

Berkeley and the spatiality of vision (to appear in the *Journal of the History of Philosophy*)

Rick Grush
University of California, San Diego

1. Introduction

Berkeley's *Essay Towards a New Theory of Vision*¹ presents a theory of various aspects of the spatial content of visual experience that attempts to undercut not only the optico-geometric accounts of e.g., Descartes² and Malebranche³, but also elements of the empiricist account of Locke⁴. My tasks in this paper is to shed light on some features of Berkeley's account that have not been adequately appreciated.

This paper is organized as follows. Section 2 will discuss Locke's account of the spatiality of vision in Book 2 of the *Essay*. While the optico-geometric approach of, e.g. Descartes and Malebranche, credits subjects (or their visual systems) with *a priori* geometrical knowledge by way of which the spatial features of their environments are deduced from, *inter alia*, the nature of the immediate visual input, the distance between the eyes, and the eyes' vergence angle, Locke's empiricism motivates an approach according to which spatial features of

¹ George Berkeley, *The Works of George Berkeley, Bishop of Cloyne*, eds. A.A. Luce and T.E. Jessop, (London: Thomas Nelson and Sons, Ltd.) All references to Berkeley's *Essay Towards a New Theory of Vision* are from Volume One of the Luce and Jessop edition, and will be cited in the text as *NTV* followed by the section number and, where appropriate, page number; all references to Berkeley's *A Treatise Concerning the Principles of Human Knowledge* are from Volume Two of the Luce and Jessop edition, and will be cited in the text as *PHK* followed by the section number and, where appropriate, page number.

² Descartes' views on vision are expressed in his *Optics*, one of the essays published along with the *Discourse on the Method*. All references to this work are from Rene Descartes, *The Philosophical Writings of Descartes*, trans. Cottingham, Stoothoff and Murdoch (Cambridge: Cambridge University Press, 1985), and will be cited as *Optics*, followed, where appropriate, by the page number.

³ Nicolas Malebranche, *The Search after Truth and Elucidations of the Search After Truth*, trans. Thomas Lennon and Paul Olscamp (Columbus, OH: Ohio State University Press, 1980), hereafter cited as *ST*, followed by book, chapter, and, where appropriate, page number. Malebranche's views on vision are discussed in *ST* 1, Chapters 6-9.

⁴ John Locke, *An Essay Concerning Human Understanding*, Peter H. Nidditch, ed. (Oxford, Clarendon Press, 1975), hereafter cited in the text by *Essay* followed by book, chapter and section number, and, where appropriate, page number of the Nidditch edition.

visual perception are either directly given in perception — *viz.* spatial information relating to features of the visual environment in the breadth and height dimensions —, or are learned through experience — as when features such as shading gradients available on the two-dimensional 'sense datum screen' come to be associated with the dimension of depth. Using the Molyneux question as a fulcrum, I will introduce some descriptive apparatus that will help to explain Locke's position, and that will be of use in examining Berkeley's.

In Section 3 I turn to Berkeley's *NTV*. After a brief preliminary discussion in Sections 3.2 of the initial sections of *NTV* that deal with the issue of depth in a way that is essentially a more detailed version of Locke's account, Section 3.3 discerns two conflated but distinguishable considerations that Berkeley provides to the effect that depth is not a proper object of vision. I then turn in Sections 3.4 to 3.6 to what is the central issue of this paper, Berkeley's discussion of the spatial axes of breadth and height. This is where the problems arise, for on the one hand, Berkeley is motivated to deny Locke's assumption that we are through vision immediately aware of a two-dimensional sense datum plane, for the spatiality of this plane, as a common sensible available to both vision and touch and hence not the exclusive province of either, would be an abstract idea.⁵ But on the other hand he frequently uses language that suggests he is crediting vision with just such planar content. The first major strand of the critical discussion of Berkeley in Section 3.4 will be a critique of his negative account to the effect that planar content is not directly given through the modality of vision, where it will be argued that Berkeley's argument fails because of an unnoticed ambiguity — the same ambiguity that was shown in Section 3.3 to be present, but relatively harmlessly so, in his discussion of depth. The second major strand, spanning sections 3.5 and 3.6, concerns Berkeley's positive account of the *apparent* planar

⁵ An abstract idea in a bad sense. In the 'Introduction' to the *PHK* (pages 25-40), Berkeley's main sustained discussion of abstract ideas, several ways in which an idea might be thought to be abstract are discerned, only one of which is, according to Berkeley, coherent. For rich and nuanced discussion of Berkeley on abstract ideas, see George S. Pappas, "Abstract ideas and the *New Theory of Vision*," *British Journal for the History of Philosophy* 10 (2002): 55-70.

content of visual experience. It is unclear that his positive account of vision can be formulated in such a way that is both adequate as an account of actual human vision and does not make surreptitious appeal to precisely the planar content the dismissal of which is its goal — and this is true even on the sympathetic reconstructions that have been offered recently by Atherton⁶ and Schwartz⁷.

In a brief final Section 4 I make explicit what alterations would have to be made to Berkeley's position in order to render it viable, and underscore the respect in which it has been vindicated by recent work in perception.

2. Locke

A lively facet of early modern philosophical and scientific theorizing concerned the mechanisms of visual perception, and one common tack taken was to credit the human perceptual system with a sort of innate geometry that would, for example, allow one to deduce the distance of a seen object on the bases of its projections on the two retinae and the distance between the eyes and their vergence angle.⁸ Descartes' philosophical views had no problem

⁶ Margaret Atherton, *Berkeley's Revolution in Vision* (Ithaca: Cornell University Press, 1990).

⁷ Robert Schwartz, *Vision: Variations on some Berkeleian themes* (Cambridge, MA: Blackwell, 1994).

⁸ Descartes' account, as expressed in the *Optics* is this:

... when our two eyes A and B are turned towards point X, the length of the line AB and the size of the two angles XAB and XBA enable us to know where the point X is. (*Optics*, page 170)

The parallel account by Malebranche in *The Search After Truth* reads:

The first, most universal, and sometimes the surest means we have of judging the distance of objects at a short distance is the angle made by the rays of our eyes with the object as its apex, that is, where the object is the point where these rays meet. When this angle is very great, we see the object as very near; and when, on the other hand, it is very small, we see it as very remote. And the change that occurs in the state of our eyes according to the changes in this angle is the means the soul employs in order to judge the remoteness or proximity of objects. For just as a blind man could touch a given body with the ends of two straight sticks of unknown length and judge its approximate distance according to a kind of natural geometry by the position of his hands and the distance between them, so might the soul be said to judge

accommodating such an innate geometry by which we make inferences about distance, or the abstract ideas of extension that figure in the premises and conclusions of such inferences. For the British empiricists, however, things were not so easy. While Berkeley's views represent the most radical break from these optico-geometrical accounts, it will be useful to say a word about Locke's account as expressed in Book 2 of his *Essay*, and especially the discussion leading up to and including the Molyneux question.

First, to the proper object(s) of sight, which for Locke is something like a two-dimensional tessellation of color patches.⁹ As Locke puts it:

When we set before our Eyes a round Globe of any uniform colour, v.g. Gold, Alabaster, or Jet, 'tis certain that the *Idea* thereby imprinted in our Mind, is of a flat Circle variously shadow'd, with several degrees of Light and Brightness coming to our Eyes. (*Essay* Book 2, Chapter 9, Section 8, p.145, emphasis original)

It will be noted that what is *not* on this list is *depth* or *distance*. Depth, as one of the three dimensions of extension, is an aspect of the content of the proper objects of touch for Locke, but not of vision. Locke thus recognizes a stage of psychological information processing involving flat images in the mind's eye corresponding to the flat retinal images in the body's eye. But for Locke, judgments about the three-dimensional spatial characteristics of our surroundings made on the strength of visual input are not the result of geometrical reasoning that takes these images,

the distance of an object by the disposition of its eyes, which varies with the angle by which it sees the object, that is, with the distance of the object. (*ST*, Book 1, Chapter 9, page 41)

Both Descartes and Malebranche mention additional means by which a seen object's distance is determined, including the feeling of eye muscle strain involved in changing the shape of the eye to clearly see objects that are very close (*Optics*, p.170; *ST*, Book 1, Chapter 9, p. 42-43).

⁹ In an article that exhibits great resourcefulness and textual knowledge, Laura Berchielli has, quite astonishingly, argued that it is not true that for Locke the proper objects of vision are exclusively two-dimensional, but at least sometimes have genuinely three-dimensional content (Laura Berchielli, "Color, Space, and Figure in Locke: An Interpretation of the Molyneux Problem," *Journal of the History of Philosophy* 40 (2002): 47-65). At the other end of the spectrum Martha Bolton articulates a non-spatial reading of Locke according to which even two-dimensional contents of vision are the product of judgment and not ground-level proper objects (Martha Bolton, "The Real Molyneux Problem and the Basis of Locke's Answer," in *Locke's philosophy: content and context*, ed. G.A.J. Rogers (New York: Oxford University Press, 1994), see especially pages 80-81). Locke is not the topic of this paper, so I will adopt the standard interpretation without extensive defense.

eye vergence angles, etc., as premises and produces distance assessments as conclusions. Locke needs a different positive account of the spatial content of the judgments we end up with (he cannot simply appeal to abstract ideas of extension as does Descartes) as well as how we get to those contents from the initial visual input (he cannot appeal to an innate geometry).

On Locke's view our sensory modalities deliver to us information in various formats: we receive ideas of colors¹⁰, shades, distances, pressures, felt resistances, tones, and so forth. We can put these kinds of sensory information into three (or two) groups. The first is *spatial*; the second *quasi-spatial*; the third (if anything falls in the group) I will call *punctate*. *Spatial* is, initially anyway, straight-forward — I can see that one point of light is between two others, that it is closer to one than the other; I can get similar information via touch. The primary contrast here is with *quasi-spatial*. Many of the channels of information we receive through sensation are such that the ideas they occasion have features that can vary along one or more dimensions, but these dimensions are not genuinely spatial dimensions. Sounds can vary along the continuous dimension of pitch, and also along the continuous dimension of volume; colors can vary along three continuous dimensions of saturation, hue and brightness; a felt surface can feel more or less solid as it offers more or less resistance to pressure. I will call these qualitative continua *quasi-spatial manifolds*. The continuum involved in genuinely spatial content I will call a *spatial manifold*. Confusions between spatial and quasi-spatial manifolds can result from the fact that it is common to refer to dimensions of these quasi-spatial manifolds and points or regions on them by means of expressions with prototypically spatial application. Thus we speak of *high* and *low* pitches, turning the volume *up* or *down* on the stereo, *low* pressure, a *big* or a *small* headache, and so forth. But these metaphors notwithstanding, it is crucial to keep the distinction in mind. In

¹⁰ As an expressive convenience, I will often use phrases of the form 'idea of X' as shorthand for 'idea that has X as an aspect of its content'.

the third category, punctate, will be those channels of information or aspects of qualitative content that are not naturally orderable along a dimension of variation. It's not clear whether there are any such things, but I leave it as a possibility by reserving the third category. None of my points will hinge on whether this category has any membership.¹¹

