

# Intermodal binding awareness

Casey O'Callaghan

To appear in *Sensory Integration and the Unity of Consciousness*, MIT Press  
eds. David Bennett and Christopher Hill

It is tempting to hold that perceptual experience amounts to a co-conscious collection of visual, auditory, tactual, gustatory, and olfactory episodes. If so, each aspect of perceptual experience on each occasion is associated with a specific modality. This paper, however, concerns a core variety of multimodal perceptual experience. It argues that there is perceptually apparent *intermodal* feature binding. I present the case for this claim, explain its consequences for theorizing about perceptual experience, and defend it against objections. I maintain that just as one thing may perceptually appear at once to jointly bear several features associated with the same sense modality, one thing also may perceptually appear at once to jointly bear features associated with different sense modalities. For instance, just as something may visually appear at once to be both red and round, or to have a red part and a green part, something may multimodally perceptually appear at once to be both bright and loud, or to have a red part and a rough part. The main lesson, I argue, is that perceiving is not just co-consciously seeing, hearing, feeling, tasting, and smelling at the same time. And perceptual phenomenal character is not on each occasion exhausted by that which is distinctive to or associated with a given modality, along with that which accrues thanks to simple co-consciousness. Not all ways of perceiving are modality specific. I defend this account against three main objections: that singular content theorists avoid my conclusions; that apparent infusion of perceptible features is required for perceptually apparent binding but does not occur intermodally; and that the diversity of objects across modalities makes perceptually apparent intermodal binding rare.

## 1 Feature binding awareness

Humans are able consciously to perceive *things* and their *features*. You may *see* a baseball in addition to its dominant color or its laces. You may *hear* a sound and its pitch or its duration. I am speaking of cases in which a subject through the use of the senses consciously experiences or becomes sensorily aware of that which is

perceived.<sup>1</sup>

Perceptible *feature bearers* include *individual* things and happenings. Their perceptible *features* may include sensible *attributes*, such as qualities or properties, and *parts*, such as surfaces or segments. That humans perceive individual things does not imply that *bare particulars* are perceptible. And here I remain neutral on the metaphysical nature of perceptible *attributes*.

Perceptible feature bearers commonly are perceived *to have* or *as having* their perceptible features. You may see a baseball's being cream-colored or its having laces. You may hear a sound to be high-pitched or to have a long duration. I'll ignore the difference between perceiving a thing's being (or having) *F* and perceiving a thing to be (or to have) *F*. In illusion, perceptible feature bearers are perceived to have or as having features they lack.

Individual things may be perceived at once to jointly have or as jointly having *multiple* features. You see the baseball at once to be cream-colored, spherical, and laced. You hear the sound at once to be loud and shrill. You feel the cool surface to have a rough part. In each case, the features are perceived to jointly belong to one thing. Perceptible feature bearers may in differing ways illusorily seem perceptually to jointly bear distinct features.

When you consciously perceive features' jointly belonging to the same thing, say that you consciously perceive those features to be *bound*. Call this a case of *feature binding awareness*. Allow that feature binding awareness may be illusory, so that you may for some reason consciously perceive things merely *as* jointly bearing distinct features. To simplify, I'll assume that feature binding awareness sometimes is veridical perception. The right adjustments accommodate cases of misperception.

My concern is not in the first instance the apparent relation between a single feature and its bearer. A single feature perceptibly belonging to an individual does not suffice for feature binding awareness as I understand it. I mean to focus on cases in which differing features perceptibly belong at once to the same thing. This is the standard concern of empirical work on feature binding.

Feature binding awareness may involve perceiving *attributes'* (properties or qualities) jointly sharing a bearer or being coinstantiated, or it may involve perceiving *parts'* jointly belonging to or composing (perhaps partly) the same whole. So, feature binding awareness need not involve a single uniform qualitative character. Nonetheless, feature binding awareness may depend upon a common type of mechanism. Anne Treisman's influential work, in particular, expressly tests for common mechanisms in property binding and part binding.<sup>2</sup>

(i) [T]he *the binding of properties*—the integration of different object properties, such as the color red and the shape + to form a red cross; (ii) the

---

<sup>1</sup>For instructive discussion of experiential awareness, see Hill (2009, esp. Chapter 3). For a differing perspective on sensory awareness, see Johnston (2006).

<sup>2</sup>See, for example, Treisman and Gelade (1980); Treisman (1988, 1996, 2003).

*binding of parts*—the integration in a structured relationship of the separately codable parts of a single object, such as the nose, eyes, and mouth in a face. . . . The first two seem to me to be closely related and to depend on the same binding mechanism. (Treisman 2003, 98)

“Binding” commonly is used to refer to the perceptual *process* by which information concerning distinct perceptible features (such as color and shape, or distinct components) is bundled together as information concerning a common perceptible item.<sup>3</sup> Talk of feature binding in this sense risks conflating information and its subject matter. This paper concerns perceptual *awareness*, so I characterize feature binding in terms of perception and its objects. Feature binding occurs when differing features perceptibly belong to a common individual. Feature binding awareness involves differing features perceptually appearing to belong to a common entity—to be coinstantiated by an individual, or to be parts of the same whole.

Feature binding awareness presumably depends upon feature binding processes. I say “presumably” because a feature binding process, as described in the first sentence of the previous paragraph, may require that features are detected or analyzed separately by subpersonal perceptual mechanisms. There is powerful evidence for this claim. But it is possible that what I have characterized as feature binding awareness could occur without such a feature binding process. In any case, my topic is the conscious upshot rather than the process. This does not set my discussion off entirely from traditional debates about feature binding, which often focus upon the psychological and explanatory relations between feature binding processes, which may be subpersonal, and phenomenal characteristics of conscious awareness.<sup>4</sup> My main topic is conscious perceptual awareness.

## 2 Intramodal feature binding awareness

This paper presupposes that conscious episodes of feature binding awareness sometimes occur. Paradigm feature binding awareness is *intramodal*. Visual feature binding is best understood. The egg looks whitish and ovoid, and “Q” has a visible part “O” lacks. A rich experimental literature—concerning, for instance, the role of attention in binding, illusory conjunctions of features, serial versus parallel search, object-specific preview advantages and penalties, and the role of pre-attentive segmentation and grouping—has investigated intramodal visual feature binding processes and their relationship to conscious visual feature binding awareness.

Audition also involves feature binding awareness. The sound audibly is high pitched but wavering, and an utterance of “overtly” has an audible part an utterance of “overt” lacks. Research on *auditory scene analysis* tracing especially to

---

<sup>3</sup>See, for instance, Treisman (1996, 2003).

<sup>4</sup>See, for instance, Treisman (1982, esp. 212–13); Treisman (1988, esp. 204); Treisman (2003, esp. 109ff.); Mitroff et al. (2005).

Bregman (1990) has illuminated the mechanisms responsible for auditory feature binding awareness.

Touch, too, involves feature binding awareness. The surface feels smooth and warm to the touch.<sup>5</sup> Taste, gustation, and olfaction are more difficult cases, but each might also enable you to perceive individuals and bound features. The cookie in your mouth tastes sweet and salty. You might smell the odor to be jointly rancid and intense.<sup>6</sup>

Not all modalities reveal the same individuals and features. But individuals and bound features are part of the structure revealed by perceptual awareness in several exteroceptive sensory modalities.<sup>7</sup> Intramodal feature binding awareness occurs in more than one modality.

### 3 Intermodal feature binding awareness

Is there *intermodal* feature binding awareness? If so, features consciously perceived through different modalities can perceptually appear to be bound and thus to belong to the same thing.

