

Vision as Dance? Three Challenges for Sensorimotor Contingency Theory*

Andy Clark

School of Philosophy, Psychology and Language Sciences
University of Edinburgh

Abstract

In *Action in Perception* Alva Noë develops and presents a sensorimotor account of vision and of visual consciousness. According to such an account seeing (and indeed perceiving more generally) is analysed as a kind of skilful bodily activity. Such a view is consistent with the emerging emphasis, in both philosophy and cognitive science, on the critical role of embodiment in the construction of intelligent agency. I shall argue, however, that the full sensorimotor model faces three important challenges. The first is to negotiate a path between two prima facie unsatisfactory readings of the central claim that conscious perceptual experience is constituted by knowledge of patterns of sensorimotor dependence. The second is to convince us that the sensorimotor contribution, in such cases, is actually constitutive of perceptual experience rather than merely causally implicated in the origination of such experienceⁱ. And the third is to respond to the important challenge raised by what I will dub 'sensorimotor summarizing' models of the relation between conscious experience and richly detailed sensorimotor routines. According to such modelsⁱⁱ conscious perceptual experience only rather indirectly reflects the rich detail of our actual sensorimotor engagements, which are instead lightly sampled as a coarse

guide, optimized for planning and reasoning, and geared and filtered according to current needs and purposes.

1. The Central Claim

Dance is, without doubt, a kind of skilful bodily activity. Subtract the body and its activity and – modulo only the severest forms of contemporary exploration – the dance itself must disappear from view. The central claim of Alva Noë's important, stylish and challenging treatment is that perception is more like dance than philosophers and cognitive scientists have (mostlyⁱⁱⁱ) noticed. For like dance "perceiving is a kind of skilful bodily activity" (Noë (2004)^{iv} p.2). What can this mean?

It does not mean, merely, that you need a body (or at least, some sense organs and a brain) to perceive. Rather, it means that skilful bodily action and perception are intimately entangled. The key to this shared intimacy is the idea that conscious perceptual experience consists in a perceiver's implicit knowledge of 'sensorimotor contingencies': rules or regularities relating sensory inputs to movement, changes and action. Implicit knowledge of such contingencies typically amounts to having a set of expectations concerning the "way sensory stimulation varies as a result of movement" (75). Both the character (the 'what it is like' of vision, touch, hearing etc) and the contents (concerning space, color, shape etc) of our perceptual experiences are said to be determined by our implicit knowledge of SMC's (sensorimotor

contingencies), or (as Noë now prefers to say) by our 'sensorimotor expectations').

To illustrate this, consider a visually presented horizontal line and a looming ball heading at alarming velocity towards your face. These two distinct experiences correspond, on Noë's account, to two distinct signatures in sensorimotor space. Thus, if you move your eyes along the straight line, the retinal stimulation is invariant, whereas if you move your eyes up or down relative to the line, there is a sudden shift. A stationary ball would display a very different profile to this, while the moving ball presents a distinctive looming pattern that can (and probably should) be terminated by an act of ducking. The first claim, as I shall understand it, is that differences in what we perceptually experience correspond to differences in the sensorimotor signature associated with certain objects, properties and states of affairs. If two things look different, they do so because in encountering them we bring to bear (rightly or wrongly) different sets of sensorimotor expectations. But despite such differences, for all visually presented objects there will be some large parts of the sensorimotor signatures in common. It is these commonalities that make the experiences visual (rather than, say, auditory). For example, vision (unlike audition or touch) samples the front or facing sides of objects, and so on. The visual attributes of sensed objects are thus that subset of the signature sensorimotor contingencies that pertain to the distinctive ways that the visual sense can sample the real properties of objects. Thus, the very same real property (e.g. size) may be apprehended by vision or sometimes (for small objects) by touch. But the mode of sampling varies dramatically, and with it the associated sensorimotor contingencies.

The visual properties of an object are, on this account, nothing but certain objective properties (e.g. size, shape and color) sampled in a distinctive way. Of course, some properties (such as color) are available to humans only via one modality (vision). But we can imagine e.g. color judgments made by a prosthetic sensor and reported via audible tones. What makes our normal experience of color *visual* thus remains the details of a certain mode of sampling.

