

High-Level Perception and Multimodal Perception

For Heather Logue and Louise Richardson (eds.), *Purpose and Procedure in Philosophy of Perception*.

Oxford University Press.

Final revision

Dan Cavedon-Taylor, The Open University

dan.cavedon-taylor@open.ac.uk

Abstract: What is the correct procedure for determining the contents of perception? Philosophers tackling this question increasingly rely on empirically-oriented procedures in order to reach an answer. I argue that this constitutes an improvement over the armchair methodology constitutive of phenomenal contrast cases, but that there is a crucial respect in which current empirical procedures remain limited: they are unimodal in nature, wrongly treating the senses as isolatable faculties. I thus have two aims: first, to motivate a reorientation of the admissible contents debate into a multimodal framework, charting its various significances. The second is to explore whether any experimental studies of multimodal perception support a so-called Liberal (or ‘high-level’ or ‘rich’) account of perception’s admissible contents. I conclude that the McGurk effect and the ventriloquist effect are both explicable without the postulation of high-level content, but that at least one multimodal experimental paradigm may necessitate such content: the rubber hand illusion. One upshot of this argument is that Conservatives who claim that the Liberal view intolerably broadens the scope of perceptual illusions, particularly from the perspective of perceptual psychology, should pursue other arguments against that view.

I. Introduction

Two central topics in contemporary philosophy of perception are the admissible contents of perception debate and increasing recognition of perception’s multimodality. Puzzlingly, each is typically addressed in isolation from the other. Notably, the question of which properties are admissible in perceptual content is largely debated within a unimodal, visual framework. This

is surprising, since perceiving the world via only one sense tends to be a rare exception. Multimodal perception is the norm, and so philosophers interested in which properties are admissible in perceptual content owe it to themselves to explore the issue from a multimodal perspective. This paper has two goals: the first is to motivate such a reorientation of the admissible contents debate, charting its potential significances. The second is to explore the possibility that at least one empirical study of multimodal perception, the rubber-hand illusion, supports a so-called Liberal (or ‘high-level’ or ‘rich’) account of perception’s admissible contents.

One reason that Liberals should welcome the argument developed here is that it constitutes a reply to the following objection to their view: including high-level properties, like natural and artefactual kinds, in perceptual content intolerably broadens the scope of what is to count as a perceptual illusion. Alex Byrne, for instance, writes:

Visual illusions, as the object of study in the visual sciences, concern properties like shape, motion, colour, shading, orientation and the like, not properties like *being tired*, *belonging to Smith* or *being a lemon*. There is thus no immediate reason to take (visual) perceptual content to include the proposition *that o is a lemon*, and the like. (2009, p.449; see also Byrne & Siegel 2016)

If what I say here is right, then we have reason to think that perceptual content can include such high-level properties precisely on the basis that there is a perceptual illusion, the rubber-hand illusion, that, as an object of study by empirical psychology, does not concern low-level properties alone.

First, what is the admissible contents debate about? Suppose one thinks that perceptual experience has content, that it attributes properties to objects and so represents the world as being some way. Someone who held this view must take a stand on what properties can figure in perceptual content; that is, *which* ways can perception represent the world as being? Two positions can be distinguished.

According to Conservatives (Byrne 2009; Price 2009; Brogaard 2013a; Tye 1995), perceptual experience is limited to representing so-called ‘low-level’ sensibles:

Vision: colour, shape, motion, spatial location and spatial orientation.

Audition: loudness, pitch, timbre and duration.

Touch: pressure, vibration, temperature, weight, texture, shape and spatial location.

Taste: saltiness, sweetness, sourness, bitterness and umami.

Olfaction: camphorousness, pungency, floralness, muskiness and earthiness.¹

According to Liberals (Bayne 2009; Nanay 2011a and 2011b; Siegel 2006 and 2011), perceptual content can also include what have come to be called ‘high-level’ properties. A tomato may not simply be represented as red and round; it may be represented as being the kind of thing it is: *a tomato*. Similarly, tables and chairs may be represented as tables and chairs, not just 3-D expanses of colour and shape at various locations and orientations. According to the Liberal, the representation of such natural and artefactual kinds, i.e. ‘being a tomato’, ‘being a table’, etc. is not always a matter of post-perceptual inference, but can be a matter of how one literally sees objects to be.²

What about those who deny the representational account of perception? Relationalists hold that perception acquaints subjects with particular objects and/or property-instances, rather than representing such object-property complexes (Martin 2004; Campbell 2002; Brewer 2011; Fish 2009). For relationalists, the debate cannot be framed in terms of what properties perception *represents*, but should be framed in terms of what properties perception can *acquaint* one with. Here I will treat the Liberal/Conservative debate as one concerning the contents of perception. However, in assuming such representationalist terminology, I am simply adopting the terminology in which this issue has, for better or worse, so far been debated. Although some

¹ Insofar as low-level properties are typically identified with a conjunction of both the special and common sensibles of the perceptual modalities, there is some debate to be had about precisely which properties are low-level, since there is debate about whether all such properties are really represented in each modality. Consider: should we think of thermal perception as part of touch? What of nociception? The sensibles of olfaction are even less well-defined and typically fail to reflect language used by ordinary people (Kaeppler & Mueller 2013). Here, I simply list some odour terms from biochemist John Amoore’s (1977) influential classification.

² Although much of the debate between Conservatives and Liberals centres on the admissibility in perception of natural and artefactual kinds, the issue may be debated for a host of other properties too, including causal properties (Siegel 2009) as well as action-based and dispositional properties (Nanay 2011a and 2011b).

have expressed scepticism about the possibility of framing the debate in terms friendly to the relationalist (Brogaard & Chomanski 2015; Cavedon-Taylor 2015), for the purposes of this paper I will assume that ‘represents’ and its cognates can be treated by relationalists as ‘acquaints one with’ and its cognates (see Siegel 2006; Silins 2013; Reiland 2014; Logue 2013; Bengson 2013; Nanay 2013; Toribio 2018; Di Bona 2017). Indeed, a number of relationalists have weighed in on the debate, some in favour of Liberalism (Fish 2009, 2013; Johnston 2006), others in favour of Conservatism (Martin 2010) and some arguing that there is no fact of the matter (Logue 2013).

Why does this debate matter? First, consider the low-level/high-level distinction. It is notoriously difficult to precisely distinguish low-level properties from high-level ones. It might be that the best we can do is say that low-level properties are the ones that we all agree are perceived, while high-level properties are the ones that we debate about (Logue 2013, pp.1-2; see also Siegel 2006 and Macpherson 2012; but see Martin 2010). But if the low-level/high-level distinction is this ambiguous, one might wonder why anyone should care about the debate. Some recent arguments conclude that it matters for the epistemology (Siegel 2006) and metaphysics (Cavedon-Taylor 2015) of perception, along with the question of perception’s penetrability by cognition (Macpherson 2012) and the possibility of perceptually-based action (Nanay 2013). But the debate impacts issues further afield as well. For instance, some hold that we can know others’ minds via perception (McDowell 1982; McNeill 2012; Cassam 2007) and some claim that we can know what is the morally right thing to do via perception (Audi 2013; Cuneo 2003; McBrayer 2010). It is not always clear whether such philosophers think of mental and moral properties as represented in perception, rather than in perceptual belief (though see Werner 2016; Brogaard 2016; Toribio 2018; and Newen 2017). But insofar as any philosopher would wish to articulate a fully *perceptual* epistemology of others’ minds and moral knowledge, they would thereby require the truth of Liberalism. Getting the right answer to the question of perception’s admissible contents has the potential to impact a number of issues in philosophy of mind, epistemology and moral philosophy.

I proceed as follows: in section II, motivate the importance of transposing the admissible contents debate into the multimodal framework. In section III, I explore two widely-studied multimodal illusions, the McGurk effect and the ventriloquist effect, and conclude that neither support Liberalism. In section IV, I explore a multimodal illusion that on balance does, the rubber hand-illusion. In section V, I consider objections. Section VI concludes.

II. Beyond Vision and the Unimodal Paradigm

Philosophers of perception are increasingly aware of the distorting effect that ‘visuocentric’ bias has had on their field. In theorising about perception in general, it is becoming clear that we shouldn’t restrict our focus to vision in particular. To take just one example, consider the relation between our senses and space. Vision is a spatial sense, but not in the same way that touch is (Martin 1992). What we see, we see to be located in space at a certain distance and direction from our own body; not so when we touch those things and matters are similarly complex in the case of audition (Nudds 2009). Olfaction, by further contrast, has only very minimal spatial content. What we smell, we typically smell to be located an undifferentiated location of ‘here’ rather than smelling those things to be distally located (Batty 2011; Matthen 2005). Concerning the relationship between our senses and space, we will have the wrong view of that relationship if we generalise from vision alone.

