Marek McGann

# *Perceptual Modalities*

## *Modes of Presentation or Modes of Interaction?*

**Abstract:** *Perceptual modalities have been traditionally considered the product of dedicated biological systems producing information for higher cognitive processing. Psychological and neuropsychological evidence is offered which undermines this point of view and an alternative account of modality from the enactive approach to understanding cognition is suggested. Under this view, a perceptual modality is a stable form of perception which is structured not just by the biological sensitivities of the agent, but by their goals and the set of skills or expertise which they are deploying at a given time. Such a view suggests that there is no such thing as an experience that is purely visual, auditory, or otherwise modal and that our attempts to understand consciousness and the mind must be conducted within a framework that provides an account of embodied, goal-directed adaptive coping with the world.*

Keywords: Modality, perception, sensation, enaction, skill theory, consciousness

### Introduction

This paper is about perceptual modalities — how we should conceive of them and what their relationship is to our bodies and to our consciousnesses. Addressing such a topic, however, first demands that we get straight a common distinction between 'sensation' on the one hand, and 'perception' on the other.

Cognitive scientists tend to interpret the distinction between sensation and perception in one of two ways. The first makes a strong

Correspondence:
Marek McGann, Department of Psychology, Mary Immaculate College, University of Limerick. *Email: marek.mcgann@mic.ul.ie*

distinction between the two, making the former a physiological process and the latter a psychological one. In this case, sensation is a characteristic of various surfaces and cells of the body which react in particular ways when they make contact with particular forms of energy such as light, pressure, or sound, or with certain chemicals. Perception, on this strong view, is the psychological process, enabled by sensation, by which we come to some form of direct knowledge of the world.

The second form of the sensation-perception distinction found amongst cognitive scientists makes the division between the two less clear. In this case, sensation is the bare basis of later perception, the recognition of elementary aspects of a given object or stimulus such as its form, colour or pitch. Perception, then, is the process by which these elements are assembled or bound into coherent but complex wholes, transforming the sensations of redness and roundness into the perception of a ripe tomato, for instance. This second form of the distinction, though both popular and quite intuitive, begs for confusion. The physiological aspects of sensation are less clearly articulated, though generally remaining at the foundation of such basic aspects of perception as recognition of colour and form. In this second case, then, the concept of sensation straddles the division between physiology and psychology, and thus, I suggest, obscures some very important aspects of that division.

The focus of this paper will be on perceptual modalities, and their relationship to sensation in that first, clear sense of physiological response. The differences between modalities form some of our most basic understandings of the structure of consciousness, and indeed the very nature of consciousness. The traditional view of the nature of perceptual modalities sees not only a strong continuity between the physiological response to a stimulus and the resultant perception of it, but a near identity between the two — an inextricable relationship illustrated by the popularity of the second, more vague form of the sensation-perception distinction. In this paper I will draw on both psychological and neuroscientific research to undermine some of our most fundamental intuitions concerning the relationship between sensation and perception and thus to undermine the long-standing traditional account of perceptual modalities, an account which can trace its lineage directly to Johannes Müller's doctrine of 'specific nerve energies', put forward in 1826. In the place of this traditional account I will argue for an enactive account of perceptual modality, one which draws heavily on the work of Kevin O'Regan, Alva Noë and Erik Myin amongst others (Myin & O'Regan, 2002; Noë, 2004; O'Regan

&; Noë, 2001a) but which offers a more coherent and more complete story of the various elements of a modality.

The enactive approach to cognitive science has been knocking around for a while now, but has thus far been largely relegated to the domain of 'low level' cognition, the basic aspects of perception and immediate bodily action. The account of modality that I will advance later in this paper will provide grounds to reject a clear distinction between lower and higher cognition, and attempt to show how an enactive approach can equally and fruitfully be applied to our understanding of the richer and more complex forms of cognition.

The term 'enactive' has become something of a buzz word in the recent Cognitive Science literature, and as such has developed a few different meanings. In the present paper, I am applying the term in the specific sense proposed by Varela, Thompson, & Rosch (1991), and developed in subsequent works by several authors including Varela (1997), Di Paolo (2005), Thompson (2007), Di Paolo *et al*. (in press) and others. This approach is in many ways a continuing development of the phenomenology of Merleau-Ponty, and the integration of some of his insights with more recent developments in the dynamics of biology and autonomous systems (see Varela, 1979; Weber & Varela, 2002; Thompson, 2007).

The emphasis on biological autonomy and the dynamic interaction between such an autonomous system and its environment in the explanation of cognition means that the enactive approach argues for a continuity between the basic processes that underlie life, adaptive response to the world, and cognition. The mind is seen not as comprising a set of distinct processes driving perception, cognition and action, but a complex of skills which allow an autonomous agent to maintain itself and achieve its goals in interaction with its environment. Perception, cognition and action are three facets of this single process of adaptively coping with the world (what the enactive literature refers to as 'sense-making') rather than being distinguishable links in a chain of processes that begin with 'input' at the sensory surfaces and end with 'output' at the muscles. Though the mind is driven by the autonomy of the agent, all of its activities occur in engagement with those aspects of the world around it that impinge upon it. The description of either an agent or its environment will therefore always perforce involve reference to the other – the cognitive agent and its environment are inextricably entwined and an analysis must appreciate not just one element or the other, but how the two interact and are interrelated.

Such a perspective makes the approach closely related to what I will refer to as 'dynamic sensorimotor' accounts such as those of Susan Hurley (1998; see also Hurley & Noë, 2003) and the theory of perception developed by O'Regan & Noë (2001a) and Noë (2004). Much of this paper is a development of the implications of that theory. But the enactive approach's emphasis on the autonomy of the acting agent and the emergence of cognition in the interaction between that agent's autonomous values and its environment, is not clearly laid out in a dynamic sensorimotor view. Though the enactive approach used here would involve a dynamic sensorimotor perspective, it is not clear that the dynamic sensorimotor theories would adopt all aspects of an enactive view.

The approach used here has a somewhat converse relationship with the work of Natika Newton (1996), who strongly emphasises the goals and values of cognitive agents and their essential (and often overlooked) role in accounts of cognition. Newton, however, also describes cognition as occurring in the embodied, sensorimotor representations in the brain, whereas the enactive approach used in this paper locates cognition only in the actual interaction between the agent and its environment (for a more thorough exploration of the 'where' of enaction, see Di Paolo, 2009). Such an enactive view, as we shall see, provides some dramatically counter-intuitive implications for even basic concepts of mind such as perceptual modality.