The central idea is thus that the contents and character of perceptual experience are determined by implicit knowledge of various types of what I am calling 'sensorimotor signature'. These signatures can include various elements, some having to do with the expected effects of our own movement on the input, others concerning the way changes to external conditions (lighting etc in the case of color) will affect the input. Some of the elements that enter into sensorimotor signatures thus turn on objective features of the world, while others turn on idiosyncratic features of our own sensory apparatus (the curve of the eyeball, the spacing of photoreceptive cells, etc). Uniting all the cases is the guiding idea that perceived content and character depends on expectancies concerning the future unfolding, under various conditions, of patterns of sensory stimulation.

The first challenge can now be stated. What does it mean to speak of expectancies concerning the future unfolding of patterns of sensory stimulation? In particular, are we to understand 'expectancies concerning sensory stimulation' as a personal or sub-personal phenomenon? It would be a personal level phenomenon if our expectancies concerned sensory experiences themselves (e.g. we expect the ball to (in

one sense) look bigger and bigger as it approaches our head). It would be a sub-personal phenomenon if what was at issue was some neural network's being able to predict the increasing area of some pattern of sensory stimulation defined at, say, the retina.

Both readings face difficulties. The first reading, as Noë (87, 228) notes, courts circularity. It assumes we already have perceptual experiences of object appearances (e.g. the way the plate looks elliptical from an angle-see p.84) and then builds further kinds of content (e.g. the plate's also looking round) from our knowledge about how e.g. that elliptical look would vary as a result of movement around the plate. The experience of roundness just is, on this account, the active deployment of our implicit understanding of how the various *looks* would alter as a result of motion.

Put like that, the story assumes 'ways things look visually' and does not explain them in any fundamental sense. Instead, what is explained is our grasp of a certain kind of visual content (roundness) on the basis of other kinds of visual content (regular variations in elliptical looks, etc).

To avoid the threat of circularity, Noë suggests we should understand what it is to experience a look as nothing but our drawing on a certain set of more basic sensorimotor skills. The elliptical look of the plate is cashed out in terms of my ability to move my hand in a certain manner were I to try to indicate the shape as it appears in my visual field. Noë writes that 'In this way, my sensorimotor skill is drawn on to constitute my experience of the shape' (89). Looks are thus relations between the sensorimotor repertoire of embodied agents and objects. The virtue of this proposal is that it

avoids a phenomenal (and hence circular) model of looks while (it seems) keeping the story operating at the personal level i.e. at the level of our actual understanding of our own sensorimotor space.

The obvious drawback with this proposal is that it leaves it unexplained why knowledge concerning the relevant sensorimotor space should result in the experience of anything at all. Perhaps this gap can be filled (see Clark (2001) and Pettit (2003) for some attempts). But nothing in Noë's account appears apt to plug the gap.

A second reading, more in line with the earlier account laid out by O'Regan and Noë (2001) would pitch the whole story at the sub-personal level. What is doing the work, on this reading, is knowing (non-consciously, implicitly) how the energy patterns impacting your sense organs will vary in response to your own actions and movements. (This is, I note, just the sort of knowledge that would be acquired by a simple recurrent neural network trained to predict the next sensory input from information concerning the present sensory state plus efferent copy of a motor command). This reading also avoids circularity (and hence was preferred in the earlier work- see Noë p.228) but seems to have been dropped in Noë's book because of a lack of 'phenomenological aptness' (228). The worry seems to be that unconscious knowledge (of sensorimotor contingencies) provides no obvious basis for phenomenal consciousness, which is what Noë seeks to explain. But as we just saw, neither does the version that invokes actual experiences of looks. By offering an account apparently pitched at the personal level that nonetheless denies that it is trading in

phenomenality as such (but only in our own grasp of our sensorimotor relations to objects), Noë may be hoping to somehow find a safe haven between the two readings. But I am not convinced that any stable intermediate line is actually displayed in the text. For either our own grasp of our sensorimotor relations to objects is something that figures in our experience or it is not. If it is, then the central claim looks circular (as an account of visual experience). If it is not, then we lose our grip on 'phenomenological aptness' and a gap looms between possessing these bodies of skill and actual visual experience. What is lacking is a persuasive account of *why it is* that certain patterns of sensorimotor knowing (understood in a staunchly non-experiential way) should make it the case that a creature has some form of perceptual experience.