The skeptical position is that there is intramodal but not intermodal feature binding awareness.<sup>8</sup> Humans only *associate* features perceived through different modalities or *infer* that they belong to the same object. Fulkerson (2011), for instance, claims that only unimodal perceptual experiences involve apparent feature binding. According to Fulkerson, “the predication or assignment of distinct features to perceptual objects” is “a distinguishing feature of unisensory perceptual experiences” (504–5).

Multisensory perceptual experiences do not involve the direct predication of features onto individual perceptual objects. Instead, there is an association *between* experiences. . . . What we experience is a higher-order association between sensory experiences. (Fulkerson 2011, 506)

Fulkerson denies that a multimodal perceptual experience can ascribe to a perceptible object distinct features associated with different modalities. Instead, he thinks distinct unimodal experiences of worldly objects are *associated* in a higher-order multisensory perceptual experience. Fulkerson thus rejects intermodal binding awareness as I have characterized it.

---

<sup>5</sup>See, for example, Fulkerson (2011).

<sup>6</sup>See, for example, Batty (2011).

<sup>7</sup>See, in particular, Clark (2000), who maintains that different features are attributed to individual locations. See, also, Matthen (2005, 277–92), who holds that vision and audition, but not olfaction, involve perceptually attributing multiple features to objects (though not to commonplace material objects in the case of audition).

<sup>8</sup>This should be distinguished from the differing skeptical position that there is no feature binding awareness at all, which §2 set aside.

Spence and Bayne (2014, esp. §7) argue that there is good evidence for unimodal but not multimodal feature binding awareness.

But are features belonging to different modalities bound together in the form of MPOs [multimodal perceptual objects]? ... [W]e think it is debatable whether the “unity of the event” really is internal to one’s experience in these cases, or whether it involves a certain amount of post-perceptual processing (or inference). In other words, it seems to us to be an open question whether, in these situations, one’s experience is of a MPO or whether instead it is structured in terms of multiple instances of unimodal perceptual objects. (Spence and Bayne 2014, ms 27, 29)

Spence and Bayne are skeptical whether perceptual consciousness includes awareness as of unified objects that bear features associated with different modalities. They propose to admit only apparent unity stemming from *post-perceptual* processing or inference rather than apparent unity that is “internal to one’s experience.” Thus, Spence and Bayne express skepticism about intermodal binding awareness.<sup>9</sup>

Here I want to present a case for intermodal binding awareness, to spell out its consequences, and to defend it against objections. My aim is not to refute the determined skeptic. And I do not claim that every variety of intramodal binding awareness occurs intermodally. Instead, my aim is to show that we should prefer a position that recognizes certain forms of intermodal binding awareness.

My case for a non-skeptical position begins with a contrast between (1) and (2).

- (1) Perceiving a thing’s being *F* and a thing’s being *G*.
- (2) Perceiving a thing’s being both *F* and *G*.

An instance of (2) requires that a single thing perceptibly has both features. However, an instance of (1) does not require that. Feature binding awareness occurs

---

<sup>9</sup>Connolly (2014) is a more complicated case. He endorses elements of skepticism. At the outset, he seems to reject intermodal binding awareness.

Are some of the contents of perception fused multimodal units (fused audio-visual units, for instance)? I think that the answer is no. ... [W]e need not hold that the content of *Q1* involves a fused audio-visual property, since we can explain that phenomenal type in terms of an auditory and a visual property. (Connolly 2014, proof 5).

And he says multimodal episodes can be explained in terms of “a conjunction of an audio content and visual content” (proof 13) but do not involve “fused audio-visual content” (proof 5–6). However, he later states that perceptual experiences may have additional *amodal* contents, involving individuals, objects, or events, characterized in modality-independent terms, of which modality-specific features may be predicated. He does not say outright whether such perceptual episodes involve mere association or intermodal binding awareness. While Connolly’s account raises questions beyond my paper’s scope, it means Connolly need not endorse thoroughgoing skepticism about intermodal binding awareness. Connolly, in correspondence, has suggested a preference for an account in terms of association rather than binding awareness.

just in case the difference between (1) and (2) sometimes is reflected in perceptual awareness.

Consider an intramodal example. Hearing a thing's being loud and a thing's being high-pitched differs from hearing a thing's being both loud and high-pitched. Hearing a thing's being both loud and high-pitched requires that a single thing perceptibly is loud and high-pitched; hearing a thing's being loud and a thing's being high-pitched does not require that. If intramodal feature binding awareness occurs, an intramodal episode of (2) may differ from an intramodal episode of (1) in a way that is reflected in perceptual awareness. Since we have assumed that intramodal feature binding awareness occurs, the difference between (1) and (2) may be reflected in perceptual awareness.

My claim is that the difference between (1) and (2) may be reflected in multimodal episodes of conscious perceptual awareness. There are intermodal episodes of (2) that are not merely episodes of (1). For instance, there are episodes of consciously perceiving a thing's being both bright and loud that are not just episodes of consciously perceiving a thing's being bright and a thing's being loud. The difference is reflected in perceptual awareness. This has important consequences for theorizing about perception. Not every aspect of perceptual awareness is associated with a specific modality or accrues thanks to simple co-consciousness.

Before I present the evidence and discuss the consequences, three further clarifications are needed. First, Tye (2003, 2007) uses an argument with a similar structure to argue that experiences associated with different modalities are co-consciously unified. But Tye's concern differs from mine. He contrasts, for instance, having a visual experience and an auditory experience at the same time with having an experience that is both auditory and visual. Some of Tye's examples involve a phenomenally unified multimodal experience as of a common object, but he does not draw the contrast I emphasize in this paper. This contrast holds between pairs of co-consciously phenomenally unified multimodal experiences. It holds, for instance, between a phenomenally unified audiovisual experience as of hearing a thing's being *F* and seeing a thing's being *G* and a phenomenally unified audiovisual experience as of a thing's being both *F* and *G*.

Next, I have assumed for simplicity that humans sometimes do consciously perceive objects and features and thus may consciously perceive features to be bound. But, as much as possible, I aim to be neutral regarding theories of perception. Representational content theorists may prefer an alternative formulation of the contrast.

(1') Perceptually representing that a thing is *F* and a thing is *G*.

(2') Perceptually representing that a thing is both *F* and *G*.

Finally, philosophers have focused on perceptible attributes and their apparent bearers, but feature binding awareness also may involve parts. Relations among

perceptible parts and wholes are part of the apparent structure revealed by perceptual awareness. Thus, a version of the contrast involves parts.

(1'') Perceiving a thing's having *a* as a part and a thing's having *b* as a part.

(2'') Perceiving a thing's having both *a* and *b* as parts.

The differences matter, but this paper's guiding concern is what the differing versions of the contrast have in common. Each involves a contrast between conscious episodes in which a subject is perceptually aware of a common item's jointly bearing features which are perceived at once using different senses and episodes in which a subject need not be perceptually aware of any such common item. In short, each contrast requires that a common individual may be perceptible as such across different senses.

#### 4 Evidence for intermodal feature binding awareness

##### *Perceptual judgment*

Five sources of evidence converge to support intermodal feature binding awareness. The first concerns perceptual judgment. In many ordinary cases, perceptual evidence does not support an *immediate perceptual judgment* that a thing is both seen and felt, as when you see an airplane and touch a baseball, or when you touch a baseball and unknowingly see it reflected in a mirror. However, perceptual evidence may support an immediate perceptual judgment that what is seen is what is felt, or that something perceived bears both visible and tactual features. Imagine holding a baseball while looking at it. Normally, it would be silly to judge on appearances that the object you see is numerically distinct from the object you feel. The simplest explanation is that you perceive the sphere in your hand at once to be jointly white and red, smooth and leathery; it appears perceptually that there is a white and red, smooth, leathery sphere in your hand. That is what you tend to judge.

