited settings, such as specially designed search tasks, synaesthetic experience is helpful, it does not in

²⁶ For further background and discussion, see Baron-Cohen and Harrison 1997; Cytowic 1998, 2002; Harrison 2001; and Macpherson 2007.

²⁷ See Ramachandran and Hubbard 2001, 2003.

²⁸ See Baron-Cohen et al. 1996 and Simner et al. 2006, respectively.

general result from principled perceptual strategies tailored to deal with the world.

Perhaps cross-modal illusions are simple quirks of processing that result from accidents of sensory "wiring". This explanation simply misses what makes the cross-modal illusions I have discussed remarkable. To start, the cross-modal illusions are common across the population of normal perceivers. Neither drugs nor a special innate condition is prerequisite to experiencing these illusions. But what is most noteworthy about the cross-modal illusions is that they are intelligible responses to unusual or extraordinary circumstances. For instance, the location of an audible sound does not frequently diverge from that of its visible source. The ventriloquist's impressive trick exploits this. The sound-induced flash illusion, on the other hand, relies on the fact that visible events that unfold in time usually correspond in number to their acoustic signs. Reconciling information that conflicts with these general regularities makes sense from the standpoint of a system deployed in determining the arrangement of one's surroundings. These kinds of responses in fact result in non-illusory experiences when applied across a wide range of natural and artificial circumstances. Given that sensory stimulation is noisy and fallible, information conflicts are best resolved.²⁹ It is safer to assume that a sensory receptor has gone amiss than to violate a natural constraint or a general principle concerning the organization of the perceptually available world. Thus, it appears that the cross-modal illusions result from perceptual organizing strategies or principles that in general are adaptive and advantageous.

Consider, however, what makes these interactions and the illusions they produce intelligible as adaptive and useful. In the first place, they are fallible since sometimes, as

²⁹ Most often, spatial conflicts are resolved in favor of vision, and temporal conflicts are resolved in favor of audition. Vision's spatial accuracy exceeds audition's, and audition's temporal accuracy exceeds vision's. See, e.g., Guttman et al. 2005 for discussion.

when cross-modal illusions occur, they produce non-veridical experiences. They nonetheless appear to involve rules for modulating experiential responses to sensory information, rules that deal with important regularities. The sensory responses in question, however, must be understood to stem from a *common environmental source* for these rules to make sense as ways for coping with an environment. That is, given divergent auditory and visual stimulation, it only makes sense to attempt in a principled manner to reconcile them if they are assumed to share a common source or cause. Otherwise, the notion that there is a conflict that requires resolution is unintelligible. Explaining cross-modal perceptual interactions thus requires "the supposition that the intersensory bias is a result of an attempt by the perceptual system to maintain a perceptual experience consonant with a unitary event."³⁰

This idea has been reflected in what have been called 'unity assumptions' for cross-modal interactions. Such assumptions modulate how a perceptual unit is formed on the basis of both auditory and visual stimulation according to principles analogous to those involved in Gestalt formation within a single modality. For instance, just as spatial continuity and cohesion govern the perceptual experience of visual objects, even when parts are occluded, temporal coincidence and spatial proximity are part of what regulates which auditory and visual features belong together. Thus, they regulate which cross-modal interactions, recalibrations, and illusions occur. Explaining bias and influence across the modalities invokes the perceptual "assumption" that an auditory and a visual stimulus belong to a unified environmental source.

Two things are significant about this result. First, deploying such unity assumptions amounts to exercising a kind of perceptual grasp upon items in the

³⁰ Welch and Warren 1980, 638.

environment that are available to multiple modalities, or upon common perceptual objects. We perceptually track such particulars in a way that is at once resistant to confounding "noise" and responsive to countervailing information from different sensory pathways. Second, however, such a perceptual grasp involves a *multi-modal* or *modality-independent* way, which is not proprietary to a given modality, of grasping or representing such environmental particulars since the items in question must be understood as those items which lead to *both* auditory and visual stimulation. These unity assumptions governing cross-modal interactions thus amount to multi-modal or modality-independent assumptions concerning the common sources of sensory stimulation. And so, we perceptually track things in a way that goes beyond what might be characterized in modality-specific terms, such as responsiveness to 'visual objects' or 'auditory streams'. Instead, we grasp or represent a class of items in terms that reflect their multi-modal significance. We grasp or represent them as multi-modal or modality-independent individuals, objects, or events.

