Grades of Multisensory Awareness

Casey O'Callaghan

Mind & Language

November 25, 2015, final version

words: 10865 inclusive

Abstract: Psychophysics and neuroscience demonstrate that different sensory systems interact and influence each other. Perceiving involves extensive cooperation and coordination among systems associated with sight, hearing, touch, smell, and taste. Nonetheless, it remains unclear in what respects conscious perceptual awareness is multisensory. This paper distinguishes six differing varieties of multisensory awareness, explicates their consequences, and thereby elucidates the multisensory nature of perception. It argues on these grounds that perceptual awareness need not be exhausted

---

Warm thanks to Tim Bayne, Brian Keeley, Peter Ross, and anonymous reviewers for *Mind & Language*. Each offered extensive, valuable comments that helped me to improve this paper. I presented this material at the Universities of London, Toronto, Helsinki, and Milan, Pitzer College, and the 2015 Cognitive Science Society Meeting in Pasadena. Many thanks to audience members on those occasions for questions and discussion, and to the organizers for their efforts.

Address for correspondence: Department of Philosophy, One Brookings Drive, Saint Louis, MO 63130–4899, USA

Email: casey.ocallaghan@wustl.edu

> by that which is associated with each of the respective sensory modalities along with whatever accrues thanks to simple co-consciousness.

Theorizing about perception has been shaped to a remarkable extent by attention to vision and visual forms of awareness. Recently, philosophers have worked to remedy this by focusing on other senses. There are now mature philosophical contributions addressing hearing, touch, smell, and taste (see Matthen, 2015b, part III). Such work aims to translate, extend, challenge, and unify our understanding of perception across its sensory modalities. Attention to non-visual senses is a thriving interdisciplinary research program. This is a promising development for the philosophy of perception.

But it does not go far enough. There remains a tempting thought: *Perceptual awareness amounts to a collection of visual, auditory, tactual, gustatory, and olfactory episodes.* So, once we have told the story about perceiving for each modality, we will have said all there is to say about exteroceptive sensory perception.

Behind this tempting thought is an assumption about how the individual sense modalities work.

> [V]isual perception . . . is best viewed as a separate process with its own principles and possibly its own internal memory . . . isolated from the rest of the mind except for certain well-defined and highly circumscribed modes of interaction. (Pylyshyn, 1999, p. 364)

However, one of the most fascinating lessons to emerge from recent psychophysics and neuroscience is that different sensory systems interact and influence each other. Recognizing and exploring this has spurred dramatic development in the cognitive sciences of perception during the past two decades. What we have learned is that perceiving does not just involve visual, auditory, tactual, olfactory, and gustatory systems working in parallel and in isolation. It involves extensive cooperation and coordination among the senses. So, theorizing about individual modalities and treating them as explanatorily independent risks failing to appreciate the ways in which perceiving with one sense depends upon and affects how we perceive with the others.

What remains mysterious is how all of this interaction and coordination is reflected in the conscious lives of perceiving subjects. Claims about perceptual processes and mechanisms notoriously do not translate neatly and uncontroversially into claims about perceptual experience (see, e.g., Macpherson, 2011; Deroy et al., 2014).

In this paper, I focus on the implications concerning perceptual awareness. I distinguish six differing ways in which conscious perceptual awareness may be multisensory. Each marks an increasingly rich grade of multisensory involvement in perceiving. Each grade requires increasingly rich explanatory resources to accommodate it within an account of perceptual awareness. Each requires a greater departure from the sense-by-sense approach. Each has correspondingly stronger consequences for how we understand and theorize about the nature of perception.

My aim here is neither to refute skeptics about multisensory awareness, such as Spence and Bayne (2015), nor to settle disputes among experimentalists. Instead, I describe the evidence for each differing variety and advance the case for the non-skeptical position. This provides the tools for future debates. My accounting is not exhaustive, and it leaves open to which degree perceptual awareness is multisensory. Together, however, these varieties of multisensory awareness enable us to see how the tempting thought that perceptual awareness must be structured as a mere collection of visual, auditory, tactual, gustatory, and olfactory episodes is mistaken. It fails because perceptual awareness on each occasion need not be exhausted by that which is associated with each of the respective modalities along with whatever accrues thanks to simple co-consciousness. In distinguishing these six varieties of multisensory awareness and explicating their consequences, this paper thereby elucidates the multisensory nature of perception.

1. Grade 1: Minimally Multisensory Awareness

People see, hear, touch, smell, and taste. They do so at the same time, and they do so co-consciously. So, perceptual awareness is at least minimally multisensory. By this I mean that it is possible for a subject to undergo episodes of co-conscious perceptual awareness associated with more than one exteroceptive sensory modality at a time.

This is the **1st grade** of multisensory awareness. It is relatively innocuous, but it is not entirely innocuous. Spence and Bayne (2015) are skeptical whether perceptual experience is, even in this very minimal sense, multisensory. They argue that perceptual consciousness at any moment is unisensory and switches quickly back and forth between senses.

I reject the unisensory view. It is most plausible if consciousness requires attention and if attention is restricted to one modality at each time. I set aside the controversy about whether consciousness requires attention. If it does, whether consciousness is unisensory is a trivial consequence if attention is unisensory. However, it is plausible that attentional resources can be allocated to different modalities at one time. For instance, a simultaneous sound can diminish visual attentional blink, repetition blindness, and backward masking, as reviewed in Deroy et al. (2014). In these multisensory conditions, devoting attentional resources to audition affects how they are devoted at once to vision. In addition, it is plausible that there can be multisensory objects of attention (see, e.g., Kubovy and Schutz, 2010). Even so, there is a more direct argument. There need not be an apparent temporal gap between experiences that are associated with distinct modalities—one sometimes seems seamlessly to follow another. And since the temporal grain of the experienced present sometimes is coarser than that of such rapid conscious shifts between modalities, temporal parts of experiences associated with different senses sometimes seem to fall within the same experienced present. Thus, they seem to overlap or to be simultaneous. Since seemingly simultaneous experiences typically are co-conscious, it follows that there are times during which experience is at least minimally multisensory.

As it stands, this is a weak claim. Failing to find evidence for further grades of multisensory awareness does not show that perceptual consciousness is not at least minimally multisensory (cf., Spence and Bayne, 2015). But we can strengthen it and capture the tempting thought. Say that perceptual awareness at each moment is exhausted by that which is associated with each of the respective modalities, along with whatever accrues thanks to mere co-consciousness (cf., O'Callaghan, forthcoming). Perceptual awareness then just is the co-conscious sum of its modality-specific parts or features or aspects. This captures the tempting thought.

2. Grade 2: Coordinated Multisensory Awareness

Cross-modal perceptual illusions challenge the explanatory independence of the senses. These are cases in which stimulation to one sensory system impacts and reshapes experience associated with another in a way that leads to misperception. Familiar examples include: ventriloquism, an auditory spatial illusion produced by vision; the McGurk effect, in which vision impacts speech perception; the rubber hand illusion, involving visual capture of proprioceptive location; the sound induced-flash effect of audition on vision; and the parchment skin illusion, an auditory influence on touch.

Just as visual illusions teach us about visual processing and the organization of visual perception, crossmodal illusions illuminate multisensory processes and the organization of multisensory perception. Unlike cross-sensory synesthesia, these effects are widespread, and they result from principled perceptual strategies that are intelligible as adaptive and as epistemically advantageous (see O'Callaghan, 2012). The leading hypothesis is that they improve accuracy and enhance the overall reliability of perception.

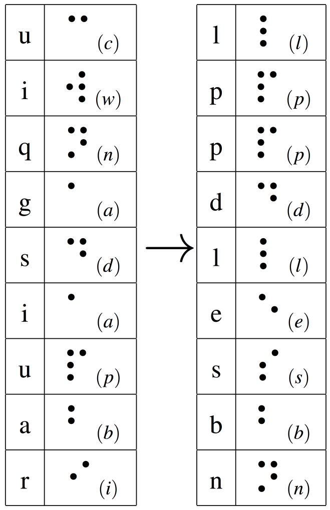
> Altogether these findings suggest that in carrying out basic perceptual tasks, the human perceptual system performs causal inference and multisensory integration, and it does so in a fashion highly consistent with a Bayesian observer. This strategy is statistically optimal as it leads to minimizing the average (squared) error of perceptual estimates; however, it results in errors in some conditions, which manifest themselves as illusions. (Shams and Kim, 2010, p. 280)

Crossmodal perceptual illusions are evidence of a **2nd grade** of multisensory awareness beyond the first. This involves coordinated perceptual awareness across the senses. The senses are not working entirely independently from each other. Instead, there are mechanisms for recalibrating and coordinating their responses in relation to each other. Such coordinated awareness requires but is not entailed by minimally multisensory awareness.

