Part V

PERCEPTION AND PSYCHOLOGY

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192978 Published online by Cambridge University Press

A Defense of a Non-Computational, Interactive Model of Visual Observation

Bonnie Tamarkin Paller

California State University–Northridge

In a recent paper[1] I argued that observation can be thought of as the result of a properly composed and functioning observation system. In brief, such a system must be composed of a source object and one or more devices which interact with a receptor. If we think of observation as the result of interaction among components are differing functions, we can approach the problem of describing human sense perception as that of describing one type of observation, differentiated from other types by the organic and species specific nature of the receptor devices. Receptor devices are typed by their specific range of sensitivity, which is best thought of as a limit on the type of devices with which the receiver can interact. One consequence of this view is that observation can be achieved in all sorts of ways by all sorts of organic and non-organic systems as long as they satisfy functional conditions. More needs to be said about these latter conditions. But at least, human sense perception is but one general type of observation; human cognition is a further achievement which requires one or more successful observation systems.

An account of observation systems which describes perception as the result of a set of properly interacting components denies some assumptions made by computational theories of perception. For instance, this account denies that receptor devices are typed by the types of signals which can be received from a source. Perception is not literally the result of the successful transmission of a signal from a source to a receiver. Second, if visual observation requires interactive components, then it does does require that a primitive signal get coded, organized with other signals and interpreted by physical mechanisms.

Physical theories of sense perception from Aristotle through our most contemporary computational models attempt to describe necessary components and their interaction. Consider, for example, that current psychological and optical theories agree that human visual perception enables humans to gather a certain range of information from their environment. Visual perception, like all other types of observation, is a systematic achievement which results in information gathering. The sensitivity and interactability of the receptor which lies at one terminus of the active information gathering system limits the type and range of information which can be gathered. Where theories of visual perception disagree is in their descriptions of the components involved in perception. And they can disagree at two points which are connected. First, according to traditional optically-based views, visual perception results from a causal process wherein light photons stimulate photoreceptor cells. These might be called micro-process accounts of perception since the trans-

PSA 1988, Volume 1, pp. 135-142

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192978 Published online by Cambridge University Press

mission devices and the receptors are described using micro-physical and cellular theory. One assumption of such a view is that a complete explanation of visual perception requires causal description at the level of physical micro-processes. It is worth noting that computational theories of perception are not required to make this assumption. Pylyshyn, for example, replaces the micro-physical condition with a looser condition that the physical description be "any description couched in the vocabulary used in physics and the allied natural sciences" (Pylyshyn 1984, p.166).

A second and different claim is that perception requires inference of varying sorts. For example, micro-process accounts require that a pattern of light stimulations be inferred from a sufficient number of individual photoreceptor firings. These and computational accounts require that the organism infer properties of its environment from properties of the patterns of light stimulations. The inference process claim is different from the micro-process assumption and will be separated in what follows.

Yet the inferential claim is controversial. J.J. Gibson, for example, argues that neither sort of inference is involved in perception. This represents a significant difference between Gibson and computationalists such as Fodor and Pylyshyn (Fodor and Pylyshyn 1981, Pylyshyn 1984) who argue that inferential processes are required in order to explain perception.

J.J.Gibson's ecological optics is arguably the leader in promising non-inferential, non-micro-physical views. But his view, as well as any other non-micro-process view, must show that it is possible to provide an explanation which satisfies two general constraints. First, since the explanation will not refer to micro-processes or events, it will face the problem of showing that it is not a redundant explanation. The problem of redundancy for non-microphysical explanation arises in the following way: For any observation which can be causally accounted for at a complex, e.g. ecological, level, there will be another causal description of micro-processes which, seemingly, result in that same observation. Thus, one might argue, the ecologically describable macro-processes reduce to some micro-physical and neurological processes which provide the actual explanation of the observation. The causal description of macro-processes is redundant and unnecessary for an explanation of visual perception. This significant problem is not addressed directly by this paper.