The discussion so far tells against the claim that perceptual systems corresponding to the sense modalities are strongly modular. Since information from one sense modality regularly impacts processing and experience associated with another sense modality, even at quite early stages, the senses are not entirely causally and informationally encapsulated systems.

It also shows that there is a subpersonal grasp, at the level of sensory or perceptual processing, on sources of stimulation that must be understood in multi-modal or modality-independent terms. If you are willing to attribute content to subpersonal perceptual states, the corresponding states possess multi-modal content. But this itself

does not strictly speaking imply anything concerning how things seem from the point of view of the subject. Claims concerning subpersonal perceptual mechanisms are notoriously difficult to connect with claims concerning the content and phenomenology of perceptual experience.

6. Cross-modal perceptual experience

A gap exists between the claims concerning subpersonal perceptual processes (and their explanation) and claims concerning the content and phenomenology of perceptual experience. One concern is that positing a subpersonal grasp upon environmental particulars which must be understood or represented in multi-modal or modality-independent terms does not imply that this grasp, or corresponding representational state, ever is manifested as such in perceptual experience. The cross-modal illusions demonstrate that perceptual processes ordinarily associated with one sense modality affect processes associated with another sense modality and thus *causally* impact perceptual experience ordinarily associated with that modality. This by itself implies neither that the resulting experiences do not occur in distinct modalities nor that they cannot be characterized exhaustively in modality-specific terms. The cross-modal interactions might thus shape or causally determine the character of perceptual experience without making perceptually apparent their modality-bridging influence. There is therefore a gap between the claim that multi-modal organizing principles causally influence subsequent experience and the claim that the composite snapshot conception of perceptual experience fails.

I wish to argue that understanding cases of cross-modal perception grounds an

argument for the claim that there exist consciously accessible aspects of perceptual experience that are not unique or specific to a given experiential modality and that may be shared across modalities. The argument proceeds in two stages. The first aims to show that there is a dimension or component of perceptual content that must be characterized in multi-modal or modality-independent terms. This component either is shared by both vision and audition or outstrips both the visual and the auditory. The second aims to show that such content is phenomenologically apparent -- there is an important aspect even of the phenomenology of seeing that it must share with hearing. The phenomenology of perceptual experience therefore cannot be characterized exhaustively in modality-specific terms.

It is worth noting that I am offering a philosophical argument for the existence of common contents across perceptual modalities, as traditionally understood, that is grounded in empirical evidence and the explanation for empirically detectable perceptual phenomena. It is not based simply upon the reflective philosophical intuitions standardly mustered in favor of this claim. My primary aim, however, is to establish common phenomenology, which runs strictly counter to the composite snapshot conception of perceptual experience.

Consider the correctness conditions for a given experience in which subpersonal cross-modal principles are invoked, such as the experience of a single flash-beep pairing.³¹ What are the conditions under which such an experience is veridical? The experience is *not* entirely veridical in a situation in which the visual flash and the

³¹ The argument goes through with each of the cross-modal cases I have discussed, such as seeing and feeling the location of one's hand, experiencing the McGurk effect, or undergoing the ventriloquist illusion. Note that I do not have in mind just those cases in which illusion takes place. Rather, I include the many instances in which cross-modal organizing principles lead to veridical experience.

