Modularity in Knowledge Representation and Natural-Language Understanding

edited by Jay L. Garfield

The MIT Press

From The MIT Press

MITCogNet

Second printing, 1989

This book was set in Palatino by Asco Trade Typesetting Ltd., Hong Kong, and printed and bound by Halliday Lithograph in the United States of America.

Library of Congress Cataloging-in-Publication Data

Modularity in knowledge representation and natural-language understanding.

"A Bradford book."
Bibliography: p.
Includes index.
1. Psycholinguistics. 2. Cognition. 3. Languages—Philosophy. 4. Vision.
5. Neurolinguistics.
I. Garfield, Jay L., 1955–
P40.M6 1987 401′.9 86-34415
ISBN 0-262-07105-3

18

Marr's Theory of Vision

Tyler Burge

In recent years I have given several arguments against a common doctrine in the philosophy of psychology called *individualism*. Individualism is the view that if one fixes a person's (or an animal's) physical history in nonintentional terms and in such a way that one makes no assumptions about the nature of the subject's environment beyond the subject's surfaces, then it is impossible for the subject's intentional states or events to have been different from what they are. The strategy of my arguments against this view has been to show that a subject's concepts or mental states are individuated in ways that lean on assumptions about the environment. For example, a person's concept of aluminum is typically what it is partly because the person has interacted (perhaps indirectly, through other people) with instances of aluminum in the world. One can coherently imagine varying the environment, without relevantly affecting an individual's physical makeup, in such a way as to vary the person's concepts and mental states.[1] Most of the arguments that I have given to this end rely on higher cognitive faculties. It would be interesting to apply the strategy of the argument to relatively primitive and thoroughly modular cognitive faculties. That is what I propose to do here.

Not all aspects of psychology involve the ascription of intentional states and events. I do not care to argue over whether these aspects are individualistic. I shall focus purely on aspects of psychological theory that are intentional. The general picture of these aspects of psychological explanation is as follows.

Ascription of intentional states and events in psychology constitutes a type of individuation and explanation that carries presuppositions about the specific nature of the person's environment. Moreover, states and events are individuated so as to set the terms for specific evaluations of them for truth or other types of success. We can judge directly whether conative states are practically successful and cognitive states are veridical. For example, by characterizing a subject as visually representing an X and specifying whether the visual state appropriately derives from an X in the particular case, we can judge whether the subject's state is veridical. Theories of vision, belief formation, memory, learning, decision-making, categorization, and perhaps even reasoning all attribute states that are subject to practical and semantical evaluation by reference to standards partly set by a wider environment.

Psychological theories are not themselves evaluative theories. But they often individuate phenomena so as to make evaluation readily accessible because they are partly motivated by such judgments. Thus, we judge that in certain delimitable contexts people get what they want, know what is the case, and perceive what is there. And we try to frame explanations that account for these successes, and the correlative failures, in such a way as to illumine as specifically as possible the mechanisms that underlie and make true our evaluations.

I want to illustrate and develop these points by considering at some length David Marr's theory of vision. I choose this example primarily because it is a very advanced and impressive theory, and because it admits of being treated in some depth. Its information-processing approach is congenial with mainstream work in cognitive psychology. Some of its intentional aspects are well understood—indeed, some of them are conceptually and mathematically far ahead of its formal (or syntactical) and physiological aspects. Thus, the theory provides an example of a mentalistic theory with solid achievements to its credit.

The theory of vision maintains a pivotal position in psychology. Since perceptual processes provide the input for many higher cognitive processes, it is reasonable to think that if the theory of vision treats intentional states nonindividualistically, other central parts of cognitive psychology will do likewise. Information processed by more central capacities depends, to a large extent, on visual information.

Certain special aspects of the vision example must be noted at the outset. As I remarked, the arguments that I have previously published against individualism (see note 1) have centered on "higher" mental capacities, some of which essentially involve the use of language. This focus was motivated by an interest in the relation between thought and linguistic meaning and in certain sorts of intellectual responsibility. Early human vision makes use of a limited range of representations—representations of shape, texture, depth and other spatial relations, motion, color, and so forth. These representations (percepts) are formed by processes that are relatively immune to correction from other sources of information; and the representations of early vision appear to be fully independent of language. So the thought experiments that I have previously elaborated will not carry over simply to early human vision. (One would expect those thought experiments to be more relevant to social and developmental psychology, to concept learning, and to parts of "higher" cognitive psychology.) But the case against individualism need not center on higher cognitive capacities or on the relation between thought and language. The anti-individualistic conclusions of the previously published arguments can be shown to apply to early human vision. The abstract schema articulated by those thought experiments also applies.