To argue that macro-physical explanation is not redundant is not the same as arguing that perception is genuinely non-computational, but there is some connection. Both Gibson and Fodor and Pylyshyn agree that some properties are directly pick-up or transduced; and they agree that direct pick-up or transduction is characterized as a process which does not require inference. Above, I noted that there are two general sorts of inference that perception might require. The first type of inference is from primitive input to complex perceptual output; the other type is from perceptual content to the organism's environment. If there is genuine macro-physical explanation of perception, then transduced properties may be physically complex. Thus if macro-physical explanation of perception is not redundant, the first type of inference will not be required in order to explain the perception of those physically complex properties which can be transduced. There is an interesting tension here; as complexity of transduced properties increases and the greater the range of physically complex properties which can be transduced, the less inference and computation which will be required on the part of the physical mechanisms.

Thus in order to provide some alternative to the view that visual perception is the result of information processing and computation from primitive neurophysical units, as computational theories suppose, it must be shown both that visual perception can be explained in some way other than as requiring rule governed inference processes and that the transduction of physically complex properties is possible and not redundant.

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192978 Published online by Cambridge University Press

The goal of this paper is to argue for an interactive account of visual perception which shows that perception does not require inference processes. I will not argue here that an explanation of visual perception may describe physical components of complex macro-organization.

1. An Interactive Model of Visual Perception

According to J.J. Gibson (Gibson 1979), the problem with traditional theories of perception is that they describe perception as the result of operations on the discrete deliverances of the senses. There are two things wrong with this picture. The first problem (which will not concern us here) is that this picture assumes that human awareness of the sensa out of which sensations and perceptions are constructed is required for perception. The second problem is that it requires discrete sensa as the primitive units which are processed or computed in order for perception to occur. Gibson's rejection of computational models of perception is interesting here.

What would sense perception be like if computational models of perception were correct? First, there would be discrete and primitive sensory inputs. Gibson's primary target is the somewhat extreme claim that inputs are primitive neuron firings. Each neuron firing is a raw and fundamental bit, some set of which is necessary for a causal sequence which will result in a perception, if the computation occurs correctly. In order for a sense perception event to occur, then, there must be both the neuron firings as the primitive inputs and some physiological machinery which operates on them. The activity of the physiological machinery on the neuron firings would have to be two-fold. First the neuron firings must be processed into an integrated or coherent product and then the product must be interpreted. In the case of visual perception one might suppose that the firings of photoreceptor cells are processed into an image and then the image is interpreted. This latter activity requires stored images which can be retrieved as well as an inference system which can utilize the stored images when it sets about to interpret a current visual image. (Gibson 1979, pp.251-253.)

The problem with computational models is that unless we are to understand such models metaphorically, there must be some one or more computational mechanisms which are both physical and have the ability to interpret the product of integrated neuron firings according to some set of rules. The possibilities are not limitless–either the computational mechanism along with its rules are born into the organism (innate) or there is some physiological apparatus which acquires the rules and hence the ability to render the required computations and interpretations. In either case, the rules endow the physiological apparatus with first, the ability to discriminate and organize the neural signals and second, these rules allow the processor to infer the correct interpretation of the organized neuronal pattern. Micro-process accounts of visual perception need to describe how complex visual events can arise from primitive units such as the firing of photoreceptor cells. These theories must posit a physical mechanism which computes neuron firings into complex neurophysical events.

Gibson rightly finds that this sort of computational model of perception implies a circularity which is especially troublesome. The processor must be able both to integrate the primitive neurological inputs and interpret the integrated product with a high degree of reliability, if not infallibility. These reliable acts require cognitive ability and knowledge, though not self-consciousness, on the part of the computational mechanism. In Gibson's words, "knowledge of the world cannot be explained by supposing that knowledge of the world already exists. All forms of cognitive processing imply cognition so as to account for cognition." (Gibson 1979, p.253) If these abilities are acquired, then some physiological apparatus must have learned the program which enables it to carry out the function of these activities successfully. It is the learning of these reliable inference rules which Gibson rightly finds problematic if not mysterious. To say that these are innate rules will do nothing to mollify the problem for, barring non-natural intervention, the difficulty would be to explain how the organism could learn or otherwise acquire reliable inference rules in the first place.

So, the general problem with any micro-computational view, it seems, is that it requires that there is a set of rules or algorithms in the physical mechanism which governs the computation. This assumption of rules seems not only implausible, but also unnecessary, as I hope to show.

