

The view that recognition is the only goal of visual perception

[Psychology](#)



This piece of writing will firstly draw upon two computational theories of visual perception and a number of neuropsychological studies in an attempt to evaluate the view that recognition is the only goal of visual perception. It will go on to outline Gibson's ecological approach to perception and evaluate the degree to which this perspective opposes the idea that perception is for recognition. Finally, evidence from the field of cognitive neuroscience will be presented and the wealth of information it provides on the topic will be discussed.

According to Epstein and Rogers, (1995) perception is the set of processes by which individuals recognise, organise, and make sense of the sensations they receive from the external world. It is the modality of visual perception, that is, perception by means of the eyes, and more specifically visual recognition which is of particular interest within the context of this work. Visual recognition has fascinated psychologists for decades, and can be described as the matching of the retinal image of an object to a description or representation of the object which is stored in the memory (Farah & Ratcliff, 1994).

In order to make sense of the way animals perceive their world, cognitive psychologists adopt a computational perspective, suggesting visual perception is mediated by internal processing mechanisms. One theory of visual perception in recognition which exemplifies a computational perspective is that of Marr (1982). He proposed a theory in which a series of explicit computational stages contribute to retinal stimulation, which is gradually developed into the perception of an object.

Four modules or representations which Marr termed 'sketches', help the viewer to elaborate the structure of light stimulation sensed from the environment into a percept. Raw and unrefined primal sketches of the light structure in the environment are gradually built up using features in the environment such as edges, blobs and terminations, into 'full primal' sketches, and later '2.5D' representations are derived which include information about depth and distance (Braisby & Gellatly, 2005). The final '3D' stage of the computation allows the observer to identify an object via the use of stored internal representations of the world in memory.

A number of studies have attempted to test elements of Marr's (1982) theory. Marr and Hildreth (1980) for example attempted to test the raw-primal sketch aspect of the theory using a computer programme. They reported that when the programme was applied to images of everyday scenes, the algorithm was to a degree, successful in locating the edges of objects. Furthermore, in an attempt to test the three-dimensional aspect of Marr's theory, Enns and Rensick (1990) found that test subjects in their experiment were able to extract and make use of 3-dimensional information in a series of grouping activities.

Exactly how 3D representations are derived from a 2.5D sketch is complex. According to Marr and Nishihara (1978), three dimensional objects can be described by one or more generalized cones, that is, any 3D shape that has a cross-section of a consistent shape throughout its length (Braisby & Gellatly, 2005), such as a cylinder or rectangular block. The object that is to be recognised is recovered by recovering the generalised cones of which it is

composed; this is done by using several sources of information, such as the major axis of the generalised cones, and the contour outline of the object and its concavities.

One particular problem then in recognising objects arises when these features are difficult to locate. In Lawson and Humphreys (1996) study, findings suggest that whilst the rotation of an object does not affect the test subjects ability to recognise objects, such ability is affected when the major axis were harder to locate, such as when the object is tilted towards the viewer and the major axis appear foreshortened. Additionally, a number of neuropsychological studies also provide evidence to support Marr and Nishiara's (1978) theory of object recognition.

For instance, Warrington and Taylor (1978) found that patients with right hemisphere damage could recognise objects when presented in a typical view, but failed to do so when objects were presented in an unusual view. Furthermore, these patients were unable to determine whether two photographs were the same object when presented in typical and unusual views, thus suggesting that the principle axis is a vital component in object recognition. Biederman's (1987) recognition-by-components model or geon theory is another dominant and widely accepted computational theory of object recognition.

According to this theory, objects are composed of a series of geons, or three-dimensional-shape concepts such as a block, cylinder, funnel or wedge.

Biederman suggests that these simple geometric 'primitives' can be combined to produce more complex ones. Bierderman and Gerhardstein

<https://assignbuster.com/the-view-that-recognition-is-the-only-goal-of-visual-perception/>

(1993) termed this distinctive arrangement of parts a 'geon structural description', which is extracted from the visual object and is matched in parallel against stored representations of the 36 geons that make up the basic set.

The identification of a visual object is thus determined by whichever stored representation provides the best fit. Biederman (1987) attempted to strengthen his claim by presenting test subjects with an arrangement of two or three geons that had been occluded, rotated in depth, or extensively degraded. His findings report that test subjects were able to identify degraded versions of the objects in a condition where the critical features were still present but not in the test condition where the critical features had been obscured.

