

Direct vs. Indirect Perception in the Works of Rock, Gibson and Churchland

Angela Tompkins 200108306

Philosophy 332

J.J. Gibson, Irvin Rock and Patricia Churchland each defend a concept of how we get from stimulus to full blown perception. Gibson presents a theory of direct perception, where percepts are not mediated by any cognitive processing in the brain, rather information is 'picked up' directly from the environment. The notions presented by Rock and Churchland (via David Marr) are indirect theories of perception. Rock's more traditional model is of an intelligent perceptual system capable of using thought-like cognitive processes (inferences, choices, rules) to enrich and make sense of visual stimuli. Churchland finds fault with the idea of a intelligent perceptual system like that of Rock. She views the perceptual system as being merely a clever organization of stupid elements that interact to produce intelligent seeming results.

In "Inference in Perception", Irvin Rock presents the view that perception is intelligent and involves many of the same cognitive operations as does thought. What separates perception from thought, imagination, or dreaming is that "perception is the mental representation of external objects and events that is *based upon or in some way corresponds to* the stimuli reaching our sense organs" (Rock, 1982). As Helmholtz earlier did, Rock also presents a theory of inferential perception, of non-conscious "interpretations" made from the stimulus by way of memory, assumptions, rules and other cognitive operations that lead to perceptions. However, while Helmholtz thought inferences must be made through induction from prior experience, Rock doesn't believe this is a necessary assumption. He is willing to allow the cognitive operations supporting such inferences to be hard-wired via evolution, if necessary. As evidence for his theory of intelligent perception, Rock gives examples of studies where

cognitive processing has been 'tricked' by extraordinary stimuli. Demonstrations of the intelligence of the perceptual system come in the form of illusory contours, and constancy in size perception and orientation perception. He writes about a study of the 'false-depth paradox': where, using a stereogram, a false image of the depth of two overlapping rectangles is given where one rectangle appears to be in front of the other (the image is actually two dimensional). When the viewer moves his head, the rectangle that appears to be further away seems to move in the direction opposite of the head movement. In normal cases of viewing a three dimensional image, a head movement would result in a parallax change. Since no parallax change occurred, Rock says, the mind infers that something must be occurring to compensate for this phenomenon, namely that the object being viewed is *moving* with the head. Rock states these types of illusions as evidence that the mind is following some sort of inference rule. When the perceptual system is presented with two alternative explanations for the same stimulus, it prefers the explanation that best embodies the *non-accidentalness rule*. The situation that best explains the stimulus with the least need to reference coincidence is to be preferred. In the case of apparent motion, it more likely that a similar object appearing in one place and then a next is the same object moving, rather than one coincidentally disappearing and another just like it then appearing, so we see the object as having moved. If perception is intelligent, then we can question why it seems to be impervious to knowledge in the case of illusions. Even when the viewer is aware that he is being tricked, it does not help to dispel the misperception. Rock makes a distinction between that sort of knowledge, and knowledge in the form of inference rules that may be learned unconsciously from prior visual experiences.

In quite a departure from traditional indirect theories of perception, “A Theory of Direct Visual Perception” and *The Ecological Approach to Visual Perception* have J.J. Gibson advocating a *direct* theory of perception: a model of perception which is not mediated by any type of cognitive processing.

Traditional theories of perception have required inferences or cognitive processing because the input stimuli has always been assumed to be inadequate. Cognitive intermediaries must enrich, organize and make sense of such insufficient data. In Gibson's understanding, the stimuli are not nearly so impoverished. He writes that there is a rich amount of data available to perceptual systems in the form of a 'structured array of ambient light'. This optic array is characterized as having different intensities of light traveling in different directions. Taking into account what structures the light, the invariant properties of the stimulus flux can be said to carry information about the world. Gibson writes that “the invariant properties of an optic array specify the permanent properties of the outer world; the variant or changing properties of an optic array specify the changes in the world and the movements of the observer himself. The invariants in stimulation explain the "constancies" in perception” (Gibson, 1962). To show that such an array is necessary for visual perception, Gibson gives two examples of conditions under which there can be no visual perception. In the darkness there is no vision because ambient light is a necessary condition of sight. In the cases of dense fog or an empty overcast sky, as Gibson describes, there is no visual perception because the ambient light is homogeneous, not structured, and so provides no differentiated information to the visual perception system. So having a structured array of ambient light available to a suitable perceptual system is necessary for visual perception. Unlike in the traditional or “Establishment” model (as explicated in Fodor & Pylyshyn, 1981), no inferences or top down processing is required for perception exactly because the input stimulus contains all information necessary for visual perception. Gibson rejected the theory that meaning is provided through long term memory. Rather, he argued that the potential uses of an object are directly perceivable, as 'affordances' that can be picked up by the visual system. The issue then (as illustrated by Fodor & Pylyshyn) is how to adequately constrain the meaning of 'direct pickup'.

