Internalization: A metaphor we can live without

Michael Kubovy and William Epstein

Department of Psychology, The University of Virginia, P.O. Box 400400,
Charlottesville, VA 22904–4400
kubovy@virgina.edu   we7v@virginia.edu
www.virginia.edu/~mklab

Abstract: Shepard has supposed that the mind is stocked with innate knowledge of the world and that this knowledge figures prominently in the way we see the world. According to him, this internal knowledge is the legacy of a process of internalization; a process of natural selection over the evolutionary history of the species. Shepard has developed his proposal most fully in his analysis of the relation between kinematic geometry and the shape of the motion path in apparent motion displays. We argue that Shepard has made a case for *applying* the principles of kinematic geometry to the perception of motion, but that he has not made the case for *injecting* these principles into the mind of the percipient. We offer a more modest interpretation of his important findings: that kinematic geometry may be a *model* of apparent motion. Inasmuch as our recommended interpretation does not lodge geometry in the mind of the percipient, the motivation of positing internalization, a process that moves kinematic geometry into the mind, is obviated. In our conclusion, we suggest that cognitive psychologists, in their embrace of internal mental universals and internalization may have been seduced by the siren call of metaphor.

Keywords: apparent motion; imagery; internalization; inverse projection problem; kinematic geometry; measurement; metaphors of mind

Theorists of perception face two fundamental questions: (1) How does the visual system resolve the inverse projection problem? This problem has been stated as follows: "In classical optics or in computer graphics the basic problem is to determine the images of three-dimensional objects, whereas vision is confronted with the inverse problem of recovering surfaces from images" (Poggio et al. 1985, p. 314). The inverse projection problem is difficult because the environment to stimulation ($E \rightarrow S$) mapping is *noninvertible* (Durbin 1985, Theorem 2.2): for every feature of S there is a corresponding feature in E (i.e., the $E \rightarrow S$ mapping is *onto*), but in all scenes countless different features of S could produce identical features on the retina (i.e., $E \rightarrow S$ is not *one-to-one*). (2) Why is the visual system's solution to the inverse projection problem successful, that is, what accounts for the commonplace fact that the solution typically accords with the way things are?

Shepard responded to these questions with regard to the perception of motion by postulating an internalized kinematic geometry. The most fully developed example of this approach is his treatment of apparent motion. According to Shepard (1994) the fact that apparent movement is perceived at all is owing to the "internalized principle of object conservation" (p. 4). But he is interested in more than the fact that movement is seen. His attention is drawn to the shape of the path adopted by the apparently moving object. Although the number of candidate movement paths joining the two locations is infinite, the perceptual system regularly settles on one. According to Shepard, the preferred path conforms to the principles of kinematic geometry.

In this article, we will claim that Shepard's theory can be divided into two sets of assertions: (1) assertions that the perception of motion is modeled by kinematic geometry, and (2) assertions about the *internalization* of kinematic geometry. We are persuaded by the first set of assertions, and believe that they represent an important advance, but have reservations about the second.

1. The inverse projection problem and internalization

Before we turn to Shepard's notion of internalization, a brief review of the history of this idea will provide a useful frame. Our brief excursion will visit Helmholtz, Transactionalism, Rock's cognitive constructivism, Marr's computational theory, and Gibson's ecological approach.

1.1. Precursors

Although Helmholtz did not label the problem he was solving as the problem of inverse projection, his theory of per-

MICHAEL KUBOVY is Professor of Psychology at the University of Virginia. He has published empirical and theoretical work on the psychology of visual and auditory perception, with a particular emphasis on Gestalt phenomena and perceptual organization. He has also written about the "pleasures of the mind," the psychology of art, and decision making.

WILLIAM EPSTEIN is Professor of Psychology Emeritus at the University of Wisconsin-Madison and Visiting Professor of Psychology at the University of Virginia. He has published empirical and theoretical work on problems of visual perception.

ception is an early attempt to solve it. He proposed (von Helmholtz 1866/1965, p. 153) that perception involves an unconscious deductive inference:

> An astronomer . . . comes to . . . conscious conclusions . . . when he computes the positions of stars in space . . . from the perspective images he has had of them at various times . . . . His conclusions are based on a conscious knowledge of the laws of optics. In the ordinary acts of vision this knowledge of optics is lacking. Still it may be permissible to speak of the psychic acts of ordinary perception as *unconscious conclusions*. . . .

In perception as Helmholtz sees it, one premise of the deductive inference is a law of optics that governs the relation between stimulation (S) and the environment (E), $S \Rightarrow E$, and is acquired by visual learning guided by tactile experience. The inference follows the *modus ponens* rule of first order predicate calculus: $S \Rightarrow E$, S, therefore E. For example, the major premise might be a law regarding the way points in the environment give rise to binocular disparity; the minor premise would affirm the occurrent proximal state (S), the disparity between the right and the left eye images. The conclusion is an assertion about the world, for example, at a certain position in the environment there is an object of such and such a shape (E).

