1989). As to the visual system, the pattern of connectivity and chemoarchitecture of the two major ascending visual pathways to the forebrain is closely comparable in birds and mammals (Güntürkün 1996). At present, it is unclear whether all these similarities reflect homoplasy (convergent evolution) or homology (shared evolutionary ancestry). But, whatever the answer, the possibility that some very general computational rules are conserved, or reinvented, in classes whose lineages have been separated some 250–300 million years ago is intriguing. It suggests that there are important constraints related to the geometrical and physical properties of the world that must have been incorporated in the design of any efficient biological visual system.

Toward a generative transformational approach to visual perception

Douglas Vickers

Psychology Department, Adelaide University, Adelaide, SA 5108, Australia.
Douglas.Vickers@adelaide.edu.au
http://www.psychology.adelaide.edu.au/members/

Abstract: Shepard's notion of "internalisation" is better interpreted as a simile than a metaphor. A fractal encoding model of visual perception is sketched, in which image elements are transformed in such a way as to maximise symmetry with the current input. This view, in which the transforming system embodies what has been internalised, resolves some problems raised by the metaphoric interpretation.
[HECHT; SHEPARD]

SHEPARD has argued that the human brain has evolved in such a way as to internalise the most general principles that operate in the physical world. Although acknowledging the sweep and suggestiveness of this idea, most commentators, such as **HECHT**, question the usefulness of the metaphor of internalisation and argue that it is falsified when made specific, and unhelpful when it remains general. It is with some trepidation, therefore, that I suggest that a quite concrete instantiation of these most general principles can provide useful insights into the way in which perceptual and cognitive processes might conceivably operate.

Independently of the validity of internalisation, it seems obvious that a comprehensive and coherent explanation of the interactions between brain processes and the physical world (and between different sets of brain processes) must be in terms of a common conceptual framework. In agreement with proponents of nonlinear dynamical analyses (e.g., Port & Van Gelder 1995), I would argue that the most general common framework is that of geometry.

Within this framework, two of the most general and powerful notions for explaining and guiding our understanding of the physical universe are: (1) symmetry under some set of transformations; and (2) some form of minimum (i.e., optimisation) principle. These are the two constraining principles, in particular, that **SHEPARD** argues have been internalised. In proposing that, in default conditions, human perception tends to conform to the principles of kinematic geometry, SHEPARD attempts to map these concepts directly onto mental functioning.

I wish to suggest that a more useful interpretation of **SHEPARD**'s position is that perceptual and cognitive processes operate as if such general principles as symmetry and optimisation had somehow been internalised. In order to justify this belief, I should like to present what is still more a thought experiment than a well worked-out hypothesis. This idea takes its inspiration from recent developments in fractal geometry and, in particular, from the use of fractal encoding to compress and generate visual images. First, a very brief explanation of these ideas may be helpful.

Fractal geometry is concerned with the analysis of complex lines, surfaces, and objects that are comparable in complexity to the outlines and composition of natural objects and textures. As an example, a fractal curve can be generated by taking a simple seed element, such as a straight line (—), and making two reduced copies of it (– –). These reduced copies are subjected to a set of transformations (e.g., ∨). Two reduced copies are then made of each line element in this output and the same set of transformations is again applied. The process can continue indefinitely, but, in practice, five or six iterations are sufficient to approach the resolution of the system used to display (or to view) the resulting curve. More complex sets of networked probabilistic generation processes can also be used to create highly complex images.

Fractal curves and surfaces are interesting for several reasons. First, such curves and surfaces frequently resemble natural objects and environments (Mandelbrot 1983; Prusinkiewicz & Lindenmayer 1990). Second, like certain ferns, fractal objects and images exhibit "self-similarity" (i.e., their structure is statistically similar at different scales of magnification, so that a small part resembles the whole structure). Third, all the information necessary to generate a fractal curve can be represented by the parameters of a set (or collage) of some half dozen transformations.

It can be proved that any image, however complex, can in principle be represented by the parameters of a set of fractal generation processes. Because of this, as Peitgen et al. (1992, p. 259) remark, "Fractal geometry offers a totally new and powerful modelling framework for such encoding problems. In fact, we could speculate that our brain used fractal-like encoding schemes." My proposed thought experiment is that we try to work out some of the consequences of taking this suggestion seriously.

