Toward a Universal Law of Generalization for Psychological Science

Roger N. Shepard

A psychological space is established for any set of stimuli by determining metric distances between the stimuli such that the probability that a response learned to any stimulus will generalize to any other is an invariant monotonic function of the distance between them. To a good approximation, this probability of generalization (i) decays exponentially with this distance, and (ii) does so in accordance with one of two metrics, depending on the relation between the dimensions along which the stimuli vary. These empirical regularities are mathematically derivable from universal principles of natural kinds and probabilistic geometry that may, through evolutionary internalization, tend to govern the behaviors of all sentient organisms.

The tercentenary of the publication, in 1687, of Newton's *Principia* (*1*) prompts the question of whether psychological science has any hope of achieving a law that is comparable in generality (if not in predictive accuracy) to Newton's universal law of gravitation. Exploring the direction that currently seems most favorable for an affirmative answer, I outline empirical evidence and a theoretical rationale in support of a tentative candidate for a universal law of generalization.

Primacy of Generalization

Because any object or situation experienced by an individual is unlikely to recur in exactly the same form and context, psychology's first general law should, I suggest, be a law of generalization. Learning theorists have seemed to suppose that a principle of conditioning (through contiguity or reinforcement) could be primary and that how what is learned then generalizes to new situations could be left for later formulation, as a secondary principle. Unfortunately, a full characterization of the change that even a single environmental event induces in an individual must entail a specification of how that individual's behavioral dispositions have been altered relative to any ensuing situation. Differences in the way individuals of different species represent the same physical situation implicate, in each individual, an internal metric of similarity between possible situations. Indeed, such a metric exists at birth, when habituation to one stimulus already exhibits unequal generalization to different test stimuli (*2*).

Recognition that similarity is fundamental to mental processes can be traced back over 2000 years to Aristotle's principle of association by resemblance. Yet, the experimental investigation of generalization did not get under way until the turn of this century, when Pavlov found that dogs would salivate not only at the sound of a bell or whistle that had preceded feeding but also at other sounds—and more so as they were chosen to be more similar to the original sound, for example, in pitch (*3*). Since then, numerous experimenters have obtained empirical "gradients of stimulus generalization," relating the strength, probability, or speed of a learned response to some measure of difference between each test stimulus and the original training stimulus.

However, methods yielding reliable gradients of generalization were not perfected until the middle of this century. In 1956, Guttman and Kalish (*4*) demonstrated that Skinner's operant conditioning technique of intermittent reinforcement (*5*) could be used to obtain orderly gradients of generalization for animals. A pigeon that was only intermittently permitted access to grain for pecking a translucent key illuminated by light of a particular wavelength would continue to respond long after termination of all reinforcement (*6*). Guttman and Kalish could then measure stable rates of responding to many different test wavelengths. And between 1955 and 1958, I established that orderly gradients of generalization could be obtained from humans during identification learning—in which subjects acquired, through correction of incorrect responses, a one-to-one association between *n* stimuli (Munsell color chips, for example) and *n* arbitrarily assigned verbal responses (*7–9*). The frequency with which any stimulus led to the response assigned to any other provided the measure of generalization between those two stimuli.

Apparent Noninvariance of Generalization

In striving to establish psychology as a quantitative science, researchers had traditionally preferred to choose, as the independent variable, a physical measure of stimulus difference—such as the difference in wavelengths of lights, frequencies of tones, or angular orientations of shapes. However, quantification does not in itself guarantee invariance. Probability (or rate) of a generalized response reliably decreased with physical difference from the training stimulus. However, the way it decreased varied from one training stimulus, sensory continuum, or species to another. Generalization could even exhibit a nonmonotonic increase between stimuli separated by certain special intervals—for example, between tones separated by an octave (*10*), between hues at the opposite (red and violet) ends of the visible spectrum (*11*), and between shapes differing by particular angles related to inherent symmetries of those shapes (*12*).

