

# Cognitive Critique



## GIBSON'S REASONS FOR REALISM

C. WADE SAVAGE

*Department of Philosophy  
University of Minnesota  
Minneapolis, Minnesota*

EMAIL: savag001@umn.edu

Accepted August 11, 2011

### KEYWORDS

Gibson, philosophy, perception

Over a period of some fifty years, in three books and a mountain of articles, James J. Gibson developed what he called a theory of direct visual perception, a theory which, he believed, *makes reasonable the common sense position that has been called by philosophers direct or naïve realism* (Gibson 1967, p.168). His theory is novel, iconoclastic, and vastly important both for psychology and philosophy. I am as eager as he to defend some version of direct perceptual realism, and as dissatisfied as he with most theories currently in vogue. But I am not yet persuaded that his theory is what he claims it to be, and I would like to present my doubts in this paper. Like Gibson, I will concentrate on visual perception.

In brief, Gibson's theory is that visual perception is not a process of inferring from or organizing visual sensations produced by light falling on the retina, but rather a process in which the total visual system extracts (*picks up*) information about the environment from the light at the eye(s) of the organism as it explores its environment. He objects to theories that base perception on sensations (sense impressions, sensedata) or postulate some operation that converts sensations into percepts. His alternative *information-based* theory of perception assumes that *sensory impressions are occasional and*

*incidental symptoms of perception* and are not required or normally involved in perception.

“It is therefore not obliged to postulate any kind of operation on the data of sense, neither a mental operation on units of consciousness nor a central nervous operation on the signals in nerves. Perception is taken to be a process of information pickup” (Gibson 1967, p.162).

Gibson believes that the process in which the visual system picks up or extracts information from light stimulation can be understood only in terms of what he calls *Ecological Optics*. Physical, geometrical optics reduces objects and surfaces in the environment to points or atoms, and posits rays of light issuing from these points that produce patterns on the retina of the eye of a seeing organism. Since a cross section of any pencil of such rays has no form or pattern, it is impossible to understand in terms of physical optics how light stimulation can contain information that specifies the objects and surfaces of the environment. Ecological optics, in contrast, maintains the following.

“At any point in a medium there will exist a bundle of *visual solid angles* corresponding to the components or parts of the illuminated environment. The *faces* and *facets* of the reflecting surfaces are such components; what we call *objects* are others. Note that the bundle of solid angles postulated above is not the same as a pencil of rays, as in geometrical optics. The cross section of a solid angle always has a *form*, no matter how small, whereas the cross section of a ray is a formless point. And the cross section of a bundle of solid angles always has a pattern whereas the cross section of a pencil of rays does not” (Gibson 2002, p. 81).

Each of the points in the medium is a point of potential observation, and a light array “can be said to exist at a point of observation whether or not an eye is stationed at that point. In this respect the array is quite unlike a retinal image, which occurs only if a chambered (vertebrate) eye is put there and aimed in a certain direction” (Gibson 2002, p. 81).

An organism moving in space successively occupies a group of these observation points. The resulting transformations in the op-

tic array at its eye(s) specify the movement of the organism. Such transformations are reversible, since the organism can retrace its path in space, and they are distinct from irreversible transformations in the optic array, such as those that occur when the object in the environment rather than the observer moves.

Describing visual stimulation in the ecological-optical manner has the advantage of explaining the normal veridicality of visual perception. For under this description the light array at the retina contains information that specifies the objects and the features of the environment. For virtually every perceivable object or feature of the environment there is a variable in the optic array that corresponds to that object or feature. Consider depth for example. For centuries psychologists and philosophers have puzzled about how we are able to see a third dimension when the retinal image that gives rise to vision is a two-dimensional mosaic of stimulation. Gibson believes his theory removes this puzzle by showing that it rests on a false problem. He argues that in perceiving depth we do not perceive the third dimension of the Cartesian coordinate system: this dimension is not *a phenomenal fact of perception* (Gibson 2002, p. 85). Rather, *we perceive the layout of the environment* (Gibson 2002, p. 85), and this layout is specified in the optic array available to the eye.

For example, the variables in the optic array that specify perceived depth are gradients of texture, like those we can see in the picture of a plowed field, where the pictures of clods at the bottom are larger and farther apart than the pictures of clods at the top.