Visuocentricism, while distorting, is sometimes said to be understandable (O’Callaghan 2011a). We often don’t pay attention to how things in our immediate environment feel (e.g., your feet in your socks), how they sound (e.g., the tick of the clock), how they taste (e.g., the lunch you wolfed down before a meeting), and so on. Whereas with vision, things are constantly there, immediately salient to one. All it takes is being awake! But visuocentricism, while understandable, remains problematic. The senses are hugely diverse, both in number and nature (see the papers in Macpherson 2011 and Stokes et al. 2014). We will have a skewed account of the senses if we try to shoe-horn all of them into a visual mould.

One way that we can break with the visuocentric perspective is to simply ask questions about non-visual modalities that we have historically asked about vision (Di Bona 2017). But this procedure fails to correct the chief wrong of visuocentricism, and a more radical break from it is available. For what is objectionable about visuocentricism is not simply that there are senses it overlooks, but that it is symptomatic of a unimodal outlook which misconstrues the nature of perception, both at the experiential and sub-personal levels. In our typical, everyday interactions with the world, it is almost inevitable that more than one sense will be stimulated at a time. And, very often, more than one sense is stimulated by one and the same object. For instance, the things we see are often the things we smell, as when we savor the aroma of coffee before a first sip; the things we touch are sometimes the things we hear, as when we drum our

fingers on a table; and the sound of another's voice is something one both sees and hears to originate from their lips. Experientially, this is not always a matter of co-conscious awareness; that is, mere simultaneous representation of the world via one sense modality and also via another. Typical episodes of multimodal perception involve a stronger degree of unification than mere temporal simultaneity of distinct unimodal experiences (O'Callaghan 2015; Bayne 2014). Although some 'minimally' multimodal experiences may be characterised in this way, e.g., when hearing voices in the café and also feeling the hardness of one's coffee cup, other multimodal experiences cannot. When we talk face-to-face my experience of your lips is a single audio-visual one rather than an overall experience in which I perceive your lips twice over, once via vision and again via audition. All this is mirrored in sensory processing, where the binding of information from distinct senses, and which creates a single "multisensory event" (Vatakis & Spence 2007, p.744), is exceptionally well-studied by cognitive psychology (Bedford 2001; Calvert et al. 2004; Driver & Noesselt, 2008; Welch & Warren 1980).

So the true, multimodal nature of perception is not something that will be at all reflected in the outcomes of a procedure that simply asks questions of non-visual senses in isolation traditionally asked of vision (see also O'Callaghan 2016). Moreover, in following such a procedure, we risk simply reorienting existing debates around one of those other senses, all the while continuing to ignore how perception actually functions for us, i.e. multimodally. Indeed, the Liberal/Conservative debate is increasingly audition-centred (see Brogaard 2018; Di Bona 2017; Nes 2016; and Reiland 2015). Multimodal perception remains overlooked, never mind olfaction, touch and gustation.

Is the neglect of multimodal perception justifiable in the case of the Liberal/Conservative debate? Perhaps it is. Indeed, there are reasons for thinking that unimodalism is the best strategy to pursue here.

First, unimodal perception is, by its nature, simpler than multimodal perception. By sticking with a unimodal approach, we simplify the Liberal/Conservative debate, making the issue more tractable than it would be from a multimodal perspective. The question of the admissible contents of multimodal perception, being more complex and challenging than its unimodal cousin, can then be usefully informed by what we have discovered in the simpler, unimodal cases.

Second, addressing the Liberal/Conservative debate from a multimodal perspective appears to be a redundant procedure. For once we've examined this issue from the perspective of the individual senses in isolation, we then seem to have our answer as to whether there is high-level content in multimodal perception. No separate multimodal investigation is required. For instance, say the outcome of a completed unimodal investigation is that vision and audition have high-level content, while olfaction, touch and gustation do not. It then follows that for any multimodal experience in the traditional five senses (of a suitably mature perceiver) we have an answer as to whether high-level content is present. A multimodal visual-olfactory experience of a spraying skunk would have high-level content due to vision's having high-level content: one would see the skunk to be a skunk, while smelling only a pungent odour. By contrast, a multimodal tactile-olfactory experience of touching and smelling a rose in the dark would have only low-level content: one would feel there to be an object which is soft in some places, spikey in others, and which is vaguely floral-smelling, but the experience would be incapable of representing the object to be a rose *per se*. Thus, framing the admissible contents of perception debate in multimodal terms looks unable to tell us anything that we wouldn't already have gleaned from the results of a completed unimodal procedure.

The difficulty with these motivations for unimodalism is that they presuppose multimodal perception to be a mere conjunctive blend of the contents of visual, auditory, tactual, etc. experiences. Granted, this may sometimes be the case. For instance, Frederique de Vignemont (2014) usefully distinguishes between *additive* and *integrative* multimodality. Multimodal perception is *additive* when it combines experiences in two individual modalities and when those experiences concern different properties of the perceived object. For instance, I may see a car's wheel spin on the tarmac and also hear its high-pitched chirp as it drives away. Multimodal perception is *integrative* when, by contrast, at least one instance of the same common sensible is represented by each modality. For instance, in grasping a ball, one sees and its feels its sphericity. In additive and integrative cases, multimodal perception is a conjunctive blending of the properties represented by the component experiences, either non-redundantly or redundantly. But one may further distinguish what Robert Briscoe (2016) calls 'generative' multimodal perception and O'Callaghan (2016) calls 'constitutive' multisensory perception. In these cases, the properties represented multimodally do not appear in one or other of the multimodal experience's component experiences. A key example is flavour perception, which combines gustation, olfaction as well as tactual and trigeminal sensations to produce experiences of novel and emergent properties, e.g., 'cherryiness,' 'peatyness,' 'astringency', etc.

and which are not themselves found in any of flavour's component experiences, including taste (Smith 2015).

Generative/constitutive multimodal experiences are highly significant for the admissible contents debate. For suppose that a completed unimodal investigation yielded the result that high-level properties are not represented in, e.g., unimodal visual experience, nor unimodal auditory experience, nor unimodal tactual experience, and so on for the other senses. From such a unimodal perspective, this is a situation in which Conservatism has been proved true and Liberalism proved false. But from a multimodal perspective, this conclusion does not follow; given the existence of generative/constitutive multimodal experiences, it would remain open that high-level content appears exclusively in multimodal perception in the way that flavour properties appear exclusively in flavour perception. Thus, the two above motivations for unimodalism not only wrongly assume that multimodal perceptual experience is simply the sum of its modality-specific parts, but by doing so they stack the deck, albeit modestly, in favour of Conservatism.³ Generative/constitutive multimodal perception should motivate Liberals in particular to take very seriously the idea of pursuing the contents of perception debate in a multimodal framework.

Importantly, the above is not to say that multimodal perception, by its nature, favours Liberalism. It would be a mistake, given the existence of generative/constitutive multimodality, to reason from the putative absence of high-level content in unimodal perception to its absence in multimodal perception. But it would similarly be a mistake to reason that the additional complexity of multimodal perception supplies reason to think that high-level content will be found there. Yet this latter reasoning seems implicit in influential work on high-level content. Consider the follow remarks of Susanna Siegel's:

[The debate] is less interesting when applied to certain multimodal experiences, compared with purely visual experiences. For instance, the idea that... a visual and kinesthetic experience represents that your hand is pulling open a zipper... is somewhat

³ Is flavour perception just high-level multimodal perception? Not obviously. Due to the ambiguous nature of the high-/low-level distinction, Conservatives may plausibly deny that flavour properties are high-level and may claim that such properties ought to be considered low-level for the flavour modality. After all, such properties pass the test of being ones that we all agree are represented by flavour.

less surprising than corresponding theses about what visual experience represents.
(2011, pp.25-26)

Siegel goes on to say that focusing on unimodal cases, vision in particular, “will maximise the punch” (p.26) of Liberalism. If what I have said above is correct, then the reverse is true, since this procedure characterises perception in an artificial manner that has little bearing on how we actually perceive. In addition, consider Siegel’s claim that unimodalism is the more interesting perspective from which to examine the admissible contents debate, on the implied basis that it is already intuitive that high-level content is present in multimodal perception.