Do cross-modal illusions show more? I have argued previously that these processes stereotypically involve reconciling conflicting or discrepant information from different senses. For instance, in the McGurk effect, conflicting visual and auditory information about a spoken utterance is resolved. In ventriloquism, the disagreement resolved concerns space. But, conflict requires a common subject matter. Even merely apparent conflict requires the presumption of a common subject matter. So, doing conflict resolution demonstrates a perceptual concern for common features or sources of stimulation to multiple senses. Implementing principles of conflict resolution is being differentially sensitive to common items or features across modalities. In the McGurk effect, the common concern is a vocal gesture or phonological feature. In ventriloquism, it is location. This suggests that there is a way of perceiving or representing common features or sources that cannot be characterized in sense-specific terms—a shared perceptual grasp of speech or of space (O'Callaghan, 2012).

This line of reasoning has two important limitations. First, doing conflict resolution does not require a way of perceiving or representing common features or sources that is shared between senses, and it does not require perceiving or representing common features or sources as such.

To illustrate this, consider a simple system that takes as input any (*Roman*, *Braille*) character pair and yields as output a correctly matched (*Roman*, *Braille*) character pair. It does conflict resolution. Figure 1 represents sample input and output from such a system. In Figure 1, the output is the rounded average of the alphabetical positions of the inputs. Other systems may conform to other principles, such as deference to the Braille. In resolving conflicts, this system implements a grasp on the common letters picked out by Roman and Braille characters. But it need not include any *shared* representations of common letters, and it need not rely on representing *that* the Roman and Braille characters pick out a common letter. It could work by brute force using a lookup table or by a simple set of if–then rules relating inputs to matched outputs. It need not include or make use of representations of common letters as such.



[[FIGURE 1 HERE, CAPTION: Figure 1. *Conflict resolution*. Sample inputs (left) and outputs (right) of a system that reconciles conflicts between mismatched Roman and Braille characters (corresponding Roman character shown in parentheses).]]

By analogy, in explaining multisensory perception, we only need to appeal to modality-specific ways of perceiving or representing things that in fact may be common targets of multiple senses, along with mechanisms for coordinating and bringing them into alignment.

There is a second limitation. I emphasized above that claims about perceptual processes do not translate straightforwardly into claims about perceptual awareness. Even if crossmodal perceptual processes target common features or sources of stimulation as such, perceptual awareness might just be structured as a co-conscious collection of coordinated but modality-specific experiences. From the point of view of the conscious subject, all perceptual awareness might remain sense specific.

Given these two limitations, coordinated perceptual awareness across the senses therefore is compatible with multisensory perceptual awareness being exhausted by that which is associated with each of the respective modalities along with whatever accrues thanks to mere co-consciousness. The tempting thought remains safe in the face of the second grade.

3. Grade 3: Intermodal Binding Awareness

Are there any core, irreducibly multisensory varieties of perceptual awareness? A critical case is that of perceptually apparent intermodal feature binding. This is a **3rd grade** of multisensory awareness. It is critical because it marks the point at which perceptual awareness can no longer be characterized in modality-specific terms (see also O'Callaghan, 2014).

Humans perceive individual things and their features. Perceptible individuals include objects and events, and among perceptible features are attributes and parts. Individuals can be perceived at once to have multiple features. When you consciously perceive multiple features jointly to belong to the same individual or to be coinstantiated, call that a case of *feature binding awareness*.

The paradigms of feature binding awareness are intramodal. A visible figure may look jointly reddish and square. 'E' has a visible part 'F' lacks. A developed experimental literature deals with visual feature binding and its relation to visual awareness (see, esp., Treisman and Gelade, 1980; Treisman, 1996, 2003). Binding also occurs in other modalities. A piercing alarm sounds high-pitched and loud. Fresh oysters feel cool and clammy to the touch. Fried chicken tastes of salt and oil. After a flood, carpet smells mildewy and pungent.

Skeptics about intermodal feature binding awareness say that awareness of features' belonging to something common results from associations between sensory experiences or from 'post-perceptual processing (or inference).' For instance, Fulkerson (2011, pp. 504–6) thinks distinct unimodal experiences are associated in a higher-order multisensory experience, and Spence and Bayne (2015, §7, p. 119) admit only extra-perceptual apparent unity (cf., Bayne, 2014). Connolly (2014, pp. 354–5, 362) says multimodal episodes can be explained in terms of 'a conjunction of an audio content and visual content' rather than 'fused multimodal units.' Deroy et al. (2014, p. 8) propose to capture the impression of a multisensory object without multisensory awareness, maintaining, 'awareness remains unimodal.'

In opposition to this, my case for a non-skeptical position relies on a contrast between intermodal episodes of (1) perceiving something's being both *F* and *G*, and (2) perceiving something's being *F* and something's being *G*. An episode of (1) requires that a single thing perceptibly has both features, but (2) does not require that. My view is that it can be perceptually apparent that features perceived through different modalities are bound and thus belong to the same thing. So, for example, you might visuo-tactually perceive a brick's being jointly red and rough. And this contrasts with just perceiving something red and something rough, as when you see a stop sign while feeling sandpaper. Or, you might audio-visually perceive an explosion at once to be jointly loud and bright. This contrasts with just perceiving something loud and something bright, as when you hear a trumpet and see a camera flash. The difference between (1) and (2) may be reflected in conscious episodes of multisensory perceptual awareness.

What is the evidence? First of all, recent experimental research on multisensory perception reports that perceptual systems bundle or bind information from different senses to yield unified perceptions of common multimodally accessible objects or events.

> [I]t is reasonable to suppose that the organism should be able to bundle or bind information across sensory modalities and not only just within sensory modalities. (Pourtois et al., 2000, p. 1329)

> There appear to be specific mechanisms in the human perceptual system involved in the binding of spatially and temporally aligned sensory stimuli. (Vatakis and Spence, 2007, p. 754)
>
> . . . a particularly powerful form of binding that produces audio-visual objects. (Kubovy and Schutz, 2010, p. 42)
>
> The binding of AV speech streams seems to be, in fact, so strong that we are less sensitive to AV asynchrony when perceiving speech than when perceiving other stimuli. (Navarra et al., 2012, p. 447)

Typically, cross-modal illusions and recalibrations are cited as evidence. The intersensory discrepancy paradigm is used to generate a cross-modal illusion and thus to establish a multisensory interaction. The fact that sensory responses are recalibrated against each other when two senses target a common source is taken as evidence that perceptual systems discern and treat that sensory information as concerning something common. Treating information as having a common source means that a critical condition for binding is satisfied.

> The bias measured in such experimental situations is a result of the tendency of the perceptual system to perceive *in a way that is consonant with the existence of a single, unitary physical event*. . . . Within certain limits, *the resolution may be complete, so that the observer perceives a single compromise event*. (Welch and Warren, 1980, pp. 661, 664, my emphasis)

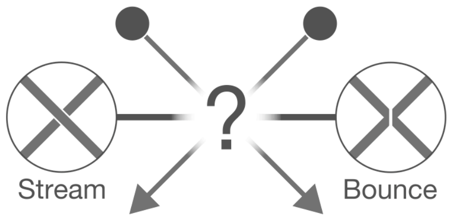
However, I want to emphasize that it is not enough to appeal to cross-modal illusions and recalibrations to establish that intermodal feature binding has taken place. The problem is that there is a gap between perceiving in a way that is consonant with a single event and perceiving something as a single event. The senses can be coordinated and brought into conformity without identifying common targets as such. Conflict resolution does not guarantee either integration or binding.

Nevertheless, standard measures of intramodal feature binding do also provide evidence for intermodal feature binding. For instance, illusory feature conjunctions (especially outside focal attention), object-specific preview effects (benefits and penalties), object and event files (temporary episodic representations of persisting objects and events), and superadditive effects all have been studied and reported in a variety of intermodal contexts (see Figure 2). For example, Cinel et al. (2002, pp. 1244–1245) say, 'These results demonstrate that ICs [Illusory Conjunctions] are possible not only within the visual modality but also between two different modalities: vision and touch,' and conclude, '[I]nformation converges preattentively for binding from different sensory modalities . . . this binding process is modulated by the parietal lobe.' Jordan et al. (2010, p. 501) report 'a standard, robust OSPB [Object Specific Preview Benefit]' between vision and audition and say their data 'explicitly demonstrate object files can operate *across* visual and auditory modalities.' Zmigrod et al. (2009, pp. 682–683) support 'episodic multimodal representations' rather than mere intermodal interactions and conclude that feature binding occurs across modalities. This experimental work reveals that perceptual processes show signs of tracking and representing individual feature bearers as common across sensory modalities and as bearing features perceptible with different senses.

	intramodal	*intermodal*
illusory conjunctions	Treisman and Schmidt (1982)	Cinel et al. (2002)
object-specific preview effects	Kahneman et al. (1992)	Jordan et al. (2010)
object/event files	Kahneman & Treisman (1984)	Zmigrod et al. (2009)
superadditive effects	Baumgartner et al. (2013)	Stein et al. (2010)
**preview effects vs. awareness*	Mitroff et al. (2005)	Zmigrod and Hommel (2011)

[[FIGURE 2, CAPTION: Figure 2. *Experimental measures of binding and awareness*. Intramodal and intermodal examples.]]

