

Synthese
<https://doi.org/10.1007/s11229-019-02156-2>

S.I. : BETWEEN VISION AND ACTION



Bodily awareness and novel multisensory features

Robert Eamon Briscoe^{1,2} 

Received: 1 February 2018 / Accepted: 21 February 2019
© Springer Nature B.V. 2019

Abstract

According to the decomposition thesis, perceptual experiences resolve without remainder into their different modality-specific components. Contrary to this view, I argue that certain cases of multisensory integration give rise to experiences representing features of a *novel type*. Through the coordinated use of bodily awareness—understood here as encompassing both proprioception and kinaesthesia—and the exteroceptive sensory modalities, one becomes perceptually responsive to spatial features whose instances couldn't be represented by any of the contributing modalities functioning in isolation. I develop an argument for this conclusion focusing on two cases: 3D shape perception in haptic touch and experiencing an object's egocentric location in crossmodally accessible, environmental space.

Keywords Multisensory perception · Proprioception · Kinaesthesia · Egocentric space · Haptic touch · Crossmodal perception

1 Introduction

The term “multisensory integration”, I have elsewhere suggested (Briscoe 2016, 2017), is used in both philosophical and scientific literature to refer to two computationally distinct processes: optimizing multisensory integration (*O-integration* for short) and what I shall refer to as non-optimizing or “generative” multisensory integration (*G-integration* for short). In what follows, I begin by examining the distinction between these two types of multisensory integration (Sects. 2, 3). Cases of G-integration, I'll then argue, have the potential to provide a strong challenge to a

✉ Robert Eamon Briscoe
rbriscoe@gmail.com

¹ Department of Philosophy, Ohio University, 202 Ellis Hall, Athens, OH 45701, USA

² Department of Philosophy, Centre for the Study of Perceptual Experience, University of Glasgow, School of Humanities, 67-69 Oakfield Avenue, Glasgow G12 8QQ, UK

long-standing view in the philosophy of perception that Tim Bayne (2014) dubs the “decomposition thesis”. According to this view, perceptual experiences resolve without remainder into their modality-specific components. Consider, for example, your overall experience of ringing a doorbell: you see the way your hand is moving; you hear the chime; and you feel, among other things, the pressure applied by the tip of your finger. If the decomposition thesis is correct, then this is essentially the whole story. Your experience can be exhaustively factored into episodes of seeing, hearing, touching, that in the relevant instance happen to be co-conscious, but that might in other circumstances have occurred without the others. The stream of perceptual consciousness contains parallel, but separate tracks for each of its modalities.

Bayne doesn’t equip us with a definition of *modality-specificity*, but discussions by Casey O’Callaghan (2014, 2015, 2017) are helpful. According to O’Callaghan, the thesis that all perceptual experience is modality-specific is best understood as the claim that the phenomenal character of any perceptual episode “is exhausted by that which, for each of its respective modalities, could be the phenomenal character of a corresponding mere experience of that modality” plus whatever results from simple co-consciousness (2014, p. 151). A *mere* experience of a modality, in turn, is defined as one that “allows prior perceptual experiences of other modalities, but requires while it occurs that its subject’s overall perceptual experience remain wholly or solely of one modality” (2014, p. 151).

Contrary to the decomposition thesis, I will argue that certain cases of G-integration give rise to experiences representing features of a *novel type*—features that couldn’t be instantiated by any mere experience of a contributing modality. O’Callaghan has recently made a case for this conclusion with respect to flavor perception (Sect. 4). No mere experience of taste, retronasal olfaction, touch, or any other modality contributing to the perception of flavor, he suggests, is, for example, an experience of *mintiness*. “There is a distinctive, recognizable, and novel quality of mint... that is consciously perceptible only thanks to the joint work of several sensory systems” (O’Callaghan 2017, p. 174). I’m going to argue for an analogous conclusion in this contribution, focusing on two additional examples. I will argue, first, that haptic touch involves a form of G-integration because neither cutaneous touch, nor proprioception, nor kinaesthesia operating by itself is a potential source of perceptual information about the 3D shapes of objects external to the subject’s body (Sect. 5). And I will argue, second, that no mere experience in any modality is an experience of an object’s egocentric location in crossmodally accessible, environmental space (Sect. 6).¹

¹ My arguments for these conclusions are intended to remain neutral with respect to foundational debates in the epistemology and metaphysics of perception, including the debate between externalism and internalism and between representationalism and relationalism/naïve realism. Thanks to an anonymous referee for requesting me to be clear on this point.

2 Optimizing multisensory integration

The initial estimates of a property produced by different modalities may sometimes conflict with one another. Vision and touch, for example, might produce conflicting initial estimates of an object's 3D shape or orientation. Alternatively, vision and proprioception, might produce different initial estimates of the location at which a part of the body is located (Stratton 1899; Harris 1965; Botvinick and Cohen 1998). Such intersensory discrepancies can arise for a variety reasons. Initial estimates, for example, may conflict due to the noisiness of neural computation (for discussion, see Knill and Pouget 2004). Or one modality might simply have a finer spatial or temporal grain when it comes to discriminating a type of feature. Vision is generally better at answering *where* questions than audition, while audition is generally better at answering *when* questions than vision.

According to a recently influential Bayesian approach to multisensory perception in cognitive science, initial estimates of a property provided by different modalities are weighted by their relative reliability and combined in a way that optimizes, i.e., reduces the variance in, the final perceptual estimate of that property.² Since this final estimate is a compromise between the different initial estimates, such optimizing multisensory integration (O-integration) also serves to reduce intersensory discrepancy at the level of conscious perception.

Numerous illusions can be explained as perceptual consequences of O-integration. In the *ventriloquism effect* (Bertelson 1999), for example, initial visual and auditory estimates of an object's direction are at odds. Typically, auditory localization of the object is strongly biased toward the initial visual estimate, and when initial estimates are not highly discrepant, the upshot is that the sound you hear non-veridically appears to be coming from the object you see. This effect is sometimes referred to as “phenomenal fusion” (Radeau and Bertelson 1977).

Other well-known illusions exemplifying processes of O-integration include:

The McGurk effect: In this illusion (McGurk and MacDonald 1976), inputs from vision influence the contents of auditory experience. When subjects watch a video of a speaker articulating the sound /ga/ dubbed with a recording of a speaker pronouncing the sound /ba/, they report hearing the sound /da/ instead (a kind of phonological compromise).

Visual capture of touch and proprioception: When initial visual and haptic estimates of an object's spatial properties, e.g., its shape, size, or orientation, are experimentally set in conflict, the final, conscious haptic estimate is strongly biased in the direction of the initial visual estimate (Gibson 1933; Rock and Harris 1967). Such visual dominance is also found when visual and proprioceptive estimates of the position of a body part are made to conflict (Hay et al. 1965; Welch and Warren 1980; Botvinick and Cohen 1998; Samad et al. 2015).

² See Rohde et al. (2016) and the essays collected in Trommershäuser et al. (2011) for useful overviews.

The parchment skin illusion: To elicit this illusion, experimenters recorded the sounds produced while subjects rubbed their palms together (Jousmäki and Hari 1998). These sounds were played back to the participants through headphones, “dubbing” the tactile stimulation they received. When high frequencies were accentuated, participants reported that their skin felt dry and paper-like. Jousmaki and Hari propose that this illusion reflects an “omnipresent intersensory integration phenomenon, which helps the subject to make accurate tactile decisions about the roughness and stiffness of different textures they manipulate” (R190).

For present purposes, the important point is that in O-integration multiple modalities provide distinct and potentially conflicting, initial estimates of a *single property*—e.g., 3D shape, orientation, or texture—that they jointly attribute to a perceived object or event. The end-product of O-integration is a revised and, when all goes well, optimized estimate of that property. Hence, the total number of property types represented by different modalities remains the same after O-integration has taken place.

This point can be brought out by reflecting on the role of O-integration in the ventriloquism effect. In the ventriloquism effect, the apparent location of an auditory event can be strongly biased in the direction of a simultaneous visual event. Although new auditory information is thus produced by interaction with the visual system, it is clearly information that the auditory system could have produced on its own in different circumstances (Macpherson 2011b). Analogously, in the McGurk effect, neither vision nor audition by itself represents the speaker as pronouncing the sound /*da*/, but /*da*/ is obviously a sound that could have been represented by means of audition alone under other conditions.

3 Generative multisensory integration

In ventriloquism, the McGurk effect, and other cases of O-integration, the information produced by multisensory interaction is new only in the sense that it is an optimized revision of the initial estimate provided by one or more of the contributing modalities. Combining estimates of different types of properties across distinct modalities, some philosophers have recently argued, however, may give rise to the representation of a genuinely novel or “emergent” property, one that couldn’t be represented by any of the contributing modalities functioning in isolation. In what follows, I’ll refer to this alternative form of multisensory processing as “generative” multisensory integration (or *G-integration* for short).

Before proceeding, it is necessary to clarify two different ways in which G-integration could support the representation of a novel type of feature. First, a feature *F* could be novel, but only relative to the representational powers of the specific modalities that contribute to the relevant G-integration process. Here, although no mere experience of any of the contributing modalities could be an experience as of something *F*, some other, non-contributing modality is capable of representing feature *F* on its own. In this type of case, we can say that *F* is a feature of a *weakly*

novel multisensory type. 3D shape, I shall argue in Sect. 5, is an example of such a feature. 3D shape is novel relative to the representational powers of the modalities that contribute to G-integration in haptic touch (cutaneous touch, proprioception, and kinaesthesia), but familiar relative to the representational powers of vision.

Alternatively, and more dramatically, G-integrating information from different modalities may result in the representation of a feature F that isn't ever perceptible unimodally. If so, then no instance of F could be represented outside the context of the relevant G-integration process. To use O'Callaghan's language, no mere experience of any modality could be an experience as of something F . In this type of case, we can say that F is a feature of a *strongly novel* multisensory type—a feature only revealed through the coordinated use of different senses. Location in crossmodally accessible, egocentric space, I shall argue in Sect. 6, is an example of such a feature. Location in egocentric space is novel relative to the representational powers of *any* modality working by itself.