The Transactionalists (Ittelson 1960; Kilpatrick 1961), led by Ames, gave the first clear expression of the inverse projection problem which they labeled the problem of "equivalent configurations." The Transactionalists proposed that the percipient settles on a particular candidate representation by drawing on "assumptions" about the world which assign likelihoods to the candidate solutions.

The Transactionalists' language of assumptions reflects a commitment they shared with Helmholtz which was later promoted by Rock (1983; 1997). It is a commitment to unconscious cognitive processes. According to Rock's cognitive constructivist theory, the perceptual system is stocked with laws (e.g., principles of central projection), and rules (e.g., nonaccidentalness), that serve to direct the solution of the inverse projection problem in an "intelligent" manner. Rock presumes that the knowledge base is represented *in* the perceptual system; it is therefore *internal*.

On the matter of internalization, Rock insisted on distancing his position from Helmholtz's. As noted above, Helmholtz, who was a determined empiricist, postulated that the major premises were themselves the product of individual learning. Rock took a different position. He did not posit that the knowledge base is learned by the individual, but that it is available without prior *individual* learning as the result of learning over the history of the species. Rock proposed to have the cognitive cake without swallowing the indigestible bits of classical empiricism.

The computationalist approach makes the assumption of internalization less urgent. Its goal (Marr 1982; Ullman 1979) is to identify plausible constraints on the environment that can make the $E \rightarrow S$ mapping invertible.

> The main idea for "solving" ill-posed problems, that is for restoring "well-posedness," is to restrict the class of admissible solutions by introducing suitable a priori knowledge. A priori knowledge can be exploited, for example, under the form of either variational principles that impose constraints on the possible solutions or as statistical properties of the solution space. (Poggio et al. 1985, p. 315)

If one can identify the constraints, one has shown how the visual system dissolves the inverse projection problem: When the environment satisfies the posited constraints, the

$E \rightarrow S$ mapping is invertible. (It should be noted, as Edelman 1997 has shown, that some important perceptual tasks, such as recognition, may not require inversion.)

It may seem that the computationalists' constraints are merely reincarnations of the Transactionalists' assumptions or of the rules in Rock's (1983) neo-Helmholtzian account. Indeed, some expressions of the computational approach may encourage this interpretation. Nevertheless, it is important to distinguish between the status of assumptions and rules on the one hand, and constraints on the other. For the transactionalists and the cognitive constructivists, assumptions and rules are lodged in the mind of the perceiver. Even though they are not available for conscious assessment they are represented and are causally active in the perceptual process.

Although it makes sense to ask how assumptions and rules are represented and who it is that uses them, these questions are not properly asked about constraints. Computational constraints are *environmental* regularities that have prevailed in the ecological niche of the species over the course of evolution of the perceptual system. As such, they have shaped the design of computational modules so that their output, given optical input under ordinary conditions, is adaptive. The computationalist theorist needs to know the relevant constraints to proceed to the algorithmic level of explanation. This should not be mistaken to mean that the perceiver needs knowledge of the constraints to see the world as she does. The difference between the cognitive constructive stance and the computational stance may be summarized simply: For the cognitive constructivist, the perceptual system *follows* rules; for the computationalists, the system *instantiates* them.

Gibson (1950; 1966; 1979) took a more radical step of denying the need for assumptions, rules, or constraints, therefore making the question of internalization moot. He thought that the so-called inverse projection problem is a pseudoproblem, owing its origin and persistence to a mistakenly narrow construal of both the objects of perception and the nature of stimulation. He argues that the object of perception is a fully furnished world, not objects detached from settings (or isolated from any setting), and that the stimulations that count for a perceptual system are dynamic structures of light and not static slices of spatiotemporal optical structures. Whoever adopts this (ecological) stance, has described the organism's ecology (E') so that the $E' \rightarrow S$ mapping is both one-to-one and onto, and is therefore invertible. This is a world in which the inverse projection problem does not appear.

Although Gibson's explicit position on the inverse projection problem and internalization may appear to be very different from the stance of the computationalists, his implicit position is actually similar to theirs. His assertion that stimulation and the environment are unequivocally linked, or that stimulation carries "information," is in fact tacitly contingent on the satisfaction of environmental constraints. Perhaps the reason that Gibson was reluctant to give this contingency the prominence it later received in the writings of the computationalists was a fear that talk of constraints so readily slips over into talk about mental entities. Had Gibson become convinced that there was a noncognitive formulation of constraints he might have admitted them explicitly into his theory.

