journal of interdisciplinary music studies
season 2020, volume 10, art. #20101201, pp. 01-16

Embodiment in the history of depth perception

Don Oxtoby

Department of Philosophy, Rice University

Background in philosophy. The philosophy of perception has traditionally focused on visual perception. In recent decades, philosophical thought about auditory perception has gained momentum, seeking to supplement visual models of perception (O'Callaghan, 2011, pp. 143-160). A key difference between vision and audition concerns spatial acuity. Vision is usually thought to present the location, especially depth, of what is seen with greater acuity than audition presents the location of what one hears (Culling & Akeroyd, 2010, pp. 128-130). Rethinking differences between visual and auditory depth perception offers fertile ground for thought about spatial perception in general. Rationalism and empiricism are two philosophical traditions which offer competing explanations of depth perception, with rationalists like Descartes arguing for "geometrical views", and empiricists like Berkeley arguing for views based on the role of "experience".

Background in psychoacoustics. Being surrounded by sound is popularly considered a desirable quality for audiences of musical performances and film screenings, evident in the demand for specialized sound recording and reproduction technology, marketed in the U.S. circa 1978 as "Surround Sound". Some novel circumstances of auditory depth perception arise from surround sound arrangements and forms of musical performance, such as "quadraphonic musical performance", whereby four performance stages are arranged to surround an audience. Theorising depth perception in this context relates to the ability of a creature to orient its body upon hearing a sound, and locate the sound's source using vision and bodily movement, as when perceiving the width of a sound's source (Whitmer et al., 2013), or "feeling surrounded by sound".

Aims. This paper offers support for the claim that empiricist views of depth perception in the history of philosophy provide an earlier historical precursor for embodied cognitive science than has been appreciated in the literature.

Main contribution. Forms of experience introduced in Berkeley's 1709 *New Theory of Vision* provide support for the claim that historical empiricist views offer a precursor to issues addressed in contemporary embodied cognitive science. Surround sound arrangements and quadraphonic music performance provide specific examples which connect empiricist views of depth perception to embodied cognitive science, psychoacoustics, and musical performance.

Implications. Empiricist views of depth perception isolate forms of experience with implications for embodied cognitive science, psychoacoustics, and musical performance, including experience of perception in multiple modalities, and experience of bodily movement. Continuity between empiricism and embodied cognitive science suggests that current attention to such forms of experience may be fruitful for our contemporary understanding of spatial perception and for future research. This paper also discusses implications of embodied views of auditory depth perception for spatial aspects of aesthetic experience and musical performance, as in the notion of "feeling surrounded by sound".

Keywords: depth perception, embodiment, empiricism, Berkeley

•*Correspondence:* Don Oxtoby, Department of Philosophy, Rice University, Houston, Texas; E-mail: oxtoby@rice.edu

•*Received:* 31 January 2020*; Revised:*14 July 2020*; Accepted:* 16 September 2020

•*Available online:* 29 January 2021

•doi: 10.25364/24.10:2020.1.1

This paper examines the extent to which eighteenth-century empiricist views of depth perception foreshadowed contemporary developments in embodied cognitive science.[1] This position is consistent with, for example, Noë (2004, pp. 96-97, 100) and Schroeder (2002, pp. 87-92), which explicitly connect elements of empiricist views - such as Berkeley's - to current research programmes. Against Cartesian rationalism, eighteenth-century empiricists maintained that there should be a greater explanatory role for what they referred to as "experience" – and that certain forms of experience arising concurrently with locomotion and bodily movement are implicated in confirmation of perception across senses. This led empiricists to push for explanations of auditory depth perception to drive theories of visual depth perception, rather than vice versa. The particular explanatory role given to such experience, especially during perceptual learning, connects empiricist views with contemporary studies on the role of multisensory integration in embodied cognitive science.

The first part of this paper discusses historical issues relevant for explaining the contribution of empiricist views to the study of depth perception, introducing the weight and relevance accorded to experience in such accounts. Part two explains what is meant by "experience" in empiricist views of depth perception, and which forms of experience are salient for embodied cognitive science. Part three explains the connection between empiricist views of depth perception and embodied cognitive science, where examples of confirmation across senses and perceptual learning explain the role of experience in visual and auditory depth perception. The role of bodily movement is discussed with reference to visual depth perception in Eleanor Gibson's (Gibson & Walk, 1960) "visual cliff" experiments, and compared with auditory depth perception in "surround sound" and "quadraphonic" musical performance. Both empiricist and embodied views argue that a creature's ability to orient its body to the source of a sound plays a key explanatory role in auditory depth perception, in both ecological and specifically musical settings.