At midcentury, influential behavioral scientists (including the neurophysiologist Karl S. Lashley and the mathematical learning theorists Robert R. Bush and Frederick Mosteller) were reaching the discouraging conclusion that there could be no invariant law of generalization (*13*). If we took physical difference as the independent variable, gradients of generalization, reflecting properties of the

The author is professor of psychology at Stanford University, Stanford, CA 94305.

particular animal as much as the physically measured differences between the stimuli, could not be expected to be uniform or even monotonic. If, instead, we sought a psychological measure of difference as the independent variable, the most basic such measure would surely be the generalization data themselves—apparently rendering the attempt to determine a functional law entirely circular.

Invariance in Psychological Space

What is sometimes required is not more data or more refined data but a different conception of the problem. Newton arrived at universal laws of motion only by departing from Aristotle's and Ptolemy's choice of the concretely given earth as the fixed reference and by choosing, instead, an abstractly conceptualized absolute space, with respect to which all objects including the earth move according to the same laws (*1*). And 230 years later, in order to ensure that the laws of physics are invariant for all observers regardless of their own relative motions, Einstein had to replace Newton's Euclidean space with an even more abstract four-dimensional Riemannian manifold (*14*).

Analogously in psychology, a law that is invariant across perceptual dimensions, modalities, individuals, and species may be attainable only by formulating that law with respect to the appropriate abstract psychological space. The previously troublesome variations in the gradient of generalization might then be attributable to variations in the psychophysical function that, for each individual, maps physical parameter space (the space whose coordinates include the physical intensity, frequency, and orientation of each stimulus) into that individual's psychological space. If so, a purely psychological function relating generalization to distance in such a psychological space might attain invariance.