Let us summarize what we have said on the issue of the inverse projection problem and internality. According to

the approach of Marr and Gibson, constraints are neither (1) lodged in the mind, nor (2) are they active constituents in the perceptual process. They are the conditions which the world must satisfy if the computational algorithms are to go through (Marr 1982) or if the claims for information in the spatiotemporal optical structures (Gibson 1966; 1979) are to be sustained. As such, constraints are passive guarantors or underwriters that are external to the perceptual process (representational transformation for Marr; "pickup" of information for Gibson). We should therefore not view them as internal, and hence not internalized. According to the Transactionalists and Rock, assumptions and knowledge are mental contents that are active in the perceptual process. They hypothesize that assumptions and knowledge direct the process of selecting the best fitting (most likely) distal attribution. It is but a small step from that hypothesis to internalization.

1.2. Locating Shepard

How should we locate Shepard's position in this theoretical landscape? At various times he has aligned himself with Helmholtz's stance (see Shepard 1990b, for example); at other times he has resonated to Gibson's resonance metaphor (Shepard 1984, for example). In the target article he does not signal his position. However, our assessment is that Shepard's position is in the neighborhood staked out by Helmholtz and Rock. Like Helmholtz's premises and Rock's rules, Shepard's universals are deemed by him to be mental contents actively engaged in the perceptual process.

What are Shepard's grounds for positing perceptual universals? In the main, perceptual universals are inferences from behavior: observations concerning the preferred shapes of the paths of movement over many studies of apparent motion lead Shepard to attribute kinematic geometry to the visual system. In this respect, his tactics are similar to the procedures adopted by the Transactionalists in inferring the action of assumptions and by Rock (1983) in his *Logic of Perception.*

Why does Shepard suppose that the universals are *internal*? He provides support for this claim by showing that when we purge all support for a percept from external stimulation, the preferred perceptual solutions conform to the putative universals. The paradigm case for Shepard is the invariant period of the earth's rotation mirrored in the circadian period of activity exhibited by animals maintained in a laboratory environment of invariant illumination and temperature. The analogue in human perception is the apparent motion display. In this case all the normal supports for motion have been removed. When under these circumstances, observers nonetheless perceive motion and the paths of movement they see vary in shape in certain systematic ways, Shepard postulates the action of invisible internal principles.

How do these principles of the geometry of motions become internal? Like Rock, Shepard's answer is to posit a process of "internalization"; a process of gradual acquisition driven by natural selection over evolutionary history.

2. Internalization as theory

2.1. Critique of internality

2.1.1. Kinematic geometry as a model. We will couch the first part of Shepard's answer in slightly more formal terms

than he does. Why? Because we wish to avoid confusion by rigorously maintaining the distinction between models and phenomena, which we have done by thinking of kinematic geometry as a measurement model for the perception of motions.

In everyday language the idea of measurement is so deeply embedded, that we do not make the distinction between the fact that object a is heavier than object b and the fact that object a weighs more than object b, that is, numbers are assigned to objects, for example, $\phi(a)$ to a, such that if a is deemed in some empirical fashion to be heavier than b, then $\phi(a) > \phi(b)$. Measurement theorists have shown (Krantz et al. 1971; Roberts 1979/1984; Scott & Suppes 1958; Suppes & Zinnes 1963) that to understand measurement we must focus on the properties of the numerical assignment. In order to do so, we distinguished (as illustrated in Fig. 1): (1) between empirical objects (e.g., different objects, a, b, . . .) and mathematical objects (e.g., real numbers, $\phi(a)$, $\phi(b)$, . . .) And (2) between empirical relations (e.g., "heavier than," $\succ$, and "placed in the same pan of the beam balance," $\oplus$), which apply to physical objects, and mathematical relations (e.g., "larger than," $>$ and "addition," $+$), which apply to elements of the set of real numbers. "Measurement may be regarded as the construction of homomorphisms (scales) from empirical relational structures of interest into numerical relational structures that are useful" (Krantz et al. 1971, pp. 8–9). If we have derived certain fundamental measurement theorems (representation, uniqueness, and meaningfulness) from a set of plausible axioms about the relations that hold among physical objects, then we can guarantee that statements such as the following are true:
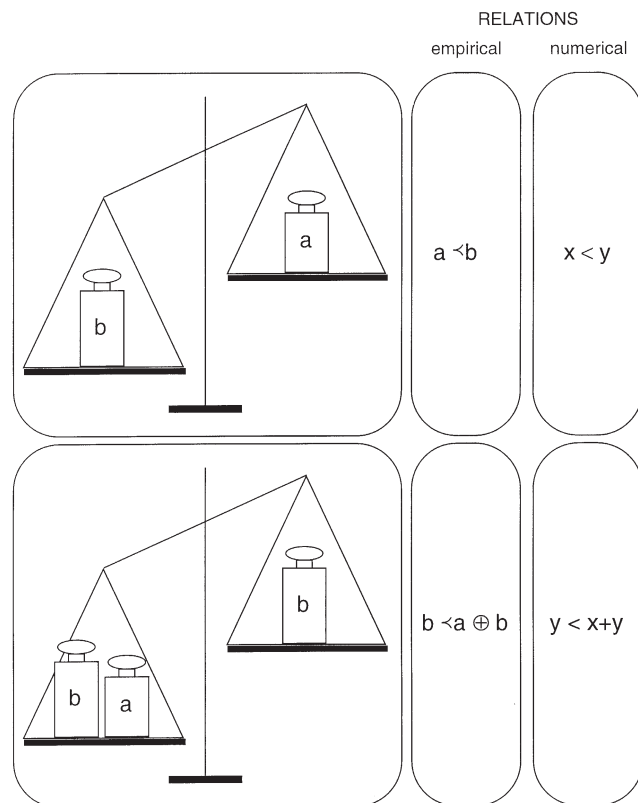

