What Was Molyneux's Question a Question About?

Jonathan Cohen (University of California, San Diego)

and

Mohan Matthen (University of Toronto)[1]

Nearly nine hundred years ago, the Andalusian Muslim statesman and philosopher, Abu Bakr ibn Tufail (1105-1185), conjectured that all sensory qualities were material, and inferred from this that they would have at least some characteristics that can be discerned independently of sensory experience. Thus, he said, if a newly sighted man encountered colours for the first time, he would experience them in ways consonant with (at least some) descriptions of colours he had been given while still unsighted. Or perhaps he meant something stronger, namely that the newly sighted man could actually recognize the colours on the basis of the descriptions he had been given while still unsighted.[2] Of course, on either interpretation, ibn Tufail's claim is controversial: for, granting that all sensory qualities are material, it may be that at least some material qualities are known to us *only* by our sensory experience of them—and so the newly sighted man might have no prior knowledge on which to hang the descriptions he was given of colour. Nevertheless, the conjecture prepares us for the realization that some material qualities might not be available to *any* sense, and others to more than one.

The translation of ibn Tufail's book into Latin by Edward Pococke in 1671 sparked a lively discussion of the empirical construction of concepts. (See Goodman, this volume.) William Molyneux pursued ibn Tufail's conjecture in a narrower and less controversial

---

[1] This is a fully collaborative work of the two authors.

[2] In this formulation, ibn Tufail's conjecture concerning colour does not turn on amodal or inter-modal knowledge. In Lenn Goodman's translation (this volume), the crucial passage is

> Suppose . . . his eyesight were restored and he could see. He would walk all through the town finding nothing in contradiction to what he had believed, nor would anything look wrong to him. The colors he encountered would conform to the guidelines that had been sketched out for him.

This leaves it open whether or not visually based "guidelines" could enable him, in Molyneux's words, to "distinguish and identify by sight alone" which of an apple and a leaf was red and which was green. For suppose he had been told something like: "Red is a dark, warm colour that is very unlike green, which is bright and cool." C.L. Hardin's (1988) discussion of spectral inversion (*ibid*. 134-54) could be taken to imply that he might actually be able to identify red and green on this basis. If so, then, perhaps ibn Tufail was right.

form.[3] Famously, he asked Locke (who would likely have recognized the allusion) whether a newly-sighted man could, by sight alone, distinguish and identify a globe and a cube, which he previously knew only by touch. This is a specific problem about qualities that can be sensed in more than one modality—Is there an inter-modally transferrable, or amodal, element in the identification of three-dimensional solids? Though derived from ibn Tufail, this question leads in a different direction. Our purpose here is to ask how it generalizes. What was the theoretical nerve that it is supposed to touch? Why is it so revealing?

The problem that guides us in this paper is more pointed than Molyneux's about sensory knowledge in a newly acquired modality. We ask: How are ideas formed in each modality, and, given how they are formed, what cross-modal correspondences can we expect to find? This question invites a more granular line of inquiry—one that takes us into details of the construction of specific qualities. But it makes ibn Tufail's conjecture and Molyneux's question unexpectedly relevant to a number of unconnected, but familiar problems that arise out of contemporary theorizing about perception. For, in contrast with many familiar approaches, ours poses questions about cross-modal divergence and convergence that range more widely than those envisaged by eighteenth century philosophers. These questions are neutral about sensory carry-over of the kind queried by Molyneux.[4] In short, it turns out that there are more questions and more answers on this terrain than Molyneux, Locke, and most of their successors realized.

I. On the origin of general ideas

Very likely, Molyneux and Locke thought of the question like this. We know certain shapes by touch. We also know the very same shapes by sight. Consider, then, the *idea* of a shape given to us by touch—the mental representation of a shape that we arrive at by touching it ("the Object of the Understanding when a Man thinks," as Locke puts it). Is this the same as the idea of the same shape given to us by sight? Or are these ideas different? To

[3] Whether he came across the topic by reading ibn Tufail or through the ensuing discussion is not clear.

[4] In fact, one way that our approach is neutral is that it does not turn on any particular demarcation of the sense-modalities. In Matthen and Cohen (forthcoming), we even note the possibility of within-modality Molyneux Questions.

repeat: these questions are not the same as asking whether there can be tactual and visual ideas of the same shape. There can be different ideas of the same thing.

Molyneux's question is about the ability to identify a shape, or to distinguish it from others. Whether or not he and Locke were clear about it, this is a question about general ideas, i.e., ideas of repeatable types—Can the newly-sighted man visually recognize this particular thing here as an instance of a general idea that he earlier arrived at by tactual experience of other things? According to Locke, a general idea is an abstraction created by discarding irrelevant characteristics of particulars. For instance, the general idea of a triangle is "neither Oblique nor Rectangle, neither Equilateral, Equicrural nor Scalenon; but all and none of these at once." (*Essay* IV.vii.9) The question Molyneux poses is not about all aspects of visual or tactual experience of globes and cubes. A particular globe might look red, shiny, and about a foot across; at the same time, it might feel cold, dry, and about a foot across. Some of these attributes of the globe are purely visual, some purely tactual; one is (apparently) shared. None of them, however, are essential to the general, repeatable idea, GLOBE, as such;[5] they are all discarded in the psychological act of generalization that terminates in that general idea. This invites the question: What qualities are retained when a blind man forms a general idea of a globe from his tactual experience?

Now: It is clear that in at least some cases, one can, by discarding components of an idea originally experienced by touch, arrive at a residue is not specifically tactual—the general ideas THREE-DIMENSIONAL, SYMMETRICAL, EQUALLY SLOPED are examples of this. So when we consider the general idea, GLOBE, formed from touch, we should ask: Are the retained qualities specifically tactual, or does the resultant general mental representation subsume the visual presentation as well, as a presentation of the very same repeatable type? And this naturally raises the question whether the general idea that results from normal visual processing is the same or different in content from the idea that the blind man possesses.

Molyneux's Question has sometimes been thought to be about the "differences between the qualitative or phenomenal character" of visual and tactual experiences of the

[5] We capitalize terms that stand for ideas and italicize those that stand for properties.

same thing (see e.g., Schwenkler 2019). For the subjective *experience* of exploring a globe by touch is obviously quite different from that of examining it by sight[6]—and this is clearly a relevant consideration. But Molyneux's Question is not about the particular experiences that a newly-sighted man suffers when he looks at a globe. Rather, it is about this man's ability to apply a general idea—i.e., a representation of a repeatable perceptual property—that he obtained by touch to objects newly encountered in visual experience. Since general ideas omit certain features of particular experiences, we have to ask whether the retained characteristics are inter-modally comparable. In other words: Can the subjectively different experiences of the same shape thrown up by two different sense modalities be subsumed under the same general idea, or at least general ideas that can be compared across phenomenal differences?

Locke answered Molyneux negatively. And perhaps he did so because he was not alert (in that context) to how vision and touch can arrive at the same general ideas by discarding specifically modal aspects of experience.[7] But even if we disagree with Locke about this, there is another question that arises for anybody who was inclined to think that general ideas of shape are cross-modally comparable. Allow that a general idea arrived at by touch alone could in fact be the same in content as that arrived at by vision. We still do not know whether it is equally apt to be *applied* to a particular idea by vision. Take the idea BOUNDED OBJECT. This is clearly an idea that has application in more than one modality. Yet, its method of application might differ in different modalities: specifically, in touch, vision, and audition. So, it could very well be that a newly sighted man might be hesitant in identifying bounded objects visually.

Coming back to the general idea, GLOBE, the point to be emphasized here is that shared content is not the only question of interest that Molyneux raised. Concede, in other words, that there is only one general idea in play, or that there are two but with significant cross-modal structural correspondences of content. The question still remains: Can one just

[6] But see Campbell 1996 (304), who thinks (on naïve realist grounds) that the shape itself constitutes, indeed individuates, any experience of it.