Experience characterizable in terms of such manifolds is delivered to us via various sense modalities. Through touch we receive three dimensions of spatial information, ideas of resistance, ideas of heat, and so forth; through vision we receive two dimensions of spatial information, and ideas of hue, brightness, saturation, and so forth. It will be noted that both modalities deliver ideas with *genuinely spatial* content, even though through vision this spatial content is (strictly speaking) limited to two dimensions.¹²

A bit of notation will help make the following points more clearly and succinctly. We can break genuinely spatial content into three dimensions, adequately expressed in terms of the three behaviorally relevant axial asymmetries of up/down, left/right, and front/back, which define the dimensions of height, breadth, and depth respectively.¹³ And we can break these dimensions into

¹¹ The closest example I can manage of such punctate qualitative content comes from the modality of olfaction, a lively if small area of chemical and neuroscientific research (see e.g. G. Ohloff, *Scent and Fragrances: The Fascination of Odors and their Chemical Perspectives* (New York: Springer-Verlag, 1994); Philip Kraft, Jerzy A. Bajgrowicz, Caroline Denis and George Frater, "Odds and Trends: Recent Developments in the Chemistry of Odorants," *Angewandte Chemie International Edition* 39 (2000): 2980-3010.). Apparently (Ralph Adolphs, personal communication) it is possible to synthesize odorants that, to paraphrase the description I was given, don't smell at all like anything else you have ever smelled (for the point of finding such odorants, see e.g. Tony W. Buchanan, Daniel Tranel, and Ralph Adolphs, "A Specific Role for the Human Amygdala in Olfactory Memory," *Learning and Memory* 10 (2003): 319-325; A.K. Anderson, K. Christoff, I. Stappen, D. Panitz, D.G. Ghahremani, G. Glover, J.D. Gabrieli, and N. Sobel, "Dissociated neural representations of intensity and valence in human olfaction," *Nature Neuroscience* 6 (2003): 196-202). Whether there are genuinely punctate qualia is not central to the task of this paper, and so I won't pursue this further.

¹² This is a terminological point of some weight. I will be using 'spatial' in such a way that not only 3 dimensional space, but also 2 dimensional space (e.g. a plane) are *spatial*. Some commentators appear to use 'spatial' in such a way as to imply three dimensions. But it will be crucial to be able to keep clearly in mind the *difference* between the kind of relation that two entities on a Lockean visual plane bear to each other, and the relation that the auditory entities of Middle C at 35 decibels and High A at 50 decibels bear to each other. The first is a spatial relationship involving genuinely spatial distance; the second is only quasi-distance, though it supports true statements such as 'High A at 50 dB is closer to Middle C at 35 dB than is High F at 70dB' (provided we correctly understand, as we all do, the metaphorical nature of 'closer'). Of course distances between points in the visual height/breadth plane may not correspond to any determinate measure of physical three-dimensional space, but that is a separate issue. When I direct my friend's attention towards a particular star by saying that it is the one closest to the horizon, the assessment of distance to the horizon, which is made relative to a two-dimensional construal of the scene, is silent on the three-dimensional distance that the star bears to the horizon. I will touch on this again in Section 2.1.

¹³ This means of distinguishing spatial dimensions was, to my knowledge, first articulated with clarity in the modern era by Kant (Immanuel Kant, *Concerning the Ultimate Foundation of the Distinction of the Directions in Space*, in *The Cambridge Edition of*

(i) a depth element and (ii) a planar height-breadth element. For Locke, the planar element is available through both vision and touch, and so I will abbreviate it as p_B , for 'planar bi-modal' content. Depth, however, is restricted to touch, and so it will be abbreviated as d_T , for 'depth touch' content.

In addition to these spatial elements, vision provides a number of quasi-spatial dimensions of content that come to serve as cues for depth (this will be discussed more in a bit). These elements I will abbreviate as δ_V . They are visual, hence the 'V' subscript, but are not themselves *really* ideas of depth, though they come to serve as cues for depth judgments, hence the Greek (rather than Roman) ' δ '.

We can now summarize one possible interpretation of Locke's account of how vision comes to provide three-dimensional spatial content, and at the same time introduce two technical terms that will figure in the remaining discussion: *correlation* and *coordination* (the third, *calibration*, will be introduced shortly). This proposal concerning the spatiality of visual experience I will call **Option One**. The breadth and height dimensions of visually apparent space are simply *given* directly through the senses, and they are identical to the breadth and height dimensions given through touch (p_B). Touch also includes a third dimension, tangible depth (d_T). Furthermore, vision provides us with a number of quasi-spatial manifolds (δ_V) — especially brightnesses, hues, and gradients of these — which *are* proper objects of sight and are apparent on the two-dimensional sense-datum plane. Now the elements of δ_V , though they are not strictly spatial, are nevertheless *correlated* with various spatial features of our visual environment. What this means is that, as a contingent feature of the way the world and our sensory systems work, depth-related features of objects reliably cause certain patterns of variation in elements of δ_V . A

darkening shading gradient is correlated with a surface that is curving away from the viewer, for instance. But these correlations are not at all the sort of thing that is either apparent from the bare nature of the elements of δ_V themselves, nor anything that can be deduced from them *a priori*. Nevertheless, because these correlations obtain, they can be learned through experience, specifically through bimodal exploration of one's environment during which the correlations are made manifest. Once learned, these correlations allow the apparent δ_V features to provide sufficient grounds for reliably judging concerning the comportment of the objects seen with respect to the dimension of depth — including the judgment that one is looking at a globe on the basis of the variously colored circle¹⁴ which *is* the proper object of our vision. One comes to consistently notice, for example, that a certain kind of shading gradient is consistently associated with a surface that curves away from oneself, and more or less automatically judge that what one is seeing is a globe.

The subject has, at this point, *coordinated* the relevant quasi-spatial visual manifolds δ_V with d_T . A coordination obtains between two manifolds if there is a *prima facie* relation of purport between elements of those manifolds — shading gradients come to have the purport of a surface curving in depth, for example.¹⁵ Because they rely on the exploitation of learned coordinations and are not given directly through the senses, Locke calls the ideas of depth one arrives at on the strength of the two-dimensional visual input and appropriate associations *ideas of judgment*.

¹⁴ Technically what we would be seeing on Locke's account is a *disk*, not a *circle*. A circle is the set of points equidistant from some point; a disk is the two-dimensional region within and including those points. I will follow Locke's loose usage here, and will adopt a similarly lax approach to 'square', 'sphere', and so forth.

¹⁵ Notice that as I am using the expressions, coordinations are psychological in a way that correlations are not. A coordination is a psychological process or the result of that process, whereas a correlation is an informational relationship between how things are and how they appear. The link between them is that learned coordinations are typically underwritten by correlations. In other words, it is because the correlations obtain that we learn the coordinations.

In an attempt to bring this point home, Locke quotes a question posed to him by William Molyneux:¹⁶

... I shall here insert a Problem of that very Ingenious and Studious promoter of real Knowledge, the Learned and Worthy Mr. *Molineux* [sic.], which he was pleased to send me in a Letter some Months since; and it is this: *Suppose a Man born blind, and now adult, and taught by his touch to distinguish between a Cube, and a Sphere of the same metal, and nighly of the same bigness, so as to tell, when he felt one and t'other, which is the Cube, which the Sphere. Suppose then the Cube and Sphere placed on a Table, and the Blind Man be made to see: Quaere, Whether by his sight, before he touch'd them, he could now distinguish and tell which is the Globe, which the Cube?* To which the acute and judicious Proposer answers, *Not. For though he has obtain'd the experience of, how a Globe, how a Cube affects his touch; yet he has not yet attained the Experience, that what affects his touch so or so, must affect his sight so or so; Or that a protuberant angle in the Cube, that pressed his hand unequally, shall appear to his eye, as it does in the Cube.* (*Essay* Book 2, Chapter 9, Section 8, pages 145-6, emphasis original)

Locke agreed with Molyneux's answer and his reasoning. But questions immediately arise. The first has to do with what the conditions of success are for the subject in the thought experiment (let's call her 'Molly'). While there are more than two ways to go here, we can get by with a strong and a weak reading of the question. On the strong reading, in order for Molly to pass the test, she must, upon opening her eyes, see the cube *as a cube*, and see the sphere *as a sphere*, and *on this basis* correctly determine which is the cube and which is the sphere. On this reading of the question, a proponent of Option One should answer negatively as did Molyneux and Locke, for Molly will not, upon having her vision restored, be in a position to have any ideas of judgment annexed to the proper objects of her sight, and will therefore only have, through vision, ideas with spatial import limited to p_V and δ_V . And there is no depth purport carried by any of δ_V . But her ideas of cubes and spheres are ideas of essentially three-dimensional solids.

¹⁶ Molyneux's inherited wealth allowed him freedom for a number of pursuits. Not only did he translate philosophical texts into English (including Descartes' *Meditations* (Rene Descartes, *Meditationes de prima philosophia*, ed. and trans. William Molyneux (London: Benjamin Tooke, 1680)), his major work *Doptrica Nova* (William Molyneux, *Dioptrica nova: a treatise of dioptricks in two parts* (London: Benjamin Tooke, 1692), hereafter cited in the text as *Dioptrica Nova*) made substantial contributions to the mathematical physics of telescopes and the functioning of lenses. His wife Lucy went blind shortly after their marriage in 1678, perhaps partially explaining his interest in the capacities of the blind.

Therefore, initially at least, Molly is incapable of *seeing* the cube as a cube, or the sphere as a sphere. Therefore, she would fail the strong version of the test. After learning, of course, the ideas of judgment Molly would have on the basis of δ_V and its coordination with d_T would allow her to pass the strong version of the test. This coordination, once established, would imbue various elements of δ_V with depth purport, with the result that Molly would then be able to see the cube *as a cube*, etc.

On a *weak* reading of the question, however, Molly passes the test simply if she can determine which object is the sphere, and which the cube — whether through perception, excogitation or divine guidance. If these are the success conditions, then it would seem as though Molyneux's and Locke's negative answer is difficult to jibe with an account of visual perception as per Option One, even pre-learning. For it would seem (however, see the discussion of *calibration* below) that the visual ideas Molly would receive upon opening her eyes, even if they had only two-dimensional spatial import, would be sufficient for her to pass the weak test. For surely one aspect of her tactile experience with cubes is that they have square faces, and an aspect of her tactile experience with spheres is that spheres do not have square faces, but have circular circumferences. This alone would seem to be enough to strongly, even decisively, suggest to Molly which object she is looking at is the cube, and which the sphere, since the visual idea produced in her by the sphere will be a circle, and that by the cube a square (or at least three quadrilaterals, one of which could be square, depending on the viewing angle).