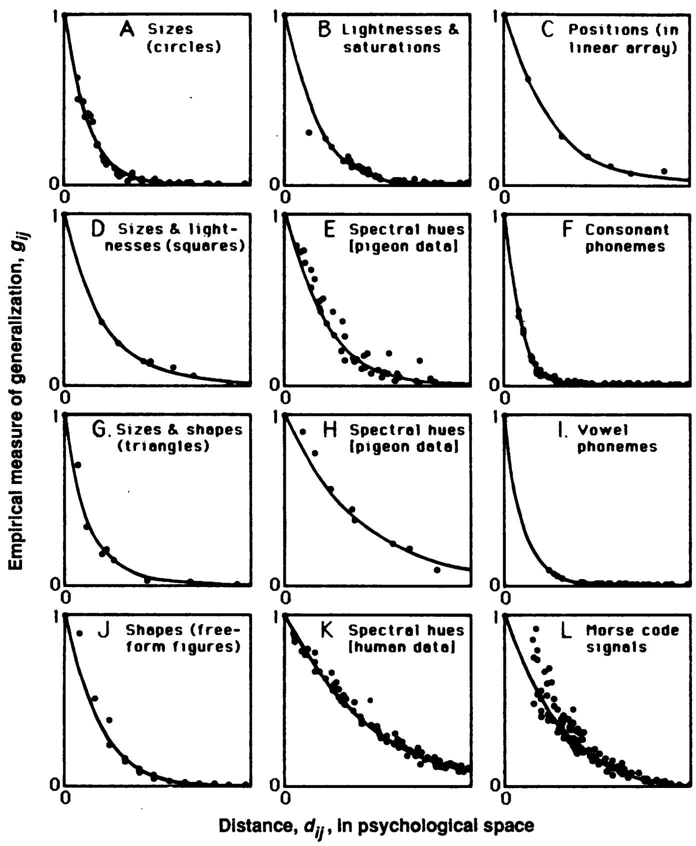

Fig. 1. Twelve gradients of generalization. Measures of generalization between stimuli are plotted against distances between corresponding points in the psychological space that renders the relation most nearly monotonic. Sources of the generalization data (g) and the distances (d) are as follows. (**A**) g, McGuire (*33*); d, Shepard (*7, 18*). (**B**) g, Shepard (*7, 17*); d, Shepard (*7, 18*). (**C**) g, Shepard (*17*); d, Shepard (*8*). (**D**) g, Attneave (*25*); d, Shepard (*8*). (**E**) g, Guttman and Kalish (*4*); d, Shepard (*11*). (**F**) g, Miller and Nicely (*34*); d, Shepard (*35*). (**G**) g, Attneave (*25*); d, Shepard (*8*). (**H**) g, Blough (*36*); d, Shepard (*11*). (**I**) g, Peterson and Barney (*37*); d, Shepard (*35*). (**J**) g and d, Shepard and Cermak (*38*). (**K**) g, Ekman (*39*); d, Shepard (*18*). (**L**) g, Rothkopf (*40*); d, Cunningham and Shepard (*41*). The generalization data in the bottom row are of a somewhat different type. [See (*39*) and the section "Limitations and Proposed Extensions."]

Instead of starting with a physical parameter space, I proposed to start with the generalization data and to ask: Is there an invariant monotonic function whose inverse will uniquely transform those data into numbers interpretable as distances in some appropriate metric space? The requirement that the resulting numbers approximate distances in a metric space breaks the circularity (*7, 15*). Thus, in a K-dimensional space, the distances between points within each subset of $K + 2$ points must satisfy definite conditions, expressible, in the Euclidean case, in terms of certain Cayley-Menger determinants (*16*). Moreover, the lower the dimensionality of the space, the stronger these constraints become. In a one-dimensional space, the distances must satisfy the following very strong additivity condition (*9, 15, 17*): For each subset of three points, the distance between the two most widely separated points equals the sum of the distances of those two points to the third point that lies between them.

The uniqueness of the function that satisfies such constraints is implicit in the following geometrical consequence of those constraints (*18, 19*): Provided that the number, n, of points in a space is not too small relative to the number of dimensions of the space, the rank order of the $n(n - 1)/2$ distances among those n points permits a close approximation to the distances themselves, up to multiplication by an arbitrary scale factor. Through Monte Carlo investigations I found that for random configurations of ten points in a two-dimensional space, distances inferred from their rank orders had an average correlation with the true distances of 0.998, and that for 45 points, the correlation exceeded 0.999,999 (*19*).

The actual determination of the unknown function (and, hence, of the associated distances) implied by a matrix of generalization data is achieved by numerical methods developed by Shepard (*18*) and Kruskal (*20*) and known as "nonmetric" multidimensional scaling. In a specified type of space, such methods move n points representing the n stimuli (usually by steepest descent) until the stationary configuration is achieved that minimizes an explicitly defined measure of departure from a monotonic relation between the generalization measures g_{ij} and the corresponding distances d_{ij}. Configurations can be obtained in spaces with different numbers of dimensions, and even with different metrics, until the most parsimonious representation is found for which the residual departure from monotonicity is acceptably small. The plot of the generalization measures g_{ij} against the distances d_{ij} between points in the resulting configuration is interpreted as the gradient of generalization. It is a psychological rather than a psychophysical function because it can be determined in the absence of any physical measurements on the stimuli.

Intimations of an Exponential Law

For a given set of n stimuli, an appropriate generalization experiment yields, for every ordered pair of these stimuli, an empirical estimate of the probability p_{ij} that a response learned to stimulus i is made to stimulus j. The multidimensional scaling method is usually applied to an $n \times n$ symmetric matrix of generalization measures, g_{ij}, obtained from such probabilities through a normalization such as $g_{ij} = [(p_{ij} \cdot p_{ji})/(p_{ii} \cdot p_{jj})]^{1/2}$, where p_{ii} and p_{jj} are