## Perceptual Modalities

In traditional, and intuitive, ways of thinking, a perceptual modality is a 'mode of presentation' of a particular stimulus. We might encounter an object in many different ways, but one basic and simple aspect of any perception will be the mode — visual, auditory, haptic etc. — in which it is presented to us as perceivers.

The theory of modalities currently dominant in the cognitive sciences is effectively that originally put forward by Johannes Müller on 'specific nerve energies'. The essential idea is that sensory neurons are responsive to particular forms of energy and it is this specificity in neuronal response that gives a modality its character. The cells of the retinae, for example, are specifically sensitive to light, and not, for example, to temperature or sound, while the neurons of the cochlea and ear are sensitive to the pressure waves of sound, but not to light. This is a physiological explanation of perceptual modality that fits well with our intuitions of clear distinct and basic modes of perception. Each modality, in this view, will have its own dedicated anatomical organs — sight has the eyes, hearing the ears, taste the tongue and so forth.

Modern neuroscience uses a version of this concept that is only slightly modified, and is seen to have been ‘conclusively demonstrated’ (Kandel *et al*., 1995, p. 371). This slight modification extends the organs of interest from the sensory surfaces deep into the brain, where we can find more anatomy particularly sensitive to and apparently dedicated to dealing with, stimulation of a specific kind. These neural organs appear to take sensory information in its raw form at the sensory surfaces and pass it up to more complicated cross-modal processing streams which are the medium of central, multi-modal cognition. It is processing in these more narrowly dedicated systems that is believed to give a perception its specific modal character.

This standard view of modalities has more than a century and a half of support and is pretty much unquestioned in current mainstream Cognitive Science.

Just how many such modalities there are is not quite so clear, though, even in the mainstream. Whatever the truth of the matter, it would appear that we have more than the five traditional modalities. Proprioception, the vestibular sense and kinaesthesia, for example, extend beyond the typical conceptions of touch, and our perceptions of texture, temperature and pain are all apparently underpinned by separate neural systems. But this debate on the number of modalities has not undermined the strong continuity between sensation and perception inherent in the traditional view. In fact, the idea that the senses can be distinguished and counted depends on the traditional model, which make modalities modular — separated by separate organs and neural bases, without cross-talk at least at the low levels of early perceptual processing.

Research in both psychology and neuroscience, however, would lead us to re-evaluate these basic elements of the traditional account of modalities. In doing so, we find ourselves facing some important and intriguing questions about the structure and form of consciousness.

## Questioning Modality Modularity

Several cross-modality illusions indicate the presence of inter-modality influence. This begins to beg questions of the traditional view: if the nerve energies are that specific, then there should be significant insulation between modalities at least at the lowest levels. Modality modularity should be a given. Probably the most commonly known cross-modal illusion is the McGurk effect, first reported by McGurk

& McDonald (1976). The McGurk effect is a phonological misrecognition based on conflicting auditory and visual information. Quite simply, while hearing one phoneme (such as 'ba') repeatedly and watching someone mouth another (such as 'ga') at the same rate, we can misperceive the sound as an intermediate phoneme (such as 'da'). There are a number of further findings about the McGurk effect, including that it occurs not only in adults but in children and that some sounds are better than others for producing it. Because it is associated with the articulation of phonemes, however, it is not the best evidence for truly low level interaction between perceptual modalities. After all, language is at least implied here, and thus the interaction is likely to be the result of learning the speech sounds of our native tongues. Such learned phenomena are usually taken to be higher level psychological phenomena, and will convince very few that we need to doubt the kind of modularity of modalities that the specific nerve energies theory tends to imply (though the McGurk effect has also been shown in pre-linguistic children; Rosenblum *et al.*, 1997). There are other inter-modality interactions, however, some of which are more starkly problematical for a traditional view. Shams *et al.* (2000; 2001) have reported a cross-modal illusion they refer to as the 'illusory flash effect'. In the case of the illusory flash effect, a single brief flicker of a visual stimulus is presented simultaneously (within 100ms) with two auditory beeps. The result is the perception of two flickers of the visual stimulus. Like many illusions, this effect is robust, resisting practice and prior knowledge on the part of the perceiver. Violentyev *et al.* (2005) describe a similar illusion induced by tactile stimuli rather than auditory ones.

Shimojo & Shams (2001) review this phenomenon amongst a range of others and conclude that the traditional modular view of sensory modalities doesn't hold up to proper inspection. Their conclusive title, 'Sensory modalities are not separate modalities', seems a bit strong to our intuitions though — while we perceive the world as a rich weave of sensations, visual, auditory, tactile and more, we do not perceive it as a smear of indistinguishable modes of sensation. Such inter-modal influence, however low level, may still not convince us that our experience of perceptual modalities is not simply dependent on the different ways in which sensory information is transformed by our nervous system as it is processed towards some multi-modal integration in the brain. Striking work by the neuroscientist Walter Freeman, though, makes a deeper and more profound challenge to a simplistic specific nerve energies approach to modality. If Freeman is correct, then at the very early stages of perceptual processing, the details of sensory stimulation are lost, disrupting another aspect of the traditional view, our intuitions about the continuity between sensation and perception.

## Questioning Specific Physiological Bases

### *Neural dynamics, sensation and perception*

Because in the traditional view the particular qualities of a modality somehow arise from the specificity of nervous sensitivity, we should expect unique neural pathways to maintain that specificity of sensitivity into the brain — the modern extension of Müller's specific nerve energies doctrine. The modalities of perception are therefore, under the traditional view, entirely dependent on the character of sensation. However, the work of Walter Freeman and his colleagues on the neural dynamics of perception cast some doubt on this possibility. In particular, Freeman has shown a Rubicon in perceptual processing, a point where in the modalities of smell, touch, sight and hearing, the sensory signal is completely lost, and perception determined not by the pattern of stimulation at the sensory surface, but by its perturbation of the inherent activity of the brain itself — it's not just what the sensory organs are doing, but what the brain is already doing, that is involved in perception. Freeman's work was conducted on rabbits, and began with investigations into the neurophysiology of smell (Freeman, 1991; 2000). Using internal measurement of the olfactory bulb, Freeman showed that the activation of the bulb for similar stimuli in different reinforcement contexts was remarkably different. That is, similar sensations did not necessarily produce similar activations of the olfactory bulb. What is more, the patterns of activation for some stimuli would change when *other* stimuli were given value through reinforcement.