auditory beep have entirely separate and unconnected environmental sources. (That is why ventriloquism and visual capture of a rubber hand involve more than just locational illusion.) Given the way the (spatial and temporal) relationships between the visual and auditory stimuli modulate your experience thanks to cross-modal organizing principles, the experience ordinarily is accurate in case it depends upon a single event -- a single event that is responsible for the auditory and the visual experience. Coincidental beeps and flashes, however, do not suffice to characterize the circumstances conveyed by the perceptual experience. If, for instance, there had been two beeps, you would not still have seen a single flash, as would be anticipated with entirely unconnected stimuli. If the flash had occurred to the left, the beep would have been heard to the left, indicating a common source. To capture all that the experience purports to reveal, there must exist a single thing that is both seen and heard -- a single environmental event connected with the visual and auditory experiences. The very same multi-modal or modality-independent content therefore either is shared by vision and audition or outreaches them both.

Furthermore, the commonality is evident even at the level of phenomenology. Some might object that in the above example, if we consider phenomenology alone, we can imagine the perceptual experience to be entirely veridical even if the visual and auditory experiences share no common source. Though I find this unconvincing in the case above and, especially, in speech perception and visual capture, others might be less impressed by the introspective, counterfactual, and behavioral evidence. So, to illustrate the point, consider a cross-modal experience that involves *inter-modal binding*. Just as you might see a certain individual to be both red and round due to the effects of *intra-modal* feature binding, you also might perceive something to be both red and rough, or

red and solid, or loud and bright due to the effects of inter-modal binding. The same object or event seems to possess both visible and tactile features, or visible and audible features, thanks to the perceptual assignment of different modality-specific perceptible features to a common sensory individual or item. This explains how we perceptually experience the same item to both resist squishing and occlude a visual angle. It also explains how we perceptually experience the high-pitched noise to belong to the small, silvery cylinder nearby and the low-pitched noise to belong to the large, brassy coil across the room. A single item must bear both visible and tactile or visible and auditory features to account for perceptual identification and feature grouping across modalities. A case in which there exists one item with the visible features and another distinct item with the tactile features fails to exhaust the veridicality conditions for such an experience. It therefore fails as a complete characterization of the phenomenology of the experience. There must be a phenomenologically accessible aspect of perceptual experience that corresponds to the experience as of a common particular and thus is shared across modalities or goes beyond what is specific to any given modality. You experience the cube you hold as the very same particular you see. You experience the blip you hear to stem from the same event as the flash you see. Perceptual experience therefore has a dimension that cannot be captured by a composite of *modality-specific*, proper, or unimodal snapshot-like components or contents. An element of experience shared by vision and audition, or that outstrips strictly visual or strictly auditory aspects of experience, is required to capture the phenomenological feel of perceptual experiences in which inter-modal binding takes place.

7. The puzzle and its source

I suggested earlier that the puzzle concerning audition -- how hearing could involve experience as of something extra-acoustic, or as of the sort of thing you might see -- stems from the traditional understanding of the senses as distinctive and discrete avenues of awareness. According to what I have called the composite snapshot conception of perceptual experience, one's total perceptual experience consists in and is exhausted by distinctive modality-specific varieties of experience. The arguments of the last section, in particular, demonstrate that this conception is inadequate as a characterization of perceptual experience. I have argued that characterizing perceptual experience requires appeal to multi-modal or modality-independent content and phenomenology.³² This entails either that your auditory, visual, etc., experiences do not exhaust your total perceptual experience or that your auditory, visual, etc., experiences cannot be characterized entirely in proprietary or modality-specific terms.³³

Consider, first, the claim that perceptual experience is exhausted by visual, auditory, olfactory, gustatory, and tactile-kinesthetic experiences. Since each such experience may include a component that could be shared across perceptual modalities, experience associated with a given modality cannot be characterized entirely in modality-specific terms. It follows that not every experience in a given modality such as vision or audition, or every aspect of such an experience, is entirely distinctive of that modality. If perceptual experience is exhausted by experiences corresponding to the different sense

³² My view, therefore, is less extreme than Tye's (2003, 2007), since I do not deny that one's overall perceptual experience consists in component experiences. I believe that phenomenological unity of perceptual experience is compatible with different kinds and measures of disunity, within and among modalities.