It is also important before proceeding to distinguish the claim that G-integration can support the representation of weakly novel types of features from the claim that certain relational features have *instances* that are only perceptible multisensorily. O'Callaghan (2014, 2015, 2017) discusses cases involving intermodal feature binding, causation, timing, and meter perception. Importantly, in each of these cases, the relevant relational feature is independently perceptible by each of the contributing modalities. Consider the case of intermodal meter perception. A study by Huang et al. (2012) found that auditory and tactile sequences were coherently grouped by musically trained subjects performing a meter recognition task. Meter, however, can be perceived by means of either audition or touch alone. Meter isn't novel relative to either of the modalities that contribute to whichever multisensory process is responsible for audio-tactile meter perception. In contrast, a type of feature F is weakly novel just in case it has some instances that are perceptible by a single modality, but is not independently perceptible by any of the modalities that contribute to the relevant G-integration process. F is novel relative to the representational powers of those modalities even though F is familiar from other unimodal contexts. The cases discussed in the next three sections will hopefully help to clarify this point.

4 Flavor perception

Consider, first, the case of flavor perception. Flavor properties aren't detected by any single set of sensory receptors functioning in isolation: they depend instead on the combination of inputs from taste and retronasal olfaction; thermal and somatosensory cues; as well as sources of information concerning chemical irritation and nociception (for discussion see, Auvray and Spence 2008; Spence et al. 2014; Smith 2015).

There are at least three influential accounts of flavor perception on the market. First, flavors could be strongly novel phenomenal features that are perceptible only through the coordinated use of different perceptual modalities. On this view, as Smith writes, “we have a category of perceptual quality for which Aristotle's classification made no room. Flavours are not common sensibles accessible by more

than one sense; we need many senses—chemical and contact—working together to produce flavour perceptions” (2015, p. 320). That is just one possibility. Alternatively, apparent flavors might be better understood as fragmenting into non-structured collections of gustatory, olfactory, and tactual properties that are referred to a single substance in the mouth. On this view, flavor perception doesn’t challenge the decomposition thesis. Apparent flavors resolve into a number of unimodally accessible components. Yet a third possibility is that flavor perception isn’t, strictly speaking, multisensory at all. Instead, what we ordinarily refer to as the “sense of taste” may just be the flavor system, where the latter utilizes a variety of dedicated transducer types beyond those found on the tongue. On this view, flavors properties are tracked by a single perceptual modality (Auvray and Spence 2008; Matthen 2015).

Casey O’Callaghan (2014, 2015, 2017) makes a persuasive case for the first possibility: flavor perception, he argues, involves multisensory awareness of a strongly novel type of feature, one whose instances are only perceptible through the coordinated operation of distinct senses. “Flavor experiences”, he suggests, “may have *entirely novel* phenomenal features of a type—even a qualitative type—that no unimodal experience could instantiate and that do not accrue thanks to simple co-consciousness” (2015, pp. 567–68, emphasis added). I’m not going to evaluate the considerations that O’Callaghan adduces in support of this view. Instead, I’ll argue in the next two sections that G-integrating the contents of bodily awareness with information derived from exteroceptive modalities may make novel spatial features accessible to perceptual consciousness. More specifically, I want to defend two proposals.

The first is that experiencing an object’s 3D shape (solid figure) by means of haptic touch involves the G-integration of cutaneous information delivered by receptors embedded in the subject’s skin with proprioceptive-kinaesthetic information about way her fingers and hand are moving. Haptic touch is a form of G-integration, I propose, because neither cutaneous touch, nor proprioception, nor kinaesthesia operating by itself is a potential source of perceptual information about the 3D shapes of nonbodily objects. Since 3D shape is a feature some of whose instances are perceptible by vision alone, however, 3D shape is a *weakly* novel type of multisensory feature: it is novel only relative to the representational powers of cutaneous touch, proprioception, and kinaesthesia.³

The second proposal is that no mere experience of any modality is an experience of an object’s egocentrically-specified location—that is, its subject-relative distance

³ An anonymous referee expresses skepticism that “vision on its own suffices for the perception of voluminous, 3D shape.” I don’t think that such skepticism is well-motivated. It is true that multiple fixations (glances) may be required in some cases to reveal an object’s overt 3D structure and, hence, that proprioceptive-kinaesthetic information about eye movements may sometimes play a role in shape processing. It isn’t possible to generalize from this, however, to the conclusion that the perception of 3D shape *always* depends on proprioceptive-kinaesthetic information. Indeed, an opaque object’s shape may be completely revealed, I take it, to a passive observer under certain conditions. (Imagine, for example, viewing a cube rotating on an axis passing through the cube’s center and two of its vertices.) There is also a large body of psychophysical evidence that stimulus-driven mechanisms of amodal completion generate representations of self-occluded, 3D object structure (Tse 1998, 1999, 2017). For philosophically oriented discussions, see Briscoe (2008, 2011, 2018).

and direction—in crossmodally accessible, 3D space. More precisely, neither vision, nor audition, nor touch, in the absence of contributions from bodily awareness, is capable of representing an object's location in a subject-relative, but non-modality-specific spatial frame of reference. Since egocentrically-specified location is novel relative to the representational powers of any modality operating by itself, egocentrically-specified location is a *strongly novel* type of multisensory feature.

Before proceeding to the next section, a few remarks are in order. First, I should note that I use the term “bodily awareness” here in a way that encompasses both *proprioception* (the sense of how the parts of one's own body are spatially configured) and *kinaesthesia* (the sense of bodily movement). Second, I assume that proprioception and kinaesthesia either independently qualify as perceptual modalities or together constitute a distinct perceptual modality. There is significant debate about the proper criteria for individuating the senses, but candidate perceptual modalities Ritchie and Carruthers (2015) plausibly suggest can be evaluated by the degree to which they have the following features:

1. Sensitivity to a set or range of related physical properties.
2. Detector mechanisms that transduce these physical properties into an informational signal to the central nervous system.
3. The evolved function of detecting and representing the physical properties specified by (1).
4. Representations with nonconceptual content and a mind-to-world direction of fit.
5. Utilization of the informational signals by the organism (typically through integration with other sensory inputs) to guide intentional action.

Proprioception and kinaesthesia plausibly embody all five of these features and, so, constitute fully credentialed perceptual modalities (Fridland 2011; Schwenkler 2013; de Vignemont 2016). A piece of further support for this view is that there exist cases of accurate representation, illusion, and hallucination for proprioception and kinaesthesia, just as for the traditional Aristotelian senses (Macpherson 2011a).⁴ Some authors subsume proprioception under kinaesthesia (for example, Bastian 1888, who coined the latter term, and Donaldson 2000), while others subsume kinaesthesia under proprioception (Roll et al. 1991; Proske and Gandevia 2012; Gallace and Spence 2014; Ritchie and Carruthers 2015). For my purposes here, however, it doesn't matter whether proprioception or kinaesthesia is understood as the more encompassing form of bodily awareness or whether they should instead be thought of as distinct senses (McCloskey 1973). What matters is that bodily awareness is a source of perceptual information that can contribute to both types of multisensory integration discussed here.

⁴ Helpful reviews of experimental findings include Roll et al. (1991), Lackner and DiZio (2005), and Taylor et al. (2017).

5 Haptic touch

Haptic touch—also known as active or exploratory touch—is a promising case of G-integration. In order to perceive the 3D shape (solid figure) of an object by means of haptic touch it is necessary to combine cutaneous information delivered by receptors embedded in your skin with kinaesthetic and proprioceptive information about specific way your fingers and hand are moving. The same cutaneous information when combined with substantially different kinaesthetic and proprioceptive information would result in perceiving the object as filling out space in a different way. Haptic touch is a form of multisensory, G-integration, I propose, because neither cutaneous touch nor bodily awareness operating by itself is a potential source of perceptual information about the 3D shapes of objects external to the body. Since 3D shape is perceptible by (normal, fully developed) vision operating on its own, however, 3D shape is a *weakly novel* type of multisensory feature. It is only novel relative to the representational powers of cutaneous touch and bodily awareness. To use O’Callaghan’s terminology, no mere experience of either cutaneous touch, or kinaesthesia, or proprioception could be an experience of a nonbodily object as, for example, cuboid, or pyramidal, or teapot-shaped.

This assessment is in keeping with what I take to be the establishment view of haptic touch in cognitive science. According to this view, haptic touch, unlike passive, cutaneous touch, is not a single, unified, perceptual modality:

It is important to emphasize that active haptic sensing per se constitutes a legitimate form of multimodal perception inasmuch as this form of sensing uses both tactile and kinesthetic sensory inputs (Jones and Lederman 2006, p. 76; quoted by Fulkerson 2014a, b, p. 18).

...[T]actile sensations can be combined with proprioceptive information about body configuration to produce a spatial, volumetric description of a tactile object.... This may simply be an instance of the general bottom-up process of combining sources of information through multisensory perception (Serino and Haggard 2010, p. 233).

Given that what is commonly referred to as the sense of touch actually appears to encompass several distinct sensory systems..., it should perhaps be considered more of a multisensory (i.e., derived from the integration of different sensory inputs) rather than a single sensory modality (Gallace and Spence 2014, p. 5).

The establishment view of haptic touch has been recently challenged by Matthew Fulkerson (2014a, b). According to Fulkerson (2014a, b), haptic touch is a single, unified sensory modality. In keeping with his *sensory pluralism*—the view that there are different explanatory dimensions along which a perceptual experience or system can be categorized as either unisensory or multisensory—this claim is relativized to a specific criterion for multisensory awareness:

Multisensory experiences do not involve the direct predication of individual features onto perceptual objects. Instead, there is an [associative relation]

between already bound or grouped sets of features.... The senses provide structured information about object features, and the diversity of multisensory interactions serve to coordinate that information (Fulkerson 2014a, b, p. 36).

In what follows, I shall refer to this necessary condition for multisensory experience as the *association criterion*. The association criterion allows for many different types of multisensory interaction and coordination. An associative relation, Fulkerson writes, “implies some connection between experiences, without specifying the nature of this connection.... In virtue of such associations, experiences may suppress one another, or enhance one another, or one may dominate the other, etc., and there is no reason to think that any single mechanism is involved in every case of multisensory involvement” (2011, p. 494).

Fulkerson argues that perceiving by haptic touch doesn’t involve an associative relation between already bound or grouped sets of perceptual features and, hence, that it ought not to be categorized as multisensory in nature. He also defends this verdict on independent, phenomenological grounds. In what follows, I shall make the case for an opposing view. In particular, I’ll argue that haptic touch can be properly characterized as a case of G-integration.