Depth perception and the empiricist tradition

Historically, philosophy of perception focuses on visual perception. A key driver in this trend is a concern for solving problems about what exactly perceiving subjects see, and how vision secures knowledge of what is seen. For example, prominent philosophers in both ancient Eastern and Western traditions typically explore visual perception by distinguishing perception or "sensation" from judgement (Hatfield, 2001). By the early eighteenth century, competition between two broadly different approaches to problems of visual perception became an important topic of debate: according to rationalists, problems of what exactly is seen were to be solved principally by understanding how the eyes, optic nerves, and brain rely upon

1 "Cognition is embodied when it is deeply dependent upon features of the physical body of an agent, that is, when aspects of the agent's body beyond the brain play a significant causal or physically constitutive role in cognitive processing . . . Embodied cognitive science encompasses a loose-knit family of research programs in the cognitive sciences that often share a commitment to critiquing and even replacing traditional approaches to cognition and cognitive processing" (Wilson & Foglia, 2015, para 1).

geometrical principles to secure knowledge of the size, shape, direction, and distance of visible objects or properties. Such views were described variously as “rationalist views”, “geometrical views”, and (by their opponents) “intellectualist views” (Schwartz, 1994, pp. 84-86). Such classification reflected the epistemological reliance of these views upon principles independent of conscious experience – namely, principles of geometry, like the law of sines, for triangulating depth in visual perception. According to so-called empiricists, on the other hand, the same puzzles and more could be resolved by understanding how various forms of experience play a role in securing knowledge of the size, shape, and spatial properties of what one sees (Atherton, 1990, pp. 77-86). Empiricist views were considered radically different because they originate partly in criticism of a basic belief held by rationalists, namely that experience itself can play a positive role in securing perceptual knowledge, against the rationalist belief that principles which are independent of experience are the correct ones for securing perceptual knowledge.

Following precedents in medieval philosophy, both rationalist and empiricist views of the eighteenth century explained perception using inference as a model (Hatfield, 2015). Inference was usually understood in accord with logical inference, whereby a statement of conclusion can be derived from a set of premises according to certain rules. Explaining perception using such a model might typically treat the knowledge of what one sees as being the conclusion of an inference performed by one's visual system. This inference would take sensory stimuli as input (premise), issuing perceptual knowledge output (conclusions) in terms of properties like size, shape, and distance. Modelled as an inferential process, the activity of the visual system could be described as unconscious (Hatfield, 2002); this much averted the attribution to subjects of conscious inferences through reflective thought – a charge historically levelled at rationalists for “over intellectualizing” perception (Schwartz, 1994, pp. 84-86).

Rationalist theories of depth perception, like those of Descartes or Kepler (Hatfield, 2002), applied the model of perception as unconscious inference to the classic “inverse problem” of visual depth perception. This problem was construed during the early modern period as one that any serious theory of visual perception ought to be able to solve. In visual depth perception, the inverse problem raises the issue of how depth can be accurately perceived on the basis of retinal stimulation associated with a flat two-dimensional image projected on the retina that is always consistent with multiple possible configurations of the environment. The rationalist formulation, then, would ask, “How does the visual system use unconscious inference from retinal stimulation to draw conclusions about depth?”

Among proponents of rationalist views, Descartes’ was one of the most widely known and criticized by empiricists (Atherton, 1990, pp. 54-57). In order to secure knowledge of the distance of visible objects, Descartes combined the model of perception as unconscious inference with the rationalist view that the visual system relies upon geometrical principles that are independent of conscious experience.[2] In

2 The term *visual system* can be ambiguous between merely sensing, and sensing with an “intellectual contribution”. For discussion of Descartes, Berkeley, and their interlocutors on the “visual system” see

the *Dioptrics*, Descartes argues that depth perception is carried out by way of “natural geometry” occurring in the visual system (Descartes, 1965, p. 106). While there is some debate about how to interpret Descartes' notion of natural geometry, the problem presented by empiricists is common to all major interpretations (Hatfield, 2015). An understanding of what Descartes meant by “natural geometry” helps to reveal exactly how empiricists distinguished their approach from rationalists’. In the *Dioptrics*, Descartes explains depth perception according to geometrical principles related to triangulation, where the distance of an unknown point, representing a perceived object or property, is perceived by way of geometrical calculation from the location of two known points – those situated at the location of a creature's two eyes (in binocular vision) (Atherton, 1990, pp. 26-29). Descartes uses the analogy of a person who is blind reaching forward with a pair of crossed sticks:

> . . . we know distance by the relation of the eyes to one another. For just as our blind man, holding the two sticks, AE, CE, of whose length I am assuming that he is ignorant, and knowing only the interval in between his two hands A and C, and the size of the angles ACE, CAE, can from that, as if by a natural geometry, know the location of the point E . . . (Descartes, 1965, p. 106)

Figure 1. “Natural geometry,” from Descartes' *Dioptrics* (Descartes, 1965, p. 105). From Bellis, 2016.[3]

According to rationalist views like Descartes', it is the ability of a visual system to triangulate which explains depth perception. Triangulation in the visual system has three salient features: (i) a point which lies at a distance unknown to the subject, (ii) two points whose location are known, located at each of the subject's eyes, and (iii) the visual system's use of geometrical rules in order to produce an “idea” or conclusion about the distance of the initial point. Relevant geometrical features include angles formed by lines relating the location of the eyes to the distance of the point where the thing seen is located. Essential to Descartes’ rationalist claim is the

Atherton (1990, pp. 50-52); Schwartz (1994, pp. 84-86; pp.104-109); Hatfield, (2002, pp. 14-26; 2015, p. 129).

3 For discussion of Descartes' analogy of the blind man illustrated in Figure 1 see, Noë (2004, p. 1; Chemero (2015).

premise that once the visual system's form of unconscious inference is understood, all that is essential to depth perception is understood. No explanatory role of similar importance is given to “experience”, in the sense in which empiricists would argue for.

