Formal Theories of Visual Perception

Edited by

E.L.J. Leeuwenberg

H.F.J.M. Buffart

Formal Theories of Visual Perception

Edited by

E. L. J. Leeuwenberg and H. F. J. M. Buffart

Psychology Department
Katholieke Universiteit
Nijmegen, The Netherlands

JOHN WILEY AND SONS
Chichester · New York · Brisbane · Toronto

Library of Congress Cataloging in Publication Data:

Main entry under title:

Formal theories of visual perception.

Includes index.
1. Visual perception—Addresses, essays, lectures.
I. Leeuwenberg, Emanuel Laurens Jan. II. Buffart, H. F. J. M.
BF241.F63 152.1 77-12441
ISBN 0 471 99586 X

Photoset and printed in Malta
by Interprint (Malta) Ltd

List of Contributors

A. A. BLANK	*Department of Mathematics, Carnegie-Mellon University, Schenley Park, Pittsburgh, Pennsylvania, U.S.A.*
H. F. J. M. BUFFART	*Department of Experimental Psychology, University of Nijmegen, Nijmegen, The Netherlands*
T. CAELLI	*Department of Psychology, University of Melbourne, Parkville, Victoria, Australia*
J. DRÖSLER	*Department of Psychology, University of Regensburg, Regensburg, Federal Republic Germany*
D. H. FOSTER	*Department of Communication, University of Keele, Keele, Staffordshire, England*
C. D. FRITH	*Division of Psychiatry, MRC Clinical Research Centre, Harrow, Middlesex, England*
H. G. GEISSLER	*Department of Psychology, Humboldt University, Berlin, German Democratic Republic*
S. GROSSBERG	*Department of Mathematics, Boston University, Boston, Massachusetts, U.S.A.*
W. C. HOFFMAN	*Department of Mathematical Science, Oakland University, Rochester, Michigan, U.S.A.*
G. JOHANSSON	*Department of Psychology, University of Uppsala, Uppsala, Sweden*
B. JULESZ	*Bell Laboratories, Murray Hill, New Jersey, U.S.A.*
F. KLIX	*Department of Psychology, Humboldt University, Berlin, German Democratic Republic*
E. L. J. LEEUWENBERG	*Department of Experimental Psychology, University of Nijmegen, Nijmegen, The Netherlands*
H. LINDMAN	*Department of Psychology, University of Melbourne, Parkville, Victoria, Australia*
J. R. POMERANTZ	*The John Hopkins University, Baltimore Maryland, U.S.A.*

F. RESTLE — *Department of Psychology, Indiana University, Bloomington, Indiana, U.S.A.*

U. I. SCHEIDEREITER — *Department of Psychology, Humboldt University, Berlin, German Democratic Republic*

H. A. SIMON — *Department of Psychology, Carnegie-Mellon University, Schenley Park, Pittsburgh, Pennsylvania, U.S.A.*

A. S. WATSON — *Department of Psychology, University of Stirling, Stirling, Scotland*

General Introduction

In the area of psychology it is extremely difficult to think of a topic which is not somehow related to the problems of classification, concept formation and recognition. Even the study of issues concerned with human thought cannot bypass the question as to how invariants may possibly be established in the multitude of data. Such themes invariably take the kernel position in the study of perception. Perhaps no other subject has been approached with such a large variety of method. It is undoubtedly the lack of a convincing overall picture which has led to the great diversity of disciplines engaged in the study of perception. Of old, the process of abstraction has been studied in the context of *epistemology*. Especially from the ranks of philosophers and psychologists, scientists came forward in the early 20th century, directing attention to *'Gestalt' phenomena*. Over-simplified notions on human perception they swept aside; their own findings, however, were often not more penetrating than their motto, 'The whole is more than the sum of its parts'.

The emergence of *Selective Information theory* around 1950 has given rise to an impressive amount of research to establish its relevance for psychology. However, insufficient knowledge of 'perceptual alphabet' was the reason that the attempts were not as successful as had been expected. The latter disadvantage also applied to theories which in certain respects deal in a more complex way with the formation of concepts. Approaches of this type are: *learning theories*, *decision theories* and *cybernetics*. The disordered picture of the results of feedback theories, in which pre-existing knowledge plays a crucial part, compelled the researchers to express their ideas in terms of *network models*. Ideally, these models themselves should yield the relevant distinctive features leading to the classification of patterns. Unfortunately, however, until now no network model has been able to accomplish this task in a perceptually relevant manner. This statement, however, is not intended to refute the perceptual relevance of selective information theory or the idea that feedback and pre-existing knowledge play a role in perception.