The circularity problem and the problem of rules is nowhere better illustrated than in the case of visual perception. (Gibson 1979, pp.58-61) Traditional optical theories describe rays of light which are focused by the lens to form an image on the retina. Early optical theories attempted to describe how the retinal image was interpreted by the mind or brain. Later optical theories do not suppose that the retinal image is the object to be interpreted but rather the photoreceptor firings caused by the impact of light photons. Both generations of optically-based visual perception theories require the reliable organization and interpretation of physical events. Gibson calls this the fallacy of the "little man in the brain." Where the retinal image is the object for interpretation "there has to be a little man, a homunculus, seated in the brain who looks at this physiological image." (Gibson, 1979, p.60) Even if we agree that there are no knowledgeable homunculi, some knowledgeable something must perform the computational tasks. Later optical theories of perception shift the object to be interpreted from the retinal image to sets of photoreceptor firings which send signals to the brain/interpreter. How *do photoreceptor firings yield percepts? "The currently fashionable answer is, by computerlike activities of the brain on neural signals.*" (Gibson 1979, p.60) Even this more sophisticated theory "has the lurking implication of the little man in the brain. For these signals must be in code and therefore have to be decoded; signals are messages, and messages have to be interpreted."(Gibson 1979, p.61) In either case, reliable rules for interpretation are required.

In their criticism of Gibson, Fodor and Pylyshyn[2] seemingly fail to see that there is any problem in explaining how there can be knowledgeable rules for computational theories of visual perception. Their effort, in fact, is to show that a theory of perception requires rule-guided inference processes and that Gibson's effort to show that perception is not mediated by inference processes fails.

Fodor and Pylyshyn argue that Gibson's theory of direct perception, though intended to be non-inferential, turns out to require inference in at least two ways. According to Gibson, there is information about the environment contained in the ambient light array[3] and perception occurs when this information is "picked-up." Fodor and Pylyshyn argue that any attempt to construe the notion of "information in the ambient array" which is consistent with Gibson's theory will require inference. They see the problem in the following way. Information which is picked-up is information which requires no inference. What sort of information could be directly perceived, i.e. perceived without inference? Gibson would agree that the most plausible candidates to serve as the mechanisms for direct perception are perceptual systems since these respond directly to properties of the ambient light array. Fodor and Pylyshyn argue that even if perceptual systems are mechanisms which respond directly to the properties that they detect, i.e. transducers in their terminology, at best this allows Gibson to conclude that properties of the ambient light array are directly perceived. Significantly, Gibson cannot conclude that properties of the environment are directly perceived, since properties of the ambient light array are not properties of the environment. Properties of the ambient light array may be correlated with, and thus represent, properties of the environment but since transducers cannot respond to the latter, these properties cannot be directly perceived. Thus Gibson must recognize that an adequate theory of visual perception will identify processes which enable the organism to infer from properties of the ambient light array to properties of the environment. Let us call this type of inference "representational inference" since an account of perception which required this

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192978 Published online by Cambridge University Press

type of inference would also claim that transducer states are distinct from but represent properties of the environment.

Fodor and Pylyshyn also find that perception requires a second type of inference. The above discussion of representational inference assumed that properties of the environment are to be inferred from states of the transducer and that states of the transducer are non-inferentially caused by properties of the ambient light array. But there is little reason to think that properties of the ambient light array are as complex as the properties of the environment with which transducer states must correlate in order for representational inference to occur. Rather, it is argued, transducers respond to fairly simple properties of light. To get a correlation between transducer states and properties of the environment, the primitive properties must be constructed into complex properties. Since the primitively detected features underdetermine the complex features they are constructed into, inference processes are necessary to guide the construction. These types of inference may be called "constructional inferences" since they would guide the construction of complex transducer states from simple ones.

The argument for representational inference assumes that because properties of the ambient light array are not properties of the environment, perception requires inference from the former to the latter. I will argue that this aspect of perception no more requires inference rules for its explanation than does the fact that perception occurs even though transducer states are not the same as properties of the ambient light array.

Fodor and Pylyshyn's objections to Gibson are generated by their prior commitment to an account which conceives of perceptions as the results of information bearing signals which are transmitted from the environment through the ambient array (transmission device) to transducers (receptors). Since the signals must be encoded, transformed and organized, interpretation is required at every step of the one-way causal process. Inference processes are posited in order to perform these functions. If this signal-transmission model of perception were not already assumed, inference processes would not be required for these tasks. But a signal-transmission picture is not the only one available. An interactive picture does not require inference processes or the rules implied by them. While this picture is similar to Gibson's resonator picture, there is sufficient difference so that it does not succumb to objections which Fodor and Pylyshyn raise.