One of the most influential criticisms of rationalist views during the early modern period appeared in Berkeley's 1709 *New Theory of Vision*, henceforth abbreviated *NTV*. *NTV* features prominently an attack on the claim that explaining depth perception is essentially a matter of explaining the use of geometrical principles by the visual system to triangulate depth. Berkeley instead draws readers' attention to the role of what he calls “experience” and certain forms of perceptual learning that make use of experience, referring to geometrical views like Descartes' as “overly intellectual” (Schwartz, 1994, pp. 84-86). While not necessarily rejecting the use of inference to characterize perception, Berkeley argues that rationalist accounts leave out the role of experience in explaining visual perception generally, and especially depth perception, and certain visual illusions like the moon illusion (Berkeley, 2002, sect LXVII-LXXVIII).

In *NTV*, Berkeley aims to explain the role of experience in what he differentiates along with his contemporaries as two kinds of depth perception; depth perception concerning nearby “proximate” distances, and depth perception concerning further away “remote” distances.[4] First, Berkeley directs readers' attention to visual perception of remote distances:

> I find it also to be generally acknowledged that our estimate of the distance of considerably remote objects is an act of judgement based on experience rather than of sense. When I perceive many intermediate objects – houses, fields, rivers, and the like – which I have experienced to take up considerable space, this leads me to judge or conclude that the object I see beyond them is at a great distance. (Berkeley, 2002, sect III)

Upon seeing objects at a remote distance, we are said to “judge or conclude based on experience” rather than “sense” that the object seen lies “at a great distance”. In his example, Berkeley imagines a case where one sees many intermediate objects between oneself and the distant object one is looking at (“houses, fields, and rivers”). These intermediate objects are things that one has “experienced to take up considerable space” in the past. Berkeley then arrives at a question: is one's visual perception of remote distance “an act of judgement based on experience” or “[an act of] sense”? The answer he endorses, and which is also held by his opponents (like Descartes) when it comes to remote distances, is that impressions of remote distance are explained by past experience of learning the location of familiar landmarks, like the intermediate objects one has previously seen, and judging the distance of a remote object on the basis of this past experience (Hatfield, 2015, p. 140). Whereas visual perception of *remote* distance is held to explicitly involve conscious judgement

4 For discussion of remote vs. proximate distance perception in early modern philosophy see Hatfield (2015, p. 140).

according to both rationalists and empiricists, Berkeley emphasizes the role of past experience over geometrical reasoning, and eventually uses this point of agreement with rationalists to put pressure on rationalist views of *proximate* depth perception.

In order for empiricist views of depth perception to truly put pressure on rationalist views, it is necessary for experience to play a role in explaining not just remote but proximate depth perception as well. Whereas the role of experience as an input to judgements of remote distance was relatively uncontroversial, finding a significant explanatory role for experience in proximate depth perception broke ground for empiricist views. Finding such a role may have seemed unlikely in one respect, since the role of experience in remote depth perception is taken to be as an input to judgement, and proximate depth perception is not understood to depend upon judgement as remote depth perception is. What role does that leave for experience in proximate depth perception?

Berkeley takes proximate depth perception to be based on muscular sensations associated with certain forms of visual focus, as well as experiences of successfully identifying the distance of what one sees in perceptual exploration (Berkeley, 2002, sect XVI-XX). The muscular sensations of visual focus Berkeley is interested in are indicated by his example of focusing on an object which one gradually moves closer to one's face, and feeling the strain of one's eyes crossing as the visible surface of the object decreases in distance. The more strained one's focus, the less distance one has an impression of between oneself and the thing seen.[5] Moreover, according to Berkeley, upon successfully identifying the distance of something one sees with input from further sensory exploration, like reaching out, walking up, or looking around, one learns and gets accustomed to feelings of visual focus that are correlated with the actual distance of what one sees.[6] Forms of experience rooted in the phenomenon of visual focus during sensory exploration and learning, especially with correlated bodily movement, are at the heart of what connects empiricist views like Berkeley's to embodied cognitive science. Later in *NTV*, Berkeley extends this model to auditory depth perception, discussed further below (Berkeley, 2002, sect XLVI).

“Experience” in empiricism and embodied cognitive science

The role of experience in *NTV*'s discussion of depth perception provides a way of connecting *NTV*'s empiricism about experience to embodied cognitive science. *NTV* introduces one form of experience to explain remote depth perception, and another to

---

5 Interestingly, both Berkeley and Malebranche make use of the rotation of the eyes in binocular depth perception, yet whereas Malebranche (a rationalist) focuses on the change of angle during eye movement as a means for computing the distance of what we see, Berkeley (an empiricist) focuses on muscular sensations as a form of experience associated with the rotation and focusing of the eyes (like accommodation as a cue for visual depth) (Atherton, 1990, pp. 82-84). See also Malebranche (1997, book I, chapter 9, section 3). Copenhaver (2010) discusses how such forms of experience are relevant for Reid’s theory of acquired perception, which bears greater similarity to Berkeley’s (especially pp. 289-303).

6 For contemporary discussion of philosophical thought about the role of experience in perception related to Berkeley's see Mulligan (1995) and Siewert (2015).

explain proximate depth perception. While these forms of experience are distinguished for the purpose of explaining the role of experience in proximate and remote depth perception, they are not mutually exclusive. The explanatory role of experience in proximate depth perception is understood to compliment the explanatory role of experience in remote depth perception, as well as complimenting the role of long-term memory and reflective thought in depth perception and judgments of depth. The latter is seen in remarks about the role of perceptual learning in coming to be able to accurately judge the depth of familiar landmarks. The connection between empiricist views of experience and embodied cognitive science lies in experience of bodily movement with multiple modes of sensory input. These features are revealed in the experience of remote and proximate depth perception, on Berkeley's view, and in the way experience is supposed to figure in perceptual learning.