The schema rests on three general facts. The first is that the entities in the objective world one normally intentionally interacts with, in the employment of many representational (intentional) types, affects the semantical properties of those representational types, what they are, and how one individuates them.[2] A near consequence of this first fact is that there can be slack between the way a subject's representational types apply to the world and what that person knows about (and how he or she can react to) the way they apply. It is possible for representational types to apply differently without the person's physical reactions or discriminative powers being different. These facts, together with the fact that many fundamental mental states and events are individuated in terms of the relevant representational types, suffice to generate the conclusion that many paradigmatic mental states and events are not individualistically individuated but may vary while a person's body and discriminative powers are conceived as constant. By the second fact one can conceive of the way a person's representational types apply to the objective world as varying while that person's history, nonintentionally and individualistically specified, is held constant. By the first fact such variation may vary the individuation of the person's representational types. By the third fact such variation may affect the individuation of the person's mental states and events. I shall illustrate how instances of this schema are supported by Marr's theory of vision.[3]

Marr's theory subsumes three explanatory enterprises: a theory of the computation of the information, an account of the representations used and of the algorithms by which they are manipulated, and a theory of the underlying physiology. Our primary interest is in the first level and in that part of the second that deals with the individuation of representations. These latter parts of the theory are fundamentally intentional.

The theory of the computation of information encompasses an account of what information is extracted from what antecedent resources and an account of the reference-preserving "logic" of the extraction. These accounts proceed against a set of biological background assumptions. It is assumed that visual systems have evolved to solve certain problems forced on them by the environment. Different species are set different problems and solve them differently. The theory of human vision specifies a general information-processing problem: that of generating reliable representations of certain objective, distal properties of the surrounding world on the basis of proximal stimulations.

The human visual system computes complex representations of certain visible properties on the basis of the values of light intensities in retinal images. The primary visible properties that Marr's theory treats are the shapes and locations of things in the world, but various other properties—motion, texture, color, lightness, shading—are also dealt with in some detail. The overall computation is broken down into stages of increasing complexity, each containing modules that solve various subproblems.

The theory of computation of information clearly treats the visual system as going through a series of intentional or representational states. At an early stage, the visual system is counted as representing objective features of the physical world.[4] There is no other way to treat the visual system as solving the problem that the theory sees it as solving than by attributing intentional states that represent objective, physical properties.

More than half of Marr's book is concerned with developing the theory of the computation of information and with individuating representational primitives. These parts of the theory are more deeply developed, both conceptually and mathematically, than the account of the algorithms. This point serves to correct the impression, often conveyed in recent philosophy of psychology, that intentional theories are regressive and all the development of genuine theory in psychology has been proceeding at the level of purely formal, "syntactical" transformations (algorithms) that are used in cognitive systems.

I now want, by a series of examples, to give a fairly concrete sense of how the theory treats the relation between the visual system and the physical environment. Understanding this relation will form essential background for understanding the nonindividualistic character of the theory. The reader may skip the detail and still follow the philosophical argument, but the detail is there to support the argument and to render the conception of explanation that the argument yields both concrete and vivid.

Initially, I will illustrate two broad points. The first is that the theory makes essential reference to the subject's distal stimuli and makes essential assumptions about contingent facts regarding the subject's physical environment. Not only do the basic questions of the theory refer to what one sees under normal conditions, but the computational theory and its theorems are derived from numerous explicit assumptions about the physical world. The second point is that the theory is set up to explain the reliability of a great variety of processes and subprocesses for acquiring information, at least to the extent that they are reliable. Reliability is presupposed in the formulations of the theory's basic questions. It is also explained through a detailed account of how, in certain specified, standard conditions, veridical information is derived from limited means. The theory explains not merely the reliability of the system as a whole, but the reliability of various stages in the visual process. It begins by assuming that we see certain objective properties and proceeds to explain particular successes by framing conditions under which success would be expected (where the conditions are in fact typical). Failures are explained primarily by reference to a failure of these conditions to obtain. To use a phrase of Bernie Kobes, the theory is not success-neutral. The explanations and the kinds of the theory presuppose that perception and numerous subroutines of perception are veridical in normal circumstances.