However, someone might object that, even granting their veridicality, perceptual appearances leave room for doubt whether features perceived with different modalities belong to the same thing. Thus, the identification does not hinge *just* upon perceptual appearances or looks; it is not simply a matter of endorsing appearances. If so, the coinstantiation of features need not be perceptible intermodally; instead, it may be cognized only through contributions from further post-perceptual resources. Therefore, the claim is that the identification is neither perceptual nor an immediate perceptual judgment but instead belongs fully to extra-perceptual cognition.

### *Perception-guided action*

We can make progress by recognizing that common *perceptually guided actions* suggest that you sometimes are sensitive to the identity of things perceived through different modalities in a way that does not require perceptual judgment. Imagine crossing a street and hearing something rapidly approaching from your left. You may reflexively jump out of the way, or you may turn quickly to look for it. But it makes little sense to jump from or to look for a sound. Your actions instead suggest sight and hearing share objects. Moreover, once you've picked it up by sight, you track and respond to it *as* a unified perceptible thing or happening, accessible to sight and hearing, rather than as distinct individuals. Another example involves seeing a baseball coming at you and visually "guiding" it into your mitt. Your activities coordinate sight and touch in a way that suggests you implicitly recognize the ball as a common perceptible target. This ability extends to novel circumstances, so it generalizes. An additional example involves using sight to orient yourself so that you can better listen to the source of a sound. Slightly angling your face away from a source often improves how it sounds.

Such activities involve responsiveness, orienting, and tracking across modalities. They suggest you *perceptually* identify or are sensitive to the identity of what's seen with what's heard or felt. The manner in which multimodal perception guides action supports intermodal binding awareness.

However, someone might object that such actions could depend on pure (but fancy) reflexes, on sophisticated learned associations and coordinated predictions, or on snap judgments and implicit inferences rather than on perception. Moreover, perception for action may be functionally distinct from perception for recognition and awareness. Thus, even if intermodal perception for action identifies common objects, a subject might still wholly lack intermodal binding awareness.

### *Empirical research*

A third source of evidence supports the claim that the identification of common objects is not limited to perception for action. A great deal of recent empirical work on multisensory perception claims that perceptual systems integrate and bind information from different senses to yield unified perceptual awareness of common multimodally accessible objects or events. Here are four representative passages that concern audio-visual binding.

When presented with two stimuli, one auditory and the other visual, an observer can perceive them either as referring to the same unitary audiovisual event or as referring to two separate unimodal events. . . . *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, 744, 754, my italics)

As an example of such privileged binding, we will examine the relation between visible impacts and percussive sounds, which allows for *a particularly powerful form of binding that produces audio-visual objects*. (Kubovy and Schutz 2010, 42, my italics)

In a natural habitat information is acquired continuously and simultaneously through the different sensory systems. As some of these inputs have the same distal source (such as the sight of a fire, but also the smell of smoke and the sensation of heat) it 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. For one such area where intermodal binding (IB) seems important, that of concurrently seeing and hearing affect, *behavioural studies have shown that indeed intermodal binding takes place during perception*. (Pourtois et al. 2000, 1329, my italics)

[T]here is undeniable evidence that the visual and auditory aspects of speech, when available, contribute to an integrated perception of spoken language... *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, 447, my italics)<sup>10</sup>

The main source of empirical evidence for intermodal binding is that sensory systems interact and share information. Cross-modal recalibrations are effects in which a stimulus presented to one sensory system impacts experience associated with another sense modality. Sometimes this generates an illusion. For instance, compelling ventriloquism involves an auditory spatial illusion produced by the visible location of an apparent sound source—the visual stimulus affects auditory spatial experience. In the McGurk effect, video of a speaker uttering /ga/ presented with audio of /ba/ leads subjects to mistakenly hear the utterance as /da/. So, processes associated with one sense sometimes interact causally with processes associated with another sense, and this can alter experience from what it otherwise would have been.<sup>11</sup>

Explaining such cross-modal effects as mere causal influence misses something important. Welch and Warren (1980) say:

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, 661, 664, my italics)

---

<sup>10</sup>See also, for example, Bushara et al. (2003); Bertelson and de Gelder (2004); Spence and Driver (2004); Spence (2007); Stein (2012).

<sup>11</sup>See O'Callaghan (2012) for a catalog and review of cross-modal illusions and recalibrations.

For instance, in ventriloquism, visual and auditory spatial information may be recalibrated to produce a concordant spatial experience. In the McGurk effect, alveolar /da/ is a compromise between the visible velar /ga/ and the audible bilabial /ba/. So, discrepant or conflicting information from different sensory systems is reconciled in order to reduce or resolve conflict. But conflict requires a common subject matter. Thus, if perceptual processes resolve conflicts between the senses, they treat information as if it has a common subject matter or shares a source. This requires discerning whether or not different sensory messages concern the same thing and thus belong together as candidates for reconciliation. (The alternative to attributing incompatible features to one item is attributing differing features to distinct items.) So, among perceptual strategies and mechanisms responsible for intermodal recalibrations and illusions, those that reduce and resolve conflicts require the capacity to treat information from different sensory systems as stemming from a common source—as concerning the same things or features. A unified sub-personal grasp upon common perceptible objects may in turn ground unified perceptual awareness as of a single event with visible and audible features.

However, one might object. Grant that there is a pattern of causal influence across sensory systems that conforms to principles of conflict resolution, and grant that information is transmitted between senses. This does not require a common or unified representation, and it does not by itself constitute a unified grasp or representation as of a common object or feature bearer. Perceptual mechanisms might effectively resolve conflicts between distinct information streams without integrating or binding them together. The performance of effective conflict resolution need not involve explicitly tracking or representing any common sources as such. Thus, further empirical evidence is needed for intermodal feature binding.

In fact, standard empirical measures of intramodal feature binding also provide evidence for intermodal feature binding. For instance, *multisensory integration*, *illusory conjunctions*, *object-specific preview effects*, *multimodal object files* and *intermodal event files* (temporary episodic representations of persisting real world objects and events) have been studied and reported in a variety of intermodal conditions.<sup>12</sup> The

---

<sup>12</sup>On *multisensory integration*, see, e.g., Stein and Stanford (2008); Stein et al. (2010); Stein (2012). On intermodal *illusory conjunctions*, see Cinel et al. (2002). On intermodal *object-specific preview benefits and penalties*, *object files* and *event files*, see Zmigrod et al. (2009) and Jordan et al. (2010). The remainder of this note describes a selection of these results in additional detail.

Stein et al. (2010) characterize *multisensory integration* as “The neural process by which unisensory signals are combined to form a new product” (1719). For instance, *superadditive effects* occur when the multisensory neural or behavioral response to a stimulus is significantly greater than the sum of the modality-specific responses to that stimulus. Such effects are evidence that perceptual processes do not merely reconcile conflicts. Instead, multisensory processes sometimes integrate information concerning a common source and generate a novel response to it as such.

A traditional source of support for intramodal feature binding is the existence of *illusory conjunctions* (ICs) of features, especially outside focal attention (see, e.g., Treisman and Schmidt 1982). Unattended perceptible features may mistakenly appear coinstantiated. For instance, an unattended red square and green circle may mistakenly cause the perceptual impression of a red circle. An

important upshot of this experimental work (which footnote 12 describes in additional detail) is that perceptual processes indeed do involve tracking or representing individual feature bearers as common across sensory modalities and as bearing features perceptible with different senses. This addresses the objection raised in the previous paragraph.