NTV draws readers' attention to what is experienced when exploring and confirming the distance of what one sees: experience of colour, shape, and size vary with experiences of visual focus and input from other senses (especially touch). These variations are then associated over time with how things visually appear when they are successfully identified at certain distances during perceptual exploration (Atherton, 1990, pp. 84-86). Inputs to depth perception from experience create expectations about how input from various senses should change over time given that something lies at a certain distance.[7] In proximate depth perception, input from experience is primarily limited to features produced by the relationship between vision and bodily movement that are significant for stereopsis, whereas remote depth perception further draws upon experience of how familiar landmarks look at various distances, and how those appearances change during bodily movement. Confirmation of the expectations created by these inputs from experience are understood as partly explaining the ability to accurately perceive depth. Even without the opportunity to explore a perceived scene, in a momentary look at the scene, the ability to perceive depth is understood as drawing upon previous experience in a way that contributes to accurately perceiving depth in a moment's look. Expressed as an inference, empiricist views claim that even momentary depth perception with an individual sense should be explained as an inference with inputs (premises) which draw upon previous experience of confirmation across the senses and perceptual learning.

NTV's view of how experience informs perceptual learning connects empiricists' notion of experience to embodied cognitive science by explaining individual sensory capacities, like visual depth perception, partly in terms of input from other senses, cues related to movement of both the perceiver's body and perceived objects, and long-term memory based upon instances of perceptually exploring a scene (especially becoming familiar with landmarks). Since rationalists already accept a version of the view, held in *NTV*, that learning the location of landmarks has a role to play in explaining remote depth perception (Atherton, 1990, pp. 77-83), *NTV* can be read as

7 There is some similarity between empiricist views and later phenomenological views like Husserl's when it comes to the role of experience in generating expectation. For discussion see Mulligan (1995); Siewert (2015, pp. 147-149).

extending the explanatory role of experience in remote depth perception to proximate depth perception, which rationalist views had explained primarily as a phenomenon occurring within the visual system according to principles of geometry, independent of experience.

NTV also compares visual depth perception to auditory depth perception in a way that illustrates the role of one's body in generating salient forms of experience for perceiving depth. In addition to bringing the role of experience in remote depth perception to bear upon explaining visual depth perception generally, *NTV* brings to bear the role of experience in explaining auditory depth perception for explaining depth perception generally speaking, including problems in *visual* depth perception. Discussion of certain cues for depth perception raised in the following passage of *NTV* illustrate the explanatory role of experience in empiricist views of depth perception generally:

> From what we have shewn it is a manifest consequence, that the ideas of space, outness, and things placed at a distance are not, strictly speaking, the object of sight . . . Sitting in my study I hear a coach drive along the streets. I look through the casement and see it. I walk out and enter into it . . . the ideas intromitted by each sense are widely different, and distinct from each other; but having been observed constantly to go together, they are spoken of as one and the same thing. By the variation of the noise, I perceive the different distances of the coach, and know that it approaches before I look out. Thus by the ear I perceive distance, just after the same manner, as I do by the eye. (Berkeley, 2002, sect XLVI)

Berkeley claims that “By the variation of the noise, I perceive the different distances”, alluding to the role of audible changes in sound over time in his understanding of auditory depth perception. Berkeley describes previous instances of perceptual exploration that connect variations of noise to perceived distance by way of vision and past experience: “I look through the casement and see [the coach]. I walk out and enter into it . . . the ideas intromitted by each sense are widely different, and distinct from each other; but having been observed constantly to go together, they are spoken of as one and the same thing.” Auditory depth perception is explained here by the constant conjunction of ideas of the coach's visual appearance, for example the visual look of the coach drawing nearer, with ideas of the coach's sounds, like the sound of the coach approaching. Past experience is said to contribute cues to auditory depth from across senses (vision, kinaesthetic perception during movement), and in virtue of learning from past experience, like learning how something sounded the last time it was at a certain distance, or the last time it was approaching. Bodily movement is also said to make a key contribution to auditory depth perception by causing variation in sensory input over time that provides cues for the distance a sound is coming from. While any form of movement which causes variation in sensory input over time (like movement of the sound's source) provides cues for auditory depth, movement of one's own body is taken to be the most relevant form of movement because of its role in perceptual exploration and successfully learning where sounds are coming from since early stages of development. (Below, I refer to self-generated bodily movement simply as “bodily movement”).

While cues from perceptual learning based upon experience, like “familiar size” (Predebon, 1992, pp. 985-988), are well understood in contemporary thought about depth perception, empiricist views provide historical reason for reflecting more broadly upon the role of experience in providing cues for depth perception generally and within individual senses like vision or audition. Cues for depth perception that involve experience, where the explanatory role of experience depends upon the nature and use of a creature's body during perceptual exploration, also draw a salient connection between empiricist views of depth perception and embodied cognitive science, discussed below.

