



Molyneux's Question and multisensory integration

Berit Brogaard^a  (brogaard@gmail.com)

Dimitria Gatzia^b  (degatzia@uakron.edu)

Abstract

Molyneux's question is, very roughly, that of whether a blind person's prior acquaintance with shape properties by touch alone would suffice for visually identifying those properties, if her sight were restored. The question is at least in part an empirical one, and various scientific attempts have been made to answer it not only for sight and haptic touch but also for various other sensory modalities. This paper is not aimed at answering Molyneux's original or related questions. Rather, our aim here is to bring his question to bear on the nature of multisensory integration. In particular, we argue that variants on Molyneux's question spell trouble for a dominant view of multisensory integration and develop an alternative account that is immune to these problems. Finally, we show that the answer to Molyneux's original question has interesting consequences for whether our alternative account is true of all cases of multisensory experience.

Keywords

Amodal integration · Auditory experience · Common sensibles · Crossmodal integration · Modal integration · Molyneux's question · Multisensory integration · Multisensory perception · Perceptual demonstratives · Sensory modalities · Tactual experience

This article is part of a special issue on "Molyneux's question today", edited by Gabriele Ferretti and Brian Glenney.

1 Introduction

Suppose that a newly sighted person was asked to distinguish two objects she was previously familiar with through touch. Would she be able to do so? This is the very question that the Irish scientist and politician, William Molyneux, posed in the now famous letter to John Locke on July 7, 1688 (Molyneux, 1978; cf. Locke, 1978):

^a University of Miami.

^b University of Akron.

Dublin July. 7. 88

A Problem Proposed to the Author of the Essai Philosophique concernant L'Entendement

A Man, being born blind, and having a Globe and a Cube, nigh of the same bignes, Committed into his Hands, and being taught or Told, which is Called the Globe, and which the Cube, so as easily to distinguish them by his Touch or Feeling; Then both being taken from Him, and Laid on a Table, Let us Suppose his Sight Restored to Him; Whether he Could, by his Sight, and before he touch them, know which is the Globe and which the Cube? Or Whether he Could know by his Sight, before he stretch'd out his Hand, whether he Could not Reach them, tho they were Removed 20 or 1000 feet from Him?

If the Learned and Ingenious Author of the Forementioned Treatise think this Problem Worth his Consideration and Answer, He may at any time Direct it to One that Much Esteems him, and is,

His Humble Servant William Molyneux High Ormonds Gate in Dublin.
Ireland

Impressed by its ingenuity, Locke decided to include Molyneux's question in the second edition of his *An Essay Concerning Humane Understanding* (1694), which has led to an ever growing interdisciplinary interest in answering this or related questions (e.g., Ferretti & Glenney, 2020; Glenney, 2013; Gregory & Wallace, 1963; Held et al., 2011; Matthen & Cohen, 2020; Schwenkler, 2019). While previous theoretical and experimental inquiries into Molyneux's question have mainly focused on the original or related questions, this is not the aim of this paper. Rather, our primary aim here is to use Molyneux's question as a springboard for highlighting some flaws in the received philosophical view of multisensory (or crossmodal) integration and advancing an alternative.

One of the reasons Molyneux was curious about whether newly sighted subjects could visually recognize objects with which they had prior acquaintance through touch can, it seems, be traced back to Aristotle's view of common sensibles. Aristotle argued that at least some properties, which he referred to as "common sensibles," can be perceived by more than one sensory modality (Aristotle, *Sense and Sensibilia*),¹ for instance, shape properties (e.g., *being cone-shaped*), size properties (e.g., *being short*), texture properties (e.g., *being squishy*), numerosity (e.g., *being distinct*), and movements (e.g., *swinging*).

From this perspective, Molyneux's original question can be seen as providing a way to ascertain whether there are common sensibles. Indeed, the properties integral to his question, viz., *being cubical* and *being spherical*, just are common sensibles. Accordingly, a negative answer to Molyneux's original question would imply that those apparent common sensibles are not common sensible after all.

¹ In (Barnes, 1984).

But, as we will see, these particular shape properties are not central to the idea underlying Molyneux's thought experiment. Related questions could be asked with other common sensibles and either the same or different sensory modalities. If related questions asked about the various other apparent common sensibles would tend to yield a "no" answer, that would suggest that there are no common sensibles.

Taking that as our springboard, we argue that MQ-like questions can be used to challenge the received view of multisensory experience. On the received view, properties are first attributed to objects or events in the individual sensory modalities, and those attributions are then integrated amodally. This view encounters difficulties, we argue, in MQ-like cases in which the partial restoration of a person's sensory abilities in one sensory modality makes her attribution of qualities she can identify in that modality dependent on her identification of an object or event in a second modality. To account for such cases, we argue, an alternative view of multisensory integration is needed.

We proceed by developing an alternative modal view of multisensory integration, according to which multisensory integration at least sometimes occurs modally by identifying properties in one sensory modality, which are then attributed to an object or event identified in another. This account, we argue, doesn't encounter the difficulties facing the received view.

Finally, we argue that an amodal integration model is needed only in cases of multisensory integration involving attributions of apparent common sensibles. But, as we will see, this has interesting consequences. If the answer to Molyneux's question and related questions involving other apparent common sensibles is "no," the apparent common sensibles are not common sensibles after all, in which case there is no need for an amodal integration model. As a result, the received view is expendable, and all cases of multisensory integration take place modally.