³³ Nudds (2001) appears to endorse the first disjunct, but rejects the second. I am more sympathetic to the second, as I suggest in endnote 34.

modalities, then it is not the case that, for each item you can experience through a given modality, there exists a modality-specific way of experiencing that item. If exhaustiveness holds, distinctiveness fails.

Suppose, on the other hand, that experiences in the different sense modalities uphold the principle of distinctiveness. That is, for any item that you can experience with a given modality, there must be a distinctive, modality-specific way of experiencing that item. Given that perceptual experiences may attribute features associated with multiple sense modalities to a common particular, some aspect of perceptual experience must explain your experience as of a single item that causes or bears features or constituents experienced through different modalities. If distinctiveness holds, some characteristic of perceptual experience that is not unique to experiences associated with a given modality must serve to represent or identify a common environmental source. Your overall perceptual experience therefore is not exhausted by experiences that are distinctive to the different sense modalities. If distinctiveness holds, exhaustiveness fails.

I conclude that the composite snapshot conception of perceptual experience involves commitment to mutually inconsistent claims.^{34, 35}

³⁴ Forced to choose between exhaustiveness and distinctiveness, I prefer exhaustiveness. Consider perceptually experiencing visual and tactile features to belong to some particular. Vision and touch each ascribe visible or tactile features to something, to some sensible individual -- neither seeing nor touching is mere awareness of sensible qualities. Since visual and tactile experiences intuitively implicate the same individual, however, the cost of retaining distinctiveness is multiplying perceptible individuals. Retaining exhaustiveness involves the more attractive prospect of recognizing sharing between vision and touch. Nothing in principle, however, rules out different explanations for different pairings.

³⁵ Though it is not my concern in this paper, one might choose simply to reject that perceptual experience comprises auditory, visual, etc., experiences, where each involves a distinct modality of sense perception. If the sense modalities cannot successfully be individuated, perhaps in part due to considerations I have addressed in this paper, then the composite snapshot conception of perceptual experience fails if it commits one to distinct sense modalities. Concerning the individuation of sense modalities, see, e.g., Grice 1962; Heil 1983; O'Regan and Noë 2001; Keeley 2002; Nudds 2003.

We now have a foothold toward a solution to the puzzle about audition I set out earlier. There is a component of the content and a dimension of the phenomenology of perceptual experience that cannot be captured in terms that are distinctive or specific to a given modality, but must invoke multi-modal or modality-independent terms. We might therefore hear, or perceive in virtue of hearing, as of things that cannot be captured in purely auditory terms, and see, or perceive in virtue of seeing, as of things that cannot be captured in purely visual terms. Either auditory and visual experience share more than we thought, or perceptual experience exceeds what is distinctively auditory or visual.

The important consequence is that, either way, perceptual experiences are capable of reaching beyond awareness as of intentional objects distinctive to each of the modalities of sensory awareness. For instance, auditory perceptual experiences are not limited to sounds, their pitch, timbre, loudness, and other distinctively audible qualities. We might thus enjoy experiences as of something, such as a sound *source*, a sounding *object*, or a sound-generating *event*, which can bear extra-acoustic features. We might hear sources, objects, and events, and not just sounds, pitches, and timbres. We also might experience ordinary material objects and events, and not just visual objects, colors, and light, thanks to vision. We might feel surfaces and solid objects, and not just pressure, texture, and warmth, thanks to touch. But we hear, see, and feel as of such common individuals because the senses do not act as isolated systems that deliver only neat, modality-specific experiences from which we somehow learn to infer the presence of ordinary objects and events. It seems fair to suppose that the sense in which perceptual experience connected with any modality seems to involve ordinary objects or events accessible through multiple perceptual modalities stems from the multi-modal principles

responsible for organizing cross-modal experiences, including the cross-modal illusions. Part of the world-involving character of perceptual experience therefore stems from inter-sensory processes.