Figure 1. shows patterns derived from EEG readings of the bulb for a rabbit before and after a new reinforcement contingency was introduced the for the scent of amyl. Note particularly that the pattern associated with air changes considerably after the change in reinforcement condition despite the fact that no change in the value of the smell of air was involved (the smell of air, in this case, is simply the background smell of the rabbit's cage). Freeman uses these and related findings to argue that the patterns of response of the bulb to a given stimulus have a lot more to do with the state of the bulb than they do the actual sensory activity evoked by the stimulus. Because of the rich feedback dynamics within the neurons of the olfactory bulb the activity on the sensory surfaces cannot strongly determine the activity of the bulb. Rather, sensory activity perturbs the continuously ongoing activity of

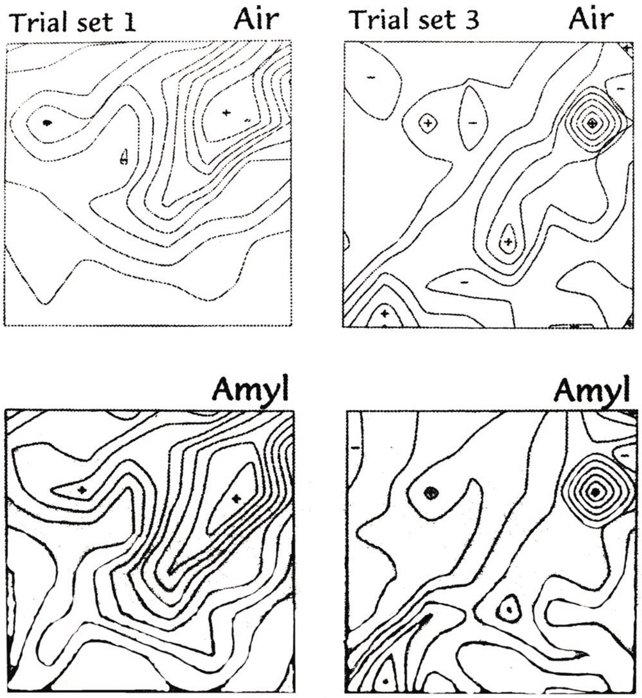


*Figure 1.* Surface EEG patterns for olfactory bulb responses to air and amyl nitrate under different reinforcement conditions (adapted with permission from Freeman & Schneider, 1982, ‘Changes in spatial patterns of rabbit olfactory EEG with condiitoning to odors’, *Psychophysiology*, 19(1), pp. 44–56). Note that no change in reinforcement to air was involved, rather the reinforcement of amyl alters the patterns of activation associated with air.

the neurons. The patterns of activity in the bulb are largely the product of its own activity and history. All sensation does is give it a nudge across its own landscape of attractors. That perturbative nudge must occur in a context of both recent activity of the bulb and the fact that the landscape of its dynamical phase space is constantly changing. The important point here is that what appears to have determined what was smelled by the rabbits in Freeman’s studies was the rabbit’s own history and the history of various reinforcement contingencies. The sensory information *per se* plays a much attenuated role in forming higher level neural patterns. What is processed by the perceptual cortex is categorical, the meaning of the stimulus for the individual, not the raw stimulus itself (Freeman, 1991). Perception does not appear to be strongly determined by the specifics of the sensory level.

Freeman & Barrie (1994) generalised this point to the neural processing of touch, audition and vision, once again in rabbits. What this work shows is a Rubicon between sensation and perception which means that wherever smell gets its odour, it is not from the activity on the sensory surfaces of the nose. We may speak of 'sensory modalities' but *these are not the same modalities in which we perceive*. This is a profoundly counter-intuitive point, but it is one which we will see is implied by the enactive approach to perception and which I am suggesting is implied by the work of Walter Freeman and his colleagues. The neurodynamics of sensation and perception undermine our basic beliefs concerning the modalities of perception. The weak conception of the sensation-perception distinction is challenged by these findings, and deep problems open up for any account of perception which relies heavily on an implicitly held continuity between the two. Note that suggesting that the neural organs specific to different forms of sensory stimulation are widely distributed, rather than anatomically localised in the brain does nothing to attenuate the implications here. What matters to the specific nerve energies view is that the neural pathways are dedicated to the processing of the sensory information, but Freeman's studies show that that sensory information is irrevocably lost once it comes in contact with the complex, multiply connected and autonomous activity that is already occuring in the brain.

What is more, that activity is not insulated from the broader interaction between the animal and its environment. Nothing changes in the amyl nitrate chemical itself during this research — it still interacts in the same way with the olfactory receptors. But the implications of that interaction, its value to the animal and its consequences for what the animal is doing transforms the way the olfactory bulb responds not just to that specific sensory signal, but to others too (as illustrated, the background smell of the rabbit's cage appears affected, despite not being involved in a change in reinforcement value). This suggests that what is going on is not some 'top-down' hypothesis of prediction about what the signal might be, but a coordination of the rabbit's activity with the sensation. Perception is a complex of motivations and on-going activities of the cognitive agent that is contextualised by the animal's needs, its behaviour and its history – it is not the output of a progressive chain of processing, an assembly-line of sensations. Nor is it a neutral prediction of what the sensation might be, given that the nervous system's responses alter not to changes in smells, but to changes in their implication for the animal.

All this said, nor is perception arbitrary or unrelated to sensation. Perception appears to involve coordinating with the world, not imposing any old interpretation upon it. We don't just get to smell what we want to, or even what we expect to. The question arises then, if the character of perception is not determined by the sensory organs being stimulated, what does determine it?

Freeman's research suggests that somehow, it is not the eyes that matter for vision, nor the nose for smell, but the right kinds of interaction between an acting perceiver and the world. This is not to say that normal vision does not depend on the eyes — that much is obvious — but is to suggest that that even in the normal case vision is not just a matter of what happens in the movement of neural activation upstream from the retinae, and it is not *necessarily* dependent on optical input at all. Evidence in support of such thinking, which I believe leaves the traditional account of perceptual modalities wide open for assault by the enactive approach, is the much quoted matter of sensory substitution.