“The so-called *cues* for the perception of depth are not the same as the *information* for the perception of layout. The former are called *signs* or *indicators* of depth, or *clues* for an inference that depth exists in the world. The meaning has to be learned by association. They are sensations in the visual field of the observer, noticeable when he inspects the latter. The available kinds of information are *specifiers* of layout, not signs or indicators of clues. They have to be distinguished or discriminated, but their meaning does not have to be learned by association” (Gibson 2002, p. 86).

The process of picking up information from the optic arrays available to the eye consists in *extracting invariants* from these arrays. Suppose, for example, a penny is standing upright on a table and that I move around the table keeping my eyes fixed on the penny. I will pass through a series of observation points, the optic

array for each differing from the one that preceded it. One of these arrays – the one defined by an observation point directly in front of the penny – will contain a circular pattern. Another will contain a slightly elliptical pattern, another an even more elliptical pattern, and so on as I move around the coin. Although these arrays differ from one another, they have something in common: each of them contains the closed curved pattern defined topologically as an ellipse (a circle being an ellipse with equal axes). This common feature is the *invariant* in the varying optic arrays. When the visual system is *tuned* to pick up invariant features, it perceives a circular coin: a penny. When tuned to pick up variant features it perceives an elliptical shape. In general, the perception of enduring objects and permanent features of the environment consists in extracting invariants of stimulation from the optic arrays available to the eye.

“The invariant properties of a changing stimulus array correspond to the invariant properties of the environment. What about the *variant* properties? The child must learn to separate the invariants from the variants more and more precisely as he grows up, and to focus his attention on them if he is to learn more about the world. He typically does so by exploration, that is, by changing the stimulus patterns on his eyes and skin so as to isolate what remains unchanged” (Gibson 1967, pp 165-166).

In order for the visual system to pick up the invariants of optical stimulation, these must be *attended to*, and the ability to attend *may depend on the maturation of the system, and on practice in looking, and even on the education or training of attention* (Gibson 2002, p. 86). Gibson consequently does not deny that we learn to perceive. But he does deny that such learning consists in associating present visual sensations with remembered visual sensations. Perceptual learning consists in *attuning* the visual system in such a way that it *resonates* to the invariant features of the optic arrays available to the eye.

It is now possible to present Gibson's view of the role of sensations in perception. Sensations, according to him, are inessential to perception and are merely its occasional symptoms, appearing only under special conditions of attention. Just as an observer can attend to the invariants of optical stimulation, so she can attend to the variants. (The invariants of stimulation are for the most part variables of higher order than the variants.) To employ the coin example again, I can attend to the common, invariant spatial feature of the optic

arrays produced by the coin (their general elliptical shape), or I can attend to their distinct, variant features (varying from round to specifically elliptical). When I attend to the invariants I do not have visual sensations: I see a round penny. Only when I attend to the variants in the optical stimulation are visual sensations produced in me: the sensation of a round patch followed by the sensation of a slightly elliptical patch and then a more elliptical patch, and so on. Since the two types of attention do not occur at the same time, visual sensations are not produced in normal perception.

Furthermore, Gibson would have us understand, the visual sensations that arise by attending to variants of the optical stimulation are not the intervening, private, mental entities postulated by philosophers and many psychologists of perception as the basis for the perceiver's inferences about the external world. For the changing pattern in my visual arrays as I move round the penny indicates that my body is moving, and so attending to these changes constitutes perception of the movement of my body, i.e., proprioception (Gibson 2002), not perception of a private visual sensation. Consequently, not only are private sensations normally not involved in visual perception, the sensations that sometimes are involved are not private sensations but rather perceptions of the perceiver's body, proprioceptions.

Our sketch of Gibson's theory now complete, we turn next to questions and criticism.

