Molyneux's Question Within and Across the Senses

John Schwenkler
Florida State University

§1. Introduction

Molyneux's question (see Locke 1694/1975: II, ix, 8) asks us to imagine a man born blind who has learned to identify certain three-dimensional shapes by touch. We are to suppose that this man is now 'made to see', and that the very shapes he can identify through touch are placed before his eyes. The question is: Would the man be able immediately to identify the shapes that he then saw?

On a first pass it may seem that contemporary philosophers' answers to this question should break down neatly according to their views of the metaphysics of perceptual experience, as follows. If perception is a presentation or representation of external objects and their properties, then since shape is a property that can be perceived through vision and touch alike, there should be no need for an associative process before a person can identify shapes that are presented or represented through sight as the same as those that are presented or represented through touch. By contrast, if perceptual experience consists in qualia or internal sensations, then since the qualitative character of visual experience is so different from that of the experience of the world through touch, we should not expect the identity of what is seen with what is touched to be transparent to a naïve perceiver.

But this would be too quick. Where *A* and *B* are both representations of the same object or property *X*, the identity *A=B* may not be transparent to a rational subject in possession of both representations as long as *A* and *B* present *X* under different modes of presentation.[1] Thus, for example, if two objects of the same color are shown under different lighting conditions, it may be no trivial thing to recognize that their colors are the same, and the ability to recognize this might depend on prior experience in the way that lighting affects the appearance of color. If there is such a difference between the modes of presentation of shapes in vision and in touch, the answer to Molyneux's question could be negative even according to an 'externalist' account of perceptual content.

Notice that in order to get this concern off the ground we need something more than the observation that there are very *general* differences between the qualitative or phenomenal character of visual experience and that of the experience of the world through touch—

---

[1] For a helpful discussion of perceptual modes of presentation, see Rescorla (forthcoming).

such that a person who is 'made to see' for the first time may say that he didn't know before what *that* was like, and such that there will be no question for a person familiar with what it is like to sense through a given modality whether a certain experience is an experience in *that* modality rather than another. This is because these general differences could have a lot to do with there being ranges of properties—like color on the one hand, and temperature on the other—that are perceived through sight but not through touch, and vice versa, whereas Molyneux's question concerns a property that is be perceivable by sight and touch alike. The idea that motivates a negative answer to Molyneux's question is that even if we abstract away from those general differences and consider only a 'common sensible' like shape or size, there is enough difference in how *a particular property* of this sort 'appears' or is presented to the perceiving subject between vision and touch that it is open to question whether being able to recognize the property in one of these sensory manifestations would guarantee the ability to recognize it in the other one.

In the present case, the most obvious reason to take this possibility seriously has to do with the great difference between how information about object shape is obtained through sight as opposed to touch. Whereas vision depends on light reflected from the facing side of a visible object, in touch the source of sensory information is the stimulation of receptors which may be in contact with several different sides of an object at once—as when a person places an object in her mouth or encompasses it with one or both of her hands. In addition, motion seems to play a more essential role in touch perception than in vision, as there is no counterpart in touch to the way we can identify the shape of a visual stimulus simply 'at a glance' and without needing to move in respect to it.[2] Even if we accept that many of the same properties are made manifest (presented, represented) as we perceive the world through these two sensory modalities, the differences between the modalities in *how* we perceive these properties through them seem to lead to differences in how they manifest the properties to us.

However, this last phenomenon—that of a given property's being manifested differently across distinct sensory perceptions of it—does not arise only when those perceptions involve different sensory modalities. For example, imagine feeling the shape of a small object with your hand and then putting that object in your mouth to perceive its shape by moving your tongue around it—or, for that matter, feeling the shape of an object with one hand and then feeling it again with the other. There is a difference between how a shape feels to your hand and fingertips, compared to how it feels to your tongue and the roof of your mouth. Even more obviously, an irregular shape will *look different* when it is seen from one viewing angle compared how it looks when seen from another one. In each case the possibility may arise of perceiving the same shape accurately on separate occasions without being able to tell if it is the same shape that one perceives, even though

---

[2] On the essential role of motion in tactile perception, see Fulkerson 2014.

neither case involves any comparison between perceptions in two different sensory modalities.