Example 1

In an early stage of the construction of visual representation, the outputs of channels or filters that are sensitive to spatial distributions of light intensities are combined to produce representations of local contours, edges, shadows, and so forth. The filters fall into groups of different sizes, in the sense that different groups are sensitive to different bands of spatial frequencies. The channels are primarily sensitive to sudden intensity changes, called *zero crossings*, at their scales (within their frequency bands). The theoretical question arises: How do we combine the results of the different-size channels to construct representations with physical meaning—representations that indicate edge segments or local contours in the external physical world? There is no *a priori* reason why zero crossings obtained from different-size filters should be related to one physical phenomenon in the environment. There is, however, a physical basis for their being thus related. This basis is identified by *the constraint of spatial localization*. Things in the world that give rise to intensity changes in the image, such as changes of illumination (caused by shadows, light sources) or changes in surface reflectance (caused by contours, creases, and surface boundaries), are spatially localized, not scattered, and not made up of waves. Because of this fact, if a zero crossing is present in a channel centered on a given frequency band, there should be a corresponding zero crossing at the same spatial location in larger-scale channels. If this ceases to be so at larger scales, it is because two or more local intensity changes are being averaged together in the larger channel (e.g., the edges of a thin bar may register radical frequency changes in small channels but go undetected in larger ones) or because two independent physical phenomena are producing intensity changes in the same area but at different scales (e.g. a shadow superimposed on a sudden reflectance change; if the shadow is located in a certain way, the positions of the zero crossings may not make possible a separation of the two physical phenomena). Some of these exceptions are sufficiently rare that the visual system need not and does not account for them, thus allowing for possible illusions; others are reflected in complications of the basic assumption that follows. The spatial-coincidence constraint yields the *spatial-coincidence assumption*: If a zero-crossing segment is present in a set of independent channels over a contiguous range of sizes, and the segment has the same position and orientation in each channel, then the set of such zero-crossing segments indicates the presence of an intensity change in the image that is due to a single physical phenomenon (a change in reflectance, illumination, depth, or surface orientation). Thus, the theory starts with the observation that physical edges produce roughly coincident zero crossings in channels of neighboring sizes. The spatial-coincidence assumption asserts that the coincidence of zero crossings of neighboring sizes is normally sufficient evidence of a real physical edge. Under such circumstances, according to the theory, a representation of an edge is formed.[5]

Example 2

Because of the laws of light and the way our eyes are made, positioned, and controlled, our brains typically receive similar image signals originating from two points that are fairly similarly located in the respective eyes or images, at the same horizontal level. If two objects are separated in depth from the viewer, the relative positions of their image signals will differ in the two eyes. The visual system determines the distance of physical surfaces by measuring the angular discrepancy in position (disparity) of the image of an object in the two eyes. This process is called stereopsis. To solve the problem of determining distance, the visual system must select a location on a surface as represented by one image, identify the same location in the other image, and measure the disparity between the corresponding image points. There is, of course no *a priori* means of matching points from the two images. The theory indicates how correct matches are produced by appealing to three physical constraints (the first of which is not made explicit but is relied upon): that the two eyes produce similar representations of the same external items, that a given point on a physical surface has a unique position in space at any given time, and that matter is cohesive (i.e., it is separated into objects, the surfaces of which are usually smooth in the sense that surface variation is small in comparison with the overall distance from the observer). These three physical constraints are rewritten as three corresponding constraints on matching: that two representational elements can match if and only if they normally could have arisen from the same physical item (e.g., in stereograms, dots match dots rather than bars), that nearly always each representational element can match only one element from the other image (except when two markings lie along the line of sight of one eye but are separately visible by the other eye, causing illusions), and that disparity varies smoothly almost everywhere (this derives from the third of the physical constraints, which implies that the distance to the visible surface varies approximately continuously except at object boundaries, which occupy a small fraction of the area of an image). Given suitable precisifications, these matching constraints can be used to prove the fundamental theory of stereopsis: If a correspondence is established between physically meaningful representational primitives extracted from the left and right images of a scene that contains a sufficient amount of detail (roughly 2 percent density for dot stereograms), and if the correspondence satisfies the three matching constraints, then that correspondence is physically correct and hence unique. The method is, again, to identify general physical conditions that give rise to a visual process, and then to use those conditions to motivate constraints on the form of the process that, when satisfied, will allow the process to be interpreted as providing reliable representations of the physical environment.[6]

These examples illustrate theories of the computation of information. The critical move is the formulation of general physical facts that limit the interpretation of a visual problem enough to allow one to interpret the machinations of the visual system as providing a unique and veridical solution, at least in typical cases. The primary aim of referring to contingent physical facts and properties is to enable the theory to explain the visual system's reliable acquisition of information about the physical world—that is, to explain the success or veridicality of various types of visual representation.

I now turn to a third point that is a natural corollary of the second and that will be critical for my argument that the theory is nonindividualistic: The information carried by representations—their intentional content—is individuated in terms of the specific distal causal antecedents in the physical world that the information is about and that the representations normally apply to. The individuation of the intentional features of numerous representations depends on a variety of physical constraints that our knowledge of the external world gives us. Thus, the individuation of intentional content of representational types presupposes the veridicality of perception. Not only the explanations but also the intentional kinds of the theory presuppose contingent facts about the subject's physical environment.