Consider a simple observation system comprised of two components: a type of source object (C_1) and a type of receptor(C_2). To function as an observation system these components must be able to interact with each other. Interactability requires that each component must be capable of differentiated states and that state changes in one component can cause state changes in components with which it is causally connected. With the simple observation system described above, the causal connections are from the direction of C_1 to C_2. Where a state change in any C_n causes a state change in any C_{n+1}, the state C_{n+1} is an indicator for the state C_n. Each component must be capable of a set of possible states and the range and depth of this set will mark various types of sensitivity.[4] Clearly, the possible states for C_n need not map 1-1 onto possible states for C_{n+1}. Some interesting features of interactions between components in an observation system can be noted:

1. Where the set of possible states for C_n is much greater that for C_{n+1}, there will be a many-to-one mapping. This means that a more sensitive C_{n+1} would be able to covary more directly with state changes in C_n. Expressed in another way, C_{n+1} would be able to elicit more information from C_n.

2. Limits of sensitivity of C_{n+1} are not constrained by the set of possible states for C_n. A C_{n+1} may be capable of a greater number of states than would ever be actualized by its interaction with C_n.

3. Where each distinct state change in C_n results in a distinct state change in C^{n+1}, the states of C_n covary with states of C_{n+1}. The mapping is 1-1.There can be lawful connections between C_n states and C_{n+1} states in any of these cases.

The interactive model is plausible only if C_{n+1} need not know about its mapping relation to C_n in order to lawfully respond and if it is possible to describe the state changes in a way which does not assume signal transmission.

Color vision is a good example of such an interactive observation system.[5] The primary components of this system consist in light and the visual receptor. A sample of light typically consists of a collection of lightwaves of different wavelengths and intensities (and not of lightwaves of all the same wavelength). Monochromatic or pure light has all lightwaves of the same wavelength. Pure light lies at one end of the saturation spectrum; at the other end of the saturation spectrum lies white light, or light which has no dominant wavelength. Most samples of light are somewhat saturated, that is, almost all wavelengths are distributed near some dominant wavelength.

Different wavelengths produce different color perceptions. Lightwaves from 455nm-485nm produce the perception of blue, 500nm-550nm produce the perception of green, 570nm-590nm produce the perception of yellow and above 625nm, the perception of red is produced. But color perception does not map 1-1 with these wavelength ranges. First, many colors do not lie on the spectrum of possible light wavelengths. Pink, brown, silver and iridescent colors do not correlate with wavelengths. Second, different perceptions can be produced by mixed collections of lightwaves of quite different distributions. For example, if monochromatic green mixes equally with monochromatic red, the result is perceived as yellow even though the spectral yellow of around 580nm is completely absent.

The human visual receptor for color consists of photoreceptors (rods and cones) which connect with bipolar cells, which connect with ganglion cells, which connect with channels. Horizontal cells connect rods and cones. Most color phenomena can be explained in terms of the response of three cone types to lightwaves of different wavelength frequencies. S-cones respond to short wavelength light, I-cones respond to intermediate wavelenth light and L-cones respond to (you guessed it) long wavelength light. Since each and every wavelength causes a unique type of cone to respond, every sample of light causes a unique set of three-cone responses. It is this "thumbprint" of mixed cone stimulations which varies with color phenomena.

For my purposes, this example indicates that a physical event which might be construed as requiring that some inference mechanism compare different photoreceptor responses can also be explained without the rule guided interpretation of a transmitted signal. Of course, we can speak metaphorically. Human visual perception can be described as if it is rule guided and computational activity. But, to quote John Searle, "[y]ou don't need to suppose that there are any rules on top of the neurophysiological structures."(Searle 1984, p.51)

The objections that Fodor and Pylyshyn (Fodor and Pylyshyn 1981) raise for Gibson's theory are not directed to the issue of whether a non-computational explanation of perception is possible. Rather, Fodor and Pylyshyn assume that since perception is computational, the primitives in perception will have to be relatively simple and not the spatially and temporally distributed properties of an optic array, as Gibson claims. But since the claim that perception can be explained in terms of macro-physical properties and processes is not the same as the claim that perception necessarily involves inference, their objections miss their mark.