However, such empirical work also raises an important objection. It concerns the relationship between experimental measures of binding and conscious perceptual awareness. In the unimodal visual case, Mitroff et al. (2005) find that, under certain conditions, object-specific preview benefits disagree with conscious visual awareness—in an ambiguous visual display, object-specific preview effects may indicate bouncing while subjects report seeing streaming (see Figure 3). Moreover, Zmigrod and Hommel (2011) report similar results in a multisensory audio-visual condition. They say, '[B]inding seems to operate independently of conscious awareness, which again implies that it solves processing problems other than the construction of conscious representations' (p. 592). Therefore, experimental measures of binding alone do not show that there is conscious perceptual awareness of binding (see also Deroy et al., 2014, p. 7).



[[FIGURE 3, CAPTION: Figure 3. *Binding and awareness may diverge.* Under certain conditions, experimental measures of binding, such as object-specific preview benefits, may disagree with conscious awareness (after Mitroff et al., 2005).]]

Consider the perceptual appearances more directly. The contrast between (1) and (2) marks a difference in how things may perceptually appear to be, whether or not you believe they are that way, and whether or not they are that way.

First, apparent binding can be illusory. Consider ventriloquism. You seem to hear a visible puppet speaking, even if you do not infer or believe the puppet talks. This contrasts with unsuccessful ventriloquism, in which it is perceptually evident that what you hear is not the puppet you see. At the movies, nothing in the theater makes the sounds you hear and is visible on screen. No single perceptible event bears those visible and audible features, so the appearance as of a common source is illusory. The illusion need not be spatial or temporal, as the speaker could be located immediately behind the movie screen—this is typical in multisensory psychology experiments. So, what seems like a case of (1) may in fact be a mere case of (2).

Conversely, you can perceive coinstantiated features as unbound, as when the ventriloquist you see makes the sounds that appear to come from the dummy. Or you can just fail to perceive coinstantiated features as bound, as when you fail to perceive the visible toe poking out from under the sheets to be your own felt toe. So, what seems like a mere case of (2) may involve perceiving features that in fact are coinstantiated.

Accordingly, intermodal binding awareness can break down. Imagine a multimedia concert recording in which the timing of the sound and video is misaligned. Maybe it is just a little bit off, as with lip syncing, in a way that is noticeable but not disturbing. If it is worse, the experience is jarring. But, if the timing is way off, the sights and sounds seem wholly dissociated. Compare this to when the sound and video are perfectly in sync. The auditory and the visual stimulation remains qualitatively the same across these scenarios, but the phenomenology differs strikingly. The phenomenological contrast it is not just a difference in the alignment of experiences. It is not a uniform, gradual shift. The categorical perceptual difference stems partly from perceiving something jointly to have audible and visible features when the sound and video coincide.

One objection is that these experiences differ in spatio-temporal respects, so controlling for spatio-temporal differences eliminates any experiential difference. My reply is that intermodal binding does not just depend on spatio-temporal cues. It also depends on other factors, such as whether and how the subject is attending, the subject's expectations, and the plausibility of the combination (Bertelson and de Gelder, 2004). For instance, a female face more easily binds a female voice than does a male face (Vatakis and Spence, 2007). So, fixing spatio-temporal features does not by itself suffice in context to fix whether intermodal binding occurs. Thus, Vatakis and Spence (2007) say:

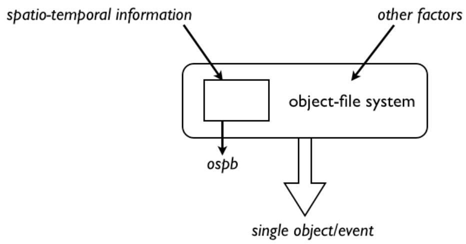
> [T]he perceptual system also appears to exhibit a high degree of selectivity in terms of its ability to separate highly concordant events from events that meet the spatial

> and temporal coincidence criteria, but which do not necessarily 'belong together.' (Vatakis and Spence, 2007, p. 754)

Furthermore, the capacity for specific forms of intermodal binding can be selectively disrupted. Here are three examples. First, individuals with autism have difficulty integrating cues about speech and emotion from vision and audition (see, e.g., de Gelder et al., 1991; Mongillo et al., 2008). Second, recent work reports 'zapping' multisensory integration performance using brain stimulation. For instance, Pasalar et al. (2010) use fMRI-guided transcranial magnetic stimulation to selectively disrupt visuo-tactile integration (see also Kamke et al., 2012; Zmigrod and Zmigrod, 2015). And, third, Hamilton et al. (2006, 2012) describe a patient who cannot integrate auditory and visual information about speech. 'We propose that multisensory binding of audiovisual language cues can be selectively disrupted' (Hamilton et al., 2006, 66).

Intermodal binding awareness is not fixed by perceptually apparent spatio-temporal features. Therefore, the appearance of binding can vary while perceptually apparent spatio-temporal features do not. Controlling for spatio-temporal differences thus need not dissolve the apparent difference in perceptual experience between an episode of (1) and of (2).

This also gives us a way to deal with the objection about the relation between empirical measures and binding awareness. If the system responsible for tracking objects (the so-called 'object-file' system) incorporates mechanisms that are responsive just to low-level spatio-temporal features, and if such mechanisms are selectively probed during creative experimental interventions, then the appearance of binding may disagree with the verdicts of some of these low-level components of the overall system that is responsible for apparent intermodal binding (Figure 4 is a schematic diagram).



[[FIGURE 4, CAPTION: Figure 4. *Binding and awareness*. Schematic of an 'object file' system that accommodates disagreement between object-specific preview benefits and binding awareness.]]

I conclude that intermodal binding awareness is a third grade of multisensory awareness beyond the second. This is significant because it means that perceptual awareness is not just minimally multisensory. Some ways to be perceptually aware of an individual thing require identifying it across modalities and so cannot be analyzed just in terms of ways you could be perceptually aware using specific sensory modalities all on their own.

For instance, visuotactually perceiving a thing's being jointly round and rough is not just co-consciously seeing a thing's being round while feeling a thing's being rough, where it just happens to be the same thing seen and felt. Being perceptually sensitive to or perceptually appreciating the identity of what is seen and felt is not something that can occur unimodally. And it is not a way of perceiving that boils down to jointly occurring episodes of seeing and feeling that could have occurred independently from each other. And it does not accrue thanks to mere co-consciousness. It is a distinct perceptual act or achievement. It is not factorable without remainder into co-conscious, modality-specific components that could have occurred independently from each other. Therefore, overall perceptual awareness is more than minimally multisensory. The tempting thought fails. Perceiving is not just co-consciously seeing, hearing, feeling, tasting, and smelling at the same time.

Not even all phenomenal character is modality specific. Given that there can be contrasting episodes of (1) and (2), visuotactually perceiving a thing's being jointly red and rough can have phenomenal features that no corresponding wholly visual or wholly tactual perceptual experience (of redness or of roughness) could have under equivalent stimulation, and that do not accrue thanks to mere co-consciousness (O'Callaghan, 2014, §5).

At this point, someone might respond: Binding awareness is an aspect of the structure of perceptual awareness. (Maybe it is due to synchronous processing, dimensional coding and distinct hyperplanes, or mere attention.) It does not involve a novel perceptible feature of the world that is accessible only through multisensory awareness. And it need not involve any novel qualitative features of conscious perceptual experience (e.g., qualia). Instead, it is just a structural characteristic of the perceptual experience itself. If so, perceptual awareness may be exhausted by that which is associated with each of the respective modalities, along with whatever accrues thanks to its intermodal binding structure and mere co-consciousness.