My own view is that as it stands, the account on offer is best viewed as a proposal concerning how conscious perceptual experiences get their contents, rather than an account of how there come to be conscious perceptual experiences at all. Considered in this light, we can perhaps say that experiences of looks get their content from our basic repertoires of sensorimotor skills (or orienting, grasping, pointing etc) while other experiences (e.g. seeing the top of the cup as circular rather than elliptical etc) get their content from our knowledge of how those looks will vary with motion and other conditions. But even considered as a story about content fixation, important challenges remain, as we shall now see.

2. Constitutive Force?

A specific dance, at some appropriate level of description, might plausibly be identified with a specific pattern of bodily motions. It may make no sense to suppose that one could know the dance without knowing something of those specific patterns of bodily motion. In this way, we espy something like a conceptual link between the dance and the details of its embodied realization.

At first blush, however, the link between our sensorimotor knowledge and skills and the contents of our perceptual experience looks less direct, as if the sensorimotor routines might be the hooks that reel in the contents, but in a merely causal fashion. The very same contents, one might well suspect, could be present in systems whose sensorimotor routines were very different to our own, or perhaps even in systems that were sensorimotor inert. I take it that these are the sorts of static internalist intuition that Noë (admirably) wants to unseat.

One way to unseat them would be to embrace a radical option that Noë (p.89 and elsewhere) rejects, describing it as unacceptably behaviorist. The option would be to depict at least some basic forms of perceptual content as necessarily involving dispositions to act in certain ways (see Evans (1985) for a classic version of such a line). This seems plausible for e.g. certain egocentrically defined contents, such as that of a sound's appearing to come from *over there*. Entertaining such a content may in part consist in a swathe of dispositions to orient towards the sound. Subtract those dispositions to act and you must (if such an account is right) subtract the perceptual content itself. Noë depicts his own view in contrast to this, and as involving, in the case of a visually perceived flicker on the right, only 'practical

knowledge of how *movement* would bring the thing into view' (89).

By moving us squarely back into the realm of knowledge, however, Noë runs a risk of letting internalism in through the back door, and creates an internal tension between two components of his own account.

One component is the oft-repeated idea that what counts are our expectations (or our implicit knowledge) concerning the way sensory stimulation will unfold. The other component is the idea that we try to bridge the explanatory gap (between physical goings-on and conscious experience) "by expanding our conception of the substrate in terms of which we hope to explain consciousness" (226). This expansion looks to the way neural activity supports embodied action as the missing link^v in explanations of perceptual consciousness. Here, the general claim is that "what determines phenomenology is not neural activity set up by stimulation as such, but the way the neural activity is embedded in a sensorimotor dynamic" (227). In this way:

"Experience is not caused by and realized in the brain, although it depends causally on the brain. Experience is realized in the active life of the skillful animal. A neuroscience of perceptual consciousness must be an enactive neuroscience—that is, a neuroscience of embodied activity, rather than a neuroscience of brain activity" (226)

But once again, it is not clear that Noë can have this both ways at once. If we are to avoid behaviorism by stressing the role of knowledge and expectations (concerning smc's), why isn't that all 'in the brain'? How do we distinguish the radical gap-bridging view from the more conservative idea that the embodied activity is just a causal precondition of setting or re-setting parameters in neural structures that encode the kinds of knowledge of sensorimotor contingencies that Noë likes to stress? Once set (or reset) and activated, these neural structures do indeed (it might be argued) realize the conscious perceptual experience. Giving away still more, perhaps they realize the experience not in a snapshot moment but only in a temporal evolution. We would thus have a time-extended, sensorimotor knowledge constituted, but fully internal vehicle of the conscious experience. Such a temporal evolution might normally be causally scaffolded by real-world action and self-selected real-world input, but that need not always be the case (e.g. in dreams and vivid hallucinations). As far as I can see, nothing in the concrete cases described in the book favors Noë's radical account over the more conservative and internalist option just described (for example, both would predict and explain the key patterns of experiential plasticity, including the case of the qualitative similarity of vision and TVSS (tactile visual sensory substitution)). The appeal to knowing how (rather than knowing that) seems only to focus attention on a type of internally coded knowledge: one causally beholden to real-world embodied action but not conceptually tied up with it.