The present chapter will explore how our understanding of Molyneux's question, and of the possibility of an experimental resolution to it, should be affected by recognizing the complexity that is involved in reidentifying shapes and other spatial properties across differing sensory manifestations of them. I will argue that while philosophers today usually treat the question as concerning 'the relations between perceptions of shape in different sensory modalities' (Campbell 1995, 301), in fact this is only part of the question's real interest, and that the answer to the question also turns on how shape is perceived within each of sight and touch individually.

§2. The Looks of a Shape

My argument is that the answer to Molyneux's question turns on at least two issues, namely (i) the relationship *between* visual and haptic perceptions of shape and (ii) the way that these properties are perceived *within* each sensory modality. Call (i) the *intermodal* question, and (ii) the *intramodal* one. The relevance of the intermodal question is easiest to see. Suppose for example that the visual system represents object shape in a proprietary 'code' which is separate from the code that is used for representing felt shape, and that cross-modal shape comparisons depend on relating the perceptual representations that are in these two distinct systems.[3]

In this case the answer to Molyneux's question is likely to be negative, except in the unlikely event that the proprietary codes of sight and touch happen to be the same, or that the translation of these codes to one another or to the postulated common code is achieved by an innate or otherwise 'hard-wired' mechanism.

If, however, the code in which sight and touch carry information about shape is common between the two modalities, then should we expect an affirmative answer instead? Not necessarily—for either of sight and touch might fail to represent shape in the way that would be required for a cross-modal comparison. One way this might happen is if, as some of the early empiricists supposed, vision represents only 2D object shape while touch represents shape in three dimensions.[4] In this case the range of spatial *properties* that are represented in each of the two senses will be too dissimilar to affect comparisons

---

[3] For earlier discussion of this 'common-coding' proposal and its relation to Molyneux's question, see Altieri 2015, Schwenkler 2015.

[4] Smith (2000) argues convincingly that this is not an idea we should take seriously any longer.

between them—even if the information about these properties is carried in a common language.
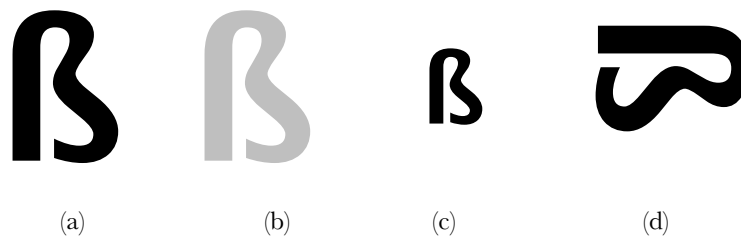

(a)   (b)   (c)   (d)

Figure 1. Four appearances of a shape.

To bring another way this second sort of thing could happen, consider the four tokens of the shape 'ß' that are shown in Figure 1. Do they have the same shape? You will answer yes if you can see that the left-hand shape is no different from the right-hand one except in how they are rotated. But that is no trivial achievement. In order to recognize these shapes as the same you need to represent the shape of each in a way that is indifferent to the shape's relation to your eyes and head when you look at the figure. And I think we can imagine a perceiver that could only recognize sameness and difference of shape in a way that was not so indifferent to egocentric orientation. Such a creature could recognize the sameness of the shapes in (a), (b), and (c) but would *not* be able to recognize that the shape in (d) was the same as these. It could identify a pair of identical shapes as the same only if they had the same orientation relative to its eyes and head.

One might object that the creature just imagined should not be described as representing the sameness and difference of *shape* at all, but only sameness and difference in a relevant dimension of sensory stimulation. And there is something to this worry, as shape is an *abstract* property in the sense that objects can share the same shape despite differences in size, color, location, orientation, and so on. This means that a perceiver cannot count as being able to recognize a given *shape* across different perceptions of it, as opposed to being able merely to respond selectively to a certain pattern of sensory stimulation, unless she can abstract away from some of these differences and tell that the shape is the same despite them.

Still, we should not require as a condition of being able recognize a certain shape by sight that one be able to recognize this shape in *every* visible manifestation of it. And the creature I asked us to imagine *has* learned to apply the same concept to the shape 'ß' in several quite different modes of visual presentation, insofar as the shapes in (a) and (c) are different in color from (b), while (c) differs from (a) and (b) in its size, and each differs from both of the others in its location on the page. What these stimuli have in common is nothing but their shape. If our creature can recognize this commonality, she will thereby have recognized that the same *shape* appears in each.