Notice that on the weak interpretation of the question, success is possible even in cases more radical than the comparison of ideas having two-dimensional content with those having three-dimensional content. Molly might be able to succeed on weak versions of the Molyneux question involving comparisons between entities that invoke *any* spatial or quasi-spatial

manifolds. While it might be the case that individual qualia from different domains have no natural likenesses or similarities (the feeling of luke warmth is not more like Middle C than it is like High A — in the terminology of this paper, they have no shared *purport*), it is nevertheless apparently true that, because these qualia are elements within homogenous qualitative continua, *patterns of variation* of qualia from different modalities *can* be compared. Consider Molly's hand placed on a heating element. The plate starts off feeling room temperature, and then slowly warms up until it is very warm and stays very warm. Next consider a case where the plate initially feels room temperature, then suddenly becomes very warm and stays very warm. Compare these two patterns of change of thermal intensity to first, a sound beginning at pitch Middle C, and then slowly rising in pitch until it reaches High A; and a case where the sound starts off at Middle C and then quickly jumps to High A. As described there is an obvious similarity between the first tactile and first auditory case, and an obvious similarity between the second tactile and second auditory case. And even the fact that both the gradients exploited in this case are temporal gradients can be abstracted from: cases involving the comparison of spatial and temporal gradients are easy to construct.¹⁷ The possibility of exploiting patterns such as these in order to make reliable cross-modal matches was behind Leibniz's affirmative response to Molyneux's question, in which he remarked that the tangible cube, unlike the tangible sphere, will have “eight points which are distinguished from all the others”,¹⁸ which will suggest that the two-dimensional projection of the cube is correct. And Berkeley himself remarked that a visual square is fitter than a visible circle to represent a tangible square because the former has distinct parts (*NTV* 142, p. 228-9).

¹⁷ The easiest and most obvious would be a curved line that, as we follow it from left to right, is horizontal, and then slowly curves upward and rises to some higher horizontal level and flattens out; and a similar line that starts out flat and horizontal and quickly rises to a higher level then flattens out. Here the temporal gradients have been swapped for spatial gradients.

¹⁸ Gottfried Wilhelm Leibniz, *New Essays on Human Understanding*, ed. and trans. Peter Remnant and Jonathan Bennett (Cambridge: Cambridge University Press, 1981), Chapter 9, pages 135-139. Leibniz is quite explicit about adopting the weak interpretation of the test, and admits that on a strong interpretation Molly would fail.

Everything I have said so far is fairly standard Locke interpretation supplemented with some terminology and notation. The new twists come with the introduction of the notion of *calibration*. Calibration concerns, for lack of a better expression, the *correctness* of the correspondences established by a coordination.¹⁹ There are three kinds of calibration: *orientation calibration*, *gradient calibration*, and *translation calibration*, based on three respects in which a coordination can be correct or incorrect. I will use the expression *dyscalibrated* (and its cognates) to refer to manifolds that *are coordinated*, but incorrectly calibrated, and *uncalibrated* (and its cognates) for manifolds that are not coordinated at all. The different sorts of calibration will be easiest to illustrate in the case of spatial manifolds. Glasses with inverting prisms reverse the *orientation* calibration of vision and touch. Inverting lenses can be worn that, as the wearer would describe it, make everything look upside down. In such a case, what looks to be *up* through the visual modality will feel to be *down* in the tactile modality. Inversions of left and right are also possible. A *gradient* dyscalibration is effected by fish-eye lenses. With such lenses, the correct orientation of the visual field and the tactile field remains intact. What is changed are the spatial gradients. Specifically, spatial intervals that feel to be equal through the tactile modality no longer look to be equal through the visual modality exactly because things in the center of the visual field are disproportionately magnified at the expense of the visual periphery. Lenses which simply make everything look to be shifted some constant amount to the left or right (or up or down, or closer or farther) would effect a *translation* dyscalibration.²⁰ One can easily imagine lenses effecting discontinuous gradient dyscalibrations that would make a circle look like a square, for example.²¹

¹⁹ The notion of calibration as I use it is not the same as the notion of calibration discussed by Schwartz in *Vision*.

²⁰ Strictly speaking, gradient dyscalibrations are special cases of translation dyscalibrations, cases in which the translation dyscalibration is not affine.

²¹ That is, possible to make a single circular object in the center of the visual field appear square. I suppose it might be possible to make lenses that would make a grid of circles appear to be a grid of squares, but this would be much trickier. See Thomson

There are three crucial points about calibration. The first is that even though the examples I employed were all spatial, calibration problems can certainly occur between quasi-spatial dimensions as well. A gradient dyscalibration analogous to a fish-eye lens in the case of felt temperatures might exaggerate the felt difference between temperatures near the luke-warm part of the continuum at the expense of discriminatory capacity elsewhere. And a pronounced version of this might make what would otherwise feel like a slow continual increase in warmth feel like a constant felt temperature that quickly jumps to a warmer temperature and remains there.

The second point is that although the examples I used involved tweaking sensory input via the addition of external devices such as lenses, it should be obvious that even in normal paraphernalia-free cases calibration is an issue. Presumably the fine-tuning of these calibrations mostly occurs during infancy, and so for normal adults re-calibration is typically only needed in special circumstances, such as when one gets a new pair of glasses. But even in normal humans gradient dyscalibrations can occur. Malebranche observed²² that for some people objects look to be different sizes when viewed through the left or the right eye (a linear gradient dyscalibration of the spatial content delivered through the right and left eyes).

The third point is the most important for what follows. There are two different kinds of case in which two manifolds can fail to be correctly calibrated: either they a) are not coordinated at all, and *a fortiori* not correctly calibrated (*uncalibrated*); or b) they *are* coordinated but calibrated incorrectly (*dyscalibrated*). Because dyscalibrated manifolds are coordinated there is a *prima facie* correspondence between the elements and gradients of the manifolds, but these

(Thomson, Judith Jarvis Thomson, "Molyneux's Question", *Journal of Philosophy* 71 (1974): 637-650), who discusses the issue of whether, e.g., a world where squares appear to be circles, is possible. I agree with Evans (Gareth Evans, "Molyneux's Question" in Gareth Evans, *Collected Papers* (Oxford: Oxford University Press, 1985) that Thompson's suggestion that such considerations were part of Molyneux's thinking cannot be correct (though Thomson's article is otherwise quite interesting).
²² *ST* Book 1, Chapter 6, Section 1, p. 28. Malebranche cites as evidence of this "...observations reported in the *Giornale de' letterati*, January 1669." I have been unable to discover what these observations were, but in a footnote he adds that "One of my friends always sees the letters of a book larger with the right eye than with the left."

correspondences are not correct. Calibrated manifolds are not merely coordinated, but correctly coordinated. When we get to Berkeley, we will see how illicit sliding between uncalibration and dyscalibration underwrites a bad argument.

All of the above examples of calibration gone awry (inverting and fish-eye lenses, etc.) were cases of dyscalibration. What is crucial is that in cases of dyscalibration it remains true that *there is an apparent calibration, exactly because there is a coordination*. When you don inverting lenses the visual up/down axis is (orientation) dyscalibrated with the tangible up/down axis. But it remains true that there is an apparent calibration. Your feet look to be *up* — not in some unspecified ‘direction’, nor not in any direction at all, but *up*. The problem is that what visually appears to be *up* is really *down*. If you were to don lenses with non-constant curvature, a straight line would look curved — not of *no shape*, but *curved*. A case of *uncalibration* between vision and touch would be one in which, e.g., your feet appeared visually to be in *some* direction, but this direction was not tangible up, down, left, right, or any other (tangibly specifiable) direction — whether this is intelligible will be discussed below. A better example of uncalibration would be the manifolds of pitch and hue. While each of these manifests in a continuous dimension of variation, it is not the case that these dimensions even purport to be aligned in any way.²³

Now that calibration issues have been introduced, it should be clear that *even* on the weak interpretation of MQ, and *even* if we limit the experiment to two-dimensional shapes, a positive answer is not guaranteed. Before Molly has sufficient bimodal experience to *calibrate* her vision and touch, there is no assurance that the sphere won't cause ideas that have a squarish content, or that the cube won't cause ideas with, to put it loosely, quite rounded sides. Interestingly, at a few

²³ It must be kept in mind that some modalities might be such that though some aspects of their content are not calibrated, others are. With sound and light, for example, changes in timbre and changes in hue may be uncalibrated, but changes in volume and changes in brightness are calibrated to some extent, and that *prima facie* calibration is even lexicalized in English by means of the modality independent 'intensity'.

spots in his *Dioptrica Nova*, Molyneux finds himself teetering on the edge of addressing just a calibration issue. For example in Book 1, Proposition 31, titled “Concerning the Apparent Place of Objects seen through Convex Glasses”, Molyneux, in an explanation as to why the apparent location of two points will lie along a certain straight line (the details of the example and Molyneux’s explanation are not pertinent), he remarks parenthetically: “unless perhaps Convexes on very small spheres, will represent the object crooked or bowed, but of this we shall take no notice... .” (*Dioptrica Nova*, p. 116) So he was well aware that the physiological apparatus of vision could project an image of a bowed line when the seen object was straight. And given the apparently ubiquitous assumption that the two-dimensional mental image was a reflection of the two-dimensional retinal image, we see Molyneux brushing past but purposefully failing to take notice of gradient calibration as an issue. But with calibration as an issue, before *some* experience, even Leibniz’ gambit may not tip Molly off correctly. This brings us to the second option concerning the relation between the spatiality of visual experience and tactile experience.

Option Two: The breadth and height dimensions of visually apparent space are simply *given* directly through the senses, and they are identical to the breadth and height dimensions given through touch (p_B). However, there is no presumption that vision and touch are correctly calibrated, before learning, with respect to these two dimensions. What looks up might feel to be down, what feels straight might look curved, and so on. Correct calibration between vision and touch with respect to these dimensions is learned through bimodal experience. Furthermore, touch includes a third dimension, depth (d_T), and vision provides us with a number of quasi-spatial manifolds (δ_V) — especially brightnesses, hues, and gradients of these — which *are* proper objects of sight and are apparent on the two-dimensional sense-datum plane. In short, the

difference between Options One and Two is that One, but not Two, just assumes that the planar content of vision and the planar content of touch are not only coordinated, but correctly calibrated, *ab initio*. Option Two only assumes a coordination — some *prima facie* purport — but holds that experience is needed to effect a correct calibration between planar visual and planar tactile spatial content.

Given Option Two, another possibility can be envisioned as to why Locke and Molyneux provided a negative answers to MQ. Perhaps Locke and Molyneux recognized that calibration is an issue, and so even on a weak interpretation of the question they hold that Molly will fail. Given the potential dyscalibration of the height/breadth plane of visual experience with any tangible spatial dimensions, there is no guarantee that the globe will induce a visual idea of a circle, or that the cube won't cause visual ideas of smooth curved lines lacking any sharp discontinuities, and so short of divine guidance or (so to speak) blind luck, Molly can't be expected to succeed even on the weak version of MQ.

The exegetical situation points pretty clearly, though not entirely unambiguously, towards the attribution of Option One to Locke and Molyneux as correct. On this reading depth is the issue of chief interest, and the lack of coordination between tangible *depth* and anything directly given in pre-learning vision is what guarantees Molly's failure on the strong reading of MQ. There is ample reason to think that Locke (and Molyneux) were especially concerned with depth. Locke's treatment of this topic starts with a description of the visual idea caused by a globe as being a flat circle, and the discussion from there clearly implies that the disparity between this and what we take ourselves to see — a globe — is the problem being addressed. And Locke's causal assuredness that the idea caused by the globe is a flat circle would seem to indicate that calibration issues were not on the docket.