It may be that Noë needs to bite (or at any rate sample) the behaviorist bullet, by arguing that nothing could *count* as the relevant know-how unless it supported certain dispositions

to behave. But that would be to fall back on the Evans-style account that was explicitly rejected in chapter 3. In general, it seems to me that securing genuine constitutive force for knowledge of smc's in the construction of perceptual experience may indeed require some such move, behaviorist or not.

3. Sensorimotor Summarizing.

So far, I have been concerned with what might be seen as internal worries and tensions: ones arising *within* Noë's ambitious (and immensely illuminating and informative) account. I should finally confess, however, to harboring a deeper and darker suspicion. It is the suspicion that the full sensorimotor model is empirically flawed, in a way that distorts the true role of sensorimotor skills in the human cognition. It is, to put the matter starkly, the thought that conscious perception may be more akin to reason than to dance.

The backdrop to this worry is work on the so-called 'dual visual systems hypothesis' (Milner and Goodale (1995)^{vi}). According to this hypothesis, our capacities of conscious visual awareness (which they sometimes call capacities of visual *perception* and contrast with capacities for visually-guided *action*) depend on a specific visual processing stream - the ventral stream - that is said to operate semi-independently of the processing stream (the dorsal stream) that guides fine-tuned motor action in the here-and-now. The dorsal stream is thus said to be specialized for fluent motor interaction while the ventral stream deals with enduring object properties and subserves explicit

recognition and semantic recall. As a kind of corollary, the ventral stream is said to take over whenever the real-world object is not present-at-hand: actions in respect of imagined or recalled objects are under ventral stream control (see Milner and Goodale (1995) pp. 136-138). Evidence for the ventral/dorsal dissociation comes in three main varieties: deficit data concerning patients with damage to areas in either the dorsal or ventral streams; performance data from normal human subjects (using experimental paradigms such as the Titchener Circles illusion); and computational conjectures concerning the inability of a single encoding to efficiently support both visual form recognition and visuomotor action. I shall not rehearse these bodies of evidence here (see Clark (1999) (2001)) but will instead locate the dual visual systems hypothesis in the wider setting of what I will now call "sensorimotor summarizing" accounts of the contents of conscious perceptual experience.

Sensorimotor summarizing accounts depict the function of conscious visual perception as the delivery of various forms of optimized representation: representations optimized to aid performance in ecologically normal circumstances. In line with Milner and Goodale's emphasis on the neural division of visual labor, such accounts understand the relevant optimizations (in the case of ventrally dominated conscious visual perception) as *ones geared to the role of visual information in planning, thought and reasoning*, rather than the fine-grained (dorsally dominated) control of visuomotor action. In, for example, the Titchener Circles experiment, the conscious visual representations would be optimized (just as Milner and Goodale suggest) to guide the choice of *which* disc to pick up, and the choice of what *kind* of grip to deploy (one apt for picking up and not, e.g. for throwing). The

conscious illusion (of one circle's being larger than another) may then be best explained by the visual system's delivering a representation enhanced in the light of information about relative size: a trick that is effective for reasoning and choice in most ecologically realistic situations, but that cannot be tolerated by fine sensorimotor control systems.