[7] In *Essay* II 5, Locke writes, "we can receive and convey into our minds the ideas of the extension, figure, motion, and rest of bodies, both by seeing and feeling." For discussion of this and related points, see Matthen and Cohen, forthcoming.

as easily use the tactually formed idea to "distinguish and identify" a shape as a globe (rather than a cube) when looking at it as when touching it? And: Is the perceptual process by which retained qualities are identified analogous or different?

II. Experience, ideas, and properties

When addressing Molyneux's Question, one can ask about at least four different contrasts between visual and haptic perception.

(1) Experience of seeing a globe vs experience of feeling it by touch.

(2) Idea generalized from visual experience vs idea generalized from touch.

(3) The property revealed by vision vs the property revealed by touch.

(4) The ability to re-identify by vision the property revealed by touch.

Much of the philosophical dialectic regarding MQ revolves around the connections among the first three contrasts. We'll follow this dialectic for the most part, though we will make mention of the additional contrasts with (4).

On one hand, it is clear that the experiences (1) are different. Some are inclined to think that this settles the question about the ideas, (2). An empiricist might argue that one cannot erase the sensory source of one's general ideas in the relevant cases. If so, then a negative answer to Molyneux's Question follows: the general ideas are different.

On the other hand, *sphere* is a geometric or spatial property that entails qualities such as symmetry and roundness. Some writers (e.g., Bennett, 1965, and Evans, 1985) hold that these qualities are inter-modally comparable. According to them, there is no relevant difference between the items in (3). And this, they think, settles, or at least severely constrains, the question about ideas—SPHERE must be the same general idea across vision and touch, or at least structurally similar enough to distinguish it from CUBE.

Both sides err by neglecting the bridge from their starting point to the middle term—that middle term being (2) *general ideas*. For both, the question is: Is sensory experience of a geometrical solid sufficient to form a modality-neutral subjective representation of the solid? An empiricist like Locke answers 'No' because of the difference in the experiences; Bennett and Evans answer 'Yes' because of the alleged non-modality of the sensed properties. But it's not clear exactly where the answer turns from negative to positive when we go from (1) to (3). As against Locke, there's the possibility that all tactually specific dimensions of the idea are bleached out by the generalization process before we get to (2). As against Bennett and Evans, there's the possibility that even if our ideas of shape represent an amodal, or intermodally comparable, property, the general ideas formed from modally specific experience are sufficiently tainted by their modal origins as to be not only distinct, but also difficult, or even impossible, to apply in vision when they originate in touch.

This sort of question can be asked about other general ideas as well. Ibn Tufail's conjecture was that colour could be known (at least descriptively) by touch. This is implausible, because colour is a specifically visual idea (but see note 2 above). But what about other features known by more than one sense? Not all of these are shapes; indeed, not all are spatial. For instance, what about number? If the newly-sighted man were shown one globe on his left and two on his right, could he distinguish and identify the cardinality of the two collections by sight alone? And (as Gareth Evans asked) what about temporally extended experiences: if the newly sighted man were shown a steady light on the left and a pulsating one on the right, could he tell which was which? One could ask: is there a general recipe for answering questions of this sort? Given that we don't know in advance what will be omitted in any given process of generalization, it doesn't seem that there could be.

In earlier work (Matthen and Cohen forthcoming), we argued that though a negative answer is warranted in a wide variety of cases, this conclusion can only be reached empirically and piecemeal. Both empiricists like Locke and their opponents like Evans are wrong to treat Molyneux's Question as a single question about the origin of ideas in different modalities.[8] Rather, it is properly posed as a question about specific ideas in specific contexts. Thus, we posited "Many Molyneux Questions," and argued for a classificatory scheme of these by the spatiotemporal dimensionality of the property represented by the idea in question.

In this paper, we venture beyond this conclusion in three ways.

---

[8] Cf. Glenney 2013.

We argue, first (sections III-V), that in each modality, complex ideas may differ structurally from the corresponding property. This sheds light on how (contrary to the conclusions of Gareth Evans) *non*-correspondent ideas of the same shape might be constructed from an inter-modally shared representation of space. In these cases, the answer to Molyneux's Question is No: the newly sighted man does not have anything on which to base visual identification.

Second, we show (sections VI-VII) that in certain cases, different modalities contingently construct structurally correspondent ideas of shape- and space-related properties. In these cases, we get a somewhat surprising Yes, if not to Molyneux's Question as posed, then at least to the question whether there are significant cross-modal correspondences. (Of particular interest in this argument: the "new" modalities created by sensory substitution.)

Finally (section VIII), we ask (but leave unanswered) new questions in the style of ibn Tufail and Molyneux regarding a significant class of non-spatial ideas, namely those involved in evaluative perception and perceptual affect.

III. A New Question: Analyzing ideas

Let's now consider a different path into Molyneux's Question, similar in some ways to the one originally explored by Denis Diderot (1749/1951) and discussed by Gareth Evans (1985). Locke distinguishes between simple and complex ideas. Simple ideas are not put together from other ideas; complex ideas are. For any complex idea, one can ask: How is it constructed from simples? Now, suppose (as many philosophers have) that the simple ideas of touch and vision are always point-located qualities.[9] Then, it would follow that shape-ideas, such as GLOBE and CUBE, which are about spatial distributions of points, are complex. So, if one could establish that the process of constructing the idea SPHERE from point-located qualities is structurally different in touch and in vision, this would give us reason for thinking that this idea is modally specific.

To start with, the question whether a newly sighted man is able to recognize a tactual complex idea *without using touch* is different from the question whether he is able to

[9] This is the assumption of the colour mosaic theory sketched by Lewis 1966.

recognize it *by vision*. For consider the complex idea, RS=ROUND AND STATIONARY. Hold in abeyance the question whether the newly sighted man would visually be able to identify the simple ideas, ROUND and STATIONARY. Allow this man whatever time and experience is sufficient to learn these ideas. And now put the question: what more would it take for him to be able to identify instances of RS? Presumably nothing more. AND is an analytic operation, and therefore not one whose acquisition depends on sensory experiences that the newly sighted man might be thought to lack. So, it is extremely plausible to think that he needs nothing more to learn RS than how to identify its simple components. But this argument does not show that RS can be applied *visually*. Of course, we have reason for believing that the conjunctive idea AND is deployed in cognition. But it is a substantive empirical question, not something that can simply be taken for granted, that AND can be applied straightforwardly in the construction and deployment of ideas within vision.

Second, note also that one cannot safely conclude that AND is, indeed, applied in the formation of a visual idea merely from the fact that the idea picks out a property that has a conjunctive analysis. Some small animals preyed on by large birds have a distinctive behavioural response (they freeze) to visually detected instances of the conjunctive property *dark and looming* (but not to visually detected instances of the property *dark* or of the property *looming*). Presumably they have a visual representation/idea of (i.e., whose content is) the complex, conjunctive property. But it does not follow that this idea is a complex construction with DARK, LOOMING, and AND as constituents: it could well be that their idea of the property *dark and looming* is an uncomposed atom. (As an aside: this would cast doubt on the notion that all simple ideas are point-located qualities.) In sum, it is wrong to think that if a property can be analysed as a conjunction, then the *idea* of that property will be conjunctive (and hence complex). The logical or mathematical structure of a property need not be reflected in the structure of the sensory idea of that property (cf. Hopkins, 2005).