Embodiment in depth perception

How specifically do forms of experience cited by empiricists like Berkeley figure in *embodied* views of depth perception? In what follows, embodied views of auditory depth perception are considered in the context of musical performance. Music cognition is a relatively well-developed research program within contemporary embodied cognitive science (Leman & Maes, 2014), and bears an interesting connection to empiricist views of the role of experience in depth perception: auditory depth perception involves development and learning from experience of orienting one's body to the location of a sound's source.

Bodily movement and locomotion are key features in the forms of experience offered to explain remote and proximate depth perception in *NTV* and in embodied cognitive science. One example is Eleanor Gibson's (Gibson & Walk, 1960) landmark studies in depth perception, using the “Visual Cliff” apparatus, which could be considered a relatively early example of embodied cognitive science. The connection between visual and kinaesthetic perception in experiments like Gibson's also forms an important basis for explaining the development of auditory depth perception for empiricists like Berkeley. In a series of experiments, Gibson probes the level of development at which different species, including human infants, acquire certain features of visual depth perception. Individual infants (age 6-14 months) were placed in the centre of a raised platform with one side of the platform constructed from an opaque, obviously solid material, and the other side constructed from a sheet of transparent glass in order to safely simulate a drop from a cliff. A mother stood across from her child with the glass portion of the platform between her and her child, and called for the child to come. Infants crawled across the opaque region, and upon crawling up to the edge of the transparent glass (“the cliff”), infants typically exhibited an avoidance response: either they stopped at the edge and started crying, or they crawled away from the edge and started crying.[8] Avoidance responses in “visual

8 In the history of depth perception, another good example of embodied cognitive science concerns changes in the position of the eyes over time due to anatomical development of the eyes and head. Eighteenth century views knew the importance of the exact position of the eyes for the visual system to triangulate depth (discussed in part one above). Changes to the position (or size) of the eyes, like

cliff" style experiments were explained with reference to the development of bodily movement in human infants and other creatures tested:

> All of these observations square with what is known about the life history and ecological niche of each of the animals tested. The survival of a species requires that is members develop discrimination of depth by the time they take up independent locomotion, whether at one day (the chick and the goat), three to four weeks (the rat and the cat), or six to ten months. (Gibson & Walk, 1960, p. 67)

By the time creatures tested on the visual cliff are capable of locomotion, they are capable of acquiring potentially fatal injuries from locomotion. In human infants, Gibson and Walk (1960) emphasize the possibility of falling from a high place like that simulated by the Visual Cliff apparatus. Permutations of the experiment suggest that it is necessary to experience locomotion and vision concurrently to develop the kind of visual depth perception involved in the visual cliff experiments (Gibson & Walk, 1960, p. 71).[9] From an evolutionary perspective, the simultaneous development of depth perception with locomotion would provide creatures capable of bodily movement with a means of avoiding potentially fatal falls and injuries from locomotion. This is the basis for an important connection between depth perception, bodily movement, and embodied cognitive science.

The claim that the body plays an important explanatory role in depth perception can be connected to embodied cognitive science in different ways, and some give a more salient role to the body than others. According to one way of thinking, bodily movement with concurrent sensory input can be understood as involving multiple senses (proprioception, vision, hearing, etc.), where what senses a creature is thought to possess, which could figure in an explanation of phenomena like depth perception, depends to an extent on what kind of body they have. This way of thinking could be interpreted to draw a somewhat superficial connection between depth perception and embodiment, since on this way of thinking, a creature's body is relevant for explaining depth perception because it is relevant for explaining what senses the creature has, but there is little further explanatory role for the body. For example, by having a body with two eyes, humans become sensitive to certain cues for visual depth, and it is possible to theorize that the human binocular visual system unconsciously infers depth independently of a significant explanatory role for the body.

Empiricist and embodied views find a more significant explanatory role for the body in explaining depth perception, introducing a role for the body based upon the experience of changes in sensory input concurrent with bodily movement. Berkeley's views - taken to reflect also those of his empiricist contemporaries - introduce a role for such experience to explain depth perception within individual sense modalities,

---

anatomical or structural changes undergone during development, are in that sense a good example of how the body of the creature plays an explanatory role in their depth perception.

9 For further evidence aligning with Gibson's conclusions from the "visual cliff" experiments, and which connects strategies for remote depth perception to proximate depth perception see Granrud (2009) as well as Granrud and Schmechel (2006).

rather than emphasizing the role of modalities like vision or audition to explain depth perception independent of bodily movement. For example, normally developed *visual* depth perception, in the sense of seeing how far away something is without moving one's body or sensing cues in other modalities, is understood as dependent upon past experience of bodily movement concurrent with input from other senses. The visual cliff experiments can be taken to show such a link between normally developed visual depth perception and bodily movement during development. “Embodied” views of depth perception, involving individual sense modalities like vision, result from allotting a more significant explanatory role to the creature's body than antecedent, rationalist views.

Contemporary embodied views draw upon empiricist views of the role of experience in depth perception, and often seek to refine the sense in which perception is said to be “embodied” or “enactive”. For example, according to Alva Noë,

> The enactive account of experience can be thought of as a generalization of Berkeley's account of the spatial content of touch to vision and other sense modalities . . . spatial content is available to other modalities such as vision in just the way that it is available to touch, namely in terms of its immediate significance for movement and action. (Noë, 2004, pp. 96-97)

Noë emphasizes the respect in which perception is connected to action, and draws upon Berkeley's thought to argue that bodily movement is crucial for explaining spatial aspects of visual perception (Noë, 2004, pp. 97, 100). While Noë isolates the respect in which perception is connected to action, accounts like Shapiro, 2004 provide theses for distinguishing different respects in which perception can be “embodied” (Shapiro, 2004). To what extent does an author view a perceptual process as realizable in other kinds of brains and bodies? To what extent is a perceptual process insulated from other bodily processes?[10] Such questions are at the heart of empiricist accounts like Berkeley's when challenging rationalist views.