The plan for the paper is as follows. In section 2, we detail how Molyneux's original question bears on common sensibles, in Aristotle's sense. In section 3, we show that the received view of multisensory integration is jeopardized by MQ-like scenarios involving multisensory integration where attributions of qualities in the individual sensory modalities are interdependent. In section 4, we develop an alternative modal view of multisensory integration and show that it avoids the pitfalls of the received view. Finally, we tie our argument for the modal integration view to Molyneux's original question, arguing that how we answer it is of utmost importance to the nature of multisensory integration in general.

2 Molyneux's question and common sensibles

Using "touch" as shorthand for "haptic touch," Molyneux's original question can be glossed as follows:

MQ: Would a person who was born blind and who has learned to distinguish and name a globe and a cube by touch be able to distinguish and name these objects by sight alone, if his sight were restored?

The majority of the research on multisensory integration has focused on answering MQ. Most scholars have responded in the negative but for a host of different reasons. By far the most common reason for a negative answer to MQ is that treatments for congenital blindness normally do not *immediately* enable patients to identify the shapes of objects by sight alone. In congenital blindness, total restoration of sight is often not possible, and partial restoration requires extensive post-operative training to be successful (Gregory & Wallace, 1963; Held et al., 2011; although see Fine et al., 2003). It is therefore not surprising that most of these approaches have yielded a negative answer.

While the prospects of *immediately* restoring vision are dim, this is not a good reason for thinking that the answer to MQ is negative. For, Molyneux did not ask whether the newly sighted person would be able to visually identify globes and cubes *immediately* following treatment. So even if restoring sight *requires* training, it is entirely coherent and intuitively plausible that one can restore sight to a blind person without training them on visual stimuli that include globes and cubes (see e.g., Cheng, 2015; Held et al., 2011; Ostrovsky et al., 2009; Sikl et al., 2013). Indeed, for most of the known cases where sight has been, at least partially, restored, there has been no training apart from the subjects themselves trying to improve their visual abilities (see e.g., Gregory & Wallace, 1963; Kurson, 2008).

The point that restoring sight may require training has not gone unnoticed by researchers (e.g., Connolly, 2013; Fine et al., 2003; Ostrovsky et al., 2009; Schwenkler, 2012, 2013, 2019; Sikl et al., 2013; although see Cheng, 2015). Indeed, recent explorations have seen MQ as a gateway to novel treatments of blindness (Humayun et al., 2024; Zrenner, 2013). Blindness is a highly complex, individualistic, and multifactorial condition that is often partial rather than complete and often acquired rather than congenital. Accordingly, recent studies of MQ have focused on exploring MQ-like questions with greater relevance to the prospects of restoring sight in people with varying degrees and kinds of vision loss (e.g., López-Bendito et al., 2022; Matthen & Cohen, 2020) as well as determining the cut-off age at which sensory modalities retain their plasticity (Ostrovsky et al., 2009).

Our aim here is not to answer either MQ or MQ-like questions. Rather, our primary focus is rooted in the issue that seems to have piqued Molyneux's curiosity.² His question seems to have been sparked by a concern about whether any perceptible properties are really discernible by more than one of our senses. Perceptible properties of this kind are also known as "common sensibles," a term originally coined by Aristotle (Aristotle, *Sense and Sensibilia*).³ Certain shape properties (e.g.,

² Elsewhere we argue that the empirical evidence aimed at providing an answer to Molyneux's question requires distinguishing between visuo-sensory seeing and visuo-epistemic seeing (see Brogaard et al., 2020).

³ In (Barnes, 1984).

being rectangular), size properties (e.g., *being large*), texture properties (e.g., *being bumpy*), numerosity (e.g., *being triplets*), and movements (e.g., *vibrating*) are prime examples of perceptible properties that philosophers unhesitatingly would classify as common sensibles (see e.g., Tye, 2007). Aristotle lists all the properties under these labels as common sensibles (Aristotle, *Sense and Sensibilia*), but surely that is overkill. For example, if size properties are common sensibles, they must be discernible by sight and touch. Yet *having a 3,958.8 mi radius* or *being gigantic for a black hole* are at best discernible by sight (with technology) but not plausibly by touch.

Molyneux's thought experiment can be seen as outlining a way to test – in theory at least – whether shape properties are common sensibles, as suggested by Aristotle. A positive answer to MQ would suggest that at least some properties, such as *being cubical* and *being spherical*, can be discerned by sight and touch and are therefore common sensibles. It follows that restoring sight to a blind person would allow them to visually perceive cubes as cubical and globes as spherical – at first sight, so to speak. (Recall that even if restoring sight to a blind person *requires* training, it is entirely coherent and intuitively plausible that one can restore it without training them on visual stimuli that include cubes and globes).

A negative answer to MQ, by contrast, would suggest that at least some properties – in this case the shape properties *being cubical* and *being spherical* – are discernible by touch but not by sight and thus are not common sensibles after all. It follows that restoring sight to a blind person would still only give them the ability to discern *being cubical* and *being spherical* by touch but not by sight. As it is intuitively plausible that we can restore sight to a blind person without training them on visual stimuli that include cubes and globes, a negative answer to MQ would rule out that shape properties generally (if ever) are common sensibles.