On reflection, it seems that Diderot's approach to Molyneux's Question also leans on an analysis of the property, though in a somewhat different way. Diderot famously holds that vision and touch must have different ways of apprehending the mathematical structure of shape-properties. He assumes (mistakenly, as we contend) that since shape-properties are spatial distributions of points, shape *ideas* must also be constructed from ideas of spatial distributions. He reasons that, since one can at a single glance form an impression of how spheres look, the visual idea can be formed by a simultaneous apprehension of the spatial relationships of constituent points. In contrast, he supposes that the tactual idea of these relations can only be formed by a temporally extended process: one cannot apprehend the whole sphere simultaneously by touch, and so one has to rely on moving one's hands over it. He concludes that the tactual idea of a sphere is different from the visual idea—the former has a kind of temporal structure that the latter lacks.

But this line of argument is mistaken. It is true that vision and touch differ in the spatial range over which they deliver information at a single instant. However, this is precisely the sort of difference that is put aside in the abstractive process that leads to general ideas. For consider: if a sphere is sufficiently small relative to the spatiotemporal range of vision then it can be visually apprehended at a moment, and one can then immediately form from it a visually generated general idea of sphericality.[10] (In this process, one discards size information, inter alia.) And exactly the same is true of touch: if a sphere is sufficiently small relative to the spatiotemporal range of touch, one can apprehend its shape at a single instant and form from it a haptically-generated general idea of sphericality (which discards size information, inter alia).[11]

Of course, Diderot is right that if the sphere is too large to be haptically apprehended at a time, then one needs to engage in temporally extended haptic exploration to form from it a general idea of sphericality. But the same is true of vision, with the difference that the spatiotemporal extent of the visual glance is larger than that of touch. But this is irrelevant, given that size information is discarded when we form the idea, SPHERE. It would seem, then, that, with respect to the perceptual acquisition of general ideas, the modalities of vision and touch are on a par: Diderot's alleged structural contrast between the two rests entirely on a modal difference in spatiotemporal range that is properly ignored in consideration of the perceptual acquisition of general ideas. He simply overlooks the possibility that a general

[10] See Klatzky and Lederman (1995) on the haptic glance.

[11] In neither case does the apprehension at issue demand perceptual contact with all of the points in the sphere; rather, what is required is presumably just the sort of perceptual contact with facing surfaces that supports acquisition of a general shape idea. This is another way of reinforcing the point that the structure of our perceptually derived ideas need not be the same as that of the corresponding properties.

idea of SPHERE could be constructed from an experience of a sphere that is small enough to be apprehended at an instant, whether it is a hand-size sphere apprehensible by touch or a much larger one apprehensible by vision.

Moreover, and even putting the last point aside, Diderot neglects the possibility that even if touch apprehends particular spheres temporally, the general tactual idea of a sphere could be formed by discarding temporality. (Evans gives some anecdotal evidence in favour of this possibility.) But even then, it would not follow that touch and sight are equally adept at making the comparison. Even if both modalities have a non-temporal idea, they might have different ideas. Or they may have the same idea (or structurally comparable ones); nonetheless, blind subjects may still not be able to make the cross-modal comparison in a way that enables them immediately to visually recognize spheres when they recover sight. And, as a matter of fact, Pawan Sinha and colleagues report empirical evidence that indicates an inability of newly sighted people to recognize shapes (Held et. al. 2011; Ostrovsky, Andalman, and Sinha 2006; Ostrovsky et. al 2009). In short, we have no grounds for a *general* answer, or even grounds to formulate a general question.

To summarize our line of thought in this section: It is wrong to assume that a sensory idea of a shape must reflect the analysis of the shape-property. A sphere is a certain distribution of points in space, but it doesn't follow that the visual idea of a shape is a representation just of that distribution—it may instead reflect gross characteristics distributed across a surface, for example, a smooth lighting gradient or a certain kind of textural variation. For this reason, we can't assume that touch and vision recognize spheres in the same way. If each relies on clues and indications rather than the actual geometrical analysis of a sphere, they might rely on different cues. Or they might rely on comparable cues (see section VI). The question is open.

This informs our general approach to inter-modal transfer: the question has to be put for specific ideas of specific shapes, and not for whole modalities. Interestingly, the answers given by Locke, Diderot, and many contemporary philosophers (including Gareth Evans) follow something much closer to a whole-modality strategy. However, we suggest, reflection on Molyneux's question about globes and cubes motivates a finer-grained approach.

IV. Space and the Modalities: Sceptical Remarks

On its face, the question Molyneux poses is about shape. But Evans attempts to reduce this to an issue about *space*—specifically, about "the relation between the perceptual representation of space attributable to the blind, and the perceptual representation of space available in visual perception." This is wrong for a number of reasons. The shift from shape to space occurs against the background of Diderot's argument. As we have seen, Diderot thought that the blind can perceptually register shape only by temporally extended tactile exploration, and that their idea of shape is therefore that of a process. Evans rightly rejected this argument.

Evans is right to reject a facile inference from temporally extended experience of shape to temporally extended representation of space, but his critique does not touch Diderot. For a follower of Diderot could concede that the blind operate with a "simultaneous" conception of space when they are locating individual points, while still insisting that temporal-process notions have to be collated over time in order to get ideas of shape. Non-simultaneous shape-representations are compatible with an underlying simultaneous space-representations.

More significantly, both Diderot and Evans overlook the fact that there could be more than one way to keep track of position (Tolman 1948, O'Keefe and Nadel 1978, Gallistel, 1990). Some animals use travel metrics ("Walk steadily for ten minutes, slow down, turn right, . . ." etc.) to measure position, as inertial navigation systems do. This doesn't preclude them taking in spatial positions in a single glance. The test of whether an organism is sensitive to allocentric spatial layout is its ability to get from C to A after first going from A to B and B to C, where A, B, and C are not collinear. Similarly, the test of whether it represents an object's shape is its ability to trace the shape from multiple different starting points—top, bottom, middle. Touch and vision might be able to keep track of position relative to a starting point without employing the same pre-existing representation of space. In fact, there are significant structural differences between touch and vision regarding how they represent space (Matthen forthcoming). So, coordinating the two senses is not straightforward.

A historical aside. Ibn Tufail's conjecture about the newly sighted man occurs in his story, *Haiy ibn Yaqzan*, which is a philosophical allegory about a boy cast away at birth on an island. Brought up by a gazelle, and utterly deprived of human contact, this boy comes to know everything that advanced cultures know, from the manufacture of clothing, to morals and religion, to cosmology. In the seventeenth century, this book was widely read, supposedly influencing Spinoza, Hobbes, and Locke (independently of Molyneux's letter), and inspiring Defoe's *Robinson Crusoe*. The newly sighted man is a thought experiment in the Introduction to the book; it elaborates the theme of knowledge in the absence of exposure to the most direct sources of knowledge: "Colours were such as he had before conceiv'd them to be, by those descriptions he had receiv'd." Molyneux cleverly plucked this anecdote from the book (or from the contemporary discussion surrounding it) and sharpened it. But since it is couched in such specific terms, it is difficult to know exactly how he viewed the problem. Locke himself didn't get the point at first glance, though he probably knew ibn Tufail's book. Presumably, Molyneux intended to query ibn Tufail's theme of knowledge without direct experience. Cross-modal carry-over isn't the only consideration relevant to this theme, nor (as Evans realized) is associative learning vs innateness. This is why we cast the net wider.

Given this historical background, it is entirely possible that Molyneux thought about his problem about cross-modal comparisons and associations simply as an instance of a much wider puzzlement. You have all sorts of non-spatial perceptual ideas of familiar things—of your loved ones, of the place where you live, of your clothes, and so on. You have many feelings evoked by perception: remember the terror induced in small animals by looming shadows. Are these non-spatial ideas transferrable from one modality to another, or comparable between them? We'll return to these questions below, but for now let's just acknowledge that there is no reason to think that Molyneux's Question was ultimately about just shape. (In fact, Molyneux's original formulation of the question to Locke explicitly invokes distance: "Or Whether he Could know by his sight, before he stretchd out his Hand, whether he Could not Reach them, tho they were Removed 20 or 1000 feet from him?") And even if it was, there is, more importantly, no reason to think that it reduces without remainder to a problem about the representation of space.