Fulkerson’s main argument against the establishment view is based on the idea that haptic touch doesn’t satisfy the association criterion. “Touch, like vision and other major senses”, he writes, “does not involve any associative relation. It involves the direct predication or binding of sensory features onto individual objects” (2014a, b, p. 40). Fulkerson takes vision, for example, to have the following existential, merely property-involving structure:

There is a visual object O, and O has certain features.

The ventriloquism effect, by contrast, would qualify as an example of a genuinely multisensory experience, on Fulkerson’s account, because it can be modeled as involving an associative relation between two distinct, existentially quantified assignments of features to objects (2014a, b, p. 35):

[There is a visible object O, and O has certain features] → [There is a sound S, and S has certain features].

The arrow here indicates that the relevant associative relation between vision and audition leads to an alteration in the content of the auditory experience (in particular, a shift in the apparent direction of the sound).

In contrast with the ventriloquism effect and other cases of O-integration, haptic touch, Fulkerson proposes, has a non-associative, feature-predicating structure:

There is a haptic object O, and O has certain features.

I don't think that this analysis is correct. To perceive a nonbodily thing's 3D shape by means of haptic touch it is typically necessary to represent the changing joint configuration of your hand. The relevant proprioceptive-kinaesthetic representations, it seems clear, directly assign certain spatial features to an object, namely, a part of the body. They have their own feature-predicating structure.⁵ In addition, it is also necessary to form cutaneous tactual representations in response to events that are taking place on the surface of your hand, for example, changes in pressure or temperature. These tactual representations also directly assign certain features to a perceived object, whether we identify the latter with the relevant part of your body, or the distal thing you are touching, or, as on some accounts, with both.⁶

Jones and Lederman (2006) review evidence concerning the range of properties attributed by cutaneous touch to its objects. Their list includes pressure, temperature, vibration, motion, texture, compliance, orientation, size, curvature, and weight. Importantly, experiences of cutaneous touch also assign certain *somatotopic locations* to the objects/events that elicit them. Afferent projections from cutaneous receptors to primary somatosensory cortex (SI), as Serino and Haggard write, "preserve the spatial organization of receptors in the skin, so that SI contains an essentially spatial representation of the physical body surface. That is, the physical body structures tactile sensation, because the receptor surface covers the physical body" (Serino and Haggard 2010, p. 225). Finally, there is a growing body of evidence that passive, cutaneous touch represents objects and events in accord with Gestalt grouping principles such as *proximity*, *similarity*, *common fate*, *good continuation*, and *closure* (for a review, see Gallace and Spence 2011). For example, tactual elements placed closer together are more likely to be grouped together (Chang et al. 2007), and subjects perceptually complete the gaps between two separate sequentially-presented vibrotactile stimuli when these gaps are filled with "tactile noise" (Kitagawa et al. 2009). These empirical results provide further support for the view that the contents of cutaneous touch possess genuine, feature-predicating structure.

According to Fulkerson, cases of multisensory experience involve some sort of associative relation between two or more perceptual states, each of which has its own feature-predicating structure. But this, I have suggested, is just what we find in haptic touch: an association between feature-predication in bodily awareness, on the one hand, and feature-predication in cutaneous touch, on the other.

Now, one possible response to what I've said so far is that although subpersonal haptic processing may depend on an associative relation between distinct, unimodal perceptual states with feature-predicating structure—and, so, may qualify as multisensory relative to the association criterion—conscious haptic experience may nonetheless display a unisensory phenomenal character. The unity of haptic touch may be

⁵ I'm not going to take a stand here on whether that structure is merely property-involving, as Fulkerson suggests, or instead involves elements that function to pick out environment particulars, e.g., "singular applications" in the sense of Burge (2010).

⁶ de Vignemont (2014), following Weber, defend a pressure theory of cutaneous touch. On their account, cutaneous touch is the direct perception of a symmetrical relation in which one feels pressure between a part of the body and some external object. Hence, cutaneous touch typically assigns features both to the object exerting pressing on a region *R* of the bodily surface and to *R* at the same time.

defensible, in other words, relative to a *phenomenological* criterion for unisensory experience. Fulkerson suggests just this: “Haptic touch... does not seem, at least to introspection, to involve an association between separate experiences (with their own perceptual objects)... Instead, one has a unified experience with different constituent elements...” (2014a, b, p. 40).

Consistent with this assessment, Fulkerson maintains that perceptual awareness of distal object features normally excludes simultaneous awareness of the state of the subject’s body in episodes of haptic touch. “Our typical bodily awareness during touch is *entirely transparent*, hidden in the background while we experience external and even distal objects and properties” (2014a, b, p. 103, my emphasis). Consciously accessing bodily information in experiences of haptic touch, he suggests, requires a perceptual switch akin to the one that occurs when we alternate between different aspects of the Necker Cube.

The analogy with the Necker Cube doesn’t seem introspectively strong, however. Contrary to Fulkerson, when we are engaged in haptically exploring an object, we are typically aware not only of the object’s own properties, but also of *which* part of our body is in contact with the object and, at least to a limited extent, *how* that body part is moving. Consider the experience of exploring the shape of a dinner plate using the tip of your index finger. Perceptual awareness of the plate as *circular* does not seem to exclude perceptual awareness of the way your hand and arm are moving as your finger traces over the plate’s rim (for example, whether your hand is moving in a clockwise or counter-clockwise direction). Similarly, awareness of a surface’s texture in haptic touch does not seem to exclude awareness of how much pressure you are exerting. Lightly running the tip of your index finger over a wooden desktop results in one experience of exploratory touch. Forcefully moving your finger over the desktop results in another saliently different experience. Both experiences reveal the surface as *smooth*, but the phenomenological contrast between the two experiences seems well explained by the difference in the accompanying feelings of pressure (de Vignemont and Massin 2015; see Katz 1925/1989, p. 51).

Considerations like these have motivated many philosophers to maintain that haptic touch has a “dual” or “bipolar” phenomenal character (Katz 1925/1989; Gibson 1962; Martin 1992; O’Shaughnessy 1995; Bermúdez 1998; Scott 2001; Brogaard 2012; de Vignemont and Massin 2015). Bermúdez writes:

The best description of the phenomenology of touch is that tactile experience is always both exteroceptive and proprioceptive. Attention can be directed either proprioceptively or exteroceptively, and it can be shifted from one to the other, but this should be viewed as an alteration of the balance between focal and peripheral awareness. When attention is directed exteroceptively toward the spatial properties of an object, the perceiver remains peripherally aware of the spatial properties of the relevant limb, and vice versa (1998, pp. 138–39).

A phenomenologically adequate account of the relation between object-directed and body-directed experience in haptic touch should not be modeled on the switch between different aspects of the Necker Cube, but rather, as Bermúdez suggests, on the distinction between focal and peripheral perceptual awareness. That we are in most cases *more* attentive to nonbodily object features when we are engaged in

haptic exploration, if this is right, doesn't mean that bodily awareness is entirely transparent or "hidden in the background". Experiences of haptic touch aren't, in general, phenomenologically unified in the manner suggested by Fulkerson.

Summing up, a strong case can be made for the view that haptic touch qualifies as multisensory both relative to the association criterion and on phenomenological grounds. Furthermore, haptic touch is plausibly a form of what I am calling G-integration because no mere experience of either cutaneous touch or bodily awareness is a potential source of perceptual information about a nonbodily object's 3D shape. 3D shape is a weakly novel type of multisensory feature.

An anonymous reviewer has objected that there is at least one instance of 3D shape that can be perceived by means of proprioception alone, namely, the 3D shape of the subject's own body (or the shapes of her bodily parts). So, 3D shape is not weakly novel, as I claim it is, relative to the representational powers of bodily awareness and cutaneous touch.

This is an interesting objection, but it is premised on the assumption that proprioception, like vision, is a self-standing source of perceptual information about the body's relatively constant structure and metric properties. This assumption has been disputed by a long line of philosophers (O'Shaughnessy 1980, 1995; Gallagher 2003; de Vignemont 2014, 2018; Wong 2017) and cognitive scientists (Longo and Haggard 2010; Longo 2015). Gallagher, for example, writes: "although proprioception provides a sense of posture and limb position, it does not deliver a sense of body shape. If I hold my arm out to the side of my body, I can tell this posture by proprioception, but proprioception is not such to tell me that my arm is of cylindrical shape rather than rectangular. I know this from other senses" (2003, p. 61). A similar point is made by de Vignemont, in a passage worth quoting at length:

[A] problem with the body senses is that they do not directly carry information about the shape of the various parts of the body, their size, and their spatial configuration. For example, the facts that we have two arms, that they are cylinder-shaped, that they are of a certain length, and that they are connected to the torso on one end and to the hands on the other end, cannot be easily derived from the body senses.... Active exploration of each body part by haptic touch seems to fare better and to be more specific. However, this involves complex tactile-proprioceptive processing, and that in turn requires taking into account the size of the exploratory body parts (e.g. fingers). Hence, the scope of information that the body senses directly carry is limited. Yet, information such as body metrics is needed in order to locate body parts in space. For instance, the arms can be crossed or not with the very same joint angles depending on their size and on the width of the shoulders (2014, p. 991).

de Vignemont argues that joint angle information provided by our muscle spindles, joint receptors, and Golgi tendon organs is normally supplemented and structured by a previously constructed representation of the configuration and metrics of the body's parts—a representation variously referred to in the philosophical and scientific literature as the "long-term body image" (O'Shaughnessy 1980), the "implicit body model" (Longo and Haggard 2010; Tsakiris 2010; Longo 2015; Wong 2017; Tamè et al. 2019), and the "body map" (de Vignemont 2014, 2018).

This stored, relatively stable representation of the subject's anatomy enriches and imposes spatial order on information provided by proprioceptive receptor systems (see O'Shaughnessy 1980, chapters 7–8). In particular, it provides information about bodily metrics needed for purposes of locating bodily parts in space and performing object-directed actions. The important point, relative to the reviewer's objection, is that proprioceptive receptor systems do not on their own provide information about the 3D shapes of the body's parts. Proprioception does not take the measure of bodily shape and size, but rather is structured by a stored representation or model of the body's dimensions.⁷

I'll turn now to examining the role of bodily awareness in perceiving the spatial layout of the distal environment.