---

unattended “O” and “R” may mistakenly cause the perceptual impression of a “Q”. Cinel et al. (2002) presents experimental evidence supporting *cross-modal illusory conjunctions* between vision and touch. For instance, an unattended felt texture may be perceptually ascribed to the wrong visible shape. The authors say, “These results demonstrate that ICs are possible not only within the visual modality but also between two different modalities: vision and touch” (1245). They argue based upon a series of studies that illusory conjunctions of visible and tactual features are “perceptual in nature” rather than effects of memory or extra-perceptual cognition (1261).

Taken together, the evidence is consistent with the idea that information converges preattentively for binding from different sensory modalities and that this binding process is modulated by the parietal lobe. ... The present evidence for cross-modal ICs suggests that there is multimodal integration of sensory information in perception so that misattributions of modalities arise under conditions of inattention. (Cinel et al. 2002, 1244, 1261)

So, the existence of intermodal illusory conjunctions supports intermodal feature binding in perceptual processes.

Another critical diagnostic for intramodal feature binding stems from *object-specific preview effects* (see Kahneman et al. 1992). Kahneman et al. (1992, see esp. 176) propose that visual object perception involves deploying *object files*, which are temporary episodic representations of persisting real world objects. Object files integrate information about distinct perceptible features. Previewing a target affects one’s capacity to recognize it again when its two appearances are “linked” perceptually to the same object (*reviewing*). If an object’s features match at two times, reviewing it enhances recognition; if its features do not match, reviewing it hampers recognition. Object-specific preview effects are used to determine whether or not feature binding occurs. A preview *benefit* requires that matching feature combinations are ascribed to a common object; no object-specific preview benefit accrues for features not initially attributed to the reviewed object. And a preview *penalty* requires mismatching feature combinations ascribed to a common object. Zmigrod et al. (2009, 675) say, “interactions between stimulus-feature-repetition effects are indicative of the spontaneous binding of features and thus can serve as a measure of integration.”

Object-specific preview benefits and penalties occur *intermodally*. Zmigrod et al. (2009, 674–5) report that patterns of interaction that characterize unimodal feature binding occur intermodally between audition and vision, and between audition and touch. They argue, for instance, that color–pitch pairs may be bound, since presenting *color*<sub>1</sub> with *pitch*<sub>1</sub> followed by *color*<sub>1</sub> with *pitch*<sub>2</sub> impairs recognition in a way that differs from what is predicted by modality-specific object files and binding alone. The authors report that perceptual processes involve “episodic multimodal representations” rather than mere intermodal interactions (682) and that feature binding occurs across modalities (683).

In addition, Jordan et al. (2010) report “a standard, robust OSPB” between vision and audition.

Although object files are typically discussed as visual, here we demonstrate that object-file correspondence can be computed across sensory modalities. An object file can be initially formed with visual input and later accessed with corresponding auditory information, suggesting that object files may be able to operate at a multimodal level of perceptual processing. (Jordan et al. 2010, 491)

The authors report that their data “explicitly demonstrate object files can operate *across* visual and auditory modalities” (501).

But it calls attention to another worry. The relationship between perceptual processes that involve feature binding—as operationalized by such experimental measures—and conscious perceptual awareness is not clear. In the intramodal visual case, for instance, Mitroff et al. (2005) report that object-specific preview benefits disagree with conscious visual percepts. Therefore, in some cases, object trajectories as determined by the object-file system may diverge from those apparent in conscious perceptual awareness. Moreover, in intermodal audio-visual cases, Zmigrod and Hommel (2011) claim that implicit measures of intermodal feature binding from event-specific preview effects may disagree with conscious perceptual awareness of audible and visible features as belonging to the same event. Event-specific preview effects can tell one story, and measures of conscious perceptual awareness can tell another. The authors say, “binding seems to operate independently of conscious awareness, which again implies that it solves processing problems other than the construction of conscious representations” (592). Thus, it is risky to draw conclusions about conscious perceptual awareness just from experimental work on feature binding.

Here is where things stand. If there is intermodal feature binding awareness as I have characterized it, some mechanisms are responsible. It remains plausible that empirical research on multisensory integration and binding of information concerning a common subject matter should play a critical role in explaining intermodal feature binding awareness. For instance, it helps to show that sensitivity to the identity of things perceptible through different sense modalities is not wholly cognitive and is not limited to perception for action. However, current empirical work does not definitively account for the relation between integration and binding processes and feature binding awareness. So, we cannot take experimental work on intermodal feature binding at face value as direct support for intermodal feature binding awareness. One may doubt psychologists’ interpretations of their own results, but that is not the issue. Psychological explanations of perceptual mechanisms and processes involving feature binding just do not translate neatly and uncontroversially to claims about conscious perceptual awareness.<sup>13</sup>

### *Perceptual appearances*

The contrast between (1) and (2) marks a difference in how things may appear perceptually *to be*. This difference may be apparent whether or not you *believe* things are as they appear and whether or not things *are* as they appear. When all is going well, the contrast corresponds to a difference in whether or not you are perceptually sensitive to the coinstantiation of features by a common individual perceptible through different senses.

The argument stems from misleading appearances and the possibility of error. On one hand, apparent binding can be illusory. Take a compelling case of ventril-

---

<sup>13</sup>Treisman (2003, esp. 109–11) is emblematic.

oquism. You may seem to hear the visible puppet speaking, even if you are not taken in. Contrast this with a poor attempt at ventriloquism, in which it is perceptually evident that the visible puppet is not what you hear. Or consider movies. Nothing in the theater utters the words you hear *and* is visible on screen. In the psychology lab, you wear headphones and watch video of discs apparently colliding with a clack. Since there is no particular perceptible event with those visible and audible features, the appearance as of a common source is an illusion. The illusion need not be spatial or temporal, since the speaker could be placed right behind the movie screen. And it does not require belief. A mere case of (1) may simply seem like a case of (2), where the difference concerns that to which you are perceptually sensitive.<sup>14</sup>

On the other hand, visible and audible features can appear to belong to distinct individuals, or not appear bound, even if you know they belong to one thing. In successful ventriloquism, the sounds appear to come from the dummy but in fact come from the ventriloquist you see. Or take the trick in which you cross one wrist atop the other, weave your fingers together, twist your hands inward and up, visually target a finger, and try to raise it. When the trick works, before you move anything, the seen but visually untargeted finger you surprisingly raise seems distinct from that finger as it is felt. Perceiving features that are coinstantiated seems like a mere case of (1). You fail to be sensitive perceptually to the identity of an individual and to the coinstantiation of its features.

So, apparent intermodal binding can be illusory, and features of one thing can mistakenly perceptually appear to be features of distinct things (or may simply not perceptually appear to be bound). These possibilities support the claim that there are cases in which intermodal feature binding is perceptually apparent that differ in what is presented in experience from other cases in which it is not. This provides the materials for a reply to Spence and Bayne. Each of these effects decouples from what you think and what you are inclined to judge on extra-perceptual grounds. Therefore, the differing appearances are not due wholly to extra-perceptual cognition or inference. These cases involve differences in conscious perceptual awareness. It also provides the materials for a reply to Fulkerson. The differences concern what you may be consciously perceptually aware of, not simply relations between experiences. A mere conscious association between experiences cannot in itself be an illusion or misperception. However, suppose that such associations between experiences ground a difference in how things seem perceptually to be and thus may be accurate or illusory. If so, in order for skepticism to have teeth, merely seeming to be associated or to tend to co-occur must differ from seeming to belong to something common. But, if seeming to be associated or to tend to co-occur does

---

<sup>14</sup>It is noteworthy that Austin mentions ventriloquism as an example of illusion: "Then again there are the illusions produced by professional 'illusionists', conjurors—for instance the Headless Woman on the stage, who is made to look headless, or the ventriloquist's dummy which is made to appear to be talking" (Austin 1962, 22–3).

not guarantee seeming to share a common object or source, then appearing merely as being associated or as tending to co-occur is too permissive to capture the relevant distinctions among the cases discussed above. For instance, a sound and an image may seem merely to be associated or to tend to co-occur without seeming perceptually to share a common source. A rough surface and a red surface may seem to be associated without their seeming perceptually to be one surface or to belong to one object. Mere associations thus do not suffice for an account of that to which one may be multimodally perceptually sensitive, and they do not suffice for an account of multimodal perceptual awareness.