1. Gibson's assertion that visual sensations arise from attending to variants of optical stimulation is troublesome for two reasons. The first is that we often perceive states and changes of objects, including our own body, by attending to variants of stimulation produced by those changes. When I see the coin revolving (instead of myself) I attend to its changing elliptical shape, although other changes – in illumination, edges, etc. – offset these. When I see myself revolving on a barstool, rather than the room around me, I attend to changes in the flow of visual and, of course, cochlear stimulation. Gibson calls such bodily perception *proprioception* and he says that although proprioception accompanies perception it is not the same thing (Gibson 2002). However, awareness of one's own body seems enough like awareness of objects in the environment to classify both as species of perception. Since perception seems to consist in attending both to variants and to invariants of stimulation, Gibson's explanation of sensations seems incomplete, if not simply incorrect. One possible reply is that proprioception consists in at-

tending to invariants of optical stimulation of a different sort than those, attention to which constitutes perception of external objects. But this would require a theory of what the difference consists in.

The more obvious difficulty in Gibson's assertion that visual sensations arise from attending to variants of stimulation is that perception of the external environment often consists in attending to variants of stimulation. Consider the case in which an inflated balloon slowly decreases in size because of a hole in its wall. An observer looking at the balloon will see a deflating balloon; and his seeing it will involve attending to a decrease in the size of a circular pattern in the optic array at his eye. In this case his seeing the balloon consists in attending to a variant-not an invariant-of stimulation.

One possible reply is that the variants of a deflating balloon are of a different type than those of a receding balloon. But what is the difference? Another possible reply is that it is the ratios between variants – not simply variants – that are attended to when a perceiver sees a deflating balloon. But ratios between which variants?

It seems that Gibson can meet these objections only by conceding that normal perception can consist in attending to variants, as well as invariants, of stimulation. But such a reply is inconsistent with his theory that visual sensations are produced by attending to variants of stimulation. If he adheres to this theory, then he may be forced to admit that perception sometimes involves having sensations; and then it is more likely that perception *always* involves having sensations, some or most of which are not attended to, at least not consciously. It may be that his only possible solution is to maintain that, although abnormal perception (afterimage perception, drug induced perception, etc.) may involve having sensations, normal perception – perception under normal conditions – *never* does.

2. The question above leads naturally to another. It is not clear why Gibson wishes to deny that perception involves having sensations. What would be objectionable in the following modification of his theory? The optic arrays at the eye give rise to visual sensations in the observer: a single visual sensation if – as is virtually impossible – the eye and the observer remain completely motionless, a sequence of such sensations if, as is usually the case, the eye or the observer is in motion, or if changes are occurring in the observer's environment. The observer can attend either to the variant features or to the invariant features of his visual sensations. When he at-

tends to the latter he perceives objects and features of the external environment; when he attends to the former he perceives (proprioceives) his own body.

The only difference between this theory and Gibson's is that the process of information pickup occurs at a large stage in the process of perception, the stage at which sensations are produced. The process in all other respects can be described exactly as Gibson describes it. He might complain that the alternative theory I have proposed makes perception less direct than does his own. But there is merely a difference in degree, and not a difference in kind between the two theories. In neither theory is it asserted that the observer *directly* attends to the objects and features of his environment, or to his own body. On Gibson's theory the observer directly attends to optic arrays; to visual sensations on my modification of his theory.

3. A concern parallel to the previous one is that it is not clear why Gibson wishes to deny that information about the world is contained in *retinal* images and extracted from these images by the nervous system. He might suppose that the retina does not contain all the information registered in the optic array at the eye, so that action based on it would require a risky inference by the central cortex. But it would seem that the retina must be capable of registering all the information that is in fact extracted and employed by the perceptual system. The retina may not be able to register all the features contained in the optic array; but it *must* be able to register enough of the contours, textures, and other invariant features of the optic array to make normal perception of the environment possible. The same point applies to the neural image projected by the optic nerve into the visual cortex of the brain. The image must contain all the information provided by prior processes that is required for the perception of the subject's body and environment of which she is capable.

Perhaps Gibson fears that if his theory is modified to say that it involves processing retinal images by the brain it will lose its *direct* character. But if direct visual perception is assumed to be perception that involves neither a retina or some other system of photoreceptors, nor a brain or some other system that employs (processes) neural signals from the photoreceptors in order to produce actions or modifications of the creature in question – animal, bird, insect, etc. – then direct visual perception does not exist. Often when Gibson claims that perception is direct he means that it is not mediated by

(conscious?) sensations. This claim may be true even if perception involves the processing of retinal images by a brain.