Does the multimodality of perception *a priori* favour Liberalism in this way? Liberals such as Siegel should refrain from being so cavalier. Conservatives deny that there is high-level content in perception when looking at a zipper; it looks to be a small, grey object with a certain shape and which bears visually perceptible spatial relations to other objects. Will Conservatives find it intuitive that any high-level property is perceptually attributed to the zipper simply by virtue of the subject’s also having *touched* the zipper? Surely not.⁴ So while a unimodal perspective on the debate may favour Conservatism over Liberalism, it is not the case that the situation is reversed once one adopts a multimodal perspective. While Liberals owe it to themselves to explore the possibility that there is high-level content in multimodal perception, thinking beyond the unimodal perspective, they should not simply assume that high-level content will be found there.

To sum up so far, we have good reasons to distrust the results of any unimodal investigation into the admissible contents of perception. At best, that approach will be artificial. At worst, it will be necessarily incomplete in ways that potentially favour Conservatism. But on the other hand, we should not assume before investigation that there is high-level content in multimodal perception; this is a further, substantive issue yet to be settled.

⁴ The high-level property allegedly represented here, according to Siegel, is ‘pulling’, a causal property. For simplicity I focus on the kind property ‘being a zipper.’ Conservatives are no less apt to deny that the causal property is perceptually represented merely by virtue of touching the zipper simultaneous with seeing it.

Finally, it is worth noting the role of empirical findings in the debate. The initial arguments developed in favour of Liberalism took an armchair form insofar as they relied upon intuitions about so-called phenomenal contrast cases, particularly the contrast between expert and novice experiences. Simplifying, there is, intuitively, a contrast in phenomenal character between a perceiver's perceptual experience of a type of object before and after acquiring a disposition to recognise objects of that type. For instance, seeing pine trees in a forest for the first time seems experientially different from seeing them once one learns to recognise such trees. The best explanation for this experiential difference was thought to be that acquiring a disposition to recognise pines is a matter of the property 'being a pine' coming to figure in the content of one's perceptual (visual) experience of them. Gaining a disposition to recognise pines, "you can spot the pine trees immediately: they become visually salient to you." (Siegel 2011, p.100)

The difficulty with phenomenal contrasts is now well-recognised: there are alternative explanations for the phenomena that appear no less plausible. Maybe the recognitional disposition manifests itself in perception by modulating one's attention to previously experienced low-level properties rather than producing the perceptual representation an additional, high-level property (Block 2014; Nanay 2011b; Briscoe 2015; Price 2009; Fish 2013; Prinz 2013). Alternatively, the phenomenal change may be at the level of cognitive, not perceptual phenomenology (Price 2009; Tye 1995). Another option is that the change takes place at a level between perception and cognition, one involving a state called a 'seeming' (Brogaard 2013b; Lyons 2005; Reiland 2014). Phenomenal contrast cases were introduced in order to counter the stalemate that would likely arise between Liberals and Conservatives were they to attempt to settle the debate via a direct appeal to introspection (Siegel 2007). Hence, contrasting mature, 'expert' perceptual experience with 'novice' perceptual experiences, i.e. where no high-level properties are represented, seemed necessary. But as things have panned out, this method too seems to have produced an impasse.

One of the most significant developments in this debate is the turn by Liberals to clinical disorders and experimentally-observed effects in order to support their view, including: visual associative agnosia (Bayne 2009), unilateral neglect (Nanay 2012), adaptation after-effects (Block 2014; Fish 2013; Di Bona 2017) and rapid categorisation or gist-perception (Fish 2013; Bayne 2016). The empirically-grounded arguments for Liberalism that have resulted, some of which retain aspects of the method of phenomenal contrast, have been thought to represent an

improvement over purely armchair ones.⁵ Still, their success remains limited insofar as they assume a unimodal perspective. To better support their view, Liberals should look to results from not merely the psychology of perception, but to the psychology of multimodal perception.

III. Multimodal Perception and the Hunt for High-Level Content

The standard way in which psychology has sought to study multimodal perception is via perceptual illusions.⁶ I will concentrate on three of the most widely-studied examples:

- (i) the McGurk effect;
- (ii) the ventriloquist effect, and
- (iii) the rubber hand illusion.⁷

In this section, I will argue that (i) and (ii) fail to support Liberalism. Then, in section IV, I will present considerations in favour of high-level perceptual content being a necessary background condition of (iii), the rubber hand illusion.

III.i The McGurk Effect

The McGurk Effect (McGurk & MacDonald 1976) is an example of an auditory-visual multimodal illusion. In one version of the illusion, an audio-recording of a tokening of the /ba/ phoneme, when paired with a visual recording of a person moving their mouth to produce the /ga/ phoneme, results in an overall multimodal, audio-visual perception of a tokening of the /da/ phoneme.

⁵ These arguments still leave room for dissent (see Briscoe 2015; Brogaard 2013a; Raftopoulos 2015; Helton 2016). Fish (this volume) discusses how scientific claims are susceptible to philosophical re-interpretation.

⁶ The definition, and existence, of perceptual illusions (in the philosopher's sense) is increasingly debated (Travis 2014; Kalderon 2011; Fish 2009; Brewer 2011). I do not wish to prejudice those issues. If it turns out that we should say that subject's perceptual responses to stimuli in these experiments are veridical, nothing I say here should be affected.

⁷ This is just a small sample of multimodal illusions. I ignore here some particularly well-known examples which, like the stream-bounce illusion (Sekuler et al. 1997) and double-flash illusion (Shams et al. 2000), seem explicable entirely in terms of the perception of low-level properties, e.g., shape and motion, along with beeping tones and other simple auditory stimuli.

How might Liberalism figure in the explanation of this illusion? The effect is commonly described in the psychological literature as a matter of speech perception. However, one should not infer from this the effect necessitates hearing high-level properties; in this case, semantic properties of speech, i.e. the meaning of words. Indeed, the effect involves no such thing. Phonological features of a language, like phonemes, are to be regarded as the audible ‘building blocks’ of language. Phonemes like /ba/, /ga/ and /da/ are semantically relevant, but are not semantically significant by themselves. Perceptual experiences of them are not high-level (O’Callaghan 2011b, p.801).

However, the Liberal may nonetheless argue that the effect depends upon vision’s having high-level content. In particular, they may claim that the effect requires vision to represent the object making the /ga/ movements to be *lips* or *a mouth*. The McGurk effect is an instance of facial perception no less than speech perception. So perhaps it is here that one can locate high-level content.⁸ This strategy claims that the representation of high-level properties is a necessary *background* condition for the illusion, granted that the property illusorily represented in multimodal perception, the /da/ phoneme, is low-level.

Experimental results tell against this account of the effect. Rosenblum & Saldana (1996) induced the McGurk effect in subjects who were entirely unaware that they were looking at faces and facial features. In their experiment, the visual stimuli were dot-lights, e.g., of the sort used in motion-capture and CGI, placed on a darkened subject’s lips, teeth, tongue, chin, and cheeks. In the video-display, only the dots and their motions were visible against a black background; the face and its features could not be seen. When the image of the dot-points was static, none of the participants correctly guessed that they were looking at a mouth. After the experiment was concluded, most participants correctly identified the stimulus, but nearly a quarter guessed incorrectly, showing themselves to be unaware that they were looking at a face. (They guessed the darkened stimulus to be ‘an owl’, ‘a butterfly’, and ‘sheep’.) Crucially, questionnaires filled out by these subjects indicated that the McGurk effect had nonetheless occurred: such subjects had a multimodal, auditory-visual experience of the /da/ phoneme, but no representation in vision (or belief) of a face—only moving points of light. It would thus seem that the McGurk effect does not depend on one’s representing the visual stimulus in a

⁸ There is room for debate about whether seeing faces *as such* is a high-level perceptual phenomenon (Brogaard & Chomanski 2015) or a low-level one (Lyons 2009).

high-level way, e.g., as a face. Rather, it simply depends on the visual representation of spatial-motion properties. Such properties are canonically low-level.

III.ii The Ventriloquist Effect

A second multimodal illusion which features prominently in the experimental literature is the ventriloquist effect (Howard & Templeton 1966), a familiar one in which the sound of a ventriloquist's voice at one location, x , is heard to come from a distinct location, y , by the mouth of their 'dummy.'

Contrary to popular wisdom, this effect does not rely on the ventriloquist 'throwing' their voice. Rather, it involves what psychologists call the 'visual capture of sound.' When the illusion is successful, the dummy's mouth 'captures' the sound of the ventriloquist's voice due, in part, to temporal synchrony between movements of the dummy's mouth and the production of speech by the ventriloquist. The two become crossmodally 'feature-bound' or 'phenomenally fused', at the location of the dummy.