NTV's discussion of depth perception also aims to shift thought about spatial perception away from exclusive focus on vision: “From what we have shewn it is a manifest consequence, that the ideas of space, outness, and things placed at a distance are not, strictly speaking, the object of sight . . .” (Berkeley, 2002, sect XLVI). What is the “same manner” in which we visually and aurally perceive depth? *NTV* explains both visual and auditory depth perception as depending upon past experience of bodily movement with concurrent input from multiple senses over time in the example,

---

10 See the “multiple realizability thesis” and “separability thesis” defined in, Shapiro, 2004, pp. i-ix. van der Schyff and colleagues (2018) hold that: “Considering the mind as embodied means rethinking the boundaries between the neural and extra-neural (e.g. metabolic, thermodynamic, and muscular, among others) factors that drive cognitive processes. From this perspective, the brain becomes a part of a larger network that involves the nervous systems and the sensorimotor capacities of the entire organism (e.g. Gallagher, 2005; 2011)” (van der Schyff, et al., 2018, p.5). For discussion in the context of evolution and culture see van der Schyff and Schiavio (2017).

> Sitting in my study I hear a coach drive along the streets. I look through the casement and see it. I walk out and enter into it . . . the ideas intromitted by each sense are widely different, and distinct from each other; but having been observed constantly to go together, they are spoken of as one and the same thing. By the variation of the noise, I perceive the different distances of the coach (Berkeley, 2002, sect XLVI).

In episodes of perceptual exploration, bodily movement is said to play a key role in forms of experience contributing to depth perception, including depth perception with individual senses. In visual depth perception, bodily movement with concurrent visual experience provides cues for the depth of what is seen, like motion parallax, cues from constancies like size or shape constancy, and cues from familiarity like familiar size (Predebon, 1992, pp. 985-995) or familiar orientation (like seeing an edge as a "depth edge"). In auditory depth perception, bodily movement with concurrent visual experience provides cues for the depth a sound is coming from (its source) based on variations in loudness correlated with the distance from a sound's source (Culling & Akeroyd, 2010, pp. 126-130), acoustic cues like where a sound's source is located based upon echoes or other acoustic qualities resulting from the environment (Culling & Akeroyd, 2010, pp. 137-138), and cues from familiarity like familiar sound or spectra (Davies, 2010, pp. 375-376 and pp. 383-400).

The role of experience in empiricist views of auditory depth perception is of special interest for embodied cognitive science, psychoacoustics, and interdisciplinary music studies. This is because of some key differences of auditory depth perception from visual depth perception connected to experience and bodily movement. Novel properties connected to the source of a sound are presented in auditory perception on the basis of past experience and bodily movement during perceptual exploration, like hearing the width of a sound's source by hearing its sound, and feeling "surrounded" by sound (especially in omni-directional auditory perception). Such properties are of particular interest for music studies and aesthetic experience, in regard to more deeply understanding why subjects experience the spatial aspects of sound the way that they do.

Auditory depth perception is omnidirectional, in the sense that it is capable of presenting sounds as coming from any direction roughly speaking, whereas vision is restricted to the familiar front facing visual field (in human vision). Such differences in auditory depth perception, accessible from the first-person perspective, create a unique role for experience in the development of auditory depth perception, and the role of audition in depth perception generally. Despite typically presenting space with less acuity than vision, audition enables depth perception in every direction because of the capacity of the auditory system to detect sounds from any direction (Culling & Akeroyd, 2010, pp. 123-128). This allows auditory perception to take on unique value for enabling a creature to orient its body to sounding objects in its environment, like orienting away from something that makes an alarming sound, or toward something that makes a desirable sound. (Glatz & Chuang, 2019). Auditory depth perception, on empiricist and embodied views, is deeply connected to bodily orientation because past experience and development of auditory depth perception involves learning how to move one's body to find the source of a sound, no matter what direction the sound is coming from (though there are some differences in acuity depending on direction) (Culling & Akeroyd, 2010, pp. 127-128).

This aspect of auditory depth perception is "adaptive", allowing one to better cope with one's environment, and also invites reflection upon the sense in which musical experience is shaped by the role of the body and bodily movement in a creature's development.[11] For example, feeling surrounded by sound involves auditory depth perception which presents the opportunity to reorient one's body in any direction, in order to learn more about those sources of sound. In aesthetic experience, one aspect of feeling surrounded by sound is therefore likely to consist in feeling surrounded by sources of sound. In sound reproduction, surround sound takes advantage of this when it aims to create an environment which can simulate a sound source positioned at any point around the hearer (Davies, 2010, pp. 404-405). Quadraphonic musical performance goes further in some respects, arranging musical performances into four (usually equidistant) stages around the audience. In quadraphonic environments then – contra to the case of surround sound - part of the experience of feeling surrounded by sound may therefore be informed by one's reaction to feeling surrounded by sources of sound, and what kind of sources one hears (like human performers).