Figure 1.   The distinction between empirical objects and relations and mathematical objects and relations.

Given an empirical relation $>$ on a set A of physical objects and a numerical relation $>$ on the real numbers $\Re$, a function ϕ from $A = \{a, b, \dots\}$ into $\Re$ takes $>$ into $>$ provided the elements $a, b, \dots$ stand in relation $>$ if and only if the corresponding numbers $\phi(a), \phi(b), \dots$ stand in relation $>$. Furthermore, the function ϕ takes $\oplus$ into $+$: $\phi(a \oplus b) = \phi(a) + \phi(b)$.

Kinematic geometry is considered to be a branch of mechanics. What does this mean? It means that we can construct a homomorphism from an empirical relational structure of physical motions into a geometrical relational structure called kinematic geometry. It is somewhat more complicated to say what it means when we say that kinematic geometry is a model for the perception of motion. It requires us to substantiate four claims (Fig. 2):

1. There is a homomorphism k from an empirical relational structure of *physical motions* into a geometrical relational structure called *kinematic geometry;*

2. There is a homomorphism p from an empirical relational structure of *physical motions* into an empirical relational structure of *perceived motions;*

3. There is a homomorphism q from an empirical relational structure of *perceived motions* into a geometrical relational structure which is a *model of the perception of motions;* and

4. There is a homomorphism m from *model of the perception of motions* into a geometrical relational structure, *kinematic geometry.*

There are two ways in which conformity between perceived motion paths and kinematic geometry may be construed. One might give the claim a more modest reading and say that kinematic geometry is a model of the perception of motions, as summarized by the list of four homomorphisms just listed.[1] This means that the visual system proceeds *as if* it possessed knowledge of kinematic geometry. We do not think that Shepard wants to be read this way. We think that he wants to persuade us that kinematic geometry is (1) internal, and (2) has been internalized. This is where our disagreement begins.

2.2. Questioning internalization

Shepard believes that by removing the percipient from ordinary contact with stimulation we can show that kinematic geometry is internal. With the exception of J. J. Gibson and the ecological realists, there is agreement among students of perception that this is a useful, perhaps indispensable strategy for making the invisible visible. Despite this consensus, there is a problem in *exclusive* reliance on nonrepresentative settings. Consider apparent motion. The apparent motion experimental setting – caveats aside, Shepard (1994, p. 10), for example – is unlike the typical conditions of real motion. So while presumably the special conditions of apparent motion may disclose the operation of hidden
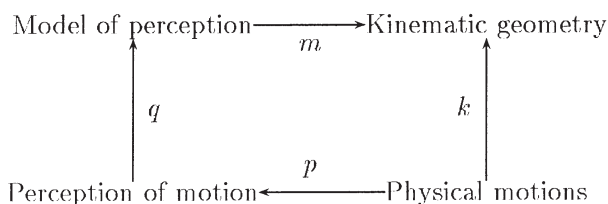

Figure 2.   The simplest construal of the relation between kinematic geometry and perception.

principles, they cannot – taken alone – establish that these hidden principles are implicated in perception of real motion.[2] Most investigators who adopt the tactic argue that what is uncovered under the special conditions applies generally, including also the representative conditions. But we should sign on to this petition only when there is evidence to support it. This requirement can, in fact, be satisfied in certain cases. However, as matters stand currently, there is no evidence in Shepard's work that shows convincingly the involvement of *internal* principles of kinematic geometry in the determination of perceived real motion.[3] And, in places Shepard seems to suggest that he considers the geometry-compliant solutions to be default solutions that should not be expected to appear under conditions of ordinary motion perception: " . . . (under favorable viewing conditions) we generally perceive the transformation that an external object is actually undergoing in the external world, however simple or complex, rigid or nonrigid. Here, however, I am concerned with the default motions that are internally represented under unfavorable conditions that provide no information about motion . . . " (Shepard 1994, p. 7).

