JOURNAL OF MATHEMATICAL PSYCHOLOGY **41**, 145–170 (1997)
ARTICLE NO. MP971154

The Structure of Perceptual Categories

Jacob Feldman

Department of Psychology, Center for Cognitive Science, Rutgers University

When presented with a small set of sample objects, human observers have the striking capacity to induce a more general class. Generalization can even proceed from a single object (''one-shot categorization''). The inference is apparently guided by the principle that a good categorical hypothesis is one in which the observed object would be a typcal, ''non-accidental,'' or *generic* example; this idea is formalized here as the Genericity Constraint. In the theory proposed here, each categorical hypothesis is a ''generative model,'' a sequence of transformations by which the object is interpreted as having been created; objects are considered to be in the same category if they were created by the same set of operations. The set of all available category models can be explicitly enumerated in a lattice, an explicit structure that partially orders the models by their degree of regularity or genericity—more abstract models are higher in the lattice, and more regular or constrained models are lower. The Genericity Constraint dictates that among all the models on the lattice that apply, the observer should choose the one in which the observed object is generic, which is simply the lowest in the partial order. A series of experiments are reported in which subjects are asked to generalize from simple figures. The results corroborate the role of the lattice and the Genericity Constraint in subjects' interpretations.

1. THE STRUCTURE OF A GENERALIZATION

This paper investigates the way human observers, presented with a small set of objects, can generalize to a larger class, which constitutes the intuitively "natural" category from which the examples appear to have been drawn. Human observers exhibit a remarkable degree of unanimity in this sort of perceptual abstraction, even when using only a single object as an example (*one-shot categorization*). Since such an inductive problem is inherently ambiguous to an extreme degree, this competence suggests the presence of enormous constraint on the generalization faculty. Yet even a few simple examples suffice to show that observers seem to pick and choose quite selectively about which properties of an observed object are fit to be abstracted over, and which retained, in the inferred general class. The constraints on generalization are heavy, but they seem to be deft.

The induction problem investigated in this paper is a narrow one. The objects are all simple, readily parameterizable geometric figures, and there is usually only one sample object (three in some experiments). Yet this problem captures many of the essential difficulties common to all induction problems, deriving from the complete absence of a priori constraint on the solution. The goal of the paper is to construct an explicit inference theory for the problem, specifying the logic by which inductive hypotheses are constructed and selected. The logic here takes the tangible form of a lattice, an explicit algebraic enumeration and ordering of categorical hypotheses. Selection of a hypothesis from the lattice is dictated by the Genericity Constraint, a rule based on the principle that sound generalizations are those that minimize unexplained coincidences in the observations.

A few examples illustrate the approach (Fig. 1). In each row, consider the example object given on the left. What is the category? Some possible "intuitive" answers are given toward the right.

In the top row, it seems natural to infer the category "dot on line." That is, we infer that in other examples of the same category, the dot will always appear somewhere *on* the segment, rather than at some arbitrary point in nearby space, although its position *along* the line might vary (as might the overall position, orientation, and size of the figure). Indeed, about 90% of subjects make exactly this generalization (see Section 3), a remarkable consensus. However "reasonable" such an inference may seem, it entails an enormous inductive leap which needs to be explained. Observers apparently guess that the fact that the dot is in contact with the line segment is some kind of *invariant* property of the category—and hence is liable to be "causally meaningful"—while the

Address correspondence and reprint requests to: Jacob Feldman, Department of Psychology, Center for Cognitive Science, Rutgers University, Busch Campus, New Brunswick, New Jersey, 08903. E-mail: jacob@ruccs.rutgers.edu.

This paper is derived from a doctoral dissertation in the M.I.T. Department of Brain and Cognitive Sciences. (The original dissertation is available as Technical Report 6 of the Rutgers Center for Cognitive Science.) I am especially grateful to Whitman Richards, who supervised the original dissertation and inspired the work. I also thank Allan Jepson, Steven Pinker, Shimon Ullman, and Alan Yuille for extensive comments on earlier drafts. I am extremely grateful to Leola Alfonso-Reese, In Jae Myung, and F. Gregory Ashby whose comments on a revision enormously strengthened the manuscript. Thanks go to Marie Coppola, Jane Gwiazda, Tatiana Schnurr, Sara Solnick, Karin Stromswold, and Josh Tenenbaum for help in collecting the data. The research was supported by AFOSR Grant 89-0504 to Whitman Richards and was revised while the author was at the Center for Cognitive Science at Rutgers University, New Brunswick, NJ.

0022-2496/97 $25.00

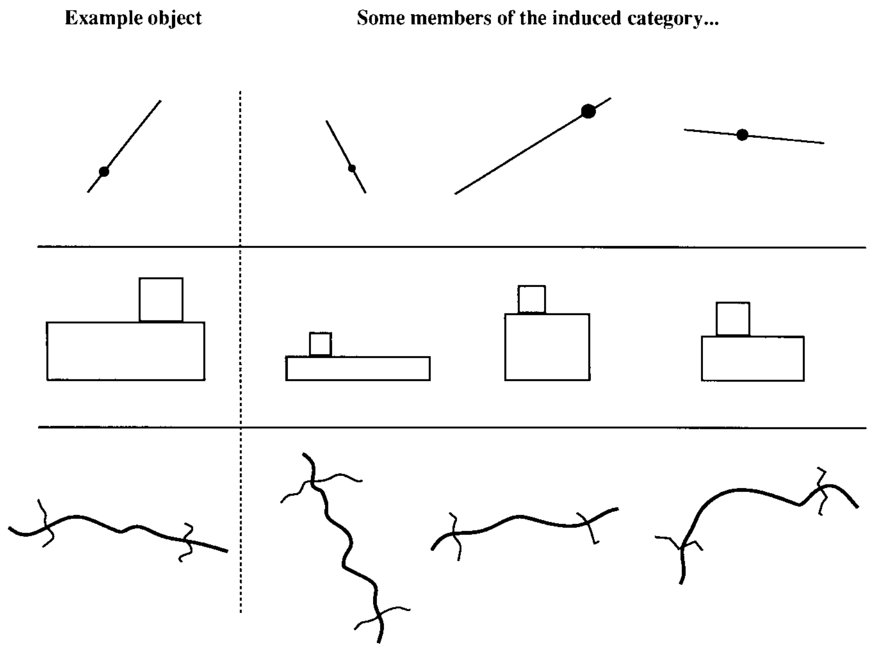

FIG. 1. "One-shot categorization" is shown. Given each example in the left column, most observers can immediately induce a more general (perhaps infinte) class, some members of which are shown to the right. In the top row, for example, observers induce *dot on line*, meaning that they guess that novel examples of the category will always have the dot touching the segment, although its position along the segment may vary.

dot's precise position along the line seems more likely to be an ephemeral property of the particular example. Other examples are given in the figure. In each case certain „suspicious" (i.e., unlikely to be accidental) features are preserved in the generalization, while others are abstracted over.

Following an argument of Leyton (1984, 1988, 1989, 1992), such a generalization can be regarded as the inference of a *generative model* for the category: a process or procedure that generates legal members by applying a limited set of transformations to some primitive object. In the dot-on-line example, one can imagine that the induced set might be produced by applying the transformation *translation along the line segment* to some simpler form, such as a line segment with the dot at its endpoint. Seen in this light the generalization competence can be viewed as the identification of the observed object with an appropriate generative model. The puzzle is that the inferred generative model is more abstracted than the observed object (it can produce objects in which the dot is at positions different from in the original example), but it is not *completely* abstracted (for example, it cannot produce objects in which the dot falls completely off the line segment).

Note that the transformations seemingly allowed in the generalizations are different from what is usually thought of as "shape" in the perceptual literature. The transformations that preserve category membership (beyond translation, rotation, and scale) change from category to category. In the dot-on-line example, translation of the dot along the line is allowed, though it is not even definable in the other two examples. In the middle example, the figure is not rotated freely (because the example seems to be sitting flush on the ground, a special configuration). In the third row, smooth bending of contours seems to be allowed, but not in either of the upper rows. Moreover, there are different *numbers* of free parameters in each of the solutions, resulting in categories with different dimensionalities. This sort of class-dependent conception of legal transformations is clearly distinct from the usual idea of a fixed shape representation. Rather, the object is regarded as a point in a generalized parametric space, and the category as a region in the space—the "consequential region" of Shepard (1987). But what kind of region? Not just any point set will do. What is the characteristic geometry of "good" categorical regions? That is, what is a category?

Regularity and Property Covariation in the World. It is now a common (though not entirely uncontroversial) view in cognitive science that categories derive their existence from regularity that obtains in the world—the fact that some properties tend to occur with greater consistency in the presence of certain other properties (see Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976; Smith and Medin, 1981; Quine, 1985). In a strict view, it is actually the equivalence classes formed among these intercorrelated properties that constitute the "prototypes" of the natural categories in the world.

The theory proposed below for perceptual categories is too narrow in scope to account for classically problematic cognitive categories such as "furniture" or even "apple" (but see Feldman, 1992b, for steps in this direction). But it shares the same underlying explanatory principle about the ontogeny of categories, namely that they derive their existence directly from regular constraints on the way properties of objects tend to interrelate. The theory begins, as any regularity-based account should, with a model of what general forms regularities in the world are liable to take—that is, what types of commonalities among disparate objects in the world might reasonably be thought of as "causally coherent" (cf Witkin and Tenenbaum, 1983). The proposed form, the *constraint manifold*, specifies the geometrical structure that useful categorical properties (viewed as regions of the object space) ought to exhibit.

However, a theory of one-shot categorization in which the categories are based on property covariation faces a peculiar paradox. In the one-shot case, by definition there is no variation of properties, and hence no *co*variation. Before any property variation has been observed at all, how can any guess about future variation be justified? The answer posed here is to build categories out of features that are in a certain generalized sense nonaccidental—that is, whose presence in even the single sample object would constitute an unexplained "suspicious coincidence" (Barlow, 1989) unless they in fact are typical of the entire class. When suitably formalized (as a particular node on the lattice of available interpretations) the least coincidental or "generic" interpretation turns out to be unique (Theorem 1).

2. REGULARITIES, LATTICES, AND MODELS

2.1. Regularities as Manifolds

In the dot-on-line example from Section 1, the entire set of possible line–dot relationships (in the plane) can be thought of as a 2D space—for example, expressing the position of the dot in a Cartesian coordinate frame centered on the line segment.[1] This space, the *configuration space*, expresses the possible relative configurations of the two objects in a completely general way; that is, it expresses this relationship between their positions prior to the imposition of any constraint or regularity on this relationship. By contrast, a constraint can naturally be thought of as a *subset* of the points in this space, that is, as a requirement that the point must fall in a particular part of this space, and is not free to fall "just anywhere." However, not all subsets make equally good sense as potentially meaningful categories.

One particularly reasonable type of subset is a *curve* passing through the configuration space. Such curves satisfy two very general constraints on the form of categories: *coherence* and *nonaccidentalness*.

[1] Note that this space can be parameterized in many different ways; it is critical in what follows that the definition of a regularity does *not* depend on this choice, although characteristics of a particular regularity are most readily expressed in a particular parameterization.

Loosely, coherence means that points in the subset all obey the some common, uniform constraint. Nonaccidentalness means that objects are unlikely to satisfy the constraint by accident, just as points in a space are unlikely to fall on any given curve. This idea was developed in the 3D shape literature (Lowe, 1987; Binford, 1981) and embedded in more general probabilistic (measure-theoretic) machinery by Bennett, Hoffman, and Prakash (1989). In the more general form of "nongenericity" (Jepson and Richards, 1991; Richards, Jepson, and Feldman, 1996), this notion provides a critical kind of inferential leverage. Because the feature is unlikely to be an accident, it must have some reason—i.e., it must occur by dint of some causal force. Hence the observer's best interpretation is to incorporate the nonaccidental feature right into the model. As part of the model, the feature is now expected—that is, generic. This idea is formalized below.

Assume a configuration space Φ, each point of which corresponds to one object. For concreteness, assume that Φ is parameterizable by R^d.[2] The number d is the dimension of the space, so that each point in Φ can be uniquely indexed by a set of d real parameters $\phi_1, \phi_2, ..., \phi_d$.

One way a set of points $\rho \subset \Phi$ can all obey a common, uniform constraint is if they all satisfy a single function f_ρ, expressed in implicit form as

$$f_\rho(\phi_1, \phi_2, ..., \phi_d) = 0. \tag{1}$$

It is conventional to restrict attention to *smooth* functions (functions that have continuous derivatives of arbitrary order). This restriction is quite general, in that it captures a wide range of constraints that actually turn out to exist in physically realizable domains[3] (Saund, 1987; Farouki and Hinds, 1985). Moreover, smoothness captures the basic idea of "uniformity." Corners, creases, and their higher-dimensional analogs—nonsmooth boundaries between smooth regions—signify changes in the underlying "generative process" (Zucker and Davis, 1988; Feldman, 1996).

When f_ρ is smooth, the resulting constraint ρ is called a *manifold*, hence the term "constraint manifold." It turns out that such a manifold passing through a d-space is itself parameterizable—not by R^d, but rather by R^{d-1}. Locally, that is, the constraint manifold is isomorphic to Euclidean $(d-1)$-space. Just as the configuration space has a certain well-behaved structure reflected in the fact that it can be

[2] In fact, the assumption here that the space is Euclidean is overly restrictive. But laying out the theory in the most general form would muddy the presentation unnecessarily. Note that here and throughout the paper, we will be concerned only with the *local* topology of the configuration space, not its global structure; that is, we will not be concerned with such global issues a space has the connectivity of a plane, a cylinder, or a torus, etc., but rather only with its structure in a small neighborhood.

[3] Moreover, this restriction allows us to state conditions sufficient to guarantee that ρ actually can be expressed implicitly at a given point, namely that the tangent map to ρ is a function.

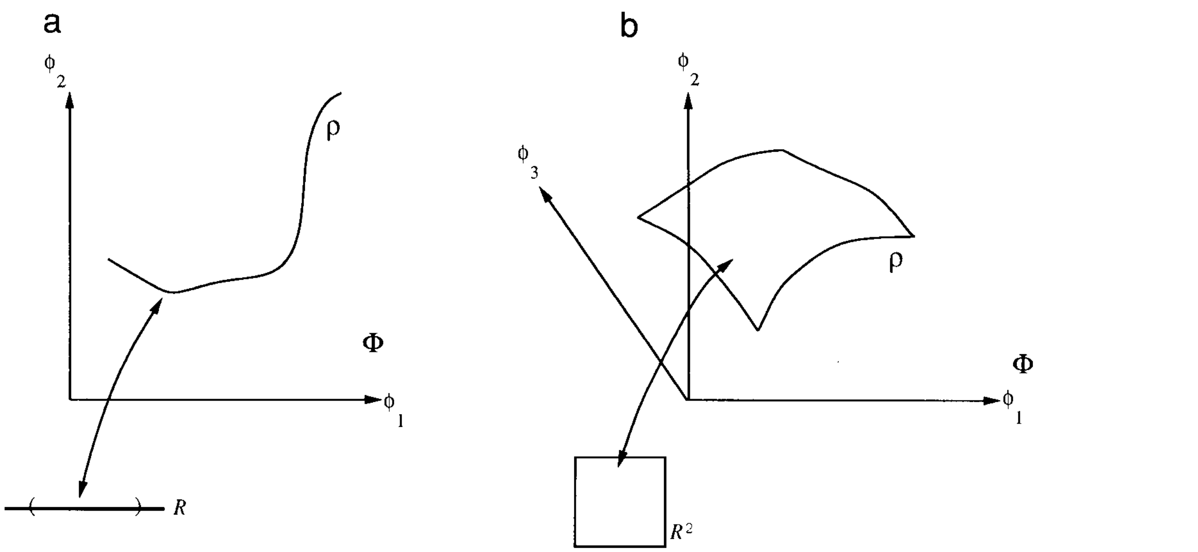

FIG. 2. Two examples of constraint manifolds, illustrating the fact that such a manifold can be parameterized by R^{d-1}, are shown. The arrows indicate differentiable mappings: (a) dim $\Phi = 2$, dim $\rho = 1$; (b) dim $\Phi = 3$, dim $\rho = 2$.

parameterized, the constraint manifold has an analogous structure, only of lower dimension and possibly curved.