4. Grade 4: Multisensory Awareness of Novel Feature Instances

Spence and Bayne (2015, p. 121) say admitting multisensory awareness in cases of binding is compatible with 'severe limitations on the degree to which consciousness can straddle distinct sensory modalities.' Richer forms of multisensory awareness ground the case for gentler restrictions. For instance, some features have instances that could only be perceived using more than one sense—such feature instances are accessible only multisensorily (see O'Callaghan, forthcoming, §4). Perceptual awareness of any such feature instance need not be exhausted by what is associated with each of the respective modalities along with that which accrues thanks to mere co-consciousness. What is novel is not just a new way of experiencing the same old features, and it is not just a matter of intermodal binding. It is not just tracking something across modalities. Instead, through the coordinated use of multiple senses, one becomes perceptually responsive to a novel, previously unperceived feature instance. This is not simply a matter of co-consciously seeing, hearing, touching, smelling, and tasting—plus binding. It is a **4th grade** of multisensory awareness.

Let me describe some examples. There are relational feature instances that could only be perceived through multisensory episodes. One important type of case involves temporal relations that hold between things experienced with different senses. Most subjects can quickly and accurately judge temporal order between modalities (see, e.g., Spence et al., 2003). Given their speed and accuracy, cross-modal temporal order judgments may reflect perceptual judgments driven by perceptually apparent intermodal temporal relations. This has practical applications. Umpires in baseball tell whether a baserunner is safe or out by watching his foot strike the bag and listening for the sound of the ball hitting the fielder's mitt. In close calls, vision alone is unreliable due to the distance between the base and the mitt. The umpire does not simply perceive each one and then work out the relation. He multisensorily perceives the temporal relation, order, or interval between the visible and audible events.[1]

[1] Given that umpires already are looking at the base, multisensory prior entry may impact temporal order judgments in a way that makes granting apparent ties to the runner suspect. See, e.g., Spence et al. (2001), whose Experiment 1 nonetheless provides support for accurate multisensory temporal order judgments (roughly 90 percent) under divided attention for stimulus

Why think that these cases involve perceived intermodal relations rather than co-conscious but modality-specific spatial and temporal location experiences? A rich experimental literature has addressed apparent intermodal temporal relations. For instance, there is extensive work on intermodal synchrony perception. Müller et al. (2008, p. 309) say, 'A great amount of recent research on multisensory integration deals with the experience of perceiving synchrony of events between different sensory modalities although the signals frequently arrive at different times.' This is a sophisticated achievement—Keetels and Vroomen (2012, p. 170) describe it as 'flexible and adaptive.' It requires accommodating timing differences introduced by the external world and by the body. For instance, the sound waves from clapping hands reach your ears well after the light reaches your eyes. When I touch your toe, the tactual signal takes longer to reach your brain than the visual signal.

> To perceive the auditory and visual aspects of a physical event as occurring simultaneously, the brain must adjust for differences between the two modalities in both physical transmission time and sensory processing time. . . . Our findings suggest that the brain attempts to adjust subjective simultaneity across different modalities by detecting and reducing the time lags between inputs that likely arise from the same physical events. (Fujisaki et al., 2004, p. 773)

Stone et al. (2001) define the audio-visual *Point of Subjective Simultaneity* as the timing at which a subject is most likely to indicate that a light and tone begin simultaneously. They found that this point varies across subjects but is stable for a given observer. Typically, it required the light to precede the tone, by an average (across subjects) of about 50 milliseconds (see also,

---

onset asynchronies above 100 milliseconds. Thanks to an anonymous referee for drawing my attention to this literature.

e.g., Zampini et al., 2005; Arrighi et al., 2006). Spence and Squire (2003) suggest that a 'moveable window' for multisensory integration and a 'temporal ventriloquism' effect contribute to perceptually apparent synchrony. The experimental results provide evidence that perceptual systems are sensitive to the relative timing of events across the senses.

A skeptic will object that subpersonal coordination disclosed by experimental work revealing sensitivity to temporal relations just yields ordered or synchronous experiences rather than perceptual experiences as of order or synchrony. At this point, the debate about awareness threatens to reach a stalemate. To reply, we need better evidence that a distinctively multisensory response drives perception of a novel feature instance.

Intermodal meter perception currently offers the best reply. Meter is the structure of a pattern of rhythmic musical sounds—its repeating framework of timed stressed and unstressed beats. Meter can be shared by patterns of sounds whose rhythm differs. A piece's time signature indicates its meter. Meter is perceptible auditorily and tactually. Huang et al. (2012) demonstrate that it is also possible to audio-tactually discriminate a novel musical meter that is present neither audibly nor tactually. 'We next show in the bimodal experiments that auditory and tactile cues are integrated to produce coherent meter percepts.' They assert, 'We believe that these results are the first demonstration of cross-modal sensory grouping between any two senses' (Huang et al., 2012, p. 1). To illustrate this type of phenomenon, consider a simple case of intermodal meter perception using an audio-tactual rhythm pattern. Suppose you hear a sequence of sounds by itself. Next, suppose you feel a different sequence of silent vibrating pulses on your hand. Now combine the two. You hear a sequence of sounds while feeling a differing sequence of pulses on your hand. You can attend to the sounds or to the vibrations. But it is also possible to discern and attend to the metrical pattern formed by the audible sounds and the tactual pulses—

the audio-tactual duples or triples. Perceiving the intermodal meter differs, and it differs phenomenologically, from perceiving either of the unimodal patterns in isolation. It also differs from experiencing two simultaneous but distinct patterns. The intermodal meter pops out.

An intermodal meter is a novel feature instance of which you may be perceptually aware. Perceiving an intermodal meter is not just co-consciously perceiving distinct unimodal meters. Perceptual awareness of an intermodal meter requires the coordinated (and not merely contemporaneous) use of multiple senses. It extends one's perceptual capacities.

Intermodal meter perception suffices to demonstrate the fourth grade. Other cases suggest fertile ground for future research. By analogy with temporal relations, consider simple spatial relations. Cross-modal interactions and recalibrations demonstrate that information about space from different senses is coordinated across modalities. Matthen (2015a) defends the Kantian thesis that space is pre-modal on the grounds that modality-specific spatial maps require such coordination. Thus, it may be possible to perceive spatial relations that hold between things experienced with different senses. For instance, you might attend to the spatial offset between an audible sound coming from just to the left of a visible speaker. Or you might perceive a visible feature and a tangible feature to be co-located—to be located in the same place. You might experience a sound paired with a light oriented vertically to grab your attention when presented following a sequence of sound and light pairs oriented horizontally. You see a located feature, you hear a located feature, and you multisensorily perceive the novel intermodal spatial relation that holds between those features.

Moreover, intermodal motion may be perceptible. You could hear a sound and then see a spot moving from left to right and intermodally perceive its motion to be continuous. Because this might seem to involve just a sum of unimodal movements, more persuasive evidence requires a novel pattern of motion that differs from both the audible and visible movements. For instance, imagine a perceptible intermodal zig-zag comprising orthogonal diagonal unimodal motion patterns, or perceptible clockwise circular motion comprising linear audible and visible movements (see Figure 5).

A skeptic will want evidence that such novel intermodal motion is perceptible rather inferred. As in the unimodal case, merely apparent or illusory intermodal motion is a good test. Some researchers have reported intermodal apparent motion. Harrar et al. (2008) claim that there is visuo-tactile apparent motion between lights and touches. Others agree:

> Apparent motion can occur within a particular modality or between modalities, in which a visual or tactile stimulus at one location is perceived as moving towards the location of the subsequent tactile or visual stimulus. . . . For example, with an appropriate time interval between a visual stimulus at one location and a tactile stimulus at another location, the participants would perceive some kind of motion stream from the first to the second location. In this kind of intermodal apparent motion, the motion stream is composed of stimuli from two different modalities. (Chen and Zhou, 2011, pp. 369, 371).

Chen and Zhou (2011) and Jiang and Chen (2013) report that auditory and visuo-tactile apparent motion influence each other.

The reports of Allen and Kolers (1981, cited by Spence and Bayne, p. 112) are intriguing. They find no evidence of apparent motion for an integrated, traveling, hybrid audio-visual object (p. 1320). However, in a heteromodal condition involving a light and a sound in different locations, the authors do find evidence of apparent intermodal motion.