### *Sensory substitution*

In work much cited within the enactive literature (e.g. (e.g . O'Regan & Noë, 2001a,b; Noë, 2004; Hurley & Noë, 2003), neuropsychologist Paul Bach-y-Rita (1972, see also 1984) has produced vision-like experiences in both blind and normal participants, based entirely on sensation on the skin. This remarkable effect is achieved using a system referred to as TVSS (tactile-visual-substitution-system). An array of vibrating contacts are placed on the skin (on the belly, back or tongue) of the person. The activity of the array is controlled by a camera which the person wears on their head. Bach-y-Rita (1972) describes how, given a period of time actively exploring their environment with this system, participants (blind or blind-folded sighted people) have reported experiencing depth, occlusion, basic object recognition, looming and even one form of visual illusion (the waterfall illusion, Bach-y-Rita, 1984). It is crucial that in the learning phase, the participant have control over the camera — *the sensation alone is not enough*. The person wearing the camera and vibrator array must have the opportunity to integrate their tactile experiences into a deliberate and on-going interaction with their environment.

Due to the limitations of skin sensitivity, the resolution of the vibrator array is never very high (about 20x20; see Bach-y-Rita, 1983, for discussion) and as a result the 'vision' it provides is weak indeed. Nevertheless, the experience involved does appear to be visual (or at least distal and spatial) in character, and is certainly not experienced as tactile by the individuals in question. This despite the fact that they can, if they wish, pay attention to the vibrations on their skin. They generally don't pay much attention after the training period, though, and these sensations seem to be relegated to the same level of perception we normally have for the clothing or watch that we might be wearing.

Neural dynamics work by Freeman and his colleagues, along with such phenomena as TVSS, constitute a significant challenge to our traditional conceptions of perceptual modalities. The question then arises of how we might cope with this challenge. If the specific nerve energies account of perceptual modalities is wrong, what kind of account might we use to replace it? An enactive approach to perception, which would endorse such work as that of O'Regan & Noë (2001a), Noë (2004), O'Regan *et al.* (2005) offers us a new means of conceiving the differences between the modalities of perception.

## The Enactive Approach to Perceptual Modalities

The basic conception of perception within the enactive approach is as a sensitivity to the aspects of the world that have meaning for the agent's actions. At first blush, that seems like either a very tall order or a rather vague promissory note. This general idea, however, makes clear the fundamental and irreducible circularity of the relationship between perception and action. We perceive, not as a broad hoovering up of available information from the world, but in order to act. What we perceive is determined as much by what we are trying to achieve as what our sensory systems are capable of. We don't just wait for the world to come to us, nor do we passively accept information available at the sensory surfaces.

A complete account of perception within the enactive approach is beyond the scope of the present paper. We can briefly introduce ourselves, though, to the aspects of such an approach that have a bearing on our conception of modalities.

The dynamic sensorimotor theory of vision put forward by O'Regan & Noë (2001a; see also Noë. 2004) is a strong foundation on which to build a thoroughly enactive approach to perception. They claim that to perceive is to exercise a mastery of the sensorimotor contingencies of a given situation. Sensorimotor contingencies (SMCs) are regularities in the interactions between our bodily movements and our sensations. A very simple example is if you move your eyes to the left, the pattern of stimulation on the retinae moves to the right. Another is if you move toward an object, then the retinal stimulation from points of texture on the object move outwards from the centre toward the periphery of the retinae. The important point here is that the regularities that constitute these sensorimotor contingencies are not simply part of the world, nor part of our sense organs, but arise because both the world and our sense organs are relatively stable, and therefore the interactions between the world and our sense organs will have relatively stable characteristics. When we exercise a mastery of these SMCs, when we can confidently and reliably guide our bodily movements appropriately to perform actions in a given environment, we are perceiving.[1]

In discussing sensorimotor contingencies, O'Regan & Noë (2001a) distinguish between SMCs based on the object and those based on the sensory apparatus. An example of an object-based visual SMC would be that as a viewpoint moves around a mug, the handle appears, distorts and disappears in characteristic ways. An example given of an apparatus-based visual SMC is that as the eyeball moves a straight line produces different forms of curves of stimulation on the retina because of the shape of the eye and retina. O'Regan & Noë (2001a, p. 946) identify apparatus-based SMCs as a significant original contribution, something not present in previous action-focused accounts of perception such as McKay's and Gibson's. The concept offers a strong reminder of the importance of embodiment and bodily interaction with the world.

Finally, O'Regan & Noë (2001a, p. 943) suggest that the difference between sensation and perception is grounded in the difference between object- and apparatus-based sensorimotor contingencies. While it would appear (from evidence such as TVSS) that they are correct in throwing into relief the role and the details of the physical body for perception, the object-based versus apparatus-based SMCs distinction doesn't quite hold up to scrutiny. Take their example of the distortion of stimulation on the retina given movements of the eyeball — as the eyeball moves upward the stimulation on the retina from a straight line in front of the person transforms from a straight line (when the eye is focused directly on the line) to a curved one (when the line is below the point of focus) because of the shape of the

[1] In more recent work, Noë (2004) prefers the phrase 'knowledge of sensorimotor regularities'. Reference to knowledge along with mastery is also made in the original O'Regan & Noë (2001a) paper. Precisely what kind of knowledge is involved, however, remains somewhat unclear, and has been a source of criticism for the approach (e.g. Hutto, 2005; Rowlands, 2005). I have opted for the more skill-focused concept of 'mastery' as I find it the more compelling account. The implications concerning perceptual modalities developed later in the paper do not depend completely on this choice, however, and can also be drawn out from a more knowledge-focused description of this dynamic sensorimotor account of perception.

eyeball. Those distortions of stimulation cannot be reliably contingent without at least implicit reference to objects, or some external thing. In the example given, the reference is an explicit one, to a straight line. And there will always have to be such reference: while the idea of the distinctiveness of apparatus-based versus object-based is sensorimotor contingencies is a useful one to make clear the role of the details of embodiment in perception, it should not be asked to do any real theoretical work. There is a distinction between sensation and perception, but it is not captured by the distinction between object- and apparatus based SMCs. This does not cause immediate problems, however, because O'Regan & Noë (2001a) do not really ask their definition to do important theoretical work.