Surround sound and quadraphonic musical performance invite one to question: in indoor concert venues, when sound waves originating from a source located at a single point, but which produces sound waves which "arrive at the listener from every conceivable direction" (Culling & Akeroyd, 2010, p. 135), why is there still a premium placed upon surround sound? For what reasons do people commonly report feeling more "surrounded by sound" or "overwhelmed" in auditory environments like these, as opposed to by sound waves produced by a speaker located at a single point?[12] While there are undoubtedly many explanations, one dimension related to bodily movement in the development of auditory perception is especially interesting for further consideration here.

Sound waves travel omni-directionally from their source and are reflected by surrounding surfaces, creating what can be described as a "sound field with a certain acoustic energy density (the energy per unit volume) which can change over time as the field builds up or dies away" (Davies, 2010, p. 393). Surround sound arrangements are not usually created to more thoroughly surround subjects with sound or sound waves, but to more accurately simulate the conditions of an environment where subjects would actually be surrounded by the sources of the sounds they hear, as they hear them (Davies, 2010, pp. 404-405). In addition to being surrounded by sound, surround sound environments offer subjects the chance to be surrounded by sound sources, like musicians or speakers. On empiricist and embodied views, thinking more deeply about how orienting one's body to sound sources explains auditory depth perception more deeply than consideration of the auditory system independently of past experience involving bodily movement and orientation.

11 For discussion of music cognition as "adaptive" see Reybrouck (2005, pp. 229-266).
12 For discussion see Culling and Akeroyd (2010, p.135); Davies (2010, pp. 379-382 and pp. 389-400).

Quadraphonic musical performance presents a special case of surround sound, in the respect that the goal of such a performance is not usually to accurately reproduce what it is like to be surrounded by sources of ecological sounds, as when surround sound is used in a theatre setting to create realistic experiences. Reflecting on empiricist and embodied views of depth perception suggests a certain way of explaining aesthetically interesting features of auditory depth perception in quadraphonic musical performance. Engagement with the sources of sounds in auditory depth perception can be understood to be explained by creatures' bodily movement and experience during perceptual development, learning, and exploration: feeling “surrounded” or “overwhelmed” in a surround sound setting, including quadraphonic or like musical performance, is partly a matter of hearing the depth of multiple sources of sound surrounding one. Auditory depth perception places one in a position to reorient one's body to the sources of sound that one hears, linking the role of bodily movement directly to auditory depth perception. Especially when the source of sound is something pleasurable or interesting, embodied views of auditory depth perception explain why the spatial aspect of auditory perception features so prominently in auditory aesthetic experience: embodied views claim that the spatial aspect of auditory perception is explained by the adaptive capacity of creatures to reorient their body. In the context of musical performance, a consequence is that perceptual experience of auditory depth perception among audience members can be more deeply understood with reference to the sources surrounding the audience, especially when it comes to feeling surrounded by performers in the case of quadraphonic musical performance.

Conclusion

This paper offers support for the claim that empiricist views of depth perception of the eighteenth century provide an earlier historical precursor for embodied cognitive science than is often acknowledged in current literature. Contemporary technologies and research concerned with auditory depth perception in musical performance scenarios provide new insights to examine the claim. Experiments in depth perception including Eleanor Gibson's “visual cliff” offer early examples of embodied cognitive science in the twentieth century, and important points related to the role of bodily movement in those experiments are salient for eighteenth century empiricist views like Berkeley's in *NTV*. Whereas empiricist accounts supplemented earlier rationalist accounts, creating a new role for “experience” and cues that experience makes available for explaining depth perception, embodied cognitive science can be understood as supplementing empiricist accounts by examining more deeply the role of the body and bodily movement in perception. In the context of aesthetic experience and musical performance, bodily orientation may be thought to play a special explanatory role in auditory depth perception because of the way in which experience and bodily movement inform depth perception for empiricist and embodied views.

Acknowledgements

I would like to thank Charles Siewert, Alex Morgan, James Pomerantz, and two anonymous referees from the *Journal of Interdisciplinary Music Studies* for very helpful feedback on versions of this paper. I am indebted to James Pomerantz for discussion of “The 'Visual Cliff” in the context of embodied cognitive science. Thank you also to participants at the Conference on Interdisciplinary Musicology 19, and the 55th meeting of the Western Canadian Philosophical Association, especially Taro Okamura for comments.

References

Atherton, M. (1990). *Berkeley's Revolution in Vision*. Cornell University Press.

Bellis, D. (2016). The Perception of Spatial Depth in Kepler's and Descartes' Optics: A studyof an Epistemological Reversal. In K. Vermeir and J. Regier (Eds.), *Boundaries, Extents and Circulations: Space and Spatiality in Early Modern Natural Philosophy*, Springer, 41. https://doi.org/10.1007/978-3-319-41075-3_5

Berkeley, G. (2002). *An Essay Towards a New Theory of Vision* (D. R. Wilkins, Ed.). Trinity College Dublin, accessed 07/2020 via <https://www.maths.tcd.ie/~dwilkins/Berkeley/Vision/1709A/Vision.pdf>. (Original work published 1709).

Chemero, A. (2015). The World at the End of the Cane. *The Philosopher's Magazine*. 68(1), 84-87. https://doi.org/10.5840/tpm20156817

Copenhaver, R. (2010). Thomas Reid on Acquired Perception. *Pacific Philosophical Quarterly*, 91, pp. 285-312. https://doi.org/10.1111/j.1468-0114.2010.01368.x

Culling, J. F., & Akeroyd, M. A. (2010). Spatial Hearing. C. J. Plack (Ed.) *Oxford Handbook of Auditory Science*, 3. Oxford University Press.