Can we, then, imagine a perceiver who could do this much but could not extend this ability to the appearance of the shape in (d) as well? It seems to me that we can indeed conceive this—perhaps not 'from the inside' by trying to imagine *what it would be like* to experience the world in this way, but insofar as we appreciate that the capacity to reidentify a shape across differences in its orientation requires something other than the capacity to reidentify it across differences in color, size, and location. That is, the representation of visible shape according to an object-centered reference frame that abstracts away from differences in egocentric orientation, such that the shape in Figure 1(d) can be recognized as the same as that in Figure 1(a), requires something additional to the viewer-relative form of visual spatial representation that would suffice to identify the shape in (a) as the same as those in (b) and (c). A perceiver who could do only the latter thing could reidentify an irregular shape across differences in color, size, and location, but could not do the same across differences in its spatial orientation of the shape relative to her eyes. For her, differences in the viewer-relative orientation of an irregular shape would make it appear to be a different shape altogether.

§3. Shape Across the Senses

Return to the question that I raised at the start of the previous section. What has to be true of sight and touch individually, i.e. beyond whatever sort of connection must exist between them, in order for a shape to be recognized as the same across these senses? The cases that we've just considered help to bring this question into focus insofar as it's quite implausible that the explanation of why someone might fail to recognize that the shape in Figure 1(d) is the same as that in Figure 1(a) would have anything to do with the lack of a suitable *connection* between these two representations. (At least not in the way that we imagined the case. It would be different if, say, the failure were due to an inability to communicate visual information between the two hemispheres.) The idea was rather that while these representations shared a common code, nevertheless the *information* contained in the two representations wasn't sufficient to identify the represented shapes as the same. And it should not be hard to see how something like this could lead to difficulty in comparing seen shapes to felt ones.

Insufficient attention to this matter has led to difficulties in some recent attempts to answer Molyneux's question experimentally. Consider for example a study carried out by Held et al., which for a while was thought to have resolved the question once and for all.[5]

[5] For example, an article in *The New York Times* reported that the Held study 'appears to show definitively that Locke was right' in answering Molyneux's question negatively (Nicholas Bakalar, 'Science of Vision Tackles a Philosophy Riddle', *The New York Times*, April 25, 2011). The criticism of the study that I offer below builds on the argument I made earlier in Schwenkler (2012; 2013).

Having performed cataract surgery that gave sight to five individuals who had been congenitally blind, Held and colleagues had them perform a shape matching task[6] in three conditions: 'vision-to-vision' (VV), in which all the stimuli were displayed visually; 'touch-to-touch' (TT), in which all were placed in the hand but hidden from view; and 'touch-to-vision' (TV), in which the first stimulus object was placed in the hands and then the others were presented to the eyes. Strikingly, while participants performed very well in the VV and TT conditions, in the TV condition they were barely above chance—just 58% of matches were identified in the cross-modal comparison, compared to 98% in touch-to-touch and 92% in vision-to-touch matching, respectively. The authors took this to show that 'the answer to Molyneux's question is likely negative', as '[t]he newly sighted subjects did not exhibit an immediate transfer of their tactile shape knowledge to the visual domain … Whatever linkage between vision and touch may pre-exist concomitant exposure of both senses, it is insufficient for reconciling the identity of the separate sensory representations' (Held et al. 2011, 552).

As we have seen, in order to accept this conclusion we would have to rule out the possibility that something other than a suitable *linkage between* visual and tactile representations was responsible for the participants' failure in the cross-modal task. Could the failure be due instead to inadequacy in the representation of object shape *within* one of these modalities? The VV and TT conditions were supposed to account for this: the researchers found that the participants *could* match seen shape to seen shape and felt shape to felt, and thus they suppose that the poor performance in the TV condition must have been due to the lack of a suitable cross-modal connection.

This is too quick, however. In Held et al.'s VV condition, the participants sat still in a chair while objects were displayed from a single viewpoint. And for reasons we have already explored, in such a condition it is possible to classify objects according to their shapes *without* representing those shapes according to a perspective-invariant, object-centered frame of reference. Yet while a merely eye- or head-centered visual representation of object shape would have been enough for performance in the VV task, such a representation would not have the sort of content necessary to compare it reliably with a representation of shape obtained through active touch.