### *Perceptual phenomenology*

A skeptic nevertheless might question whether the difference between (1) and (2) itself may be marked by a difference in the phenomenology of perception.

Imagine watching a movie with a compelling, immersive sound track. You hang on the actors' words and jump from your seat at the explosions. It sounds like planes flying up behind you and overhead. Now imagine the sound track's timing is off. It could be just a little bit, so that it is noticeable but not disturbing. It could be even more, so that the experience is jarring. Or it could be a lot, so that the sights and sounds appear wholly dissociated. In each of these four cases, the auditory and visual stimulation independently remain qualitatively the same, but the phenomenology differs unmistakably. The *alignment* matters. The dramatic phenomenological difference between the perfect soundtrack and the very poorly aligned soundtrack stems in part from perceiving audible and visible features as belonging to something common in the coincident case but not in the misaligned case. The contrast is between apparent intermodal episodes of (2) and of (1).<sup>15</sup> A similar argument applies to dubbed foreign language films. In that case, the fine-grained structures mismatch.

Someone may object: These experiences differ in *spatio-temporal* respects; once you control for spatio-temporal differences, such as those involving apparent temporal or spatial relations between what's audible and visible, any experiential difference dissolves.

Notice that in this respect my case parallels that of perceiving causality. Stimulus features that cue perceptual awareness as of causality also are responsible for the scene's apparent spatio-temporal features. The main features that indicate causation just are spatio-temporal. So, it is difficult to control for perceptually apparent spatio-temporal features.

---

<sup>15</sup>The phenomenological difference between the jarring third case and the far off fourth case also may involve a contrast between apparent episodes of (2) and of (1). Perhaps what makes it jarring is the sense that the misaligned features belong to something common, and thus should be aligned. But perhaps the jarring third case is a better candidate for seeming merely to be associated, in respect of which it differs from the slightly misaligned but not jarring second case, which involves apparent intermodal binding awareness.

In the case of intermodal binding awareness, there is a clear way forward. Intermodal binding awareness may depend not just on spatio-temporal cues, but also on factors such as whether and how the subject is attending, the plausibility of the combination or the compellingness of the match, and whether the subject expects one event or multiple events to occur.

The binding versus segregation of these unimodal stimuli—what Bedford (2001) calls the object identity decision; see also Radeau and Bertelson (1977)—depends on both low-level (i.e., stimulus-driven) factors, such as the spatial and temporal co-occurrence of the stimuli (Calvert, Spence, & Stein, 2004; Welch, 1999), as well as on higher level (i.e., cognitive) factors, such as whether or not the participant assumes that the stimuli should “go together.” This is the so-called “unity assumption,” the assumption that a perceiver makes about whether he or she is observing a single multisensory event rather than multiple separate unimodal events. ... (Vatakis and Spence 2007, 744)

Fixing spatio-temporal features does not by itself suffice to fix whether intermodal binding awareness occurs.

*At the same time, the 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, 754, my italics)*

Thus, it is possible to tease apart the appearance of intermodal feature binding from perceptually apparent spatio-temporal features. Fixing apparent spatio-temporal features need not fix whether or not intermodal feature binding is perceptually apparent. Take a pair of cases that controls for spatio-temporal features and for other aspects of perceptual phenomenology. A case in which you “get” the perceptual effect of intermodal binding awareness may contrast in character with an otherwise similar one in which you do not.

In addition, the capacity for intermodal binding can be disrupted. Individuals with autism have difficulty integrating cues about emotion from vision and audition. But mechanisms for integrating information from different sets of senses or even features may be dissociated, so localized deficits or brain damage may not cause a wholesale inability to perceive features as bound intermodally. Instead, *specific forms* of intermodal feature binding awareness may fail. For instance, Pasalar et al. (2010) shows that transcranial magnetic stimulation can disrupt visuo-tactile sensory integration. Hamilton et al. (2006) report a patient who is unable to 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).

Controlling for spatio-temporal differences—even apparent ones—therefore need not dissolve the phenomenological difference in perceptual experience.

Higher-level cognitive factors sometimes may play a role in determining whether or not intermodal feature binding awareness occurs. This implies neither that intermodal feature binding is extra-perceptual nor that the phenomenology of intermodal binding awareness is wholly cognitive. Cognition may causally but not constitutively influence perception, and intermodal binding awareness need not involve awareness of the relevant cognitive factors.

One complication concerns the role of attention. I am attracted to the idea that intermodal attention is required for intermodal feature binding awareness. So, suppose there are differing ways of deploying attention. For instance, you might maintain distinct intramodal streams, or you might sustain a single intermodal focus. If so, phenomenological differences associated with these differing ways of deploying attention might account for phenomenological differences between apparent cases of (1) and of (2) that otherwise are alike. Nevertheless, given that perceptual attention targets individual objects or groups whose members are treated as parts of a unified perceptible entity, a single intermodal focus may require recognizing a common perceptible item. Such attended items may perceptibly bear features associated with different senses.

### *Summary*

Perceptual judgment, perception-guided action, empirical research on multisensory perception, perceptual appearances, and perceptual phenomenology together provide good evidence that intermodal episodes of (2) may contrast with intermodal episodes of (1), that intermodal episodes of each occur, and that the difference sometimes is reflected in perceptual awareness. Humans can be perceptually aware as of something's jointly having both visible and audible features. This may differ from seeing as of something's having visible features and hearing as of something's having audible features. Only the latter is compatible with their being apparently distinct individuals. Thus, perceptually apparent intermodal feature binding occurs. There is intermodal feature binding awareness.

## **5 Consequences**

### *Perception is not just minimally multimodal*

Intermodal feature binding awareness has noteworthy consequences. It follows that consciously perceiving an individual object or event is not always a modality-specific episode. Some *ways* to perceive individuals cannot be analyzed just in terms of ways in which you could perceive with specific modalities on their own. For instance, visuotactually perceiving a thing's being jointly *F* and *G* is not merely co-consciously seeing a thing's being *F* and feeling a thing's being *G*, where it just

happens that the same thing is *F* and *G*. Perceptually appreciating or being sensitive to the identity of what is seen and felt cannot occur unimodally. So, visuotactually perceiving a thing's being *both F* and *G* is not a way of perceiving that boils down to jointly occurring episodes of seeing and feeling that could have occurred independently.

Thus, overall perceptual awareness is not just a matter of co-consciously seeing, hearing, feeling, tasting, and smelling. Where *F* and *G* are perceived thanks to different senses, an attentive sensory episode of perceiving a thing's being both *F* and *G*, in which you are sensitive to and able to recognize the identity of what is *F* with what is *G*, may not be factorable without remainder into co-conscious modality-specific components that could have occurred independently from each other.