As for Molyneux, the issue of depth also seems to be primary. This is explicit in an earlier version of the letter that Locke did not quote.²⁴ But there are two hints that Molyneux may have been aware, even if vaguely, that calibration was one of the issues faced by Molly. First, one of the few diversions from mathematical and practical optics in Molyneux's *Nova Dioptrica* (proposition 28) was a statement of the inverted retinal image "problem".²⁵ To the extent this is a problem, it is an orientation calibration problem: what is projected on the top of the retina is actually what is at the bottom of the visual scene. But an orientation dyscalibration wouldn't be enough to foil Molly on the weak version of the test — upside down circles and squares can be as easily associated with globes and cubes as their correctly oriented counterparts. Second, there is possibly a hint of concern with gradient calibration in the last clause of Locke's quote, which reads "... or that a protuberant angle in the cube, that pressed his hand unequally, shall appear to his eye as it does in the cube ...". This at least possibly suggests recognition of the fact that a felt right angle might correspond to a gentle visual curve. There will be a discontinuity of the pressure felt in the hand when pressed against an angle — a single point of great pressure surrounded by little or no pressure. The suggestion might be that this could correspond to a visual image lacking any such sharp discontinuity — such as a gentle visual curve. Especially if

²⁴ Molyneux's had actually sent 2 letters to Locke, the first of which was sent in 1688, two years before the publication of the first edition of the *Essay*. Molyneux had read a 50 page abridgment of the *Essay*, written by Locke and published in 1687 in the *Bibliothèque universelle et historique*, a French periodical operated by Locke's friend Le Clerc, which also published an announcement of Molyneux's *Doptrica Nova*. Locke appears not to have responded to the first letter (Bodleian MS Locke c. 16, fol 92r). The second letter, of 1693, is the one that Locke quotes in later editions of the *Essay*. The versions of MQ in the two letters are nearly identical except that the first includes an additional component of the question. The relevant additional question reads: "... Or Whether he could know by his sight, before he stretched out his Hand, whether he could not Reach them, tho they were Removed 20 or 1000 feet ... ". For interesting discussion of the two letters, see Park (Desiree Park "Locke and Berkeley on the Molyneux Problem," *Journal of the History of Ideas*, 30 (1969): 253-260, see especially p.254 n1.)

²⁵ Molyneux, *Dioptrica Nova*, Proposition 28, section 4, pages 105ff. Interestingly, Molyneux's 'solution' to the retinal image problem is very similar to the one Berkeley himself would later offer in *NTV*. For example, Molyneux offers a redefinition of 'inverted' and 'erect' as being relative to the Earth, so that any image of a man that has his feet touching the earth and his head away from it is, by definition, erect, not inverted (*Dioptrica Nova*, page 105), while the parallel move in *NTV* is "... a man born blind... could mean nothing else by the words higher and lower than a greater or lesser distance from the earth..." (*NTV* 94, page 209). More interestingly still, Berkeley cites Molyneux as an *opponent* on this topic in *NTV* 89, p. 208. He credits to Molyneux an impulse theory, according to which impulses felt on the bottom of the retina are judged to be from objects above the eye in accord with geometric and optical principles (Molyneux's discussion here is in *Dioptrica Nova*, Book 2, Chapter 7, page 289). Berkeley fails to mention, or notice, that Molyneux much earlier in the book (*Dioptrica Nova*, page 105) offers a solution that is quite similar to his own.

the lens on the eye is appropriately shaped (recall Molyneux's observations, mentioned earlier, about the distorting effects of lenses with a small radius of curvature). But aside from the last clause of the letter Locke quoted Molyneux does not, to my knowledge anyway, seem to recognize gradient calibrations as presenting a problem for visual perception, and so this clause would seem to be an isolated lapse into this particular insight.

Certainly isolated from Locke, who nowhere seems at all concerned with gradient calibration issues, and seems to have included the last clause of Molyneux's letter without recognizing that it introduced a layer of complexity entirely unaddressable by his own account. If gradient dyscalibrations are possible, then it is by no means obvious that, pre-learning, the image corresponding to a globe would be a circle rather than an ellipse or even a square. It is possible that had the issue of gradient dyscalibration been brought to Locke's attention he would have recognized it as a problem.²⁶ But following such speculations is beyond the scope of this paper.

3. Berkeley

3.1 Introductory

Berkeley's *NTV* was explicitly aimed at optico-geometric accounts, but in some key respects it was a response to Locke's account as well. Like Locke, Berkeley could make no room in his psychological theory for an innate geometry of the sort Descartes and Malebranche appealed to in their account of perceptual distance/depth judgments. But unlike Locke, Berkeley could also make no room for ideas or concepts of extension that were common to more than one

²⁶ For an interpreter who thinks it clear that Locke would not have considered gradient calibrations a problem, see Mackie (J. L. Mackie, *Problems from Locke*, (Oxford: Clarendon Press, 1976), p. 30). Mackie does not, of course, phrase the issue as concerning gradient dyscalibrations, but rather in terms of whether or not Locke would also answer negatively to a two-dimensional version of MQ. Mackie feels that Locke would accept that Molly would be able to match the visible and tangible square, and this would be possible only if gradient dyscalibrations were not an issue.

modality. While Locke explicitly says that ideas of space are common to more than one modality (*Essay* Book 2, Chapter 5, page 127), Berkeley clearly thinks that any such ideas would be *abstract* ideas, and he rails against such ideas at length, blaming philosophers' appeal to abstract ideas for most of the difficulties in which they find themselves ensnared.²⁷

The offending ideas in this case are the dimensions of breadth and height, which according to Locke are common between vision and touch. The aspect of Berkeley's proposal that we will focus on is that part dealing with our visual access to breadth and height, and so it will be useful to be a bit more explicit about this dimension of presumed visual content, p_V (for *visual planar*) that has so far been lumped under p_B of Options One and Two. The central idea is that of a two-dimensional visual field conceived as a region within which the proper objects of vision appear. There is good reason to think that the visual field is a psychological reality,²⁸ even if analyzing it in terms of a screen on which inner objects are projected, as Locke seems to, has difficulties. Of the three dimensions manifest in normal adult vision, the height-breadth plane has features not interchangeable with the axis of depth. For instance, entities in the visual field can easily be located and assessed with respect to the height-breadth plane while ignoring depth. While looking at the sky at night, I might direct my friend's attention to a particular star by describing it as 'the one just to the left of that treetop'. The naturalness and effectiveness of such a description is surely to be explained by the ease with which the visual field can be treated as a

²⁷ For useful extended discussion of Berkeley's views on abstract ideas, see George S. Pappas, *Berkeley's Thought*, (Ithaca, NY: Cornell University Press, 2000); for discussion of Berkeley's views on abstract ideas specifically in the context of *NTV*, see Pappas "Abstract Ideas". It was not just abstract ideas of extension that Berkeley found objectionable, but abstract ideas of any sort, especially the primary qualities including not only extension, but motion and number. Berkeley takes these on in *NTV* as well, but extension, indeed only a couple aspects of extension, will be my concern in this paper.

²⁸ The size (in degrees) of the visual field can be measured, both during normal perception and even during mental imagery. The former sort of assessment is straight-forward enough. For the latter, one sort of technique involves having subjects, with their eyes closed, produce visual imagery of common objects like a bus or meter stick, and then imagine 'approaching' the object until the ends are no longer within the imaginary visual field. Results are that the imaginary visual field is roughly the same size as the overt visual field, a result that interestingly tracks changes to cortical damage (see M.J. Farah, M. J. Soso, and R.M. Dasheiff, "Visual angle of the mind's eye before and after unilateral occipital lobectomy," *Journal of Experimental Psychology: Human Perception and Performance* 18 (1992): 241-6).

sort of two-dimensional spatial manifold.²⁹ Attempting similar assessments within a plane spanned by either of the other two pairs of axes (depth-breadth or depth-height) while ignoring the third is not at all psychologically natural. Of course, the ease of such descriptions does not imply that these are the only or best or even primary descriptions. The point is just that this is a conception of the visual field that is *not* a mere construct of modern philosophy, but is a standard part of our common understanding, one that has a firm physiological basis.³⁰ The visual field considered as a two-dimensional expanse, where the two dimensions are breadth and height, I will call the *2D visual field*.³¹

A further point of clarification about the 2D visual field is in order before we proceed. I follow standard interpretations here in treating ‘distance’ and ‘depth’ as more or less synonymous, and thus do not agree with the claims of Atherton (*Berkeley’s Revolution*) and Schwartz (*Vision*) that distance should be kept distinct from depth or ‘bulginess’. The Schwartz/Atherton point seems to be that objects can be seen as ‘bulgy’ or as having *depth* even

²⁹ Consider ‘Polaris is closer to the horizon than Betelgeuse’ made as a natural and clear assessment of distance as ascertained on the two-dimensional visual field. To respond that in fact Betelgeuse is over one hundred light years closer to the horizon than Polaris would not be to correct an error, but to indulge in a sort of cute impertinence.

³⁰ Not only is the retina two-dimensional, but more importantly there are areas of visual processing and ocular motor control, such as the superior colliculus, lateral geniculate nucleus, and early visual cortical areas, that operate on what is essentially the two-dimensional visual plane. The superior colliculus, for example, is centrally involved in saccades and gaze direction. And as far as moving the eyes to foveate something goes, the location of the stimulus on the two-dimensional plane is of primary importance. The assessment that ‘The Polaris is closer to the horizon than is Betelgeuse’ can be understood physiologically (though this is not the only way to understand it) as: when foveating the horizon, it is a shorter eye movement to foveate Polaris than to foveate Betelgeuse.

³¹ Thomas Reid (in *An inquiry into the human mind on the principles of common sense*, ed. Derek Brookes (Edinburgh: Edinburgh University Press, 1997)) describes the 2D visual field more fully:

... let us distinguish betwixt the *position* of objects with regard to the eye, and their *distance* from it. Objects that lie in the same right line drawn from the centre of the eye, have the same position, however different their distances from the eye may be: but objects which lie in different right lines drawn from the eye’s centre, have a different position; and this difference of position is greater or less, in proportion to the angle made at the eye by the right lines mentioned. Having thus defined what we mean by the position of objects with regard to the eye, it is evident, that as the real figure of a body consists in the situation of its several parts with regard to one another, so its visible figure consists in the position of its several parts with regard to the eye; and as he that hath a distinct conception of the situation of the parts of the body with regard to one another, must have a distinct conception of its real figure; so he that conceives distinctly the position of its several parts with regard to the eye, must have a distinct conception of its visible figure. (Reid, *Inquiry*, Chapter 6, Section 7, page 143)

This conception of visual experience (aided and abetted by the well-known ability of artists to suggest depth by their placement of pigments on a flat canvas, and by the discovery that ‘flat’ images of the environment are projected onto the retinae) is undoubtedly behind Locke’s view, where he takes vision to properly consist of a flat two-dimensional arrangement of flat color patches.

in cases where we are not in a position, on the basis of our visual experience, to assign any determinate distance to points on the apparently bulgy surface. I can see that a bowling ball located 20 meters away is spherical, and hence that its surface is differentially oriented in depth, even though the difference in distance between the closest point of the ball and a point on the outer edge is not one I can reliably judge at that distance (this example is mine, not theirs).

Atherton gives the following example, which she credits to Schwartz: "... a picture of, say, just a globe seen through a stereoscope will look bulgy or in depth but, without any other distance cues, it won't look to be at any particular distance" (Atherton, *Berkeley's Revolution*, page 75).