Figure 2 shows two examples of constraint manifolds. Note how each manifold expresses a relationship that amounts to a kind of *covariation*: as one parameter changes, the others must change accordingly in order to remain on the manifold. This idea captures the essence of constraint manifolds seen as coherent categories—potentially independent parameters varying in concert with each other in order to satisfy a uniform constraint. Note moreover that the constraint that these points all satisfy is an extremely strong one, in that only a measure-zero subset of the points in Φ obey it (because it is of lower dimension than Φ). Loosely, points in Φ are very unlikely to satisfy the constraint ρ "by accident."

The above arguments suggest a simple definition of a regular constraint, called a *regularity*:

Definition 1 (Regularity). Given a configuration space Φ of dimension d, a regularity $\rho \subseteq \Phi$ is a manifold of dimension less than d.

Note that configuration spaces and constraint manifolds have analogous structures, only with different dimensions. In fact, the whole configuration space itself may be thought of as a constraint manifold—only with respect to some even higher-dimensional space. That is, any manifold is a regularity only with respect to some less constrained space. This observation leads naturally to a hierarchy of nested manifolds, each of which is a regularity with respect to the larger manifold in which it is embedded. Such a hierarchy will be developed explicitly below.

Because a manifold is a regularity in virtue of the difference between its dimension and that of the space in which it is embedded, it is convenient to refer explicitly to this difference, called the *codimension* (see Poston and Stewart, 1978, for a discussion of geometric aspects or Jepson and Richards, 1992, for an application to perceptual theory). The codimension represents the number of independent constraints or conditions that points in the space must satisfy in order to be on the manifold, thus amounting to a measure of the degree of specialization or regularity exhibited by each categorical hypothesis. Each unit of codimension corresponds to one "degree of freedom" having been removed from the overall configuration space, with the codimension 0 "regularity" being simply the configuration space itself. Hence a regularity is simply a submanifold of codimension greater than zero.[4]

2.2. Intersecting Regularities: The Regularity Lattice

Now consider the more general case in which multiple regularities—a fixed *regularity set* $\Pi = \{\rho_1, \rho_2, ..., \rho_N\}$—inhabit a single configuration space, possibly intersecting. Because the intersection (as always, locally) of two manifolds is another manifold, it is natural to consider each of these intersections as in effect another regularity, possibly with a higher codimension; and likewise the intersections of the intersections, and so forth. The complete set of regularities can be enumerated and diagrammed in the

[4] Nevertheless, it will sometimes be convenient to refer to the configuration space itself as the "improper regularity" that exhibits no actual regularity.

regularity lattice, a discrete graph structure in which larger, more inclusive regularities appear towards the top, and smaller, more constrained regularities appear towards the bottom (Jepson and Richards, 1992; Feldman, 1991; Richards, *et al.*, 1992; Feldman, 1992b).

To see how the regularity lattice is derived, consider the *closure under intersection* of the regularity set (plus, for the sake of completeness, the configuration space itself, thinking of it as the "improper regularity"). This closure set thus includes the configuration space Φ itself; each member of the regularity set; and each point set (subset of Φ) that can be produced by intersecting regularities. One of these will be the empty set $\varnothing$ whenever any two of the regularities are disjoint (do not intersect). This closure set is called the *regularity lattice* (the name is explained below). It is formally defined as follows:

DEFINITION 2 (Regularity Lattice). For a configuration space Φ with regularity set $\Pi = \{\rho_1, \rho_2, ...\}$, the *regularity lattice* $\mathscr{L}_{\Phi,\Pi}$ is the smallest set that

(a) contains Φ,

(b) contains each $\rho \in \Pi$, and

(c) is closed under intersection.

Loosely, the regularity lattice is simply the set obtained by taking the set including all the regularities plus the configuration space, and closing it under intersection. Consider the internal structure of this closure set. Some of its members are subsets of others, giving it a natural partial order defined by subset inclusion. More specifically, this partial order is a *lattice*. A lattice $\mathscr{L}$ is distinguished from other partial orders by the fact that, for each of its subsets, it contains both the greatest lower bound (called the "meet" and denoted $R \wedge S$ for $R, S \in \mathscr{L}$, or $\bigwedge \{R, S...\}$ for subset of $\mathscr{L}$) and the least upper bound (the "join", $R \vee S$ or $\bigvee \{R, S...\}$) under the partial order[5] (see Davey and Priestley, 1990). In the regularity lattice, the meet of a set of regularities is simply their intersection, and the join of a set of regularities is the smallest regularity in the lattice that contains them all.[6] The partial order can also be expressed directly as $R \leqslant S$ and $R \geqslant S$ to indicate $R \subseteq S$ and $R \supseteq S$ respectively. The join of the entire set $\mathscr{L}$ is simply the space Φ itself, which contains all the others; and the meet of the entire set is $\bigcap_i \rho_i$, which is a subset of all the others. This will be $\varnothing$ if they do not all intersect, i.e., if there is at least one pair of disjoint regularities (as in Fig. 3b below). In this case the regularities actually "over-determine" the points in the space, so that no point can satisfy them all. In other cases the greatest lower bound will be an isolated point at which the regularities all intersect.

Whenever it would be unambiguous, $\mathscr{L}_{\Phi,\Pi}$ will be shortened to $\mathscr{L}$, usually referring to the (po)set, but sometimes to the partial order itself when no confusion is created. The notation $[R, S]$ indicates the interval $\{X : R \leqslant X \leqslant S\}$ of the partial order. All elements $R, S, ... \in \mathscr{L}$ will be referred to as "regularities," reserving $\rho_1, \rho_2, ...$ for the originating regularities from which the lattice was constructed. The lattice containing only one object (whose partial order is empty) will be referred to as the trivial lattice.

EXAMPLES. Figure 3 shows sample configuration spaces with various configurations of regularities. Figure 3a shows the crossed case, in which two codim-1 regularities intersect. The resulting lattice (shown on the right in the figure) has four regularities: the space Φ itself, ρ_1, ρ_2, and their intersection $\rho_1 \wedge \rho_2$. The corresponding lattice, shown on the right, connects those regularities that differ by one unit of codimension. (The edges in the lattice diagram are shown dotted in the configuration space as well.) The codim-0 case at the top, the configuration space itself, is the *generic* case; that is, the "typical" case for points in the space. Lower cases are nongeneric in the space because they obey some regular constraint. At the bottom is the unique, completely regular case satisfying all regularities.

Figures 3b and 3c show slightly more complex examples. In (b), the regularities ρ_2 and ρ_3 do not intersect. Hence beneath the two codim-2 cases ($\rho_1 \wedge \rho_2$ and $\rho_2 \wedge \rho_3$) is the empty case $\varnothing$. In (c), the regularity ρ_2 is contained entirely within ρ_1, resulting in only three cases: Φ, ρ_1, and $\rho_2 = \rho_1 \wedge \rho_2$.

Categorization Hypothesis. The hypothesis is that objects in the space are categorized entirely in terms of what regularities they obey. That is, objects are regarded as categorically equivalent if they obey the same set of regularities, and categorically distinct if one obeys a regularity that the other one does not. Now, depending on the geometry of the regularities as embedded in the configuration space, only certain combinations of regularities actually occur; the regularity lattice lists them. And because (by hypothesis) each of these combinations corresponds precisely to one psychological category, the regularity lattice lists all possible categories for a given configuration space and regularity set, partially ordered by their degree of regularity. For each category on the lattice, each lower neighbor (i.e., each node that hangs off it) is a category that is exactly one "notch" more specialized—a special case in which one degree of freedom has collapsed.

2.3. Constraint on the Choice of Regularities: Modality

The above theory leaves the choice of regularity set completely unconstrained. However, some constraint on the

[5] Here and elsewhere the domain of the partial order is assumed to be finite.

[6] Note that this is *not* the union. The join of R and S is the smallest point set *on the lattice* that contains R and S, not the smallest point-set in Φ that contains them, which would be the union. Hence the lattice operation meet $\bigwedge$ and set intersection $\bigcap$ may be thought of as interchangeable, but the join $\bigvee$ and set union $\bigcup$ cannot be.

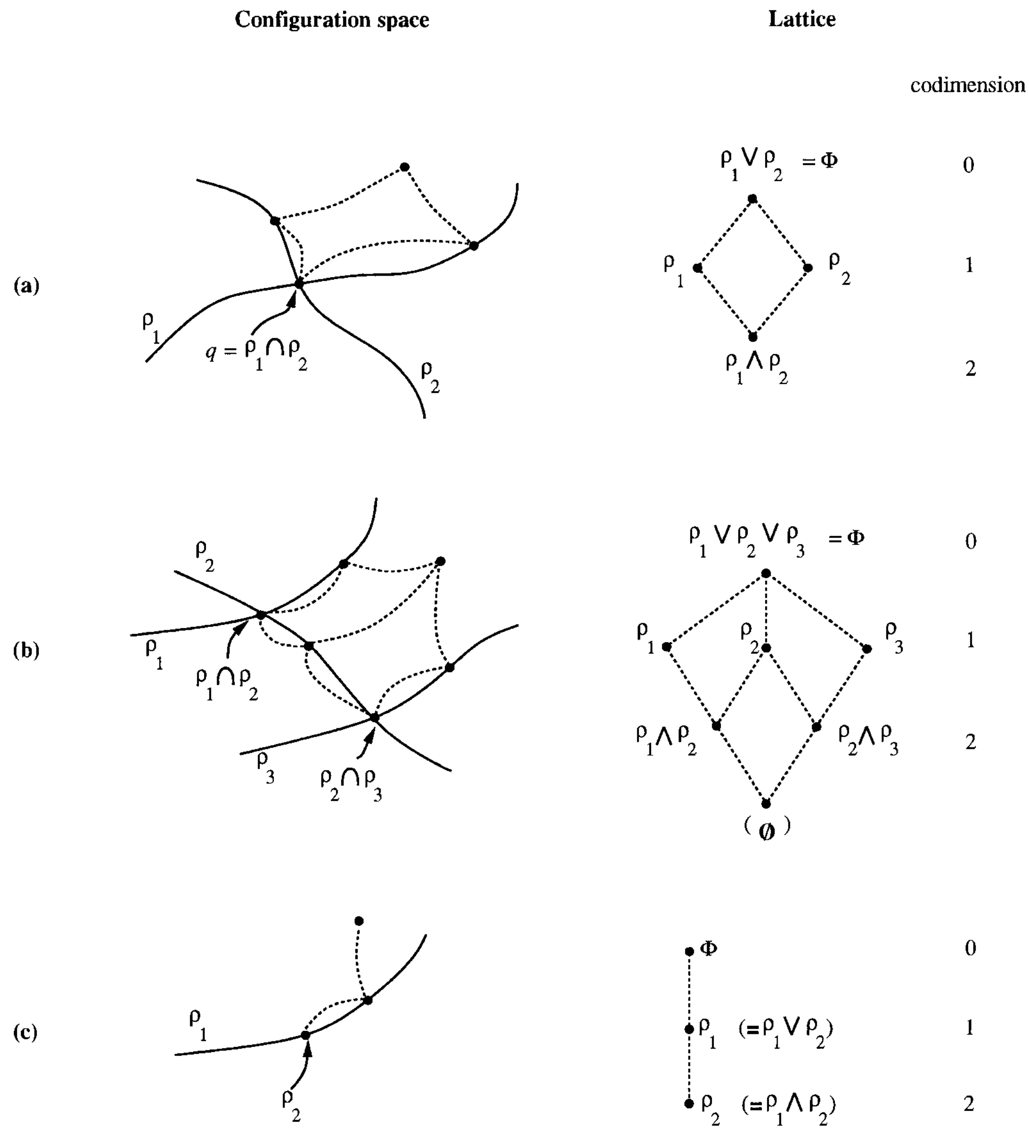

FIG. 3. Several configuration spaces and regularity sets (left) and the corresponding regularity lattices (right). In each configuration space, dotted lines are shown corresponding to the edges in the lattice partial order.

precise form that regularities can take is necessary. Without constraint, for example, *any* set of points can be regarded as obeying a common regularity (simply by passing a curve through them all). Hence, without constraint any pair of objects—no matter how similar or dissimilar—are members of the same number of common categories (a variant of the "Ugly Duckling Theorem" of Watanabe, 1969; see also Watanabe, 1985).

2.3.1. *Generative Models of Objects*

One proposal for how regularities can be constrained comes from a consideration of what the parameters $\phi_1, \phi_2, \ldots$ of Φ actually measure. In principle, these numbers need not measure anything physically meaningful. The following proposal requires that regularities should be such that the parameter changing along them does in fact correspond to something concrete, the degree to which some *transformational operation* was carried out during the creation of the object.

Leyton (1984) (see also Leyton, 1992) has pointed out that shapes can be modeled using a fixed algebraic sequence of group-theoretic transformations and has also proposed that human subjects use such a scheme to represent shape, recursively modeling objects as deformations of successively simpler objects (Fig. 4). In this scheme, the observer's representation of shape acts as a model of the *causal history* of the object—an interpretation of how the shape appears to

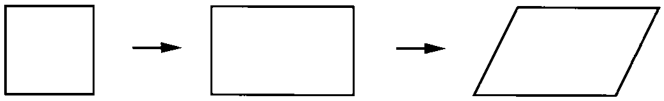

FIG. 4. In Leyton's (1984) example, a parallelogram is seen as a deformation of a rectangle, which in turn is seen as a deformation of a square.

have come into existence. Note that the conceptualization (as in the figure) of a parallelogram as a slanted, stretched square—as opposed to simply representing its shape "in and of itself"—though intuitively compelling, is by no means logically necessary. Rather, this conceptualization amounts to a kind of *theory* about the causal origins of the rectangle (see also Leyton, 1989, on this point).

The scheme is quite general. In addition to such simple mathematical operations as stretching and shearing, object shapes can be varied by using simple translations, rotations, and scale changes (operations that are normally treated as invariant when they act on an entire object) acting on just one part of an object with respect to the main body. Thus one part can be turned or lengthened with respect to the main axis—corresponding to, say, bending an arm, lengthening a neck, and so forth. Tapering, bending, and twisting may also be modeled by simple matrix multiplication (Barr, 1984), and hence make for mathematically convenient generative models. A number of simple mathematical operations have been shown to closely model natural growth of various kinds, which is convenient for the identification of these mathematical operations with natural processes whereby objects' shapes are determined. D'Arcy Thompson (1942) has noted that the shape change induced by gross size changes between structurally similar animal species often takes the surprisingly simple form of an affine transformation. More recently, Mark, Shapiro, and Shaw (1986) have noted that the mathematical operation called cardiodal strain, when carried out upon head-like pictures, is perceived as cranial growth.