Example 3

In building up information or representational primitives in the primal sketch, Marr states six general physical assumptions that constrain the choice of primitives. I shall state some of these in order to give a sense of their character: that the visible world is composed of smooth surfaces having reflectance functions whose spatial structure may be complex, that markings generated on a surface by a single process are often arranged in continuous spatial structures (curves, lines, etc.), and that if the direction of motion is discontinuous at more than one point (along a line, for example) then an object boundary is present. These assumptions are used to identify the physical significance of—that is, the objective information given by—certain types of patterns in the image. The computational theory states conditions under which these primitives form to carry information about items in the physical world (Marr 1982, pp. 44–71). The theory in example 1 is a case in point: Conditions are laid down under which certain patterns may be taken as representing an objective physical condition—as being edge, boundary, bar, or blob detectors. Similar points apply for more advanced primitives.

Example 4

In answering the question of what assumptions we reasonably and actually apply when we interpret silhouettes as three-dimensional shapes, Marr motivates a central representational primitive by stating physical constraints that lead to the proof of a theorem: that each line of sight from the viewer to the object grazes the object's surface at exactly one point, that nearby points on the contour in an image arise from nearby points on the contour generator on the viewed object (i.e., points that appear close together in the image actually are close together on the object's surface), and that the contour generator lies wholly in a single plane. Obviously, these are conditions of perception that may fail, but they are conditions under which humans seem to do best at solving the problem of deriving three-dimensional shape decriptions from representations of silhouettes. *Definition:* A generalized cone is a three-dimensional object generated by moving a cross section along an axis; the cross section may vary smoothly in size, but its shape remains the same (e.g., footballs, pyramids, legs, and stalagmites are, or approximate, generalized cones). *Theorem:* If the surface is smooth and if the physical constraints given just above hold for all distant viewing positions in any one plane, then the viewed surface is a generalized cone. The theorem indicates a natural connection between generalized cones and the imaging process. Marr (1982, pp. 215–225) infers from this, and from certain psycho-physical evidence, that representations of generalized cones—that is, representations with intentional content concerning generalized cones—are likely to be fundamental among our visual representations of three-dimensional objects.

Throughout the theory, representational primitives are selected and individuated by considering specific, contingent facts about the physical world that typically hold when we succeed in obtaining veridical visual information about that world. The information or content of the visual representations is always individuated by reference to the physical objects, properties, or relations that are seen. In view of the success orientation of the theory, this mode of individuation is grounded in its basic methods. If theory were confronted with a species of organism reliably and successfully interacting with a different set of objective visible properties (by perhaps a different set of optical laws), the representational types that the theory would attribute to the organism would be different, regardless of whether the individual organism's physical mechanisms were different.

I am now in a position to argue that the theory is not individualistic: (1) The theory is intentional. (2) The intentional primitives of the theory and the information they carry are individuated by reference to contingently existing physical items or conditions by which they are normally caused and to which they normally apply. (3) Thus, if these physical conditions and, possibly, attendant physical laws were regularly different, the information conveyed to the subject and the intentional content of his or her visual representations would be different. (4) It is not incoherent to conceive of relevantly different physical conditions and perhaps relevantly different (say, optical) laws regularly causing the same nonintentionally, individualistically individuated physical regularities in the subject's eyes and nervous system. It is enough if the differences are small; they need not be wholesale. (5) In such a case (by point 3), the individual's visual representations would carry different information and have different representational content, though the person's whole nonintentional physical history (at least up to a certain time) might remain the same. (6) Assuming that some perceptual states are identified in the theory in terms of their informational or intentional content, it follows that individualism is not true for the theory of vision.

I shall defend the argument stepwise. First, I take it that the claim that the theory is intentional is sufficiently evident. The top levels of the theory are explicitly formulated in intentional terms. And their method of explanation is to show how the problem of arriving at certain veridical representations is solved.

The second step of the argument was substantiated through examples 3 and 4. The intentional content of representations of edges or generalized cones is individuated in terms of specific reference to those very contingently instantiated physical properties, on the assumption that those properties normally give rise to veridical representations of them.

The third step in the argument is supported both by the way the theory individuates intentional content (see the preceding paragraph and examples 3 and 4) and by the explanatory method of the theory (see the second point illustrated above, and examples 1 and 2). The methods of individuation and explanation are governed by the assumption that the subject has adapted to his or her environment sufficiently to obtain veridical information from it under certain normal conditions. If the properties and relations that normally caused visual impressions were regularly different from what they are, the individual would obtain different information and have visual experiences with different intentional content. If the regular, lawlike relations between perception and the environment were different, the visual system would be solving different information-processing problems, it would pass through different informational or intentional states, and the explanation of vision would be different. To reject this third step of our argument would be to reject the theory's basic methods and questions. But these methods and questions have already borne fruit, and there are no good reasons for rejecting them.