6 The experience of egocentric space

Processes of O-integration are initiated only when there is sufficient evidence that a single object (or event) is the causal antecedent of stimulation arriving in different modalities. One important source of such evidence, in many contexts (but not all, see Spence 2013), is *spatial congruence*: signals that arrive from a common location in space, other things being equal, are more likely to have a common cause than those that don't.

The puzzle, from a neural point of view, is that the senses initially encode spatial information using different, body-relative frames of reference (Stein and Meredith 1993; Driver and Spence 1998a, b; Clark 2011). An object's location is initially encoded in eye-centered (retinocentric) coordinates by vision, head-centered coordinates by audition, and hand-centered, somatotopic coordinates by manual touch. The receptor surfaces for different senses, moreover, don't stand in invariant spatial relations to one another: the relative position of your eyes, head, and hands change from one moment to the next. How, then, is a stable, crossmodally accessible representation of external space that can support O-integration achieved?

The solution to the puzzle involves combining modality-specific sources perceptual information about an object's location with proprioceptive-kinaesthetic information about the current spatial configuration of the body. Driver and Spence write:

When taken in isolation, each modality signals stimulus location with respect to its own receptor surface only (e.g. on the eye for vision, or on the skin for touch). Since the receptors for each modality can move freely relative to external objects (as in eye- or hand-movements) and can also move relative to each other (as when making an eye-movement but no hand-movement, or vice versa), a single modality alone cannot provide a stable representation of external space. Instead, this requires integration of information from multiple sensory modalities, so that *current posture is taken into consideration*, as well as the location of stimuli on receptor surfaces (Driver and Spence 1998a, p. 254, emphasis added).

⁷ See, however, Longo and Haggard (2010, 2011), Longo et al. (2012), and Longo (2014) for evidence that this representation is systematically distorted in surprising ways.

Generating more useful representations of space (e.g., in order to derive stable visual position across saccades) requires spatial information to be integrated from more than one sense (e.g., retinal position to be combined with extraretinal signals about eye-position). Thus, even when attending to a particular location in just the “visual” world, it may be that more than one sense is involved in generating a stable representation of that location (Spence and Driver 2004, pp. 180–81).

Supporting this proposal, there is a large body of neuroscientific evidence that neurons in posterior parietal cortex compute the subject-relative locations of objects in external space by integrating incoming exteroceptive information with sources of proprioceptive-kinaesthetic information about relevant bodily effectors (Zipser and Andersen 1988; Andersen et al. 1997; Cohen and Andersen 2004; Azañón et al. 2010; for a review, see Whitlock 2017). In the visual case, retinocentric information about an object’s location can be combined with extraretinal signals specifying eye position in order to generate a representation of the object’s location in a head-centered reference frame. The relevant computational process is known as *coordinate transformation*. Importantly, the process of coordinate transformation can be iterated by further combining proprioceptive information about the rotation of the perceiver’s head and the joint angles of her shoulder, elbow, and wrist to work up a representation of the object’s location relative to her torso or hand. Similar computations are possible for the other exteroceptive modalities.

When combined with sources of proprioceptive information, visual, auditory, and tactile signals can thus all be used to represent locations in external space using common, body-relative reference frames (Bermúdez 1998, chap. 6; Spence and Driver 2004; Grush 2007; Briscoe 2009; Clark 2011; Wu 2014; Briscoe and Schwenkler 2015; Tamè et al. 2019). And coordinate transformation thus makes it possible to assess whether the sources of signals arriving different modalities are spatially congruent as is often required for purposes of non-generative, optimizing integration (Briscoe 2016). For example, combining visual signals with information about eye position enables the perceptual system to compute the location of something *seen* relative to the perceiver’s head and, hence, the distance between that location and the location of something *heard*. This, in turn, may lead to the ventriloquizing of the latter in the direction of the former.

Because the computations involved in coordinate transformation depend on the availability of up-to-date proprioceptive information of various kinds, it is not surprising that illusions of limb movement and movement can result in corresponding *illusions of object motion*. Roll et al. (1991), for example, found that vibratory stimulation of muscles in the neck induces both an illusory sensation of upward head movement as well as an experience of upward, apparent motion for a stationary, visually fixated target. More strikingly, Lackner and Levine (1978) found that when a target light was attached to a subject’s stationary hand and illusory motion of the forearm induced by mechanical muscle vibration, the light appeared to displace in the direction of illusory motion even though both the target and the subject’s eyes remained motionless. “[T]hese findings”, Lackner and Levine write, “provide further support for the notion that the computation of visual direction involves the

operation of spatial constancy mechanisms that utilize information from retinal, oculomotor, and postural sources in assigning the apparent direction of an object in relation to the head and body” (1979, p. 282).

Also relevant are studies of *perceptual adaptation* involving devices that displace, invert, or reverse the retinal image [Welch (1978) and Redding and Wallace (1997) offer comprehensive overviews]. Some of the earliest experiments on optical rearrangement were performed by von Helmholtz (1867/2005), who practiced reaching to targets while wearing prisms that displaced the retinal image to the left. The initial tendency was to reach too far in the direction of lateral displacement, but after numerous trials reaching gradually regained its former level of accuracy. Further, Helmholtz discovered that immediately after removing the prisms from his eyes errors in were made in the opposite direction, that is, when reaching for a target, he now moved his hand too far to the right.

The dominant explanation of this after-effect involves the recalibration of proprioceptive signals arriving from one or more points of articulation in the body (Stratton 1899; Harris 1965; Welch 1978, chap. 3). After adapting to laterally displacing prisms, outputs from receptors that previously signified deviation of the eyes to the left may come to signify that they are pointing straight ahead. Hence, an object located somewhere on the plane of the bodily mid-line may visually appear to be located to the right. Similar recalibrations and after-effects are possible for proprioceptive signals concerning head rotation. For present purposes, the point is that work on perceptual adaption provides further evidence that sources of information about the configuration and movement of the body play a major role in perceptually localizing objects in external space.

A third set of relevant empirical findings concern the skillful deployment of *spatial attention*. Numerous studies conducted by Jon Driver, Charles Spence, and others have found that attending to a region of space in one perceptual modality can attract attention to the same region in other modalities (for helpful overviews, see Spence 2010; Spence et al. 2014). Hearing a sound on one’s left, for example, can enhance the speed and accuracy with which targets subsequently presented in the same leftward region of space are discriminated by vision or touch.

Driver and Spence’s studies indicate that the crossmodal links in selective attention are mediated by representations of cue and target locations within non-modality-specific frames of reference. What matters, for example, in tactile-visual attentional facilitation isn’t simply whether cutaneous stimulation is delivered to the skin on the left or right side of your body (and, so, whether the tactile cue is registered by the left or right hemisphere of your brain). What matters rather is whether the stimulated region of the bodily surface is located to the left or right of the direction of gaze at the time of stimulation, even if that region is not currently visible (Kennett et al. 2002). Parallel points hold with respect to visual-tactile facilitation: when you are looking straight ahead with your arms crossed in front of you, a flash on the right side of your visual field will advantage tactile discriminations with your left hand, which is now closer to the light. Austen Clark in an insightful discussion comments: “Notice that this ‘closer to’ is a *cross-modal relative distance*: When a subject’s arms are crossed, tactile stimuli on the left hand are now closer to the light source than are tactile stimuli on the right hand. This ‘closer to’ cannot be assessed

by touch alone or by vision alone; it requires representation of the place of the light and the place of the hands in a *common space*” (2011, p. 386, my emphasis). In order to estimate and adaptively respond to the distance between a tactile cue and a visual target in a “common” or crossmodally accessible space, attentional systems must take into account proprioceptive-kinaesthetic information of various kinds. For example, to estimate the distance between a tactile stimulus delivered to the left hand and a light displayed to the right of the bodily midline, attentional systems must be able to access shoulder and elbow joint angle information. “A third modality (here proprioception)”, as Driver and Spence write, “can apparently influence the attentional interactions between two other modalities” (1998b, p. 1322).

The above findings concern crossmodal links in spatial attention. It is important to emphasize, however, that unimodal spatial attention may also in many contexts make use of non-modality-specific frames of reference. Attending to a visible location need not be attending to a location in the *visual field*—that is, to a location whose azimuth and elevation are represented using retinocentric coordinates. Similarly, attending to the location of a tactile stimulus need not be attending to a somatotopic location on the skin. Driver and Grossenbacher (1996), for example, performed a study in which blindfolded subjects were required to discriminate between different vibrotactile stimuli presented to one hand, while ignoring concurrent vibrotactile distractors on the other hand. Distractor vibrations were found to interfere significantly less when subjects’ hands were placed far apart than when placed close together, even though the stimuli and distractors were respectively presented to exactly the same locations on the bodily surface. This finding suggests that “some forms of proprioceptive-tactile integration can take place prior to attentional selection” (Driver and Grossenbacher 1996, p. 232). In other words, tactual attention can be directed to locations that are represented in a frame of reference that isn’t somatotopic and, so, could potentially be used to also represent the locations of objects that are seen or heard. A similar conclusion can be extracted from experimental studies of the so-called “crossed-hands deficit”, in which a subject’s ability to judge the temporal order of tactile stimuli presented to her hands is impaired when her arms are crossed (Yamamoto and Kitazawa 2001; Shore and Spence 2002; Heed et al. 2012; Azañón et al. 2016).

Studies of patients with *unilateral perceptual neglect* further back-up the view that unimodal spatial attention need not be directed to locations represented using a spatial coding system that is proprietary to the sense in question. Whether a neglect patient will become aware of a visually or tactually presented stimulus can depend not only on the latter’s position in relation to the eye or the surface of the body, respectively, but also on the spatial configuration of the patient’s body (Kooistra and Heilman 1989; Karnath et al. 1991; Driver et al. 1997; Vuilleumier et al. 1999). Summarizing the neuropsychological evidence, Driver and Vuilleumier write:

in neglect patients, the same visual stimulus at a fixed retinal position may be neglected or detected depending on the current orbital position of the eye. For instance, a left visual field stimulus that was neglected with the head and eyes directed forwards may be detected when presented at the same point on the retina but with the eyes and/or head turned right. Passively twisting the trunk

towards the left, while leaving the eyes and head facing straight ahead, can also bring an otherwise neglected left visual field stimulus back into awareness, again suggesting that extraretinal factors can influence visual neglect. This is a remarkable finding; whether the patient sees a stimulus depends not only on the visual information entering the eyes, *but on body posture also* (2001, p. 46, emphasis added).