Such findings provided empirical support for his theory. However, cognitive scientists such as Marr (1982) and Biederman (1987), who posit comprehensive theories of object recognition, did not specify different types of representations or processes for different types of stimulus. Instead they described a single type of system capable of recognizing many varieties of stimuli. In contrast, neuropsychological data suggests a very different view of object recognition.

They noted that brain damage can sometimes impair the recognition of certain categories of stimuli relative to others and focused on a particularly extraordinary problem: that of impaired recognition between living and nonliving 'things'. Individuals with a condition known as agnosia can successfully recognise certain categories of objects but not others (Sheridan

<https://assignbuster.com/the-view-that-recognition-is-the-only-goal-of-visual-perception/>

& Humphreys, 1993; Forde, Francis, Riddoch, Rumiati & Humphreys, 1997; For example Warrington and Shallice, (1984) report patient JBF, who had suffered temporal lobe damage and was able to recognise 90% of inanimate objects shown to him, but only 6% of living things.

He was able to provide concise descriptions of nonliving artefacts (describing a briefcase as a 'small case used by students to carry papers') however was unable to do so with living things, describing a snail as an 'insect animal' and even failing to recognise a parrot entirely. Further clinical evidence in category specificity in visual recognition is outlined by Farah, Wilson, Drain and Tanaka (1998) who summarise a severe disorder of face recognition or prosopagnosia.

Farah et al go on to describe a patient who experiences great difficulty in recognising famous and familiar people by visual inspection and who furthermore was unable to recognise himself in the mirror. A patient studied by Pallis (1955) was unable to recognise his own wife and doctors and failed to identify pictures of Maryln Monroe and Hitler however promptly recognised and named objects in a series of line-drawings.

Such neuropsychological findings suggest vision for recognition is not as straight forward as it may first appear. Evidence for computational models of perception such as those presented by Marr (1982) and Biederman (1987) suggest that recognition is the goal of visual perception. One approach that to some degree opposes the idea that perception is for recognition is the ecological approach championed by James Gibson. Gibson (1979) offers a

radical, bottom-up alternative to the traditional perspectives of visual perception.

According to this view, perception is a result of animal-environment systems and information does not need to be stored in the form of memorial codes, or representations in the way that Marr and Bierderman have suggested (although Gibson does not deny the existence of cognitive processes such as memory). For Gibson, visual perception is strongly linked to action which is very much seen as an end point of perception. According to Gibson, perception is the act of picking up invariants in the environment which specify events, structures, surfaces, objects and the layout for goal-directed activity (Gibson, 1979).

Information is constantly and directly available as a consequence of the environment and the rich spatio-temporal order that it imposes upon the surrounding energy flows, particularly, that which is contained within the ambient optic array, the structure that is imposed on light reflected by the textured surfaces in the external world. Light reaches the eye after having been reflected off surfaces and objects in busy, cluttered environments as the individual moves around.

That light is reflected in straight lines and exists in the highly structured distribution of the array and a number of elements such as material composition, texture and angular interconnections of surfaces alter the flow. These variants in the environment help the viewer to perceive motion of objects travelling in the environment, as well as the layout of surfaces,

depth, distance, and the viewers own ego-motion (awareness of his/her own motion in the world).

Furthermore, the optic array plays a crucial role in perception-action coupling, a fundamental aspect of ecological theory in which informational invariants in the optic flow field become coupled with the action to control movements in an adaptive way; as the individual moves, the information flows are altered creating more information to steer the on-going action (Lee and Lishman, 1975).

One of Gibson's most significant contributions which highlight the importance of perception for movement is the concept of affordances, a concept that he considers a powerful way of combining perception and action. Within the theory of affordances, perception is an invitation to act - the term 'affordances' refers to whatever it is about the environment that contributes to the kind of interaction that occurs between an individual and that environment (Greeno, 1994).

For example, in sport, a ball invites actions such as hitting, catching or throwing whilst barriers afford leaping, stepping or hurdling. Further to the concept of affordances, is the issue of resonance. According to Gibson, the perceptual system resonates to invariant environmental information in the optic array, in the same way that a radio might tune in and pick up a specific radio station. The very concept of resonance suggests an active, exploratory engagement with the environment and further illustrates the constant orientating and moving of a perceiver within his/her surroundings.