The scenario Molyneux presents in his letter to Locke focuses on low-level, simple common sensibles for distinguishing and naming objects (e.g., *being cubical* and *being spherical*) but this is by no means essential to the puzzle. He could have presented it in terms of the high-level common sensibles like *pot-* and *pan-shaped* or *dog-* and *cat-shaped*. Or he could have framed it in terms of a different pair of sensory modalities and relevant common sensibles. For example, neurotypical people primarily recognize and experience musical sounds through hearing. Sounds cause the fine hair cells that line the inner ear to vibrate, and the vibrations are projected to the auditory cortex, which interprets the incoming auditory signals as musical sounds. Deaf people, by contrast, predominantly recognize and experience musical sounds through haptic touch (see e.g., Auer et al., 2007; Levänen et al., 1998). By touching the medium through which sound waves travel with, say, their feet (e.g., in a concert hall) or their hands (e.g., by placing them on a loudspeaker), the vibrations of music – particularly the bass notes – are transmitted through their body to their brain's auditory cortex, which then interprets the incoming sensory signals as musical sounds.

Molyneux could thus have framed his question in terms of the sensory modalities, hearing and haptic touch, and the common sensible *being musical sounds* as follows:

MQ–Music: Would a person who was born deaf and who has learned to recognize musical sounds by touch be able to recognize musical sounds by hearing alone, if her hearing were restored?

As in the case of MQ, a positive answer to MQ–Music would suggest that *being musical sounds* is a common sensible, whereas a negative answer would suggest that it is not.

3 MQ and amodal integration: A critique of the received view

Notice that if the answer to the original version of MQ is “yes,” then a newly sighted person who were to simultaneously touch and see an object would have a multisensory experience. For example, a shape property, e.g., the common sensible *being cubical*, would be both seen and felt by a newly sighted person. In this case, the attribution of *being cubical* to the cube would be perceptually attributed to the cube by both the visual and the tactual modalities. So, if a newly sighted person were to both hold and see the cube, the attributions of *being cubical* to the cube in the two sensory modalities would be integrated into a complete multisensory experience. By way of contrast, if a newly sighted person were to perceptually attribute *being cubical* to the cube by sight alone, the attribution of that feature to the cube would be limited to a single sensory modality, i.e., vision. As this experience is the result of attributing *being cubical* to the cube in a single sensory modality, we can think of it as unimodal.

In the multisensory case, the integration of the attributions in the visual and tactual modalities is neither entirely visual nor entirely tactual. It cannot, therefore, be accounted for by appealing to just one of the two sensory modalities. It is for this reason that the integration or binding of attributions that occurs in multisensory experience has been thought of as amodal. In fact, this idea lies at the core of the dominant philosophical view of multisensory (or crossmodal) integration (see e.g., Altieri, 2015; Bayne, 2014; Bourget, 2017; Briscoe, 2016; Dainton, 2002; De Vignemont, 2014; Deroy, 2014; Nudds, 2001; O’Callaghan, 2012, 2014, 2015; Rescorla, 2020; Schwenkler, 2015, 2019). Call the view that amodal integration is a distinctive mark of multisensory experience the “received view.”

Casey O’Callaghan (2012, 2014, 2015), a prominent defender of the received view, argues that amodal integration is not only what distinguishes multisensory experience from ordinary unimodal experience but also what distinguishes it from the mere co-presence of experiences. Say you are watching some people dancing across the street while petting your dog. In the envisaged scenario, you have a

visual experience of people dancing across the street and a tactual experience of your dog's fur. These two experiences are constituent parts of the total sensory experience you currently have. But they are not integrated in any substantial sense. They merely co-exist as parts of your total experience.

According to O'Callaghan, the difference between ordinary unimodal experiences and merely co-present experiences, on the one hand, and multisensory experiences, on the other, is that in the former case, the sensory character of ordinary unimodal experiences and merely co-present experiences can be fully accounted for in terms of attributions of properties to an object in the individual sensory modality or modalities in question. For example, the sensory character of the experience you undergo when looking at the people dancing while petting your dog derives from your visual experience of the people dancing and your co-present tactual experience of your dog's fur.

But, in O'Callaghan's view, this is not so in the case of multisensory experience. Say you see and hold a tennis ball and undergo a visuo-tactile experience that represents the ball as spherical. In this case, the sensory character of your experience reflects that *being spherical* is attributed to the ball in the visual and tactual modalities, and those attributions are then bound together amodally in perceptual faculties that are neither visual nor tactual in nature, for instance, in higher non-sensory areas of the brain.

More abstractly, on the received view, the integration of sensory information about a property *F* received by two sensory modalities proceed as follows: Modality 1 attributes *F* to an object on the basis of the sensory signals received by modality 1, and modality 2 attributes *F* to an object on the basis of the sensory signals received by modality 2. The attributions of *F* to objects in the two sensory modalities are then integrated amodally, which involves identifying the object instantiating *F* and presented in modality 1 with the object instantiating *F* and presented in modality 2.

On the received view, the same point, of course, carries over to multisensory experiences that attribute more than one property to an object. Suppose you are seeing and holding a beige, wooden ball and undergo a visuo-tactual experience of the ball being beige, wooden, and spherical. Here, *being beige*, *being wooden*, and *being spherical* are attributed to the ball in the visual modality, whereas *being wooden* and *being spherical* are attributed to the ball in the tactile modality. Accordingly, the sensory character of your visuo-tactual experience reflects the amodal integration of those attributions in the two separate sensory modalities.