A related argument shows that not all perceptual *phenomenal character* is modality specific. Suppose that the phenomenal character *associated* with some modality on an occasion includes just that which could be instantiated by a perceptual experience wholly of that modality under equivalent stimulation, where a perceptual experience wholly of some modality belongs to that and no other modality. For example, given a particular multimodal perceptual experience, the phenomenal character *associated* with vision on that occasion includes just that which could be instantiated by a wholly visual perceptual experience under equivalent stimulation, where a wholly visual perceptual experience is one that is visual but not auditory, tactual, olfactory, or gustatory. The previous section's arguments have as a consequence that visuotactually perceptually experiencing a thing's being jointly *F* and *G* may have phenomenal features that could not be instantiated either by a wholly visual or by a wholly tactual perceptual experience and that do not accrue thanks to mere co-consciousness. (Phenomenal features that accrue thanks to mere co-consciousness may include simple co-conscious phenomenal unity or those that supervene upon phenomenal character that is associated with specific modalities.) To demonstrate this, suppose seeing a thing's being *F* could be a wholly visual perceptual experience, and suppose feeling its being *G* could be a wholly tactual perceptual experience. Co-consciously seeing a thing's being *F* and feeling a thing's being *G*, where it happens that what's seen is what's felt, does not suffice for visuotactually perceptually experiencing as of a thing's being *both F* and *G*. So, co-consciously seeing a thing's being *F* and feeling a thing's being *G* may differ phenomenally from visuotactually perceiving a thing's being jointly *F* and *G*. Thus, the phenomenal character of a multimodal perceptual episode need not be exhausted by that which is associated with each of its modalities along with that which accrues thanks to mere co-consciousness. Therefore, not all phenomenal character on each occasion is modality specific.

While intermodal binding awareness as I have characterized it entails this conclusion, it is worth being explicit that a skeptical position about intermodal binding awareness is compatible with the conclusion. For instance, Fulkerson's account in

terms of conscious higher-level associations between modality-specific experiences entails the same conclusion while rejecting intermodal binding awareness. Nevertheless, skepticism about intermodal binding awareness is required to maintain that all phenomenal character apart from that which accrues thanks to mere co-consciousness is modality-specific.

*Phenomenal character is not locally distinctive*

Many philosophers say that perceptual experiences of a given modality have a *distinctive* phenomenal character.<sup>16</sup>

From the above it follows that not all perceptual phenomenal character is *locally distinctive* since not every phenomenal feature is distinctive to a specific modality. That is, it is not the case that each perceptual phenomenal feature could be instantiated only by perceptual episodes associated with a certain modality.<sup>17</sup>

This is not just the traditional argument from *common sensibles*. The argument from intermodal feature binding requires that it is possible at a time to perceive visible and audible features to be coinstantiated, and the argument from common sensibles does not. And, unlike the traditional argument from common sensibles, it is not feasible to escape the argument from intermodal feature binding with help from modality-specific modes of presentation or modality-inflected phenomenal character (phenomenal character that is partly a product of the modality itself, understood as a mode of intentionality). Each leaves unaddressed the phenomenal character of perceptually experiencing as of a single something's having both visible and audible features—the phenomenally apparent numerical sameness of an individual that is seen and heard.

Phenomenal character nonetheless may be *regionally distinctive* within a modality. Due to perceptually apparent proper sensibles, the overall phenomenal character associated with any given modality on any occasion may be distinctive in that it could only be instantiated by perceptual experiences of that same modality. However, this comes at a cost. Since local distinctiveness fails, you may not be able to tell what modality a phenomenal feature is associated with on an occasion. So, there may be no clear verdict concerning which phenomenal features, among many candidates, belong to the distinctive overall character that is associated with a given modality. Thus, the boundaries of the phenomenal character associated with a modality on an occasion may not be introspectible, and they are not settled just by considering what's distinctive.

---

<sup>16</sup>See, for example, Grice (1962, esp. 267); Peacocke (1983, esp. 27–28); Lopes (2000, esp. 439).

<sup>17</sup>It is false that for every perceptual phenomenal feature  $f$  there exists a unique perceptual modality  $m$  such that every possible perceptual experience that instantiates  $f$  belongs to modality  $m$ .

## 6 Objections and replies

### *Singular contents*

I have aimed to be neutral regarding theories of perception. But (1) and (2) talk about perceiving “a thing,” and the contrast between (1) and (2) is clear when read to express *existentially quantified* or *general* perceptual contents. Perceiving *that something is F and something is G* differs from perceiving *that something is both F and G*.

What if perception has *singular* or *particular* contents? Someone might object that *intermodal feature binding awareness does not show that not all perceptual experience is modality specific*. You might hear *that o is F* and see *that o is G*. This captures the identity of the individual heard and seen, but you could hear *that o is F* without seeing, and you could see *that o is G* without hearing. So, overall perceptual awareness may be just co-consciously seeing, hearing, and the rest. No parallel move exists for general perceptual contents.<sup>18</sup>

There is a good reply that helps illuminate the issue.

In principle, twin objects undetectably could be swapped. So, if *o* and *p* are distinct but you cannot by perceiving discern the difference in a way that enables you to tell which is which, you may not be able to detect the difference, in a way that enables you to tell which is which, between, for instance, seeing *that o is G* and seeing *that p is G*.

So, singular content theorists should accept:

(\*) Suppose *o* and *p* are distinct but perceptually indistinguishable in ways that would enable a subject to tell which is which. Controlling for other differences, hearing *that o is F* while seeing *that o is G* is introspectively indistinguishable in ways that would enable a subject to tell which is which from hearing *that o is F* while seeing *that p is G*.

But then the discernible difference between when features are perceptually experienced intermodally to be bound and when they are not cannot be explained by modality-specific singular contents alone. The singular content theorist in this respect has no advantage over the general content theorist.

Some singular content theorist might reject (\*) and try to capture the difference with differing modality-specific singular contents—for instance, by saying intermodal binding is perceptually apparent just in case singular contents overlap.

This is a bad idea. First, it requires accepting that any difference in visual singular content is introspectively discernible by a subject. Suppose *o* and *p* are distinct but perceptually indistinguishable to a subject. And suppose that, controlling for other differences, hearing *that o is F* while seeing *that o is G* must be introspectively distinguishable by the subject from hearing *that o is F* while seeing *that p is*

---

<sup>18</sup>Thanks to Jeff Speaks for pressing me to address this line of objection.

G. Since hearing *that o is F* is introspectively indistinguishable from hearing *that o is F*, seeing *that o is G* must be introspectively distinguishable from seeing *that p is G*. Second, it leaves no coherent way to explain illusions of identity and merely apparent distinctness in terms of modality-specific singular contents. What is the singular content of an episode of illusory intermodal binding awareness? It must be hearing *that o is F* and seeing *that o is G*, if apparent binding requires overlapping singular contents. But since the appearance is illusory, the singular contents cannot overlap, contrary to the proposal.

So, capturing the contrast and the illusions requires something further, such as the perceptual content *that o is F and G* (or, *that o is p*).

However, perceptual contents probably are not closed under conjunction. This is especially plausible for the singular content theorist who accepts (\*). Sharing a constituent of singular content (seeing and hearing the same individual) does not guarantee that a subject is able to recognize that what is seen and heard is the same individual. And it does not guarantee the subject perceives features to be coinstantiated or to be bound intermodally. Indeed, you *can* see and hear the same thing without its being perceptually apparent that something has both visible and audible features.<sup>19</sup> And, even if perceptual contents were closed under conjunction *within* a modality, different perceptual modalities are more plausibly viewed as distinct ways of entertaining contents, so it is far less plausible that contents from different modalities are closed under conjunction.

So, hearing *that o is F* while seeing *that o is G* does not guarantee perceiving *that o is F and G*. Perceiving *that o is F and G* requires a contentful perceptual episode that differs from just hearing *that o is F* while co-consciously seeing *that o is G*.