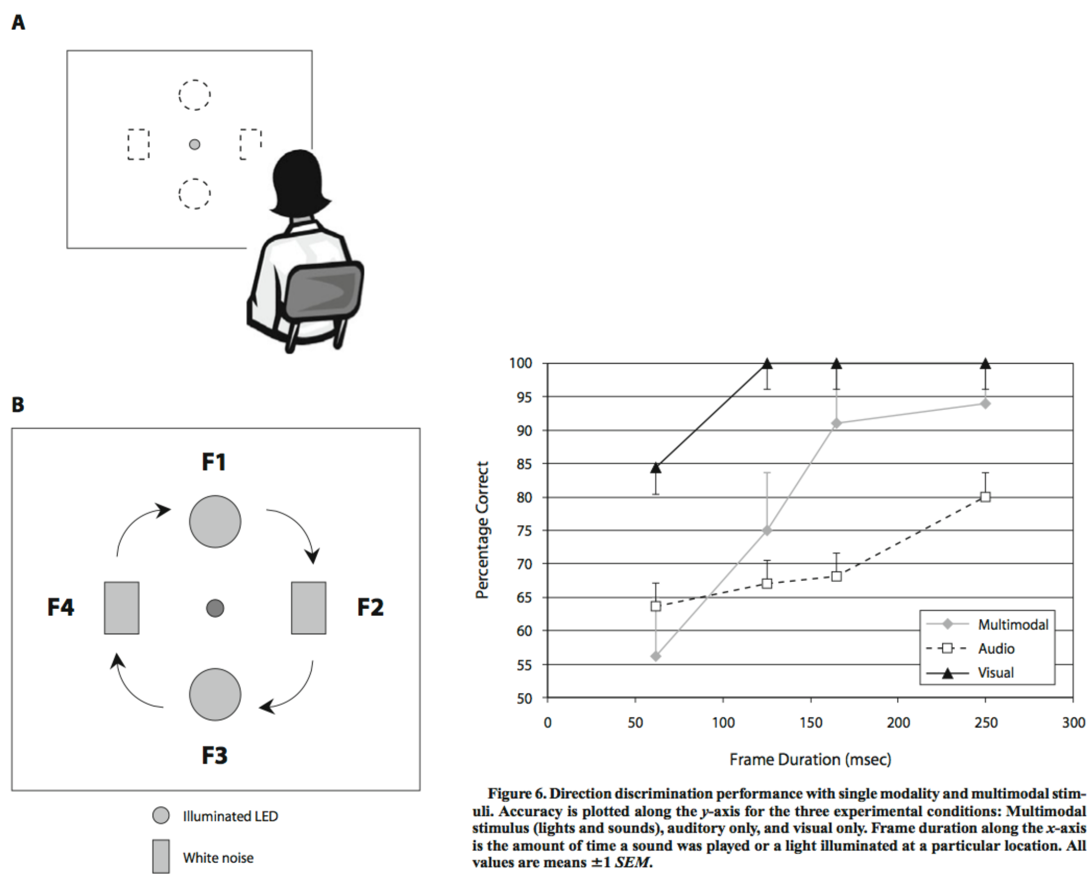
> One of the authors (Allen) once perceived what could be regarded as a sonorous light or a luminous sound in motion between a visual and an auditory stimulus. The following is an account written at the time of the occurrence:
>
>> A light breaks away from the location of the visual stimulus at the latter's onset—its trajectory can be followed for perhaps .5 meters, but a sense of its continuing to the ear is strong. The light seems to arrive there at the onset of the tone and then returns to the location of the visual stimulus, arriving there at the offset of the tone. One could ascribe a 'sonorous' quality to the light, especially on its return to the location of the visual stimulus during the onset of the tone. The phenomenon repeated perhaps 25–30 times. (Allen and Kolers, 1981, p. 1320)

Nonetheless, others have failed to find intermodal apparent motion, leading Spence and Bayne (2015, p. 112) to skepticism. For instance, Huddleston et al. (2008) test for a case of audio-visual apparent motion and find, 'Although subjects were able to track a trajectory using cues from both modalities, no one spontaneously perceived "multimodal [apparent] motion" across both visual and auditory cues' (p. 1207). The authors elaborate:

> The results of Experiment 3 provide initial evidence that, although subjects could use information from both modalities to determine the trajectory of the stimulus, the stimulus used in this experiment was not sufficient to overcome the need for spatial and temporal congruence to integrate multimodal cues for the perception of motion across modalities and, therefore, did not lead to the perception of a unified 'audiovisual' stimulus. (Huddleston et al., 2008, p. 1215)

However, the results of Huddleston et al. are inconclusive, and I want to suggest an alternative interpretation. Their studies show that subjects are able to discern audio-visual motion with good accuracy even if they do not report spontaneous perceptually apparent audio-visual motion. The authors say that subjects failed to perceive audio-visual motion because the experimental stimuli lacked the sort of spatial and temporal congruence that is needed to integrate cues across modalities, which is a requirement on intermodal motion perception. I think this is not the whole story. Perceiving motion requires identifying something as moving. Huddleston et al. use LED lights and white noise bursts at different locations over time in their multisensory condition. Lights and noises separated by space may not provide strong enough cues that a single item has traveled from one place to another. The bar for intermodal motion perception may be higher than for visual apparent motion, which tolerates robust qualitative difference across space.

This interpretation fits the evidence. In the unimodal visual condition, Huddleston et al.'s participants achieved 90 percent accuracy reporting direction of motion when each LED was presented for at least 100 milliseconds. In the unimodal auditory condition, white noise bursts were presented in the vertical frontal plane, and performance peaked at 80 percent accuracy (p. 1214, Figure 6; my Figure 5). In the multisensory condition, participants were 90 percent accurate reporting the direction of intermodal motion when each stimulus was presented for at least 175 milliseconds (Figure 5). This was better than the audition-only condition. However, auditory spatial localization is far worse than vision in the vertical frontal plane. (Notably, Allen and Kolers, 1981, p. 1319, found loudspeakers insufficient for robust apparent auditory motion, so used headphones instead.) Spatial audition (directional hearing), and thus motion determination, improves greatly in the horizontal plane centered around the subject. Thus, it is most noteworthy for my argument that accuracy in the multisensory condition matched performance in a separate unimodal auditory condition conducted in the horizontal plane using two different types of sounds: a white noise burst and a ‘distinctive’ complex sound (see my Figure 6, Unambiguous task). So, weak or absent identity cues in the multisensory condition may have affected not just performance but also awareness of apparent intermodal motion. Stronger source identity cues thus could reveal awareness as of apparent intermodal motion.



[[FIGURE 5a and 5b, CAPTION: Figure 5. *Intermodal apparent motion.* Audio-visual apparent motion using LED and white noise in the frontal plane, contrasted with visual and auditory apparent motion. (Huddleston et al., 2008, pp. 1213–1214, with permission)]]