Further to his direct theory of perception, Gibson wants to deny completely that perception is

based on the having of sensations, though such sensations may accompany perception (Gibson has no wish to deny sensations altogether, only that they play any role in perception). Instead, he wants to suggest that perception relies only on the direct 'pickup' of information from the structured array of light. What this pickup entails is not detailed.

Later theorists found much to like in Gibson's focus on the higher order variables in stimuli. He was one of the first to see real movement as a vital part of visual perception. Gibson improves upon the proximal stimulus theory (set forth by Koffka, though Koffka was arguing against it) that "things look as they look because they are what they are" (Koffka, 1935). Gibson's optic array is a vast improvement on the idea of the proximal stimulus as a static retinal field.

Where Rock's theory has an advantage over that of Gibson is in dealing with misperception. Gibson does not give a satisfactory way of accounting for non-veridical perceptions. In an inferential view like Rock's, the standard method of accounting for misperceptions is to presume that somewhere along the line, the inferences that led to the wrong perception must have failed or been inadequate in some way. This option is not open to Gibson, as he wants to deny any inferential processing on the stimulus. Because he wants to say that the stimulus contains all the information necessary for perception, it also cannot be that information is somehow missing from the stimulus. As Fodor and Pylyshyn say, the answer Gibson might give is that misperceptions are a result of a "failure to pick up 'all the relevant information'" (Fodor & Pylyshyn, 1981). This leaves Gibson open to questions of how a sample of information can be sufficient for one entity to perceive something, and not sufficient for another. Again, Rock's solution is simple: one must possess cognitive processes different from, and more suitable than the other. Gibson again has no access to that solution, since he wants to deny the need for cognitive processing on the stimulus. The answer may lie somewhere in his conception of

visual learning: that learning is a process of tuning the perceptual system to become more sensitive to information present in the stimulus. One may learn to pick up informational structures that previously went undetected (Epstein, 1989).

Churchland questions the seemingly intelligent perceptual model of Rock. She writes that what may seem to be high-level inferencing and problem-solving in perception may actually just be many unintelligent elements interacting to produce what appears to be intelligent results, in a manner engineered by evolution. She rejects Rock's inferential explanations of visual phenomena on the basis that they are better explained through reductive models like the computational or information-processing theory of vision of David Marr. Marr's view is that different parts of a task in the brain may be handled by different modules, which perform some computations and pass some data onto the next module. That way, a large problem can be divided into many smaller subproblems to be solved individually, in a divide-and-conquer style strategy. The operations of each module are to be specified to an algorithmic level. Furthermore, each module must execute its operation without resorting to bringing information down from a higher-level cognitive center. If it is necessary to hard-wire assumptions about the world to stick to this sort of bottom-up processing, then it must be done by evolution. Churchland writes that if it can be shown in this manner that a perceptual problem *could possibly* be solved in a computational way, it has not yet necessarily been shown that this *is* the way that the brain solves it. However, it is a demonstration that “perception might well be intelligent without the intelligence deriving from a bit of reasoning on the part of the subject” (Churchland, 1982). For her, a reductionist explanation is to be preferred over one like Rock's. Churchland does point out some issues with Marr's model. She remarks that, given the plasticity of the brain and the amount of learning and growing it goes through, there may need to be more inter-modular communication than Marr has allowed for. Also of interest is how learning affects perception (for example, in illusory contour

perceptions, where previous exposure makes further perceptions of such contours more likely). She supposes that there may need to be more talk between higher centers of cognition and lower modules than Marr's bottom-up computational model has supposed.

One objection that has been raised against Rock's theory of perception is this: such a model, using inferences, problem-solving and choosing among several alternatives, seems to require *something* or *someone* to make those inferences. Positing such a homuncular entity is likely to lead to an infinite regress.

What makes Rock's problem-solving approach different from Marr's computational model is its violation of bottom up processing. Since Churchland is ready to concede a certain level of top-down information flow to allow learning to affect perception, the major difference between her and Rock's views is that Rock simply doesn't go far enough in attempting to explain the underlying neurophysiology. Reducing inferences, rules, etc., to their basic neurophysiological actions (be they as Marr theorizes, or otherwise) results in a more fleshed-out, non-homuncular version of Rock's theory that may be more acceptable to Churchland.

The theories of Rock and Marr/Churchland have a decidedly constructivist cast, with the brain constructing representations of the world from an inadequate source of input. Churchland advocates a more or less bottom-up, modular way of accomplishing this, while Rock's method is intelligent; it is monolithic in nature, and allows for some top-down transfer of information from higher cognitive centers to lower ones. Gibson's model, on the other hand, is more involved with the extraction of information which already exists completely in the environment.