While the above discussion primarily has focused on how attributions of properties in the visual and tactual modalities are integrated, the received view is supposed to be applicable to other sensory modalities as well. Consider a person, Lena, who is born deaf. Lena has acquired the ability to tell when people are speaking by visually identifying lip movements as speaking events. One day Lena undergoes a procedure that partially restores her hearing, enabling her to hear sounds, including speech sounds, although she cannot hear the direction from which these

sounds are transmitted. Even so, when she is in a location in which only a single person is talking, she can infer that the speech sounds must be coming from that person. If, however, she is in a location with multiple people, her inability to hear the direction of sounds makes it impossible for her to infer on the basis of hearing alone who is making the speech sounds she can hear. The question here is whether Lena would be able to recruit her prior visual abilities to attribute the speech sounds she can hear to a particular person, despite her inability to hear the direction from which the sounds are transmitted.⁴ We can formulate this question as an MQ-like question as follows:

MQ–Speaking: Would a person who was born deaf and who has learned to attribute speaking events to particular people by sight in non-crowded conditions but not in crowded ones be able to attribute speaking events to particular people by sight and hearing in crowded conditions, if her hearing were partly restored?

Suppose the answer is “yes.” Now let’s consider our earlier scenario. Recall that even though Lena’s hearing has been partly restored, her residual hearing impairment prevents her from hearing sounds coming from a specific direction. So, in conditions where multiple objects could potentially be the source of the sounds she can hear, she cannot attribute those sounds she can hear to a particular source on the basis of hearing alone. Even under those conditions, however, she can sometimes exploit her prior ability to visually identify events as sound-producing events, which then enables her to attribute sounds she can hear to a particular sound source she can see. For example, if she is at a gathering where she is facing a person located in the right side of her visual field and a person located in the left side of her visual field, and she hears speech sounds, she cannot tell on the basis of hearing alone whether it’s the person to the right or left of her who is speaking. However, she can take advantage of her prior ability to visually identify lip movements as sources of speech to attribute the speech sounds she hears to either the person on the right or the left of her.

The received view yields the wrong prediction in this case. On the received view, a visuo-auditory experience of someone speaking is the result of visually referring to someone, aurally referring to someone, and amodally identifying the perceptual referents. So, the received view predicts that the newly hearing person (with residual hearing impairments) should not be able to take advantage of her prior ability to visually identify lip movements as sources of speech to attribute certain speech sounds she hears to a particular person. After all, her very problem to begin with is that she cannot attribute speech sounds to a particular person

⁴ Studies indicate that cochlear implants can lead to the development of multimodal (e.g., visual and auditory) brain plasticity, which enhances the ability to attribute speech to particular speakers in crowded environments. These cases can be seen as analogous to cases in which subjects regain their hearing. See, for example, Rouger et al. (2007).

on the basis of hearing alone, which would be required if the received view were correct.

The envisaged case only presents a problem for the received view if the answer to MQ-Speaking is “yes.” So, the natural response for advocates of the received view here is that since the answer to MQ-Speaking might be “no,” our counterexample doesn't present a definitive case against their view. However, while it is far from clear what the answer is to the original version of MQ, because the question asks about a person's abilities to utilize a *fully* restored sense, the answer to MQ-Speaking is much more likely to be affirmative. Indeed, partial restoration of hearing in deaf people has been demonstrated on multiple occasions. In many of those cases, the partly hearing person continues to rely on sight in order to map the sounds they can hear to sound sources (see, e.g., [Kral & Sharma, 2023](#) for a review). So, our envisaged case in fact spells doom for the received view.

4 MQ and modal integration: A solution

If our argument against the received view is cogent, then it behooves us to find an alternative view of multisensory integration. In our view, the most obvious alternative is a modal integration view ([Brogaard & Chudnoff, 2018](#)). On the latter approach, multisensory integration does not require that properties are perceptually attributed to objects or events in each of the relevant sensory modalities before integration can occur (amodally). Rather, on the modal view, integration at least sometimes occurs when properties perceived in one sensory modality are attributed to an object or event perceived in a second sensory modality. Because the integration takes place in the second sensory modality, it is modal rather than amodal. For example, in a visuo-auditory experience that represents sounds as produced by a visible source, audible qualities perceived in the auditory modality are attributed to a visible source perceived in the visual modality. Though the integration occurs within the visual modality, the resulting visuo-auditory experience represents the audible qualities as produced by the visible source. For example, if you are seeing and hearing Danny Keye sing *Civilization*, the resulting visuo-auditory experience represents the audible qualities as produced by Danny Keye moving his lips.

On the modal account, perceptual integration takes place on the basis of perceptual demonstrative reference being made by one sensory modality and anchored by another ([Brogaard & Chudnoff, 2018](#)). In the Danny Keye case, the visuo-auditory experience attributes sounding like such and such to a lip-moving event picked out by a visual demonstrative whose reference is anchored to that event by virtue of its being visible. So, a visual demonstrative reference is being made by the visual modality and is anchored by the auditory modality.

A visual demonstrative is the perceptual equivalent of demonstratives that occur in ordinary language, such as “this” and “that.” Like other referential expressions, demonstratives contribute a discourse referent and a condition to the dis-

course representation (i.e., the representation generated by the speech/discourse (Heim, 1982; e.g., Kamp, 1981). Discourse referents are a special kind of variable that are bound by discourse-wide existential quantifiers. So, demonstratives only refer to an external mind-independent individual when there is such an individual that can serve as a value of the variable introduced by the expression. Reference requires satisfying not only the conditions introduced by the referring expression and any potential demonstration (e.g., a pointing) but also those introduced by co-referring expressions, such as in cases of anaphora where a demonstrative refers back to a previously mentioned referent, as in “Jim left his dirty laundry on the floor. This annoyed Erin” or where a demonstrative refers to an about-to-be-mentioned referent, as in “It was his own fault that Curtis didn’t get to go to camp.” “This” as it occurs in “Jim left his dirty laundry on the floor. This annoyed Erin” contributes a discourse referent x and the condition that the referent must be an event or fact to the discourse representation. Here, “Jim left his dirty laundry on the floor,” and “this” refer to a particular event E only if there is a mapping from “Jim’s-leaving-dirty-laundry-on-the-floor(x)” and “this(x)” to E .