8. Perceiving across the modalities; seeing through the senses

I have presented a set of reasons to believe that proprietary ways of perceiving unique to each individual sense modality fail to exhaust the content and phenomenology of perceptual experience. Convincingly explaining cross-modal perceptual illusions requires perceptual mechanisms that modulate the impact of stimulation to one sense modality upon experience commonly associated with another sense modality. Such mechanisms are intelligible as principled perceptual strategies only if they involve a kind of grasp upon environmental items that are the sources of stimulation across multiple modalities. This alone explains the need to reconcile divergent stimuli -- indeed, taking divergent stimulation in separate sensory modalities as conflicting requires treating it as unified, or as providing information concerning a common source. Such items, however, cannot be grasped or understood in terms that are specific or proper to a given sense modality. To be deployed in the modulation of experience across multiple modalities requires that they be construed in multi-modal or modality-independent terms.

I have argued, however, that this multi-modal or modality-independent grasp is not limited to the subpersonal level. Rather, by considering the correctness conditions for perceptual experiences in cases in which cross-modal or multi-modal organizing principles might impact experience in multiple modalities, we learn that characterizing perceptual content requires invoking multi-modal or modality-independent constituents --

perceptual individuals that bear both visible and audible features. Furthermore, granting that common items might be perceived through distinct modalities does not mean we still can exhaustively characterize the phenomenology of perceptual experience by citing only peculiar, distinctive, modality-specific ways of experiencing those items. Rather, the point goes through even at the level of phenomenology. As the case of inter-modal binding demonstrates, some phenomenologically-apparent aspect of the experience must correspond to the experience of individuals whose features are available to distinct modalities but nonetheless are experienced as coinstantiated. Such an element of perceptual experience cannot be unique to any particular sense modality.

It follows from this that any perceptual snapshot associated with a specific modality already is a multi-modal sculpture infused with information derived from other sensory systems. Perceptual experience is not simply an assemblage of discrete, modality-specific component experiences, since it outstrips peculiar sense-specific ways of experiencing. Though some aspects of perceptual experience might remain proprietary, the multi-modal aspects serve to tie experience into a perceptual whole replete with features drawn from varied sources. The mark of the source, however, need not be phenomenologically apparent in each constituent of the experiential whole. The traditional composite snapshot conception of perceptual experience fails.

If this is correct, then no particular modality of sense perception can be understood or characterized entirely in isolation from the others. Input to sensory processing is not limited to the stimulation of a sensory surface associated with a given sense modality, and subperceptual processes are not causally isolated into discrete sensory pathways. Furthermore, the processes that lead to and organize perceptual

experience are intelligible as advantageous only if the contributions of the several senses are considered collectively, and, in particular, only if such processes implement a kind of modality-independent grasp upon common sources of stimulation. A similar claim concerning the intelligibility of individual modalities holds even at the level of the content and phenomenology of perceptual experience. Without appealing to other sense modalities, such as audition and touch, we are unable to characterize exhaustively even visual experience. Ignoring the non-visual modalities leaves out a component critical for understanding vision and visual experience. Not only is vision responsive to information from the other senses, but individuals accessible to non-visual modalities constitutively shape vision's content and phenomenology. Sensory atomism is false.³⁶

More importantly, its failure reveals a perilous flaw in the visuo-centric thinking from which it stems. Taking vision independently as a representative paradigm for theorizing about perception and perceptual experience not only is incomplete, it also encourages us to think we can understand each perceptual modality as an autonomous mode of awareness and domain for philosophical and scientific inquiry. This, however, leaves out what is most critical for resolving long-standing philosophical disputes about the possibility and grounds of perceptual access to the world of things and events. It thus undermines a comprehensive understanding of perception and perceptual experience. If, for instance, interactions and relationships among perceptual modalities reveal that a common multi-modal or modality-independent lexicon is shared among them, then our perceptual sense that the world comprises a complex but unified whole independent from our experiences plausibly stems from multi-modal awareness. Perhaps in grasping