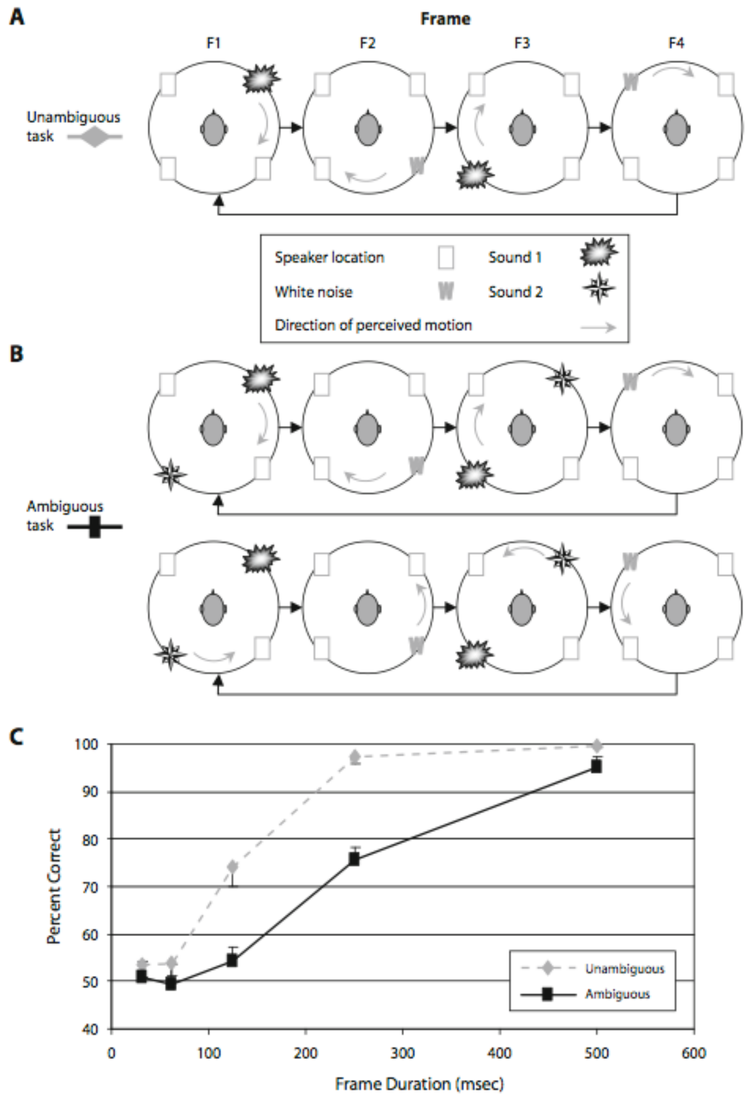


Figure 4. Experiment 2: Cued attention-based motion perception. Subjects visually fixated a marker located between speakers while sitting in the middle of a four-speaker horizontal array. (A) Unambiguous stimulus. Alternating white noise or distinctive sounds were played on a single speaker during each of four frames sequenced in a clockwise or counterclockwise direction (clockwise illustrated in figure). Subjects indicated perceived direction of motion on each trial. (B) Ambiguous stimulus. The stimulus is very similar to the unambiguous stimulus, but on frames when the first distinctive sound was played, a second (different) distinctive sound was played simultaneously at the opposite speaker. If the subject was cued to attend to the first sound, the perceived motion should rotate clockwise (direction of motion indicated by arrows between speaker locations). If the subject attended to the second distinctive sound in the same stimulus sequence, the perception should rotate in a counterclockwise direction. (C) Psychophysical performance for the unambiguous and ambiguous (requiring attention) tasks. Frame duration is the amount of time a sound was played at a particular location (interstimulus interval = 0). All values are means ±1 *SEM*.

[[FIGURE 6, CAPTION: Figure 6. *Auditory apparent motion*. Performance using one versus two sound types in the horizontal plane. (Huddleston et al., 2008, p. 1211, with permission)]]

Other types of features also may have instances that are perceptible multisensorily. Especially noteworthy are structural features. For instance, there is good empirical evidence that intermodal causal relations are perceptible (Sekuler et al., 1997; Guski and Troje, 2003; Choi and Scholl, 2006; Shams and Beierholm, 2010). And philosophers, including Nudds (2001) and Siegel (2009), have described cases in which typical humans perceptually experience causal relations intermodally.

Each of these arguments leaves room for a skeptic to resist. Suppose the experiments show that you detect such relational features. Nevertheless, it is possible that you are only ever consciously aware of the locations in space in time of the objects and events you perceive through different senses, while you fail to consciously multisensorily perceive the spatial and temporal relations among them.

Here is my reply. According to a moderately liberal general account, humans sometimes are perceptually aware of spatial, temporal, and causal relations in addition to places and times. The objection grants that evidence from psychophysics and neuroscience can show that perceptual systems detect intermodal relational features. It denies that this establishes multisensory awareness of such features. However, the philosophical arguments in the intermodal case are just as compelling as in the intramodal case. For example, you may perceptually experience the visible striking of a bell to produce or to generate its audible ringing, and this contrasts with just seeing a striking then hearing a ringing. The capacities dissociate, and the associated perceptual processes, patterns of action, perceptual beliefs, and phenomenology all differ. Multisensory awareness of intermodal causality explains the contrast in a way that resists the confounds. Moreover, given the range and flexibility of factors that influence multisensory processing, it is even more plausible that the impression of intermodal causality sometimes breaks down in typical perceivers. This is analogous to the selective breakdown of intermodal binding awareness. If so, contrast arguments are even more effective in the intermodal case than in the intramodal case. Therefore, denying that you are ever perceptually aware of any such feature intermodally relies on reasons that in turn can be used to deny that you are ever even unimodally perceptually aware of any such relational feature—spatial, temporal, or causal. However, this yields an implausibly austere account of human perceptual awareness, and it introduces an unexplained rift between perceptual capacities and perceptual judgments. According to a moderately liberal account of human perceptual awareness, there is no good reason to deny that some such relational feature instance is consciously perceptible, and just being intermodal introduces no special trouble.

The examples in this section involve multisensory perceptual awareness of relational feature instances that hold between things you perceive with different senses. Some are controversial. However, one such demonstration, as intermodal meter perception provides, suffices to make my case. It is a counterexample to the claim that all perceptual awareness on each occasion is modality specific and thus to the tempting thought with which we began. Each such case involves an episode of multisensory awareness that is not exhausted by what is associated with each of the respective modalities along with whatever accrues thanks to simple co-consciousness. Moreover, perceptual awareness as of such an intermodal relation is not merely an aspect of the structure of perceptual experience itself. Instead, it involves seeming to be acquainted with a feature of the world that is accessible only multisensorily. This demands an intentional or relational characterization. Thus, it is a fourth grade of multisensory awareness beyond the third.

There remains a limitation. Each of these perceptible features belongs to a type with instances that are perceptible unimodally. You can perceive spatial, temporal, and causal relations through vision, touch, or hearing alone. Since these feature types are familiar from unisensory contexts, perceptual awareness of their intermodal instances need not be multisensory in a deeper respect. Multisensory perceptual experiences of such features might only have phenomenal features of types that unisensory perceptual experiences can instantiate. Such phenomenal features themselves thus might belong to unimodal or to amodal types rather than to types whose members are constitutively or necessarily multisensory.

This limitation is theoretically significant. The arguments above demonstrate that we cannot exhaustively capture an episode of multisensory awareness just by mentioning features instantiated by corresponding unisensory experiences—not every multisensory episode is just the co-conscious sum of its modality-specific parts. However, they do not show that it is not possible to account for multisensory perceptual awareness, even of novel feature instances, just in terms of (unimodal or amodal) features that unimodal perceptual experiences could have. And so, we might still say that the qualitative components of phenomenological character are not in this respect deeply multisensory.

5. Grade 5: Multisensory Awareness of Novel Feature Types

So let me introduce a **5th grade** of multisensory awareness beyond the fourth. Suppose there were novel features belonging to types whose instances could only be perceived multisensorily. The capacity to access any such feature would require multiple sensory modalities. You could not be fully aware of an instance of such a type through any single sense. In this respect, such features are unlike spatial, temporal, and causal features, which only have some novel intermodal instances.

Flavor, whose perception involves taste, smell, and trigeminal somatosensation, sometimes is mentioned as a candidate for such a novel feature type (see Smith, 2015). The distinctive and recognizably minty flavor of fresh mint ice cream is perceptible only thanks to the joint operation of several sensory systems. It requires taste, olfaction, and trigeminal stimulation, but it is not fully perceptible through taste, olfaction, or somatosensation independently. Thus, flavor experiences, such as experiencing the minty quality of mint, may have entirely novel characteristics, including phenomenal features, that no unimodal experience could have and that do not accrue thanks to simple co-consciousness.

This is a rich case, but I will just mention the crux. First, flavor perception does not involve a novel sensory modality. It has no dedicated sense organ. And flavors really do involve smells, tastes, and tingles. Part of flavor is being salty or creamy or burning. This implicates taste, smell, and touch. Second, if apparent flavor is just an agglomeration—an otherwise unstructured mixture or bound collection of gustatory, olfactory, and tactual qualities—then flavors pose no special trouble. Since all of their components are perceptible unimodally, awareness of flavors may stem from simple intermodal feature binding awareness. Flavor awareness need not involve any wholly novel phenomenal feature types. Third, however, apparent flavor could involve (1) a novel sort of organization or structure among sense-specific components. Or it could be (2) an organic unity among them. Or it could include (3) a further qualitative component beyond the modality-specific features. It could involve all three. In my view, we should not rule out any of these. The case of mintiness is particularly telling. There is a distinctive, recognizable, and novel quality of mint (i.e., mintiness) that is consciously perceptible only thanks to the joint work of several sensory systems. Surely this is one aspect of the flavor of mint. There are other aspects, like being tingly and cool, that characterize the full, unified flavor of mint. If so, flavors are emergent features—even distinctive qualitative characteristics—of a type that cannot fully be perceived unimodally or thanks to simple co-consciousness and intermodal binding. Experiencing flavors such as mintiness involves phenomenal features that are not instantiated by any unimodal experience and that do not accrue thanks to mere co-consciousness or intermodal binding. Flavor awareness thus is deeply, irreducibly multisensory.

Future work should explore additional forms of deeply multisensory awareness. Speech perception and balance are promising examples.

6. Grade 6: Novel Awareness in a Modality

I'll close by speculating about a sixth and quite different variety of multisensory awareness. The discussion so far establishes that perceptual awareness is not exhausted on each occasion by what is associated with each of the respective modalities. This holds even as a claim about phenomenal character. There also may be forms of perceptual awareness that *are* associated with a given sense modality, but which would not have been possible without current or past perception through another.