And of this example, she says "It is true that in ordinary circumstances, when something looks bulgy, viewers are able to make some estimate about how far away the front of the object is from the back, but, in the absence of distance cues, mere bulginess alone wouldn't permit such an estimate. In the example of the stereoscope, the front of the cube could be any distance at all from the back." (*Berkeley's Revolution*, page 75, n. 25). But this does not seem to be correct (we have moved from 'no particular distance' in the first quote to 'any distance' in the second).

Specifically, the front simply cannot be seen to be *zero* distance from the back — it can't be seen as a flat arrangement of polygons. More generally, while it is true that specific distance information is not required to see something as bulgy, what is required is either *relative* or *proportional* distance information: the back of the cube will look to be, say, 20% further away than the front (proportional distance information); and while I cannot make reliable distance judgments about the parts of the surface of the bowling ball, I can make pretty good *relative* distance judgments: the side of the ball is a few centimeters further away than the point nearest me. Perhaps the distinction between i) absolute depth judgments, and ii) proportional or relative depth judgments, is all Atherton and Schwartz have in mind with the distinction between

distance and bulginess. If this is the case I have no quarrel; but then it does not seem to be a distinction between distance and something other than distance so much as distinguishing different kinds of distance judgment, viz. judgments of absolute, proportional, or relative distance.

So much for clarifications concerning the nature of the 2D visual field. Berkeley's genius lies in the fact that he was, to my knowledge, the first philosopher to see that the status of even height and breadth content as a proper visual sensible can be challenged in the same way that Locke had challenged the status of depth content as a proper visual sensible. The nature of his challenge was to first take over, with substantial improvements, a more or less Lockean account of the 'distance' content of visual experience; and then second, to try to do for the other two dimensions — breadth and height — what was done for depth. That is, to argue that the apparent breadth-height features of visual experience are likewise matters of judgment based on experienced correlations, and not aspects of the proper objects of vision *per se*. But genius aside, Berkeley's challenge has problems.

3.2 Visual depth: Berkeley's positive account

Sections 2 through 51 of *NTV* are concerned with the issue of our visual perception of distance/depth. Content of distance/depth (as well as breadth and height) are for Berkeley properly carried by ideas of the modality of touch, proprioception, kinesthesia, etc.³² Through vision we receive only color and light, together with some number of other sensations that

³² Berkeley, strictly speaking, needs to observe a cap of one-modality-per-content (a content obtainable through more than one modality would be a common sensible), and so whatever delivers genuine spatial content must be a single modality. On the other hand, it is not obvious that touch, kinesthesia, and proprioception are all one undifferentiated modality. Exploring this is beyond the scope of this paper. See Pappas' "Abstract ideas and the *New Theory of Vision*" for discussion of Berkeley's *heterogeneity thesis*.

accompany our visual experience, such as how we are moving our eyes or head when we see something, feelings of strain in our eyes, and so forth. That is, a visual experience of some object or scene, while not including any ideas of depth, does include a number of quasi-spatial manifolds, such as differences of shading and brightness, degrees of clarity or confusedness, amounts of felt strain of the eye muscles. Like Locke, Berkeley's idea is that these components of δ_V carry information about, because correlated with, depth/distance, and eventually get coordinated with tangible distance, and these coordinations, when formed, are the bases of the apparent depth content carried by visual experience.

The principle difference is that Berkeley recognizes more manifolds than Locke, including sensations of movements, etc., in δ_V , but otherwise this part of Berkeley's proposal is essentially parallel to Locke's account. For example, objects that are very close can be out of focus (and degree of out-of-focusness is being treated as a one dimensional manifold here, one that will figure centrally in Sections 3.5 and 3.6), or require eye strain (amount of strain is the relevant manifold) in order to be seen clearly (*NTV* 21 page 175; *NTV* 27, page 176-7); gradients of shading carry information about the orientation and curvature (in depth) of the surface; and even the feelings of the eye muscles (mainly the relative lengths of the medial recti and lateral recti, which largely control the eyes' left-right rotation and orientation, and thus jointly carry information about the eyes' vergence angle) carry information about depth (*NTV* 16, page 174). But although these quasi-spatial manifolds carry *information* about distance/depth, the correlations involved are all contingent, and hence experience is required in order to exploit this information in the formation of judgments concerning depth.

3.3 Visual depth: Two strands in Berkeley's negative account

The previous section outlined Berkeley's *positive* account of apparent visual depth. Now I want to look more carefully at Berkeley's *negative* account — his reasons for thinking that depth is not given immediately through vision. The basic form of argument is to show, for each of the manifolds that *is* given through vision and correlated with depth, that this manifold considered by itself shares no purport with the tangible manifold of depth.

Crucially, Berkeley has two different kinds of consideration for the lack of shared purport between these depth-cue manifolds and actual distance/depth, and the nature of these considerations will be important when we move on to *situation* (or position in the 2D visual field). Berkeley does not clearly recognize the differences between these two kinds of consideration, nor do all his commentators. While the difference is harmless when the topic is depth, it becomes crucial in the case of height-breadth. The *first* kind of consideration is really no more than the bare observation that none of the manifolds in question manifolds share purport with ideas of depth. Shading gradients and eye strains, by themselves, have no *depth purport*, meaning that, e.g., a felt eye muscle strain, *by itself*, does not, indeed *could* not, strike one *as a perception of depth*. That in any case is the doctrine. The lack of any purport entails that the manifolds are not intrinsically coordinated, and hence are uncalibrated, at least before relevant experience. And obviously if the manifolds in question are, considered on their own, entirely unlike each other (like hue and timbre) then there is no reason to think that any particular way of coordinating them would be preferable to any other. The lack of any shared purport entails that any learned coordinations are contingent.

The *second* kind of argument bypasses the issue of shared proprietary purport and focuses on the fact that the particular correlations that are exploited between the relevant visual manifolds and the spatial dimension of depth are contingent. So for example, it might easily have been the case that faintness would correlate differently with distance than the way in which it does (*NTV* 3). Similarly, the correlation between specific felt lengths of the muscles controlling eye orientation and eye angles and the distance of seen objects is also a contingent one. That having my eyes oriented such that their angle to the foveated object is θ degrees feels like *this* rather than like *that* is surely contingent, and could have been otherwise.³³

But notice that these forms of argument are subtly different. The second kind of argument can be marshaled even in cases where there *is*, in fact, *prima facie* purport between the manifolds under discussion. What looks up might be down, as we know from the effects of inverting prisms, and so the correlation between visual up and tangible up is contingent even though there is *prima facie* purport between the manifolds. The first sort of consideration applies to manifolds that are independently known to lack any shared purport and then concludes that any coordinations that are established are contingent. Examples exhibiting these different forms of argument are given in the next subsection.

3.4 Berkeley's negative argument concerning visual magnitude and 'situation'

Berkeley's discussion of situation is aimed at the structure of the apparent 2D visual field discussed in 3.1. I will not single Berkeley's discussion of 'magnitude' out for separate treatment

³³ This is not to imply that Berkeley thinks that what we do is to learn the correct correlations between the sensations of the eye muscles and the optical angles, and then compute distance. The feelings of the eye muscles in various contexts also correlated directly with objects' distance, and so can be associated with feelings of distance without having to go through the mediation of association with optical angles and calculations of distance.

because magnitude (both visual and tactile) is a function of situation: a visual entity's visual magnitude is determined by the relative situation of its borders. A reason Berkeley feels compelled to argue that breadth and height are in the same boat as depth is that if, as in Locke's account, visual experience has as an aspect of its content the 2D visual field as normally understood — that is, as carrying proprietary content relevant to left-right and up-down judgments — then this is an aspect of its content that it shares with touch. Locke explicitly embraced this, but for Berkeley such content would amount to an abstract idea:

But before I come more particularly to discuss this matter, I find it proper to consider extension in abstract: for of this there is much talk, and I am apt to think that when men speak of extension as being an idea common to two senses, it is with a secret supposition that we can single out extension from all other tangible and visible qualities, and form thereof an abstract idea, which idea they will have common both to sight and touch... (NTV 122, page 220)

Such abstract ideas are the philosophical ailment, and the following is Berkeley's medicine:

... the question now remaining is, whether the particular extensions, figures, and motions perceived by sight be of the same kind with the particular extensions, figures, and motions perceived by touch? In answer to which I shall venture to lay down the following proposition: *The extension, figures, and motions perceived by sight are specifically distinct from the ideas of touch called by the same names, nor is there any such thing as one idea or kind of idea common to both senses.* (NTV 127, pages 222-3, emphasis original)

This sets a main desideratum for Berkeley's account: the spatiality of vision must be explained in such a way that no appeal is made to any proprietary content that vision has in common with any other modality, touch in particular. Planar height-breadth content would be just such a taboo content.

An initial problem here is that this position just seems wildly implausible. It is one thing to claim that levels of blurriness and muscular sensations are specifically distinct from depth content, but another to claim that visual left and tactile left, or visual largeness and tactile largeness are likewise completely distinct. Berkeley's strategy of terminologically distinguishing '*visual figure*' and '*visual motion*' from '*tangible figure*' and '*tangible motion*' hardly soothes the chaffing. But few are less daunted by implausibility than Berkeley. He has arguments for this position, indeed arguments he takes to be entirely parallel to that offered in the case of depth. But they aren't quite parallel.

As in the case of depth, Berkeley points out that the correlations between visual features and their tangible counterparts are contingent and even continually altering. Thus what looks to be large might feel quite small, and indeed this might have been a regular correlation (*NTV* 63, pages 194-5); what is judged to have a constant tangible magnitude will be seen to have a variable visual magnitude as we approach or recede from it (*NTV* 55, page 191). Similar points could easily be made about the directions of up-down and left-right, as is clear from inverting prisms, though Berkeley's discussion of such matters is mostly restricted to the inverted retinal image issue.

But it will immediately be noted that in the case of depth there were *two* kinds of consideration: pointing out the evident lack of shared proprietary purport, and arguments from contingency of correlation that apply even when there is shared purport. But in the case of magnitude and left-right and up-down position, it seems that we have only arguments of the second kind. All of the considerations I discussed in the previous paragraph were arguments to the effect that the manifolds could be coordinated in ways *other than the way that they in fact are*, and hence that dyscalibration is a threat. There are no considerations of the first sort — no

claims or observations about there being an obvious lack of purport between visual planar and tangible planar content that parallel the claims about the obvious lack of purport between distance and muscle strain.

The reason is clear. Eye strains and levels of blurriness considered in themselves obviously carry no depth purport. Hence, in the case of depth, both kinds of consideration are applicable: there is no purport, and hence any correlations that are developed are contingent. By contrast, visual left does seem to have a clear purport with respect to tangible direction: tangible left. Ditto for up, down, large, small, etc. Given the *prima facie* purport, Berkeley cannot appeal to the first sort of consideration, but must resort to arguments based on the possibility of alternate coordinations.