Perceptual demonstratives function in a way analogous to anaphoric pronouns. They introduce a perceptual referent x – i.e., the perceptual equivalent of a discourse referent – and a condition, which jointly may or may not refer to an actual mind-independent object or event. Perceptual references to a perceptual referent in different sensory modalities can thus be interdependent in the way that certain co-referring expressions in different parts of speech can be interdependent. In the case of seeing and hearing someone sing, the visual aspect of the visuo-auditory experience provides a visual demonstrative that contributes a singing or lip-moving event x to the experiential content, and the auditory aspect of the experience attributes audible singing qualities to x by using the visual demonstrative. By making use of a visual demonstrative, the auditory aspect of the visuo-auditory experience thus becomes dependent on and not merely co-present with the visual aspect of the visuo-auditory experience.

The modal integration is, of course, constrained by temporal and spatial congruency. For example, if your television and your speakers were out of sync, then the movement of Danny Keye’s lips and the music would be mismatched. In this case, the phenomenology of your experience would reflect this sort of dissociation between the visual and auditory inputs.⁵

⁵ The reader might be concerned that the modal account of multisensory integration cannot account for the different manners of representation of the distinct sensory modalities (Chalmers, 2004). For example, while vision represents visually, audition represents auditorily. If that is the case, then multisensory experience fails to represent modally. It follows that the phenomenology of multisensory experience (e.g., seeing Keye sing) cannot be derived from the phenomenology of each of the relevant modalities. However, this objection assumes that manners of representation cannot be additive but must undergo change from being unisensory to becoming amodal. It is plausible, however, that when you see Keye sing, your experience represents him in a visuo-auditory manner.

Spatial and temporal congruency seems to contribute significantly to the phenomenon of ventriloquism. In ventriloquism, although we know that the ventriloquist produces the voice of the puppet in his hand, the voice visually and auditorily appears to come from the puppet's mouth. Here, we mistakenly attribute the speech sounds identified in the auditory modality to the puppet's speaking motions identified in the visual modality. We are taking advantage of this mistaken attribution in ventriloquism in many ordinary contexts, such as when we watch television.

It merits emphasis that the dependence relation in modal multisensory experience can go in both directions. This can be seen from two cases of multimodal interaction in which what we see alters what we hear or vice versa. One case is the McGurk effect in which seeing lip movements influences and alters the speech sounds we hear (McGurk & Macdonald, 1976). For example, when the auditory syllable "ba" is presented in synchrony with a speaker mouthing "ga," subjects typically report hearing "da." The other case is the double-flash illusion, in which the presentation of two brief auditory beeps makes a single flash look like two flashes (Shams et al., 2000). In these two cases of visuo-auditory experiences, the dependence relation seems to go in opposite directions, depending on what is taken to produce what. In the McGurk effect, seen lip movements are taken to produce the speech sound, which suggests that the audible qualities are attributed to the referent of a perceptual demonstrative that is anchored to lip movements by virtue of those lip movements being visible. In the double-flash illusion, the beeps are taken to produce the flashes, which suggests that the flashes are attributed to the referent of a perceptual demonstrative that is anchored to the beeps by virtue of those beeps being audible.

There are many other cases of multisensory experience that seem to involve modal rather than amodal integration.⁶ Here is a case involving vision and proprioception. You see someone hit you and feel the pain radiating from the site of impact. Here, the proprioceptively felt properties of pain are attributed to the referent of a visual demonstrative that is anchored to the seen hitting event by virtue of that event being visible. Or consider this case involving audible and tactual properties being attributed to a perceptual referent anchored to a visually pictured event. You are standing with your back to a pool and hear a splash that sounds like a person jumping in the water and feel splashy water hit your body. Here, the splash sounds are attributed to the referent of an auditory demonstrative, and the tactual

⁶ There is empirical support for this claim. For example, Shams (2000) found that auditory signals directly influence visual perception. Similarly, van Wassenhove (2005) found that visual and auditory modalities in speech perception are not separately processed and then integrated amodally but rather interact very early in the perceptual process. These findings suggest that visual cues, such as lip movements, directly modulate the neural processing of auditory speech at an early stage, indicating an intertwined visuo-auditory experience that challenges the notion of separate modality-specific referents being amodally integrated. Consistent with these results, Nath (2012) found that visual and auditory information is integrated in the superior temporal sulcus at an early stage of processing.

qualities of splashy water hitting your body are attributed to a tactual demonstrative, and the two referents of the two perceptual demonstratives are anchored to an unseen and visually pictured event, viz., that of a person's body making impact with the water.