As regards its contribution to the theory of perception, research in the field of *pattern recognition* may be regarded as having a more modest value, since the distinctive pattern features are built into the models at fore hand. This research is, indeed, chiefly aimed at the automatic classification of a very specific set of patterns. The need for wider and more flexible classes into which patterns may apparently be categorized, has led to the development of *figure grammars*. Up-till now, however, these also seem to be applicable only to a very specific set of patterns; in other words, this development may perhaps be important purely, for solving technical problems. Thus, it only seldom happens that there is no conflict between the starting-points of the figure grammar approach on the one hand and the findings in the psychology of perception on the other. Here again absolutely no refutation is implied of all the aspects of existing figure grammars or of pattern recognition models.

The growing conviction within the group of perception researchers that the human construction of a percept should once again be regarded as a conclusion which follows upon a process or reasoning has stimulated workers in the field of *Artificial Intelligence* (A. I.) to give expression to such construction of percepts in computer programs. Amongst other things, inventive '*scene-analysis*' programs have been worked out determining the three-dimensional positioning of objects from figures that are given in two dimensions. It remains unsatisfactory, however, that what could be established was at most a possible isomorphism between the input–output relation of the program and that of the perceptual process. Indeed, the multitude of interacting process elements in a program highly complicates the testing of the theoretical principles that are involved. In spite of the large amount of liberty taken and the latitude that is maintained in A.I. model construction, no simulation programmes of the perception process have yet been implemented which can summon up the Gestalt phenomena.

An area of research more amenable to testing—specifically with the aid of protocols—is that concerned with human thought processes ('*problem solving*'). This field of research, however, has yielded insights into perception which agree remarkably well with a series of theories from the field of perception research. The theories do not primarily deal with the perceptual process itself but rather with its outcomes. These *coding theories* are based on perceptual elements, the total number of which indicates the amount of structural information. These elements are not physical dimensions, which could e.g. be determined by means of data reduction techniques applied to relationship judgements between patterns; instead, they are specific relationships, which hierarchically build one upon the other. These relationships constitute the perceptual elements which correspond to structural aspects of the pattern code; in contrast, selective information theory is at most concerned with the metric or the quantitative values associated with each of the structural aspects. These coding theories, which show some resemblance to figure grammars, predict data from extrapolation experiments, complexity judgements, Gestalt phenomena, learning effects and context effects. These coding theories, in which the efficiency principle is of crucial importance are such that for overlearned sets of patterns they allow the deduction of the appropriate distinctive features as code elements.

Alongside the above-mentioned approach, there have recently been surprising developments on what might be seen as the more 'basic' perception *process*. Characteristic of the studies concerned is that the mathematical description of process aspects allows predictions on the interaction of many input data. This type of approach is indicated in the present volume under the heading: *theories* *on field effects*. These theories were already in existence at the very beginning of the 20th century; since that time they have hardly been affected by the above-mentioned movements thanks to the coherence of their concepts. Not only do these theories link up with each other on account of their mathematical nature but at the same time they exhibit points of connection with the coding theories mentioned above. Thus the manner in which a figure is scanned will determine the choice of interpretation ultimately given to the figure. Conversely, and probably of crucial importance to the status of these field theories is, amongst other things, the feedback from the final stage of the information assimilation process, as determined by means of the coding theories, to the first stages of the storage process. For example, the interpretation of a pattern can in the first instance determine the nature of the eye movements and then, in turn, these can influence the field effects of the sensory input.

The immediate occasion for coupling the two last-mentioned approaches was that between them a remarkable convergence of ideas was taking place in recent years. Once again, these two themes are:

(1) Analytical theories of field effects
(2) Coding theories of complex patterns

With regard to the latter theories, it may be noted that practically all their initiators appear in the present volume, although their contributions here deal chiefly with conclusions and implications of their coding theories.