It is not incoherent to hold that independent and different principles govern perceived motion paths under such radically different circumstances as those which prevail in the cases of apparent motion and real motion. However, in the context of Shepard's internalization hypothesis an apparent paradox arises. If it is supposed that in the case of ordinary motion – which served up the grist for the evolutionary mill and a process of internalization – there is no evidence of a role for geometrical constraints then how could our distant ancestors have internalized kinematic geometry for application to the special case of apparent motion? It is easy to imagine generalization from the ordinary case to the special case, assuming counterfactually that ordinary real motion did exhibit a determining role for geometric constraints. But it is hard to see how internalization of kinematic geometry could have proceeded independently of the same development in the case of ordinary motion.

Even when the preceding concerns are set aside, the internalization hypothesis suffers from a number of shortcomings. It has no obvious empirical content and cannot be tested experimentally. Moreover, the cash value of the "internalization" hypothesis is questionable. Functionalism, as a starting point, is not what is at stake here. We agree that questions of function (what Marr 1982 has called the "computational theory") are very important. It matters whether one supposes that the function of the heart is to pump blood or to produce audible thumps. And it matters whether one supposes that the function of the perceptual system is to deliver descriptions (representations) of the environment (Marr 1982) or to support action (Gibson 1979) or both (Milner & Goodale 1995). While an argument can be made that neither functionalism as an -ism nor evolution by natural selection is necessary in assessments of function, we will accept that an evolutionary stance helps focus attention on function. (This stipulation notwithstanding, in the history of biology few discoveries rival Harvey's discovery of the function of the heart which Harvey made without the help of evolutionary theorizing or a self-conscious functionalist stance.)

The matter at issue here is, in what way does speculation concerning the origins of a function in the remote part (Pleistocene) contribute to an understanding of the process that subserves the function?

Is there a single parade case that can be trotted out to show the power of the internalization hypothesis to reduce the number of candidate process hypotheses? We cannot think of a case in the field of perception. Consider two hypothetical knowledge states: (1) In the first state, we have come to know in great detail the brain structures and activity underlying stereopsis by studying the contemporary visual system exclusively, with no reference to phylogenetic development. (2) In the second state, we know everything that we know in the first state and in addition – based on a study of the fossil record – we also have come to know (that is, to tell a plausible story about) the evolution of the brain structures. What is the advantage of 1 over 2 when it comes to understanding how the visual system computes depth from disparity? We do not see how the knowledge of origins constrains the choice among the plausible algorithms.

In places, Shepard suggests that the internalization route to acquisition avoids the embarrassments of the rival origins story which "leaves it to each individual to acquire such facts by trial and possibly fatal error" (1994, p. 2). In this respect, Shepard and Rock are moved to posit internalization by the same considerations. But the argument from selective refutation can be misleading when it causes us to fail to notice that the surviving hypothesis has its own defects. A condition that must be satisfied for plausible postulation of internalization is that an enduring pervasive external regularity must be obvious. The exemplary case cited by Shepard is the external day-night cycle and the internal sleep-wake cycle. But where is the pervasive external regularity in the case of object motions? As Proffitt[4] reminds us, "real motion cannot violate kinematic geometry" and therefore the laws of kinematic geometry are a superset of the laws that could be extracted by the visual system from the data offered by real motions. (D in Fig. 3). Therefore, only the latter might have been internalized. As Todorovič (1996) has remarked:

> A source of the problem may be an inappropriate analogy between the operation of the perceptual system and the operation of scientific inquiry. Concretely, the idea is that the principles that the perceptual system has extracted from the environment during evolution are in certain aspects analogous to the principles, such as Chasles's Theorem, that the scientific community has formulated in the last centuries. Such an idea is intriguing, but it should be treated with caution. Perceptual and intellectual activities are to a certain extent related, but they are also quite different. (pp. 17–18)

Because, as we have just seen, the laws of kinematic geometry are a superset of the laws of real motions, Proffitt and Kaiser (1998, sect. II.D, pp. 184–90) see the relation between internalization and evolution differently. They point out that Cassirer's (1944) view of perception is a powerful argument against the belief that the conformity of perception to kinematic geometry has its roots in ontogeny, in keeping with Helmholtz's thought, or in phylogeny, in keep-
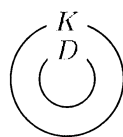
ing with Rock's ideas. In order to make sense of Shepard's views, Proffitt and Kaiser (1998) read him as if he were following Cassirer's neo-Kantian convictions. We think that too much of what Shepard (1994) says would have to be ignored to interpret him this way.

3. The pragmatics of internalization

Having read our argument against the use of the concept of internalization, the reader may ask, why are the authors trying to legislate the use of the term internalization? After all, science does not progress through a progression of terminological refinements, but through a cycle of progressively more refined data collection and theory construction.