Similarly, a study by Carrasco et al (2004) found that the allocation of attention affected the appearance of a visual stimulus, causing an enhanced contrast effect in a cued grating. Reporting on this effect Treue (2004) comments that:

" Attention turns out to be another tool at the visual system's disposal to provide an organism with an optimized representation of the sensory input that emphasizes relevant details, even at the expense of a faithful representation of the sensory input" Treue (2004) p.436-437

It is, in fact, a commonplace idea in contemporary neuroscience that conscious awareness is in general concerned with the provision of what Christof Koch (2004) terms 'executive summaries' apt to aid frontal regions in the selection of one among a set of possible types of action or response. Campbell's (2002) 'targets' view of consciousness, Jacob and Jeannerod's (2003) more nuanced version of the dual visual streams view, and Matthen's (2005) account of 'descriptive sensory systems' all share something of the flavor of this kind of view. What all these views share is the image of conscious perceptual experience as reflecting the content of representations whose cognitive role is to enable the deliberate selection of actions and action types (including, and perhaps especially, 'epistemic actions' such

as sorting, sifting, comparing and the like- see Pettit (2003), Matthen (2005)). Such representations need not, and on this account do not, typically reflect the full intricacies of our actual sensorimotor engagements with the world. Instead, they are optimized to inform reason, selection and choice, and thus reflect only the broad outlines of a space of possible targets and possible kinds of sensorimotor engagement.

I do not claim that this view is correct, or even (as yet) sufficiently clearly articulated. I do think, however, that it is suggestive of an importantly different take on the role of detailed sensorimotor knowledge in the construction of perceptual experience. For if it is at all on the right track, then detailed sensorimotor knowledge may be buffered from the realm of conscious perceptual content, informing such contents only at a fairly high level of abstraction. If this is correct, then one of the most striking implications of the full sensorimotor model may be called into question. This is the claim that all differences in embodiment (insofar as they impact sensorimotor contingencies) must thereby make *some* difference to qualitative experience (27-28). This claim will turn out to be false if what structures experience is (not the full suite of sensorimotor details but) a kind of coarse summary whose main concern is with a space of targets and of action types. The sensorimotor summarizing view thus avoids the threat of 'sensorimotor chauvinism' (Clark and Toribio (2001)), leaving it an open empirical question to what extent (if any) sameness of experience requires sameness of embodiment.

It may be, however, that neither the full sensorimotor summarizing view nor the full sensorimotor contingency

view has the ability to capture all the layers and components of perceptual experience. Perhaps, that is, perceptual experience is not a unitary thing, whose contents are to be determined by a single suite of engagements or representations. Such a mixed view is in fact suggested (in different ways) by both Jacob and Jeannerod (2003) and by Matthen (2005). What still seems likely, even if this is so, is that conscious perception is *sufficiently deeply* bound up with reason and planning for those special needs to significantly impact and buffer the role of 'raw' knowledge of sensorimotor contingencies in the construction of perceptual experience.

Conclusions: Dancing with Reasons

I have raised three challenges for Noë-style sensorimotor contingency theory.

The first challenge is to find a safe haven between two unsatisfactory readings of the central claim that perceptual experience is conditioned by expectancies concerning sensory stimulation. One reading looks circular, since it depicts the expectancies as already operating in the realm of experience. The other looks inadequate, since it depicts the expectancies as fully sub-personal (e.g. as concerning energetic impacts on sense organs) and fails to visibly bridge the ever-treacherous 'explanatory gap'. As an aside, I'd personally be strongly inclined to put my money on the latter reading, and to hope to tell some kind of additional story to (try to) dissolve the gap (see Pettit (2003) and Clark (2000) for examples of such stories). Interestingly, Noë (228)

seems to opt for the former, thus leaving the account in the uncomfortable position of perhaps taking 'a little bit of consciousness for granted' (230)

The second challenge is to fix the intended force of the central claim. Is the claim that there is a conceptual connection between sensorimotor knowing and the contents of perceptual experience? Intent to avoid (what he sees as) the behaviorism in Evans and others' appeals to actual dispositions to behave, Noë leaves this reader wondering whether anything less (such as Noë's cautious appeal to practical knowledge) can really secure more than a causal role for the sensorimotor loops themselves.

The third challenge is to accommodate (or give principled reasons to reject) the fairly extensive empirical data suggesting that the contents of conscious visual experience are optimized for selection, choice and reason, and effectively dislocated from the representations that enable the fine control of action. Such optimization and dislocation threatens to introduce a buffer zone keeping detailed knowledge of sensorimotor contingencies at arms length from the day-to-day business of conscious perceptual experience.