4. It is also not clear why Gibson denies that perception is a process in the brain. Probably his reason – or part of it – is that he does not regard perception as a process of *passively* receiving and processing signals sent to the brain. Perception, on his theory, is an active process in which the brain not only receives signals but also sends messages to the eye muscles and other muscles of the body so as to cause the organism to explore its environment and obtain further optical information. Perception is thus an active, circular process consisting of receiving and sending messages and extracting information from the multitude of stimuli thus produced. All this may be true; but none of it implies that perception is not a process in the brain. Since the brain is receiver, sender, and processor of visual excitation, it's appropriate to say that perception – albeit an active process – occurs (mainly? essentially? centrally? at least partly) the brain.

In any case, Gibson has failed to explain what role the retina, optic nerve and brain have in perception. He admits that they have a critical role in sensation (the production of sensations), but some of what he says seems to imply that these essential systems have no role in perception.

In one paper on the topic he says that *information does not consist of signals to be interpreted* (Gibson 2002, p. 79), that *vision... [is] not a photographic process of image registration* (Gibson 2002, p. 84), that *we do not have to speculate about how the brain could store the sequences of images transmitted to it* (Gibson 2002, p. 84), and that *the size, the form, and the color, of the image impressed, on the retina...are not relevant to [perceived dimensions]* (Gibson 2002, p. 87). In summary: *this theory of vision asserts that perception is direct and is not mediated by retinal images transmitted through the brain* (Gibson 2002, p. 88).

These passages seem to contain the startling suggestion that the retina, optic nerve, and brain have no role in perception! But surely no such unscientific suggestion could be intended.

Perhaps Gibson fears that if his theory is modified to say that perception is the process of picking up information from retinal stimulation or that the information, wherever obtained, is processed



by the brain, then it would lose its *realist* character. For he argues that since the optic array is public and contains public information, different perceivers can see the same things (Gibson 1967, pp 170-171). The apparent implication of this argument is that information obtained from a public optic array is public, whereas information obtained from retinal images or brain images is not public. Apparent or real, the implication is mistaken. In the sense in which different observers can sample the same optic array by successively occupying the same point of observation, they can sample the same retinal image if their retinas are similarly constructed. A realist theory of perception – a theory according to which distinct perceivers can perceive the same public, *real* object – can thus be based almost as comfortably on information extraction from retinal images as on information extraction from optic arrays. Note that any viable theory of perception must concede that if different perceivers have different retinas they will pick up information from identical optic arrays in different ways. Human visual perception cannot be so direct as to be possible without something like eyes having something like retinas (natural or artificial), nor so direct as to eliminate individual differences among perceivers and the possibility of perceptual abnormalities.

5. Gibson distinguishes his theory of perception from so-called *associationist* theories, which hold that we learn to perceive objects and features of the environment by associating the sensations produced in us by those objects and features with other sensations. Such theories assume that we remember our sensations, and remember that when they are of a certain type, an object was indicated and found to be present. From then on we infer the presence of the object from the occurrence of the sensations. Consider a penny lying on a table. As I move round the table, keeping my eyes fixed on the penny, a sequence of sensations is produced in me. One of these contains a round patch, another a slightly elliptical patch, another a more elliptical patch, and so on. I discover that, when the sequence of sensations is produced in me I receive a *roundish* tactual sensation by extending my hand and grasping the penny. From then on I infer that when that sequence of visual sensations is produced in me, there is round penny on the table.

According to Gibson's theory – at least as he sometimes presents it – no such process is involved in my learning to perceive the penny. Instead, I simply extract the invariant, closed-curve pattern in the various optic arrays at the various points of observation

of the penny and thus see a single, persisting object. But it seems unlikely that I could have acquired the ability to extract the invariant without a period of learning that requires remembering that the arrays are produced in a certain order. Thus it seems, on Gibson's theory, as well as the sensationist, some process of *association* of arrays *is* required for me to learn to attend to the invariants of optical stimulation and infer the presence of the penny. If this is correct, then Gibson's theory of perception is no less associationist than the theory he criticizes. Perhaps he would say that what is associated in his theory are optic arrays, not sensations. But then his objection is not to association, but to the view that it is sensations that are associated. And when we recall the earlier objection that his theory could be restated to say that information is extracted from visual sensations produced by the optic arrays, the difference between his and the usual associationist theory seems insignificant.