V. Shape: Foundation and Integration

Evans holds that our ideas of location and/or direction are shared by more than one modality. Suppose he is right. Let a light be switched on. Can a newly sighted man point to it, or at least tell whether it is on his left or his right? Can he make inter-modal comparisons among these presentations? It is not clear: evidence suggests that these comparisons emerge over a complex interactive developmental trajectory (Thelen, et al, 2001). Either way, this does not settle the issue about shape. Suppose that a neonate has the ability to locate each of a collection of points by both vision and touch. Suppose further that she can discriminate by touch whether these points are collinear or not. Would she be able to determine *by vision* whether these points are collinear or not? And, similarly, the newly sighted person? There is no a priori reason to think that the answer to these questions is 'yes' and no reason to think that it is 'no' either. The ability to locate points in space is not the same as the ability to determine whether they stand in complex spatial relations. The senses may share a representational framework for space without sharing a representational framework for shapes.

Given these observations, we propose to break the question of inter-modal comparability of general ideas of shapes up into parts as follows.

1. *Spatial Foundation*. Does the ability to locate points by touch have the same representational basis as the ability to locate points by sight?

2. *Integration Step*. Given a set of points that have been located by the Spatial Foundation of touch and also that of vision, is the ability to determine by touch whether they constitute an instance of a shape S structurally similar to the ability to determine this by vision? And if so, would this help, or even enable, a newly sighted person to recognize instances of S by vision alone?

Our general view is that these questions have to be tackled piecemeal. The ability to locate points may be context dependent; the formation of general ideas of shape S might employ the same integrative tools in two modalities, but this might not be the case for shape S'.

Even if Evans is right that the Spatial Foundation question should be answered positively, this does not address the question of the Integration Step. And as we argued in the preceding section, it is plausible to think that different modalities might address this differently. In Matthen and Cohen (forthcoming), we taxonomized the Integration Step by the dimensionality of shapes. This allowed us to bring a number of different empirical studies to bear on the issue, because, though they are not initially addressed to Molyneux's Question, they could be interpreted as addressing the Integration Step.

In all of the cases we reviewed, we found that touch and vision follow different strategies in the Integration Step. So, in these cases, the answer was indeed No. But we did not arrive at this conclusion by a general argument; we had to adduce cases to prove each No. Thus, there is no path to a general answer to Molyneux's Question. Now, as we'll argue in the next two sections, there are cases where the Integration Step is structurally similar across modalities. Since the integrative strategies followed by different mechanisms for perceptual feature extraction are diverse, there is no reason to expect a uniform answer across all features.[12]

VI. Beyond Shape: The Perceptual Representation of Faces, Biological Motion, and More

So far, we have largely followed mainstream philosophical tradition in construing Molyneux's question as, in the first instance, directed at the perceptual representation of shape. But our articulation of the question into foundational and integration components invites analogous questions about the foundational and integrative roots of non-shape properties as well. For any property that can be apprehended by multiple modalities, we can ask: What are the foundational elements and integrative procedures enlisted by the different modalities? And are these sufficiently similar between the modalities to allow crossmodal transfer (immediately, with certainty, etc.) for ideas of these properties?

This way of thinking of MQ is more fine-grained than our earlier dimensional taxonomy in that it allows for obstacles to crossmodal transfer between pairs of modalities even within a single spatiotemporal dimensional regime, and even when there can be other

[12] Note that our pluralism about Molyneux's Question arises from our concern about particular perceptual strategies for the recovery of individual features; Glenney (2013) and Hopkins (2005) endorse pluralism about MQ as well, but for different reasons.

cases of successful crossmodal transfer between those same two modalities and within the same spatiotemporal dimensional regime.

To see some of the potential obstacles, and with the caveat that, as we've learned from the shape MQs, we can't expect our answers to such questions to generalize automatically between cases, we want to review a number of instances where the evidence suggests that the foundational and integration steps may be modally specific, hence where crossmodal transfer may fail.

As an initial example of this type, we can consider the integrative processes at work in the visual representation of human faces.[13] Though there are a number of ways to bring out the special features of these processes, one of the most striking comes from the so-called Thatcher effect—the finding (Thompson 1980) that local anomalies in the geometric organization of faces are dramatically harder to detect in upside-down than in right-side up faces (see figure 1). This result gives us reason for thinking that at least some of the integrative processes involved in the visual representation of faces can't be understood simply as those involved the ordinary visual integration of metrical spatial information—orientation matters as well. Faces, then, are treated as a kind of spatial form, but a kind for which vision reserves special treatment. (This is yet another way of making the point we made at the end of section III: perceptually informed ideas of a property need not be composed the same way as the property itself.)

[13] Material in this and the next section overlaps with Cohen (2018).

Figure 1: The Thatcher Effect

There are several other effects suggesting that visual integration for the representation of faces is special relative to ordinary spatial form. Thus, detection times are significantly faster for faces than for upside-down faces or arbitrary assemblages of facial features in non-face arrangements (Purcell and Stewart 1988; Farah et. al. 1995; Yin, 1969). Visual face recognition is susceptible to characteristic overgeneration errors: we (mis-) recognize a face in a cloud or a mountain far more readily than other objects. It can be selectively impaired in congenital or acquired defects (prosopagnosia) that spare other aspects of visual processing (Barton 2003). And it is no surprise, given these and similar findings, that vision devotes special resources to face perception, in the form of dedicated processing (Kanwisher 2010; Sinha et. al. 2006; Sugita 2009), carried out in specialized neural areas (Liu, Harris, and Kanwisher, 2010; McCarthy et. al. 1997). (But see below.)

Suppose this is all correct — that, though human vision is capable of integrating a wide range of two-dimensional spatial forms from lower dimensional spatial information, its integration of faces enjoys certain advantages relative to that of other (equally complex) spatial forms within this dimensional regime. Should we expect this fact about vision to hold of other modalities capable of representing spatial form? Not at all. There is no *ex ante* guarantee that a second modality, *m*, that also integrates spatial information in two dimensions, will also carry out its integrations of faces in the special ways in which vision appears to do. But if not, then we should be unsurprised to find failures of crossmodal transfer between vision and *m* for representation of faces (or other properties to which the two modalities are differentially attuned).

We can test this conclusion by comparing face perception in vision against face perception in two other modalities capable of representing two-dimensional spatial form: ordinary haptic touch, and an artificial modality for sensory substitution that presents the outputs of visual transducers to (normal) audition: the Prosthesis for Substitution of Vision with Audition (PSVA) of Capelle et al, 1998.