I take it that the fourth step is a relatively unproblematic counterfactual. There is no metaphysically necessary relation between individualistically individuated processes in a person's body and the causal antecedents of those processes in the surrounding world.[7] (To reject this step would be self-defeating for the individualist.) If the environmental laws and conditions were different, the same proximal visual stimulations could have regularly had different distal causes. In principle, we can conceive of some regular variation in the distal causes of perceptual impressions with no variation in a person's individualistically specified physical processes, even while conceiving the person as well adapted to the relevant environment—though, of course, not uniquely adapted.

Steps three and four, together with the unproblematic claim that the theory individuates some perceptual states in terms of intentional content or representational type, entail that the theory is nonindividualistic.

Steps two and three are incompatible with certain philosophical approaches that have no basis in psychological theory. One might claim that the information content of a visual representation would remain constant even if the physical conditions that lead to the representation were regularly different. It is common to motivate this claim by pointing out that one's visual representations remain the same, whether one is perceiving a black blob on a white surface or having an eidetic hallucination of such a blob. So, runs the reasoning, why should changing the distal causes of a perceptual representation affect its content? On this view, the content of a given perceptual representation is commonly given as that of *the distal cause of this representation*, or *the property in the world that has this sort of visual appearance*. The content of these descriptions is intended to remain constant across possible situations in which the microphysical events of a person's visual processes remain the same while the distal causes of those processes are regularly and significantly different. It is thought that the representations themselves (and our experiences of them) remain constant under these circumstances. Thus, as the distal antecedents of one's perceptual representations vary, the reference of those representations will vary, but their intentional content will not.[8]

There is more wrong with this line than I have room to develop here. I will mention some of the more straightforward difficulties. The motivation from perceptual illusion falls far short. One is indeed in the same perceptual state whether one is seeing or hallucinating. However, that is because the intentional content of one's visual state (or representation) is individuated against a background in which the relevant state is normally veridical. Thus, the fact that one's percepts or perceptual states remain constant between normal perception and hallucinations does not even tend to show that the intentional visual state remains constant across circumstances in which different physical conditions are the normal antecedents of one's perceptions.

Let us consider the proposals for interpreting the content of visual representations. In the first place, both descriptions (*the distal cause of this representation* and the property in the world that uses this sort of visual appearence et al.) are insufficiently specific. There are lots of distal causes and lots of things that might be said to appear *thus* (e.g. the array of light striking the retina as well as the physical surface). We identify the relevant distal cause (and the thing that normally appears thus and so) as the thing that we actually see. To accurately pick out the "correct" object with one of these descriptions would at the very least require a more complex specification. But filling out the descriptive content runs into one or both of two difficulties: Either it includes kinds that are tied to a specific environment (*the convex, rough-textured object that is causing this representation*). In such a case, the description is still subject to our argument, for these kinds are individuated by reference to the empirical environment. Or it complicates the constraints on the causal chain to the extent that the complications cannot plausibly be attributed to the content of processes in the early visual system.

Even in their unrevised forms, the descriptions are overintellectualized philosophers' conceits. It is extremely implausible, and empirically without warrant, to think that packed into every perceptual representation is a distinction between distal cause and experiential effect, or between objective reality and perceptual appearance. These are distinctions developed by reflecting on the ups and downs of visual perception. They do not come in at the ground, animal level of early vision.

A further mistake is the view that perceptual representations never purport to specify particular physical properties *as such*, but only via some relation they bear to inner occurrences, which are directly referred to. (Even the phrase *the convex object causing this percept* invokes a specification of objective convexity as such.) The view will not serve the needs of psychological explanation as actually practiced. True descriptions of information are too inspecific to account for specific success in solving problems in retrieving information about the actual, objective world. Moreover, the view raises difficulties in epistemology.

The best empirical theory that we have individuates the intentional content of visual representations by specific reference to specific physical characteristics of visible properties and relations. The theory does not utilize complicated, self-referential, attributively used role descriptions of those properties. It does not individuate content primarily by reference to phenomenological qualities. Nor does it use the notions of cause or appearance in specifying the intentional content of early visual representations.[9]

The second and third steps of my argument are incompatible with the claim that the intentional content of visual representations is determined by their "functional role" in each person's system of dispositions, nonintentionally and individualistically specified. This claim lacks any warrant in the practice of the science. In the first place, the theory suggests no reduction of the intentional to the nonintentional. In the second, although what a person can do, nonvisually, constitutes evidence for what he or she can see, there is little ground for thinking that either science or common sense takes an individual person's nonvisual abilities fully to determine the content of his or her early visual experience. A person's dispositions and beliefs develop by adapting to what the person sees. As the person develops, the visual system (at least at its more advanced stages—those involving recognition) and the belief and language systems affect one another. But early vision seems relatively independent of these nonvisual systems. A large part of learning is accomodating one's dispositions to the information carried by visual representations. Where there are failures of adaptation, the person does not know that the visual apparatus is presenting to him or her. Yet the presentations are there to be understood.