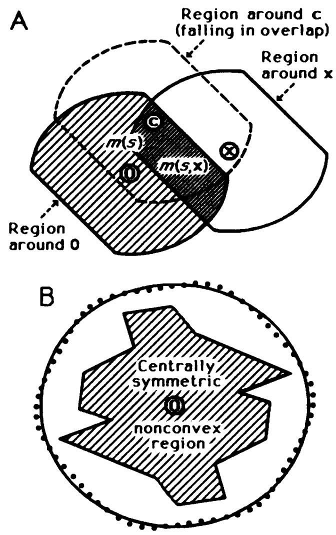

Fig. 2. (**A**) A centrally symmetric convex region shown as centered on **0**, as centered on **x**, and as having a center, **c**, falling within the intersection of the regions centered on **0** and on **x**. (**B**) For an illustrative nonconvex region centered on **0**, the locus of centers, **c**, of similarly shaped regions having a constant (approximately 20%) overlap with the region centered on **0** (dotted curve); and an ellipse corresponding to the Euclidean metric (smooth curve).

the probabilities that stimuli i and j each evoke their originally associated responses (*8, 15*).

A sample of the plots relating such generalization measures to the distances in the configurations that I obtained by applying multidimensional scaling to those measures is presented in Fig. 1. The spatial configurations themselves are presented elsewhere (*9, 11*). The data are from a number of researchers, who tested both visual and auditory stimuli, and both human and animal subjects. Yet, in every case, the decrease of generalization with psychological distance is monotonic, generally concave upward, and more or less approximates a simple exponential decay function—the smooth curve fitted to each plot solely by adjustment of its slope parameter. Moreover, I have verified that multidimensional scaling does not impose this form of the function but, by means of the assumed geometrical constraints, merely renders explicit whatever form is implicitly contained in the data (*11, 18*).

Multidimensional scaling does, however, impose monotonicity. When monotonicity was not achievable in one dimension, recourse was taken to a higher dimensional space. The increase in generalization between the red and violet ends of the visible spectrum was thus accommodated in a two-dimensional space, where the continuum of hue curves to form a circle (*9, 18*), in fact, the color circle originally described by Newton (*21*). Heightened generalization between tones separated by an octave was accommodated in a three-dimensional space, where the continuum of pitch twists into a helix (*10, 11*). And augmented generalization between all planar orientations differing by 180°, in the case of a polygon approximating central symmetry, was accommodated in a four-dimensional space, where the 360° circle of orientations deforms into the edge of a Möbius band (*12*). Only in relation to such abstract spatial representations can we achieve an invariant law.

Two Metrics for Psychological Space

When generalization data require a psychological space of more than one dimension, they also provide evidence about the metric of that psychological space (*22, 23*). For unitary stimuli, such as colors differing in perceptually integral attributes of lightness and saturation, the closest approximation to an invariant relation between generalization data and distances has uniformly been achieved in a space endowed with the familiar Euclidean metric (*17, 20, 23, 24*). For analyzable stimuli, such as shapes differing in perceptually separable attributes of size and orientation, the closest approach to invariance has generally been achieved with a different, Minkowskian metric (*23–25*), approximating what is sometimes referred to as the "city-block" metric, because distances between points in an orthogonal grid of streets conform with that non-Euclidean metric. These two metrics are also associated with what mathematicians call the L_2-norm and L_1-norm for the space. In terms of the coordinates x_{ik} (for stimulus i on dimension k) of a K-dimensional space, these metrics are obtained by setting $r = 2$ or 1, respectively, in the Minkowski power metric formula:

$$d_{ij} = \left(\sum_{k=1}^{K} |x_{ik} - x_{jk}|^r \right)^{1/r} \qquad (1)$$

In a two-dimensional coordinate space, these two metrics are distinguished by the fact that around any point, the contours of equal distance, and hence of equal generalization, are circular if $r = 2$ (the L_2-norm), and rhombic if $r = 1$ (the L_1-norm).

Are these regularities of the decay of generalization in psychological space and of the implied metric of that space reflections of no more than arbitrary design features of some terrestrial animals? Or do they have a deeper, more pervasive source? I now outline a theory of generalization based on the idea that these regularities may be evolutionary accommodations to universal properties of the world.