The advantage of the modal integration view compared to the received view is that the former doesn't run into the same sort of difficulties as the latter with respect to MQ-Speaking. In MQ-Speaking, hearing is partly restored to a deaf person, Lena, which enables her to hear sounds but leaves her with a residual hearing impairment that makes her unable to determine the direction of sounds by hearing alone. So, in a setting in which sounds she can hear are coming from a particular direction but in which the sounds themselves do not give away their source because they are sufficiently similar, she cannot attribute those sounds to a particular source. So, if Lena is facing a person located in the right side of her visual field and a person located in the left side of her visual field, and she hears speech sounds, she cannot tell on the basis of hearing alone whether it's the person to the right or left who is speaking. However, she can exploit her prior abilities to visually identify lip movements as producing speech sounds in order to attribute the speech sounds she hears to either the person on the right or the left.

This case challenges the received view, as we have seen, because, according to it, a visuo-auditory experience of someone speaking is the result of visually referring to someone, aurally referring to someone, and amodally identifying the perceptual referents. It follows that the received view incorrectly predicts that Lena, the newly hearing person with residual hearing impairments, should not be able to take advantage of her prior ability to visually identify lip movements as sources of speech to attribute the speech sounds she hears to a particular person. This sort of case does not present a difficulty for the modal integration view, as the latter view correctly predicts that Lena can take advantage of her prior abilities to visually identify lip movements as producing speech and then exploit that visual referent in order to attribute the speech sounds she can now hear to either the person on the right or the left.

Note that the received view does not face an analogous counterexample in cases of multisensory experiences in which one can make perceptual reference to an object in both sensory modalities on the basis of identifying the same distinctive characteristic of the object in the two modalities. Assuming an affirmative answer to the original version of MQ, the newly sighted person's visuo-tactual experience of a cube is exactly a case of this kind. The newly sighted person can identify the cube by sight alone only by visually identifying the quality *being cubical*. But it is by perceiving that very quality in the tactual modality that she is able to identify the cube in that modality. So, in such cases, one cannot conjure up a counterexample analogous to MQ-Speaking because if one posited that a newly sighted subject with residual vision impairments had trouble determining the identity of an object she was seeing and holding on the basis of visually identifying its shape, she could simply exploit her prior ability to tactually identify its shape to determine the iden-

tity of the object. There would be no need for the newly sighted subject to rely on both vision and touch in order to determine its identity.

Things are quite otherwise when it comes to MQ–Speaking. Here, the overall multisensory experience had by Lena, the person who had her hearing partially restored, involves attributing audible qualities to one of the two people present. But since she cannot attribute sounds to a particular source when the sounds could come from more than one potential source, audition alone is not sufficient for her to determine who the speaker is. To attribute audible qualities to one person rather than the other, Lena thus needs to rely on her prior visual abilities to pick out one of the two people on the basis of their speaking motions. In MQ–Speaking, audition is thus dependent on vision in that it can attribute the heard speech sounds to a particular person only by making use of vision to make reference to that person's speaking motions.

By contrast, assuming an affirmative answer to the original version of MQ, the newly sighted person can make perceptual reference to an object in both sensory modalities on the basis of identifying one and the same distinctive characteristic of the object in the different modalities. Accordingly, there is no need for them to rely on both sensory modalities to make up for any residual deficits in one modality. In such cases, the modal integration view thus agrees with the received view that multisensory experience arises as a result of attributing a property to a perceptual referent in each individual sensory modality and then integrating them amodally.

However, it seems that we only need to posit *amodal* integration in cases of multisensory experience in which the very same quality (or set of qualities) is perceptible in more than one sensory modality, that is, cases of multisensory experience seeming to involve common sensibles. Take, for example, a case of a visuo-tactual experience of a beige cube. Since beige can only be visually perceived, the modal integration view can treat the integration as modal rather than amodal, that is, the case can be accounted for as follows: the cube is identified by its cubicality in the tactual modality, and the property *being beige* is identified in the visual modality and is then attributed to the cube identified in the tactual modality. In this case, the integration takes place in the tactual modality by attributing the color property identified in the visual modality to the object identified in the tactual modality.

If we are right that we only need to posit *amodal* integration in cases of multisensory experience seeming to involve common sensibles, then the question of whether all multimodal experience involves modal integration turns out to depend on the answer to the question of whether there are common sensibles and hence on what the answer to the original version of MQ is. If there are no common sensibles (i.e., if the answer to MQ is “no”), then all multisensory integration is modal, and the received view is dispensable.⁷

⁷ Schwenkler (2012, 2013, 2019) argues that studies on the performance of newly sighted subjects fail to provide a positive answer to MQ because they may lack the ability to form three-dimensional representations of viewed objects, despite having the ability to form three-dimensional tactile representations of these objects (see also Connolly, 2013; for criticism see

5 Conclusion

In this paper, we began by suggesting that Molyneux's original question (MQ) is best understood as a way to inquire about whether certain properties are common sensibles. Specifically, if the answer to MQ is "yes," then a newly sighted person would be able to *visually* identify apparent common sensibles, i.e., *being cubical* and *being spherical*, with which they were previously acquainted by touch. But if the answer to MQ is "no," then a newly sighted person would not be able to *visually* identify these apparent common sensibles. So, if the answer to MQ is "no," then these apparent common sensibles are not common sensibles after all. As Molyneux's original question can be asked about other common sensibles involving the same or different sensory modalities, a consistent "no" answer to MQ questions involving apparent common sensibles suggests that there are no common sensibles after all.

Given this starting point, we went on to argue that MQ-like questions spell trouble for the received view of multisensory integration. On the received view, multisensory integration occurs by attributing properties to objects or events in each individual sensory modality, and these attributions are then integrated amodally. But, as we have shown, that view is unable to deal adequately with certain MQ-like cases in which a subject with a partly restored sensory ability needs to make use of both sensory modalities in order to attribute a property identified in one modality to a particular object or event identified in the other. We then proposed an alternative view of multisensory integration, according to which multisensory integration at least sometimes occurs by identifying qualities in one sensory modality that are then attributed to the referent of a perceptual demonstrative that is anchored to an event or object identified in the another. This view, we argued, avoids the problems plaguing the received view.