For example, as a first step in this direction, K. Preiss and I have devised a program that takes any regular fractal curve (e.g., the well-known Koch curve) and uses a plausibly constrained sequence of transformation operations to discover (by mapping the transformed patterns onto the original) the reduction factor, number of copies, translation, rotation and reflection parameters, and the number of iterations required to generate a copy of the curve. This copy may then be matched with the original. I propose that such a program can represent a highly simplified, toy model of one way that the perceptual process might conceivably operate.

To make this model more general, we might go on to speculate that the visual system carries out a collage of the simplest transformations on elements of the visual image that will maximise symmetries between the transformed output and the current image (and with changes in that image), given the rate of change in the input and the physical parameters controlling the transforming systems. That is, visual perception is conceived as the attractor-like output of a dynamic, generative transformational system that "resonates" with the current input.

From this perspective, the transforming system embodies what has been internalised. The principles of symmetry maximisation and optimisation are obvious candidates as hypotheses for describing the operation of the system. However, hypotheses in these terms are now empirically testable. Even if true as descriptions of the system, they are not used to describe the end result of perception, but the operation of the perceptual system *as it interacts with the environment.* As a result, the various examples cited by commentators as falsifying the internalisation notion, become instead empirical findings that may or may not be consistent with the predicted output of the system.

If we think of very general constraints as (possibly) applying to the system that processes external information, then the opposition between well-resolved geometric regularities and "more abstract" statistical regularities, to which **HECHT** draws attention, need not arise. Perception is now seen as determined by the distribution of active transformations that generate an output that is maximally symmetric (statistically) with the current input. That is, perceptual responses are not bound by group-theoretic requirements of perfect symmetry. At the same time, the repertoire of transformations that the system calls on may embody evolutionary developments reflecting bodily and external constraints. Thus, the system can incorporate the kind of prior knowledge (about any and every stimulus) that the Bayesian approach to perception implausibly ascribes to cognition. (It might even incorporate some aspects of HECHT's "externalisation" hypothesis.)

A further consequence of this approach is that it may be able to accommodate one of the features of responses to nongeneric views of objects that otherwise seem puzzling. This is the sudden sharp change in perception that occurs (as when a pencil is viewed end-on). To do this, however, it may be important not to underestimate the richness of the information that may be used to select the appropriate transformations.

Specifically, in our program, we used a Hausdorff distance metric to compare the transformed and the original images (Rucklidge 1996). The process of measuring the Hausdorff distance between two sets of points, A and B, takes into account the distances from each point in set A to every point in set B and the distances from each point in B to every point in A. This measuring process yields a rich landscape of information. In essence, each point in an array has a "view" of every other point.

For example, when applied to a (single) Glass figure, among the distribution of inter-point distances, one single value predominates; all that is necessary is to select the transformation that will move a point through that distance. In other words, from this perspective, some perceptual phenomena that have been seen as problematic because of a correspondence problem (which points go with which), appear to be *over*- rather than under-determined.

The generative transformational approach that I suggest is worth exploring is congruent with some other recent work in cognition, such as Feldman's (1997) treatment of one-shot categorization and Leyton's (1992) speculation that we may be sensitive to the "process-history" of objects, where this history is interpreted as a sequence of transformations from some maximally symmetric original state. However, its main merit, in the present context, is that it can provide a computational model, in terms of which **SHEPARD**'s internalisation hypothesis retains its generality. Perhaps a further appealing feature is that this is achieved by a concrete instantiation of two notions central to Shepard's thinking: transformation and a generative system.

ACKNOWLEDGMENT

I wish to thank Michael Lee for helpful comments.

What's in a structure?