Conclusion

There is a general argument that seems to me to show that a person's nonintentional dispositions could not fix (individuate) the intentional content of the person's visual presentations. The argument begins with a conception of objectivity. As long as the person's visual presentations are of public, objective objects, properties, or relations, it is possible for the person to have mistaken presentations. Such mistakes usually arise for a single sensory modality—so that when dispositions associated with other modalities (e.g. touch) are brought into play, the mistake is rectified. But as long as the represented object or property is objective and physical, it is in principle possible, however unlikely, that there may be a confluence of illusions such that all an individual person's sensory modalities would be fooled and all the person's nonintentional dispositions would fail to distinguish between the normal condition and the one producing the mistaken sensory representations. This is my first assumption. In the argument, I shall employ a corollary: Our concept of objectivity is such that no one objective entity that we visually represent is such that it must vary with, or be typed so as necessarily to match exactly, an individual's proximal stimuli and discriminative abilities. The point follows from a realistic, and even from a nonsubjectivistic, view of the objects of sight.[10]

I argued above that intentional representational types are not in general individuated purely in terms of an attributive role description of a causal relation, or a relation of appearance-similarity, between external objects and qualitative perceptual representatives of them. For present purposes, this is my second assumption: Some objective physical objects and properties are represented as such; they are specifically specified.

Third, in order to be empirically informative, some visual representations that represent objective entities as such must have the representational characteristics that they have partly *because* instances regularly enter into certain relations with those objective entities.[11] Their carrying information, their having objective intentional content, consists partly in their being the normal causal products of objective entities. And their specific intentional content depends partly on their being the normal products of the specific objective entities that give rise to them. That is why we individuate intentional visual representations in terms of the objective entities to which they normally apply for members of a given species. This is the core of truth in the slogan (sometimes misapplied) that mistakes presuppose a background of veridicality.

The assumptions in the three preceding paragraphs make it possible to state a general argument against individualism regarding visual states. Consider a person P who normally correctly perceives instances of a particular objective visible property O. In such cases, let the intentional type of P's perceptual representation (or perceptual state) be O'. Such perceptual representations are normally the product of interaction with instances of O. But imagine that for P perceptual representations typed O' are on some few occasions the product of instances of a different objective property, C. On such occasions, P mistakenly sees an instance of C as an O; P's perceptual state is of type O'. Assume that O' represents any instance of O as such (as an O), in the sense of the second premise, not merely in terms of some attributive role description. Since O' represents an objective property, it is possible, by the first premise, to conceive of P as lacking at his or her disposal (at every moment up to a given time) any means of discriminating instances of C from instances of O.

Now hold fixed both P's physical states (up to the given time) and his or her discriminative abilities, nonintentionally and individualistically specified. But conceive of the world as lacking *O* altogether. Suppose that the optical laws in the counterfactual environment are such that the impressions on P's eyes and the normal causal processes that lead to P's visual representations are explained in terms of instances of *C* (or, at any rate, in terms of some objective, visible entities other than instances of *O*). Then, by the third premise, P's visual representation (or visual state) would not be of intentional type *O′*. At the time when in the actual situation P is misrepresenting a *C* as an *O*, P may counterfactually be perceiving something (say, a *C*) correctly (as a *C*) if the processes that lead to that visual impression are normal and of a type that normally produces the visual impression that P has on that occasion. Thus, P's intentional visual states could vary while his or her physical states and nonintentionally specified discriminative abilities remained constant.

The first premise and the methodology of intentional-content individuation articulated in the third premise entail the existence of examples. Since examples usually involve shifts in optical laws, they are hard to fill out in great detail. But it is easiest to imagine concrete cases taken from early but still conscious vision. These limit the number of an individual's dispositions that might be reasonably thought to bear on the content of his or her visual states. Early vision is relatively independent of language or other cognitive abilities. It appears to be thoroughly modular.