...these studies demonstrate there is more to the spatial nature of neglect than mere primary sensory loss within fixed retinal or somatotopic co-ordinates. They imply that neglect can arise at higher levels of spatial representation, some of which involve the integration of different sources of spatial information... (2001, p. 48).

I have reviewed numerous sources of evidence in support of the view that proprioception and kinaesthesia play a critical role in crossmodal representation of the spatial environment.⁸ The proposal that I want to consider in the last part of this essay is that the kind of adaptive interaction between the senses involved in coordinate transformation qualifies as an example of G-integration. More specifically, I will argue, focusing on the case of vision, that when we experience an object's subject-relative location in crossmodally accessible space, we have an experience that represents a *strongly novel* type of perceptible feature—a feature that is only accessible multisensorily.

Before proceeding, it may avoid confusion to distinguish the problem of *visual direction perception* (VDP) from the problem of *visual space constancy* (VSC). The problem of VDP is the problem of perceptually estimating an object's egocentric direction, that is, its direction in space in relation to the observer (MacKay 1973; Swanston et al. 1987). An adequate solution to the problem of VDP must specify, among other things, the bodily frame of reference relative to which direction is perceived as well as which sources of information are necessary to compute an object's egocentric direction. For example, to compute direction in *head-centered* coordinates, as already mentioned, two variables require values: the perceptual system needs an estimate of the object's location in an eye-based or retinocentric frame of reference and, in addition, an estimate of the position of the eye in its orbit. Over a century of work on VDP suggests that there are two main sources of information about eye-position: afferent or “inflowing” signals from receptors in the six extraocular muscles and, as von Helmholtz (1867/2005) was one of the first to propose, efferent or “outflowing” information concerning the saccade command signal from the oculomotor system.⁹

Other bodily frame of reference for VDP are possible. Peacocke (1992, 2002), for example, maintains that visual experiences represent the way visible surfaces are arrayed in three-dimensional space relative to a point placed in the subject's torso.

⁸ For other empirical findings that support this conclusion, see Howard and Templeton (1966), Berthoz (2000), Lackner and DiZio (2005), and Taylor et al. (2017).

⁹ See Donaldson (2000) and Bridgeman (2010) for historically oriented reviews of psychological work on VDP. See Husserl (1907/1997) for an interesting and rich philosophical discussion of the role of kinaesthesia in visual space perception as well as Hatfield (1990), Gallagher (2005), Schwenkler (2014), Briscoe (2014), and Briscoe and Grush (2017) for further philosophical assessments.

“The appropriate set of labeled axes”, he writes, “captures distinctions in the phenomenology of experience itself. Looking straight ahead at Buckingham Palace is one experience. It is another to look at the palace with one’s face still toward it but with one’s body turned toward a point on the right. In this second case the palace is experienced as being off to one side from the direction of straight ahead, even if the view remains exactly the same as in the first case” (1992, p. 62). The visual system, however, initially encodes an object’s location relative to the eye, that is, in retino-centric coordinates. Representing an object’s location in visual experience relative to the torso thus presupposes prior integration of visual information about the object’s eye-relative location with proprioceptive information about the direction of gaze and the orientation of the head.

The problem of visual space constancy (VSC), by contrast, concerns the experienced *stability* of the visual world across eye-movements. When we execute a saccade, the image of the world projected on the retina displaces in the direction of eye rotation, yet objects in front of us do not appear to move. A recent treatment by Wayne Wu offers a nice characterization of the phenomenon: “Eye movements... generate massive changes in the position of objects relative to the direction of gaze and the foveal and fixation point positions that define it. Yet despite saccadic eye movement, the world remains spatially stable. How then do we represent spatial constancy?” (2014, p. 393). The problem of VSC concerns which sources of information support such experienced stability, how they are used, and the identification of supporting brain areas.

Two points are worth emphasizing. First, the problem of VDP has explanatory priority relative to the problem of VSC. Indeed, if VSC is the experience of constancy with respect to an object’s egocentric location—“the spatial constancy that we experience every time the eye moves”, Wu writes, “is the spatial constancy of objects relative to us, say relative to our head, torso or another otherwise immobile body part” (2014, p. 394)—then even to coherently formulate the problem of VSC we first need to understand *what it is* to experience an object’s egocentric location. Second, work in vision science suggests that the problem of VDP and the problem of VSC may have rather different solutions. In particular, there is on-going debate concerning whether the visual system needs to compensate for retinal image motion and maintain VSC by taking proprioceptive/kinaesthetic information about eye-position into account.¹⁰

To prevent misunderstanding, I am not dealing with the problem of VSC in what follows. For my purposes, it is the problem of VDP that matters, and there is good evidence that VDP is enabled by the coordinated use of vision, proprioception, and kinaesthesia.

Now, it may be conceded that subpersonal processes of coordinate transformation enable representation of an object’s egocentric location within one or another

¹⁰ “The role of efference copy and of extraretinal signals generally...,” Bridgeman suggests, “appears to be to inform the brain about static eye position during visual fixation, the time between saccades when the retina is transducing the visual world reliably. It does not support space constancy” (2010, p. 105). For other assessments, see Bridgeman et al. (1994), Deubel (2004), and Wurtz (2017).

non-modality-specific spatial frame of reference. It may also be conceded that processes of coordinate transformation generate strongly novel perceptual representational contents: neither vision, nor proprioception, nor kinaesthesia is by itself capable of representing an object as occupying, for example, a certain head or torso relative location in surrounding space. Nonetheless, it may be denied that *perceptual awareness* of an object's location has content that goes beyond the contents associated with each of the modalities contributing to the coordinate transformation process. In general, claims about multisensory interaction at the subpersonal, information-processing level do not straightforwardly entail claims about theoretically interesting forms of multisensory experience (Deroy et al. 2014; Macpherson 2011b; Mudrik et al. 2014; O'Callaghan 2008). Our awareness of a visible object *O*'s subject-relative location in space, phenomenologically speaking, may decompose into awareness of *O*'s retinocentric position plus co-conscious awareness of the eye's position in its orbit and/or the head's orientation relative to the torso. If so, *O*'s location relative to part(s) of the body other than the eye may be inferred rather than consciously perceptible, and our experience of *O* wouldn't possess any novel content. Call this view *decompositionalism about VDP*.

There are at least three lines of response here. To begin with, decompositionalism about VDP isn't phenomenologically well-motivated. The locations of which we seem most directly and non-inferentially to be perceptually aware aren't locations represented in multiple, distinct, modality-specific frames of reference, but rather locations in a single, crossmodally accessible system of places and distances (Evans 1982; Bermúdez 1998, chap. 6; Driver and Spence 1998a; Kaplan 2007; O'Callaghan 2008; Clark 2011; Bayne 2014; Matthen 2014, 2017; Nudds 2014). For example, we may be aware that something we *hear* is above us and to the left and, at the same time, that something we *see* is down and to the right, across many different changes in eye-position and bodily posture. Such awareness presupposes the ability to keep track of where objects are located around the body using one or more non-modality-specific spatial frames of reference.

According to the decompositionalist, the experience of VDP comprises two introspectively distinguishable, modality-specific components: awareness of the perceived object's position in an eye-centered frame of reference, and, in addition, co-conscious awareness of the eye's position relative to the head. Even if we are able introspectively to distinguish between these two experiential components—and this is my second response to the objection—the experience of VDP could still involve awareness of an *additional*, non-modality-specific component.

Consider by way of analogy the so-called “duality” or “bipolarity” of haptic touch. Experiences of haptic touch have a bipolar phenomenology and content because, when you are engaged in haptically exploring an object, it is possible to attend to the object's tangible properties and, at the same time, to how the configuration of the relevant part(s) of your body is changing. One way this could be true would be if, as I argued in the previous section, experiences of haptic touch involve *both* modality-specific cutaneous tactual, proprioceptive, and kinaesthetic content as well as non-modality-specific object-representing content. Experiences of haptic touch, I proposed, do not resolve into their modality-specific components,

but, in addition, make weakly novel multisensory features accessible to perceptual consciousness.

This view contrasts with a decompositionalist account of the relation between haptic touch perception and bodily awareness that Fulkerson (2014a) dubs *Constitution Tactual Bodily Dependence* (Constitution-TBD). According to Constitution-TBD, experiences of haptic touch are just complex sets of concurrent proprioceptive, kinaesthetic, and cutaneous tactual experiences. Object-directed contents in haptic touch, in other words, are constituted or defined by bodily contents. Constitution-TBD confronts a battery of objections (Scott 2001; de Vignemont and Massin 2015; Richardson 2013). Here I discuss two. First, when we haptically explore an object, we may be focally aware of the object's tangible properties and only peripherally aware of the bodily experiences that result from our exploration (Bermúdez 1998, chap. 6). This, however, seems inconsistent with the claim that experiences of haptic touch are fully constituted by their modality-specific proprioceptive, kinaesthetic, and cutaneous components: were this the case, focally attending to the object's shape or texture, for example, would amount to focally attending to the relevant complex of bodily experiences or aspects thereof. A second problem with Constitution-TBD is that many introspectively different sets of bodily experiences can accompany representation of the *same* tangible property. Fulkerson writes:

I can... feel a surface as hard by pressing it with my thumb, or forefinger, or forehead, or toes, or back, and so forth. Each of these is an entirely different state of bodily awareness (with distinct contents), but all assign the same property, *hardness*, to the surface. It does not seem likely that externally directed touch just is a specific set of bodily experiences. While bodily contents *are* required, they are a diverse and disjointed lot rather than contents that partly *constitute* or *define* external contents. Consider, what is the associated general bodily content that defines feeling something as *squishy*? Or *slimy*? Or *elastic*? It seems entirely plausible that we can experience an object as squishy in many ways through touch, and certainly not through a precise, well-defined bodily content.... The problem isn't that we have no conception of what our bodies feel like when we touch something slimy; it's that such feelings do not constitute *what it is* for something to be [experienced as] slimy (2014a, p. 85; see also Scott 2001, pp. 153–154).¹¹

Something analogous is plausibly true of haptic shape perception. Numerous, significantly different combinations of proprioceptive, kinaesthetic, and cutaneous tactual input are capable of revealing the 3D shape of a coffee mug, which shape is experienced as constant and distinct from the contributions made by those modalities to perceptual consciousness. There is non-modality-specific perceptual constancy or

¹¹ David Katz, in *The World of Touch* (1925/1989), poses a similar anti-decompositionalist objection with respect to the role of kinaesthesia in texture discrimination: "Make a movement over your desk pad, and feel its quality of 'soft roughness.' This impression builds up during the movement, without having even the slightest trace of movement as a component in itself.... To be sure, one subjectively experiences the movement of the touch organ when it seizes upon the texture of a surface, *but the movement does not thereby become a component of the surface touch*" (Katz 1925/1989, p. 80, emphasis added).