The sensorimotor aspects of the relationship between the agent and their environment cannot be the whole story. That sensorimotor contingencies enable perception is certainly consistent with the enactive approach, but it lacks any reference to the endogenously driven activities of the agent — the valued, goal-oriented nature of an agent's actions. Perceiving is not sensorimotor contingencies alone, but the exercise of their mastery. It is for this reason that O'Regan & Noë's view is considered a form of 'skill theory' of vision.

So how are modalities constituted within such an account? O'Regan & Noë (2001a) claim that SMCs are governed by relatively stable sets of laws, and that it is these more or less coherent sets of laws that constitute a perceptual modality.[2] As with the contingencies themselves, the laws governing contingencies emerge because of the details of our specific embodiment and the kinds of actions which our environment enables. Some movements will affect the pattern of sensation on the retinae, but have no effect on other sensations, e.g. a movement of the head, which produces dramatic changes in retinal stimulation but will have no effect on the feeling of a wine bottle in our hands. Other movements will affect our sensations of touch (such as moving our hands over that wine bottle), but have no effect on vision. Below is a table, taken from an on-line draft of O'Regan *et al.* (2005) which gives some examples of sensorimotor contingencies for vision and hearing, and how these contingencies can offer an explanation of the differences between the two modalities.

---

[2] O'Regan & Noë (2001a) actually refer to the laws governing SMCs as constituting *sensory* modalities, but draw no explicit distinction between sensory and perceptual modalities. It seems clear, however, that if sensation is the transduction of specific kinds of energies into neural activity then sensory modalities will be physiological and organ-specific – much closer to Müller's view. O'Regan & Noë appear to acknowledge this in pointing out that there are two different kinds of contingencies, apparatus-based contingencies and object-based contingencies, as mentioned.

| Action | Seeing | Hearing |
|---|---|---|
| Blink | Big change | No change |
| Moves eyes | Translating flowfield | No change |
| Turn head | Some changes in flow | Left/right ear phase and amplitude difference |
| Move forward | Expanding flowfield | Increased amplitude in both ears |

*Table 1.* Some sensorimotor contingencies associated with seeing and hearing (taken from a draft version of O'Regan *et al.*, 2005, available online at http://nivea.psycho.univ-paris5.fr/CONS+COG/CC_OREGAN.htm).

While the work of O'Regan, Noë and Myin challenges the intuitive and traditional concept of perceptual modalities and offers us a foundation for a properly enactive approach to modality, the enactive approach demands a somewhat clearer emphasis on the agent's valued actions. O'Regan & Noë's approach is not challenged directly by the evidence presented earlier on cross-talk between modalities, and the work on TVSS is directly invoked by them in support of their theory. Their account of modalities emphasises the interaction between the agent and their environment, perceiving as exploration of the world rather than a passive reception of information about the world. Given such a view, there can be no primitives of perception simply given by physiology — everything is in the interaction. What is more, that interaction is always going to involve some aspect of goal-directed activity on the part of the perceiving agent. It is such activity that drives the exploration of the world in the first place, and gives value to the interaction, the results of that exploration. It is this aspect of the process is not quite unpacked in the detail it needs in O'Regan & Noë's work (nor indeed, in Noë, 2004), and so how it impacts on the way we should consider perceptual modalities from an enactive point of view needs a little more exegesis. In particular, while O'Regan & Noë (2001a) and O'Regan *et al.* (2005) discuss the concept of modalities characerised by sensorimotor contingencies, they do not fully explore some of the more significant implications of a skills-based account.

Remember, again, that within the enactive approach perception, cognition and action, rather than being separable or clearly distinguishable processes, are more like different aspects of the one process of adaptive coping in which a goal-directed agent is continually involved. Modalities are not atomic in nature, but a product of a dynamic process which involves an embodied agent (with goals and sensitivities) and a world.

Within the traditional view, which holds modalities as basic 'modes of presentation', a perception is simply 'presented' to us as either visual, gustatory, tactile and so on. All other aspects of the perception (recognition of the object, interpretation of the event) are deemed to involve some form of further, often inferential, cognitive operation. The enactive approach, however, rails against any such stage-like description of perception, and in doing so transforms (and indeed multiplies) the modes of experience available to us. Cognition is not added to perception after the fact, because it is inherent in the process of perception itself, it is part of what continually initiates, drives and structures the act of perceiving.

An enactive approach to perception thus maintains a strong distinction between sensation and perception. Perception, wrapped up as it is in cognition, action, sense-making, is an activity embedded within, contextualised by, value-driven intentional action. Sensation is on aspect of an embodied agent's interaction with the world, an important part certainly, but not one with any veto or absolute authority as the character of experience.

The demand that perception be understood as structured by the intentional actions of the agent as well as by the sensorimotor contingencies that arise from embodied interaction with the world means that the character of our perceptions is never, and *can never be*, simply visual, or auditory, or tactile. In the abstract perceptual task of a chess-player looking at a chessboard, for example, the player is not only engaged visually with the world, but also engaged 'chessily'. Our perception of the world is imbued with goal-orientation and skilful appraisal, such that the 'mode of presentation' will be as much a matter of the skills we are deploying at the time as it is a matter of the dynamics of sensory stimulation involved. Describing the perception of the board as visual without describing it as chessy would be as incomplete a description of the perception as describing the board without making mention of its visual characteristics.

All perception, then, is inherently multi-modal. Our more abstract skills cannot be deployed without engaging the embodied skills of sensorimotor activity, but those sensorimotor skills are not encapsulated either, they are engaged (their use structured) in the context of goal-directed action. And of course, any normal perception will involve a plethora of such skills — in looking at a chess board I will also be guiding visuomotor behaviours, engaged in a social interaction and be prepared to use a whole manner of other skills depending on the details of my context. The different 'modes' of perception, then, are not sharply separated, but develop and operate together, to the point that they can have quite deep dependencies. Many of us with poor vision will have had the experience of having to put on our glasses in order to hear more clearly, the interaction between taste and smell is celebrated, in a game of soccer being able see the opening and available run is not something that comes without being a skilled player, but once you're good enough, you 'just see it'.[3] As a final example, I offer this suggestion as a phenomenological exercise for the reader. Close your eyes and spin yourself around (either while standing or, more safely, in a swivel chair) until you feel that you are becoming quite dizzy. Stop completely, then open your eyes and fixate an object as best you can. Your head movements and retinal stimulation will be almost identical to a situation in which your continuing maintenance of and sense of your balance was not so disturbed, but what is your perception like? If you are anything like me, you will find evidence that normal 'visual' perception has a surprising dependence on vestibular activity.