Virgil Whitmyer

Department of Psychology, Indiana University, Bloomington, IN 47405.
Vwhitmye@indiana.edu

Abstract: Shepard's general approach provides little specific information about the implementation of laws in brains. Theories that turn on an isomorphism between some domain and the brain, of which Shepard's is one, do not provide specific detail about the implementation of the structures they propose. But such detail is a necessary part in an explanation of mind. [**SHEPARD**]

I have a general worry about projects like **SHEPARD**'s. Consider a cousin of SHEPARD's theory which has, I take it, a similar goal. The Gestalt psychologists were out to uncover the structure of the brain. Köhler expresses a desire to explore the "*terra incognita*" between stimulation and responsive behavior (1929, p. 54). One might consider such neural territory *qua* realizer of phenomenal states, looking, for example, for opponent process cells that realize sensations of red and green, yellow and blue. On the other hand, she might be interested in the brain *qua* representer, in which case she might ask about the mapping that takes neural elements into a represented domain. Each project requires a mapping between different types of domain: in the former case from phenomenal to brain states, and in the latter from representations to represent*eds.* Köhler's principle interest was in the mapping from phenomenal to brain states. I'm not entirely sure which mapping SHEPARD is principally interested in, despite his usual talk about "representation."

In each case theorists have used the idea that the neural domain is isomorphic to what we might call the target domain (be that a set of representations or a set of phenomenal states). The idea behind such theories is that the domains *share* a structure. This identity of structure allows us to learn about the neural structure by studying the target structure. The most simplistic description of the Gestalt case has it that when we find a circular percept in our "behavioral environment" (Koffka 1935), we can infer that "concomitant physiological processes" (Köhler 1929, p. 61) share this circular structure. And while few theorists today hold that there is such a straightforward mapping, the very same principle lies behind the idea that the psychological color space with its three axes maps onto a neural coordinate space with sets of opponent process cells realizing each axis (e.g., Hurvich 1981).

The same problem haunts every version of such a theory, but it is most easily seen in the spatial case. We know that the sense in which the neural domain has a circular structure will differ from the sense in which the phenomenal domain has that structure – the neural version won't look circular to an observer gazing at the cortex. There *is* a trivial sense in which it has a circular structure just if it realizes a circular percept, but I take it a theorist interested in the nature of the brain would want a better description, like a map of the surface of the cortex. Yet, however we characterize the structure shared by the phenomenal and the neural domain, it must be abstract enough to describe both domains; so it cannot be written in neural terms. Certainly **SHEPARD**'s are not. The case is the same with structures characterized otherwise: if the terms in a functional characterization (a type of structure) do not refer "transparently" to entities and relations in the domain, then we need a translation into the particulars of a domain. Some (e.g., Fodor) maintain that psychological laws are irreducible, hence that no translation is required. But I doubt that SHEPARD is of this school.

SHEPARD has famously invoked isomorphism as well. He writes that "the default motions that are experienced in the absence of external support are just the ones that reveal, in their most pristine form, the internalized kinematics of the mind and, hence, provide for the possibility of an invariant psychological law" (target article, sect. 1.7). These laws do not directly govern brains, any more than the "next-to" relation in the visual field tells us about the next-to relation in cortical cells. Such judgments are reports about the structure of a state space of subjective states, a structure shared by the brain. Shepard has devised ingenious techniques whereby we can learn about the structures of psychological domains, and his present paper describes several. The large-scale project is the same as Köhler's: a set of laws that describe the behavior of phenomenal states (as those in accordance with Chasles' theorem) describes a *structure* of those states. In contrast to Köhler's phenomenal circle, this structure is distributed over time and counterfactual situations. But we must still say how this particular structure is realized in neural stuff. If the realization is abstract, how does this characterization constrain the possible configurations of the neural realizers? How are we to arrive at a translation from the characterization of the structure SHEPARD proposes to the description of a brain?

An important element in **SHEPARD**'s argument is that selective pressures favor the internalization of law-governed regularities. He writes of "the benefits of representing objects *as* enduring entities" (emphasis mine). The most straightforward way such internalization might proceed is to internalize a domain that directly realizes the laws that constitute the structure. But this is as implausible as the direct realization of phenomenal structures; it is more likely the case that we simulate the structure indirectly. But here the distinctions between various domains are important. We are out to find an informative description of the brain that tells us either how it represents or how it realizes phenomenal states. But the mapping from *what* is represented to *how* it is represented (i.e., to the nature of the vehicle) is no more straightforward than