“objectification” in haptic touch (Burge 2010, p. 399) across introspectively salient variation in the contents of bodily awareness. Experiences of 3D shape by means of haptic touch are characteristically accompanied by experiences of proprioception, kinaesthesia, and cutaneous touch, but don’t seem to decompose into them.¹²

Also consider again, in this connection, the case of *flavor perception*. O’Callaghan writes:

...while it incorporates such features, apparent flavor may be more than just an agglomeration of gustatory, olfactory, and tactual qualities that are attributed to something in the mouth. If flavors are not just undifferentiated mixtures or unstructured collections of features, then they may involve novel or emergent features of a type that cannot be perceived unimodally. For instance, apparent flavor could involve a structure (qualitative or temporal) among its sense-specific components, an organic unity involving them, or an additional qualitative component beyond its modality-specific features (2015, pp. 567–568).

The relevant point is that the following two claims could *both* be true: (1) flavor perception includes various modality-specific components to which we can separately attend (even if such attending may be challenging—see Smith 2015), and (2) flavor perception involves a structural or qualitative component that goes *beyond* those modality-specific components. The proposal that apparent flavor is strongly novel feature is consistent with the fact that flavors have a complex structure and that we can differentially attend to their components.

Something analogous plausibly holds, I would now suggest, when it comes to experiencing a visible object’s location (VDP). You can attend, if you make the appropriate effort, to the object’s unchanging direction from you even as you attend to the object’s changing location in your visual field or to proprioceptive feelings accompanying changes in your direction of gaze (or the orientation of your head). That you can introspectively distinguish between the modality-specific components involved in the experience of VDP doesn’t mean that VDP factors without remainder into those components. Experiences of VDP may depend *inter alia* on sources of information provided by proprioception and kinaesthesia without being fully constituted or defined by the contributions of these systems to perceptual consciousness.

Finally, decompositionalism about VDP comes at a steep price. According to decompositionalism, subpersonal processes of coordinate transformation may generate non-modality-specific representations of the way objects are positioned around us in surrounding space, but personal-level awareness of space is always modality specific. Coordinate transformation processes, however, not only support the representation of subject-relative position, they also support the representation of subject-relative *movement*. Consider an observer viewing a luminous target in an otherwise dark room. Displacement of the target’s retinal image—say, from the left to the right of the direction of gaze—could be caused in at least two different ways.

¹² Compare Thomas Reid’s denial in his *Inquiry* that our perceptual conceptions of tangible properties like *hardness* and *extension* resolve into the proprioceptive, kinaesthetic, and tactual sensations elicited in us when we touch them. For a valuable discussion, see Buras (2015).

In one situation, the target is stationary, but the observer's eyes have saccaded to a point left of the target. In the other, the observer's eyes are stationary, but the object has moved to the right. In order to distinguish perceptually between these two situations, proprioceptive-kinaesthetic information needs to be taken into account. To use the language of Bayesian psychology, the observer will experience the object as moving only when this "hypothesis" is made sufficiently probable by both the available visual evidence (left to right displacement of the object's retinal image) and proprioceptive evidence (extra-retinal signals concerning the change in eye position) available to her perceptual system. The available visual evidence, by itself, underdetermines which of the two situations obtains.¹³

In the type of case just described, proprioceptive-kinaesthetic information is utilized to interpret (disambiguate) the displacement of an object's retinal image. Displacement of an object's retinal image, however, isn't necessary to perceive the object as in motion. As mentioned above, experiments by Roll et al. (1991) have found that experimentally-induced illusions of head or limb movement can trigger corresponding illusions of movement for visually fixated targets. Or consider everyday experiences of visually tracking a moving object, for example, a bird in flight, using *smooth pursuit eye movements*. As Bridgeman and co-authors note, smooth pursuit tracking "reverses the movement conditions on the retina: the tracked object sweeps across the retina very little, while the background undergoes a brisk motion" (1994, p. 255). Nonetheless, it is the tracked object that appears to be in motion while the environment appears to be stationary. Furman and Gur (2012) refer to this as "pursuit compensation".

There is strong empirical support for the view that sources of proprioceptive-kinaesthetic information play an important role in pursuit compensation. In particular, brain-imaging and electrophysiological studies indicate that (1) the visual brain compensates for pursuit-induced retinal motion by means of efference-based information about eye movement and (2) that it encodes the changing position of the tracked object in head-centered coordinates (for a review, see Furman and Gur 2012). Neuropsychological findings relevantly indicate that failure to integrate efference-based information about eye movement leads to a breakdown in perceived background stability when engaged in smooth pursuit (Haarmeier et al. 1997; Fischer et al. 2012; for philosophical discussion, see Wu 2014).

According to decompositionism about VDP, our experience of an object's egocentrically specified position resolves into visual awareness of the object's retinocentric position plus co-conscious proprioceptive-kinaesthetic awareness of the spatial configuration of the body, in particular, the eye's changing position in its orbit. We are now in a position to see a further problem with this view beyond those already indicated: decompositionism about VDP entails a similarly reductive claim about the experience of subject-relative motion. If the experience of visual direction for a stationary object O factors without remainder into an experience of O 's position

¹³ Sources of proprioceptive-kinaesthetic information perform an analogously important function in spatial vision when both the eyes and head are moving, for example, during exercise of the vestibulo-ocular reflex (VOR). See Fetter (2007) and Medendorp (2011) for discussion.

in a retinocentric frame of reference plus co-conscious proprioceptive-kinaesthetic experiences of bodily configuration, then the same should be true of the experience of visual direction when O is moving relative to the observer. After all, to experience an object as moving, say, from your left to your right, just is to experience its egocentric direction as undergoing a certain kind of change. Thus, when tracking a moving target by means of smooth pursuit, decompositionism entails that your experience resolves into an experience of the target's stable position in your visual field plus co-conscious proprioceptive-kinaesthetic awareness of your changing direction of gaze. You may *infer* from these experiences that the target is moving along a certain path through space, but, in notable contrast, for example, with the target's color or shape or size, the target's motion is not a feature that, strictly speaking, is present in the content of your experience.

This, to put it gently, is counterintuitive. There has been long-standing debate between restrictivist and expansionist accounts of the admissible contents of visual experience (Prinz 2006; Siegel 2010; Briscoe 2015; Siegel and Byrne 2016). Restrictivists typically deny that high-level kinds such as *tiger*, *pine tree*, or *alarm clock* are represented in the phenomenal contents of visual experience. But not even the most austere of restrictivists would deny, I presume, that we are perceptually aware of an object's subject-relative motion in the case just described. On the contrary, subject-relative motion, like shape, color, and size, is typically taken to figure among the commonly admissible "low-level" contents of visual experience. Decompositionism about VDP, in short, comes at the cost of an overly spare inventory of the range of spatial features represented in perceptual awareness.¹⁴

7 Conclusion

I began above by distinguishing between optimizing and generative forms of multisensory integration. Contrary to the decomposition thesis, I then argued that generative multisensory integration (G-integration) can give rise to perceptual experiences representing spatial features of a *novel type*, features whose instances couldn't be represented by any of the contributing modalities functioning in isolation. More specifically, I defended two proposals. First, I argued that the representation of solid, 3D shape in haptic touch is a form of G-integration because no mere experience of cutaneous touch, kinaesthesia, or proprioception could be an experience of a non-bodily object as, for example, cuboid, or pyramidal, or teapot-shaped. Since solid shape is unimodally accessible to vision, however, solid shape is a weakly novel type of multisensory feature: it is novel only relative to the representational powers of cutaneous touch, proprioception, and kinaesthesia. Second, I argued that no mere experience of any modality is an experience of an object's subject-relative

¹⁴ The structure of my argument here closely mirrors the structure of one of O'Callaghan's (2017) arguments for intermodal apparent motion. I have focused in this section on the experience of object motion relative to a stationary subject. See Schwenkler (2014) for an argument that self-motion relative to a stationary object or 3D scene also figures among the admissible contents of visual experience.

location in crossmodally accessible, 3D space. More precisely, neither vision, nor audition, nor touch, in the absence of contributions from proprioception and kinaesthesia, is capable of representing an object's location in a non-modality-specific, egocentric frame of reference. As Roll and Roll (1991) nicely put it, bodily awareness is the “link between body space and extra-personal space”. Since egocentric location is novel relative to the representational powers of any modality operating by itself, egocentric location, I have proposed, is a strongly novel type of multisensory feature.¹⁵

Acknowledgements Early versions of this paper were presented at the *Multimodal Perception* workshop held at the University of Antwerp in April 2015 and the *Senses and Crossmodal Perception* conference held at the University of Berne in October 2016. I'm grateful to Fiona Macpherson, Bence Nanay, Mohan Matthen, Matthew Nudds, and especially Casey O'Callaghan for helpful discussions and encouragement. I would also like to thank four anonymous referees for insightful comments and questions that resulted in many improvements. This article was completed with the generous support of the US-UK Fulbright Commission (Grant No. C5E1C25997U), while I was Fulbright-Scotland Visiting Professor at the University of Glasgow (2017–2018).