Sensory substitution research, where sensory dynamics are transformed by augmenting or compensating technologies, also supports this view of modality. Adaptation to the technology in question only occurs when the person engages in active, goal-directed behaviour. This is the case for TVSS (Bach-y-Rita, 1972), and also for various vision-to-audition devices that have been developed, such as Meijer's (1992) The Voice, which transforms pixel information from a camera into a collection of audible frequencies (see also Auvray *et al.*, 2007, for an analysis of objection recognition and localisation with the device).

Adaptation to the technologies also appears to be activity-specific. For example, participants using the Voice, were more likely to claim that their experience more resembled vision for object localisation tasks, but more resembled hearing for object recognition tasks (Auvray *et al.*, 2007). Auvray & Myin (2009) argue that such sensory substitution research (they prefer the phrase 'perceptual augmentation') implies that perceptual modalities are a not rigid set of

[3] I am *not* a skilled soccer player – my position on the team could best be described as 'handicap', in that I'm placed on the team with better players when we can't make even numbers. I'm sure that all of us have comparable experiences where it is clear that our friends can 'just see' or 'just hear' something – it pops out to them – that is utterly obscure to us.

distinct categories but are more like a space of possibilities. Some areas of this space are more populated than others (the 'traditional' modalities), but intermediate forms of perception are also possible. Some participants in Auvray *et al*'s work with the Voice suggested that rather than experiencing something visual or auditory, their experiences when localising objects with the device was closer to a 'new sense'. In the case of TVSS, there is a similar debate over whether the new experience should be considered 'visual' *per se* (see for example, Prinz's, 2006, criticisms of Noë, 2004, but also O'Regan & Noë's 2001, p.958, p.1013 discussion of TVSS as 'quasi-vision') but this is only a concern if we have good reason to believe that modalities must be modular and discrete, which it appears they are not.

An enactive approach to modalities does not, therefore, make strong or exclusive distinctions between forms of perception. Rather, modalities are areas of stability within this space of possibilities, stabilities that form on the basis of multiple interacting constraints — the sensitivities of the individual agent, their goals, their expertise. Stability does not imply fixation or rigidity, however, and even these enacted stabilities will be in continual development, as our sensitivities, goals and expertise change. Our experiences are stable because this development is very slow relative to our on-going actions.

For many, this may all sound a bit too idiosyncratic. Fine, we might train or develop our visual or auditory skills in various ways, but there remains something fundamental, something basic, which we all share and which characterises some aspects of our perception no matter what our goals, intentions and activities. We may see in different ways in different activities, but it is precisely what is similar across the experience of those different activities that allows us to identify something like vision. Red is red is red, after all, and whether I'm seeing red because I'm checking which bishop my opponent has ready to move against my king, or because I'm judging whether it is the same colour as the car I'm thinking of buying, or whether I'm deciding if I like it, my visual perception is still of the same red.

For an enactive account too, vision is precisely a set of aspects of our experiences that are similar across different intentional contexts, but *without* those different contexts there would be nothing to identify as visual. Identifying a perceptual modality is not basic or fundamental, but is something done in the very same way as other forms of perceptual judgement — with figure against ground. The enactive approach simply makes this interdependency in perception more explicit and drives home the importance of always identifying both aspects of it.

But, a critic might continue, having identified a continuity or consistency across contexts that we can label as the 'visual' aspects of experience, can we not then state what it is that underlies that consistency or continuity? The answer is that we cannot, not in the categorical terms that such a critic would find satisfying. Perceptual modalities are modes of interaction, fuzzy sets of activities structured by skilful action, environmental affordance and biological sensitivity and irreducible to any of those components. Visual perceptions can be identified not by the organs or biology involved, but by the kinds of actions that they make possible, the kinds of activities in which they evoked. It is a characteristic not of the biology of the agent, but of the interaction in which the agent is engaged. That means that any set of skills developed to a similar extent, be they concerned with vision, hearing, driving, chess or social interaction, with have equal claim to being a modality of perception.

This emergent nature of modalities, though, is not necessarily a problem, and does not mean that we cannot effectively identify, communicate and investigate them. We human beings do share many forms of sensitivity and many basic motivations. What is more, our childhoods are long apprenticeships in which the development of our perceptual skills are shaped and channelled so that significant variations or idiosyncrasies in experience are unlikely and unusual. Our social skills play a part in the deployment of our visual, haptic and olfactory ones, and the community in which we exist hold norms of activity and norms of description which constrain our experiences. A delightful and useful exploration of the development and disciplining of experiences which acknowledges the interplay between the object, the acting perceiver and the community of perception is Barry Smith's (2007) discussion of wine-tasting. Smith examines how our experiences are guided and encouraged by a community which both demands and validates certain kinds of interactions with the objects of that community's intentions. Something might be objectively the case about the object, but not discriminated without the community's guidance and endorsement, a case where the modalities of taste and smell are transformed by the development of a new 'wine' modality supported by an interacting society of experts.

While shared biological sensitivities enable this sharing of interactions, they do not determine it. The story of how modalities are formed and stabilised over developmental time is one which involves many more players than biological maturation. Examining this question in the kind of depth it deserves is unfortunately beyond the scope of this paper. Nevertheless, the fundamentally social nature of that development is worth a brief comment. While our bodies provide very basic modes of activity from birth, the most potent of these is social activity, a form of interaction which will become the primary means by which many of our perceptual capacities will be scaffolded, calibrated and put to use. Social interaction is a complex coordination of activity between two agents (De Jaegher & Di Paolo, 2007), involving embodied interaction as well as the mastery of a host of 'social' contingencies. These contingencies are not only sensorimotor, but primarily emotional or affective in character. Due to the pervasive nature of social skill and social interaction, its foundational role in the formation of our minds and its primacy in our interacting with and learning about the world, it deserves at least as much recognition as a modality as any of those with an easily identifiable peripheral organ (see McGann & De Jaegher, 2009, for the first steps in an enactive analysis of social skill and social perception).