Concretely, and fairly generally, these operations can be thought of as the action of some group $\mathbf{g}$ of transformations. Denote the elements of this group by $\{g^1, g^2, ...\}$, using the superscripts as indices. Write xg^α for the object produced when object x is transformed by operation g^α. The group definition requires that $\mathbf{g}$ contain a unique null operation e (such that $xe = x$ for all x), and that each operation g^α has a unique inverse $-g^\alpha$ such that $x(g^\alpha)(-g^\alpha) = xe = x$. Note that the null operation e is the same across different groups: if e_i is the null operation for the group $\mathbf{g}_i$ and e_j for the group $\mathbf{g}_j$, then $xe_i = x = xe_j$, hence $e_i = e_j$. Of particular interest are the so-called "continuous groups" whose members can be parameterized by some real number. The group of rigid motions (including arbitrary rotations and translations) is of this type; so are operations that can be thought of as rigid motion of just a part of an object.

Imagine now that the observer has, as tools with which to construct interpretations of objects which it observes: (1) some fixed inventory $\mathscr{G} = \{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_d\}$ of groups of operations and (2) a unique[7] primitive object, denoted $\mathbf{0}$. The set of objects that can be produced by a group $\mathbf{g}$ from an object x, that is, the set

$$\{xg^\alpha : g^\alpha \in \mathbf{g}\}, \tag{2}$$

is called the *orbit* of x under the group $\mathbf{g}$, denoted $x\mathbf{g}$. Similarly, $x\mathbf{g}_1\mathbf{g}_2$ is the orbit of x under the two operations $\mathbf{g}_1$ and $\mathbf{g}_2$. Each orbit represents an entire class of objects, by abstracting over the precise operations used to construct the objects (that is, by dropping the indices). In this way, any subset $G \subseteq \mathscr{G}$, coupled with a primitive object $\mathbf{0}$, represents a class of objects, namely the orbit of $\mathbf{0}$ under the operations in G.

Generative operations and orbits have a natural correspondence with configuration spaces and regularities. The orbit of $\mathbf{0}$ under the full set $\mathscr{G}$ of generative operations is a configuration space, namely the full set of objects that can be produced by these operations. Assuming that the operations are all transverse (linearly independent), this space is in fact a vector space of dimension d (the number of operations), with a basis provided by $\mathscr{G}$. Likewise, the orbit of $\mathbf{0}$ under each *proper* subset of $\mathscr{G}$ is a regularity—that is, a positive-codimension subspace of the full configuration space $\mathbf{0}\mathscr{G}$. Such a regularity is called a *modal* regularity:

DEFINITION 3 (Modal Regularity). A regularity $\rho \subset \Phi$ is called *modal* if it is coextensive with the orbit $\mathbf{0}G$ of some primitive object $\mathbf{0}$ under some set $G \subsetneq \mathscr{G}$ of generative operations.

Not all regularities are modal. While any curve through the space, no matter how idiosyncratic or "bent," counts as a regularity, only those that exactly correspond to the action of a constructive operation are modal. Modality thus places a very heavy constraint on the choice of regularities, limiting the choice to those for which the observer has an explicit constructive operation. The examples given below give a more concrete idea, though less than a comprehensive theory, of what generative operations make psychologically reasonable choices for constructing shape categories. In general the idea is that operations defined with respect to the object at hand are allowed: e.g., rotations of an articulated part, but not of an arbitrary image subset; translation of a part in a direction defined by an image axis, but not in an arbitrary direction; and so forth.

2.4. The Lattice of Category Models

To complete the correspondence with regularities and the regularity lattice, it is convenient to relax the uniquenes of the primitive object. Instead, assume multiple primitive

[7] Later multiple primitive objects will be allowed.

objects $\mathbf{0}_1, \mathbf{0}_2, \ldots$. The resulting configuration space no longer corresponds to any single vector space, but this causes no problems.

Now it is possible to construct a lattice of modal regularities exactly isomorphic to the regularity lattice defined above. Each node on such a lattice, called a *category model* or *model*, is the orbit of one of the primitive objects under some subset of the generative operations, and hence corresponds to a category of objects, all of which can be constructed from the same set of operations. The meet $\bigwedge$ is again simply intersection. The join $\bigvee$ is linear span; that is

$$\mathbf{0}G \vee \mathbf{0}H = \mathbf{0}(G \cup H). \tag{3}$$

As with all regularity lattices, movement up the lattice yields larger, more inclusive object categories, while movement down the lattices produces more constrained categories, generated by smaller sets of operations. In a sense, lattices defined by regularities are defined "from the top down" (i.e., by taking meets of codim-1 cases), while those defined by generative operations are defined "from the bottom up" (by taking joins of dim-1 cases). The resulting structures are exactly isomorphic. The notion of dimension carries over as well. Note that a category model's dimension can be regarded in two ways: geometrically, as its dimension when viewed as a point-set in Φ; or simply as the number of one-parameter operations used to define it. The two senses agree. Likewise, the codimension is defined either in the usual geometric way, or as the difference between the number of operations used to define the model and the number used to define the entire space.

Figures 5–7 give some concrete examples of category lattices, explicitly showing the primitive object(s) and generative operations used. Note the exact isomorphism with the three regularity lattices exhibited above. The examples suggest what operations make reasonable choices of generative operations. Considerably more complex

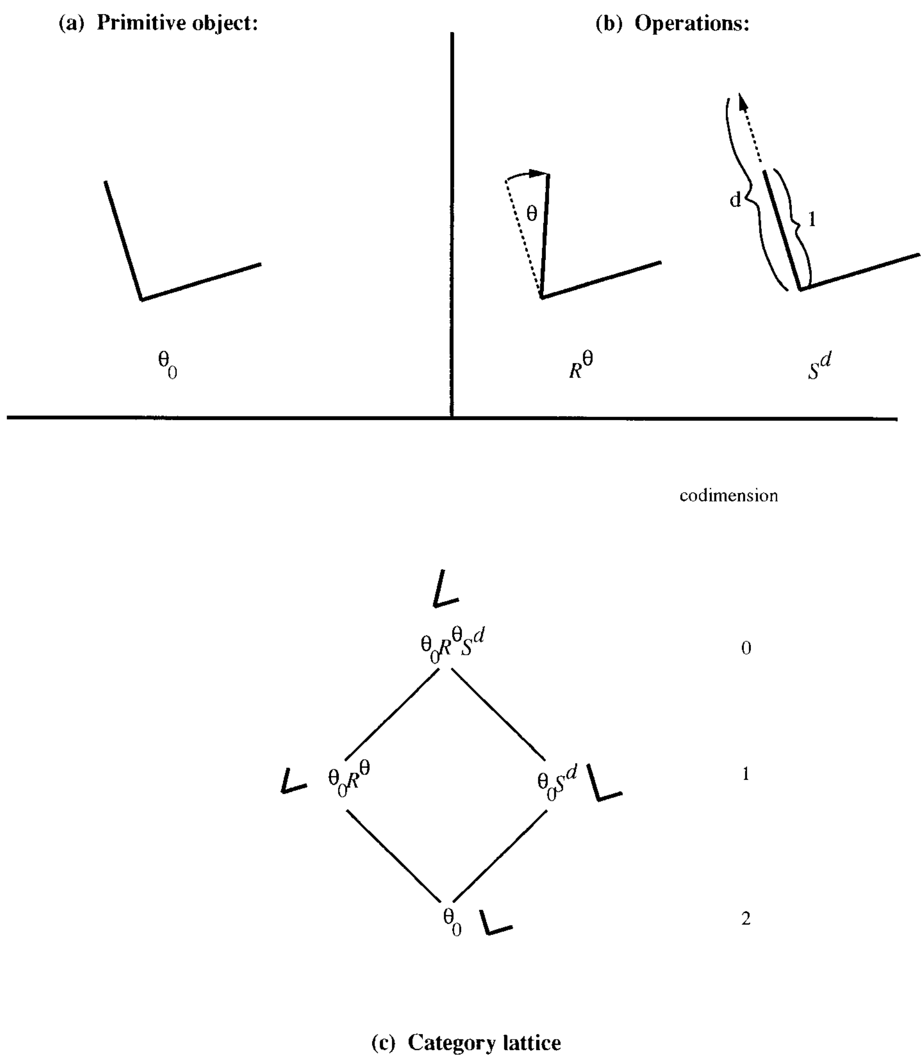

FIG. 5. (a) A primitive object and (b) generative operations that can produce Vs and (c) the resulting category lattice.

examples can be produced with a computer program (see Feldman, 1992a, 1992b, for details).

The example in Fig. 5 shows a space of "V's" (where a V is two line segments in the plane that share an endpoint). Here the (sole) primitive object, θ_0, is the right isosceles V; and the two operations are R^θ (bend the V's joint by an angle θ) and S^d (stretch one leg by a ratio d). Any V can be expressed as $\theta_0 R^\theta S^d$ for some θ, d; that is, θ and d parameterize the space of V's. There are thus four models, θ_0, $\theta_0 R$, $\theta_0 S$, and $\theta_0 RS$, which form a lattice (Fig. 5b). The four models correspond to vector spaces of dimension 0, 1, 1, and 2 respectively. Again, the lattice partially orders the vector spaces by their dimension, connecting each model to other models that differ by exactly one collapsed (or expanded) degree of freedom. The lattice is isomorphic to that in Fig. 3a.

Similarly, Fig. 6 shows the V's space with an additional primitive V, the "straight" isosceles V. The resulting partial order (Fig. 6) has six models: one of dimension 2, three of dimension 1, and the two 0-dimension primitive objects. This lattice is isomorphic to that in Fig. 3b.

Finally, Fig. 7 shows an example reminiscent of the dot-on-line example from Fig. 1. Each object consists of a line segment and a dot. The sole primitive object, r_0, is a segment with a dot at one endpoint. One operation, t_{segment}, is translation along the segment. Note that this is well-defined only when the dot is actually touching the segment, and for this reason the resulting lattice is less than the full crossing of the two operations. The other operation, $t_{r,\theta}$, is arbitrary translation in the plane. This lattice is isomorphic to that in Fig. 3c.

Categorization Hypothesis, Modal Version. The category lattice is simply the modal or "constructive" version of a regularity lattice. Extending the "categorization hypothesis" to the modal case, the thesis is now that each mode on the lattice—each set of objects that bear exactly the same generative history—is a psychological category. This idea neatly parallels Leyton's prescription that each object be described by the generative operations that brought it about; here, each object *category* is a set of objects all of which can be produced by the exact same set of operations.

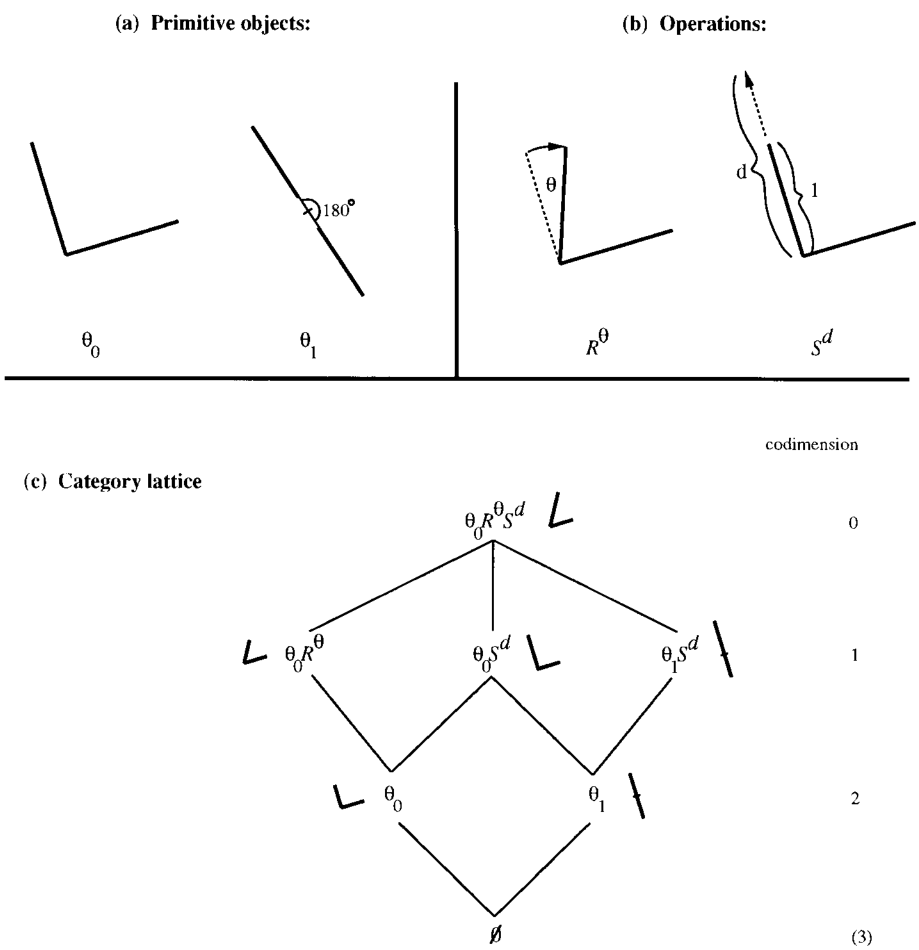

FIG. 6. (a) A larger set of primitive objects, coupled with (b) the same operations as in Fig. 5, and (c) the resulting category lattice.

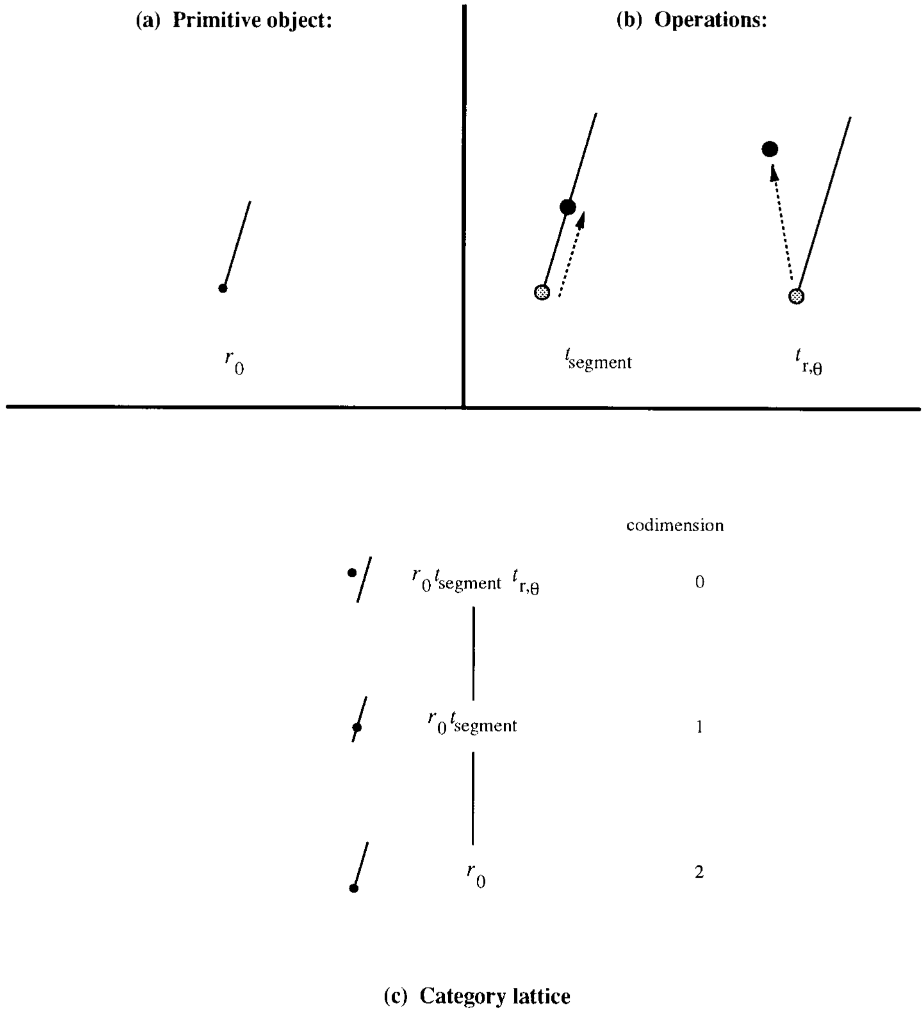

FIG. 7. (a) A primitive object consisting of a line segment with a dot at the endpoint, coupled with (b) two generative operations and (c) the resulting category lattice.