Does this illusion favour Liberalism? On the face of it, one might think the chances are no better than for the McGurk effect. After all, this is an illusion regarding a *spatial* property, i.e. the location of a sound, and spatial properties are paradigm examples of low-level properties. But recall the second route to Liberalism explored in relation to the McGurk Effect: Liberals may claim that the perceptual representation of high-level properties is a necessary, background condition for the illusion and its phenomenal fusion to occur, granted that the property illusory represented at the multimodal level, the sound's spatial location, is itself low-level.

How might this idea be developed? The reasoning might go as follows: the explanation for why visually experiencing a dummy's mouth to move in time with a ventriloquist's speech results in phenomenal fusion is that this is a case of (visual) *facial* perception and (auditory) *speech* perception. That is, the ventriloquist's voice over here is heard to be located over there with the dummy's mouth only because the voice is auditorily represented to have the high-level property 'being speech' (i.e. it is not simply heard as a collection of meaningless tones), while the dummy's mouth is visually represented to have the high-level property 'being a mouth' (i.e. it is not merely a moving object having a certain colour and shape). The two stimuli become grouped in perception only in virtue of them first being visually and auditorily represented as being the kinds of things they are, kinds which naturally go together: a mouth and speech. And

so perception represents such high-level properties as ‘being speech’ and ‘being a mouth’ as a background condition for the effect.

As promising as the above idea might seem, the empirical literature suggests it to be on the wrong track. For one, the visual capture of sound occurs when auditory stimuli are mere tones and pulses (rather than speech) and are paired with visual stimuli that are simple LED lights (rather than mouths) (Radeau & Bertelson 1976; Bertelson & Radeau 1981). Indeed, psychologists Argiro Vatakis and Charles Spence (2007, p.745) lament the fact that the ventriloquist effect tends to be studied using “arbitrary, or nonmeaningful, combinations of stimuli, such as flashing lights and brief tones” rather than “informationally rich” everyday objects. But the fact that psychologists can induce the ventriloquist effect with non-meaningful audio-visual stimuli is good news for Conservatives. It casts significant doubt on high-level content being a background condition for the illusion.

A potential reply on behalf of the Liberal would be to grant that high-level content isn’t necessary for the ventriloquist effect in general, but that in conditions when the auditory stimulus *is* speech, a mouth-like visual stimulus is necessary to capture it, suggesting that representing something as a mouth is required for visual capture of speech. However, the empirical literature doesn’t bear this out either. Radeau & Bertelson (1977) found visual capture of speech occurred in situations in which the visual stimulus was lights that flashed in time with speech.

So far we have seen reason to doubt both that the ventriloquist effect involves either the multimodal representation of a high-level property or that high-level perception in unimodal perception is a background condition for the effect. But there is another potential explanatory role for high-level content here, which is as an *enhancer* of multimodal, audio-visual fusion. On this suggestion, high-level content is not necessary for the ventriloquist effect; rather, *when* present, it aids or strengthens the effect.

This conjecture is empirically supported. Consider two experiments in one of the earliest studies of visual capture of sound (Jackson 1953). In one experiment, the visual stimulus was a kettle, an artefact kind, and the audio stimulus was a kettle’s whistle, a sound associated with an artefact kind. In another, the stimuli were bells and flashes, non-meaningful stimuli whose pairing is rather more arbitrary. In both, the visual stimuli successfully captured the auditory

stimuli; asked to point to the location of the sound, subjects reliably pointed in the direction of the visual stimulus as opposed to the actual origin of the sound. However, in comparing the results of the pointing tasks across the two experiments, participants appeared to bias audio-visual fusion when the stimuli were kettles and whistles. Asked to point to where the sound (whistle) was coming from, subjects pointed closer to the visual stimulus (kettle) than they did in the flash/bell conditions.

This enhanced strength of response towards familiar, everyday objects versus non-meaningful stimuli in the case of the ventriloquist effect is robust (see Vatakis & Spence 2007). The Liberal line of argument under consideration says that the perceptual systems of subjects in these experiments were biased towards fusing the whistle sound with the kettle because they saw and heard the stimuli in that very way, i.e. *as* kettles and *as* whistles. Granted, had they failed to represent the stimuli in a high-level way, there still would have been a degree of multimodal, audio-visual fusion, as shown by the flash/bell experiment.

Conservatives, however, can give equally plausible explanations of the bias. Everyday stimuli such as kettles and whistles can be considered ‘informationally rich’ (Vatakis & Spence 2007), relative to non-meaningful stimuli like flashes and tones. This relative richness can be understood as a greater degree of complexity in terms of low-level properties; it need not be analysed as the everyday objects being perceptually represented in Jackson’s experiments in a high-level way. Kettles, whistles, mouths, and speech tend to be more intricate and complex with respect to their shapes, colours, pitches, etc. than do mere flashing lights and beeping tones. Audio-visual fusion may not bias everyday stimuli by virtue of perceptually representing them as having high-level properties, but simply by virtue of such stimuli having a greater number of low-level properties (see also Deroy 2015, section 4.2).

How is this conjecture on behalf of the Conservative explanatory? For one, it fits with discussion in the psychological literature of the role of bottom-up factors in effecting audio-visual fusion. For as well as spatial and temporal coincidence playing a key role, another important factor is dimensions of what Spence (2011, p. 972) calls *synaesthetic congruency* between stimuli, where this is precisely a matter of “correspondences between more basic stimulus features (e.g., pitch, lightness, brightness, size) in different modalities.” Indeed, there are well-known audio-visual synaesthetic correspondences between such low-level properties as, e.g., hue and pitch, brightness and loudness, and brightness and high pitches. So, when stimuli used to test audio-

visual fusion are relatively numerous in terms of their low-level properties, as everyday objects are, one should naturally expect a greater degree of fusion than when simpler stimuli like tones and flashes are used. Everyday objects, by virtue of having a greater number of low-level properties, present greater opportunities for synaesthetic congruency. And where there is greater congruency between audio-visual stimuli, it is well-known that one's perceptual system is thereby more likely to fuse together such stimuli (Welch & Warren 1980). In addition, since it is plausible that subjects have background beliefs linking whistles to kettles, but not beeps to flashes, Conservatives may argue that the bias subjects exhibit in such experiments towards fusing information from ordinary objects simply reflects a post-perceptual decision bias as to the whistle's spatial location, rather than revealing anything about their actual perceptual experiences (see Bertelson & de Gelder 2004, pp.151-156).

In sum: we have reason to think that the ventriloquist effect is an instance of a broader audio-visual fusion effect that neither produces, nor relies upon, nor is definitively enhanced by, the perceptual representation of high-level properties.

IV. The High-Level Contents of the Rubber Hand Illusion

The rubber-hand illusion (RHI) (Botvinick & Cohen 1998) is a multimodal illusion in which tactual sensation is referred to a false, rubber hand. While in the ventriloquist effect we find audio-visual fusion, RHI involves a fusion of vision and touch. Subjects have one of their hands placed out of view, and a fake appendage placed before them where their hand would ordinarily be. Subjects are then instructed to visually fixate on the rubber hand while it is stroked with a paintbrush. At the same time, their out-of-sight hand receives matching cutaneous stimulation. Introspective reports and behavioural indicators of proprioception having 'drifted' towards the rubber hand suggest that subjects begin to experience (non-veridically) the stimulating of the rubber hand; the sensations actually occurring on their hand become spatially referred to the fake one. As the title of Botvinick and Cohen's study aptly puts it: the rubber-hand *feels* the touch that the eye sees. As in the case of both the McGurk and ventriloquist effects, the properties represented here multimodally do not seem candidate high-level ones, as again the illusion concerns a non-veridical spatial experience. It remains to be seen whether high-level perceptual contents are, however, a necessary background condition for the illusion to occur.

IV.i RHI on Liberalism and Conservatism

RHI can be regarded as an illusion in two interrelated ways. First, RHI involves an illusion of bodily ownership: one proprioceptively senses (non-veridically) the rubber hand from the inside. Second, RHI involves a tactual illusion: the cutaneous stimulation of one's actual hand is felt (non-veridically) to be located on the rubber hand. Indeed, Botvinick and Cohen (1998, p.756) describe the illusion as revealing "a three-way interaction between vision, touch and proprioception." But one might simplify Botvinick and Cohen's description of RHI, referring to it as revealing mere visual-tactual interaction, since proprioception is plausibly necessary for touch to begin with, if not partially constitutive of it (Martin 1992).⁹ From this perspective, it is no surprise that RHI is an illusion of both bodily ownership and tactual sensation. For, if proprioception is partly constitutive of touch, then we can picture the proprioceptive illusion of bodily ownership as the factor that causally explains the (non-veridical) referral of tactual-cutaneous sensation to the rubber hand.