Such a claim points to a final potential criticism of the enactive approach to modalities: the idea that despite all of these tortuous accounts of perception, the sensory organ really is some kind of final arbiter on the subject. Other ways of skilfully interacting with the world may be perceptual modalities, but certain modalities (five that we could name immediately) are *special*, because they have sensory organs dedicated and unique to them, while other skills are parasitic on those basic, *genuinely* sensory modes of exploration of the world. But this is just the same organ-centric criticism addressed above and is subject to the same response — all of that dedicated sensation gets us nowhere without reference to the appetite for exploration and action that has been emphasised above. Sensory organs may be vital, but no perception depends on, nor can be explained by, a specific organ alone. Yes, vision is obviously dependent on the eyes, but it is also non-obviously dependent in the normal case on vestibular function, proprioceptive function (eyeball movements, for example), audition and who knows what else? We should certainly take care to ensure that the obvious facts are accounted for, but there is no call to restrict our accounts to the obvious facts. Furthermore, given the active nature of perception, without an account of the intention, the skilled activity, driving the interaction between sensory organ and environment, we will never find meaning in the patter of stimulation at the sensory surface.

If, following the same organ-centric logic, we deny that the case of a skilled TVSS user is a case of vision, however attenuated (none of the vision-specific nerves have been enervated after all), we would still be very hard-pressed to defend the claim that the experience is one of touch. We can identify perceptual modalities by disciplining our discourse and calibrating our experiences with one another over developmental time and through social practice, but we will not be able to draw sharp and conclusive boundaries between them.

## Implications of Enactive Modalities

### *No theory of <Insert your favourite modality>*

One of the most dramatic implications of this approach to modalities is that while it may be possible to produce a general theory of perception (as one aspect of a general theory of mind), we will not ultimately be able to produce a general theory of any particular perceptual modality. There *is* no generic or pure vision, no generic or pure hearing, but specific and task-dependent interactions which involve visual and auditory modalities.

Most research in these various domains will actually be unaffected by this, as applied questions dominate. Where research questions and methods are developed within a particular context they will produce perfectly useful conclusions and generalisations within that domain. But if our ambitions are to produce broad theories of an entire way of being conscious — a theory of seeing, or of hearing — then we are bound to be disappointed. The context-dependent nature of perception is not an interesting quirk of the perceptual system, but a fundamental characteristic of its operation. If we have learned nothing else in the past century and a half of empirical research into the mind and consciousness, it is that context is not contextual.

Far from restricting the ways in which we study perception, though, this description of modality opens up a wide array of new perceptual questions and domains of application.

### *Extending our understanding of perception and enactive cognition*

If the modalities of perception are a matter of the skills through which we interact with the world, then the range of modalities which human beings experience is not just greater than the traditional 'five', but is in fact innumerable. Of course, we have always known that expertise transforms a person's perceptions in appropriate contexts, but for the main the research within Cognitive Psychology has treated that phenomenon as a post-perceptual process of inferential pattern-matching. Experts have tens of thousands of situations represented in memory in some way, and those patterns can short-circuit complex computational demands to provide immediate guidance for behavioural output.

Except, if the enactive approach is correct then no such pattern-matching goes on. Rather, as Hubert Dreyfus (2002) has argued, drawing on Merleau-Ponty's (1962) phenomenology of perception and Freeman's neurodynamical work, experts exhibit transformed interactions in their chosen fields. Rather than matching patterns or considering outcomes their experience of the situation is structured within their expertise, such that rather than the expert responding to an interpretation of the situation, the situation directly evokes or invites a particular response, one that is structured from its foundations by that person's expertise.

The distinction between 'lower' and 'higher' cognition is therefore a misleading one, and we should do away with it. Perception and cognition are not sequential links in a chain but are fused, inseparable and complementary aspects of the process of adaptive coping by an agent in its environment. Within the enactive literature, this process of adaptive coping is referred to as 'sense making' the production of meaning in interaction (Varela, 1991; Weber & Varela, 2002; Thompson & Stapleton, 2008). Cognition is not to be understood as a set of processes which occur in the brain over increasingly more complicated representations but is the on-going activity of an entire organism carrying out its goals within an environment that is, like a fulcrum, simultaneously constraining and enabling that activity. Any account of cognition that hopes to do the phenomenon justice must thus encompass the entire system of the biology of the agent (their embodiment), their goals, expertise and the environment. The consciousness of this system will not be easily partialled into simple or universal components, neither perceptual modality nor cognitive module. Any aspect of an agent's awareness during a particular action will have to be described and interpreted in light of the rest of the system during that same activity.

## Conclusion

Once we make a clear and consistent distinction between sensation and perception, the nature and structure of our awareness of the world around us becomes dramatically affected. Psychological and neuroscientific research have made the concept of a set of dedicated, modular systems for the processing of sensory-specific information problematic. What seems at first an issue in need of innocuous refinement in fact reveals some significant fault lines in the traditional view, and opens a series of questions which I suggest are best answered by recourse to an enactive approach to understanding the mind. Fully grasping the implications of an enactive account of modalities, we can see that the enactive approach need not, indeed cannot, be restricted to the traditionally 'lower' forms of cognition. Nor is consciousness ever a bare or simple phenomenon. Our goal-directed actions and skilful cognition infuse and form consciousness from its very foundations and our understanding of the processes of skilful, on-line embodied coping may be used to transform and guide our investigations into traditionally abstract and 'higher' cognition.[4]

## References

Auvray, M., Hanneton, S. & O'Regan, J.K. (2007), 'Learning to perceive with a visuo-auditory substitution system: Localization and object recognition with The Voice', *Perception*, **36**, pp. 416–30.

Auvray, M. & Myin, E. (2009), 'Perception with compensatory devices: From sensory substitution to sensorimotor extension', *Cognitive Science*, **33**, pp. 1036–58.

Bach-y-Rita, P. (1972), *Brain Mechanisms In Sensory Substitution* (London: Academic Press).

Bach-y-rita, P. (1983), 'Tactile vision substitution: Past and future', *International Journal of Neuroscience*, **19** (1), pp. 29–36.