2.5. Category Inference Using Lattices

The question of how the observer computes with the lattice machinery can be separated into two closely related questions: (a) given a lattice, what category should an observed object be assigned to, and (b) given an object, what lattice should be constructed? Question (a) is considered first.

2.5.1. *Inference to a Category*

From the point of view of an observer confronted with an object x, the crucial question is: To what category does x belong? In many cases, x will belong to several different categories on the lattice. In fact, if x is contained in one category R, then it is also contained in all of R's upper ancestors $S \geqslant R$ (because they are supersets of R). Hence additional constraint is required in order to render category assignment unique.

The Genericity Constraint. The organization of the lattice suggests a natural constraint, called *genericity.* Consider the point $q = \rho_1 \cap \rho_2$ in Fig. 3a. q is contained in all four categories in the lattice: Φ, ρ_1, ρ_2, and $\rho_1 \wedge \rho_2$. Nevertheless, q is a very *atypical* member of the first three categories. For example, while q is in Φ, it is unlike all other points in Φ, in that it is *also* in ρ_1 and ρ_2. Similarly, while q is in ρ_1, it is unlike all other members of ρ_1 in that it is also in ρ_2.

In giving q a categorical interpretation, it makes little sense to choose a category containing mostly points that are unlike q. Rather, points should be associated only with categories in which they are typical or "generic." An object x is generic in a category R if x is in R but not in any more constrained categories on the lattice. Formally,

DEFINITION 4 (Genericity). An object x is *generic* in a category $R \in \mathscr{L}_{\Phi,\Pi}$ iff $x \in R$ but $x \notin S$ for all $S < R$.

Given an object x and a lattice $\mathscr{L}$, there is one and only one category $R \in \mathscr{L}$ in which x is generic. This is referred to as the generic interpretation of x.

THEOREM 1 (Existence and Uniqueness of the Generic Interpretation). *Given $\mathscr{L}_{\Phi,\Pi}$, for any $x \in \Phi$,*

(i) *there exists an $R \in \mathscr{L}_{\Phi,\Pi}$ such that x is generic in R, and*

(ii) *R is unique.*

A proof can be found in the Appendix. Informally, the existence of a generic interpretation is simply a result of the fact that all objects are in Φ, while its uniqueness is a result of the uniqueness of the meet $\bigwedge$ guaranteed in any lattice. The unique generic interpretation of x given a lattice $\mathscr{L}_{\Phi,\Pi}$ will be denoted $\text{GEN}_{\Phi,\Pi}(x)$ or simply $\text{GEN}(x)$.

The uniqueness of $\text{GEN}(x)$ means that (given a regularity set) one example suffices to choose a category. While the object may be an example of many categories, it is a *generic* example of only one category. In general, this generic interpretation is the only interpretation devoid of "suspicious coincidences", i.e., the only completely nonaccidental categorization. In the modal case, where each interpretation is a constructive model, the generic interpretation is the only model in which every operation was carried out to a nonzero magnitude, i.e., the only model in which every operation invoked has a valid justification. Hence the generic interpretation is the unique generative model of which the observed object is a typical outcome, and hence is the "natural" generalization of the object.

The Partition Due to Genericity. It is worth remarking that Theorem 1 means that the generic interpretation defines a *partition* of Φ into disjoint regions. This partition is in fact schematically described by the lattice itself: each node R on the lattice corresponds to just those points that have R as a generic interpretation (i.e., all the points that satisfy R minus the nongeneric ones). The partition by genericity can be thought of as picking out the "seams" inherent to this miniature universe of objects—the natural divisions between qualitatively distinct regions of the space. Modulo the choice of regularities, it makes sense to think of this partition as picking out the "qualitative kinds" in the space.

Inference Rules. The organization of the lattice provides an extremely direct way of computing the generic interpretation $\text{GEN}(x)$. An object x is generic in an interpretation R if x is in R but x is not in any more constrained categories—i.e., is not in any categories below R on the lattice. Hence the Genericity Constraint is equivalent to the following minimum rule:

(MINIMUM RULE). *Given an object x, among all R's in $\mathscr{L}$ s.t. $x \in R$, choose the R which is the minimum in the lattice partial order.*

Pictorially, this minimum is simply the lowest node in the lattice diagram to which x belongs. It can be expressed several ways. Denote by R_x is the set of nodes in $\mathscr{L}$ that contain x, i.e.,

$$R_x = \{R \in \mathscr{L} \mid x \in R\}. \tag{4}$$

Then the minimum rule says that the generic interpretation of x is just

$$\text{GEN}(x) = \min_{\mathscr{L}} R_x, \tag{5}$$

where $\min_{\mathscr{L}}$ denotes the minimum over the lattice partial order. Equivalently,

$$\text{GEN}(x) = \bigwedge R_x. \tag{6}$$

Also equivalently, the generic interpretation is just the member of R_x with the largest codimension, yielding the alternative rule:

(MAXIMUM CODIMENSION RULE). *Given an object x, among all R's in $\mathscr{L}$ s.t. $x \in R$, choose the R with the largest codimension.*

Whichever computation rule is used, the Genericity Constraint is recognizably a variant of the "minimum principle," many versions of which are familiar from the perceptual literature (Hochberg and McAlister, 1953; Leeuwenberg, 1971; Simon, 1972; Barrow and Tenenbaum, 1981; Kanade, 1981; cf Rissanen, 1978). This rule is usually thought to descend from the principle of simplicity: draw the simplest interpretation possible. While intuitively appealing, however, this principle notoriously lacks an independent justification or even a normative definition (Quine, 1965; Sober, 1975). In particular, the apparent presumption that the world is simple has troubled many authors (Hatfield and Epstein, 1985; Leeuwenberg and Boselie, 1988). The minimum rule here, by contrast, derives only from the Genericity Constraint. There is no presumption whatsoever that the world itself tends to minimize complexity—only an assumption that certain special configurations are unlikely to happen by accident. An additional advantage is that the minimization here is not global (i.e., over all possible interpretations, with attendant difficulties) but rather is restricted to a fixed finite set of interpretations, namely those on the lattice.

2.5.2. *Inference to a Lattice*

Given an object x, what lattice should be constructed? In principle, the observer might simply use all known primitive objects and all known generative operations to produce the "full" lattice (containing all known categories) at all times. In practice, of course, this would be enormous, and in fact only a small portion of the full lattice is required in any given context. This mandatory portion of the lattice is of special interest when looking for empirical validation of the lattice theory, because its presence "in subjects' heads" provides direct evidence for the theory.

The idea is that the observer begins with a primitive object, and then builds the lattice up from the bottom, but only goes as high as necessary and no higher. The procedure stops when the top node is powerful enough to express the observed object, which it will do generically (because all lower nodes do not contain the object). Viewed as a function $\mathscr{L}(x)$ mapping objects to lattices, the desired output is simply the (inclusive) *interval* of the lattice partial order between some primitive object and the lowest category capable of generating x.

DEFINITION 5 (Category Lattice for an Observed Object). Given an object x and lattice partial order $\leqslant$ deriving from $\mathscr{L}_{\Phi,\Pi}$, the corresponding lattice $\mathscr{L}(x)$ is the interval $[\mathbf{0}, x]$ for some primitive object $\mathbf{0}$, i.e., the set

$$\mathscr{L}(x) = \{ R \mid \mathbf{0} \leqslant R \leqslant \text{GEN}(x) \}.$$

Hence $\mathscr{L}(x)$ is a subset of $\mathscr{L}_{\Phi,\Pi}$, and inherits a partial order from it.

Remark. Note that the size and complexity of $\mathscr{L}(x)$ increases monotonically[8] with the *dimension* of x. As x's dimension increases, the lattice gains nodes, but never loses any. In Fig. 3a, for example, the lattice $\mathscr{L}(q)$ of the point $q = \rho_1 \cap \rho_2$ is simply the trivial lattice containing the single node $\rho_1 \wedge \rho_2$, in which q is generic. A generic point on ρ_1, in contrast, would require the lattice

$$\begin{array}{c} \rho_1 \\ \Big| \\ \mathbf{0} \end{array}, \tag{7}$$

in whose top node it would be generic. Similarly, a generic point in Φ would require the full lattice with four nodes given in Fig. 3a.

The increase in lattice complexity with dimension is crucial, because it leads directly to an empirical prediction: when subjects are asked to categorize single objects, higher-dimensional objects lead to larger lattices. Remember that higher-dimensional objects are not *intrinsically* more complex than lower-dimensional ones—they are only more complex given a particular generative model. Hence the prediction is quite central; the core of the theory is the lattice structure itself, and this manipulation potentially allows the experimenter to uncover it, as it were, node by node, simply by varying the dimension of the sample object.

[8] Actually, is monotone non-decreasing.

2.6. Empirical Validation of the Theory

To confirm this prediction, it remains only to develop a technique for directly exhibiting the lattice employed by a subject, so that its size and complexity can be examined. This section proposes that, under some very reasonable assumptions, the lattice's structure is faithfully reflected in the probability density of objects generated by subjects as part of the "same" category as the sample object.

The general procedure employed in the experiments is to show subjects a small set of sample objects (1 or 3), accompanied by instructions intended to elicit a categorical interpretation (e.g., "this is an example of a *blicket*"). The subject is then asked to draw six more examples of the category. The set of objects that subjects produce is then coded into some parameter space, and frequencies are tallied at each point in the space. The result is a probability density of objects corresponding in some way to the subjects' "mental space of blickets."

A flat distribution in this density function would mean that the subjects generalize the sample object(s) without bias. More plausibly, one would expect a spike or mode in the neighborhood of the sample object(s); this would suggest that subjects generalize the sample by "blurring" or adding an error term to the sample along each dimension. This latter outcome is the almost universal presumption of researchers in similarity and categorization, corresponding to the seemingly obvious premise that natural categories correspond to topological neighborhoods in the appropriate parameter space—i.e., that nearby points are more likely to be in the same category than are distant points.

Nevertheless, the experiments reported below show a slightly different pattern of results, in two respects. First, as discussed in the Introduction, subjects generalize over *certain* dimensions, but not others. The distinction seems to be based on the construction of explicit generative models and is governed by the Genericity Constraint. Second, subjects' probability densities characteristically show *additional modes*, above and beyond those around the sample object.

The appearance of these additional modes might seem a bit of an enigma. But they make sense when interpreted in the following way; each node in the lattice gives rise to a distinct component of the distribution of objects in the probability density. This makes sense if the density is regarded as the subject's prior probability distribution for objects, so that each distinct category—by hypothesis, each node on the lattice—contributes one distinct class-conditional density function (Ashby and Perrin, 1988; Ashby, 1992). Hence the entire lattice is reflected in the probability density.

The appearance of nongeneric nodes in the probability densities is not predictable a priori (and in a sense is bizarre, as discussed below). Nevertheless, it appears to be quite reliable, and can be taken advantage of as a technique to examine subjects' lattices almost directly. The experiments below use this phenomenon to test the central prediction that higher-dimensional objects have larger lattices. The extra modes appearing in subjects' probability densities only with higher-dimensional objects are direct evidence in favor of the lattice theory.

For simplicity, the experiments focus on the simplest possible comparison: 1D objects vs 0D objects. Each of the configuration spaces employed can be parameterized by a single dimension. The crucial comparison then is between the probability densities resulting from *generic* and *nongeneric* sample objects. Generic objects should have lattices with at least two nodes, including both a generic node and a primitive object, and hence at least two modes in the probability density. Nongeneric objects should have trivial lattices, and hence only a single spike, at the object. This extremely qualitative empirical difference is predicted by the lattice theory, but would be mysterious otherwise, and hence can be taken as an empirical hallmark of the theory.

Note that this prediction is completely independent of the particular choice of regularities (i.e., of primitive objects and generative operations); the main contrast between multimodality in the generic case and unimodality in the nongeneric case remains stable. This is particularly important because there is no absolute a priori way of determining what generative models subjects will employ (i.e., what configurations they will regard as nongeneric).

Nevertheless, the distinction between generic and nongeneric configurations is far from arbitrary. As developed above, the main idea is that nongeneric configurations exhibit some nonaccidental property, meaning a *coincidence* (or "fuzzy identity"; see Witkin and Tenenbaum, 1983) of salient image structures—e.g., the coincidence between the dot and the endpoint of the line. This idea can be generalized somewhat when one considers the symmetry structure of the objects involved. For example, a line segment is mirror symmetric about its midpoint. This makes the midpoint a kind of "generalized endpoint," i.e., an endpoint after the symmetry group has been factored out. This makes the midpoint another plausible nongeneric configuration (see below). Similarly, consider a rhombus, which is mirror symmetric about two axes and rotationally symmetric up to 180° turns. But when the skew angle is 90° and the shape becomes a square, the symmetry group of the shape abruptly changes: it is now mirror symmetric about *four* different axes and rotationally symmetric up to 90° turns. Hence the 90° angle makes a plausible nongeneric configuration (see below). Note that this change of symmetry group entails a corresponding change in the dimensionality of the group's eigenspace (Leyton, 1984, 1992; see Leyton's first and second "interaction principles"). Hence this notion of genericity aligns nicely with the structure of the lattice and the notion of codimension.

There is one slight complication to the simple correspondence between nodes on the lattice and components of the frequency distribution. Certain categories have both positive and negative versions. Consider for example a one-parameter transformation of a 0-dim primitive object $\mathbf{0}$. Depending on the symmetries of the space, the generic object produced by applying the transformation in one direction away from the $\mathbf{0}$ may be psychologically distinct from that produced by applying the opposite transformation. In other cases, the objects produced in opposite directions are identical, meaning that there is really only one direction. Components of the probability density resulting from both positive and negative cases would be expected, rather than just the one deriving from the single generic node. Note that, for each node $R \in \mathscr{L}$, *each* lower neighbor of R can serve as the dividing point between positive and negative versions. To keep the notation simple, denote by $\mathscr{L}'$ the "augmented" set of nodes constructed by adding to $\mathscr{L}$ the negative versions of nodes when they exist.[9] Because $|\mathscr{L}'| \geqslant |\mathscr{L}|$, this complication does not disturb the main prediction.

2.7. Determining the Number of Modes

By the above argument, the crucial empirical test of the lattice theory depends on determining the number of component distributions in the probability density. Statistically, this turns out to be a difficult problem, and several assumptions must be imposed in order to make it possible.

By hypothesis, each node $R \in \mathscr{L}'$ on the lattice gives rise to separate probability density function denoted $p(x \mid R)$, which itself occurs with probability $p(R)$. Then the resulting density function is simply the mixture of these p.d.f.'s,

$$p(x) = \sum_{R \in \mathscr{L}'} p(R)\, p(x \mid R), \tag{8}$$

i.e., the sum of the p.d.f.'s weighted by their mixing proportions. The problem then is to take an empirical p.d.f. and determine the number of components that gave rise to it.