Say that a feature of a perceptual episode is associated with a given modality on an occasion only if it could be instantiated by an episode that is wholly or entirely of that modality (not any other) under equivalent stimulation (see O'Callaghan, forthcoming). So, for instance, the phenomenal character of your current multisensory experience that is associated with audition on this occasion is that which could be instantiated by a wholly auditory experience under just the same stimulation; the representational content of your current multisensory experience that is associated with vision includes only that of a wholly visual experience under equivalent stimulation.

The arguments above show that the features of an episode of multisensory awareness need not be exhausted by those that are associated with each of the respective modalities. However, there could be a difference in features between the auditory awareness of a creature who only ever had audition and the auditory awareness of a creature who has a rich background of experience with the other senses. Under equivalent stimulation, there could be a difference between a presently and historically purely auditory experience, and an experience that is currently merely, wholly, or exclusively auditory (it is not also visual, tactual, and so on) but which in the past has been multisensory. This means that there could be auditory experiences that are cross-modally dependent upon other senses.

To illustrate, here are four types of examples. (1) Cross-modal parasitism occurs whenever features are perceptible with one modality but only thanks to their being perceptible through another. For instance, Strawson (1959) says a purely auditory experience would be non-spatial; however, he also maintains that you can hear spatial features, but only thanks to your having inherently spatial visual or tactual experiences. A Berkeleyan might say that visual awareness of space is parasitic on tactual awareness of space. Or consider seeing the solidity of a statue. Synesthesia may involve an atypical, systematically illusory variety of cross-modal parasitism (see, e.g., Auvray and Deroy, 2015). This paper's focus is on typical, adaptive perceptual capacities. (2) Cross-modal completion may involve an intermodal form of so-called amodal completion. In visual amodal completion, you see an object to complete behind an occluder without seeing its hidden parts, and this affects your perception of its visible features, such as its shape. Analogously, you might auditorily perceive an event to be a thing with visible but unseen aspects, and this may affect your perception of its audible features. (3) Cross-modal perceptual learning also could yield awareness associated with one modality that is cross-modally dependent on another. (4) So could cognitive penetration with cross-modal etiology. In these examples, awareness associated with a given sense on an occasion differs from what it otherwise could have been if not for a background of awareness involving other senses.

If there is any such cross-modally dependent variety of perceptual awareness, then its instances are candidates for conscious perceptual episodes belonging to a single modality which in an important respect are multisensory. But they are multisensory in a way that differs from any of the previous grades. They involve a novel variety of perceptual awareness within a modality that is made possible only thanks to prior or concurrent awareness involving another sense. This is a **6th grade** of multisensory awareness.

Its consequences for theorizing also differ. This grade implies that it is not even possible exhaustively to characterize perceptual awareness that is associated with a given modality in terms that are wholly proprietary to that sense. In typical, adult human subjects, capturing visual awareness itself requires appealing to extra-visual forms of perception. This threatens to undercut the very project of theorizing about perceiving with one sense in isolation or abstraction from the others.

7. Conclusions

I have distinguished six varieties of multisensory awareness.

The 1st is minimally multisensory awareness. It implies that conscious perceptual awareness at any moment may be (and sometimes is) associated with more than one sense modality. So, perceptual consciousness is not always unisensory.

The 2nd involves coordinated awareness across sensory modalities, as revealed by cross-modal recalibrations and illusions. It implies that the senses do not function wholly independently from each other. Sensory awareness associated with one modality may reflect joint perceptual concerns shared with another modality.

The 3rd is intermodal binding awareness. With features perceived thanks to different senses, one may consciously perceive those features' jointly belonging to a common individual. This implies that it is possible to perceptually identify a common item as such across sense modalities. In such cases, one's perceptual awareness is not exhausted by what is associated with each of the modalities along with what accrues thanks to mere co-consciousness. According to my criterion, the features of a perceptual episode that are associated with a given modality on an occasion include only those instantiated by a corresponding unimodal episode under equivalent stimulation. Thus, a multisensory perceptual episode of intermodal binding awareness instantiates further features beyond those associated with each of the respective modalities on that occasion.

The 4th is multisensory awareness of novel feature instances, such as spatial, temporal, or causal relations that are not perceptible unimodally. This implies that there are episodes of multisensory awareness whose features are not exhausted by those associated with each of the respective modalities on that occasion along with those that accrue thanks to simple co-consciousness plus those that are merely aspects of the structure of perceptual awareness itself. Instead, the senses are used in coordination to enable perceptual awareness of a novel feature instance in the world and thus to extend one's perceptual capacities.

The 5th is multisensory awareness of novel feature types, such as flavors, that are inaccessible unimodally. It implies that there are cases of multisensory awareness whose features are not exhausted by those that may be instantiated by some unimodal perceptual episode or another along with those that accrue thanks to mere co-consciousness. That is, perceptual consciousness may involve novel types of features, including qualitative characteristics, that emerge only in multisensory awareness.

The 6th is novel awareness in a modality that depends historically or presently on another sense. These are cases of perceptual awareness associated with a given sense that would not have been possible without another sense. This implies that even perceptual awareness that is associated with a given sense modality on an occasion may have features that a (historically and presently) purely unimodal experience would lack. This grade surely fragments into differing forms of cross-modally dependent awareness. Further work is needed to distinguish them.

Grade 1 simply establishes that perceptual awareness sometimes is multisensory. It leaves open that the senses operate independently and that each conscious perceptual episode is a mere co-conscious sum of modality-specific experiences.

Grade 2 establishes that the senses interact in a principled way to yield coordinated awareness across the senses. It leaves open that all perceptual awareness is modality specific.

Grades 3 through 5 demonstrate that perceptual awareness is not a simple co-conscious sum of visual, auditory, tactual, olfactory, and gustatory episodes. Each grade introduces a capacity that is increasingly difficult to accommodate within a unisensory framework. Binding awareness might be merely structural, but awareness of novel feature instances is not. Multisensory awareness of novel feature instances might involve only unimodal or amodal characteristics, but awareness of novel feature types must involve novel, emergent multisensory characteristics. My discussion of each of these grades aims to demonstrate that multisensory perceptual consciousness may have increasingly rich characteristics beyond those associated with the respective modalities.

Grade 6 demonstrates something else. Cross-modally dependent perceptual awareness implies that not even what is associated with a given modality on an occasion can be exhaustively characterized in terms of perceptual capacities involving that sense alone. The capacities of one sensory modality may depend upon those of another. Forms of awareness associated with one modality on an occasion may depend upon forms of awareness associated with another. For instance, explicating visual content and character may require appealing to touch or proprioception. Thus, not even vision itself can be captured in wholly visual terms. No sense is an island.

The important consequence of these six forms of multisensory awareness is that not all perceptual awareness is modality specific. Some multisensory episodes require the kind of coordination that enables you to perceive novel features or to identify individuals across modalities. So, they do not just involve co-conscious episodes of seeing, hearing, touching, tasting, and smelling that could have occurred independently from each other.

A related consequence is that not even all phenomenal character is modality specific. The phenomenal character of a multisensory perceptual episode need not be exhausted by that which is associated with each of the modalities plus whatever accrues thanks to simple co-consciousness.

The significant upshot is that the assumption of explanatory independence fails even at the level of perceptual awareness. So, we should abandon the sense-by-sense approach. No fully adequate account of perceptual awareness or its phenomenal character, within or across modalities, can be formulated in modality-specific terms. Perceiving is more than just co-consciously seeing, hearing, feeling, tasting, and smelling.

Department of Philosophy

Washington University in St. Louis

References

Allen, P. and Kolers, P. 1981: Sensory specificity of apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 1318–26.

Arrighi, R., Alais, D., and Burr, D. 2006: Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision*, 6, 260–8.

Auvray, M. and Deroy, O. 2015: How do synaesthetes experience the world? In M. Matthen (ed.), *Oxford Handbook of Philosophy of Perception*, pages 640–658. Oxford: Oxford University Press.

Baumgartner, F., Hanke, M., Geringswald, F., Zinke, W., Speck, O., and Pollmann, S. 2013: Evidence for feature binding in the superior parietal lobule. *NeuroImage*, 68, 173–80.

Bayne, T. 2014: The multisensory nature of perceptual consciousness. In D. Bennett and C. Hill (eds.), *Sensory Integration and the Unity of Consciousness*. Cambridge, MA: MIT Press, 15–36.

Bertelson, P. and de Gelder, B. 2004: The psychology of multimodal perception. In C. Spence and J. Driver (eds.), *Crossmodal Space and Crossmodal Attention*. Oxford: Oxford University Press, 141–77.

Chen, L. and Zhou, X. 2011: Capture of intermodal visual/tactile apparent motion by moving and static sounds. *Seeing and Perceiving*, 24, 369–89.