Consider first the comparison to haptic touch. One significant difference between visual and haptic integrations of two-dimensional form is that the two modalities differ in the lower-dimensional foundational information from which the integrations are performed: vision works from point-colours (or point-intensities) in an external space, while haptic touch works from point-pressure in bodily space (often accompanied by awareness of willed motion of our hands and other body parts). However, despite these foundational-step differences in the lower-dimensional data available to the two modalities, it turns out that vision and haptic touch converge on the results of form integrations in a wide range of cases. Moreover, and perhaps yet more surprisingly, there is evidence that haptic touch treats facehood, in particular, as special in some of the ways that vision does. Specifically, the existence of a Thatcher effect for haptic touch that parallels the Thatcher effect for vision (Kilgour and Lederman 2006) provides strong evidence that haptic touch, like vision, is specially attuned to facehood, relative to equally complex two-dimensional spatial forms (Lederman, Klatzky, and Kitada 2010). And, since the two modalities align in apportioning importance (i.e., computational resources) to this one property of facehood, it's not too surprising to learn that there is, indeed, bidirectional cross-modal transfer for facehood between them in priming studies (Reales and Ballesteros, 1999; Easton et al.,1997a, b;

Hadjikhani and Roland, 1998; Kilgour and Lederman, 2002; Casey and Newell, 2003, 2007; Norman et al., 2004).[14]

Now consider the perception of two-dimensional form through PSVA. Here, again, the foundational data from which the modality integrates two-dimensional form information—consisting of auditory pitches organized temporally, accompanied by awareness of willed motion of the head/mounted receptor array—is quite different from the data on the basis of which vision performs its form integrations. Should one expect that there would be significant differences in the integrative processing methods employed by PSVA and vision? Our point is that what we know about vision and haptic touch gives us no firm basis for answering this question one way or the other. In particular, we cannot assume that the form integrations corresponding to faces will be favoured under PSVA (as they are in vision and haptic touch) over those corresponding to equally complex forms. And there is no evidence of a Thatcher effect for PSVA.[15]

What can we conclude from the above evidence? First, it seems that there are certain similarities between vision and touch with regard to the construction of face-ideas. This suggests that there will be some cross-modal carry-over between the two modalities. Of course, it would be far too hasty to jump to a positive response to the Molyneux-type question whether a newly sighted man could recognize the face of his lover by sight alone.[16] And, indeed, it would be too quick even to conclude that he could immediately make judgements like "That looks like a hooked nose." For as we noted earlier, similarity of content does not imply equal ease of application. Nevertheless, we can expect some cross-modal parallels between visual and haptic face perception that do not hold with respect to the visual and haptic perception of spatial form in general. It would not be surprising, for example, to

[14] Investigators have taken these findings to suggest that vision and touch processes share at least some structural representations of faces. It remains controversial whether such shared representations are more closely allied with a format used by one of the modalities (which would mean that interaction with the other modality would involve a certain amount of representational remapping), or whether they are expressed in a common and modally unspecific format.

[15] There are other sensory substitution systems that transform visual input to auditory output; consistent with our general approach, we also should not assume that any conclusions about PSVA will extend to these.

[16] Brian Glenney notes that face-blindess is a common deficit among the newly sighted (Fine et al, 2003; Cattaneo and Vecchi 2011: 98-102). Whether this is related to Molyneux's Question is unclear.

find that the newly blind person would preferentially pick up on certain cues for face recognition or be quicker to learn it under haptic presentation. So, while we would not have an answer to Molyneux's question for faces, we *would* expect some positive answers in the general vicinity.[17]

However, we cannot assume that any such correspondence would generalize to other cases. As noted, there is a mismatch between the capacities for integrating faces in vision (where such integrations are given special status relative to other integrations built from spatial foundational data) and in PSVA (where they are not). Under these conditions, one would not expect that there should be crossmodal transfer of integration for facehood between the two modalities.[18]

Having made these observations, it is now natural to ask whether and how widely the special integrative processes at work in vision (and perhaps other modalities) extend. On the evidence, there do appear to be other types of two-dimensional spatial forms, beyond human faces, to which vision (at least) applies special integrative processes. Take, for instance, the visual system's ability to classify automatically, quickly, and easily, certain but not other, equally complex, patterns of a few moving points of light as reflecting the motion of key joints in a moving organism (Johansson 1973).

[17] This is broadly consonant with the differences in learning speed for different cues reported by Sinha and colleagues for shapes. See Matthen and Cohen forthcoming: section VI.

[18] Specifically, one would not expect to see crossmodal transfer of integration for facehood from the modality attuned to the property (vision) to the modality not attuned to the property (PSVA); this leaves it open that there might be transfer in the other direction.

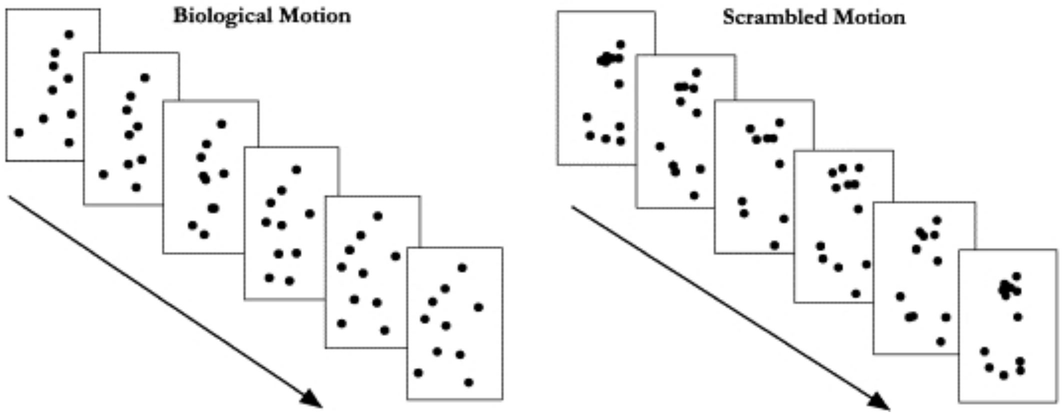

Figure 2: Subjects visually distinguish moving patterns that reflect the motion of key joints in an organism from random, but equally complex, patterns.

Like visual perception of faces, visual perception of biological motion appears to involve specialized processing (Lu 2010), to be carried out in specialized areas (Allison et al 2000; Grossman, et al 2000), and can be spared in injuries that damage gross motor and other spatial abilities (Jordan et. al 2002; Kim et. al 2008). Moreover, there is a Thatcher effect for biological motion as well: it is significantly easier to detect local anomalies (anomalies that perturb the visual classification of the motion as biological) in displays that are right side up than in displays that are inverted (Troje and Westhoff 2006; Mirenzi and Hiris 2011).

And, indeed, there is evidence that at least some of the criteria we have appealed to in the foregoing extend to yet further properties (especially with perceptual learning). Thus, Twedt, Sheinberg, and Gauthier 2007, and Wong et. al. 2010 find evidence of a Thatcher effect in vision for a range of non-face object types, including cars, buildings, bikes, and letter strings. And researchers have found face-like reduced reaction times in visual recognition tasks and disproportionate costs for recognizing inverted figures by experts with a wide range of non-face domains, including particular dog breeds (Diamond and Carey, 1986), and novel artificial ("greeble") figures (Gauthier & Tarr, 1997).[19]

[19] The generalizability of such face-like effects to other properties naturally leads to doubts about whether the allegedly face-specific processing and face-specific neural areas are as domain-dedicated as many have held.

Taken together, these results suggest that vision can become specially attuned to—can come to perform specialized integrations for—a range of form properties in a variety of ways, some perhaps innately specified, some perhaps not. However wide this range turns out to be, it does seem that vision is, in the ways we have elaborated, more attuned to some spatial form properties than others. Thus, even if it is true (and it may not be) that the ideas enlisted in the visual perception of form all share a common spatial foundation, they may nonetheless differ in the way they are integrated from their foundations. And if there can be such differences for ideas of spatial form *within* a modality, there is all the more reason for thinking that crossmodal transfer may fail. For, again, modalities might build their ideas of a spatial form from distinct foundations, and even if they share foundations, they may carry out their compositional steps in different ways, so that whatever special advantages (of speed, accuracy, etc.) exhibited in a first modality may fail to manifest in a second.