Here is why this is important. Berkeley wants to argue that the apparent spatiality of the 2D visual field is due *not* to the fact that the proper objects of vision have as a proprietary element of their content anything like spatial extension, but is due entirely to the fact that these proper objects have non-spatial elements of their content that are cues for height-breadth content. He must establish that these proper objects have no height/breadth purport. But crucially, he produces arguments only of the second kind — arguments that *admit* shared purport but conclude that, nevertheless, the correlations that hold might have been other than they are. Berkeley fails to notice, or fails to bring attention to, the fact that arguments of the first kind are not produced. There are no arguments to the effect that visual left is as unrelated to tangible left as blurriness is to distance. By not clearly seeing the difference between uncalibration and dyscalibration, he seems to have blurred the two kinds of argument together. He seems to have taken credit for making a case to the effect that planar visual and planar tactile content have no common purport. But he has shown no such thing. All he has shown is that the common purport

that is apparently there might be dyscalibrated, that the way they are normally calibrated might not be the only way. This is perhaps interesting, but it falls quite short of establishing the needed point to the effect that there is no shared or common content.

Berkeley's recent proponent Atherton also manifests no clear recognition of the difference between uncalibration and dyscalibration — her blur is codified by her expression 'conceptually unrelated'. This expression is itself ambiguous: the semantic connection between 'conceptual' and 'necessity' (as in the common philosophical assumption that conceptual truths are necessary truths) invites the use of 'conceptually unrelated' upon observing that the connection is contingent, so that even manifolds that have common purport could be 'conceptually unrelated' provided they could be coordinated in more than one way. Thus the existence of inverting prisms would entitle us to say that tangible up and visible up are conceptually unrelated in this sense. But 'conceptually related' also carries connotations of similarity of purport, in the sense that while it might seem felicitous to describe hue and pitch to be 'conceptually unrelated', describing tangible up and visual up (when wearing inverting prisms) as 'conceptually unrelated' in the same sense is far from felicitous. Atherton often slides from one to the other in a way that I think closely mirrors the sliding Berkeley was prone to in his thinking on these issues. A clear example is:

The significant difference between Berkeley's account of distance perception and the account he sets up as its rival is that, according to Berkeley, our ability to perceive distances by sight does not rely on assumptions about *necessary connections* between the immediate and the mediate objects of sight. Instead, that we can recognize by sight how far away something is from us is to be explained as a learned ability to associate visual cues with *conceptually unrelated* tangible ideas of distance. (Atherton, Berkeley's revolution, page 118)

The first italicized phrase, *necessary connections*, is clearly referring to the potential variability of calibrations between, e.g. visible largeness and tangible largeness. Necessary connections would be connections that could not be otherwise than they are, could not be differently calibrated. The second italicized phrase, however, is clearly referring to a total lack of purport, as would obtain between a degree of blurriness and tangible distance. We move from a premise about a lack of necessary connection to a conclusion about lack of common purport.

To be maximally charitable, in the surrounding context of this quote Atherton exhibits awareness of the fact that there is *some sort* of difference here. She recognizes that the conceptual disconnection between visual size and tangible size is a harder sell than that between level of visual blurriness and tangible size. But her continuing discussion doesn't, so far as I can tell, fully recognize that these are not just differences in degrees of resistance to sales pressure, but they are *completely different* ways in which two manifolds can lack correct calibration, and that the difference undercuts the entire negative argument. If the distinction is legitimate, then establishing the possibility of alternate coordinations does not establish the absence of shared purport. But establishing of the possibility of alternate calibrations is all Berkeley does in the case of breadth-height. He has not, therefore, cast any doubt on the idea that vision and touch share a common content. Only by blurring the difference between the two kinds of considerations, and relatedly the difference between uncalibration and dyscalibration, does the inference gain an undeserved look of plausibility.

For a compact example of Berkeley making exactly this slide, see *NTV* 64, page 195):

... it is manifest that as we do not perceive the magnitudes of objects immediately by sight, so neither do we perceive them by the mediation of anything which has a necessary connexion with them. Those ideas that now suggest unto us the various magnitudes of external objects before we touch them, might possibly have suggested no such thing: or they might have signified them in a direct contrary manner: so that the very same ideas,

on the perception whereof we judge an object to be small, might as well have served to make us conclude it great. Those ideas being in their own nature equally fitted to bring into our minds the idea of small or great, or no size at all of outward objects; just as the words of any language are in their own nature indifferent to signify this or that thing or nothing at all.

I conclude that, at least considered on its own, Berkeley's negative argument is unconvincing. Perhaps, though, it can be saved by working in conjunction with his positive account.³⁴

3.5 Option Three

If we use ' π_V ' to stand for the planar counterparts of δ_V , that is, the manifolds Berkeley credits to the proper objects of vision that eventually serve as signs or cues for tangible planar content p_T , we can phrase the crucial question as: What exactly is in π_V ? We know it can't be p_T , since then it would be a common sensible (whether or not it can be calibrated in more than one way). This would seem to leave two options: the content of π_V is genuinely spatial, but this space

³⁴ Atherton discerns another negative argument in *NIV*. The position claims that since spatial content is a function of kineasthetic experience, nothing not in three dimensional space could have *any* spatial content, and so since the 2D visual field, if it existed, would lack depth content and therefore not be identical to three-dimensional physical space, it so would lack height-breadth content. The relevant passage from Atherton (whose reconstruction is much clearer than Berkeley's text on this issue) is:

... the situation of an object is immediately perceived through the kinesthetic experience of reaching out and touching. But we reach out and touch only those things which are located at some distance. Since ... the Molyneux Man would not be able to read any distance information from his first experience of an array of light and color, he would have no reason to take what he is now seeing to be something toward which he could reach out and touch, and so he would not take what he is seeing to be the sort of thing to which situation terms are applicable. What he is aware of visually would seem to him to be like other nonkinesthetic ways of apprehending... (Atherton, *Berkeley's Revolution*, page 153)

But if this is Berkeley's argument, it is surely a bad one. While it is true that in order to *grasp* something that thing must have some determine location in three dimensional egocentric space, it is not true that all kinesthetically laden behavior has this requirement, since grasping is not the only kind of kinesthetically laden behavior. I can *point at* stars or the moon, even though their distance is quite indeterminate. And even something presented in a stereoscope and lacking any determinate depth content, there are different *eye movements* required to *foveate* the different parts of the image — if the depthless image is that of a traffic light, it remains true that I move my eyes *up* to foveate the red light. And I move my eyes in other ways to count the corners of a square, even a square that, because presented in a stereoscope, I could not reach out and touch.

is not identical to, calibrated with, or coordinated with, any dimensions of standard tangible space until after appropriate experience and learning; or the contents of π_V are merely quasi-spatial, and hence not identical to, calibrated with, or coordinated with any dimensions of standard tangible space until after appropriate experience and learning. The first of these is Option Three, which will be explored in this section, the second is Option Four, which will be explored in the next.

Option Three: The proper objects of vision have as elements of their content a genuinely spatial two-dimensional manifold, a 'visual space', and this either exhausts, or at least constitutes the core of, π_V (in other words, modulo complications arising from Berkeley's thesis of minimal visibilia and associated difficulties with continuity, π_V is an R^2 , or at least a finite bounded region of an R^2). But this visual space is neither calibrated with p_T , nor coordinated with it — and so experience is required to establish that visual up corresponds to tangible up, etc. Its axes initially have *no* purported relation to any of the spatial axes of touch. For example, the direction in which the minute hand in a visual image of a clock in canonical orientation that reads 12:15 points (i.e. *right_V*) is not the same direction as the direction along which one moves ones fingers to read Braille (i.e. *right_T*), nor any other tangibly or proprioceptively discernable direction. Similar remarks hold for *straight*, *curved*, etc. Learning sets up the appropriate coordinations and calibrations, and we come to be able to judge of tangible space on the strength of visual space. Quotes where, *apropos* Option Three, Berkeley makes unguarded use of the language of space, movement and extension to describe aspects of the proper objects of vision are quite numerous, but *NTV* 49 and 55 are among the more blatant.

Though Berkeley uses spatial language to describe aspects of the proper objects of vision, there are two exegetically decisive reasons to maintain that Berkeley clearly wanted to distance

himself from Option Three. First he explicitly states that the use of spatial language in this context is metaphorical (*NTV* 94, pages 209-10). Second, the final sustained argument of *NTV* is directed against the idea that there could be a 'purely visual' geometry (*NTV* 121-159, pages 219-235), and the most natural way to read this is as an argument against the idea that the proper objects of vision have anything like a genuine space, even one distinct from tangible space, as aspects of their proprietary content. It will be useful nevertheless to examine a bit more closely the reasons why Berkeley would want to deny Option Three.

How would Option Three work? Let's suppose that in a given case the subject has visual experience to the effect that there are twelve lights arranged in circular pattern around the center of the visual field, eleven of which are white and one is red. This single red light is in, let us say, visual direction A. Before learning, this visual direction, as a direction in the exclusively visual space, has no purported relation to any tangible direction: by itself it appears neither to be (tangible) left, nor right, nor up, nor down, nor ahead, nor behind. But when we execute eye movement α we notice that the red light comes into clear focus at the center of the visual field, and we also know through experience that eye movement α moves our eyes away from our feet. Thus after a period of bimodal experience we come to coordinate visual direction A with tangible up. Similar remarks hold for other directions, of course.

One problem with this proposal is that it requires us to accept that there could be a genuine two-dimensional space, such that directions in that space bear no purported relation to any directions specifiable in the real tangible three dimensional space. It requires us to make sense of the idea that the imagined red light in the above example would be seen to be in a genuinely spatial direction, but not in any direction specifiable in up, down, left, right, ahead or behind terms. If we asked the subject having this experience to give us a tangible clock direction

for the red light — is it at 12 o'clock, or 4 o'clock? — or easier still, ask her to simply point in the direction of the light, she ought to be as stymied as if she were asked to try to hit Middle C with a dart. This ought to be unacceptable for anyone, and it is demonstrably unavailable to Berkeley as a candidate solution.

First as to why it ought to be unacceptable for anyone. It is simply the case that there are exactly three genuinely spatial dimensions, and they are the same three that are apparent through normal perception, including tactile perception. There are not, in addition to these three spatial dimensions, some indeterminate number of other phantom spatial dimensions that we experience, capable as serving as the spatial content of this purported 'visual space'. I can ask you to produce a vivid visual image of a vase of flowers, and then afterwards ask you if the vase was right-side-up, or upside-down, or pointing left or right, and these questions will have determinate answers. But what you cannot do is to form a visual image of a vase of flowers that is oriented in none of these directions, but in one of these other phantom directions instead (go ahead and try).³⁵

Now while I think the proponent of Option Three should be troubled by these observations, I don't see that they are completely compelling. If she is still loyal to Option Three, and is thinking her position through clearly, she will respond that you *can* imagine the vase of flowers being oriented in one of these phantom directions, and in fact you did so the first time you imagined the vase, even though you were not aware of it. In imagining the vase, you put yourself in an as-if state of visual experience in which the vase was oriented in visual direction C

³⁵ It will be objected that it is plausible to suppose that images are spatial and also that images are not in three-dimensional physical space, therefore there are more spatial dimensions than the three found in physical space. But while images are not in physical space, their spatial content is content in terms of egocentric physical space. The imagined vase is imagined to be right side up (not upside down), in front of you (not behind you), and at a certain distance (a distance that is empirically investigable, see Farah et al., "Unilateral occipital lobectomy"). And unless your bout of imagination is one in which you are simultaneously imagining yourself being in some odd orientation (like standing on your head), then if you are imagining the vase of flowers being right side up, you are also imagining it with its base nearer to the Earth's surface than its open end.