Suppose that the relevant visible entities are very small and not such as to bear heavily on adaptive success. An *O* may be a shadow of a certain small size and shape on a gently contoured surface. A *C* may be a shallow crack of similar size. In the actual situation, P sees *O*s regularly and correctly as *O*s: P's visual representations are properly explained and specified as shadow representations of the relevant sort. Assume that P's visual and other discriminative abilities are fairly normal. P encounters the cracks of relevant size very rarely and, on those few occasions, not only misperceives them as shadows but has no dispositions that would enable him or her to discriminate those instances from shadows. We may assume, given P's actual abilities and the actual laws of optics, that P would be capable, in ideal circumstances, of visually discriminating some instances of *C* (relevantly similar cracks) from instances of *O* (the relevant sort of shadows). But our supposition is that in the actual cases where P is confronted by instances of the cracks, the circumstances are not ideal. All P's abilities would not succeed in discriminating those instances of relevant cracks, in those circumstances, from instances of relevant shadows. P may not rely on touch in cases of such small objects, or touch may also be fooled. P's ability to have such mistaken visual states is argued for by the objectivity premise.

In the counterfactual case, the environment is different. There are no instances of the relevant shadows visible to P, and the laws of optics differ in such a way that P's physical visual stimulations (and the rest of P's physical makeup) are unaffected. Suppose that the physical visual stimulations that in the actual case are derived from instances of the relevant sort of shadows are counterfactually caused by and explained in terms of cracks of relevant size. Counterfactually, the cracks take the places of the shadows. On the few occasions where, in the actual case, P misperceives shadows as cracks, P is counterfactually confronted with cracks; the optical circumstances that lead to the visual impressions on those occasions are, we may suppose, normal for the counterfactual environment.[12] On such counterfactual occasions, P would be visually representing small cracks as small cracks. P would never have visual representations of the relevant sort of shadows. One can suppose that even if there were the relevant sort of shadows in the counterfactual environment, the different laws of optics in that environment would not enable P ever to see them. But since P's visual states would be the normal products of normal processes and would provide as good an empirical basis for learning about the counterfactual environment as P has for learning about the actual environment, it would be absurd to hold that (counterfactually) P misperceives the prevalent cracks as shadows on gently contoured surfaces. Counterfactually, P correctly sees the cracks as cracks. Thus, P's intentional perceptual states differ between actual and counterfactual situations. This general argument is independent of the theory of vision that I have been discussing. It supports and is further supported by that theory.

Acknowledgments

A version of this chapter was given orally at the Sloan Conference at MIT in May 1984. I have made use of discussions with David Israel, Bernie Kobes, and Neil Stillings. Except for a few expository adjustments, this is part of a larger paper: "Individualism and Psychology," *Philosophical Review* 95 (1986): 3–45.

Notes

1. "Individualism and the Mental," *Midwest Studies* 4 (1979): 73–121; "Other Bodies," in *Thought and Object*, ed. A. Woodfield (Oxford University Press, 1982); "Two Thought Experiments Reviewed," *Notre Dame Journal of Formal Logic* 23 (1982): 284–293; "Cartesian Error and the Objectivity of Perception," in *Subject, Thought, and Context*, ed. MacDowell and Pettit (Oxford University Press, 1986); "Intellectual Norms and Foundations of Mind," *Journal of Philosophy* 83 (1986): 697–720. The aluminum argument is adapted from an argument in Hilary Putnam's "The Meaning of 'Meaning'" (*Philosophical Papers*, vol. 2 [Cambridge University Press, 1975]). What Putnam wrote in his paper was, strictly, not even compatible with this argument (see the first two papers cited in this note for discussion). However, the aluminum argument lies close to the surface of the argument he does give. The arthritis argument raises rather different issues, despite its parallel methodology.

2. *Representational type* (also *intentional type*) is a relatively theory-neutral term for intentional content, or even intentional state-kinds. See note 3. One could about as well speak of concepts, percepts, and the representational or intentional aspects of thought contents—or of the counterpart states.
3. In what follows I make use of Marr's important book *Vision* (San Francisco: Freeman, 1982). Marr writes:

   > The purpose of these representations is to provide useful descriptions of aspects of the real world. The structure of the real world therefore plays an important role in determining both the nature of the representations that are used and the nature of the processes that derive and maintain them. An important part of the theoretical analysis is to make explicit the physical constraints and assumptions that have been used in the design of the representations and processes.... (p. 43)
   >
   > It is of critical importance that the tokens [representational particulars] one obtains [in the theoretical analysis] correspond to real physical changes on the viewed surface; the blobs, lines, edges, groups, and so forth that we shall use must not be artifacts of the imaging process, or else inferences made from their structure backwards to the structures of the surface will be meaningless. (p. 44)

   Marr's claim that the structure of the real world figures in determining the nature of the representations that are attributed in the theory is tantamount to the chief point about representation or reference that generates my nonindividualistic thought experiments—the first step in the schema. I shall show how these remarks constitute the central theoretical orientation of the book.