Given the failure of conjunctive closure for perceptual contents from different modalities, an episode of perceiving *that o is F and G* need not be factorable without remainder into modality-specific contentful perceptual episodes that could occur independently from each other. And it may have phenomenal features beyond those of a wholly auditory or a wholly visual experience under equivalent stimulation. Therefore, even if contents are singular, intermodal binding awareness shows that not all perceptual experience is modality specific.<sup>20</sup>

### *Binding and infusion*

O'Dea (2008) says that features perceived through one sense can appear bound *in a manner* that features perceived through different senses cannot. Intramodally bound features may appear to qualify or to be bound to *each other*, rather than just appearing to belong to a common object.

---

<sup>19</sup>Thus, singular content theorists do have one important advantage over general content theorists. Singular contents allow for overlapping modality-specific contents without perceptually apparent binding, as when a subject fails to perceptually appreciate the overlap.

<sup>20</sup>For helpful discussion of co-conscious phenomenal unity, content, and closure, see Bayne (2010, Chapter 3).

For example, to describe a visual experience of a red square as simply an experience of an object as red and as square is to miss out something crucial, namely that it is the redness that we are aware of that we are experiencing as square-shaped. It is not the case that we see an *object* which is square *and* which is red—it is the *squareness* which is red and the *redness* which is square. (O’Dea 2008, 302)

O’Dea describes the redness as *infusing* the squareness.

If perceptually apparent feature binding requires that one feature appears to *infuse* another, and if one feature may appear to infuse only another feature perceived with the same modality, then there is no intermodal feature binding awareness.

O’Dea does not offer this argument. He rejects that features may appear infused intermodally with other features, but he allows that features may appear intermodally to belong to a common object.<sup>21</sup> Thus, according to O’Dea, intermodal binding awareness does not require infusion. And I agree. Feature binding awareness requires only that features appear jointly to be features of some common item. Perceptibly bound features may even include distinct parts that perceptually appear to belong to a common whole.

This reply is not *ad hoc*, since (as O’Dea also allows) even perceptually apparent *intramodal* feature binding does not require infusion. Features such as *speckledness*, *hen-shapedness*, and *being wattled* do not all appear to infuse each other even though all could appear to qualify one body. And, as O’Dea suggests, apparent infusion may be asymmetric.

Moreover, the criteria for infusion are obscure. Why think perceptibly bound features *never* appear infused intermodally? The booming sound of the explosion might seem to infuse its bright flash. Why *can’t* the coolness seem to infuse the blueness of the sphere? The voice you hear might seem to infuse the visible mouth movements and articulatory gestures of the speaker—qualities of sound and visible motion thus may seem bound up with *each other*. The McGurk effect demonstrates that the apparent qualities of one regularly do depend on the other. And, even if I ceased to see it, it is difficult for me to imagine my perceptible interlocutor’s vocal activity “losing all of its visible properties without affecting its audible qualities,” which is one criterion for infusion O’Dea mentions (305).

Nevertheless, another observation may block intermodal infusion. Infusion involves a dependence of your awareness of features of one type upon your awareness of features of another type. Apparent infusion thus may require that, for each pair of apparently infused features (for instance, color and shape, or timbre and duration), if you ceased entirely to perceive any feature belonging to one of those types, you would cease to perceive any feature belonging to the other type. So, ceasing to see color would render an object’s shape invisible. Thus, what is diffi-

---

<sup>21</sup>O’Dea’s thesis is that Tye’s (2003) account of the co-conscious unity of perceptual experience cannot accommodate the difference between infusion and binding.

cult to imagine is losing *visual awareness* of something's color while leaving intact *visual awareness* of its shape.

Occasions of intermodal feature binding awareness involve perceiving thanks to more than one modality. You see an event's brightness and hear an event's loudness, even while perceiving something's being jointly bright and loud. And perceiving with one modality dissociates from perceiving with another. So, even if you ceased to see, you could continue to hear an event's loudness. More generally, it is not so difficult to imagine losing *visual awareness* of each of something's visible properties while leaving intact *auditory awareness* of each of its audible qualities. And, if there is no feature such that ceasing entirely to perceive any feature of that type with one modality would render an object's features imperceptible to another modality, then there is no apparent intermodal infusion.

Notice that feature binding awareness does not require infusion, thus understood. Not even all intramodally bound features satisfy this criterion for infusion. For instance, the features belonging to a face or the parts that make up a typed letter each perceptibly appear bound but are not, according to this criterion, infused. So, this does not rule out intermodal binding awareness. It may, however, rule out intermodal instances of one particularly intimate *variety* of binding awareness that occurs intramodally.<sup>22</sup>

Infusion may be a distinctive variety of feature binding, and it deserves further attention. In particular, a clear explication of infusion and its differences from other common forms of feature binding would be valuable, as would a study of whether intermodal infusion is possible. But apparent infusion is too restrictive as a requirement on feature binding awareness, and feature binding awareness suffices to establish the conclusions of §5.

### *Multimodal perceptual objects*

I maintain that intermodal feature binding awareness requires shared objects. However, given the diversity of objects across modalities, someone might object that intermodal binding awareness is less common than I have made it seem.

Consider this contrast. You can see and touch the baseball, and you can perceive its being jointly yellowed and rough. But, you see the baseball and hear *something else*: the sound it makes when it hits the bat. Thus, one might argue that no *common* object perceptibly seems to bear both audible and visible features—the sound is

---

<sup>22</sup>One might be tempted to think that apparent infusion involves only *integral*, in contrast to *separable*, feature dimensions (in a sense stemming from Garner 1970, 1974; see also Treisman and Gelade 1980; Treisman 1986, esp. 35-6–35-7). Hue, brightness, and saturation, for example, are examples of integral dimensions. This would help explain why awareness of features of one type requires awareness of features of another type. And perhaps intermodally there are no integral, as opposed to separable, feature dimensions. However, O'Dea's examples involve paradigm *separable* features, such as color and shape. More importantly, paradigm cases of feature binding involve separable rather than integral features.

loud; the ball is round and rough. So, in this case, no intermodal binding awareness occurs.

The objection succeeds only if the features do not perceptibly belong to something common. I maintain that in many such non-obvious cases we can admit common perceptible objects. My view is that perceptual objects in general are best understood as *mereologically complex individuals* that bear perceptible features. Perceiving something requires perceiving *some* of its parts or properties. However, it does not require perceiving *all* of them. In cases of intermodal feature binding, you may perceive the same mereologically complex individual while many of its parts and properties are perceptible to one but not both modalities.

Take the case of seeing and hearing. Physical objects such as baseballs and bats are visible. They participate in events such as collisions. Such events also are visible. When such events occur in a surrounding medium, they may involve sounds. And sounds are audible. But suppose the sound is a feature—a constituent part or a complex property—of such an event that occurs in a medium, rather than a *wholly* distinct individual. The sound is a feature of the collision that occurs in the medium, and the baseball and bat are participants in that collision. The audible sound and the visible rebounding of the ball from the bat thus each are perceptible features of the collision. Events like the collision of the baseball with the bat in a surrounding medium are audible because they include sounds as features. (This does not imply that you hear the collision mediately by or in virtue of hearing the sound.) You may hear the sound, and you may hear the collision of which it is a part or property. But you need not hear all of their features. For instance, you need not hear the baseball or the bat as such; you certainly do not typically hear their colors or their facing surfaces as such. And you may see the baseball, the bat, and thus the collision, but not their hidden parts or their sound. So, you can see and hear the collision of the ball with the bat thanks to its visible and audible features. When you consciously perceive its jointly having both visible and audible features at once, that is a case of intermodal binding awareness.

In order to determine the reach of intermodal binding awareness, this strategy must be assessed case by case.

## 7 Conclusions

My main claim is that there is intermodal feature binding awareness. Features—properties or parts—are consciously perceived to be coinstantiated or to belong to the same thing—to be bound—intermodally. The argument for this claim is that evidence from immediate perceptual belief, perception-guided action, experimental research, perceptual illusions, and perceptual phenomenology converges to support contrasting intermodal episodes of (1) and (2).