We have chosen to answer this question in four ways. First, with a distinction between two kinds of terminological strictures in psychology. Second, with a case study: how an apparently innocuous choice of language, by Shepard, led to subsequent confusion. Third, with a comparison of Shepard's use of internalization with the use of cognate terms used in psycholinguistics. Fourth, with a discussion of the appeal of metaphors in the formulation of theories.

3.1. Two kinds of terminological strictures

We would like to make clear that our goal is to persuade our colleagues to take particular care in the choice of *theoretical* terms, which is different from trying to reform the use of *descriptive* terms. An example of the latter occurs in the literature on anthropomorphism. For example, Crist (1999, p. 22) quotes Barnett (1958, p. 210):

> Darwin . . . took it for granted that terms like *love, fear,* and *desire* can usefully be employed to describe the behaviour of animals – or at least of mammals – generally. He accepted the colloquial use of the word emotion. In doing so he assumed (by implication) that other species have feelings like our own. . . . Since his time it has gradually been found more convenient to describe animal behaviour, not in terms of feelings of which we are directly aware only in ourselves, but in terms of the activities which can be seen and recorded by any observer; we may also try to describe the internal processes that bring these activities about.

Barnett does not propose to purge "emotion" from the *theoretical* vocabulary of psychology:

> If the word emotion were to be used in the scientific study of animal behaviour, its meaning would have to be shifted from the familiar, subjective one: it would have to be used to refer, not to feelings, but to internal changes which could be studied physiologically. (Barnett 1958, p. 210)

In other words, if the term was found to be useful in theories about humans and other animals, there would be no harm in suggesting that other species may have subjective experiences similar to those reported by humans.

We are arguing that the use of the term "internalization" in *theories* may lead to confusion. We turn now to an example of how such confusion could arise.

3.2. Do scientific terms matter?

One of us (Kubovy 1983), in an enthusiastic review of Shepard and Cooper (1982), had one criticism of this book. It was not a criticism of the methodology or of the theory, but of the way the results were formulated. The authors of the book often used expressions such as "the rotation of men-

Figure 3. The relation between the laws of kinematic geometry (K) and the laws that can be extracted from data provided by real motions (D).

tal images." Kubovy thought that there was a danger that such expressions might mislead readers into thinking that more was being claimed about mental imagery than was implied by the data. Where was this rotation taking place? In the mind? In the brain? Kubovy suspected that locutions such as "mental rotation" could lead people to think that a black box had been pried open without so much as an EEG.

Without doubt, the brilliance of the research done by Shepard, Cooper, and their collaborators warranted its enthusiastic reception by a cognitive community that was beginning to mature and was looking for a rallying point. It was a community looking for a phenomenon that would convince psychologists and nonpsychologists alike that the cognitive approach was producing clear and important results. And so the cognitive community was a bit quick to attribute achievements to this research that went beyond the evidence given.

The problem Kubovy saw was created by the way Shepard and Cooper summarized their results: they spoke as if they could show that in their experiments a physical action ("rotation") was being applied to a mental entity ("mental image") whose content was a physical object ("random polygon"). This linguistic form surely implies that something physical is happening to the mental image. Now this sort of careless expression is not uncommon in psychology. For example, we may talk about the decay of memory, which, however subtly, suggests that something (the engram?) is decaying somewhere (in the brain?). Nevertheless, there was less opportunity for readers to be misled when they were learning about memory than when the topic was mental imagery. The reason is this: In discussing imagery, Shepard was also talking about the form of processing that was taking place: it was an analog process rather than a digital one.

We believe that the cognitive community did not only infer from these forms of expression that somewhere an image was rotating, but that the data could support a claim about the *computational implementation* of this rotation: that it was analog rather than digital. Of course this led to a great deal of controversy (Anderson 1978), but the damage was done. The controversy focused on the indeterminacy of the computational implementation, but this focus did not remove the impression that the topic under discussion was the rotation of mental images.

Can one formulate the results in this field without risking being misunderstood? According to Kubovy (1983) one should avoid locutions of the form

$$(1) * \Phi\text{-ACTION}(\Psi\text{-CONTAINER}(\Phi \text{ OBJECT})),$$

where Φ stands for "physical," and Ψ stands for "mental." Instead, we should use expressions of the form

$$(2) \Psi\text{-ACT}(\Phi\text{-TRANSFORMATION}(\Phi \text{ OBJECT})).$$

For example, instead of talking about a person "rotating a mental image of an object," one might say the she is "imagining the rotation of an object." Not only does form (2) avoid sandwiching the mental between two physical descriptions, but it also avoids implying that a mental entity is a thing by turning it into a verb ("imagine") rather than a noun phrase ("mental image").