Fodor and Pylyshyn (1981) do argue that complex properties cannot be non-computationally transduced. Presumably any property which cannot be transduced must therefore be computed, a process which requires inference. I have argued that perception occurs when components of perceptual systems interact and that interaction of this sort does not require inference. Since the state of a component is a property of it and states in general are not limited to those which are micro-physical, their objection concerns the account presented here.

As a first condition for interactability, Fodor and Pylyshyn argue that states which have property P will be able to interact (resonate, in Gibson's language) with a transducer only if there are laws about P. A P-detector is possible only if it is "a device whose states are lawfully related to the presence of P..." (Fodor and Pylyshyn, 1981, p.178) We are in agreement thus far. But while necessary, they find that lawfulness is not sufficient for detectability. In cases where property recognition requires analysis of the internal structure of the property, that is where the property is complex, there can be no property detector. (ibid., p.181)

This claim seems to have two parts. First, where P is a complex property, detection of the simpler parts which constitute P does not necessarily amount to the detection of P. This also seems unproblematic. One can detect that a pot of water is cold at t and subsequently detect that the same pot of water is hot without thereby having detected the comparative property that the pot of cold water has now gotten hot. Second, while Fodor and Pylyshyn provide no general criterion for distinguishing properties which have internal structure from those which do not, *recognition* of the former depends upon *recognition* of the latter. But this seemingly violates their perception/cognition distinction. Recognition is a cognitive activity, and while there is no doubt that inference processes are required for cognition, the very issue is whether they are required for perception.

There is a second problem with using internal structure as a mark of non-transducibility which is that there will be some point of view from which every property has internal structure. This is so even for a seemingly simple property such as color. After all, color never occurs *simpliciter* but as a property of some thing and color, now understood as color-of-that-object is a property with internal structure. If there is no privileged point of view, if there is no one correct answer to the question of whether a property has internal structure, then this criterion cannot be used in any simple way to mark out transducible properties. And if any property can be construed as complex, the use of this criterion threatens to make all properties non-transducible.

Obviously this is not their intended consequence but points to a methodological problem in marking out transducible properties. Fodor and Pylyshyn object that without some constraint on what properties can be picked-up or transduced, then we might as well assume that *whatever* the organism can perceive has been picked-up. If all properties are picked-up, then trivially there is no problem in explaining perception. This trivialization problem finds its photo-negative if all properties are non-transducible, for without some transducible properties, perception cannot be explained at all.

The problem with using complexity as a mark of non-transducibility is not a problem with complexity *per se*, but surprisingly with the method behind its proposal. Rather than suppose that we need to discover the logically necessary conditions which mark transducibility, I suggest that a better way to decide whether a property can be transduced or not is to physically analyze the detector with which it interacts. The question, "To what kinds of properties is the detector sensitive?" is an empirical question.

I have argued for the plausibility of an interactive account of visual perception in which perception does not rely on signal transmission or rule guided processes. This account is intended to go some distance in showing the limits of non-inferential processes, namely that non-inferential perceptual processes must be observation systems. It remains to offer an account which shows that perceptual processes can be used to observe macro-physical properties which does not succumb to the problem of redundancy.

Notes

[1]See Paller, B. (forthcoming).

[2]See Fodor and Pylyshyn, (1981).

[3]Gibson's notion of the ambient optic array as radiant light which has been structured by the physical environment is discussed in Gibson (1979), especially chapter five. "The Ambient Optic Array."

[4]For example, photoreceptors are stimulated by radiation from around 400nm-650nm. This is the sensitivity *range.* While photoreceptors have states which vary with 475nm and 525nm, they do not have differential response to 460nm and 475nm. This would mark an increase in sensitivity *depth.* More about light in what follows.

[5]See D. Falk, D. Brill and D. Stork (1986), especially Chapters 9 and 10, for a discussion of color vision.

References

Falk, D, Brill, D. and Stork, D (1986). *Seeing The Light.* New York: Harper and Row.

Fodor, J. and Pylyshyn, Z. (1981). "How Direct is Visual Perception?: Some Reflections on Gibson's 'Ecological Approach'." *Cognition* 9: 139-196.

Gibson, J.J. (1979). *An Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Paller, B. (forthcoming). "Visual Perception, Observation Systems and Empiricism." *Philosophical Studies.*

Pylyshyn, Z. (1984). *Computation and Cognition.* Cambridge: The MIT Press.

Searle, J. (1984). *Minds, Brains and Science.* London: BBC.