In approaching the question of whether there is high-level perception in RHI in terms of it being a background condition, let us begin by focussing on the idea that RHI involves an illusion of bodily ownership. A natural way to unpack this idea is to say that RHI involves a non-veridical, proprioceptive representation of the rubber hand as having the property 'mine.' At first glance, this might seem like a candidate high-level property. But that impression should immediately be dispelled. The property 'being mine' is best thought of as a low-level property of proprioception, given that it is ubiquitous, if not definitive, of the modality. It seems precisely the kind of property that, when proprioception is construed as a sense modality (Macpherson 2011; Schwenkler 2013), all would agree upon proprioception as representing.

Let us switch focus now from the content of proprioception in RHI to the content of touch. It is difficult to see on what grounds a Liberal could argue that high-level tactual content is a necessary background condition for RHI. This would be to claim that RHI requires subjects to tactually represent the object stimulating their out of sight hand to be *a brush* or *a paintbrush*, rather than simply representing pressure, textures, vibration, etc. on their skin. This is a fairly

⁹ RHI might be regarded as multimodal along several dimensions on the basis that touch itself is not a single sense but is multimodal in its own right. Touch combines experiences of pressure, shape and vibration, as well as temperature, and relies not only on receptors in the skin but also those located in muscles, joints and tendons; touch also seems to lack a dedicated sense organ (see Fulkerson 2016, section 2).

implausible suggestion, and so reflection on the content of touch in RHI also fails to motivate the existence of high-level content.

What of the *visual* component of RHI? Here the case for Liberalism begins to look more plausible. To see why, let us contrast how Conservatives and Liberals are likely to construe the visual content of subjects undergoing RHI.

On a Conservative account of the content of vision in RHI, subjects visually represent the rubber hand as having a certain colour, shape, size, spatial location and spatial orientation. In short: it looks to be a wavy-shaped object of a certain colour. This content is likely to be accurate, barring abnormal experimental conditions. Thus, according to Conservatives, the content of vision in RHI is veridical: although vision is a constitutive element of RHI's *multimodality*, RHI is an *illusion* in proprioceptive and tactual content alone.

Liberals, by contrast, are apt to claim that the stimulus rubber object is represented by vision, erroneously, as *a hand*. That is, Liberals are likely to think that the rubber hand is visually represented as not merely an object with a certain size, shape and colour, but as an object which has the high-level property 'being a hand.' The rubber appendage does not have this property—it's not a hand. So, RHI, on Liberalism, is a threefold illusion with misrepresentation in (high-level) visual content as well as in (low-level) proprioceptive content and in (low-level) tactual content.

Putting this together, the crucial matter is that the Liberal thinks of RHI as a multimodal perceptual illusion in which an object is non-veridically seen/proprioceived/tactually-felt to have the property 'my hand.' Vision supplies the 'hand' part of this complex, erroneous content, while proprioception and touch supply the 'mineness.' But for conservatives, RHI is at most a twofold illusion in proprioceptive and tactual content alone. In RHI, on this latter view, nothing is amiss from vision's perspective. In giving conflicting accounts of the extent of misperception in RHI, Conservatives and Liberals cannot both be right. In the remainder of section VI, I shall show why the Liberal account is to be preferred.

IV.ii RHI and the Liberal's Conjecture

If proprioception drives the referral of tactual sensation from one's actual hand to the rubber hand, then what, in turn, drives that initial, proprioceptive referral? The answer, everyone

agrees, is vision: seeing (not feeling) the rubber hand to be stroked over there, while at the same time feeling (not seeing) one's own hand to be stroked over here, constitutes a conflict of spatial information, similar to the audio-visual conflict in the ventriloquist effect. As often occurs with conflicting spatial information across two or more senses, if one of those senses is vision, then that is what tends to dominate.

Notice that the above sketch explains the tactual-proprioceptive referral wholly in terms of perceptually low-level properties, i.e. conflicting spatial information generated by temporally synchronous visuo-cutaneous stimulation. But the Liberal should insist that there is an additional explanatory factor at work. Given that they are apt to think of RHI as involving visual misrepresentation of the rubber object as being an actual hand, they surely ought to exploit this as a partial explanation of the tactual-proprioceptive referral. That is, the Liberal should be inclined to say that the tactual-proprioceptive referral occurs, in part, because the rubber appendage is visually represented as being a hand (in addition to its being represented as having the low-level properties just mentioned). Why would that be explanatory? The answer is that if the property 'being a hand' is visually attributed to the rubber appendage, then the visual system would be classifying the stimulus as a body part, something that is apt to be proprioceptively perceived and which is a suitable object of tactual-proprioceptive experience. Thus, the candidacy of the rubber appendage for tactual-proprioceptive experience is a product of it being seen in a high-level way. Call this 'the Liberal conjecture.'

It is important at this point to note that a number of bottom-up factors, which involve perceptually low-level properties, are necessary for RHI. These don't simply include the synchronous tactual stimulation of one's actual hand and visual experience of the fake appendage receiving spatially competing stimulation. There must also be congruency between the subject's hand and the rubber appendage in terms of rotation and orientation (Farnè et al. 2000) and there must be left-hand/left-rubber-hand congruency (or right-hand/right-rubber-hand congruency, as the case may be) (Tsakiris & Haggard 2005). Unless these perceptually low-level conditions are in place, RHI fails. By contrast, colour is a low-level property that matters little, since RHI is not constrained by whether the rubber hand matches subjects' skin tone (Farmer et al. 2012). Thus, the Liberal conjecture is not that visually representing the rubber stimulus (erroneously) to be a hand, even while receiving synchronous visuo-cutaneous stimulation from the experimenter, is *sufficient* for RHI. The conjecture is only that it is a

necessary one, along with the perceptual representation of the above-mentioned low-level properties.

IV.iii Empirical Support for the Liberal's Conjecture

The Liberal's account of the extent of misperception in RHI is that there is misperception not just in tactual-proprioceptive content, but in visual content as well—the rubber stimulus *looks* to be a hand—something the Conservative will deny. The Liberal conjecture being explored here is that this is necessary to explain why the stimuli object is then felt (non-veridically) to be one's own hand. On this view, high-level visual content is a necessary background condition for the illusion.

Naturally, Conservatives will deny the Liberal's conjecture and will insist that only the perception of low-level properties is fully sufficient for RHI. Worryingly for the Liberal, early studies of RHI suggest that Conservatives may get their wish. Indeed, one of Botvinick & Cohen's (1998, p.756) initial claims about RHI was that “intermodal matching” is a “sufficient” condition for the tactual-proprioceptive referral; what they likely have in mind is synchrony in terms of visuo-cutaneous stimulation, a perceptually low-level phenomenon which, if fully sufficient for RHI, would leave no room for the Liberal conjecture.

But the idea that RHI can be explained solely in terms of bottom-up, low-level factors, is now widely rejected. More recent, experimental results suggest that RHI requires top-down influences and that bottom-up, stimulus-driven cues, while necessary, are insufficient (De Vignemont, et al. 2006; Haans et al. 2008; Tsakiris & Haggard 2005; Tsakiris et al. 2008; Tsakiris et al. 2010). A recent review is emphatic on this matter:

Converging evidence from RHI studies, studies on visuo-tactile extinction on neuropsychological patients, and neurophysiological studies on monkeys suggests that correlated multisensory stimulation and spatial proximity are *necessary but not sufficient* for the integration of a visual stimulus to peripersonal space or for the experience of ownership during the RHI. (Tsakiris 2010, p.706, emphasis my own)

This appears to be good news for the Liberal, since this is exactly what they are likely to claim. They will likely say that RHI requires something extra, over and above the representation of

low-level properties, and they say precisely what that something is: visually representing the instantiation of the natural kind ‘being a hand.’¹⁰

This is only to suggest that the Liberal may be on the right track. Can one find further support for the Liberal conjecture? One way to settle the matter more decisively would be to consider a prediction that the conjecture generates: when the physical stimulus used to induce the illusion is not apt to be seen to be a hand, there will be no RHI. Strikingly, vindication of this prediction can be found in three recent studies of RHI.

First, in Tsakiris & Haggard (2005), RHI failed when the visual stimulus was a wooden stick. Similarly, Kalckert et al. (2019) failed to induced RHI using a balloon. Clearly, neither stimulus is a good candidate for being visually represented as a hand. So, from the perspective of the Liberal, it is entirely unsurprising that such studies would fail to induce RHI: sticks and balloons, unlike hands, aren’t represented by vision as falling under any kind whose instances are suitable objects for being proprioceptively sensed. The fact that tactual-proprioceptive referral failed in such conditions is as exactly as the Liberal conjecture predicts.