Determining the number of components in an arbitrary distribution turns out to be problematic (McLachlan and Basford, 1988). Consider a fixed frequency distribution which can be modeled either as single component or as the sum of two distinct components. Because the latter model is a strict generalization of the former (i.e., its parameters are a strict superset), one would normally expect to be able to use a likelihood ratio test (e.g., see Hoel, Port, and Stone, 1971) to test the significance of the added explanatory power of the additional component, and hence to discover

[9] Of course this changes the partial order as well, so $\mathscr{L}'$ will be treated only as a set.

the true number of components. However, this application of the test entails a breakdown in what statisticians refer to as the "regularity conditions" sanctioning the test. Specifically, the parameters of the one-component model (whether thought of as the mixing proportion of the additional component vanishing, or as its parameters becoming identical to the first component's) lie at the boundary of the parameter space of the two-component model, rather than in its interior (Gosh and Sen, 1985). Attempts to compensate for this problem analytically (Aitkin and Rubin, 1985) were later questioned (Quinn, McLachlan, and Hjort, 1987), and as a result the test cannot be interpreted strictly (McLachlan and Basford, 1988). For conservativeness, the latter authors suggest an approximate correction for the likelihood ratio test, which is explained and employed below. Additional safeguards must be imposed, however, because in many cases unimodal distributions are actually better fit by mixtures of multiple distributions, leading to the spurious detection of additional components.

To preserve the soundness of the likelihood test, and hence to be able to determine the number of components reliably, the form of the component distributions must be restricted. The following two simple assumptions suffice (again see McLachlan and Basford, 1988, for discussion): (1) *Each p.d.f. must be unimodal.* In modeling the data below, normal (Gaussian) components are assumed, with free mixing proportions and distinct variances (the so-called heteroscedastic case). Furthermore, (2) *the Mahalanobis distance between the means of adjacent components must be larger than two.* That is, the distance separating any two components must be more than twice as large as the (smaller) standard deviation.

Another approach to finding the number of components is to directly count the number of modes (distinct local maxima) in the frequency histogram. Obviously, if the component distributions are not unimodal, this will not accurately reflect the number of underlying component distributions. More subtly, even with Gaussian components, the number of modes may not reliably reflect the number of components. Needless to say, multiple modes may arise from single components simply due to sampling error. Conversely, under certain circumstances multiple components can yield unimodal histograms, no matter how finely sampled and regardless of the mixing proportions (see Everitt and Hand, 1981, for an extensive discussion). Unimodality will occur if the difference between the means is less than twice the (lesser of the two) standard deviations. This is ruled out by the assumptions already cited above. The two approaches thus agree on the required restrictions. As a result the terms "component" (separate distribution in the mixture) and "mode" (separate local maximum in the histogram) will henceforth be used interchangeably.

Assumption 2 places a very substantive restriction on the form that category models may take, and thus helps to ensure the falsifiability of the theory. The distance separating a generic category from a nongeneric one refers to the magnitude of a generative operation that must be applied to a primitive object in order to create a prototypical object of higher dimension. Requiring that this distance be above a certain threshold (here, two Mahalanobis units) means that multiple-component distributions cannot mimic single-mode p.d.f.'s, lending decisiveness to the empirical data. Conceptually, this required separation means that a *generic* example cannot be simply an arbitrary or random example; rather, a generic example must be far enough from the nearest nongeneric point, in this Mahalanobis sense, to avoid being conflated with it. Generic examples strictly avoid the appearance of regularity, with Assumption 2 specifying how much distance actually suffices to accomplish this. Independent theoretical arguments detailed in Feldman (1993) suggest an ideal value of about 2.255 Mahalanobis units, agreeing neatly with the statistically derived requirement.

3. EXPERIMENTS

In each of the following seven experiments, subjects were asked to induce a category from a sample object or objects, and then produce six novel examples, using paper and pencil. There were either one (Experiments 1, 2, 4, 5, and 7) or three (Experiments 3 and 6) example objects. The objects used were always either a line segment with a dot on it (henceforth, line–dots for short, and called "blickets" in the experiments) or two segments sharing an endpoint (called V's for short, and "zorks" in the experiments). Experiment 1 used generic line–dots, i.e., ones with the dot in a typical or nonspecial position, while Experiment 2 used nongeneric line–dots (dot at midpoint), the critical comparison. Likewise, Experiment 4 used generic V's, while Experiment 5 used nongeneric V's (right angles). Experiments 3 and 6 used three generic line–dots and V's, respectively. Finally, Experiment 7 used generic line–dots as in Experiment 1, but with the position of the dot varying along the line to create a flatter pooled distribution.

Figure 8 shows typical sample objects for each experiment. Also shown are $\mathscr{L}(x)$ for each sample object, as well as $\mathscr{L}'(x)$ in schematic form, to illustrate the main prediction of increased complexity in generic vs nongeneric sample objects.

General Method

Sample objects appeared on the first page of a booklet, accompanied by the words (e.g.) "Here is an example of a *blicket*" (one sample object) or "Here are some *blickets*" (three sample objects). On the second page appeared the instructions "Please draw 6 different examples of a blicket"

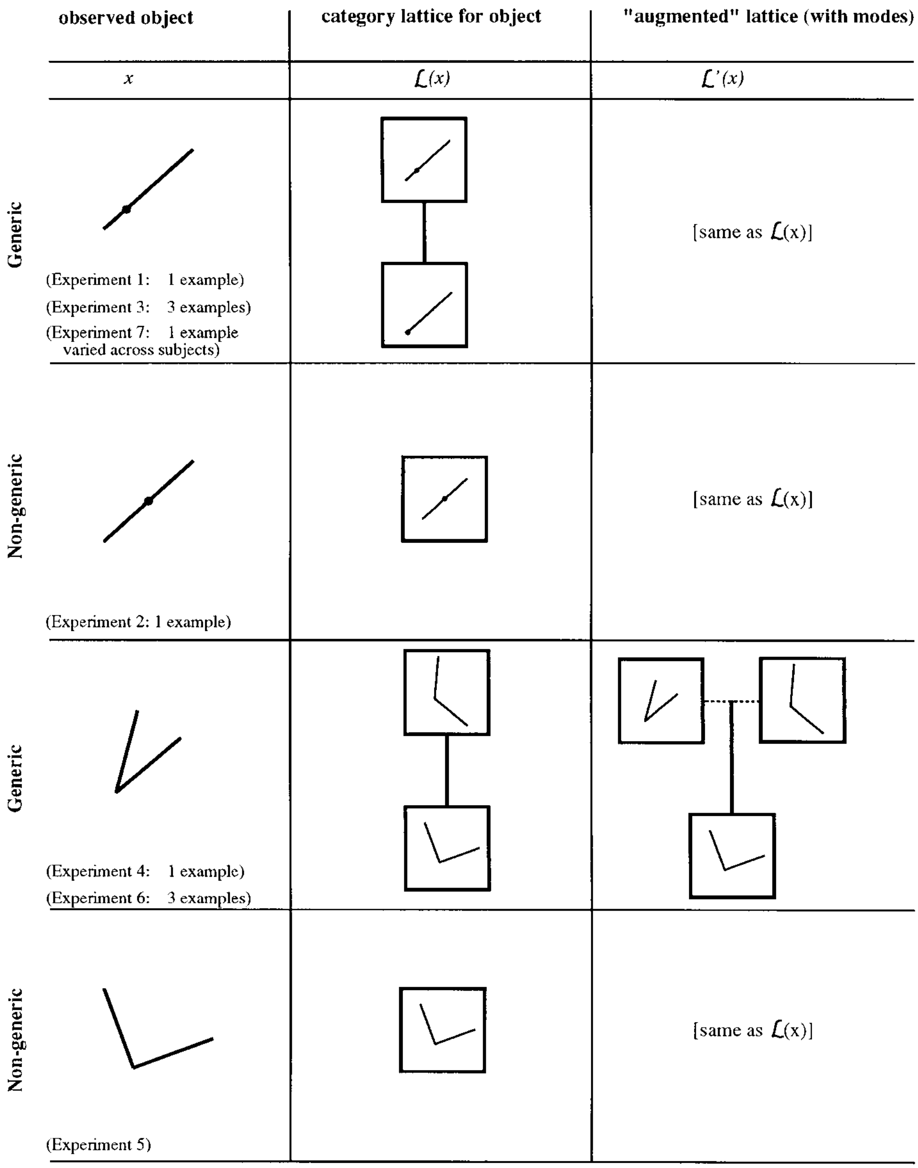

FIG. 8. A table showing the sample objects used in each of the experiments, their associated category lattices, and their augmented lattices in schematic form.

over 6 empty boxes. The phrasing of the ostension instructions was intended to cue the subject to draw a category interpretation rather than a fixed-object interpretation, yet without explicitly instructing the subject to vary some parameters in novel examples. This methodology rests on the presumption that subjects will tend to assume that names belong with kinds, identifying coherent concepts with entries in the lexicon (see Landau and Jackendoff, 1993). Subjects apparently did generally draw the category interpretation, understanding that novel examples did not need to resemble the sample object(s) exactly. For example, virtually every subject in the line–dots experiments varied the orientation of the objects freely, evidently attaching no causal importance to the absolute orientation of the sample.

The large majority of subjects were drawn from the university community at large. Some but not all subjects participated in both a line–dots and a V's session; when one subject was tested in both conditions, the order of the two conditions was determined randomly. Each session took less than 5 min.

Coding of Data

Novel objects produced by subjects were coded by hand and tallied in a histogram expressing frequency as a function of a measurement of the objects' structure.

For line–dots, the measurement was proportional to the ratio of the distance from the dot to the nearer endpoint of the line to the length of the line, i.e., the position of the dot along the line as an absolute (unit-free) number. The position of the dot was encoded as a number from 0 to 10, 0 meaning at the endpoint, and 10 meaning at the midpoint. Note that by measuring from the nearer endpoint the data are in effect collapsed about the midpoint. One unit in this scale is about the same size as the dot.

For V's, the measurement was the interior angle of the V, which was measured to the nearest 5°.

Objects that could not be coded via the above parameters were discarded. The percent of data discarded in each experiment, typically about 10% and ranging from 0 to 30.4%, is given below. The nature of uncodable responses will be discussed more explicitly below.

General Method of Analysis

After coding, objects in each experiment were tallied in a histogram $H(x)$ showing frequency as a function of the measured variable x (distance along the line or angle). First, the probability density function underlying this histogram was estimated using a Gaussian kernel (see Silverman, 1986), with the resolution of the data (1 unit for line–dots, 5° for angles) as a smoothing parameter. (In the figures this estimated density is scaled by n for comparison with the raw frequency histogram.) As discussed above, modulo certain assumptions, the number of components in the density can then be assessed by comparing the fit of models with various numbers of components. The procedure was as follows. The density was fitted (using Levenburg–Marquardt and squared error as a loss function) to a mixture of k heteroscedastic normal components,

$$H(x) = \sum_{i=1}^{k} \frac{h_i}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2\right], \tag{9}$$

where μ_i, σ_i, and h_i are the mean, standard deviation, and weight of the ith component, respectively. The null model has k equal to the number of generic modes (e.g., 1 in the line–dots experiments, 2 in the V's experiments). The hypothesized model has k one greater, i.e., now including the putative nongeneric mode.

The fits of these two models were then compared to yield the likelihood ratio λ between the larger and smaller models. Because the two models are strictly embedded, the test statistic $-2 \log \lambda$ is normally expected to be distributed as χ^2 with degrees of freedom equal to the difference in the number of parameters between the two models. As discussed above, a more conservative estimate of χ^2 suggested by McLaghlan and Basford (1988) is

$$-2\left(\frac{n-1-p-\frac{1}{2}g}{n}\right)\log \lambda,$$

where n is the number of data points, p is the number of independent variables in the p.d.f. (here 1), and g is the number of components in the larger model. This corrected χ^2 is used below. Initial values of the fitting procedure were carefully controlled to ensure that the extra mode in the larger model actually consisted of the critical nongeneric mode. As discussed above, in interpreting the analyses care must be taken so that the estimated parameters of the fitted components do not deviate from the assumptions mandating the test. Estimated parameters and R^2's are given below for each "winning" model. Each plot shows the raw frequencies, the estimated density, and the winning model.

3.1. Experiment 1: One Generic Line–Dot

Subjects. Eighteen subjects were tested in this condition.

Materials. The sample object was a line segment with a dot placed generically on it (one quarter of the way along its length, i.e., at 5 units). For half the subjects, the line segment was at a shallow oblique angle (15° from horizontal); for the other half a steep one (80° from horizontal). Orthogonally, for half the subjects the dot was nearer to the upper end, and for half it was nearer to the lower end.[10]

Results. The results for Experiment 1 are shown in Fig. 9. Some 10% of objects were uncodable. In this and all later experiments, subjects freely varied the size and overall orientation of their figures, in addition to the variables actually plotted. Sizes and orientations were not recorded, as they were not the main focus of the experiment

The 2-component model was significant over the 1-component model, corrected $\chi^2 = 10.3415$, df $= 3$, $p < 0.025$. For the winning model, $R^2 = 0.9246$. In the winning model, the generic mode is at $\mu = 4.8446$ ($\sigma = 2.3610$, $h = 92.6877$), and the nongeneric mode is at $\mu = -0.0034$ ($\sigma = 0.5458$, $h = 6.8459$). The generic mode is near the location of the sample dot (5 units), while the additional mode is at the endpoint.

3.2. Experiment 2: One Non-generic Line-Dot

Subjects. Eighteen subjects were tested in this condition.

Materials. The sample object was a line segment with a dot at the midpoint. In light of the results of Experiment 1, the endpoint might seem a more obvious choice, but it was felt that in that case subjects might not even recognize the possibility of moving the dot along the line segment, inappropriately biasing the results toward the theory. Orientations were as in Experiment 1.

[10] These manipulations, like the use of both acute and obtuse angles in the V's experiments, were an attempt to minimize the effect of nuisance variables. Only the extreme labor-intensity of the coding prevented more orientations, etc., from being used.

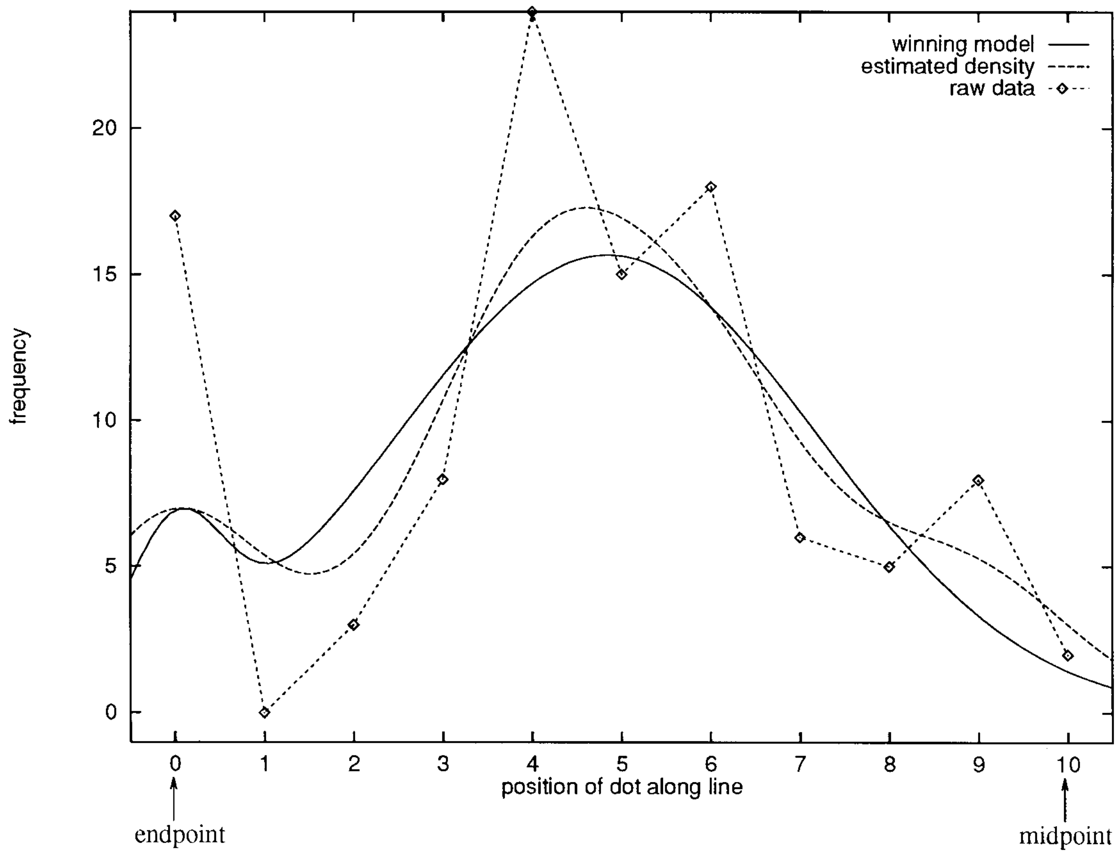

FIG. 9. Results of Experiment 1 (one generic line–dot), showing the winning model.