Interaction with the environment is at the centre of Gibson's ecological approach to perception. It emphasises the important relationship between perception and action, suggesting that it is direct and cyclical and, in many respects, it challenges the view that recognition is the goal of visual perception. Cognitive neuroscience, a field of research concerned with the study of biological mechanisms underlying cognition offers a wealth of information relevant to the current discussion.

For instance, Ungerleider and Mishkin (1982) claim that the visual cortical areas are individually organised domains, divided into two information streams: one centred on the V4 area, bringing information on object properties to the infero-temporal lobe (ventral stream), and the other centred on the MT or V5 area which brings spatial information to the parietal lobe (dorsal stream). Milner and Goodale (1992) agree on the fundamental separation of functions and the activity of two independent, parallel systems, however deny that the difference is in the resulting percept (object vs. pace).

Instead they suggest that the ventral (occipitotemporal) visual stream provides information for perception and the dorsal (occipitoparietal) stream provides important information necessary for the control of action. Whilst some authors such as Sperry (1952) remain unsatisfied with such an idea, and stress the logical difficulty of considering action and perception as separate functions, Milner and Goodale's (1992) work with their patient Dee Fletcher (DF) suggest that this is the case.

In a tragic incident in February 1988, DF was left with severe impairment of the visual system after carbon monoxide poisoning. DF never regained a full

<https://assignbuster.com/the-view-that-recognition-is-the-only-goal-of-visual-perception/>

and integrated experience of the visual world, and was unable to recognise objects or faces or even make simple visual discriminations such as between a triangle and a circle (Goodale and Milner, 2004). In one experiment, DF was presented with a series of line drawings of various objects (including an apple, a book and a boat) and asked to copy them using a pencil.

DF was unable to recognise the objects, and therefore failed to successfully recreate them. However, when asked to draw the objects from memory, DF produced reasonable renditions. Milner and Goodale (1995) suggest that their patients inability to copy the drawings was not due to a failure to control her finger and hand movements as she moved the pencil, since she had managed to produce the objects from memory in the second phase of the task, but was due to her inability to recognise shapes - when DF was later shown the objects she had drawn, she had no idea what they were.

In the 'mailbox' task, DF was asked to 'post' a card into an open slot similar to that found on a mailbox, however, the slot would be presented in a number of different orientations, not just horizontally. In the matching part of the task DF was asked to turn a hand-held card so that it matched with the orientation of the slot without reaching out toward the display. In the 'posting' aspect of the experiment she was asked to reach out and post the card through the slot.

The patient did very well in the 'posting' aspect of the task but performed almost randomly on the matching task. When posting the card DF began to rotate it toward the slot well in advance of posting it and almost always inserted it smoothly into the slot; according to Milner and Goodale (2004) DF

<https://assignbuster.com/the-view-that-recognition-is-the-only-goal-of-visual-perception/>

was using vision right from the start to guide her movements. It appeared then, that DF was unable to use her ventral system for analysing sensory input, however did have an intact dorsal stream.

Other evidence for this type of two system approach comes from Efron's (1968) study, where his patient was unable to perceive simple geometric shapes, however features such as colour, brightness and movement discrimination were preserved. Additionally, Norman (2002) attempted to further characterise the dorsal and ventral streams, and in doing so suggests a dual-process approach. According to Norman (2002) the two streams are seen as acting synergistically so that the dorsal stream is mainly concerned with perception for recognition and the ventral stream drives visually guided behaviour.

Whilst the evidence presented here, with respect to computational models of visual perception, paints a convincing picture of recognition as the only goal for perception, there is also a wealth of evidence to suggest that it is not the only objective. Findings from neuropsychological studies begin to tell a different story - with the study of brain damaged individuals, the different types of representations or processes for different types of stimulus are considered.

Gibson's approach to perception concentrates more on perception for action, whilst Marr and Biederman's theories were chiefly concerned with object recognition. Furthermore, with the advancement of technology, cognitive neuroscientific studies suggest that two quite separate visual systems exist for both perception and action. In short, whilst each approach discussed here

<https://assignbuster.com/the-view-that-recognition-is-the-only-goal-of-visual-perception/>

have their differences, it is clear that as we humans make sense of and negotiate our way around our external world, both recognition of objects and the performance of action is crucial if we are to live full and enriched lives.