Third, in the experiments conducted by Tsakiris et al. (2010), a number of stimuli were used, beginning with a wooden plank, and progressively modifying the stimulus’ shape across experiments so that it became more hand-like (see Figure 1). Crucially, of the stimuli depicted in Figure 1, only stimulus 5 induced RHI. Not even stimulus 4 was sufficient. Again, the Liberal has an explanation: only stimulus 5 was seen by subjects *to be a hand*. This is a plausible since, on the face of it, only stimulus 5 is uncontroversially visually passable, without further contextual information, e.g., jewellery, as being a hand.¹¹ No surprise, the Liberal will think:

¹⁰ Liberals need not, and probably should not, say that this is the only top-down factor. The aforementioned psychologists who claim that top-down factors are necessary for RHI typically have in mind the subject’s body schema. It should be stressed that this need not be at odds with the idea that RHI involves seeing the rubber stimulus to be a hand. Indeed the two claims may well be complimentary insofar as seeing the rubber stimulus in this high-level way could be responsible for schema-incorporation, or vice versa. I leave exploration of this idea for another occasion.

¹¹ It is plausible that with further contextual information stimulus 4 would be seen to be a hand, e.g., if jewellery and/or a watch were appropriately placed on it. Also, one is perhaps more likely to think that stimuli 3 and 4 are visually passable as hands when the stimuli are lined up together and seen concurrently with stimulus 5, as in the figure. But this was not the viewing condition of subjects in these experiments.

only stimulus 5 could induce RHI, since only that stimulus, in being apt to be seen to be a hand, could confer vision with content sufficient to fool tactual and proprioceptive awareness.

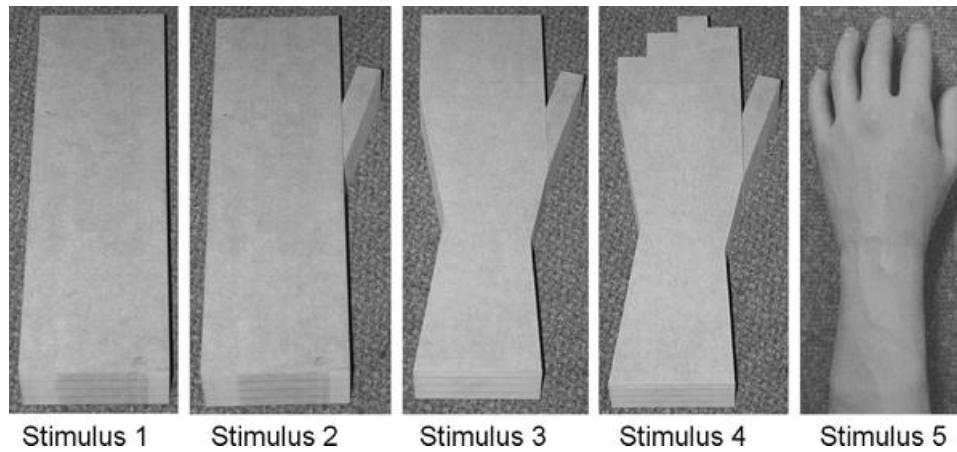


Figure 1. Stimuli used in Tsakiris et al. (2010)

Conservatives, by contrast, ought to be deeply puzzled by the results of Tsakiris & Haggard (2005), Kalckert et al. (2019) and Tsakiris et al. (2010). According to Conservatives, nothing can look to be a hand. Here's the rub: then why is it that only visual stimuli exceptionally similar to hands trigger RHI? We ought to be able to deduce from Conservatism some minimally explanatory claim about why RHI failed in these experiments, in the way that we have deduced from Liberalism a very clear answer on this matter. But the Conservative comes up short. Indeed, the Conservative is at risk of taking it to be a brute fact that only an exceptionally narrow range of groupings of low-level properties, 'gestalts', can refer proprioceptive and tactual content in RHI. However, this is far from satisfying, since it is unlikely to be by fluke that these suitable gestalts are identical to the visual gestalts typical of hands. This would amount to simply taking it for granted that visual experience of the rubber hand's size and shape properties does, whereas visual experience of other stimuli's size and shape properties does not, refer the tactual-proprioceptive sensation. Moreover, in pursuing this line of reply, the Conservative would risk ignoring the previously mentioned fact that RHI is partially effected in a top-down manner, since the low-level gestalts that they will appeal to are perceptually represented in a purely bottom-up way. Neither this, nor the explanatory paucity of the Conservative view regarding why RHI fails in the above three experiments, are acceptable. Hence RHI lends support to Liberalism.

Another potential reply would be for the Conservative to concede that the rubber hand has to be represented as a hand at some level for RHI to occur, but insist that this representation is not at the level of conscious perceptual experience. Read one way, e.g., as claiming that the rubber hand is represented as a hand in belief, the reply has no credibility; subjects are not fooled into thinking the stimulus is an actual hand. Read another way, e.g., as claiming that the rubber hand is represented as a hand in *unconscious* perceptual experience, the reply is more credible, but is a significant concession to Liberalism, if not an outright abandonment of Conservatism; Liberals, for consistency's sake, may take their thesis to be about both conscious and unconscious perception and not merely the former. Setting the matter of consistency aside, Conservatives would still owe a compelling story about why the property 'being a hand' can only be represented by perception unconsciously and not also consciously.

A third potential reply would be for Conservatives to claim that the narrowness of the range of gestalts capable of inducing RHI is not a brute fact, but is explained by the further fact that tactual-proprioceptive sensation can only be referred from one's actual hand to an object very much visually like it in terms of low-level properties. But this does nothing to help remove the mystery surrounding the Conservative account of RHI failure in the above three experiments. Without some representation in the subject's psychology of the rubber hand *as a hand* it is still inexplicably brute why such referral should occur simply on the basis of visually experiencing a gestalt that happens to be similar to those of actual hands, but which is not, in addition, represented as being in any way hand-related from the subject's point of view. As discussed in the paragraph above, there seem no plausible options for the Conservative to identify as the bearer of the relevant 'hand' content.

The Liberal's account of RHI gains a measure of further support insofar as it predicts that somatoparaphrenic subjects will be strongly susceptible to the illusion. Such subjects have reduced feelings of ownership for their bodies and typically have weakened proprioception (Vallar & Ronchi 2009). As such, they seem all the more likely to have their tactual-proprioceptive sensations dominated by the conflicting visual information that occurs in RHI. This prediction appears vindicated by at least one study of RHI in a somatoparaphrenic subject (Van Stralen et al. 2013). Strikingly, the subject reported feelings of ownership for the rubber hand simply when looking at it, prior to the onset of tactual-cutaneous stimulation. The Liberal has a partial explanation for this surprising result: the stimulus was being seen to be a hand, i.e.

a body-part, something apt for ownership, hence why a somatoparaphrenic subject reported feelings of ownership towards it, even before tactual stimulation.¹² Unless they saw the rubber appendage as a hand, their having feelings of bodily ownership for it would appear unexplained. This is precisely how matters stand on Conservativism.

V. Feeling What's Not There

Frederique De Vignemont (2014) has recently developed an argument that seems to show that RHI cannot support Liberalism. De Vignemont considers whether the possession of a sortal concept, in perception or cognition, is necessary for RHI. (Sortal concepts are typically thought to be constituents of high-level perceptual content.) They write:

[I]n the Rubber Hand Illusion, it seems unlikely that the participants feel that their hand is F, see that the rubber hand is F, erroneously judge that their rubber hand is their own hand, and then only integrate what they feel with what they see. It may rather be the reverse. Participants do not identify the rubber hand as their hand and then experience the illusion; rather, they experience the illusion and only then do they judge as if the rubber hand were their hand. The identification of the rubber hand as one's hand is not a prerequisite of visuo-somatosensory binding; it is a consequence of it. (2014, p.144)

But it is important not to misconstrue the Liberal conjecture. Liberals are likely to think that subjects in RHI perceptually represent the rubber stimulus as having the multimodal, erroneous content 'my hand', with proprioception and touch supplying the 'mineness' and vision supplying the 'handness.' This is to agree with De Vignemont that the illusion of ownership occurs prior to the identification of the hand as '*my* hand.' But it is to disagree that what follows is that sortal concepts are not involved in such multimodal fusion. What De Vignemont does not consider is whether participants first erroneously identify via perception the stimulus to be '*a* hand' (granted they do not first identify it as '*my* hand', at least not prior to the illusion having occurred and the integration of vision with proprioception and touch).

¹² Again, one shouldn't forget about various bottom-up factors needing to be in place as well. After all, somatoparaphrenic subjects do not have feelings of bodily ownership for each and every hand they observe in everyday life.