VII. Another Clue to Modality-Specific Processing: Evidence from Illusion

Another strategy for uncovering modality-specific forms of integration that may impede crossmodal transfer involves reflecting on classic perceptual illusions. Perceptual illusions within a modality can be viewed as signatures of integrative processes at work within that modality because they reflect how these processes can go wrong in reconstructing the world. That vision responds to the Müller -Lyer configuration in its usual, illusory, way, tells us that the processes underlying our visual integration of two-dimensional form lead to a characteristic type of misestimate—one that may or may not be replicated by integrative processes for the same two-dimensional form in other modalities.[20]

Now, investigators have known for some time that the Müller-Lyer illusion, and other geometric illusions (e.g., the Poggendorff, the vertical-horizontal illusion) have counterparts in haptic perception (Bean 1938, Over 1966). Commentators have sometimes concluded that such transfer between modalities will occur quite generally (e.g., across many different pairs

[20] Highlighting failures of crossmodal transfer for this class of integrations is also useful in bringing out that special features of particular modalities need not be those that possess adaptive significance. This is salutary because it takes us beyond the domain of biologically significant features such as facehood and biological motion.

of natural and artefactual modalities, especially after extended perceptual learning).[21] We want to suggest that the evidence supports a more guarded assessment: sometimes transfer occurs, sometimes it does not, and even where it does there may be interesting differences to note.

To begin with a case of successful transfer, consider the susceptibility of haptic perception to the same errors of line length estimation present in visual perception of the Müller-Lyer configuration. It is striking that the effect not only occurs in both visual and haptic perception, but that its visual and haptic manifestations are analogous in a number of more specific respects. In both cases, errors are larger for more acute angles (Over 1966, Heller et. al. 2002) and decrease with exposure time/perceptual learning (Rudel and Teuber 1963). In both cases, explicit instructions to ignore the fins in favour of focusing on body-centered cues reduces error rates significantly (Millar and Al-Attar 2002). And there is a strong positive correlation between magnitudes of visual and haptic misestimates within individual subjects (Gentaz et. al. 2004). Finally, the haptic effect seems equally present in sighted, sighted but blindfolded, sighted but low-vision, and congenitally blind subjects (Heller et. al. 2002). This series of shared features is notable, and encourages the impression that analogous integrative mechanisms may be at work across modalities.

On the other hand, we can see that this situation is far from guaranteed by considering a pair of papers by Laurent Renier and colleagues on two classic cases—the Ponzo illusion (depicted in figure 3) and the vertical-horizontal illusion (depicted in figure 4)—in subjects using a prosthesis for substitution of vision with audition (PSVA) (Renier et. al. 2005; Renier, Bruyer, and DeVolder 2006).

[21] For example, such optimism about transfer emerges from the reports of Gerard Guarniero (1977), a congenitally blind philosophy Ph.D. student who served as a subject for Paul Bach y Rita's celebrated work with Tactile-Visual Sensory Substitution (TVSS), and who went on to write a dissertation on space perception (Guarniero 1977b), and also from the approving citations of these remarks by Bach y Rita (1984) and Noë and O'Regan (2002). Unfortunately, it is somewhat unclear what to make of these claims: both because Guarniero's reports are anecdotal, and because it is not obvious that the emerging abilities are perceptual rather than post-perceptual, it is not obvious that they show anything at all about intermodal transfer.

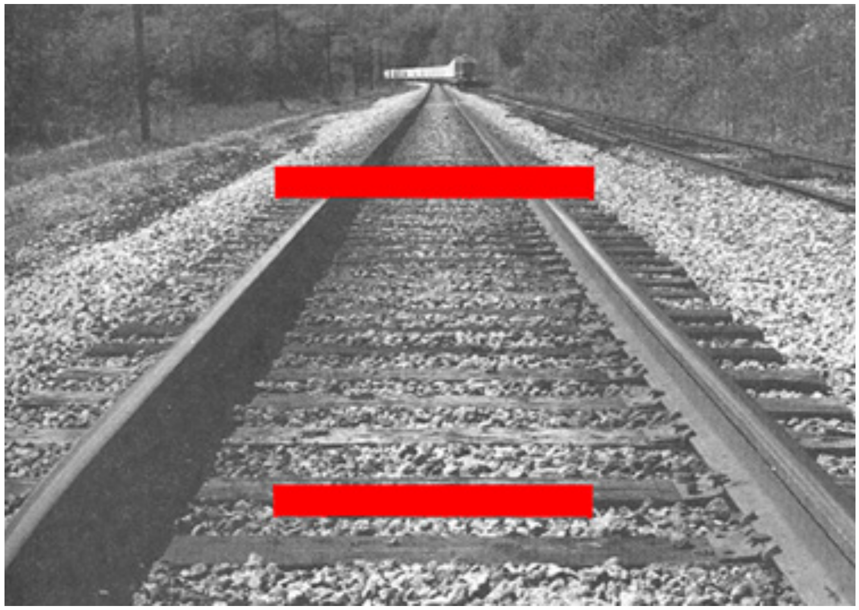

Figure 3 The Ponzo illusion: Parallel horizontal lines of equal length appear unequal when displayed over a pair of oblique lines that converge toward the top of the display.

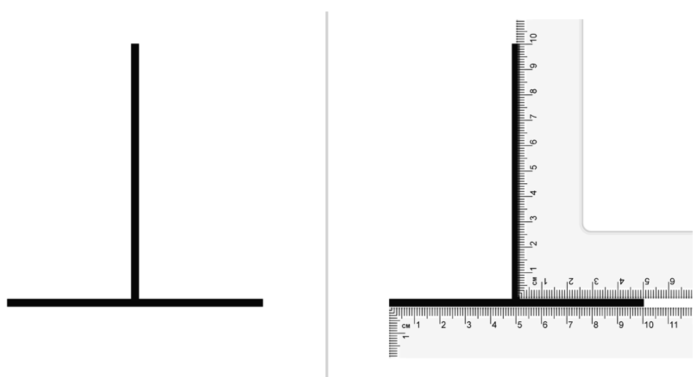

Figure 4 The Vertical-Horizontal Illusion: Subjects systematically over-estimate the length of vertical lines.

Renier et al were able to induce both the Ponzo and vertical-horizontal illusions (in some subjects) under PSVA, but only after a non-trivial period of training with the new modality. This would seem to show, by itself, that facility with the (characteristically illusionprone) integration of two-dimensional length information in vision does not transfer immediately to just any other modality capable of recovering two-dimensional form.

But putting things this way understates the obstacles to intermodal transfer uncovered by this work. In fact, Renier et. al. 2005 report that, initially, their subjects were not susceptible to the Ponzo illusion under PSVA at all, because they were able to perform the line length estimation task without attending to representations of the converging oblique lines crucial to the visual version of the illusion. Only by requiring subjects using PSVA to consider the two oblique lines of the stimuli before comparing the length of the two horizontal bars, could they make the illusion re-emerge. But even under this condition, the effect remained weak: they found the effect in sighted subjects (but still smaller in magnitude than among control sighted subjects not using PSVA), but not at all in their "early blind" subjects (subjects blind before their 20th month of age). And Renier, Bruyer, and DeVolder (2006) report a similar pattern for the vertical-horizontal illusion: the effect was strongest in sighted control subjects not using PSVA, somewhat weaker in blindfolded sighted subjects using PSVA, and completely absent in early blind subjects using PSVA.

These results illustrate different obstacles to crossmodal transfer for integration, even for a pair of modalities capable of recovering two-dimensional form properties. First, perceptual expertise for the relevant integration type may be present in one modality but not the other, with the result that transfer is impossible without (possibly quite elaborate) training. Second, because perceptual attention may be fixed in quite different ways between the two modalities, integrations that depend on which elements of a perceptual configuration are attended (or in what order) may succeed in one modality but not another. Third, and finally, the results with both the Ponzo and the vertical-horizontal configurations show that the possibility or robustness of a given integrative strategy may depend on the subject's perceptual history with a given modality; consequently, between-modality variation in this respect may block crossmodal transfer as well.