Challenges aside, *Action in Perception* is a wonderful achievement. Noë has produced a lucid, rich, elegant, and important work that raises a wide variety of truly central issues in a fresh and constantly thought-provoking way.

* Thanks to Ned Block, Susan Hurley, Josefa Toribio, Matthew Nudds, and Alva Noë for illuminating discussions of the sensorimotor view. This project was completed thanks to teaching relief provided by Edinburgh University and by matching leave provided under the AHRC Research Leave Scheme.

References

Ballard, D.H., M.M. Hayhoe, P.K. Pook, and R.P.N. Rao (1997) "Deictic codes for the embodiment of cognition," *Behavioral and Brain Sciences* 20, 4, 723–767.

Campbell, J (2002) *Reference and Consciousness* (Oxford University Press, Oxford)

Carrasco, M, Ling, S and Read, S (2004) "Attention alters appearance" *Nature Neuroscience* 7, 308 - 313 (2004)

Clark, A (1999) "Visual Awareness and Visuomotor Action" *Journal Of Consciousness Studies* 6:11-12:1-18

Clark, A (2000) "A Case Where Access Implies Qualia?" *Analysis* 60:265:2000 p.30-38

Clark, A (2001) "Visual Experience and Motor Action: Are The Bonds Too Tight?" *Philosophical Review* 110:4: 495-519

Clark, A and Toribio, J (2001) "Sensorimotor Chauvinism?: Commentary on O'Regan, J.K., and Noë, A. "A sensorimotor approach to vision and visual consciousness" *Behavioral and Brain Sciences* 24: 5: 979-980

Evans, G (1985) "Molyneux's Question" in G. Evans, *The Collected Papers of Gareth Evans* (Oxford University Press, London)

Jacob, P and Jeannerod, M (2003) *Ways of Seeing: The Scope and Limits of Visual Cognition* (Oxford University Press, Oxford)

Koch, C (2004) *The Quest for Consciousness* (Roberts and Co, NY)

Matthen, M (2005) *Seeing, Doing and Knowing* (Oxford University Press, Oxford)

Merleau-Ponty, M [1945] (1962) *The Phenomenology of Perception* trans. Colin Smith Routledge Press, London

Milner, A., and Goodale M., (1995) *The Visual Brain in Action*. Oxford University Press, Oxford

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

O'Regan, J.K. (1992) "Solving the 'real' mysteries of visual perception: the world as an outside memory" *Canadian Journal of Psychology* 46:3:461-488

O'Regan, J.K., and Noë, A. (2001) "A sensorimotor approach to vision and visual consciousness" *Behavioral and Brain Sciences* 24: 5: 939-973

Pettit, P (2003) "Looks as Powers" *Philosophical Issues* 13:221-252

Treue, S (2004) "Perceptual enhancement of contrast by attention" *Trends in Cognitive Sciences* 8:10: 435-437

Varela, F., Thompson, E., and Rosch, E (1991) *The Embodied Mind* (MIT Press, Camb. MA)

ⁱ This point was brought home to me in several conversations with Ned Block, to whom this short treatment owes a special debt.

ⁱⁱ See Milner and Goodale (1995), Clark (1999), Clark (2001), Jacob and Jeannerod (2003), Campbell (2002) and the interesting, multi-layered discussion in Matthen (2005)

ⁱⁱⁱ Exceptions include Merleau-Ponty (1945/62), Varela, Thompson and Rosch (1991), and O'Regan (1992), as well as the recent waves of work in 'active vision' (eg Ballard et al (1997).

^{iv} Henceforth, all page references are to this work unless otherwise stated

^v I do not myself see how this move alone would bridge the explanatory gap, hence some of my concerns in the previous section.

^{vi} Noë (p.19) rejects the suggestion that the 'dual visual streams' work raises problems for the sensorimotor view, on the grounds that (1) his story makes no claims about what conscious vision is for, and (2) that both dorsal and ventral activity depend on sensorimotor skills. The thought I want to explore is that the way sensorimotor information matters for conscious perception will be quite different if that information is first optimized for reasoning and planning, rather than accessed 'raw' and simply *made available to* reason and planning.