A Theory of Generalization

An object that is significant for an individual's survival and reproduction is never sui generis; it is always a member of a particular class—what philosophers term a "natural kind." Such a class corresponds to some region in the individual's psychological space, which I call a consequential region. I suggest that the psychophysical function that maps physical parameter space into a species' psychological space has been shaped over evolutionary history so that consequential regions for that species, although variously shaped, are not consistently elongated or flattened in particular directions.

The problem that a positive or negative encounter with an unfamiliar object poses for an individual is just the problem of inferring the consequential region to which that object belongs. A bird that ingested a caterpillar bearing particular coloration and markings and found it delectable or sickening, must decide whether another object of more or less similar visual appearance is of the same natural kind and should therefore be seized or avoided, respectively. Generalization is thus a cognitive act, not merely a failure of sensory discrimination. Indeed, an animal would be ill served by the assumption that just because it can detect a difference between the present and a previous situation, what it learned about that previous situation has no bearing on the present one.

In finding a novel stimulus to be consequential, the individual learns only that there is some consequential region that overlaps the point in psychological space corresponding to that stimulus. In accordance with whatever probabilities the individual imputes to nature, a priori, the individual can only assume that nature chose the consequential region at random. Such an individual can nevertheless obtain an estimate of the conditional probability, given that the consequential region overlaps the first point, that it also overlaps a second, by integrating over all (probabilistically weighted) possible locations, sizes, and shapes of the consequential region.

Mathematical Formulation

For the present, I suppose psychological space to be a coordinate space of some dimensionality, K. The space of objects differing only in color, for example, might be the three-dimensional space of lightness, hue, and saturation. I represent any test stimulus by the vector of its coordinates, $\mathbf{x} = (x_1, x_2, \cdots, x_K)$. If the coordinate system is chosen so that the origin corresponds to the stimulus found to be consequential, that stimulus is represented by the null vector $\mathbf{0} = (0, 0, \cdots, 0)$. I then make the following provisional specifications concerning what an individual assumes about the disposition of a consequential region in this space: (i) all locations are equally probable; (ii) the probability that the region has a size s is given by a density function $p(s)$ with a finite expectation μ; and (iii) the region is convex, of finite extension in all directions, and centrally symmetric.

Now, if the individual were to assume that the consequential region has some particular shape and, also, a particular size s, then the constraint of central symmetry entails that the set of such regions that overlap the original point $\mathbf{0}$ or the test point $\mathbf{x}$ would be just the set of such regions whose centers fall within a region of this size and shape centered on $\mathbf{0}$ or on $\mathbf{x}$, respectively. Therefore, the set of such regions that overlap both $\mathbf{0}$ and $\mathbf{x}$ would be the set of regions whose centers, $\mathbf{c}$, fall in the intersection of such regions centered on $\mathbf{0}$ and on $\mathbf{x}$ (see Fig. 2A). Because all locations of the consequential region are taken to be equally likely, the conditional probability that $\mathbf{x}$ is contained in the consequential region, given that $\mathbf{0}$ is, is just the ratio $m(s,\mathbf{x})/m(s)$ of the (volumetric) measure of the overlap to the measure of a whole such region (Fig. 2A).