Finally, we argued that amodal integration is needed only for cases of multisensory integration involving what seems to be common sensibles. But this has the interesting outcome that if the answer to the original version of MQ is "no," and there are no common sensibles, then the received amodal integration view is unmotivated, as all cases of multisensory experience inevitably will involve modal rather than amodal integration.

Acknowledgments

We are grateful to the guest editors of this special issue of PhiMiSci and to audiences at CUNY Graduate Center, NYU, Stanford University, and the University of Miami for the opportunity to present some of our ideas on perception and multisensory integration and for excellent discussion and feedback. We are also grateful to two anonymous reviewers for this journal for helpful comments on a previous version of this paper and to Ned Block, Alex Byrne, Elijah Chudnoff, Melvyn A. Goodale, Robert Kentridge, Azenet Lopez, Fiona

Cheng, 2015). Schwenkler (2019) further argues that MQ is likely to be negative. For, if shape properties are represented differently in vision and in touch, newly sighted subjects would be unable to compare them cross-modally.

Macpherson, Mike Martin, Casey O'Callaghan, Thomas Alrik Sørensen, and Aurore Zelazny for past insightful discussions of these and related issues.

References

- Altieri, N. (2015). Multimodal theories of recognition and their relation to Molyneux's question. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01547>
- Auer, E. T., Bernstein, L. E., Sungkarat, W., & Singh, M. (2007). Vibrotactile activation of the auditory cortices in deaf versus hearing adults. *Neuroreport*, 18(7), 645–648. <https://doi.org/10.1097/WNR.0b013e3280d943b9>
- Barnes, J. (1984). *The complete works of Aristotle, volume one*. Princeton University Press. <https://press.princeton.edu/books/hardcover/9780691016504/the-complete-works-of-aristotle-volume-one>
- Bayne, T. (2014). The multisensory nature of perceptual consciousness. In D. J. Bennett & C. S. Hill (Eds.), *Sensory integration and the unity of consciousness* (pp. 15–36). The MIT Press. <https://doi.org/10.7551/mitpress/9780262027786.003.0002>
- Bourget, D. (2017). Representationalism and sensory modalities: An argument for intermodal representationalism. *American Philosophical Quarterly*, 54(3), 251–267. <https://www.jstor.org/stable/44982142>
- Briscoe, R. E. (2016). Multisensory processing and perceptual consciousness: Part I. *Philosophy Compass*, 11(2), 121–133. <https://doi.org/10.1111/phc3.12227>
- Brogaard, B., Chomanski, B., & Gatzia, D. E. (2020). Molyneux's question and the semantics of seeing. In G. Ferretti & B. Glenney (Eds.), *Molyneux's Question and the History of Philosophy* (pp. 195–215). Routledge.
- Brogaard, B., & Chudnoff, E. (2018). Multisensory consciousness and synesthesia. In R. J. Gennaro (Ed.), *The Routledge Handbook of Consciousness* (pp. 322–336). Routledge.
- Chalmers, D. J. (2004). The representational character of experience. In B. Leiter (Ed.), *The Future for Philosophy* (p. 0). Oxford University Press. <https://doi.org/10.1093/oso/9780199247288.003.0007>
- Cheng, T. (2015). Obstacles to testing molyneux's question empirically. *I-Perception*, 6(4), 2041669515599330. <https://doi.org/10.1177/2041669515599330>
- Connolly, K. (2013). How to test Molyneux's question empirically. *I-Perception*, 4(8), 508–510. <https://doi.org/10.1068/i0623jc>
- Dainton, B. (2002). *Stream of consciousness* (0th ed.). Routledge. <https://doi.org/10.4324/9780203464571>
- De Vignemont, F. (2014). A multimodal conception of bodily awareness. *Mind*, 123(492), 989–1020. <https://doi.org/10.1093/mind/fzu089>
- Deroy, O. (2014). The unity assumption and the many unities of consciousness. In D. J. Bennett & C. Hill (Eds.), *Sensory integration and the unity of consciousness* (pp. 105–124). MIT Press.
- Ferretti, G., & Glenney, B. (Eds.). (2020). *Molyneux's question and the history of philosophy*. Routledge. <https://doi.org/10.4324/9780429020377>
- Fine, L., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., & MacLeod, D. I. A. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, 6(9), 915–916. <https://doi.org/10.1038/nn1102>
- Glenney, B. R. (2013). Philosophical problems, cluster concepts, and the many lives of Molyneux's question. *Biology & Philosophy*, 28(3), 541–558. <https://doi.org/10.1007/s10539-012-9355-x>
- Gregory, L. R., & Wallace, J. G. (1963). *Recovery from early blindness: A case study*. Heffer & Sons.
- Heim, I. (1982). *The semantics of definite and indefinite noun phrases: A dissertation*. <https://www.semanticscholar.org/paper/The-semantics-of-definite-and-indefinite-noun-:-a-Heim/99b1804689ec67dbc794e6707262901dd76085a8>
- Held, R., Ostrovsky, Y., Gelder, B. de, Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, 14(5), 551–553. <https://doi.org/10.1038/nn.2795>
- Humayun, M. S., Clegg, D. O., Dayan, M. S., Kashani, A. H., Rahhal, F. M., Avery, R. L., Salehi-Had, H., Chen, S., Chan, C., Palejwala, N., Ingram, A., Mitra, D., Pennington, B. O., Hinman, C., Faynus, M. A., Bailey, J. K., Johnson, L. V., & Lebkowski, J. S. (2024). Long-term follow-up of a phase 1/2a clinical trial of a stem cell-derived bioengineered retinal pigment epithelium implant for geographic atrophy. *Ophthalmology*, 131(6), 682–691. <https://doi.org/10.1016/j.ophtha.2023.12.028>
- Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, & M. Stokhof (Eds.), *Formal methods in the study of language* (pp. 277–322). University of Amsterdam.
- Kral, A., & Sharma, A. (2023). Crossmodal plasticity in hearing loss. *Trends in Neurosciences*, 46(5), 377–393. <https://doi.org/10.1016/j.tins.2023.02.004>
- Kurson, R. (2008). *Crashing through - the extraordinary true story of the man who dared to see*. Random House Trade Paperbacks. <https://www.penguinrandomhouse.com/books/96299/crashing-through-by-robert-kurson/>
- Levänen, S., Jousmäki, V., & Hari, R. (1998). Vibration-induced auditory-cortex activation in a congenitally deaf adult. *Current Biology*, 8(15), 869–872. [https://doi.org/10.1016/S0960-9822\(07\)00348-X](https://doi.org/10.1016/S0960-9822(07)00348-X)
- Locke, J. (1978). Letter to William Molyneux. In E. S. de Beer (Ed.), *The correspondence of John Locke* (Vol. 4). Clarendon Press.
- López-Bendito, G., Anibal-Martínez, M., & Martini, F. J. (2022). Cross-modal plasticity in brains deprived of visual input before vision. *Annual Review of Neuroscience*, 45, 471–489. <https://doi.org/10.1146/annurev-neuro-111020-104222>
- Matthen, M., & Cohen, J. (2020). Many Molyneux questions. *Australasian Journal of Philosophy*, 98(1), 47–63. <https://doi.org/10.1080/00048402.2019.1603246>
- Mcgurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748. <https://doi.org/10.1038/264746a0>