Moreover, De Vignemont mentions some RHI experiments that seem to pose a problem for views of RHI along the lines of Liberal conjecture. Notably, Guterstam et al. (2013) elicited the illusion in the absence of any stimuli, referring tactual sensation into empty, peripersonal space. The effect was that participants felt as if they had an invisible hand, yet in the absence of stimuli that could be visually represented as a hand.

What can the Liberal say about this? Worryingly for them, Guterstam et al. (2013) are not alone in inducing RHI via non-bodily stimuli. For instance, Armel & Ramachandran (2003) referred tactual-proprioceptive feedback to the corner of a table, and Ramachandran & Hirstein (1998) referred it to a shoe. However, the Liberal may take heart from the fact that these results are not widely replicated, indeed the former never so (Ma & Hommel 2015, p.76). Indeed, it is striking that Guterstam et al. (2013) failed to induce RHI when using a wooden box as stimulus. This is entirely in line with the results of Tsakiris & Haggard (2005), Kalckert et al. (2019) and Tsakiris et al. (2010) and which were marshalled above as having the potential to vindicate the Liberal conjecture. Insofar as the replicability of Guterstam et al. (2013)'s own results remain to be seen, we don't have here a knock-down objection to the Liberal conjecture. There is also a further question about whether to type this illusion as a variety of RHI or whether to type it as its own *sui generis* 'invisible hand' illusion.

Less defensively, grant that Guterstam et al. (2013) have shown that RHI can be induced in a purely bottom-up manner, via the perception of low-level properties alone. The fact that RHI can be so induced when no physical object is present does not show that it can also be induced in a purely bottom-up, low-level manner when a physical object *is* present, without any need of high-level content. Again, the results of experiments by Tsakiris & Haggard (2005), Kalckert et al. (2019) and Tsakiris et al. (2010) seem to confirm precisely this: low-level, bottom up, stimulus-driven cues are insufficient for RHI *when the visual stimulus is a physical object*. The fact that they may be sufficient when no physical object is present is therefore not to the point.

VI. Conclusion

Reflection on RHI suggests two conclusions. First, that high-level properties figure in perceptual content. Second, that Conservatives who claim that a consequence of the Liberal view is an intolerable broadening of the scope of perceptual illusions, particularly from the perspective of perceptual psychology, should pursue other arguments against the view.

References

- Amoore, J. (1977). "Specific Anosmia and the Concept of Primary Odors." *Chemical Senses* 2 (3): 267-281.
- Audi, R. (2013). *Moral Perception*. Princeton University Press.
- Batty, C. (2011). "Smelling Lessons." *Philosophical Studies* 153: 161-174
- Bayne, T. (2009). "Perception and the Reach of Phenomenal Content." *Philosophical Quarterly* 59 (236): 385-404.
- (2014). "The Multisensory Nature of Perceptual Consciousness." In D. Bennett & C. Hill (eds.) *Sensory Integration and the Unity of Consciousness*. MIT Press.
- (2016). "Gist!" *Proceedings of the Aristotelian Society* 16 (2): 107-126.
- Bengson, J. (2013). "Presentation and Content A Critical Study of Susanna Siegel. The Contents of Visual Experience." *Noûs* 47(4) : 795–807.
- Bedford, F. (2001). "Towards a General Law of Numerical/Object Identity." *Cahiers de Psychologie Cognitive* 20 (3-4): 113-175.
- Bertelson, P. & Radeau, M. (1981). "Cross-modal Bias and Perceptual Fusion with Auditory-Visual Spatial Discordance." *Perception and Psychophysics* 29: 578-584.
- Bertelson, P. & de Gelder, B. (2004). "The Psychology of Multimodal Perception." In C. Spence & J. Driver (eds.) *Crossmodal Space and Crossmodal Attention*. OUP.
- Block, N. (2014). "Seeing-As in the Light of Vision Science." *Philosophy and Phenomenological Research* 89 (1): 560-572.
- Botvinick, M. & Cohen, J. (1998). "Rubber Hands 'Feel' Touch that Eyes See." *Nature* 391: 756.
- Byrne, A. (2009). "Experience and Content." *Philosophical Quarterly* 59 (236): 429-451.
- Byrne, A. & Siegel, S. (2016). "Rich or Thin?" In B. Nanay (ed.), *Current Controversies in Philosophy of Perception*. Routledge.
- Brewer, B. (2011). *Perception and its Objects*. OUP.
- Briscoe, R. (2015). "Cognitive Penetration and the Reach of Phenomenal Content." In A. Raftopoulos & J. Zeimbekis (eds.), *The Cognitive Penetrability of Perception*. OUP.
- (2016). "Multisensory Processing and Perceptual Consciousness: Part I." *Philosophy Compass* 11 (2): 121-133.
- Brogaard, B. (2013a). "Do We Perceive Natural Kind Properties?" *Philosophical Studies* 162 (1): 35-42.

- (2013b) “Phenomenal Seemings and Sensible Dogmatism.” In C. Tucker, (ed.) *Seemings and Justification*. OUP.
- (2018). “In Defense of Hearing Meanings.” *Synthese* 195 (7): 2967-2983
- (2016). “Perceptual Appearances of Personality.” *Philosophical Topics* 44 (2): 83-103.
- Brogaard, B. & Chomanski, B. (2015). “Cognitive Penetrability and High-Level Perception.” *Pacific Philosophical Quarterly* 96 (4): 469-486
- Calvert, G, Spence, C. & Stein, B. (eds.) (2004). *The Handbook of Multisensory Processes*. MIT Press.
- Campbell, J. (2002). *Reference and Consciousness*. OUP.
- Cassam, Q. (2007). *The Possibility of Knowledge*. OUP.
- Cavedon-Taylor. (2015). “Kind Properties and the Metaphysics of Perception: Towards Impure Relationalism.” *Pacific Philosophical Quarterly* 96 (4): 487-509
- Cuneo, T. (2013). “Reidian Moral Perception.” *Canadian Journal of Philosophy* 33 (2): 229-258.
- de Vignemont, F. (2014). “Multimodal Unity and Multimodal Binding.” In. D. Bennett & C. Hill (eds.) *Sensory Integration and the Unity of Consciousness*. MIT Press.
- de Vignemont, F., Tsakiris, M., Haggard, P. (2005). “Body Mereology.” In G. Knoblich, I.M. Thornton, M. Grosjean, M. Shiffrar (eds.) *Human Body Perception from the Inside Out*. OUP.
- Deroy, O. (2015). “Multisensory Perception and Cognitive Penetration.” In J. Zeimbekis & A. Raftopoulos (eds.) *The Cognitive Penetrability of Perception*. OUP.
- Di Bona, E. (2017). “Towards a Rich View of Auditory Experience.” *Philosophical Studies* 174 (11): 2629-2643.
- Driver, J. & Noesselt, T. (2008). “Multisensory Interplay Reveals Crossmodal Influences on ‘Sensory-Specific’ Brain Regions, Neural Responses, and Judgments.” *Neuron* 2008 57 (1): 11-23.
- Farmer, H., Tajadura-Jimenez, A., Tsakiris, M. (2012). “Beyond the Colour of my Skin: How Skin Colour Affects the Sense of Body-Ownership.” *Consciousness and Cognition* 21 (3): 1242–1256.
- Farnè, A. Pavani, A. Meneghello, F. Làdavas, E. (2000). “Left Tactile Extinction Following Visual Stimulation of a Rubber Hand.” *Brain* 123 (11): 2350-2360.
- Fish, W. (2009). *Perception, Hallucination and Illusion*. OUP.
- (2013). “High-Level Properties and Visual Experience.” *Philosophical Studies* 162 (1): 43-55.