As an exercise, let us apply this improved form of speech to the question of the analog nature of mental rotation. If instead of wondering whether the rotation of the mental image is analog or digital, we can ask: does the evidence support the claim that when people imagine the rotation of an object, they imagine the object undergoing a continuous rotation? When rephrased in this way, the question of the analog nature of mental transformations seems less of a puzzle, but also less of an achievement, because phenomenology seems to provide a prima facie answer, yes. And furthermore, the data in no way contradict phenomenology.

By the time Kubovy invited Shepard to contribute to the section he edited of the *Handbook of Perception and Human Performance* (Fink & Shepard 1986), Shepard had accepted the criticism and was eager to have the terminology of the chapter conform with Kubovy's terminological suggestion. Kubovy was concerned that this stylistic constraint would lead to awkwardness of expression. We do not think it did, but we invite our readers to judge for themselves.

The topic went dark for a while until Nelson Goodman (1990) expressed concern about cognitive psychologists' talk about "pictures in the mind." In his reply Shepard (1990c, p. 370), cited Kubovy's review and concluded:

> Moreover, the latter, more carful reformulation brings out the essential feature of mental imagery as I recommend we understand it: mental imagery is of external objects, and is therefore to be defined and studied not as some strange, non-material "picture in the head" but in relation to potential test stimuli that are both external and physical.

Unfortunately, it appears that the confusion has not abated. In a dialog between two distinguished French scientists, Jean-Pierre Changeux – a neurobiologist – and Alain Cannes – a mathematician, the latter says:

> Reading your book *Neuronal Man,* I was surprised to realize how much is understood about the brain. . . . I was impressed too by Shepard's mental rotation experiments, in which a subject is asked if two objects are the same after rotating them in three-dimensional space. They show that the response time is proportional to the angle of rotation, and thus that cerebral function obeys physical laws. (Changeux & Cannes 1995, p.5)

3.3. What has Shepard been trying to do?

When one is dealing with a figure as important to the history of our field as is Shepard, one sometimes better understands the breadth and depth of the person's thinking by elucidating certain parallels that guide his thinking. We suspect that Shepard has been looking for a model of perception that shares some important features with Chomskyian linguistics.[5] In 1981, Shepard offered the diagram shown in Figure 4 to illustrate his concept of psychophysical complementarity and illustrate its application to mental rotation (Shepard 1981b). It is particularly interesting to note that he calls the internal representation "Deep Structure."

The Poggio et al. (1985) argument that the environment to stimulation ($E \rightarrow S$) mapping is noninvertible is parallel to a similar insight of psycholinguistics: "A device capable of [developing the competence of a native speaker] would have to include a device that accepted a sample of grammatical utterances as its input . . . and . . . would produce a grammar of the language . . . as its output. . . . To imagine that an adequate grammar could be selected from the infinitude of conceivable alternatives by some process of pure induction on a finite corpus of utterances is to misjudge completely the magnitude of the problem" (Chomsky & Miller 1963, pp. 276–77).

More specifically (Pinker 1984), suppose children learned

DISTAL   PROXIMAL   INTERNAL

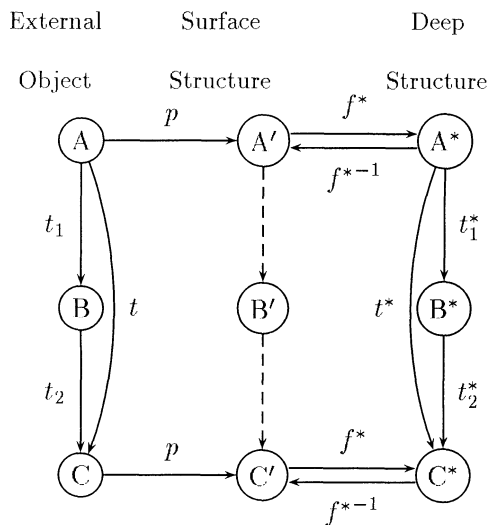STIMULUS   STIMULUS   REPRESENTATION

External   Surface   Deep

Object   Structure   Structure

Figure 4. Shepard's schema of the projective (p), formative (f^*), and transformations mappings between objects (A, B, and C), proximal stimuli (A′, B′, and C′), and internal representations (A*, B*, and C*). (Redrawn from Shepard 1981b, Fig. 10–1, p. 295.)