Results. The results for Experiment 2 are shown in Fig. 10. Again, subjects freely varied the size and orientation of their figures in addition to the manipulations shown in the plot. Some 5.8% of objects were uncodable.

The 2-component model is not significant, with corrected $\chi^2 = 1.6738$. (Uncorrected χ^2 is nonsignificant as well.) For the winning model, $R^2 = 0.9589$. The single mode in the winning model is at $\mu = 10.2906$ ($\sigma = 1.7766$, $h = 119.2412$), near the sample dot (10 units).

Discussion. The results of Experiments 1 and 2 taken together clearly confirm the increased complexity of the lattice in the dim-1 case vs the dim-0 case. Two nodes were in evidence when the sample object was generic, but only one when it was nongeneric. In the modal interpretation, the object in Experiment 1 was taken by subjects as an endpoint line–dot transformed by translation of the dot into a generic line–dot, with both categories showing up in their induced probability distributions, while the object in Experiment 2 was taken simply as itself, with no generative precursor. This is a qualitative difference between subjects' treatment of generic and nongeneric objects. Obviously, the choice of midpoint vs endpoints as the original primitive state has not been accounted for. Nevertheless, the role of the lattice in subjects' generalization is clear.

Moreover, the fact that in Experiment 1 the number of modes is greater than the number of sample objects—that subjects in effect induced *two* category populations from just *one* example—argues strongly that the observer brings a great deal of internal machinery to bear on the induction, rather than simply forming a statistical summary of observed examples. After all, the nongeneric component in the observed p.d.f. is based on no observed objects at all. The next experiment attempts to strengthen this point by presenting multiple sample objects, distributed as uniformly as possible both within and between subjects, so that again the structure in the resulting p.d.f.'s can be attributed to internal mechanisms rather than environmental frequencies.

3.3. Experiment 3: Three Generic Line-Dots

In this experiment, three different generic line–dots serve as sample objects. The idea is to present as sharp a contrast as possible between the environmental distribution and the distribution induced by subjects. Hence the pooled distribution of dot positions in the samples was constructed to mimic a uniform distribution as closely as possible, while avoiding all nongeneric locations. Thus if a nongeneric mode appears in the induced distribution, it can be attributed completely to the lattice, rather than to the

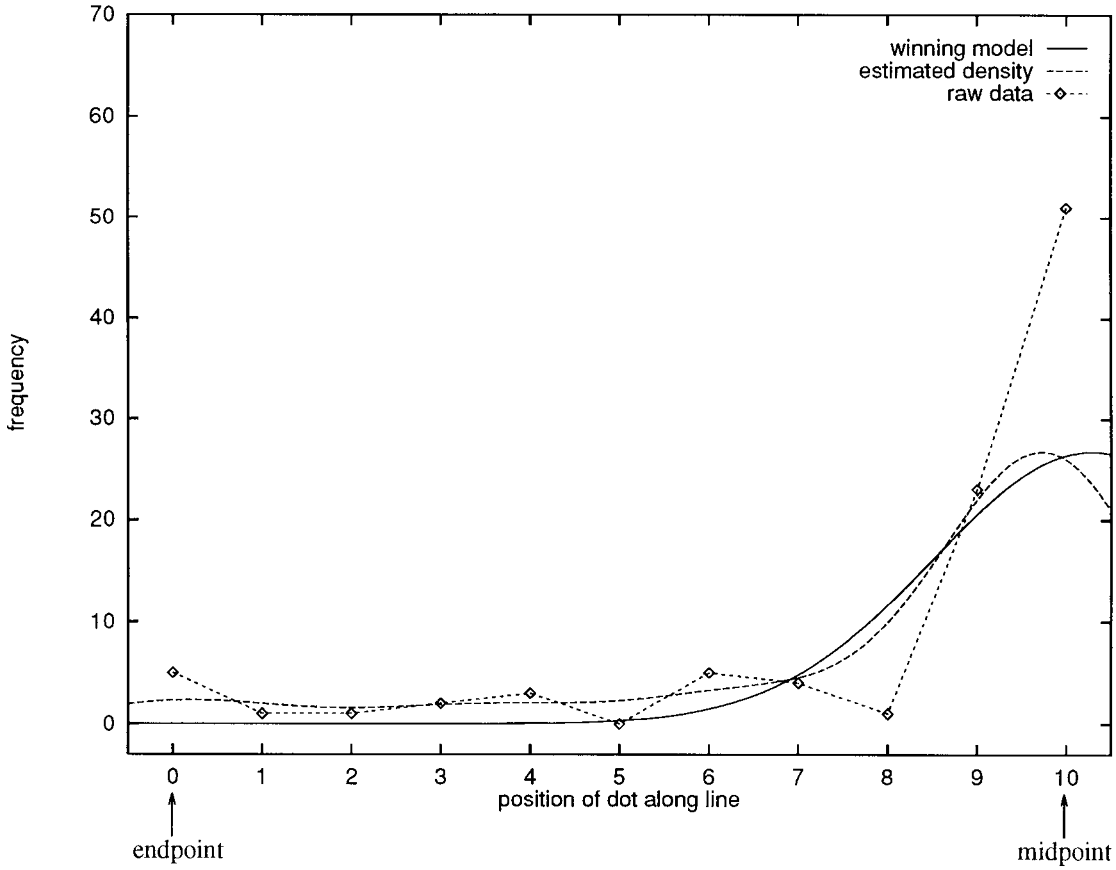

FIG. 10. Results of Experiment 2 (one nongeneric line–dot), showing the winning model.

environment. Of course, each subject only sees three points in the near-uniform distribution, so they can hardly be expected to have perceived its complete shape accurately, and hence individual subjects cannot be said to have "observed" a uniform distribution per se (see Fried and Holyoak, 1984, for a discussion of this problem). On the other hand, they certainly did not observe modes at nongeneric points either, since none were present; so the presence of such modes would tend to confirm the influence of the lattice.

Subjects. Eighteen subjects were tested in this condition.

Materials. Sample objects for each subject were 3 generic line–dots. One-third of the subjects saw objects with dot positions of 1, 2, and 7; one-third saw 3, 4, and 8; and one-third saw 5, 6, and 9. Hence the pooled distribution is flat from 1–9, while simultaneously the within-subject means are as uniformly distributed as possible at 3.33, 5, and 6.67, while also avoiding nongeneric points. The orientation of each object was determined randomly. Each subject saw objects of three different fixed sizes, permuted randomly. Half of the subjects saw two objects with dots closer to the upper half, and one with the dot closer to the lower half; the other half of the subjects had this reversed. This factor was crossed orthogonally with the distribution factor.

Results. The results for Experiment 3 are shown in Fig. 11. No objects were uncodable.

The 2-component model is significant over the 1-component model, corrected $\chi^2 = 269.0747$, df $= 3$, $p < 0.005$. For this model $R^2 = 0.9713$. The generic mode is at $\mu = 4.5933$ ($\sigma = 2.2621$, $h = 72.2813$), and the nongeneric mode is at $\mu = 9.2182$ ($\sigma = 1.2253$, $h = 34.1383$), near the midpoint.

As in Experiment 1, the presence of the additional nongeneric mode corroborates the main complexity effect. Because of the construction of the sample set, the visible modes in the distribution are due only to the lattice, and not to the environmental distribution.

Experiments 4–6 repeat the design of Experiments 1–3 with V's instead of line–dots. Note that (as discussed above) there are now two distinct generic modes expected, acute and obtuse (i.e., one on either side of the nongeneric point, assumed to be at 90°).

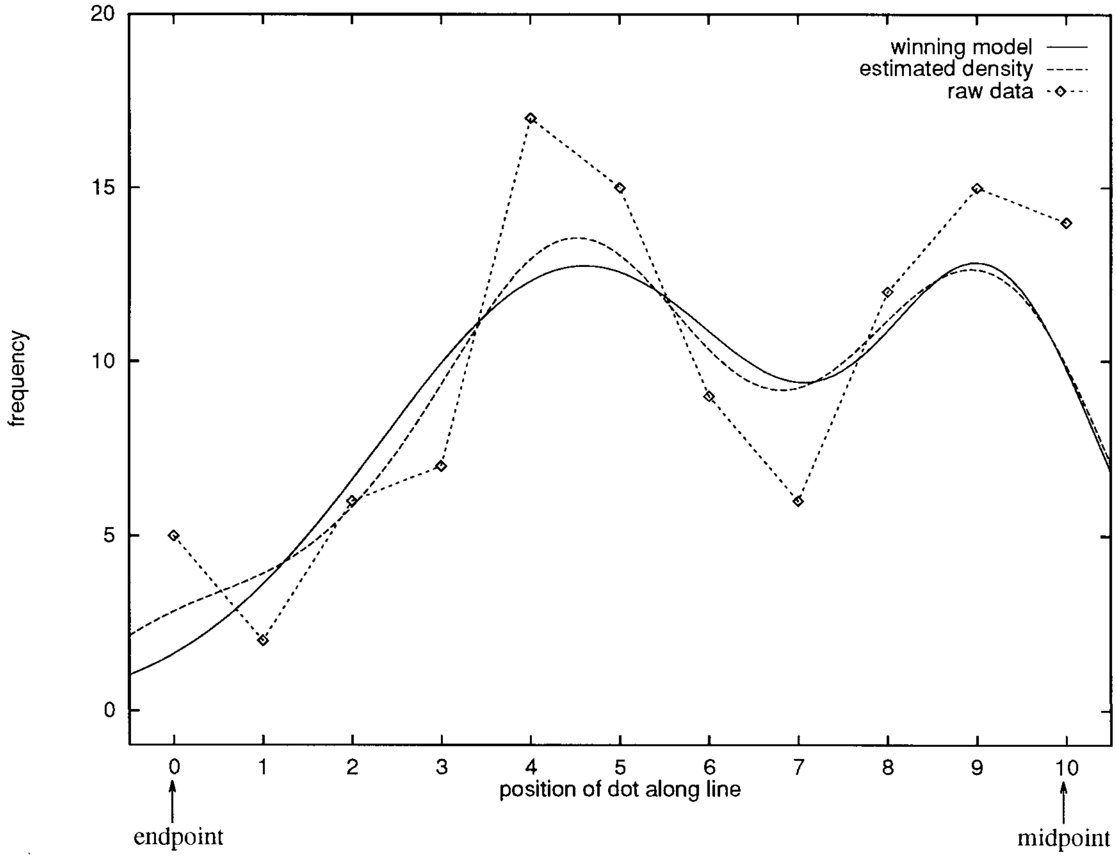

FIG. 11. Results of Experiment 3 (three generic line–dots), showing the winning model.

3.4. Experiment 4: One Generic V

Subjects. Sixteen subjects were tested in this condition.

Materials. The sample was a single generic V. Half of the subjects saw an acute (55°) version, while half saw an obtuse (120°) version, both angles chosen arbitrarily as generic-looking angles. The V's were all positioned at the same orientation, so that the (imaginary) bisector was at 45° from horizontal.

Results. The results for Experiment 4 are shown in Fig. 12. Some 30.4% of objects were uncodable.

The 3-component model is significant over the 2-component model, corrected $\chi^2 = 39.4771$, df $= 3$, $p < 0.005$. In this winning model, $R^2 = 0.9523$. One generic mode (acute) is at $\mu = 50.3824°$ ($\sigma = 11.4010°$, $h = 226.5331$); the other generic mode (obtuse) is at $\mu = 121.6172°$ ($\sigma = 13.2716°$, $h = 161.7464$); and the non-generic mode (right) is at $\mu = 84.02609°$ ($\sigma = 8.1893°$, $h = 79.6957$).

3.5. Experiment 5: One Nongeneric V

Subjects. Eighteen subjects were tested in this condition.

Materials. The sample object was a right (90°) V. Half of the subjects saw V's whose (imaginary) bisector pointed toward 60° from the horizontal, and half saw Vs pointed towards 320° from the horizontal.

Results. The results for Experiment 5 are shown in Fig. 13. Some 5.3% of the objects were uncodable.

With the 2-component model the estimation procedure converges to two superimposed modes less than 0.1° apart (90.2556° and 90.3331°). This solution violates Assumption 2 above, so it must be rejected, and the 1-component model wins. For the winning model, $R^2 = 0.9591$. The one mode is at $\mu = 90.1732$ ($\sigma = 12.1666$, $h = 451.7745$), i.e., a right angle just as in the sample object.

Discussion. The pattern in Experiments 4 and 5 is just as in Experiments 1 and 2: the density resulting from a nongeneric object is just a single mode, but that resulting from a generic object includes both generic and nongeneric modes, reflecting the entire lattice.

3.6. Experiment 6: Three Generic V's

Subjects. Twenty-four subjects were tested in this condition.