Gatzia, D., & Brogaard, B. (2024). Molyneux's question and multisensory integration. *Philosophy and the Mind Sciences*, 5. <https://doi.org/10.33735/phimisci.2024.11620>



- Molyneux, W. (1978). Letter to John Locke. In E. S. de Beer (Ed.), *The correspondence of John Locke* (Vol. 3). Clarendon Press.
- Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *NeuroImage*, 59(1), 781–787. <https://doi.org/10.1016/j.neuroimage.2011.07.024>
- Nudds, M. (2001). Experiencing the production of sounds. *European Journal of Philosophy*, 9(2), 210–229. <https://doi.org/10.1111/1468-0378.00136>
- O’Callaghan, C. (2012). Perception and multimodality. In E. Margolis, R. Samuels, & S. P. Stich (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (pp. 92–117). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195309799.013.0005>
- 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 University Press. <https://doi.org/10.1093/acprof:oso/9780199832798.003.0006>
- O’Callaghan, C. (2015). The multisensory character of perception: *Journal of Philosophy*, 112(10), 551–569. <https://doi.org/10.5840/jphil20151121035>
- Ostrovsky, Y., Meyers, E., Ganesh, S., Mathur, U., & Sinha, P. (2009). Visual parsing after recovery from blindness. *Psychological Science*, 20(12), 1484–1491. <https://doi.org/10.1111/j.1467-9280.2009.02471.x>
- Rescorla, M. (2020). Perceptual co-reference. *Review of Philosophy and Psychology*, 11(3), 569–589. <https://doi.org/10.1007/s13164-018-0411-6>
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implemented deaf patients are better multisensory integrators. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 104(17), 7295–7300. <https://doi.org/10.1073/pnas.0609419104>
- Schwenkler, J. (2012). On the matching of seen and felt shape by newly sighted subjects. *I-Perception*, 3(3), 186–188. <https://doi.org/10.1068/i0525ic>
- Schwenkler, J. (2013). Do things look the way they feel? *Analysis*, 73(1), 86–96. <https://doi.org/10.1093/analys/ans137>
- Schwenkler, J. (2015). Commentary: “Multimodal theories of recognition and their relation to Molyneux’s question.” *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01792>
- Schwenkler, J. (2019). Molyneux’s question within and across the senses. In T. Cheng, O. Deroy, & C. Spence (Eds.), *Spatial Senses* (pp. 274–283). Routledge.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, 408(6814), 788. <https://doi.org/10.1038/35048669>
- Síkl, R., Simecček, M., Porubanová-Norquist, M., Bezdiček, O., Kremláček, J., Stodůlka, P., Fine, I., & Ostrovsky, Y. (2013). Vision after 53 years of blindness. *I-Perception*, 4(8), 498–507. <https://doi.org/10.1068/i0611>
- Tye, M. (2007). The problem of common sensibles. *Erkenntnis*, 66(1), 287–303. <https://doi.org/10.1007/s10670-006-9035-9>
- Wassenhove, V. van, Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America*, 102(4), 1181–1186. <https://doi.org/10.1073/pnas.0408949102>
- Zrenner, E. (2013). Fighting blindness with microelectronics. *Science Translational Medicine*, 5(210), 210ps16. <https://doi.org/10.1126/scitranslmed.3007399>

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