   Calling the theory Marr's is convenient but misleading. Very substantial contributions have been made by many others, and the approach has developed rapidly since Marr's death. See, for example, D. Ballard, G. Hinton, and T. Sejnowski, "Parallel Vision Computation," *Nature* 306 (November 1983): 21–26. What I say about Marr's book applies equally to more recent developments.
4. It is an interesting question when to count the visual system as having gone intentional. I take it that information is, in a broad sense, carried by the intensity values in the retinal image; but I think that this is too early to count the system as intentional or symbolic. I'm inclined to agree with Marr that where zero crossings from different-size filters are checked against one another (see example 1), it is reasonable to count visual processes as representational of an external physical reality. Doing so, however, depends on seeing this stage as part of the larger system, in which objective properties are often discriminated from subjective artifacts of the visual system.
5. Marr 1982, pp. 68–70. See also D. Marr and E. Hildreth, "Theory of Edge Detection," *Proceedings of Royal Society of London* B 207 (1980): 187–217, where the account is substantially more detailed.
6. Marr 1982, pp. 111–116, 205–212; D. Marr and T. Poggio, "A Computational Theory of Human Stereo Vision," *Proceedings of Royal Society of London* B 204 (1979): 301–328; S. Ullman, *The Interpretation of Visual Motion* (Cambridge, Mass.: MIT Press, 1979).
7. As I have intimated above, I doubt that all biological (including physiological) processes and states in the human body are individualistically individuated. The failures of individualism for these sciences involve different but related considerations.
8. Descartes went further in the same direction. He thought that the perceptual system, and indeed the intellect, could not make a mistake. Mistakes derived from the will. The underlying view is that we primarily perceive or make perceptual reference to our own perceptions. This position fails to account plausibly for various visual illusions and errors that precede any activity of the will, or even the intellect. And the idea that perceptions are in general what we make perceptual reference to has little to recom-

mend it and, nowadays, little influence. The natural and, I think, plausible view is that we have visual representations that specify external properties specifically, that these representations are pre-doxastic in the sense they are not themselves objects of belief, and that they sometimes (when they result from abnormal processes) fail to correctly represent what is before the person's eyes.

9. Of course, at least in the earliest stages of visual representation, there are analogies between qualitative features of representations in the experienced image and the features that those representations represent. Representations that represent bar segments are bar-shaped, or have some phenomenological property that strongly tempts us to call them bar-shaped. Similarly for blobs, dots, lines, and so forth. (See page 211 of Marr and Hildreth 1980 for a remark on this dual aspect of representations.) These "analogies" are hardly fortuitous. Eventually they will probably receive rigorous psychophysical explanation. But they should not tempt one into the idea that visual representations in general make reference to themselves, much less into the idea that the content of objective representation is independent of empirical relations between the representations and the objective entities that give rise to them. Perhaps these qualitative features are constant across all cases where one's bodily processes, nonintentionally specified, are held constant. But the information they carry, their intentional content, may vary with their causal antecedents and causal laws in the environment.

10. There is no need to assume that the abnormal condition is unverifiable. Another person with relevant background information might be able to infer that the abnormal condition is producing a perceptual illusion. In fact, another person with different dispositions might even be able to perceive the difference.

11. Not all perceptual representations that specify objective entities need have their representational characteristics determined in this way. The representational characters of some visual representations (or states) may depend on the subject's background theory or primarily on interaction among other representations. There are hallucinations of purple dragons. (Incidentally, few if any of the perceptual representations—even the conscious perceptual representations—discussed in Marr's theory depend in this way on the subject's conceptual background.) Here, I assume only that some visual representations acquire their representational characters through interaction. This amounts to the weak assumption that the formation of some perceptual representations is empirical.

    Some of interaction that leads to the formation and the representational characters of certain innate perceptual tendencies (or perhaps even representations) may occur in the making of the species, not in the learning histories of individuals. Clearly this complication could be incorporated into a generalization of this third premise without affecting the anti-individualistic thrust of the argument.

12. What of the nonintentionally specified dispositions that in the actual environment (given the actual laws of optics) would have enabled P to discriminate *C*s from *O*s in ideal circumstances? In the counterfactual environment, in view of the very different optical laws and different objects that confront P, one can suppose that these dispositions have almost any visual meaning that one likes. These dispositions would serve to discriminate *C*s from some other sort of entity. In view of the objectivity premise, the nonintentional dispositions can always be correlated with different, normal antecedent laws and conditions—in terms of which their intentional content may be explained.

    The argument of this section is developed in parallel but different ways in my "Cartesian Error and the Objectivity of Perception" (see note 1 above).