If this is the correct explanation of the results obtained in the TV condition, then we should expect that Held et al.'s participants would also have had difficulty in matching seen shapes with seen ones across significant changes in viewing angle. While this question was not tested directly, there is *prima facie* reason to think that the newly sighted individuals would indeed have been unable to do this. For example, Ostrovsky and

---

[6] The task involved presenting a first stimulus object, followed by a pair of two more, one of which matched the original target. The instruction was to identify the match.

colleagues showed images of three-dimensional shapes to three patients whose sight had been surgically restored between two weeks and three months earlier. Presented with these stimuli, 'the recently treated subjects reported perceiving multiple objects, one corresponding to each facet. They were unable to integrate the facets into the percept of a single three-dimensional objects' (Ostrovsky et al. 2009, 1486). Similar results were observed using color photographs of common objects, as the participants 'pointed to regions of different hues and luminances as distinct objects', revealing that their visual system had partitioned the stimuli into 'meaningless regions, which would be *unstable across different views* and uninformative regarding object identity' (ibid., 1487; emphasis added). Such representations might have been sufficient, however, for participants to recognize sameness or difference in the shape of a stimulus object as long as each object was displayed consistently from the same viewpoint. They just were not robust enough to permit comparisons across changes in perspective.

§4. Simple and Complex

A natural conclusion to draw from this criticism of Held et al.'s study is that the stimuli they used to test their participants' capacity for cross-modal matching were simply too *complex* for their visual systems to handle. To get around this limitation, the participants might have been presented instead with stimulus objects whose shapes could be distinguished without having to integrate a number of distinct facets into the percept of a single object. Indeed, Molyneux's original statement of the question points us this way: he suggested using a *sphere* and a *cube*, rather than the more complex Lego shapes that Held et al. employed, as a way of testing the newly sighted man's ability to match seen shape with felt.

But there is a danger of pushing things too far in this direction, namely that of turning to stimuli that are *so* simple then they can be distinguished reliably without any need to represent their shapes as such. As before, we can motivate this difficulty just by considering an intramodal task: for example, a person might be able to distinguish a seen square from a seen circle, and match it with another square that she sees, just by being sensitive to the degree of *discontinuity* in the respective stimuli—a sensitivity that could ground the requisite judgments of sameness and difference in the absence of any genuine perception of shape. In order for performance on this sort of task to provide a meaningful measure of shape perception, it needs to involve pairs of stimuli that share in common a number of low-level spatial features, which must in turn be integrated into a single complex percept. And this is just the thing that the newly-sighted subjects were found to have a particular difficulty in doing.

As another way to appreciate this difficulty, consider the possibility that we might avoid the difficulties with Held et al.'s study by using 'two-dimensional' stimuli instead of three-dimensional ones. ('Two dimensional' is in scare-quotes because the stimuli themselves will necessarily be three-dimensional objects.) Once again, there is historical precedent here, as such a variant on the original Molyneux question was imagined by Denis Diderot in the 18th century (see Diderot 1749/1977), and more recently by Gareth Evans in his (1985). Could we, then, just replace Molyneux's cube and sphere with a square and a circle, or the Lego shapes that Held et al. used with shapes like the '**ß**' of Figure 1, either drawn on a screen and presented to the eyes or constructed as raised-line drawings which then can be felt by hand?

The foremost problem for this proposal is simply that it is *hard* for human perceivers to identify 'flat' shapes by touch—a fact that may be surprising if we take vision as our default model of what perception is like, but should not be if we recognize the essential role of exploratory movement in haptic perception, and how rarely anything like the shape of a raised-line drawing is a relevant object of haptic perception 'in the wild'. [7] Usually, when we need to use touch to perceive the shape of an object (e.g., in the dark), we do this by running our hands or fingers around its contours—a form of dynamic exploratory motion that Lederman and Klatzky (1987) term *contour following*, and which is especially well suited to the perception of 3D shape. By contrast, at least for human perceivers contour following does not work well at all as a means of perceiving the shape in a two-dimensional display: indeed, as Lederman and Klatzky observe (1987, 343), 'people can explore the contours of a two-dimensionally depicted object for as long as several minutes without being able to identify it'.

Kevin Connolly, who favors a version of the proposal in question, has suggested that this difficulty could be overcome just by making the stimuli very simple: 'since raised-line drawings typically are testing for proficient tactile identification in general, … they are often testing at a level of difficulty unnecessary for a two-dimensional test of Molyneux's question, where what is required is just proficient tactile identification for a few simple shapes' (Connolly 2013, 509). Yet we have already seen that if the shapes used as stimuli are *too* simple, then evidence that participants can discriminate between them, even to the point of making successful cross-modal comparisons, will not be evidence that they have identified the *shapes* of these stimuli at all—for they could instead be responding merely to low-level features like the overall complexity or degree of discontinuity in the stimulus objects, just as the newly sighted individuals in Held and colleagues' experiment were arguably doing in the VV task.