6. Gibson says that on his theory perception does not consist in inferring from or organizing visual sensations or sensedata, and he seems to imply that, more generally, it does not consist in inferring from or organizing *anything*. It is difficult to see how this can be correct. On his theory a perceiver extracts information from the optic arrays at her eyes, and she extracts information from this optical stimulation. What the observer attends to is light; the information she extracts is contained in light. Now light and patterns of light and dark are distinct from objects in and features of the environment that produce those light patterns. Consequently, it seems that the observer must make something like an inference in order to perceive the environment, must in some sense infer that the light patterns she attends to are produced by objects and features of the environment.

Perhaps *inference* is not the right word here. Perhaps we should say that the observer *interprets* or *organizes* the information contained in the optic array. But arguably some such processes must take place in order for him to perceive. In either case, Gibson must at least admit that some processes of *decoding* the information encoded in visual arrays is required. This process is a sort of interpretation. Furthermore, on Gibson's theory perception involves not merely the extraction of information from visual arrays, but also the feedback in neural loops required to shift one's gaze and position, thus supplying additional visual arrays from the object to assist in disambiguating information about the object. Consequently, the complete perceptual process is just as complicated on Gibson's theory as on those theories that posit inference and interpretation.

And similarly the process in some sense *intervenes* and makes his theory hardly less *direct* than theirs. Less cognitivist, perhaps, but not less direct.

7. The objections above can be gathered up and summarized in the following abbreviated form. Gibson's theory has no better claim to be a theory of direct perception than some of those he rejects. Sensationist or sensedata theories do not deny that we see the external environment. But they do deny that we *directly* see this environment, and they affirm that we are directly aware only of our own sensations or sensedata. Gibson therefore calls them theories of indirect perception and contrasts them with his own. He says,

“The doctrine that all we ever experience directly is the flow of sensedata implies that our experience of objects and events is indirect. Perception is mediated by sensation... For this doctrine we now have a substitute. There can be direct or immediate awareness of objects and events when the perceptual systems resonate so as to pick up information” (Gibson 1967, p. 168).

But on Gibson's theory perceivers are not directly aware of their environment. In the case of vision they are directly aware of optic arrays at their eyes. Sensationist theories hold that perceivers are directly aware of their own sensations; Gibson holds that they are directly aware of (directly attend to) their optic arrays. The latter may in some respect be more public and external than the former; but they are not the environment and therefore are not what is seen in ordinary sense, seen indirectly. Gibson therefore should, to be consistent, classify his own theory as one of indirect perception.

The comparison is instructive. For if Gibson's theory is not one of direct perception, then what would qualify as such a theory? There is a powerful inclination, which both philosophers and psychologists have exhibited, to say that perception is direct only if it does not depend upon any physical or physiological process connecting the perceiver with the perceived entity, only if perception is unmediated by such processes. If this is the proper analysis of direct perception, then it is clear that normal visual perception cannot be direct on any scientifically acceptable theory of perception, neither on Gibson's theory nor on any other theory compatible with the science of vision. For it is one of the most firmly established facts of vision science that normal visual perception occurs only if light from the seen object enters the eye of the observer and stimulates the rods and/or cones of his retina, which in turn stimulate

his optic nerve, and thus produces brain processes in his occipital cortex. (The visual perception of afterimages and hallucinations is abnormal and so does not conform exactly to this generalization.) Even on the ancient Greek theory that visual perception consists in a filmy copy of the perceived object entering the eye and then the brain, visual perception is indirect, since it is not the object but rather a copy that is directly perceived.

How then should we define direct perception? The feature of sensationist theories of perception that lead us to classify them as theories of indirect perception is their supposition that perceptual processes contain an act of immediate awareness of something other than the perceived external object, sensations, for example. Similarly, the feature of Gibson's theory that prompts us to regard it as a theory of indirect perception is his supposition that the perceptual process includes awareness of the optic array. If this supposition is replaced by some other, then perhaps we can say of either theory, that it is a theory of direct perception. Now it is difficult to drop the supposition in the case of sensationist theory; for if an observer has sensations, then it seems that he must be aware of them, at least unconsciously. But the supposition is more easily dropped in Gibson's theory. Instead of saying that the observer is aware of, or attends to, the optic arrays at his eyes, Gibson can say that the visual system *samples* the optic array and *extracts* information from them, thus using terms that do not so obviously suggest awareness or attention. It will then seem that on his theory it is features of the environment – not sensations or sensedata – to which the observer attends, of which she is aware.