Davies, W. J. (2010) The Acoustic Environment. In C. J. Plack (Ed.) *The Oxford Handbook of Auditory Science: Hearing*, 3. Oxford University Press.

Descartes, R. (1965). Dioptrics (P. Olscamp Trans., Ed.) *Discourse of Method, Optics, Geometry and Meteorology*. Hackett. (Original work published 1637).

Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford University Press.

Gallagher, S. (2011). Interpretations of Embodied Cognition. In W. Tschcher, C. Bergomi (Eds.) *The Implications of Embodiment: Cognition and Communication* (pp. 59-73). Imprint Academic.

Gibson, E. J., & Walk, R. D. (1960). The 'Visual Cliff'. *Scientific American*, *202*(4), 64–71. https://doi.org/10.1038/scientificamerican0460-64

Glatz, C., & Chuang, L. L. (2019). The Time Course of Auditory Looming Cues in Redirecting Visuo-spatial Attention. *Scientific Reports*, *9*. https://doi.org/10.1038/S41598-018-36033-8

Granrud, C.E. (2009). Development of Size Constancy in Children: A Test of the Metacognitive Theory. *Attention, Perception, & Psychophysics*, *71*(3), 644-654. 10.3758/APP.71.3.644

Granrud, C.E., & Schmechel, T.T.N. (2006). Development of Size Constancy in Children: A Test of the Proximal Mode Sensitivity Hypothesis. *Perception & Psychophysics, 68*(8), 1372-1381. 10.3758/bf03193736

Hatfield, G. (2001) Perception: History of the Concept. *International Encyclopedia of the Social & Behavioral Sciences*, 11202-11205. 10.1016/B0-08-043076-7/00137-6

Hatfield, G. (2002). Perception as Unconscious Inference. D. Heyer and R. Mausfeld (Eds.) *Perception and the Physical World*. Wiley. https://doi.org/10.1002/0470013427.ch5

Hatfield, G. (2015). Natural Geometry in Descartes and Kepler. *Res Philosophica*. *92*(1), 117-148. 10.11612/resphil.2015.92.1.6

Leman, M., & Maes, P. (2014). The Role of Embodiment in the Perception of Music. *Empirical Musicology Review*, *9*(3-4), 236-246. http://dx.doi.org/10.18061/emr.v9i3-4.4498

Malebranche, N. (1997). *The Search After Truth* (T.M. Lennon & P. J. Olscamp, Eds.) *Elucidation of the Search After Truth*. Cambridge University Press. (Original work published in 1674).

Mulligan, K. (1995). Perception. In B. Smith and D. W. Smith (Eds.) *The Cambridge Companion to Husserl*. Cambridge University Press.

Noë, A. (2004). *Action in Perception*. MIT Press.

O'Callaghan, C. (2011). Lessons from Beyond Vision (Sounds and Audition). *Philosophical Studies*, *153*(1), 143-160. 10.1007/sl 1098-010-9652-7

Predebon, J. (1992). The Familiar-size Cue to Distance and Stereoscopic Depth Perception. *Perception*. 22, 985-995. https://doi.org/10.1068/p220985

Reybrouck, M. (2005). A Biosemiotic and Ecological Approach to Music Cognition: Event Perception Between Auditory Listening and Cognitive Economy. *Axiomathes*, *15*, 229-266. 10.1007/s10516-004-6679-4

Schroeder, S. (2002). George Berkeley's Embodied Vision. *Philosophy in the Contemporary World*, *9*(2), 87-92. pcw20029223

Schwartz, R. (1994). *Vision*. Cambridge University Press.

Shapiro, L. (2004). *The Mind Incarnate*. MIT Press.

Siewert, C. (2015). Phenomenological Approaches. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception* (pp.136-152). Oxford University Press.

van der Schyff, D., Schiavio, A., Walton, A. Velardo, V., & Chemero, A. (2018). Musical Creativity and the Embodied Mind: Exploring the Possibilities of 4E Cognition and Dynamical Systems Theory. *Music & Sciences*, *1*, 1-18. https://doi.org/10.1177/2059204318792319

van der Schyff, D., & Schiavio, A. (2017). Evolutionary Musicology Meets Embodied Cognition: Biocultural Coevolution and the Enactive Origins of Human Musicality. *Frontiers in Neuroscience*. *11*, 1-18. https://doi.org/10.3389/fnins.2017.00519

Whitmer, W.M., Seeber, B.U., Akeroyd, M.A. (2013). Measuring the Apparent Width of Auditory Sources in Normal and Impaired Hearing. *Adv Exp Med Biol*. 787, 303-310. 10.1007/978-1-4614-1590-9_34

Wilson, R., & Foglia, L. (2015). "Embodied Cognition". E. Zalta (Ed.) *Stanford Encyclopedia of Philosophy*. Accessed 07/2020 via <https://plato.stanford.edu/entries/embodied-cognition/>.

Biography

Don Oxtoby is PhD candidate in philosophy at Rice University. His work focuses on perception, perceptual judgement, and audition. He holds an MA in philosophy from Rice University and an MA in philosophy from McMaster University.