---

[7] For further discussion of the difficulty in identifying raised-line drawings by touch, including reference to relevant experimental literature, see Cheng 2015.

In order to make progress in explaining how Molyneux's question could be addressed experimentally in a way that avoids the difficulties described above, philosophers need to propose experimental designs in which the stimuli involved are both complex enough that they can be compared successfully only through *bona fide* perception of their shapes, yet simple enough that their shapes can be perceived, both visually and through touch, by a person newly given sight. Only when we are confident that both of these conditions have been satisfied in a given experiment can we trust that its results will be genuinely informative.

§5. Conclusion

If a child is born with congenital cataracts and her sight is surgically restored, has she thereby been 'made to see'? Unsurprisingly, it depends on what this last phrase is supposed to mean. Experimental research of the sort I have surveyed makes it clear that *some* visual function is present immediately after restorative surgery, but a great deal more depends on subsequent perceptual learning. And it is an inconvenience to philosophers that in this visual learning, touch is also engaged—so that in *learning to see* a person will also have the opportunity to associate the visual appearances of things with the ways that those things feel. This inescapable fact, together with the evidence we have about the limits of newly restored sight, makes it seem unlikely that any particular experiment could answer Molyneux's question in a way that philosophers would find satisfactory.

If that is right, then the best way forward on the question is through careful phenomenological reflection that is appropriately informed by relevant perceptual science. This requires getting clear on the differences between visual and haptic shape perception that give Molyneux's question its philosophical significance—differences that have to do not with sensory quality or phenomenal character, but rather with the manner in which shape is presented when it is made manifest through vision as opposed to touch. There do seem to be some such differences. But it is doubtful whether there is so much to them that a visual representation of object shape that was robust enough to allow reidentification of that shape across differences in size, location, color, and viewpoint could not also be compared reliably with a representation of shape obtained through active touch.

REFERENCES

Altieri, N. 2015. 'Multimodal theories of recognition and their relation to Molyneux's question'. *Frontiers in Psychology* 5, 1547
Campbell, J. 1995. 'Molyneux's question'. *Philosophical Issues* 7, 301-318

Cheng, T. 2015. 'Obstacles to testing Molyneux's question empirically'. *i-Perception* 6, 1-5

Connolly, K. 2013. 'How to test Molyneux's question empirically'. *i-Perception* 4, 508-510

Diderot, D. 1749/1977. *Letter on the Blind*. Repr. in M.J. Morgan, *Molyneux's Question*. Cambridge: Cambridge University Press

Evans, G. 1985. 'Molyneux's question'. In Evans, *Collected Papers*. Oxford: Oxford University Press

Fulkerson, M. 2014. *The First Sense*. Cambridge: The MIT Press

Held, R., Ostrovsky, Y., de Gelder, B., Gandhi , T., Ganesh S., Mathur, U. and P. Sinha. 2011. 'The newly sighted fail to match seen shape with felt'. *Nature Neuroscience* 14, 551-553

Lederman, S.J., and R.L. Klatzky. 1987. 'Hand movements: a window into haptic object recognition'. *Cognitive Psychology* 19, 342-368

Locke, J. 1694/1975. *An Essay Concerning Human Understanding*, 2nd ed. New York: Oxford University Press

Ostrovsky, Y., Meyers, E., Ganesh, S., Mathur, U. and P. Sinha. 2009. 'Visual parsing after recovery from blindness'. *Psychological Science* 20, 1484-91

Rescorla, M. Forthcoming. 'Perceptual co-reference'. *Review of Philosophy and Psychology*

Schwenkler, J. 2012. 'On the matching of seen and felt shapes by newly sighted subjects'. *i-Perception* 3, 186-188

Schwenkler, J. 2013. 'Do things look the way they feel?' *Analysis* 73, 86-96

Schwenkler, J. 2015. 'Commentary: "Multimodal theories of recognition and their relation to Molyneux's Question"'. *Frontiers in Psychology* 6, 1792

Smith, A.D. 2000. 'Space and Sight'. *Mind* 109, 481-518