The important consequence is that not all perceptual awareness is modality specific. Some multimodal perceptual episodes require the kind of coordinated sen-

sitivity that enables identifying individuals across modalities. Some multimodal perceptual experiences are not just co-conscious episodes of seeing, hearing, touching, tasting, and smelling that could have occurred independently from each other. A closely related consequence is that not all perceptual phenomenal character is modality specific. The phenomenal character of a multimodal perceptual experience need not be exhausted by that which is associated with each of the modalities plus that which accrues thanks to mere co-consciousness.<sup>23</sup>

The significant upshot is that limiting inquiry to individual modalities of sensory perception and bare co-consciousness leaves out something critical. It leaves out richly multimodal forms of perceptual awareness, such as intermodal binding awareness. Therefore, no complete account of perceptual awareness or its phenomenal character can be formulated in modality-specific terms. Perceiving involves more than just co-consciously seeing, hearing, feeling, tasting, and smelling.<sup>24</sup>

## References

- Austin, J. L. (1962). *Sense and Sensibilia*. Oxford University Press, Oxford.
- Batty, C. (2011). Smelling lessons. *Philosophical Studies*, 153:161–174.
- Bayne, T. (2010). *The Unity of Consciousness*. Oxford University Press, Oxford.
- 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 141–177. Oxford University Press.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge, MA.
- Bushara, K. O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K., and Hallett, M. (2003). Neural correlates of cross-modal binding. *Nature Neuroscience*, 6(2):190–195.
- Cinèl, C., Humphreys, G. W., and Poli, R. (2002). Cross-modal illusory conjunctions between vision and touch. *Journal of Experimental Psychology: Human Perception and Performance*, 28(5):1243.

---

<sup>23</sup>This type of argument need not be limited to apparent feature binding. There may be intermodally perceptible relations, such as motion, synchrony, rhythm, and causality. See O’Callaghan (2014).

<sup>24</sup>For valuable questions, conversations, and correspondence that helped me to improve this paper, many thanks to Mike Barkasi, Tim Bayne, David Bennett, Justin Broackes, David Chalmers, Kevin Connolly, Ophelia Deroy, James Genone, Richard Grandy, Fiona Macpherson, Michael Martin, Mohan Matthen, Jesse Prinz, Barry C. Smith, Jeff Speaks, Charles Spence, Dustin Stokes, and audience members at the University of Geneva and the 2013 Central APA symposium on multimodal perception.

- Clark, A. (2000). *A Theory of Sentience*. Oxford University Press, New York.
- Connolly, K. (2014). Making sense of multiple senses. In Brown, R., editor, *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*, volume 6 of *Studies in Brain and Mind*, chapter 24. Springer.
- Fulkerson, M. (2011). The unity of haptic touch. *Philosophical Psychology*, 24(4):493–516.
- Garner, W. R. (1970). The stimulus in information processing. *American Psychologist*, 25(4):350–358.
- Garner, W. R. (1974). *The processing of information and structure*. Erlbaum.
- Grice, H. P. (1962). Some remarks about the senses. In Butler, R. J., editor, *Analytical Philosophy, Series 1*. Blackwell, Oxford.
- Hamilton, R. H., Shenton, J. T., and Coslett, H. B. (2006). An acquired deficit of audiovisual speech processing. *Brain and Language*, 98(1):66–73.
- Hill, C. S. (2009). *Consciousness*. Cambridge University Press, Cambridge.
- Johnston, M. (2006). Better than mere knowledge? The function of sensory awareness. In Gendler, T. S. and Hawthorne, J., editors, *Perceptual Experience*, chapter 7, pages 260–290. Clarendon Press, Oxford.
- Jordan, K. E., Clark, K., and Mitroff, S. R. (2010). See an object, hear an object file: Object correspondence transcends sensory modality. *Visual Cognition*, 18(4):492–503.
- Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2):175–219.
- Kubovy, M. and Schutz, M. (2010). Audio-visual objects. *Review of Philosophy and Psychology*, 1(1):41–61.
- Lopes, D. M. M. (2000). What is it like to see with your ears? The representational theory of mind. *Philosophy and Phenomenological Research*, 60(2):439–453.
- Matthen, M. (2005). *Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception*. Oxford University Press, Oxford.
- Mitroff, S. R., Scholl, B. J., and Wynn, K. (2005). The relationship between object files and conscious perception. *Cognition*, 96(1):67–92.
- Navarra, J., Yeung, H. H., Werker, J. F., and Soto-Faraco, S. (2012). Multisensory interactions in speech perception. In Stein, B. E., editor, *The New Handbook of Multisensory Processing*, chapter 24, pages 435–452. MIT Press, Cambridge, MA.

- O'Callaghan, C. (2012). Perception and multimodality. In Margolis, E., Samuels, R., and Stich, S., editors, *Oxford Handbook of Philosophy of Cognitive Science*, pages 92–117. Oxford University Press, Oxford.
- O'Callaghan, C. (2014). Not all perceptual experience is modality specific. In Stokes, D., Biggs, S., and Matthen, M., editors, *Perception and Its Modalities*. Oxford University Press, Oxford.
- O'Dea, J. (2008). Transparency and the unity of experience. In Wright, E., editor, *The Case for Qualia*, chapter 16, pages 299–308. MIT Press, Cambridge, MA.
- Pasalar, S., Ro, T., and Beauchamp, M. S. (2010). TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience*, 31(10):1783–1790.
- Peacocke, C. (1983). *Sense and Content*. Oxford University Press, Oxford.
- 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(6):1329–1333.
- Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science and Technology*, 28(2):61–70.
- Spence, C. and Bayne, T. (2014). Is consciousness multisensory? In Biggs, S., Stokes, D., and Matthen, M., editors, *Perception and Its Modalities*. Oxford University Press, Oxford.
- Spence, C. and Driver, J., editors (2004). *Crossmodal Space and Crossmodal Attention*. Oxford University Press, Oxford.
- Stein, B. E. (2012). *The New Handbook of Multisensory Processing*. MIT Press, Cambridge, MA.
- Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Alex Meredith, M., Perreault, T. J., Ramachandran, R., Röder, B., Rowland, B. A., Sathian, K., et al. (2010). Semantic confusion regarding the development of multisensory integration: a practical solution. *European Journal of Neuroscience*, 31(10):1713–1720.
- Stein, B. E. and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4):255–266.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2):194–214.

- Treisman, A. (1986). Properties, parts and objects. In Boff, K., Kaufman, L., and Thomas, J., editors, *Handbook of Perception and Human Performance*, volume 2, chapter 35, pages 1–70. Wiley.
- Treisman, A. (1988). Features and objects. *The Quarterly Journal of Experimental Psychology*, 40A(2):201–237.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6(2):171–178.
- Treisman, A. (2003). Consciousness and perceptual binding. In Cleeremans, A., editor, *The Unity of Consciousness: Binding, Integration, and Dissociation*, pages 95–113. Oxford University Press, Oxford.
- Treisman, A. and Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14:107–141.
- Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136.
- 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.
- Vatakis, A. and Spence, C. (2007). Crossmodal binding: Evaluating the “unity assumption” using audiovisual speech stimuli. *Perception and Psychophysics*, 69(5):744–756.
- Welch, R. B. and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3):638–667.
- Zmigrod, S. and Hommel, B. (2011). The relationship between feature binding and consciousness: Evidence from asynchronous multi-modal stimuli. *Consciousness and Cognition*, 20(3):586–593.
- Zmigrod, S., Spapé, M., and Hommel, B. (2009). Intermodal event files: integrating features across vision, audition, taction, and action. *Psychological Research*, 73(5):674–684.