Bach-y-Rita, P. (1984), 'The relationship between motor processes and cognition in tactile vision substitution', In W. Prinz & A. Sanders (Eds.), *Cognition and Motor Processes*, pp. 149–50 (Berlin: Springer-Verlag).

De Jaegher, H. & Di Paolo, E. (2007), 'Participatory sense-making: An enactive approach to social cognition', *Phenomenology and the Cognitive Science*s, **6**, pp. 485–507.

Di Paolo, E. (2005), 'Autopoiesis, adaptivity, teleology, agency', *Phenomenology and the Cognitive Sciences*, **4** (4),pp. 429–52.

Di Paolo, E. (2009), 'Extended life', *Topoi*, **28**, pp. 9–21.

Di Paolo, E.A., Rohde, M, & DeJaegher, H. (in press), 'Horizons for the enactive mind: values, social interaction and play', In J. Stewart, O. Gapenne, & E. Di Paolo (Eds.), *Enaction: Towards a New Paradigm of Cognitive Science* (Cambridge, MA: MIT Press).

Dreyfus, H.L. (2002), 'Intelligence without representation: Merleau-Ponty's critique of mental representation. The relevance of phenomenology to scientific explanation', *Phenomenology and the Cognitive Sciences*, **1** (4), pp. 367–83.

Freeman, W.J. (1991), 'The physiology of perception', *Scientific American*, **264**, pp. 78–85.

Freeman, W.J. (2000), *How Brains Make up Their Minds* (London: Phoenix).

Freeman, W.J. & Schneider, W. (1982), 'Changes in spatial patterns of rabbit olfactory EEG with condiitoning to odors', *Psychophysiology*, **19** (1), pp. 44–56.

Freeman, W.J. & Barrie, J. (1994), 'Chaotic oscillations and the genesis of meaning in cerebral cortex', In G. Buzs'aki, R. Llin'as, W. Singer, A. Berthoz, & Y. Christen (Eds.), *Temporal Coding in the Brain* (Berlin: Springer-Verlag).

Hurley, S.L. & Noë, A. (2003), 'Neural plasticity and consciousness', *Biology and Philosophy*, **18** (1), pp. 131–68.

[4] The author would like to thank three anonymous referees who made comments on two earlier drafts of this paper, and thanks to whose suggestions the current paper is much improved. I would also like to thank the members of the E- Intentionality seminar group at the University of Sussex, who have discussed the ideas in this paper with the author on more than one occasion.

Hurley, S.L. (1998), *Consciousness in Action* (Cambridge, MA: Harvard University Press).
Kandel, E.R., Schwartz, J.H. & Jessell, T.M. (1995), *Essentials of Neural Science and Behavior* (London: McGraw-Hill).
McGann, M., & De Jaegher, H. (2009), 'Self–other contingencies: Enacting social perception', *Phenomenology and the Cognitive Sciences*, **8** (4), pp. 417–37.
McGurk, H, & McDonald, J. (1976), 'Hearing lips and seeing voices', *Nature*, **264**, pp. 746–48.
Meijer, P. (1992), 'An experimental system for auditory image representations', *IEEE Transactions on Biomedical Engineering*, **39**, pp. 112–21.
Merleau-Ponty, M. (1962), *The Phenomenology of Perception* (London: Routledge & Kegan Paul).
Myin, E, & O'Regan, J.K. (2002), 'Perceptual consciousness, access to modality and skill theories: A way to naturalize phenomenology?', *Journal of Consciousness Studies*, **9** (1), pp. 27–45.
Newton, N. (1996), *Foundations of Understanding* (Amsterdam: John Benjamins Publishing Co.).
Noë, A. (2004), *Action In Perception* (Cambridge, MA: MIT Press).
O'Regan, J.K., Myin, E. & Noë, A. (2005), 'Sensory consciousness explained (better) in terms of "corporality" and "alerting capacity"', *Phenomenology and the Cognitive Sciences*, **4** (4), pp. 369–87.
O'Regan, J.K., & Noë, A. (2001a), 'A sensorimotor account of visual consciousness', *Behavioral and Brain Sciences*, **11** (5), pp. 939–73.
O'Regan, J.K. & Noë, A. (2001b), 'What is it like to see: A sensorimotor theory of perceptual experience', *Synthese*, **129**, pp. 79–103.
Rosenblum, L., Schmuckler, M. & Johnson, J. (1997), 'The McGurk effect in infants', *Perception and Psychophysics*, **59**, pp. 347–57.
Shams, L., Kamitani, Y. & Shimojo, S. (2000), 'What you see is what you hear', *Nature*, **408**, p. 788.
Shams, L., Kamitani, Y. & Shimojo, S. (2001), 'Sound alters visual evoked potentials in humans', *Journal of Vision*, **12**, pp. 3849–52.
Shimojo, S. & Shams, L. (2001), 'Sensory modalities are not separate modalities: plasticity and interactions', *Current Opinion in Neurobiology*, **11**, pp. 505–509.
Thompson, E. (2007), *Mind in Life: Biology, Phenomenology and the Sciences of Mind* (Harvard University Press).
Thompson, E. & Stapleton, M. (2009), 'Making sense of sense-making: Reflections on enactive and extended mind theories', *Topoi*, **28**, pp. 23–30.
Varela, F.J. (1979), *Principles of Biological Autonomy* (Norwalk, CT: Appleton & Lange).
Varela, F.J. (1991), 'Organism: A meshwork of selfless selves', In A. Tauber (Ed.), *Organism: The Origins of Self* (Dordrecht: Kluwer).
Varela, F.J. (1997), 'Patterns of life: Intertwining identity and cognition', *Brain and Cognition*, **34** (1), pp. 72–87.
Varela, F. J., Thompson, E. & Rosch, E. (1991), *The Embodied Mind* (Cambridge, MA: MIT Press).
Violentyev, A., Shimojo, S. & Shams, L. (2005), 'Touch-induced visual illusion', *Neuroreport*, **16**, pp. 1107–11.
Weber, A. & Varela, F.J. (2002), 'Life after Kant: Natural purposes and the autopoietic foundations of biological individuality', *Phenomenology and the Cognitive Sciences*, **1** (2), pp. 97–125.

Paper received May 2008, revised June 2009