(one of the phantom directions). But crucially, because of your long experience associating visual direction C with tangible ‘down’, you also experienced the imaginary vase as being oriented ‘upside-down’. Though I find the proposal that there are additional *spatial* dimensions that are the content of our visual perception other than the familiar three to be an enterprise of vacuous desperation, my so finding it is not an argument. What we need to resolve the issue is not more appeal to increasingly vague intuitions and arguments based on metaphors, but a theory of spatial content, and while providing such a theory is beyond the scope of this paper, I will say a bit about this in Section 4.

But even if Option Three could be saved in this way, Berkeley cannot make use of it (for two distinct reasons), nor would he want to. First, if π_V is an R^2 , then relations hold among its elements that can be described in geometric terms. The above description of the twelve points of light even admitted that there was a specifiable direction from the center of the visual field to the lone red light, and if there are directions, there are lines, and if there are lines in a space of more than one dimension, then there are angles, and so forth. While the suggestion that such a direction is not along any of the usual spatial directions might be barely coherent, the suggestion that there are *no* directions or lines in this visual space and yet it is a genuine space for all that is surely not. And these directions and lines would be the sort of abstract ideas that Berkeley must distance himself from — not because they are the same directions, since they are not, but because we would have a content (e.g. *line*, *angle*) that could be delivered via more than one modality. I take it that Berkeley recognized this, and his repudiation of the idea of a purely visual geometry (which takes up the last 38 sections or so of *NTV*) was his way of trying to make clear that although expressive limitations of the English language were forcing him to use spatial

expressions to describe the proper objects of vision ('visual extension', 'visual motion' etc.), this language was merely metaphorical. He was in fact embracing Option Four.

The second reason Berkeley cannot help himself to this suggestion is that it makes appeal to entities to which we have no direct transparent mental access, and this goes against what I will call Berkeley's *psychological transparency thesis*: the entities appealed to in a psychological explanation of some phenomenon must be capable of being known or consciously reflected upon by the subject. Berkeley *continually* criticizes his opponents for appealing to entities that the subject has no mental access to. For example, Berkeley claims that:

Moreover it is evident that no idea which is not itself perceived can be the means of perceiving any other idea. If I do not perceive the redness or paleness of a man's face themselves, it is impossible I should perceive by them the passions which are in his mind. (*NTV* 10, page 173)

And then concludes from this:

But those lines and angles, by means whereof some men pretend to explain the perception of distance, are themselves not at all perceived, nor are they in truth ever thought of by those unskillful in optics. I appeal to anyone's experience whether upon sight of an object he computes its distance by the bigness of the angle made by the meeting of the two optic axes? Or whether he ever thinks of the greater or lesser divergency of the rays, which arrive from any point to his pupil? Everyone is himself the best judge of what he perceives, and what not. in vain shall any man tell me various ideas of distance, so long as I myself am conscious of no such thing. (*NTV* 12, page 173)

But a 'visual space' spanning phantom directions, posited only to satisfy the demands of Option Three but being quite psychologically inaccessible is exactly the sort of thing that *NTV* 10 claims is unacceptable, and thus Option Three would seem to be in exactly the same boat as the theories of Descartes and Malebranche in this regard. Option Three should be an object of derision from the standpoint of *NTV*, not an interpretation of it.

3.6 Option Four

Option Four: The proper objects of vision have as elements of their content a number of quasi-spatial manifolds that comprise π_V . As quasi-spatial manifolds, they are neither calibrated with p_T nor coordinated with it. In a manner presumably parallel to the relation between δ_V and d_T , learning sets up the appropriate coordinations and calibrations, and we come to be able to judge of tangible space on the strength of vision. *Apropos* Option Four, Berkeley at places claims that the ‘spatiality’ of visual appearances are merely quasi-spatial, and the application of spatial expressions in their description to be metaphorical, as with high and low pitches (see e.g. *NTV* 94, pages 209-210).

Option Four would seem, all things considered, to be the one Berkeley should favor (in case this is not obvious, reasons for this will be discussed shortly). In order for his positive account of visual situation to parallel the account of visual depth, the story would have to be that, while breadth and height purport is not an aspect of the proper objects of vision, these proper objects do have as aspects one or more quasi-spatial dimensions of variation that carry information about these (tangible) spatial dimensions by being correlated with them, and these correlations become manifest in multimodal exploratory learning, the result of which is the setting up of appropriate coordinations and calibrations. But what are these quasi-spatial dimensions of variation and the correlations? Berkeley’s official story exploits eye movements and the fact that visual objects become clearer or more distinct when foveated:

...when upon turning his head or eyes up and down to the right and left he shall observe the visible objects to change, and shall also attain to know that they are called by the same names, and connected with the objects perceived by touch; then indeed he will come to speak of them and their situation, in the same terms that he has been used to apply to tangible things; and those that he perceives by turning up his eyes he will call upper, and those that by turning down his eyes he will call lower. (*NTV* 97, page 211)

And this seems to me the true reason why he should think those objects uppermost that are painted on the lower part of his eye: for by turning the eye up they shall be distinctly seen; as likewise those that are painted on the highest part of the eye shall be distinctly seen by turning the eye down, and are for that reason esteemed lowest; for we have shewn that to the immediate objects of sight considered in themselves, he would not attribute the terms *high* and *low*. It must therefore be on account of some circumstances which are observed to attend them: and these, it is plain, are the actions of turning the eye up and down, which suggest a very obvious reason why the mind should denominate the objects of sight accordingly high or low. (*NTV* 98, page 211)

There is a subtle but fatal problem here which can be brought into cleaner relief in Atherton's reconstruction. First, she reiterates that there must be something (what I have called a quasi-spatial dimension of variation, she calls it 'some way the visual array looks') present that is correlated with the two dimensions of breadth and height:

... situation is a mediate object of sight, inasmuch as what we see has come to mean the presence of specific, tangible locations. Since we are able to reach for what we see, we have learned how to recognize what tangible locations look like, or, more precisely, we have learned how to correlate the way a particular movement of reaching in some direction feels with some way the visual array looks ... (Atherton, *Berkeley's Revolution*, pages 154-155)

So far so good. Leaving aside for now issues of what 'visual array' might mean (this issue will figure in what follows), this is surely correct as an interpretation of Berkeley. But the explanation continues:

Berkeley's proposal is that it is the experience of the way the muscles of the eye feel when we turn our eyes up that suggests we are seeing 'up' and hence can serve as cues for when we must reach up ... In order for a cue to contingently suggest some movement or

other, it need do no more than covary with the movement in question. The experience of feeling the eyes move up must reliably covary with the feeling of reaching an object by moving up. ...

Thus we are not correlating what is on the bottom of the retina with something that is visually up, for nothing can be visually up. We apply the term 'up' merely analogically to what we see when we literally turn our eyes up, which is also when we bring the bottom of the retina into focus. (Atherton, *Berkeley's Revolution*, pages 155-6.)

The same explanation presumably holds for the left-right dimension as well as the up-down dimension, and hence we have a candidate solution for how the visual field becomes imbued with apparent two-dimensional content (in addition to the dimension of depth that was taken care of earlier). The problem is that it won't work. The candidate solution is this: there are times when something is seen, but indistinctly ('out of focus' in Atherton's reconstruction), but when we move our eyes in a certain direction (and we can feel this direction courtesy of our proprioceptive input from the relevant eye muscles) this entity becomes more clear — we have foveated the object. Since we don't want to load the dice, let's call this felt eye movement that brings the object into clear focus Movement α . We note, through experience, that objects that get brought into clear focus through Movement α are such that we can touch them by moving our hands (tangible) upward. We then come to judge that objects that become more distinct upon execution of Movement α are (tangible) *up*. Similar phenomena associated with Movements β , γ and δ give us visual access to objects as being *right*, *down* and *left*.

The tires on this account go flat when we realize that we have the wherewithal to make directional assessments on the basis of vision *without moving our eyes* (or head) — and indeed such directional assessments are part of the explanation for our ability to move our eyes or head in the right direction in order to orient correctly with respect to the object in the first place. With my eyes straight ahead and *not moving* I can tell that the red light is above the blue light, and the

green light is to the left of the yellow light; and that the flashing light is up and to the right and the big blue circle is down and to the left. I can even see that an object is quickly approaching from my right and hence reflexively flinch to the left. This is the case even when all the visually apparent objects are equally distant from my fovea, and hence all equally indistinct (out of focus). But on the Berkeley/Atherton theory, such eye-stationary directional assessments should not be possible.

That argument may have been a bit quick, so let's take it again more slowly by distinguishing two different versions of Option 4. The Berkeley/Atherton proposal that I have just outlined is what I will call a *minimal* version of Option 4 (though it will be apparent that in fact both Berkeley and Atherton flip flop between the minimal and robust versions I will describe). It is minimal because it credits π_V with fewer manifolds than the robust version. On the minimal version, the contents of π_V are lights, colors, muscular cues, and degrees of being out of focus (the list is supplied by Atherton, *Berkeley's Revolution*, page 185, and as far as I can tell is exactly the list Berkeley makes use of in *NTV*).

To see how directional assessments work on this proposal consider the following thought experiment. Suppose that I am in a completely dark environment in which various objects are dispersed. Each such object emits an odor, and the strength of this odor is a function of distance to the object. Odor intensity is playing a role here analogous to degree of out-of-focusness. So when I smell an object I can, by assessing how intense the odor is, determine roughly how far from this object I am. But I cannot determine just on the basis of the odor's intensity the direction in which this object lies. In order to do that, I must take at least a few steps in different directions, and take note of how the intensity fluctuates as I move. If I move in all four directions, and the intensity increases upon executing movements α and β , but decreases upon

executing movements γ and δ , I can infer that the object lies somewhere between directions A and B. And so forth for other possibilities. But what I would *not* be able to do is to smell 4 different odors, all of them at, say, half maximum intensity, and determine their directions without moving. They might all be in the same direction, they might all lie in completely different directions. All I can tell is that they are all the same distance from where I am. But until I start moving and assessing how the different odors' intensity changes, judgments about their relative locations or directions is not possible. For similar reasons, I could not smell a single odor and determine the direction in which the associated object lies before I move.

Now to what I will call the *robust proposal*. Let's stay in the smell world, and keep the assumption that objects emit odors whose intensities fall off as a constant function of distance. But in addition to discernible intensity we will assume that there are two other discernible features of each odor that are a function of direction. Let's call these odor features Feature X and Feature Y (perhaps they are mintiness and sweetness). Initially, I smell some odor that is, say at one half maximum intensity, and has a great deal of Feature X, and an average amount of Feature Y. I then start moving around, and note that as I execute movement α , the odor gets more intense. In fact, after a while it becomes apparent that all odors that have a great deal of Feature X and an average amount of Feature Y can all be made more intense by executing movement α . And I similarly learn, through exploration, that odors with an average amount of Feature X and a great deal of Feature Y can be made more intense by executing movement β . And so forth. These features of the odors have no spatial purport. But after learning through exploration as just described, these features can come to be taken as signs for distance and direction. After learning, I might smell an odor that is at three-quarters maximum intensity, and has very little of features X and Y, and thus come to judge, without moving, that the object is a

certain distance in a direction half way between the directions that of movements γ and δ would take me *if* they were executed.