Choi, H. and Scholl, B. 2006: Measuring causal perception: connections to representational momentum? *Acta Psychologica*, 123, 91–111.

Cinel, C., Humphreys, G., and Poli, R. 2002: Cross-modal illusory conjunctions between vision and touch. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1243–66.

Connolly, K. 2014: Making sense of multiple senses. In R. Brown (ed.), *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*. Dordrecht: Springer, 351–64.

de Gelder, B., Vroomen, J., and van der Heide, L. 1991: Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology*, 3, 69–86.

Deroy, O., Chen, Y., and Spence, C. 2014: Multisensory constraints on awareness. *Philosophical Transactions of the Royal Society B*, 369, 20130207.

Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. 2004: Recalibration of audiovisual simultaneity. *Nature Neuroscience*, 7, 773–8.

Fulkerson, M. 2011: The unity of haptic touch. *Philosophical Psychology*, 24, 493–516.

Guski, R. and Troje, N. 2003: Audiovisual phenomenal causality. *Perception and Psychophysics*, 65, 789–800.

Hamilton, R., Shenton, J., and Coslett, H. 2006: An acquired deficit of audiovisual speech processing. *Brain and Language*, 98, 66–73.

Harrar, V., Winter, R., and Harris, L. 2008: Visuotactile apparent motion. *Perception and Psychophysics*, 70, 807–17.

Huang, J., Gamble, D., Sarnlertsophon, K., Wang, X., and Hsiao, S. 2012: Feeling music: integration of auditory and tactile inputs in musical meter. *PLoS ONE*, 7, e48496.

Huddleston, W., Lewis, J., Phinney, R., and DeYoe, E. 2008: Auditory and visual attention-based apparent motion share functional parallels. *Perception and Psychophysics*, 70, 1207–16.

Jiang, Y. and Chen, L. 2013: Mutual influences of intermodal visual/tactile apparent motion and auditory motion with uncrossed and crossed arms. *Multisensory Research*, 26, 19–51.

Jordan, K., Clark, K., and Mitroff, S. 2010: See an object, hear an object file: object correspondence transcends sensory modality. *Visual Cognition*, 18, 492–503.

Kahneman, D., Treisman, A., and Gibbs, B. 1992: The reviewing of object files: object-specific integration of information. *Cognitive Psychology*, 24, 175–219.

Kamke, M., Vieth, H., Cottrell, D., and Mattingley, J. 2012: Parietal disruption alters audiovisual binding in the sound-induced flash illusion. *NeuroImage*, 62, 1334–41.

Keetels, M. and Vroomen, J. 2012: Perception of synchrony between the senses. In M. Murray and M. Wallace (eds.), *The Neural Bases of Multisensory Processes*, Frontiers in Neuroscience. Boca Raton, FL: CRC Press, 147–77.

Kubovy, M. and Schutz, M. 2010: Audio-visual objects. *Review of Philosophy and Psychology*, 1, 41–61.

Macpherson, F. 2011: Cross-modal experiences. *Proceedings of the Aristotelian Society*, 111, 429–68.

Matthen, M. 2015a: Active perception and the representation of space. In D. Stokes, M. Matthen, and S. Biggs (eds.), *Perception and Its Modalities*. Oxford: Oxford University Press, 44–72.

Matthen, M. (ed.) 2015b: *Oxford Handbook of Philosophy of Perception*. Oxford: Oxford University Press.

Mitroff, S., Scholl, B., and Wynn, K. 2005: The relationship between object files and conscious perception. *Cognition*, 96, 67–92.

Mongillo, E., Irwin, J., Whalen, D., Klaiman, C., Carter, A., and Schultz, R. 2008: Audiovisual processing in children with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38, 1349–58.

Müller, K., Aschersleben, G., Schmitz, F., Schnitzler, A., Freund, H., and Prinz, W. 2008: Inter- versus intramodal integration in sensorimotor synchronization: a combined behavioral and magnetoencephalographic study. *Experimental Brain Research*, 185, 309–18.

Navarra, J., Yeung, H., Werker, J., and Soto-Faraco, S. 2012: Multisensory interactions in speech perception. In B. Stein (ed.), *The New Handbook of Multisensory Processing*. Cambridge, MA: MIT Press, 435–52.

Nudds, M. 2001: Experiencing the production of sounds. *European Journal of Philosophy*, 9, 210–29.

O'Callaghan, C. 2012: Perception and multimodality. In E. Margolis, R. Samuels, and S. Stich (eds.), *Oxford Handbook of Philosophy of Cognitive Science*. Oxford: Oxford University Press, 92–117.

O'Callaghan, C. 2014: Intermodal binding awareness. In D. Bennett and C. Hill (eds.), *Sensory Integration and the Unity of Consciousness*. Cambridge, MA: MIT Press, 73–103.

O'Callaghan, C. forthcoming: The multisensory character of perception. *The Journal of Philosophy*.

Pasalar, S., Ro, T., and Beauchamp, M. 2010: TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience*, 31, 1783–90.

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., and Crommelinck, M. 2000: The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11, 1329–33.

Pylyshyn, Z. 1999: Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, 22, 341–423.

Sekuler, R., Sekuler, A., and Lau, R. 1997: Sound alters visual motion perception. *Nature*, 385, 308.

Shams, L. and Beierholm, U. 2010: Causal inference in perception. *Trends in Cognitive Sciences*, 14, 425–32.

Shams, L. and Kim, R. 2010: Crossmodal influences on visual perception. *Physics of Life Reviews*, 7, 269–84.

Siegel, S. 2009: The visual experience of causation. *Philosophical Quarterly*, 59, 519–40.

Smith, B. 2015: The chemical senses. In M. Matthen (ed.), *The Oxford Handbook of Philosophy of Perception*. Oxford: Oxford University Press, 314–52.

Spence, C., Baddeley, R., Zampini, M., James, R., and Shore, D. 2003: Multisensory temporal order judgments: When two locations are better than one. *Perception & Psychophysics*, 65, 318–28.

Spence, C. and Bayne, T. 2015: Is consciousness multisensory? In D. Stokes, M. Matthen, and S. Biggs (eds.), *Perception and Its Modalities*. Oxford: Oxford University Press, 95–132.

Spence, C., Shore, D., and Klein, R. 2001: Multisensory prior entry. *Journal of Experimental Psychology: General*, 130, 799–832.

Spence, C. and Squire, S. 2003: Multisensory integration: maintaining the perception of synchrony. *Current Biology*, 13, R519–21.

Stein, B., Burr, D., Constantinidis, C., Laurienti, P., Alex Meredith, M., Perrault, T., Ramachandran, R., Röder, B., Rowland, B., Sathian, K., Schroeder, C., Shams, L., Stanford, T., Wallace, M., Yu, L., and Lewkowicz, D. 2010: Semantic confusion regarding the development of multisensory integration: a practical solution. *European Journal of Neuroscience*, 31, 1713–20.

Stone, J., Hunkin, N., Porrill, J., Wood, R., Keeler, V., Beanland, M., Port, M., and Porter, N. 2001: When is now? Perception of simultaneity. *Proceedings of the Royal Society B*, 268, 31–8.

Strawson, P. 1959: *Individuals*. New York: Routledge.

Treisman, A. 1996: The binding problem. *Current Opinion in Neurobiology*, 6, 171–8.

Treisman, A. 2003: Consciousness and perceptual binding. In A. Cleeremans (ed.), *The Unity of Consciousness: Binding, Integration, and Dissociation*. Oxford: Oxford University Press, 95–113.

Treisman, A. and Schmidt, H. 1982: Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–41.

Treisman, A. and Gelade, G. 1980: A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.

Vatakis, A. and Spence, C. 2007: Crossmodal binding: evaluating the ‘unity assumption’ using audiovisual speech stimuli. *Perception and Psychophysics*, 69, 744–56.

Welch, R. and Warren, D. 1980: Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638–67.

Zampini, M., Guest, S., Shore, D., and Spence, C. 2005: Audio-visual simultaneity judgments. *Perception and Psychophysics*, 67, 531–44.

Zmigrod, S. and Hommel, B. 2011: The relationship between feature binding and consciousness: evidence from asynchronous multi-modal stimuli. *Consciousness and Cognition*, 20, 586–93.

Zmigrod, S., Spapé, M., and Hommel, B. 2009: Intermodal event files: integrating features across vision, audition, taction, and action. *Psychological Research*, 73, 674–84.

Zmigrod, S. and Zmigrod, L. 2015: Zapping the gap: reducing the multisensory temporal binding window by means of transcranial direct current stimulation (tDCS). *Consciousness and Cognition*, 35, 143–9.