³⁶ I thank John Doris for urging the expression 'sensory atomism'. This captures the sense in which each modality offers independent sense-specific elements that jointly exhaust perceptual experience.

multiple sensory perspectives as perspectives on a common source, perceptual independence takes hold. What is most striking and important about perceptual experience thus may depend upon the relationships and interactions among sensory modalities. Comprehending such relationships and interactions is helpful in resolving puzzles concerning audition's object- and event-involving character, but it may prove essential to any satisfactory philosophical understanding of perception. The tyranny of the visual threatens to blind us to the nature, character, and scope of perceptual experience.³⁷

References

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., and Bolton, P. (1996). Synaesthesia: Prevalence and familiarity. *Perception*, 25(9):1073–1079.

Baron-Cohen, S. and Harrison, J. E. (1997). *Synaesthesia: Classic and Contemporary Readings*. Blackwell, Malden, MA.

Batty, C. (2007). *Lessons In Smelling: Essays on Olfactory Perception*. PhD thesis, MIT, Cambridge, MA.

Bermúdez, J. L. (2000). Naturalized sense data. *Philosophy and Phenomenological Research*, 61(2):353–374.

Bertelson, P. (1999). Ventriloquism: A case of cross-modal perceptual grouping. In Aschersleben, G., Bachmann, T., and Müsseler, J., editors, *Cognitive Contributions to the Perception of Spatial and Temporal Events*, pages 347–362. Elsevier, Amsterdam.

Bertelson, P. and de Gelder, B. (2004). The psychology of multimodal perception. In Spence, C. and Driver, J., editors, *Crossmodal Space and Crossmodal Attention*, pages

³⁷ I am grateful to a number of people for conversations, comments, and encouragement that helped me to improve this paper. In particular, I thank José Bermúdez, Dave Chalmers, John Doris, Benj Hellie, Tony Jack, Brian Keeley, Adam Morton, Alva Noë, Mark Okrent, Jesse Prinz, Susanna Siegel, and Jeff Speaks. I expressly thank Mohan Matthen for his extensive and attentive comments on a draft of this paper. I presented versions of this material at Toward a Science of Consciousness 2006 in Tucson, University of Maine, Orono, University of Notre Dame, Washington University in St. Louis, and the Society for Philosophy and Psychology 2007 meeting in Toronto. Thanks to members of those audiences for their challenging questions.

141–177. Oxford University Press.

Campbell, J. (2002). *Reference and Consciousness*. Clarendon Press, Oxford.

Carroll, N. (1985). The power of movies. *Daedalus*, 114(4):79–103.

Casati, R. and Dokic, J. (1994). *La Philosophie du Son*. Chambon.

Casati, R. and Dokic, J. (2005). Sounds. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*.

Chalmers, D. J. (2004). The representational character of experience. In Leiter, B., editor, *The Future for Philosophy*, pages 153–181. Oxford University Press, Oxford.

Chalmers, D. J. (2006). Perception and the fall from Eden. In Gendler, T. S. and Hawthorne, J., editors, *Perceptual Experience*, pages 49–125. Clarendon Press, Oxford.

Cytowic, R. E. (1998). *The Man Who Tasted Shapes*. MIT Press, Cambridge, MA.

Cytowic, R. E. (2002). *Synesthesia: A Union of the Senses*. MIT Press, Cambridge, MA, 2nd edition.

Grice, H. P. (1962). Some remarks about the senses. In Butler, R. J., editor, *Analytical Philosophy, Series 1*. Blackwell, Oxford.

Guttman, S. E., Gilroy, L. A., and Blake, R. (2005). Hearing what the eyes see. *Psychological Science*, 16(3):228–235.

Harrison, J. (2001). *Synaesthesia: The Strangest Thing*. Oxford University Press, New York.