Materials. Sample objects for each subject were 3 generic V's. As in Experiment 3, an elaborate scheme was

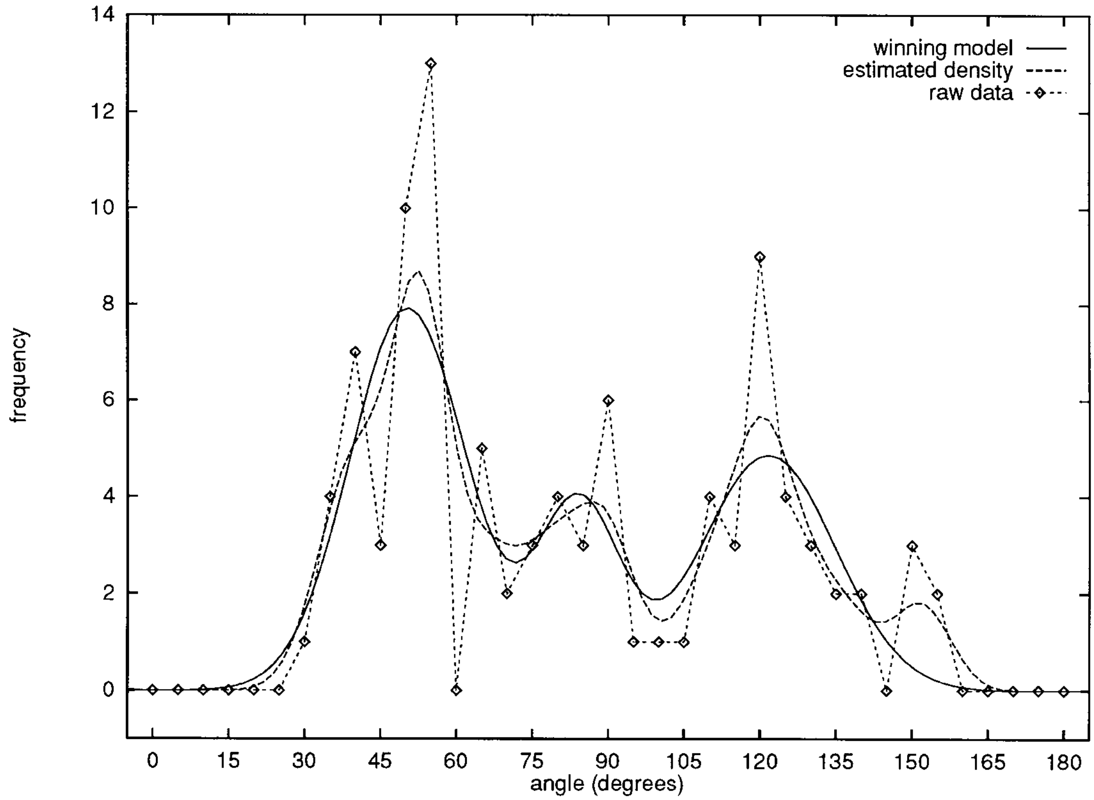

FIG. 12. Results of Experiment 4 (one generic V), showing the winning model.

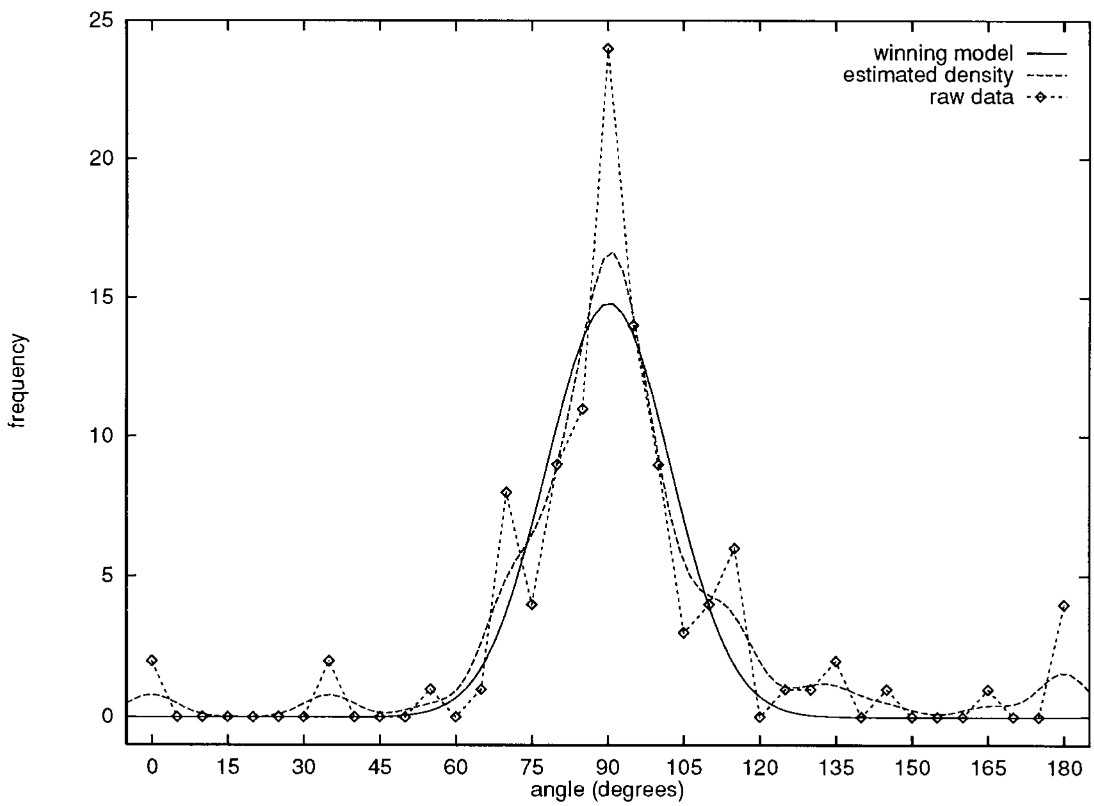

FIG. 13. Results of Experiment 5 (one nongeneric V), showing the winning model.

TABLE 1

Angles Used in Each of the Six Groups of V's Presented to Subjects

	Angles in Sample	Mean Angle of Sample
1	25°, 55°, 130°	70°
2	30°, 60°, 135°	75°
3	35°, 65°, 140°	80°
4	40°, 115°, 145°	100°
5	45°, 120°, 150°	105°
6	50°, 125°, 155°	110°

devised to make both the pooled and the within-subject means as close to uniform as possible while avoiding nongeneric points (Table 1).

One-sixth of the subjects saw each of the six sample groups. Note that each of the angles from 25–65° and from 115–155° is used exactly once over the whole set, so the overall distribution is exactly flat in these ranges. Note as well that each subject saw either two acute and one obtuse angle or two obtuse and one acute, so no subject would be tempted to conclude that all members of the class must be of one kind or the other. Each subject saw objects of three different fixed sizes, permuted randomly. All the Vs had one leg exactly horizontal, in order to force subjects to focus on the angle between the legs and thus reduce the number of "incorrect" inductions observed in the first generic V's experiment.

Results. The results for Experiment 6 are shown in Fig. 14. Some 7.7% of objects were uncodable.

The 3-component model is significant, corrected $\chi^2 = 25.3267$, $\mathrm{df} = 3$, $p < 0.005$. In this winning model, $R^2 = 0.9120$. In the model, the generic modes are at $\mu = 49.7090°$ ($\sigma = 26.0851°$, $h = 363.7483$) and $\mu = 128.7955°$ ($\sigma = 19.0470$, $h = 315.0620$), near the acute and obtuse samples respectively. The nongeneric mode is at $\mu = 89.9219$ ($\sigma = 3.9619$, $h = 29.5632$) near the right angle.

3.7. Experiment 7: One Generic Line-Dot, Variable

In Experiment 1, the generic mode was at about the same location as the sample object. It is unclear whether this location derived from the sample object per se or from some kind of principled maximization of genericity. Experiment 7 attempts to disentangle these two factors. Again a single generic line–dot is used as a sample object, but now the position of the dot is varied between subjects. As in Experiment 3, the pooled distribution is uniform over the entire generic region. Again, of course, subjects cannot be expected to accurately perceive the overall distribution, since each only sees one object. But because the location of the sample object is now manipulated freely, it is possible to probe its influence on the location of the resulting mode.

Subjects. Twenty-eight subjects were tested in this condition.

Materials. The induction example was again a line segment with a dot placed some distance along it. One-seventh of the subjects observed an object at each of the distances 2–8. The nongeneric values 0 and 10 were omitted as before. The values 1 and 9 were omitted as well, for fear that with only one sample they were too confusible with 0 and 10. The dot was always closer to the lower end.

Results. The results for Experiment 7 are shown in Fig. 15. Some 15.2% of the objects were uncodable.

The 2-component model is not significant over the 1-component model (corrected $\chi^2 = 1.0568$). The raw data (plotted in the figure) show small but evident modes at both nongeneric points (endpoint and midpoint), but they are not significant.

Split Group Analysis. The fact that the generic mode is at nearly the same location as in Experiment 1 suggests that manipulating the location of the sample does not influence its position. This is a very weak test, however, because the mean sample location across subjects is again at 5 units, confounded with the location of the generic mode. A much stronger test is as follows. The subjects were split into two groups, one with near-endpoint locations (2, 3, and 4), the other with near-midpoint locations (6, 7, and 8), omitting the subjects at 5 units. The two groups' sample means are separated by 4 units (3 vs 7). In order to determine the location of the generic modes in their p.d.f.'s as sensitively as possible, each data set was first stripped of the data at 0, 1, 9, and 10 units, and then fitted with single Gaussians. The endpoint group's mode (4.6117 units) is in fact closer to the endpoint, while the midpoint group's mode (5.7525 units) is closer to the midpoint. However, the separation between the two means is only 0.8223 units, as compared to the difference of 4 between the sample means. Moreover, the distance between the lower end of the 95% confidence interval for the 2-3-4 group and the uper end of the 95% confidence interval for the 6-7-8 group (1.4166 units) is still far less than four units, so even under the most liberal interpretation the two group modes are far less separated than are the sample means.

Thus the interpretation is mixed: the location of the generic mode is influenced by the location of the sample object, but tends to migrate substantially toward a point of maximum genericity. This maximally generic point is as far as possible from both nongeneric points (i.e., at 5 units), exemplifying a theoretical principle explored more fully in Feldman (1996).

Uncodable Data. The magnitude and variability of the amount of uncodable data call for some comment. Uncodable objects ranged from reasonable alternative inductions (e.g., a dot and line not touching) to fanciful

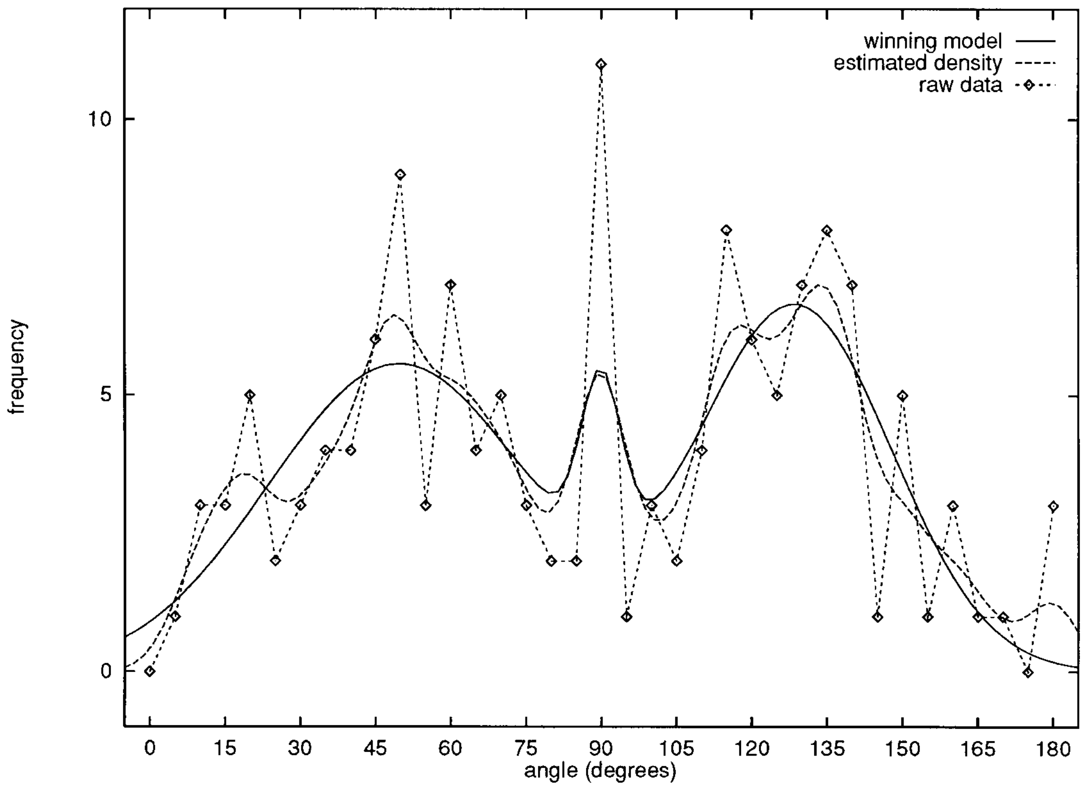

FIG. 14. Results of Experiment 6 (three generic V's), showing the winning model.

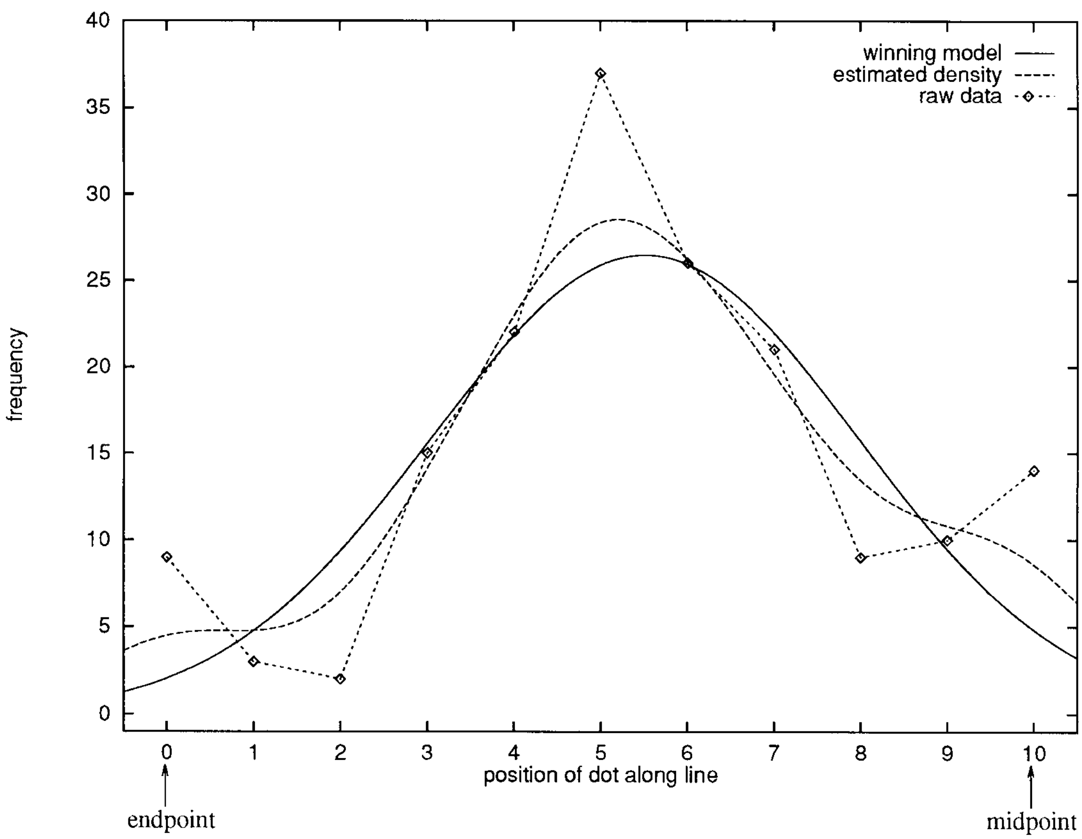

FIG. 15. Results of Experiment 7 (one generic line–dot, variable), showing the winning model.

drawings of houses, dragons, etc.—the latter perhaps reflecting subjects' awareness of the totally unconstrained nature of the induction. For example, in the condition with the greatest number of uncodable responses (Experiment 4, one generic V), seven subjects produced uncodable data. Of these, two drew unrelated cartoons; two drew V's with the dot not at the vertex, making the angle of the V too ambiguous to code; two drew V's with the two segments not touching, again making the angle too ambiguous to code; and only one drew what appears to be a legitimate alternate induction, objects with three segments coterminating instead of two. Hence more often than not the uncodable responses simply reflect somewhat haphazard drawing by the subjects; legitimate alternate inductions are infrequent and comprise far too wide a variety of forms to support any systematic theory.