References

- Andersen, R., Snyder, L., Bradley, D., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, 20(1), 303–330.
- Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition*, 17(3), 1016–1031.
- Azañón, E., Longo, M., Soto-Faraco, S., & Haggard, P. (2010). The posterior parietal cortex remaps touch into external space. *Current Biology*, 20(14), 1304–1309.
- Azañón, E., Mihaljevic, K., & Longo, M. (2016). A three-dimensional spatial characterization of the crossed-hands deficit. *Cognition*, 157, 289–295.
- Bastian, H. C. (1888). The “muscular sense”; Its nature and cortical localisation. *Brain*, 10, 1–137.
- Bayne, T. (2014). The multisensory nature of perceptual consciousness. In D. Bennett & C. Hill (Eds.), *Sensory integration and the unity of consciousness* (pp. 15–36). Cambridge: MIT Press.
- Bermúdez, J. L. (1998). *The paradox of self-consciousness*. Cambridge: MIT Press.
- Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. *Advances in Psychology*, 129, 347–362.
- Berthoz, A. (2000). *The brain's sense of movement*. Cambridge: Harvard University Press.
- Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’. *Touch That Eyes See*. *Nature*, 391(6669), 756–756.
- Bridgeman, B. (2010). Space constancy: The rise and fall of perceptual compensation. In R. Nijhawan & B. Khurana (Eds.), *Space and time in perception and action* (pp. 94–108). Cambridge: Cambridge University Press.

¹⁵ An anonymous referee brings up the case of Ian Waterman [see Cole (1991) for a detailed description and the marvelous BBC *Horizon* documentary “The Man Who Lost His Body”]. Waterman suffers total deafferentation of his body from the neck down and, thus, has no sense of touch or proprioception in his torso or limbs. “Did this lead to deficits in visual object localization,” the referee asks, “as the author's view might predict?” My reply, in brief, is that Waterman's intact proprioception and kinaesthesia from the neck up afford information about eye and head position sufficient for perceptually locating objects in an egocentric frame of reference. My view doesn't predict visuospatial deficits in this case, and Waterman doesn't exhibit them. See Gallagher (2005), chapter two, for a careful discussion of Waterman's case and further considerations in support of this assessment.

- Bridgeman, B., van der Heijden, A. H. C., & Velichkovsky, B. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, 17, 247–292.
- Briscoe, R. (2008). Vision, action, and make-perceive. *Mind and Language*, 23(4), 457–497.
- Briscoe, R. (2009). Egocentric spatial representation in action and perception. *Philosophy and Phenomenological Research*, 79(2), 423–460.
- Briscoe, R. (2011). Mental imagery and the varieties of amodal perception. *Pacific Philosophical Quarterly*, 92(2), 153–173.
- Briscoe, R. (2014). Spatial content and motoric significance. *Enactivism: Arguments and Applications*, *Avant*, 5(2), 199–217.
- Briscoe, R. (2015). Cognitive penetration and the reach of phenomenal content. In A. Raftopoulos & J. Zeimbekis (Eds.), *The cognitive penetrability of perception: New philosophical perspectives* (pp. 174–199). Oxford: Oxford University Press.
- Briscoe, R. (2016). Multisensory processing and perceptual consciousness: Part I. *Philosophy Compass*, 11(2), 121–133.
- Briscoe, R. (2017). Multisensory processing and perceptual consciousness: Part II. *Philosophy Compass*, 12(12), e12423.
- Briscoe, R. (2018). Superimposed mental imagery: On the uses of make-perceive. In F. Macpherson & F. Dorsch (Eds.), *Perceptual memory and perceptual imagination* (pp. 161–185). Oxford: Oxford University Press.
- Briscoe, R., & Grush, R. (2017). Action-based theories of perception. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of philosophy*, Spring 2017 edition. Retrieved from <https://plato.stanford.edu/archives/spr2017/entries/action-perception>.
- Briscoe, R., & Schwenkler, J. (2015). Conscious vision in action. *Cognitive Science*, 39, 1435–1467.
- Brogaard, B. (2012). What do we say when we say how or what we feel? *Philosophers Imprint*, 12(11), 1–22.
- Buras, T. (2015). Thomas Reid's *Experimentum crucis*. In R. Copenhaver & T. Buras (Eds.), *Thomas Reid on mind, knowledge, and value* (pp. 14–34). Oxford: Oxford University Press.
- Burge, T. (2010). *Origins of objectivity*. Oxford: Oxford University Press.
- Chang, D., Nesbitt, K. V., & Wilkins, K. (2007). The Gestalt principle of continuation applies to both the haptic and visual grouping of elements. In *Proceedings of the second joint EuroHaptics conference and symposium on Haptic interfaces for virtual environment and teleoperator systems* (pp. 15–20). Tsukuba, Japan: IEEE Computer Society.
- Clark, A. (2011). Cross modal links and selective attention. In F. MacPherson (Ed.), *The senses: Classic and contemporary philosophical perspectives* (pp. 375–396). Oxford: Oxford University Press.
- Cohen, Y., & Andersen, R. (2004). Multisensory representations of space in the posterior parietal cortex. In G. Calvert, C. Spence, & B. Stein (Eds.), *The handbook of multisensory processes* (pp. 463–479). Cambridge: MIT Press.
- Cole, J. (1991). *Pride and a daily marathon*. Cambridge: MIT Press.
- de Vignemont, F. (2014). A multimodal conception of bodily awareness. *Mind*, 123(492), 989–1020.
- de Vignemont, F. (2016). Bodily awareness. In *The Stanford encyclopedia of philosophy* (Summer 2016 Edition). <https://plato.stanford.edu/archives/sum2016/entries/bodily-awareness/>.
- de Vignemont, F. (2018). *Mind the body: An exploration of bodily self-awareness*. Oxford: Oxford University Press.
- de Vignemont, F., & Massin, O. (2015). Touch. In Mohan Matthen (Ed.), *The Oxford handbook of the philosophy of perception* (pp. 294–313). Oxford: Oxford University Press.
- Deroy, O., Chen, Y., & Spence, C. (2014). Multisensory constraints on awareness. *Philosophical Transactions of the Royal Society B*, 369(1641), 20130207.
- Deubel, H. (2004). Localization of targets across saccades: Role of landmark objects. *Visual Cognition*, 11(2–3), 173–202.
- Donaldson, I. M. L. (2000). The functions of the proprioceptors of the eye muscles. *Phil. Trans. R. Soc. Lond. B*, 355, 1685–1754.
- Driver, J., & Grossenbacher, P. G. (1996). Multimodal spatial constraints on tactile selective attention. In T. Inui & J. L. McClelland (Eds.), *Attention and performance*, XVI (pp. 209–236). Cambridge: MIT Press.
- Driver, J., Mattingley, J. B., Rorden, C., & Davis, G. (1997). Extinction as a paradigm measure of attentional bias and restricted capacity following brain injury. In P. Theier & H.-O. Karnath (Eds.), *Parietal Lobe contributions to orientation in 3D space* (pp. 401–429). Berlin: Springer.

- Driver, J., & Spence, C. (1998a). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, 2(7), 254–262.
- Driver, J., & Spence, C. (1998b). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 353(1373), 1319–1331.
- Driver, J., & Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition*, 79(1), 39–88.
- Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.
- Fetter, M. (2007). Vestibulo-ocular reflex. *Neuro-Ophthalmology*, 40, 35–51.
- Fischer, E., Bühlhoff, H. H., Logothetis, N. K., & Bartels, A. (2012). Human areas V3A and V6 compensate for self-induced planar visual motion. *Neuron*, 73(6), 1228–1240.
- Fridland, E. (2011). The case for proprioception. *Phenomenology and the Cognitive Sciences*, 10(4), 521–540.
- Fulkerson, M. (2011). The unity of haptic touch. *Philosophical Psychology*, 24(4), 493–516.
- Fulkerson, M. (2014a). *The first sense: A philosophical study of human touch*. Cambridge: MIT Press.
- Fulkerson, M. (2014b). Rethinking the senses and their interactions: The case for sensory pluralism. *Frontiers in Psychology*, 5, 1–14.
- Furman, M., & Gur, M. (2012). And yet it moves: Perceptual illusions and neural mechanisms of pursuit compensation during smooth pursuit eye movements. *Neuroscience and Biobehavioral Reviews*, 36(1), 143–151.
- Gallace, A., & Spence, C. (2011). To what extent do gestalt grouping principles influence tactile perception? *Psychological Bulletin*, 137, 538–561.
- Gallace, A., & Spence, C. (2014). *In touch with the future*. Oxford: Oxford University Press.
- Gallagher, S. (2003). Bodily self-awareness and object perception. *Theoria et Historia Scientiarum*, 7(1), 53–68.
- Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Clarendon Press.
- Gibson, J. J. (1933). Adaptation, after-effect and contrast in the perception of curved lines. *Journal of Experimental Psychology*, 16, 1–31.
- Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, 69(6), 477–491.
- Grush, R. (2007). Skill theory v2.0: Dispositions, emulation, and spatial perception. *Synthese*, 159, 389–416.
- Haarmeier, T., Thier, P., Repnow, M., & Petersen, D. (1997). False perception of motion in a patient who cannot compensate for eye movements. *Nature*, 389, 849–851.
- Harris, C. S. (1965). Perceptual adaptation to inverted, reversed, and displaced vision. *Psychological Review*, 72(6), 419–444.
- Hatfield, G. (1990). *The natural and the normative: Theories of spatial perception from Kant to Helmholtz*. Cambridge: MIT Press.
- Hay, J., Pick, H., & Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychonomic Science*, 2, 215–216.
- Heed, T., Backhaus, J., & Röder, B. (2012). Integration of hand and finger location in external spatial coordinates for tactile localization. *Journal of Experimental Psychology: Human Perception and Performance*, 38(2), 386–401.
- Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. New York: Wiley.
- Huang, J., Gamble, D., Sarnlertsophon, K., Wang, X., & Hsiao, S. (2012). Feeling music: Integration of auditory and tactile inputs in musical meter. *PLoS ONE*, 7, e48496.
- Husserl, E. (1907/1997). *Thing and space: Lectures of 1907* (Richard Rojcewicz, trans.). Dordrecht: Springer.
- Jones, L., & Lederman, S. (2006). *Human hand function*. Oxford: Oxford University Press.
- Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: Sound-biased touch. *Current Biology*, 8(6), R190–R191.
- Kaplan, D. (2007). The nature and bases of spatial representation. Ph.D. Thesis submitted to the Department of Philosophy, Duke University.
- Karnath, H. O., Schenkel, P., & Fischer, B. (1991). Trunk orientation as the determining factor of the “contralateral” deficit in the neglect syndrome and as the physical anchor of the internal representation of body orientation in space. *Brain*, 114, 1997–2014.
- Katz, D. (1925/1989). *The world of touch*. Psychology Press.
- Kennett, S., Spence, C., & Driver, J. (2002). Visuo-tactile links in covert exogenous spatial attention remap across changes in unseen hand posture. *Perception and Psychophysics*, 64, 1083–1094.