VIII. Perceptual Affect: Does It Carry Over?

Finally, it is worth noting yet another perceptual phenomenon that falls under ibn Tufail's rubric of learning and potentially under Molyneux's paradigm of crossmodal transfer as well. This is the link between perception and affect in what has been called evaluative perception. Introspection and a large and growing body of evidence reveal that our perceptual states have an affective dimension—certain smells are *disgusting*, certain tactile experiences *pleasant*, certain flavours *repulsive*, certain sights and sounds *beautiful*, certain pains *unbearable*, certain melodies *melancholy*, and so on. An aperçu of Ibn Tufail invites us to consider the transferability of such affect: he says that his newly-sighted man experiences "great joy" when he becomes visually acquainted with the colours—his earlier descriptive knowledge of them presumably offered no such delights. (Goodman 19xx, 7-8)

Despite significant controversy about the best theoretical account of the affective dimensions of perceptual experience, we can take it as a relatively uncontroversial starting point that these phenomena can be characterized in terms of three central psychological features: at a minimum, affective perceptual states are valenced; they have motivational force; and they confer reasons to the subjects in whom they occur.[22]

Evaluative perception raises all kinds of interesting questions, but two in particular are relevant to our present investigation. The first, which is closer to ibn Tufail's, is whether affect is a constituent part of certain perceptions or a distinct state that accompanies them. Is the perceptual experience of certain melodies itself sad or melancholy, or are these emotions mere accompaniments? Would an inexperienced listener necessarily feel the emotions of the "morning effect" when he listened to a sensitive rendering of Raga Vibhas?[23] Is startle part and parcel of hearing sudden loud sounds, or is it a distinct state that has somehow become associated with them? Would someone necessarily jump when she first heard a clap of thunder or a flash of lightning? Must a newly sighted person become erotically aroused at her first sight of a naked body?

[22] For useful overviews of philosophical work on this topic, see Aydede and Fulkerson (2015) and Bergqvist and Cowan (2018).

[23] Raga Vibhas, which is usually played before sunrise, is said to evoke a feeling of the morning in listeners. The reader can judge for herself at https://youtu.be/YAt6jW8QvJE

The second, closely related to Molyneux's, is this: Is a person who experiences something in one modality subject to the same affect as when she experiences it in another? Would a newly sighted person experience a shape or a texture as visually pleasant if she had previously found it haptically pleasant (holding relevant contextual dimensions fixed)? Could she experience it as visually pleasant though haptically without affect, or as visually pleasant and haptically unpleasant?

We see no a priori reason to believe that distinct modalities need (or for that matter need not) bear or connect with the same/analogous affective dimensions at all. Nor is it obvious, in the case that they do, that a history of experiences of a given affective type in one modality will (or won't) facilitate the acquisition or deployment of that affective type in a different modality. How soon before a newly sighted person has visual affective startle/erotic arousal/musical pleasure? Would having a history of repulsion in olfaction give the newly sighted person an advantage in acquiring affective repulsion in vision (relative to a newly sighted person without a history of olfactory repulsion)?

As in other cases considered above, there might not be a single answer across the board. We take it that the answers to these Molyneux Questions about affect are unobvious, and will depend on the local, particular, details of the psychological mechanisms implicated by the specific modalities at issue.

IX. Conclusion

As we have seen, Molyneux's question, ostensibly about the crossmodal application of sensory knowledge, is a special instance of the broader set of questions (prompted by ibn Tufail) about the fungibility of the senses—that is, about how knowledge and general representations/ideas that are normally formed in one way in one modality can be formed in another way in another modality. This broader question about fungibility leads us to ask about the respects in which such ideas do and do not converge, as well as the range of crossmodal transfers they underwrite. The variety evident in the cases we've examined shows that no general answers are forthcoming: there are instances of both crossmodal divergence and convergence; and even convergence, where it is found, is insufficient to guarantee crossmodal transfer. Given all this variety, it is likely that an adequate response to

Molyneux's question will require coming to grips with the broader set of issues raised by ibn Tufail.[24]

[24] We are grateful for comments from and discussion with Matthew Fulkerson, Ayoob Shahmoradi, and Brian Tracz, and for suggestions from the editors of this volume, all of which have much improved the paper.

REFERENCES

Allison, Truett, Aina Puce, and Gregory McCarthy (2000), "Social perception from visual cues: role of the STS region," *Trends in Cognitive Sciences*, 4, pp. 267-278.

Aydede, Murat and Matthew Fulkerson (2019), "Reasons and Theories of Sensory Affect," in *The Nature of Pain: Unpleasantness, Emotion, and Deviance*, ed. by David Bain, Michael Brady, and Jennifer Corns, Routledge, New York, pp. 27-59.

Bach y Rita, Paul (1984), "The relationship between motor processes and cognition in tactile vision substitution," in *Cognition and Motor Processes*, ed. by W. Prinz and A. F. Sanders, Springer, Berlin, pp. 150-160.

Barton, Jason J. (2003), "Disorders of face perception and recognition," *Neuro- logic Clinics*, 21, 2, pp. 521-548.

Bean, C. H. (1938). The blind have "optical illusions." *Journal of Experimental Psychology, 22*(3), 283–289. https://doi.org/10.1037/h0061244

Bennett, Jonathan 1965. Substance, Reality, and Primary Qualities, *American Philosophical Quarterly* 2/1: 1-17.

Bergqvist, Anna and Robert Cowan (eds.) (2018), *Evaluative Perception*, Oxford University Press, Oxford.

Casey, Sarah and Fiona Newell (2003), "Haptic own-face recognition," in *EuroHaptics 2003*, pp. 424-429.

— (2007), "Are representations of unfamiliar faces independent of encoding- modality?" *Neuropsychologia*, 45, 3, pp. 506-513.

Cattaneo, Zaira and Tomaso Vecchi (2011), *Blind vision: the neuroscience of visual impairment* Cambridge: MIT Press.

Cohen, Jonathan (2018), "Sensory Substitution and Perceptual Emergence," in *Sensory Substitution and Augmentation*, ed. by Fiona Macpherson, Proceedings of the British Academy, Oxford, pp. 205-235.

Diderot, Denis 1749 (1951). *Lettre sur les aveugles, à l'usage de ceux qui voient*, ed. Robert Niklaus, Genève: Librairie Droz.

Easton, Randolph D., Anthony Greene, and Kavitha Srinivas (1997a), "Do vision and haptics share common representations?: Implicit and explicit memory within and between modalities," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1, pp. 153-163.

Easton, Randolph D., Anthony Greene, and Kavitha Srinivas (1997b), "Transfer between vision and haptics: memory for 2-Dpatterns and 3-D objects," *Psychonomic Bulletin & Review*, 4, 3, pp. 403-410.

Evans, Gareth 1985. Molyneux's Question, in *Collected Papers*, ed. Gareth Evans, Oxford: Clarendon Press: 364-399.

Fine Ione, Wade Alex R., Brewer Alyssa A., May Michael G., Goodman Daniel F., Boynton Geoffrey M., Wandell Brian A., MacLeod Donald I. A. (2003) "Long-term deprivation affects visual perception and cortex," *Nature Neuroscience* 2003; 6:915-916.

Gallistel, Charles R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press

Gauthier, Isabel and Michael J. Tarr (1997), "Becoming a "Greeble" expert: Exploring mechanisms for face recognition," *Vision Research*, 37, 12, pp. 1673- 1682.

Gentaz, Edouard & Jacquet, Annie-Yvonne & Hatwell, Yvette & Camos, Valérie. (2004). "The Visual and the Haptic Müller-Lyer Illusions: Correlation Study." *Current Psychology Letters: Behaviour, Brain & Cognition* 13(2)

Glenney, Brian R. (2013). Philosophical problems, cluster concepts, and the many lives of Molyneux's question. *Biology and Philosophy* 28 (3):541-558.

Grossman, Emily, Michael Donnelly, Ronald Price, D. Pickens, V. Morgan, G. Neighbor, and Randolph Blake (2000), "Brain areas involved in perception of biological motion," *Journal of Cognitive Neuroscience*, 12, pp. 711-720.