The two classes of theories are in the first place characterized by the fact that while having perceptual relevance they also show the exactness of mathematical expression. It is unusual to find these qualities together. For the majority of the theories presented it can be said that they do not solely describe the results of specific experiments but that they offer explanations for them on a conceptual foundation. Because of these particular properties it is possible, in principle to arrive at either an assimilation of or a confrontation between the two classes of theories. The exploration of this possibility constitutes the chief aim of the present volume.

Pieces of research in the framework of the two themes of the present volume do not as yet give the impression of giving conflicting evidence. Possibly this may be due to the existing differences in method of approach. Studies on Theme 1 tend to be based more on phenomena ascertained by the peripheral sense-organs; whereas studies on Theme 2 sooner tend to concentrate on insights gained through introspective experiences. Connected to this distinction, there is another difference: research work on Theme 1 is in general more concerned with cause-to-effect (causal) relationships, whereas work on Theme 2 investigates effect-to-cause (final) relationships. Where the relevant data on brain-processes are lacking one must obviously resort to arguments like those of efficiency and optimalization if one is to describe relations between subjective experiences. This research leads to statements about rules that are concerned with the form of memory representation. These correspond to the final result of the perceptual process. The process itself is the object of study in Theme 1. Interestingly, the latter investigation, which in first instance may be more easily grasped in terms of analytical mathematics, can lead to findings that can be described in disjunctive terms. The latter terms offer possibilities for linking up with discrete mathematical expressions proper to the coding systems that arise from research work on Theme 2.

Thus it is possible for the coding rules not only to possess validity within—and appropriate to—the domain of subjective experience but also, to occupy a significant position outside that subjective domain.

Contents

Part I

Theories on Field Effects

Introduction

If we are to develop theories on the general structures of the human perceptual system we can naturally only do so by investigating it in relation to its surroundings. Such theories may be developed from two points of view.

The first viewpoint is that in which the perceiver is regarded as an element in a dynamic system. The perceiver needs to adapt himself to his environment in order, ultimately, to survive. The perceptual system, therefore, must be constructed in such a manner that it makes this adaptation possible. From the second point of view, it is not the adaptation of the perceiver that is put central, but his input–output behaviour. The theory in this case is built up in order to explain the behaviour. Naturally, theories developed on the basis of the first viewpoint must be able ultimately to explain the behaviour, and similarly the second class of theories must also be able to explain the adaptation phenomena. The two viewpoints may be typified as follows. In the first view, the theory on the perception system is deduced from the behaviour of the subject's surroundings and in the second, it is deduced from the behaviour of the subject himself.

Within the framework of a general theory on perception mechanisms, it is possible to construct models for specific phenomena. These may be tested. Of course the theory itself is not thereby tested. Certain explanations for a phenomenon are excluded from the framework of the general theory. The general theory therefore reduces the number of possible interpretations of a certain phenomenon and the phenomena are given a place within the framework.

Grossberg evolves a theory on the general structure of the perceptual system from the angle of adaptation philosophy. He visualizes the system as a network of centre-on, surround-off cells. The behaviour of the environment puts requirements on what type of network the system can be. Many different phenomena can thus be understood on the basis of but a few organizational principles. Pattern recognition, it is proposed, is feature detection by means of hierarchically ordered detectors.

Hoffman builds up his theory setting out from the perceiver's behaviour. He assumes that the perceiver receives information about his surroundings in a two-dimensional space. It is essential that the perceiver keeps giving the same reaction to some changing stimuli, i.e. that the perceiver remains aware of certain constancies such as shape or length. Hoffman uses a mathematical theory, Lie's theory of groups, which is able to indicate connections between operators that broadly transform a stimulus and local operators that effectuate an infinitely small transformation. The constancies can be expressed in these local operators.

In the sense of the Lie algebra, a complete group is made from these operators by adding new local operators. These predict new constancies. According to Hoffman, the brain contains representations of this group. Its elements are regarded as constancy processors. In the same way as the general operators are built up out of local operators, so the representations of the stimuli are built up by means of the constancy processors. Extension of the theory by making the local operators into contour tracing operators leads to the assumption that form recognition takes place by means of the structured sequence of short lengths of line. Surprisingly, this is the very concept on which Leeuwenberg's coding theories (see Part II) are based.