- Kitagawa, N., Igarashi, Y., & Kashino, M. (2009). The tactile continuity illusion. *Journal of Experimental Psychology: Human Perception and Performance*, 35(6), 1784–1790.
- Knill, D., & Pouget, A. (2004). The bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), 712–719.
- Kooistra, C. A., & Heilman, K. M. (1989). Hemispatial visual inattention masquerading as hemianopia. *Neurology*, 39, 1125–1172.
- Lackner, J., & DiZio, M. (2005). Vestibular, proprioceptive, and haptic contributions to spatial orientation. *Annual Review of Psychology*, 56, 115–147.
- Lackner, J., & Levine, M. (1978). Visual direction depends upon the operation of spatial constancy mechanisms: The oculobrachial illusion. *Neuroscience Letters*, 7, 207–212.
- Levine, M. S., & Lackner, J. R. (1979). Some sensory and motor factors influencing the control and appreciation of eye and limb position. *Experimental Brain Research*, 36(2), 275–283.
- Longo, M. (2014). The effects of immediate vision on implicit hand maps. *Experimental Brain Research*, 232, 1241–1247.
- Longo, M. (2015). Implicit and explicit body representations. *European Psychologist*, 20, 6–15.
- Longo, M., Azañón, E., & Haggard, P. (2010). More than skin deep: Body representation beyond primary somatosensory cortex. *Neuropsychologia*, 48, 655–668.
- Longo, M., & Haggard, P. (2010). An implicit body representation underlying human position sense. *Proceedings of the National Academy of Sciences*, 107, 11727–11732.
- Longo, M., & Haggard, P. (2011). Implicit body representations and the conscious body image. *Acta Psychologica*, 141, 164–168.
- MacKay, D. M. (1973). Visual stability and voluntary eye movements. In *Central processing of visual information A: Integrative functions and comparative data* (pp. 307–331). Berlin: Springer.
- Macpherson, F. (2011a). Introduction: Individuating the senses. In Fiona Macpherson (Ed.), *The senses: Classic and contemporary philosophical perspectives* (pp. 3–38). Oxford: Oxford University Press.
- Macpherson, F. (2011b). Cross-modal experiences. *Proceedings of the Aristotelian Society*, 111(3), 429–468.
- Martin, M. G. F. (1992). Sight and touch. In Tim Crane (Ed.), *The contents of experience* (pp. 196–215). Cambridge: Cambridge University Press.
- Matthen, M. (2014). Active perception and the representation of space. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 44–72). Oxford: Oxford University Press.
- Matthen, M. (2015). The individuation of the senses. In M. Matthen (Ed.), *The Oxford handbook of philosophy of perception* (pp. 567–586). Oxford: Oxford University Press.
- Matthen, M. (2017). Is perceptual experience normally multimodal? In B. Nanay (Ed.), *Current controversies in philosophy of perception* (pp. 121–136). London: Routledge.
- McCloskey, D. I. (1973). Differences between the senses of movement and position shown by the effects of loading and vibration of muscles in man. *Brain Research*, 61, 119–131.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748.
- Medendorp, W. (2011). Spatial constancy mechanisms in motor control. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 366(1564), 476–491.
- Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences*, 18(9), 488–496.
- Nudds, M. (2014). Is audio-visual perception “amodal” or “crossmodal”? In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 166–190). Oxford: Oxford University Press.
- O’Callaghan, C. (2008). Seeing what you hear: Cross-modal illusions and perception. *Philosophical Issues*, 18(1), 316–338.
- O’Callaghan, C. (2014). Not all perceptual experience is modality specific’. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 133–165). Oxford: Oxford University Press.
- O’Callaghan, C. (2015). The multisensory character of perception. *The Journal of Philosophy*, 112(10), 551–569.
- O’Callaghan, C. (2017). Grades of multisensory awareness. *Mind and Language*, 32(2), 155–181.
- O’Shaughnessy, B. (1980). *The will* (Vol. 1). Cambridge: Cambridge University Press.
- O’Shaughnessy, B. (1995). Proprioception and the body image. In J. L. Bermúdez, A. Marcel, & N. Eilan (Eds.), *The body and the self* (pp. 175–204). Cambridge: MIT Press.
- Peacocke, C. (1992). *A study of concepts*. Cambridge: MIT Press.
- Peacocke, C. (2002). Postscript to peacocke 1994. In Y. Gunther (Ed.), *Essays on nonconceptual content* (pp. 107–132). Cambridge MA: MIT Press.

- Prinz, J. (2006). Beyond appearances: The content of sensation and perception. In T. Gendler & J. Hawthorne (Eds.), *Perceptual experience*. Oxford: Clarendon Press.
- Proske, U., & Gandevia, S. (2012). The proprioceptive senses: Their roles in signaling body shape, body position and movement, and muscle force. *Physiological Reviews*, 92(4), 1651–1697.
- Radeau, M., & Bertelson, P. (1977). Adaptation to auditory-visual discordance and ventriloquism in semi-realistic situations. *Perception and Psychophysics*, 22, 137–146.
- Redding, G., & Wallace, B. (1997). *Adaptive spatial alignment*. Mahwah: Erlbaum.
- Richardson, L. (2013). Bodily sensation and tactile perception. *Philosophy and Phenomenological Research*, 86(1), 134–154.
- Ritchie, J. B., & Carruthers, P. (2015). The bodily senses. In M. Matthen (Ed.), *The Oxford handbook of philosophy of perception* (pp. 353–370). Oxford University Press.
- Rock, I., & Harris, C. S. (1967). Vision and touch. *Scientific American*, 216, 96–107.
- Rohde, M., van Dam, L., & Ernst, M. (2016). Statistically optimal multisensory cue integration: A practical tutorial. *Multisensory Research*, 29, 279–317.
- Roll, J.-P., & Roll, R. (1991). Eye and neck proprioceptive messages contribute to the spatial coding of retinal input in visually oriented activities. *Experimental Brain Research*, 85, 423–431.
- Roll, J.-P., Roll, R., & Velay, J.-L. (1991). Proprioception as a link between body space and extra-personal space. In J. Paillard (Ed.), *Brain and space* (pp. 112–132). New York: Oxford University Press.
- Samad, M., Chung, A. J., & Shams, L. (2015). Perception of body ownership is driven by Bayesian sensory inference. *PloS one*, 10(2), e0117178.
- Schwenkler, J. (2013). The objects of bodily awareness. *Philosophical Studies*, 162(2), 465–472.
- Schwenkler, J. (2014). Vision, self-location, and the phenomenology of the “point of view”. *Noûs*, 48(1), 137–155.
- Scott, M. (2001). Tactual perception. *Australasian Journal of Philosophy*, 79(2), 149–160.
- Serino, A., & Haggard, P. (2010). Touch and the body. *Neuroscience and Biobehavioral Reviews*, 34(2), 224–236.
- Shore, D. S., & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research*, 14(1), 153–163.
- Siegel, S. (2010). *The contents of visual experience*. Oxford: Oxford University Press.
- Siegel, S., & Byrne, A. (2016). Rich or thin? In B. Nanay (Ed.), *Contemporary debates in the philosophy of perception*. New York: Routledge.
- Smith, B. (2015). The chemical senses. In Mohan Matthen (Ed.), *The Oxford Handbook of the philosophy of perception* (pp. 314–352). Oxford: Oxford University Press.
- Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences*, 1191, 182–200.
- Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, 1296, 31–49.
- Spence, C., Auvray, M., & Smith, B. (2014). Confusing tastes with flavours. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 247–276). Oxford: Oxford University Press.
- Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford: Oxford University Press.
- Stein, B., & Meredith, M. (1993). *The merging of the senses*. Cambridge: MIT Press.
- Stratton, G. M. (1899). The spatial harmony of touch and sight. *Mind*, 8, 492–505.
- Swanston, M., Wade, N., & Day, R. (1987). The representation of uniform motion in vision. *Perception*, 16(2), 143–159.
- Tamè, L., Azañón, E., & Longo, M. (2019). A conceptual model of tactile processing across body features of size, shape, side, and spatial location. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2019.00291>.
- Taylor, M. W., Taylor, J. L., & Seizova-Cajic, T. (2017). Muscle vibration-induced illusions: Review of contributing factors, taxonomy of illusions and user’s guide. *Multisensory Research*, 30, 25–63.
- Trommershäuser, J., Kording, K., & Landy, M. (Eds.). (2011). *Sensory cue integration*. Oxford: Oxford University Press.
- Tsakiris, M. (2010). My body in the brain: A neurocognitive model of body-ownership. *Neuropsychologia*, 48(3), 703–712.
- Tse, P. (1998). Illusory volumes from conformation. *Perception*, 27(8), 977–992.
- Tse, P. (1999). Volume completion. *Cognitive Psychology*, 39(1), 37–68.
- Tse, P. (2017). *Dynamic Volume Completion and Deformation*. *i-Perception*, 8(6), 2041669517740368.

- von Helmholtz, H. (1867/2005). *Treatise on physiological optics* (Vol. 3), New York: Dover.
- Vuilleumier, P., Valenza, N., Mayer, E., Perrig, S., & Landis, T. (1999). To see better to the left when looking more to the right: Effects of gaze direction and frames of spatial coordinates in unilateral neglect. *Journal of the International Neuropsychological Society*, 5(1), 75–82.
- Welch, R. (1978). *Perceptual modification: Adapting to altered sensory environments*. New York: Academic Press.
- Welch, R., & Warren, D. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3), 638–667.
- Whitlock, J. R. (2017). Posterior parietal cortex. *Current Biology*, 27(14), R691–R695.
- Wong, H. Y. (2017). On proprioception in action: Multimodality versus deafferentation. *Mind and Language*, 32(3), 259–282.
- Wu, Wayne. (2014). Against division: Consciousness, information and the visual streams. *Mind and Language*, 29(4), 383–406.
- Wurtz, R. (2017). Corollary discharge contributions to perceptual continuity across saccades. *Annual Review of Vision Science*, 3(1), 215–237.
- Yamamoto, S., & Kitazawa, S. (2001). Reversal of subjective temporal order due to arm crossing. *Nature Neuroscience*, 4(7), 759–765.
- Zipser, D., & Andersen, R. (1988). A back propagation programmed network that simulates the response properties of a subset of posterior parietal neurones. *Nature*, 331, 679–684.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.