Guarniero, Gerard (1977a), "Tactile Vision: a personal view," *Visual Impairment and Blindness*, 71, 3, pp. 125-130.

Guarniero, Gerard (1977b), *The senses and the perception of space*, PhD thesis, New York University.

Hadjikhani, Nouchine and Per E. Roland (1998), "Cross-Modal Transfer of Information between the Tactile and the Visual Representations in the Human Brain: A Positron Emission Tomographic Study," *The Journal of Neuroscience*, 18, 3, pp. 1072-1084.

Hardin, C. L. (1988) *Color for Philosophers: Unweaving the Rainbow* (Indianapolis: Hackett).

Held, Richard, Yuri Ostrovsky, Tapan K. Gandhi, Suma Ganesh, Umang Mathur, and Pawan Sinha 2011 The newly sighted fail to match seen with felt, *Nature Neuroscience* 14/5: 551-553.

Heller, Morton A., Brackett, Deneen D., Wilson, Kathy, Yoneyama, Keiko., Boyer, Amanda, & Steffen, Heather (2002). "The Haptic Müller-Lyer Illusion in Sighted and Blind People". *Perception*, 31(10), 1263–1274. https://doi.org/10.1068/p3340

Hopkins, Robert (2005). "Molyneux's Question," *Canadian Journal of Philosophy* 35 (3):441-464.

Johansson, Gunnar (1973), "Visual perception of biological motion and a model for its analysis," *Perception & Psychophysics*, 14, 2, pp. 201-211.

Jordan, Heather, Jason E Reiss, James E Hoffman, and Barbara Landau (2002), "Intact perception of biological motion in the face of profound spatial deficits: Williams syndrome." *Psychological Science*, 13, 2, pp. 162-7

Kanwisher, Nancy (2010), "Functional specificity in the human brain: A window into the functional architecture of the mind," *Proceedings of the National Academy of Sciences*, 107, pp. 11163-11170.

Kilgour, Andrea R. and Susan J. Lederman (2002), "Face recognition by hand," *Perception & Psychophysics*, 64, 3, pp. 339-352.

Kilgour, Andrea R. and Susan J. Lederman (2006), "A haptic face-inversion effect," *Perception*, 35, pp. 921-931.

Kim, Jejoong, Randolph Blake, Sohee Park, Yong-Wook Shin, Do-Hyung Kang, and Jun Soo Kwon (2008), "Selective impairment in visual perception of biological motion in obsessive-compulsive disorder," *Depression and Anxiety*, 25, 7, E15-25.

Lederman, Susan J., Roberta Klatzky, and Ryo Kitada (2010), "Haptic Face Processing and Its Relation to Vision," in *Multisensory Object Perception in the Primate Brain*, ed. by Marcus Johannes Naumer and Jochen Kaiser, pp. 273- 300, DOI: 10.1007/978-1-4419-5615-6_15.

Lewis, David 1966. Percepts and Colour Mosaics in Visual Experience, *Philosophical Review* 75/3: 357-368.

Liu, Jia, Alison Harris, and Nancy Kanwisher (2010), "Perception of face parts and face configurations: An fMRI study," *Journal of Cognitive Neuroscience*, 1, pp. 203-211.

Lu, Hongjing (2010), "Structural processing in biological motion perception." *Journal of Vision*, 10, 12, p. 13, http://www.biomedsearch.com/nih/Structural-processing-in-biological-motion/21047745.html.

Matthen, Mohan (forthcoming) "The Dual Structure of Touch," in F. de Vignemont, H. Y. Wong, A. Farné and A. Serino (eds) *The World at Our Fingertips* (Oxford: Oxford University Press).

Matthen, Mohan and Jonathan Cohen (forthcoming). *Australiasian Journal of Philosophy*: forthcoming. https://doi.org/10.1080/00048402.2019.1603246.

McCarthy, Gregory, Aina Puce, John C. Gore, and Truett Allison (1997), "Face- Specific Processing in the Human Fusiform Gyrus," *Journal of Cognitive Neuroscience*, 9, 5, pp. 605-610.

Millar, Susanna and Zainab Al-Attar (2002), "The Müller-Lyer illusion in touch and vision: Implications for multisensory processes"

Mirenzi, Aaron and Eric Hiris (2011), "The Thatcher effect in biological motion," *Perception*, 40, 10, pp. 1257-1260.

Noë, Alva and J. Kevin O'Regan (2002), "On the brain-basis of visual consciousness: A sensorimotor account," in *Vision and Mind: Selected Readings in the Philosophy of Perception*, ed. by Alva Noë and Evan Thompson, MIT Press, Cambridge, Massachusetts, pp. 567-598.

Norman, J. Farley, Hideko F. Norman, Anna Marie Clayton, Joann Lianekhammy, and Gina Zielke (2004), "The visual and haptic perception of natural object shape," *Perception & Psychophysics*, 66, 2, pp. 342-351.

Ostrovsky, Y., A. Andalman, and P. Sinha 2006. Vision After Extended Congenital Blindness, *Psychological Science* 17/12: 1009-1014.

Ostrovsky, Y., E. Meyers, S. Ganesh, U. Mathur, and P. Sinha 2009. Visual Parsing After Recovery from Blindness, *Psychological Science* 20/12: 1484-1491.

Over, Ray. "A Comparison of Haptic and Visual Judgments of Some Illusions." *The American Journal of Psychology* 79, no. 4 (1966): 590-95. doi:10.2307/1421295.

Purcell, Dean G. and Aan L. Stewart (1988), "The face-detection effect: Configuration enhances detection," *Perception & Psychophysics*, 43, pp. 355- 366.

Reales, José Manuel and Soledad Ballesteros (1999), "Implicit and explicit memory for visual and haptic objects: Cross-modal priming depends on structural descriptions," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 3, pp. 644-663.

Renier, Laurent, Raymond Bruyer, and Anne G. DeVolder (2006), "Vertical- horizontal illusion present for sighted but not early blind humans using auditory substitution for vision," *Perception & Psychophysics*, 68, 4, pp. 535- 542.

Renier, Laurent, C. Laloyaux, O. Collignon, D. Tranduy, A. Vanlierde, Raymond Bruyer, and Anne G. De Volder (2005), "The Ponzo illusion using auditory substitution of vision in sighted and early blind subjects." *Perception*, 34, pp. 857-867.

Schwenkler, John (2019). "Molyneux's Question Within and Across the Senses," in T. Cheng, O. Deroy, and C. Spence (ed.) *Spatial Senses: Philosophy of Perception in an Age of Science* London: Routledge.

Sinha, P., B. Balas, Y. Ostrovsky, and R. Russell (2006), "Face recognition by humans: Nineteen results all computer vision researchers should know about," *Proceedings of the IEEE*, 94, pp. 1948-1962.

Sugita, Yoichi (2009), "Innate face processing," *Current Opinion in Neurobiology*, 19, pp. 39-44.

Thompson, Peter (1980), "Margaret Thatcher: A new illusion," *Perception*, 9, 4, pp. 483-484.

Troje, Nikolaus F. and Cord Westhoff (2006), "The Inversion Effect in Biological Motion Perception: Evidence for a "Life Detector"?" *Current Biology*, 16, pp. 821-824.

Twedt, Elyssa, David Sheinberg, and Isabel Gauthier (2007), "Comparing Thompson's Thatcher effect with faces and non-face objects," *Journal of Vision*, 7, 9, p. 508.

Wong, Yetta K., Elyssa Twedt, David Sheinberg, and Isabel Gauthier (2010), "Does Thompson's Thatcher Effect reflect a face-specific mechanism?" *Perception*, 39, pp. 1125-1141.

Yin, Robert K. (1969), "Looking at upside-down faces," *Journal of Experimental Psychology*, 81, pp. 141-145.