Foster proposes a theory which is also based on the perceiver's behaviour, though it describes a less extensive field. His theory deals with apparent motion. He gives an outline of a theory on apparent motion, setting out from the idea that perceived phenomena result from the minimalization of a quantity after the presentation of a stimulus. In specific models still to be determined an indication will naturally have to be made as to which quantity should be minimalized. This could, for instance, have some connection with the functions introduced by Buffart and Watson. Foster reports that his theory cannot fit in with a theory such as Hoffman's. The implication is that at least one of the two theories is false.

It is attractive to imagine visual space as metric space. This is certainly not obvious at first sight. As has been stated in the introduction, cognitive factors can affect visual space and in so doing disturb the metric character. However, one can accept this as such and yet attempt to find out whether in the absence of strong cognitive influence visual space may be regarded as metric space and if so what the metrics are. In order to determine experimentally whether there is cognitive influence on the nature of the space, subjects may be instructed in a manner such as could be expected to affect the possible metric character of the space. If one is able to demonstrate this influence and if one can manipulate it in a consistent fashion, then one can hope only that the space free of cognitive influences, is metric space.

If one regards visual space as metric space, one must investigate whether and to what extent objects in the physical space determine these metrics. The fact that there are stimuli that are perceived differently from their physical existence gives rise to considering visual space as non-Euclidean space. The most familiar non-Euclidean geometries are the Riemannian geometries(1). Various theories on the nature of the visual space that are presented below (Blank, Caelli, Drösler, Watson) are in fact Riemannian geometries.

Blank discusses the geometry of the binocular visual space. He demonstrates that the visual space is not Euclidean but Riemannian. It follows from the fact that the space possesses in first order the Desarguesian property that it is a Riemannian space with a constant curvature. This is probably hyperbolic. There are subjects, however, who react in a Euclidean manner. According to

Blank, they have remembered the Euclidean interpretation of space.

Caelli et al. conclude from Hoffman's theory a model for the time-space (two dimensional) relation in the case of apparent motion. The basis of the model is the idea, that the signal propagation in the brain can not be faster than a velocity c. The space describing the relation is an elliptic space.

Drösler's theory is based on the perception of length, area and velocity in monocular, two-dimensional space. His treatment of the perception of these quantities as properties of a geometric space leads him to discard the well-known psychometric power laws. He offers a different explanation for these observations and arrives at the conclusion that the monocular visual space is a Riemannian space with a constant positive curvature, an elliptic space.

A concept in disagreement with Caelli's and Drösler's is Watson's concept. He emphatically states that the monocular visual time-space is Riemannian but does not possess a constant curvature. Although his theory deals only with monocular space, it can have consequences for perception in binocular space. Watson conceives of the Riemannian space as a deviation of Euclidean space by presuming that each brightness contrast in the space introduces a certain field, in principle perpendicular to this contrast.

This idea does not necessarily have to be seen as an ad hoc assumption, since Buffart in his paper on brightness perception has also introduced such fields (weighting coefficients). Possibly these are the same. According to Watson, the fields cause deviations from Euclidean space, so that each stimulus induces its own metrics in the visual space. On the basis of this, Watson makes predictions on the perception of visual illusions and aftereffects. There is, however, the possibility that his theory does not apply to illusions under special circumstances; a Poggendorf illusion, for instance, at a certain orientation is not an illusion.

Buffart presents a theory on contrast processing and brightness perception. According to him the visual system only reacts to luminance changes in a stimulus. He indicates how a stimulus is processed by the retina and how the outputs of the two retinas induce fields in the cortex. These fields form the basis of the brightness perception and the binocular interaction. The theory explains depth perception in random-dot stereograms, binocular rivalry, monocular brightness perception and brightness illusions. If the suggestion is correct that the fields introduced by Buffart and the fields proposed by Watson are identical, then the above mentioned problem of Watson's theory may be solved, for in that case eye movements can affect the Watson field and thus the metrics of space.

Note

(1) These are spaces in which the shortest distance between two points is not a straight line in the ordinary sense. Thus, for example, the shortest route from Sydney to Amsterdam goes over the surface of the earth, approximately via Calcutta. If one

were to ask someone who considered a line over the earth's surface as a straight line to prolong the Sydney to Calcutta line, he would arrive in North-West Europe. If one were to ask the same of somebody who measures in terms of Euclidean geometry, he will arrive somewhere in outer space.