Hay, J. C., Pick, H. L., and Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychonomic Science*, 2:215–216.

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

Howard, I. P. and Templeton, W. B. (1966). *Human Spatial Orientation*. Wiley, London.

Keeley, B. L. (2002). Making sense of the senses: individuating modalities in humans and other animals. *The Journal of Philosophy*, 99(1):5–28.

Lewis, D. (1966). Percepts and color mosaics in visual experience. *The Philosophical Review*, 75:357–368.

Lycan, W. (2000). The slighting of smell. In Bhushan, N. and Rosenfeld, S., editors, *Of*

- Minds and Molecules: New Philosophical Perspectives on Chemistry*, pages 273–89. Oxford University Press, Oxford.
- Macpherson, F. (2007). Synaesthesia. In de Caro, M., Ferretti, F., and Marraffa, M., editors, *Cartographies of the Mind: Philosophy and Psychology in Intersection*, volume 4 of *Studies in Brain and Mind*. Kluwer, Dordrecht.
- Martin, M. (1992). Sight and touch. In Crane, T., editor, *The Contents of Experience*. Cambridge University Press, Cambridge.
- Matthen, M. (2005). *Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception*. Oxford University Press, Oxford.
- McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264:746–748.
- Noë, A. (2004). *Action in Perception*. MIT Press, Cambridge, MA.
- Nudds, M. (2001). Experiencing the production of sounds. *European Journal of Philosophy*, 9:210–229.
- Nudds, M. (2003). The significance of the senses. *Proceedings of the Aristotelian Society*, 104(1):31–51.
- O’Callaghan, C. (2007). *Sounds: A Philosophical Theory*. Oxford University Press, Oxford.
- O’Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24:939–1031.
- Pasnau, R. (1999). What is sound? *Philosophical Quarterly*, 49:309–324.
- Pasnau, R. (2000). Sensible qualities: The case of sound. *Journal of the History of Philosophy*, 38:27–40.
- Pick, H. L., Warren, D. H., and Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception and Psychophysics*, 6:203–205.
- Ramachandran, V. S. and Hubbard, E. M. (2001). Synaesthesia — a window into perception, thought, and language. *Journal of Consciousness Studies*, 8(12):3–34.
- Ramachandran, V. S. and Hubbard, E. M. (2003). The phenomenology of synaesthesia. *Journal of Consciousness Studies*, 10(8):49–57.
- Rock, I. and Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, 143(3606):594–596.

- Russell, B. (1912). *The Problems of Philosophy*. Oxford University Press, London.
- Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature*, 408:788.
- Shams, L., Kamitani, Y., and Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14:147–152.
- Siegel, S. (2006a). Subject and object in the contents of visual experience. *Philosophical Review*, 115(3):355–388.
- Siegel, S. (2006b). Which properties are represented in perception? In Gendler, T. S. and Hawthorne, J., editors, *Perceptual Experience*. Oxford University Press, New York.
- Siegel, S. (2008). The visual experience of causation. *Philosophical Quarterly*.
- Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., Scott, K., and Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, 35(8):1024–1033.
- Smith, A. D. (2002). *The Problem of Perception*. Harvard University Press, Cambridge, MA.
- Snowdon, P. (1992). How to interpret ‘direct perception’. In Crane, T., editor, *The Contents of Experience*, pages 48–78. Cambridge University Press, Cambridge, UK.
- Tye, M. (2003). *Consciousness and Persons: Unity and Identity*. MIT Press, Cambridge, MA.
- Tye, M. (2007). The problem of common sensibles. *Erkenntnis*, 66:287–303.
- Vroomen, J., Bertelson, P., and de Gelder, B. (2001). Auditory-visual spatial interactions: Automatic versus intentional components. In de Gelder, B., de Haan, E., and Heywood, C., editors, *Out of Mind*, pages 140–150. Oxford University Press, Oxford.
- Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., and Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, 31(3):1247–1256.
- Welch, R. B. and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3):638–667.