their language (which we will denote T for "target") by induction. Before the child has been exposed to T, it could hypothesize what the rules of the language might be (let us denote the set of hypothesized rules by H). At that point there will be no overlap between T and H (Fig. 5a): none of the child's utterances is grammatically well formed. For example, the child never says *We went,* or *We broke it,* and always says *We goed* and *We breaked it.* Because the items of T (marked "+") to which the child is exposed can serve as *positive evidence* that H is incorrect (because these items are not members of H), H will grow and come to overlap with T (Fig. 5b): The child might say *We went,* but persist
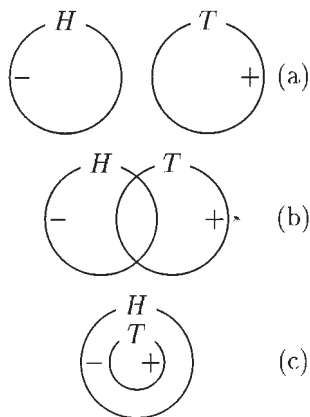
Figure 5. Three situations a child could be in while learning a language. Each disk represents the set of sentences constituting a language. H stands for "hypothesized language," T stands for "target language."

in saying *We breaked it.* But now the child faces two problems: learning the further rules of T, and eliminating the incorrect hypotheses of H (marked "−"). This could occur only if parents corrected their children's incorrect grammar. But, according to Pinker (1984), parents do not provide the required *negative evidence.* If this were the case, children would end up knowing a superset of T (Fig. 5c). This is one reason why Pinker argues that children are endowed with inborn constraints about the possible form of linguistic rules.

Shepard's argument about the internalization of kinematic geometry is parallel to the linguistic argument. This linguistic argument runs: children would not be able to learn a grammar unless they were endowed with inborn linguistic constraints. These constraints are internalized. Analogously, since kinematic geometry is a superset of what can be observed in real motions (Fig. 3), human beings could not have acquired it unless they were endowed with inborn geometric constraints that corresponded to kinematic geometry. Therefore, the argument runs, kinematic geometry is internalized. The difference between the two domains is this: In language acquisition, the inborn constraints insure that eventually the size of H will diminish and coincide with T. In perception, the mind comes equipped with K, and loses nothing by using kinematic geometry to resolve ambiguities that result from missing information in the sensory input.

3.4. Metaphors of mind

Confusing forms of expression with respect to topics in cognitive science could be due to the metaphorical nature of many abstract concepts, as Lakoff and Johnson (1990) showed persuasively in their book *Metaphors We Live By.* Lakoff and Johnson (1999, pp. 235–36) summarize the evidence that "there is an extensive subsystem of metaphors for mind in which the mind is conceptualized as a body." When the MIND IS A BODY metaphor is applied to thinking, two metaphors of smaller scope come into play: THINKING IS PHYSICAL FUNCTIONING and IDEAS ARE ENTITIES WITH AN INDEPENDENT EXISTENCE. When the two metaphors are combined, we get THINKING OF AN IDEA IS FUNCTIONING PHYSICALLY WITH RESPECT TO AN INDEPENDENT EXISTING ENTITY.

In particular, confusion with respect to mental rotation may be attributed to a common form of the metaphor that equates thinking with functioning physically and ideas with independent existing entities. This is the metaphor THINKING IS OBJECT MANIPULATION: "ideas are objects that you can *play with, toss around,* or *turn over in your mind*" (Lakoff & Johnson 1999, p. 240). So it is natural to speak of rotating a mental image. And because more often than not, the metaphorical nature of our thinking is unconscious, it is not easy for the scientist to see how such expressions could be misleading.

The metaphor of internalization is rooted in another member of the same family of metaphors: ACQUIRING IDEAS IS EATING (Lakoff & Johnson 1999, pp. 241–43). This metaphor leads us to compare an interest in ideas to an appetite (e.g., *thirst for knowledge*). We refuse to *swallow* bad ideas, but we *can really bite into* ideas that are *meaty.* The term internalization is but a slightly disguised synonym of ingestion.

4. Assessment and recommendation

Although the notion of internalization is an appealing metaphor, it does not add to the power of Shepard's theory according to which certain important aspects of perception are captured by kinematic geometry. This and other episodes in the history of our field lead us to recommend that we strip our scientific writing of metaphors we can live without, and, until theoretical and empirical progress suggest and support metaphorical terms, that we formulate our theories in as neutral a language as we can.

NOTES
1. The reader will surely agree that to demonstrate that kinematic geometry is a model of perception in this sense would be a major achievement.

2. We hasten to note that there actually is considerable evidence to support the contention that the same mechanisms underlie real and apparent motion. Our argument here is about scientific strategy, not about apparent motion.

3. Writing elsewhere on the evolution of principles of the mind, Shepard (1987a, p. 266) remarks that "the internalized constraints that embody our knowledge of the enduring regularities of the world are likely to be most successfully engaged by contexts that most fully resemble the natural conditions under which our perceptual/representational systems evolved."

4. In comments on a draft of this article, personal communication, March 10, 1999.

5. This section may appear to some readers as an exercise in hermeneutics, but since we were enlightened by it, we thought that our colleagues might be too, and therefore might tolerate our heretical use of a method that would be listed in the *Index of Forbidden Methods* if such an index were prepared.