One very salient trend is that the experiments with multiple sample objects (Experiments 3 and 6) produced fewer uncodable objects than those with only one. This is a natural result of subjects' tendency to build categorical interpretations out of the apparently nonaccidental aspects of the induction samples, as hypothesized by the theory. Say the induction sample has some special property (e.g., the dot touching the line) which can be expected to occur at random with probability ε. When there are three such examples, the probability that they all exhibit the property independently falls to ε^3, with $\varepsilon^3 \ll \varepsilon$. When multiple sample objects are observed, all exhibiting the same regularity, the inference is all the more reliable and unambiguous, and subjects' responses are consequently more unanimous.

3.8. Discussion

The above experiments paint a picture of object generalization in which observers build generative models of objects and then choose the model for which the observed object is a generic outcome. Objects thus belong in the same category just when they are generic in the same model, that is, when they share the same generative histories. The formal instruments for accomplishing this are the regularity lattice, which enumerates the available models and partially orders them by genericity, and the Genericity Constraint, which dictates which interpretation to choose.

The structure of the lattice is revealed by the contrast of subjects' generalizations of single generic sample objects (Experiments 1 and 4) vs single nongeneric objects (Experiments 2 and 5). A nongeneric object has only a trivial one-node lattice, and the resulting induced distribution has only one mode; no more complex generative model is justified. A generic object, on the other hand, has multiple nodes on its lattice, and the resulting distribution has multiple modes. In a sense, the presence of nongeneric modes in the probability densities is peculiar: after all, subjects presumably do not regard such nongeneric objects as good examples of the induced category. Nevertheless, their presence is instrumental in laying bare the structure of the lattice.

Admittedly, there is still something unclear about how subjects choose the primitive objects from which to begin generative models, and hence how the locations of nongeneric modes are determined. In the line–dots experiments, for example, the primitive object was sometimes the endpoint version and sometimes the midpoint. One possibility is that subjects build object models up from a psychologically salient nearby form; hence only when both near-endpoint and near-midpoint generic objects are used as examples, as in Experiment 7, do both nongeneric modes show up (albeit only slightly). Certainly, biases specific to the object class may come into play, and a much more elaborate theory of generative models would be required to completely answer this question.

Another unexplained phenomenon concerns the widths (σ's) of the various modes. Generally, nongeneric modes tend to be narrower than generic ones. This is consonant with the motivating view of lower-dimensional categories as manifolds passing through the higher-dimensional space. Such categories are essentially infinitesimal punctures of the embedding space and would be expected to more closely resemble "spikes" or Dirac pulses in the resulting probability distribution. Nevertheless this issue is not completely clear and nothing in the main prediction depends on it.

Perhaps the most important outcome of these experiments is simply the overwhelming unanimity of subjects' inductions. Given the unconstrained nature of the task, subjects were free to make arbitrarily bizarre or idiosyncratic generalizations (and, of course, in a small fraction of cases they did so). Even restricting attention to salient parameters (e.g., the location of the dot in line–dots), subjects agreed about which transformations made sense and which did not. The dot could be moved along the line, but not off it. The V's angle could be changed, but the two line segments could not be separated.

These decisions all reflect the Genericity Constraint. If the category in Experiment 1 contained objects with the dot off the line, then the example object was certainly a very atypical example. Hence this category makes a poor hypothesis. Likewise, if the category in Experiment 4 were simply "two line segments, not necessarily touching," then the touching line segments from the sample object would be an extremely "suspicious coincidence" (Barlow, 1989). Even in one-shot generalization, one can make guesses about what aspects of the sample are likely to be shared in other members of the category. Genericity provides the inferential leverage.

4. GENERAL DISCUSSION

The lattice theory, coupled with the Genericity Constraint, provides a concrete and computable method of performing inductive generalization in the context investigated here (using simple visual figures and a very small number of sample objects). Moreover, this method of generalization instantiates a very natural and powerful principle, namely that unobserved members of a category are likely to have been generated by the same causal process that brought about the observed examples.

The pattern of modality seen in the experiments provides a prima facie confirmation of the main parts of the theory, namely the structure of the lattices and the Genericity Constraint. Admittedly, subjects employ a considerable amount of knowledge that is not accounted for here, including detailed guesses about what types of spatial operations should be invoked in each different type of object model. Moreover, many of the assumptions that gave rise to the theory (the continuity of categories, the group-theoretic generation of object models, etc.) were not tested, and are in fact probably not testable. Rather, they are to be regarded as motivating axioms, and are of interest only for the concrete formal structures they give rise to.

The theory highlights the degree and subtlety with which internal mental operations can be brought to bear on induction problems when (as in these experiments) observers have very little data to go on. Indeed, in some of the experiments (e.g., 1 and 4), there were more modes in the subjects' heads (so to speak) than in the world. Admittedly, it is obvious that without a rich environment, observers are left to their own internal devices in constructing categories. Nevertheless, it is worth contrasting this point with the assumptions underlying many other generalization theories, e.g., connectionist nets, which in one form or another draw statistical summaries of the input in a relatively unbiased manner. In such theories the best generalization of a single data point would simply be centered on the point itself. By contrast, the construction of generative models—and the resulting distinction between object models of various degrees of dimensionality—imposes a heavy bias.

These experiments bear on a recent controversy concerning the nature of concepts and categories. A tremendous amount of evidence suggests that membership in human psychological categories is not definite, but rather is graded by typicality. Yet some have argued that this seemingly unassailable empirical fact makes problematic the *composition* of categories into new more complex categories, which is presumably necessary so that complex concepts can be learned, expressed, and understood (Fodor, 1994). Yet the seemingly clear dichotomy between graded categories and symbolic composable concepts is called into question in the lattice theory. The basic category definition (manifolds and generative models) is eminently qualitative and discrete—a point is either on a manifold or not, either spanned by a model or not. In fact, the lattice itself is nothing other than an explicit compositional system for combining primitives to form categories. Yet the empirical expression of the lattices certainly includes degrees of membership, as evidenced by the smooth distributions in the probability densities above.

5. CONCLUSION

In a classic essay, the mathematician Weyl (1952) noted our ubiquitous tendency to ascribe symmetries to the world and to describe the world in such a way as to emphasize its symmetry. Similarly, Poston and Stewart (1978) have noted that many global world symmetries and regularities that are in principle unverifiable (such as the multifarious symmetries of Euclidean space) may have very high (even infinite) codimension in the set of all possible worlds—and yet we take them for granted. Indeed, the over-regularity of the physical world is the common thrust of many attempts to account for and justify the human capacity to generalize, ranging from Hume's "Principle of the Uniformity of Nature" to the "Principle of Natural Modes" (Bobick, 1987; Richards and Bobick, 1988).

The contribution of this paper is to show how a related notion of "regularity" can be applied to category inference. Despite the inductive ambiguity inherent in generalization, observers expect unknown examples of a category to be largely predictable from known examples. This principle underlies the Genericity Constraint, which requires in effect that the examples and the category are systematically related to one another. Moreover, categorical hypotheses can be ordered by the degree to which they satisfy this principle, an ordering made explicit in the regularity lattice. Hence the generic interpretation is in a sense the categorization that is most in harmony with the observer's underlying model of regularity in the world.

APPENDIX: PROOF OF THEOREM 1

THEOREM 1 (Existence and Uniqueness of the Generic Interpretation). *Given $\mathscr{L}_{\Phi,\Pi}$, for any $x \in \Phi$,*

(i) *there exists an $R \in \mathscr{L}_{\Phi,\Pi}$ such that x is generic in R, and*

(ii) *R is unique.*

Proof. (i) (*Existence*). By contradiction. We know that x is contained in at least one $R \in \mathscr{L}$, since Φ is in $\mathscr{L}$, and $x \in \Phi$. Denote by R_x the set of regularities in $\mathscr{L}$ that contain x (i.e., $R_x = \{R \in \mathscr{L} \mid x \in R\}$). We know that $\bigcap R_x$ exists and is in $\mathscr{L}$, since $\mathscr{L}$ is closed under intersection. x is in $\bigcap R_x$, and in fact is generic in it, since there are no other regularities not in R_x that contain x.

(ii) (*Uniqueness*). By contradiction. Assume x is generic in two distinct regularities $R, S \in \mathscr{L}$, $R \neq S$. If $R \subset S$, then x is not generic in S, contrary to hypothesis; similarly if $S \subset R$. If neither $R \subset S$ nor $S \subset R$ (R and S are unordered in the partial order), then $R \cap S$ is nonempty and $x \in R \cap S$, so that x is not generic in either R or S, again contrary to hypothesis.

REFERENCES

Aitkin, M., and Rubin, D. (1985). Estimation and hypothesis testing in finite mixture models. *Journal of the Royal Statistical Society B*, **47**(1), 67–75.

Ashby, F. G. (1992). Multidimensional models of categorization. In F. Ashby (Ed.), *Multidimensional models of perception and cognition*. Hillsdale, NJ: Lawrence Erlbaum.

Ashby, F. G., and Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychology Review*, **95**(1), 124–150.

Barlow, H. (1989). Unsupervised learning. *Neural Computation*, **1**, 295–311.

Barr, A. (1984). Global and local deformations of solid primitives. *Computer Graphics*, **28**(3), 21–30.

Barrow, H. G., and Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence*, **17**, 75–116.

Bennett, B. M., Hoffman, D. D., and Prakash, C. (1989). *Observer mechanics: A formal theory of perception*. London: Academic Press.

Binford, T. (1981). Inferring surfaces from images. *Artificial Intelligence*, **17**, 205–244.

Bobick, A. (1987). *Natural object categorization* (Technical Report 1001). Cambridge, MA: Massachusetts Institute of Technology Artificial Intelligence Laboratory.

Davey, B., and Priestley, H. (1990). *Introduction to lattices and order*. Cambridge: Cambridge University Press.

Everitt, B., and Hand, D. (1981). *Finite mixture distributions*. London: Chapman and Hall.

Farouki, R., and Hinds, J. (1985). A hierarchy of geometric forms. *IEEE Computer Graphics and Applications*, 51–78.

Feldman, J. (1991). Perceptual simplicity and modes of structural generation. In *Proceedings of the 13th Annual Conference of the Cognitive Science Society*.

Feldman, J. (1992a). Constructing perceptual categories. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.*

Feldman, J. (1992b). *Perceptual categories and world regularities* (Ph.D. thesis). Massachusetts Institute of Technology; also available as Technical Report 6, Rutgers Center for Cognitive Science.

Feldman, J. (1995). Perceptual models of small dot clusters. In *Proceedings of the DIMACS Workshop on Partitioning Data Sets* (DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 19, pp. 331–355).

Feldman, J. (1996). Regularity vs genericity in the perception of collinearity. *Perception*, **25**, 335–342.

Fodor, J. (1994). Concepts: a potboiler. *Cognition*, **50** 95–113.

Fried, L. S., and Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**(2), 234–257.

Ghosh, J., and Sen, P. (1985). On the asymptotic performance of the log likelihood ratio statistic for the mixture model and related results. In L. Le Cam and R. Olshen (Eds.), *Proceedings of the Berkeley Conference in the Honor of Jersy Neyman and Jack Kieffer* (Vol. 2). Monterey: Wadsworth.

Hatfield, G., and Epstein, W. (1985). The status of the minimum principle in the theoretical analysis of visual perception. *Psychological Bulletin*, **97**(2), 155–186.

Hochberg, J., and McAlister, E. (1953). A quantitative approach to figural "goodness." *Journal of Experimental Psychology*, **46**, 361–364.

Hoel, P., Port, S., and Stone, C. (1971). *Introduction to statistical theory*. Boston: Houghton Mifflin.

Jepson, A., and Richards, W. A. (1991). *What is a percept?* (Occasional Paper 43). MIT Center for Cognitive Science.

Jepson, A., and Richards, W. A. (1992). A lattice framework for integrating vision modules. *IEEE Transactions on Systems, Man, and Cybernetics*, **22**(5), 1087–1096.

Jepson, A., and Richards, W. A. (1992). What makes a good feature? In L. Harris and M. Jenkin (Eds.), *Spatial Vision in Humans and Robots*. Cambridge University Press.

Kanade, T. (1981). Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence*, **17**, 409–460.

Landau, B., and Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, **16**(2), 217–265.

Leeuwenberg, E. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, **84**(3), 307–349.

Leeuwenberg, E., and Boselie, F. (1988). Against the likelihood principle in visual form perception. *Psychological Review*, **95**, 485–491.

Leyton, M. (1984). Perceptual organization as nested control. *Biological Cybernetics*, **51**, 141–153.

Leyton, M. (1988). A process-grammar for shape. *Artificial Intelligence*, **34**, 213–247.

Leyton, M. (1989). Inferring causal history from shape. *Cognitive Science*, **13**, 357–387.

Leyton, M. (1992). *Symmetry, Causality, Mind*. Cambridge, MA: M.I.T. Press.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, **31**, 355–395.

Mark, L., Shapiro, B., and Shaw, R. (1986). Structural support for the perception of growth. *Journal of Experimental Psychology: Human Perception and Performance*, 149–159.

McLachlan, G., and Basford, K. (1988). *Mixture models: Inference and applications to clustering*. New York: Marcel Dekker.

Poston, T., and Stewart, I. N. (1978). *Catastrophe theory and its applications*. London: Pitman Publishing Limited.

Quine, W. (1965). On simple theories of a complex world. In M. Foster and M. Martin (Eds.), *Probability, confirmation, and simplicity: Readings in the philosophy of inductive logic*. New York: Odyssey Press.

Quine, W. (1985). Natural kinds. In H. Kornblith (Ed.), *Naturalizing Epistemology*. Cambridge, MA: M.I.T. Press.

Quinn, B., McLachlan, G., and Hjort, N. (1987). A note on the Aitkin-Rubin approach to hypothesis testing in mixture models. *Journal of the Royal Statistical Society B*, **49**(3), 311–314.

Richards, W. A., and Bobick, A. (1988). Playing twenty questions with nature. In Z. Pylyshyn (Ed.), *Computational Processes in Human Vision: An Interdisciplinary Perspective*, pp. 3–26. Norwood, NJ: Ablex Publishing Corporation.

Richards, W. A., Feldman, J., and Jepson, A. (1992). From features to perceptual categories. In *Proceedings of the British machine vision conference*.

Richards, W. A., Jepson, A., and Feldman, J. (1996). Priors, preferences, and categorical percepts. In D. Knill and W. A. Richards (Eds.), *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, **14**, 465–471.

Rosch, E., Mervis, C., Gray, W., Johnson, D., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382–439.

Saund, E. (1987). Dimensionality-reduction using connectionist networks. (A.I. Memo 941). Massachusetts Institute of Technology Artificial Intelligence Laboratory.

Shepard, R. (1987). Toward a universal law of generalization for psychological science. *Science*, **237**, 1317–1323.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Simon, H. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, **79**(5), 369–382.

Smith, E., and Medin, D. (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.
Sober, E. (1975). *Simplicity*. London: Oxford University Press.
Thompson, D. (1942). *On growth and form*. Cambridge: Cambridge University Press.
Watanabe, S. (1969). *Knowing and guessing: a quantitative study of inference and information*. New York: John Wiley.
Watanabe, S. (1985). *Pattern recognition: human and mechanical.* New York: John Wiley.
Weyl, H. (1952). *Symmetry*. Princeton: Princeton University Press.
Witkin, A. P., and Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope, and A. Rosenfeld (Eds.), *Human and machine vision* (pp. 481–543). San Diego: Academic Press.
Zucker, S. W., and Davis, S. (1988). Points and endpoints: a size/spacking constraint for dot grouping. *Perception*, **17**, 229–247.

Received: May 9, 1995