- Fulkerson, M. (2016). "Touch." E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition). Available online at:
<https://plato.stanford.edu/archives/spr2016/entries/touch/>.
- Guterstam, A. Gentile, G. & Ehrsson, H. (2013). "The Invisible Hand Illusion: Multisensory Integration Leads to the Embodiment of a Discrete Volume of Empty Space." *Journal of Cognitive Neuroscience* 25 (7): 1078-1099.
- Haans A., Ijsselsteijn, W., de Kort Y. (2008) "The Effect of Similarities in Skin Texture and Hand Shape on Perceived Ownership of a Fake Limb." *Body Image* 5 (4): 389-94.
- Helton, G. (2016). "Recent Issues in High-Level Perception." *Philosophy Compass* 11 (12): 851-862.
- Howard, I. & Templeton, W. (1966). *Human Spatial Orientation*. Wiley.
- Jackson, C. (1953). "Visual Factors in Auditory Localization." *Quarterly Journal of Experimental Psychology* 5: 52-65
- Johnston, M. (2006). "Better Than Mere Knowledge? The Function of Sensory Awareness" In T. Gendler & J. Hawthorne (eds.), *Perceptual Experience*. OUP.
- Kalckert, A., Bico, I., & Xi Fong, J. (2019). "Illusions With Hands, but Not With Balloons – Comparing Ownership and Referral of Touch for a Corporal and Noncorporal Object After Visuotactile Stimulation." *Perception* 48 (5): 447-455.
- Kalderon, M. (2011). "Color Illusion." *Noûs* 45 (4): 751-775.
- Kaeppler, K. & Mueller, F. (2013). "Odor Classification: A Review of Factors Influencing Perception-Based Odor Arrangements." *Chemical Senses* 38 (3): 189-209
- Logue, H. (2013). "Visual Experience of Natural Kind Properties: Is there any Fact of the Matter?" *Philosophical Studies* 162: 1-12.
- Lyons, J. (2005). "Clades, Capgras and Perceptual Kinds." *Philosophical Topics* 33 (1): 185-206.
 ----- (2009). *Perception and Basic Beliefs*. OUP.
- Ma, K. & Hommel, B. (2015). "Body-Ownership for Actively Operated Non-Corporeal Objects." *Consciousness & Cognition* 36 75-86.
- McBrayer, J. (2010). "A Limited Defense of Moral Perception." *Philosophical Studies* 149 (3): 305–320
- McDowell, J. (1982). "Criteria, Defeasibility and Knowledge." *Proceedings of the British Academy* 68: 455-479.
- McGurk H. & MacDonald J. (1976). "Hearing Lips and Seeing Voices." *Nature* 264: 746–748
- McNeill, W. (2012). "On Seeing that Someone is Angry." *European Journal of Philosophy* 20 (4): 575-597

- Macpherson, F. (ed.) (2011). *The Senses: Classic and Contemporary Philosophical Perspectives*. OUP.
- (2012). "Cognitive Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism." *Philosophy and Phenomenological Research* 84 (1): 24–62.
- Martin, M. (1992). "Sight and Touch." In T. Crane (ed.), *The Contents of Experience*. CUP.
- (2002). "The Transparency of Experience." *Mind and Language* 4 (4): 376-425.
- (2010). "What's in a Look?" In B. Nanay (ed.) *Perceiving the World*. OUP.
- Matthen, M. (2005). "Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception." OUP.
- Nanay, B. (2011a) "Do we Sense Modalities with our Sense Modalities?" *Ratio* 24 (3): 299-310.
- (2011b). "Do we See Apples as Edible?" *Pacific Philosophical Quarterly* 92 (3): 305-322.
- (2012). "Perceptual Phenomenology." *Philosophical Perspectives* 26 (1): 235-246.
- (2013). *Between Perception and Action*. OUP.
- Nes, A. (2016). "On What we Experience when we Hear People Speak." *Phenomenology and Mind* 10: 58-85.
- Newen, A. (2017). "Defending the Liberal-Content View of Perceptual Experience: Direct Social Perception of Emotions and Person Impressions." *Synthese* 194 (3): 761-785.
- Nudds, M. (2009). "Sounds and Space." In M. Nudds & C. O'Callaghan (eds.) *Sounds and Perception: New Philosophical Essays*. OUP.
- O'Callaghan, C. (2011a). "Lessons From Beyond Vision (Sounds and Audition)." *Philosophical Studies* 153 (1): 143-160.
- (2011b). "Against Hearing Meanings." *Philosophical Quarterly* 61 (245): 783-807.
- (2015). "The Multisensory Character of Perception." *Journal of Philosophy* 112 (10): 551-569.
- (2016). "Enhancement Through Coordination." In B. Nanay (ed.) *Current Controversies in Philosophy of Perception*. Routledge.
- Price, R. (2009). "Aspect-Switching and Visual Phenomenal Character." *Philosophical Quarterly* 59 (236): 508-518.
- Prinz, J. (2013). "Siegel's Get Rich Quick Scheme." *Philosophical Studies* 163 (3): 827-835.
- Radeau, M. & Bertelson, P. (1976). "The Effect of a Textured Visual Field on Modality Dominance in a Ventriloquism Situation." *Perception and Psychophysics* 20: 227-235.
- (1977). "Adaptation to Auditory-Visual Discordance and Ventriloquism in Semirealistic Situations." *Perception and Psychophysics* (2): 22.

- Raftopoulos, A. (2015). "What Unilateral Visual Neglect Teaches Us About Perceptual Phenomenology." *Erkenntnis* 80 (2): 339-358.
- Reiland, I. (2014). "On Experiencing High-Level Properties." *American Philosophical Quarterly* 51 (3): 177-187.
- (2015). "On Experiencing Meanings." *Southern Journal of Philosophy* 53 (4): 481-492.
- Rosenblum, L. & Saldana, H. (1996). "An Audiovisual Test of Kinematic Primitives for Visual Speech Perception." *Journal of Experimental Psychology: Human Perception and Performance* 22 (2): 318-331.
- Schwenkler, J. (2013). "The Objects of Bodily Awareness." *Philosophical Studies* 162 (2): 465-472.
- Shams, L., Kamitani, Y. & Shimojo, S. (2002). "Visual Illusion Induced by Sound." *Cognitive Brain Research* 14: 147-152.
- Sekuler R., Sekuler A., & Lau, R. (1997). "Sound Alters Visual Motion Perception." *Nature* 385: 308.
- Siegel, S. (2006). "Which Properties are Represented in Perception?" In T. Gendler & J. Hawthorne (eds.), *Perceptual Experience*. OUP.
- (2007). "How Can We Discover the Contents of Experience?" *Southern Journal of Philosophy* 45 (1): 127-42.
- (2009). "The Visual Experience of Causation." *Philosophical Quarterly* 59 (236): 519-540.
- (2011). *The Contents of Visual Experience*. OUP.
- Silins, N. (2013). "The Significance of High-Level Content." *Philosophical Studies* 162 (1): 13-33.
- Spence, C. (2011). "Crossmodal Correspondences: A Tutorial Review." *Attention, Perception, & Psychophysics* 73 (4): 971-995.
- Stokes, D. Matthen, M. & Biggs, S. (eds.) (2014). *Perception and Its Modalities*. OUP.
- Smith, B. (2015). "The Chemical Senses." In M. Matthen (ed.) *The Oxford Handbook of Philosophy of Perception*. OUP.
- Toribio, J. (2018). "Visual Experience: Rich but Impenetrable." *Synthese* 195 (8): 3389-3406.
- Travis, C. (2004). "The Silence of the Senses." *Mind* 113 (449): 57-94.
- Tsakiris, M. (2010). "My Body in the Brain: A Neurocognitive Model of Body-Ownership." *Neuropsychologia* 48 (3): 703-12.

- Tsakiris, M., Carpenter, L., James, D. & Fotopoulou, A. (2010). "Hands Only Illusion: Multisensory Integration Elicits Sense of Ownership for Body Parts but not for Non-Corporeal Objects." *Experimental Brain Research* 204 (3): 343-52.
- Tsakiris, M., Costantini, M., & Haggard., P. (2008). "The Role of the Right Temporo-Parietal Junction in Maintaining a Coherent Sense of one's Body." *Neuropsychologia* 46: 3014–3018.
- Tsakiris, M. & Haggard, P. (2005). "The Rubber Hand Illusion Revisited: Visuotactile Integration and Self-Attribution." *Journal of Experimental Psychology: Human Perception and Performance* 31 (1): 80-91.
- Tye, M. (1995). *Ten Problems of Consciousness*. MIT Press.
- Vallar, G & Ronchi, R. (2009). "Somatoparaphrenia: A Body Delusion. A Review of the Neuropsychological Literature." *Experimental Brain Research* 192 (3): 533-51.
- van Stralen, H., van Zandvoort, M., Kappelle, L., Dijkerman, H. (2013). "The Rubber Hand Illusion in a Patient with Hand Disownership." *Perception* 42 (9): 991-993.
- Vatakis, A. & Spence, C. (2007). "Crossmodal Binding: Evaluating the 'Unity Assumption' Using Audiovisual Speech Stimuli." *Perception & Psychophysics* 69: 744-756.
- Welch, R. & Warren, D. (1980). "Immediate Perceptual Response to Intersensory Discrepancy." *Psychological Bulletin* 88 (3): 638-667.
- Werner, P. (2016). "Moral Perception and the Contents of Experience." *Journal of Moral Philosophy* 13 (3): 294-317