But consider: by the means just described even a theory of perception which maintains that the observer attends to his retinal images can be converted into a theory of direct perception. For *attending to the retinal image* one simply substitutes (*nonconscious*) *sampling of and/or extracting information from the retinal image*. Such a theory will be hardly less direct than Gibson's. There is, consequently, no reason, at least not from a preference for direct realism, to prefer a theory that posits the extraction of information from the optic array to one that posits the extraction of information from the retinal image or from sensations that accompany the retinal image. By the device suggested a theory of perception that maintains that the observer unconsciously attends to his sensations and unconsciously infers external entities from these becomes a theory of direct perception by revising it to say that the observer

unconsciously *extracts information* from his sensations. The only theory that is more direct realist than these is, again, that of the ancient Greek philosophers, on which a copy of the perceived object enters the brain. And even then it is not the object, but only a copy, that is directly perceived.

It thus appears that the *direct realism* of Gibson's theory consists primarily in the analogies and tendentious language he employs. He defines the visual stimulus, not as a brain, optic nerve, or retinal process, but as a set of optic arrays, thus moving the stimulus *outside* the organism and bringing it *closer* to the thing seen. He says that the active senses are analogous to *tentacles and feelers*, thus suggesting that the perceiver is in *direct contact* with the thing perceived. He calls attending to the optic array *information pickup and extraction*, suggesting that sensations are unnecessary. He says that perceptual learning is the *attuning* of the perceptual system to the invariants of stimulation, and that perception occurs when the system *resonates* (like a tuning fork) to those variables. This language gives his theory a direct realist feel. But if – as it seems – the alternative theory, that vision consists in the processing of retinal images by the brain, can be converted into a direct-realist theory by similar devices, then Gibson's theory is *philosophically* less significant than he and some of his expositors have supposed, its *direct realism* consisting chiefly in the analogies and suggestive language employed in its presentation. Its advantages for devising experiments and suggesting alternative psychological theories – whether *direct* or *indirect* – may nonetheless be substantial. It may well be the best *psychological* theory of visual perception yet devised.

## AFTERWORD

The first version of this paper was written long ago, when I was working hard on Gibson's theory of perception and teaching courses on it, just after his *Senses Considered as Perceptual Systems* appeared. It is not much influenced by his *The Ecological Approach to Visual Perception*, which appeared later and which I have had difficulty assimilating. Producing this second, revised version of the paper has led me to think that my interpretation of Gibson's theory must be corrected in the following way.

On Gibson's theory normal visual perception consists in extracting information from the optic array. It may or may not involve attending to the entity perceived. It does not involve attending to the

retinal image thereby produced, nor to any sensation or sense-datum that might be produced, not even one corresponding to the retinal image. A shift in attention is required to attend to the retinal image, and when this occurs visual sensations (sense-data) are produced. But visual sensations do not occur in normal perception.

It is a mistake to suppose – as I do in the paper above – that Gibson attempts to make perception (more) direct by removing from it the intervening process of attending to retinal images or to sensations or sense-data corresponding to these images. On his view sensation is a process entirely distinct from and incompatible with perception and hence one that could not be a component of perception.

## REFERENCES

- Gibson JJ (1950) *The perception of the visual world*. Houghton Mifflin Company, Boston, MA
- Gibson JJ (1966) *The senses considered as perceptual systems*. Houghton Mifflin Company, Boston, MA
- Gibson JJ (1967) New reasons for realism. *Synthese* 17:162-172
- Gibson JJ (1979) *The ecological approach to visual perception*. Houghton Mifflin Company, Boston, MA
- Gibson JJ (2002) A theory of direct visual perception. In: Noe A, Thompson E (eds) *Visions and mind: selected readings in the philosophy of perception*. MIT Press, Cambridge, MA, pp 77-89