Both the minimal and the robust proposal start out with olfactory manifolds none of whose dimensions have spatial purport. And both claim that the subject comes to extract spatial information from these quasi-spatial odor manifolds by learning correlations that are made manifest during movement and exploration. Where they differ is on the richness of the dimensions of the initial manifolds, and relatedly whether or not movement and exploration continue to be needed after learning. The robust proposal claims that there are quasi-spatial dimensions of variation *in the odors* apart from movement that can be coordinated and correctly calibrated with direction, and once sufficient movement and exploration have taken place to learn these coordinations/calibrations, the manifolds themselves can be taken as cues for directional judgments even without concurrent movement. The minimal proposal does not recognize these additional dimensions of variation, and as a result, recognizes nothing in the odors themselves that could, without concurrent movement, be learned to correlate with direction. Rather, the minimal proposal requires that each directional assessment is made on the basis of noticing how the one manifold (intensity) changes over time as a function of movement. And each new stimulus must have its direction established in this way.³⁶

Clearly, the *official* Berkeley/Atherton account is a version of the minimal proposal. In the passage quoted earlier in this section, Atherton explicitly tells us that “In order for a cue to contingently suggest some movement or other, it need do no more than covary with the movement in question. The experience of feeling the eyes move up must reliably covary with the feeling of reaching an object by moving up.” The cue here is one that includes movement, and

³⁶ I believe that this distinction is the one discussed by Gareth Evans in his paper “Molyneux's Question” by means of his distinction between 'simultaneous' and 'sequential' representations. Evans' paper is as underappreciated as it is unclear.

what it is a cue for is tangible direction. By invoking only on a single scalar manifold — degree of out-of-focusness — the account needs to rely on concurrent movements and explorations in order to establish the breadth-height direction of each stimulus. But we can make such assessments without eye or head movement. Therefore the minimal proposal can't be correct as an account of visual breadth-height content.

So though we can see now that the official Berkeley/Atherton line won't work, can this line be modified in such a way as to render it adequate? Clearly some sort of robust version of Option Four is required, and if one could be formulated, it would still be in a broadly Berkeleyian spirit insofar as it would recognize that bimodal exploration is required to gain fully normal spatial vision — but is only required during the learning stage. The trick is to articulate it in such a way that it is not just Option Three (which of course can support a version of the robust proposal, using directions and locations in 'visual space' as the discriminable features that go beyond movement and degree of out-of-focusness — see the description of Option Three at the beginning of Section 3.5) plus mumbles to the effect that 'spatial' is metaphorical. The obvious move would be to try to identify some other *quasi*-spatial dimensions of variation that could serve. But it is not clear to what we might appeal. Given the restrictions of the transparency thesis, these manifolds have to be things we notice, or at least can take notice of. And once we eliminate spatial location, we are left with only qualities that, so it seems, can all apply equally well to objects regardless of location. Once the out-of-focusness/eye-muscle-movement gambit has been discredited, there are simply no candidates other than objects' apparent locations. But this is what we are trying to explain.

3.7 Discussion

Let's take stock. I have singled out Berkeley's account of the height-breadth component of visual perception, and discussed two aspects of it. First, a negative argument to the effect that the proper objects of vision lack proprietary height-breadth planar content; and a positive argument that attempts to explain how visual experience becomes imbued with apparent height-breadth planar content.

The negative argument was unsuccessful, I argued, because it failed to notice the difference between uncalibration and dyscalibration. Failure to notice this difference makes the negative argument for the height-breadth aspect of visual spatial content appear to be parallel to the negative argument for the depth aspect of visual spatial content, when in fact it is not parallel at all. In the case of depth, the obvious lack of purport between, e.g. level of eye muscle strain and distance to a seen object, was offered as a reason to conclude that the coordinations that result are contingent and must be learned. The result is that a *prima facie* plausible case can be made to the effect that the proper objects of vision have no proprietary depth content. But in the case of breadth-height, no such conclusion established with any plausibility. The only examples are dyscalibrations that obtain between manifolds that nevertheless appear to have shared purport. Interesting though it may be, the possibility of dyscalibrations does not establish, but indeed relies upon, shared purport between the dyscalibrated manifolds. The illusion of a coherent argument is maintained only by sliding from the fact that the relevant *prima facie* calibrations could be different (what looks visual-up might have been tangible-down), to the ill-gotten conclusion that there is no shared content (what looks visual-up shares no content with any tangible direction).

The positive account also failed because of an illicit slide. Four options concerning the relation between the spatiality of vision and touch have been articulated. Berkeley's avoidance of abstract ideas is sufficient reason to rule Options One and Two out as positions he might bring himself to agree with. Option Three is also ruled out for Berkeley, appealing as it does to both 'visual' and 'tangible' spaces, since even in this terminologically segregated sense they are enough to trigger his abstract idea gag reflex. Furthermore, the dimensions of such a visual space would appear to be inaccessible to the subject as such, and hence run afoul of the transparency thesis. This leaves Option Four.

The problem is that it does not appear as though a minimal version of Option Four is workable, as instanced by the inadequacy of the Atherton/Berkeley proposal discussed in Section 3.6, nor does it appear that a robust version that is otherwise kosher by Berkeleian standards (i.e. not a veiled version of Option 3) is constructible. The search for dimensions of variation sufficient to support a robust version of Option Four, and sufficient to account for the full range of coordinations and calibrations established between vision and touch — a search which as we saw cannot end with degrees of blurriness and sensations of eye or head movements, or other introspectively available materials — invites a return to crediting visual objects with dimensions of variation described in spatial terms (note, e.g., Atherton's casual invocation of the expression 'array' in the quote from *Berkeley's Revolution*, page 154-5 above). But then we are back at Option Three, and its problems.

I don't think there is any escape from this oscillation for Berkeley. When articulating his positive proposal in a context where the goal is to show how it appeals only to processes and states that are immune to the kinds of criticism he heaps on competing accounts, degrees of focus and muscular cues are the descriptions of choice for π_V . But when the context is one in which the

goal is to make it sound as though the proposal has a plausible chance of doing justice to the facts of human vision, appeal to 'visual space' (Berkeley), or 'visual grid' (Atherton) is close at hand along with metaphor disclaimers. But the disclaimers can't be honored. The illusion of there being a coherent and adequate theory on offer is maintained only by poorly-placed faith in the innocence of such context-driven slides.

4. Conclusion

Given the extent to which I have been critical of Berkeley's account of the spatiality of vision, it might seem that I think it is entirely a non-starter. This is not correct. I will end this critical essay with a gesture towards what I take to be the salvageable core of Berkeley's account — a core that has been entirely vindicated by recent theorizing on perception. I will bring attention to this core by stating what Berkeley would have to change about his overall position in order to yield a workable, and more or less correct, theory.

First, the transparency thesis and the moratorium on abstract ideas must be abandoned — both are highly questionable by contemporary lights anyway. With these impediments gone, a robust version of Option Four is at hand. The account would be that as part of the process that explains the spatial content of our visual experience s_V , (since this account will treat all three dimensions similarly, there is no need for a division between visual depth d_V and visual planar content p_V), there is a set of manifolds σ_V that are exclusively manifest at the sub-personal level — they are features discerned by the perceptual system that are not introspectively accessible by the subject. Because these manifolds are not psychologically accessible, they are not, on their

own, contentful at all. They carry neither spatial nor (not being qualia at all) qualitative quasi-spatial content.

Schematically, we can call this process — by which physiological sensitivity to σ_V during perception gives rise to psychological receptivity of spatial perceptual experience in the form of s_V — a function f that maps σ_V to s_V : $s_V = f(\sigma_V)$. This function serves much the same role as a coordination does in Berkeley's official account. A coordination is a mapping from one manifold to another that serves to imbue the target manifold with content properly carried by the source manifold, e.g. to imbue ideas of eye muscle strain with tangible depth content. The difference is that this function, unlike a coordination, can take as *input* a manifold that is not psychologically accessible and produce as output one that is.

Next, giving up the moratorium on abstract ideas open the possibility that the *output* of this function is not restricted to some set of manifolds that are given directly through modality-specific experience, but could be a multimodal or supra-modal brand of content. And therefore there is no need to single out one of the sensory modalities as the proprietary provider of spatial content as Berkeley did with touch-kinaesthesia-proprioception. In particular, it means that the same function might account for the spatial content of all of our modalities, schematically $s_A = f(\sigma_M)$, where σ_M is the set of sub-personal manifolds employed by modality M that come to undergird the spatial content of personal-level perceptual experience, and s_A is amodal spatial content (perhaps supra-modal would be a better description).³⁷

³⁷ Those familiar with Wilfrid Sellars' account of Kantian sensibility (Wilfrid Sellars, *Science and Metaphysics: variations on Kantian Themes*, (London: Routledge & Kegan Paul, 1968), Chapter 1) will see the image of that account here. In accord with Sellars' project of morphing Kant's distinction between appearances and things-in-themselves into a distinction between the manifest and scientific images, the core of his discussion is a distinction between, on the one hand, spatial features of the objects of perceptual experience, and on the other hand, analogous σ -features that are features of how our perceptual apparatus is affected by the 'transcendental' (scientifically described) objects corresponding to the objects of experience. The 'spatial' features are aspects of the manifest image, and the σ -features are features of the 'impressions' that our perceptual systems, as scientifically described, pick up on when causally affected by those objects. For Sellars, the function f is a learned 'conceptual response'.

It might seem that such a proposal has shed its last vestiges of Berkeleianism, since it not only disposes of the transparency thesis, but it recognizes a supra-modal spatial content that applies equally fundamentally to many modalities of experience. These are tough pills for a Berkeleian to swallow. But whether or not the last vestiges of Berkeleianism are shed depends on details of the function f . Recent work in both philosophy and cognitive neuroscience has suggested that spatial perception is constitutively tied to sensorimotor processes.³⁸ In particular, that a sensory input gets imbued with spatial significance insofar as, and only insofar as, that input is taken up by processes that coordinate the subject's sensorimotor behavior. In other words, f looks in fact to be a process by which a sensory state, understood as some sub-personal quasi-spatial manifold, is taken up by processes whose job is to cue and guide the organism's overt behavior. So while the link between spatial perception and motor action is not quite as direct as Berkeley thought — spatial perception does not reduce to touch and motor action, and f is not just the coordination of visual manifolds with touch and kinaesthesia — the connection appears to be there, just at a deeper, more subtle level. It is his recognition of this connection between spatial content of perception and sensorimotor control, even if he miswielded this insight in some ways, that is the lasting contribution of Berkeley's work on vision.

Acknowledgments:

This paper benefited from two separate rounds of discussion at the UC San Diego Philosophy Department's History of Philosophy Roundtable, and I am indebted to the participants of that group for invaluable feedback. An anonymous referee also provided extremely useful comments and suggestions, for which I am grateful.

³⁸ For details and references, see Rick Grush, "Skill and spatial content," *Electronic Journal of Analytic Philosophy* 6 (1998); Rick Grush, "Self, world and space: the meaning and mechanisms of ego- and allocentric spatial representation," *Brain and Mind*, 1 (2000): 59-92; Rick Grush, *The Machinery of Mindedness*, (in preparation).