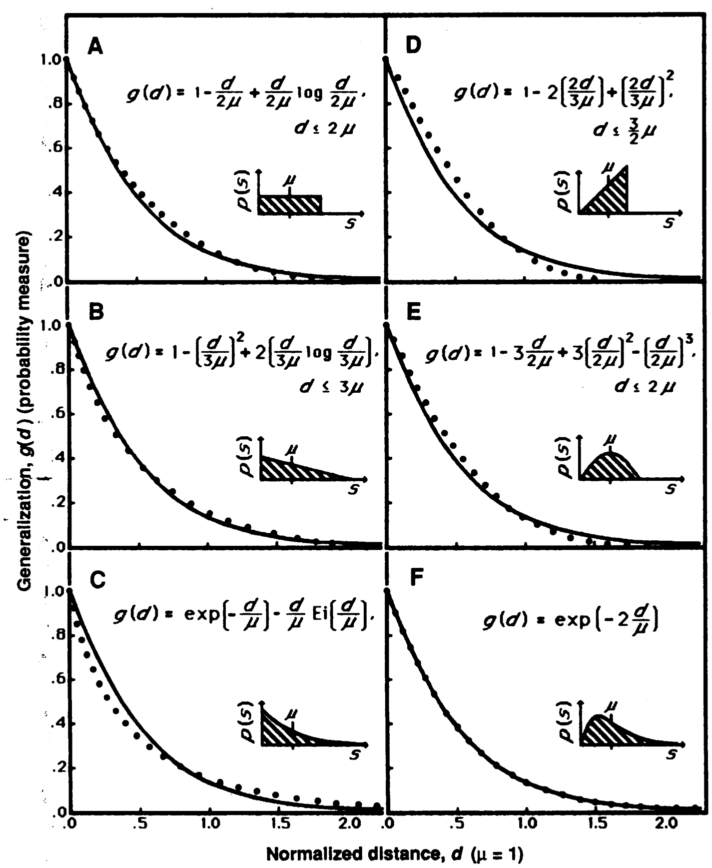

Fig. 3. Six generalization functions, $g(d)$, relating probability of generalization to normalized distance in psychological space, derived by substituting into Eq. 6 the functions $p(s)$ shown in the shaded insets, and integrating (dotted curve); and the corresponding simple exponential decay function (smooth curve). In (C), the function Ei is defined as follows

$$\mathrm{Ei}(x) = -\log x - \gamma - \sum_{n=1}^{\infty} (-1)^n \frac{x^n}{n \cdot n!}$$

By hypothesis, however, the individual does not know the size, s, of the consequential region. In order to obtain the individual's estimate of the conditional probability that $\mathbf{x}$ falls in this region, given that $\mathbf{0}$ does, the product of the ratio $m(s,\mathbf{x})/m(s)$ and the individual's corresponding a priori probability $p(s)ds$ (that the size lies between s and $s + ds$) must be integrated over all possible sizes, s. I take the result to be the probability $g(\mathbf{x})$ that a response learned to the stimulus $\mathbf{0}$ will generalize to $\mathbf{x}$

$$g(\mathbf{x}) = \int_0^\infty p(s) \frac{m(s,\mathbf{x})}{m(s)} ds \tag{2}$$

Because the size of the consequential region cannot be negative and is assumed to have finite expectation μ, $p(s)$ is zero for all $s < 0$, and (in addition to being nonnegative itself) satisfies the two conditions

$$\int_0^\infty p(s)\, ds = 1 \tag{3}$$

$$\int_0^\infty s \cdot p(s)\, ds = \mu < \infty \tag{4}$$

Derivation of the Exponential Law

In the unidimensional case, a convex consequential region is simply an interval of a certain length, $m(s) = s$, and the measure of the overlap $m(s,x)$ is then $s - |x|$, if $s \ge |x|$, or zero, if $s < |x|$. Accordingly, Eq. 2 reduces to

$$g(x) = \int_{|x|}^\infty p(s) \frac{s - |x|}{s} ds \tag{5}$$

The distance between the two stimuli $\mathbf{0}$ and $\mathbf{x}$ is now just $d = |x|$. Separating terms and successively differentiating with respect to d, we obtain, for $g(d)$ and its first and second derivatives,

$$g(d) = \int_d^\infty p(s)\, ds - d \int_d^\infty \frac{p(s)}{s} ds \tag{6}$$

$$g'(d) = -\int_d^\infty \frac{p(s)}{s} ds \tag{7}$$

$$g''(d) = \frac{p(d)}{d} \tag{8}$$

Regardless of the form assumed for the probability density function $p(s)$, then, generalization $g(d)$ has unit value at $d = 0$ (Eqs. 3 and 6), monotonically decreases with increasing d (Eq. 7), and is concave upward, unless rendered linear in those intervals, if any, where $p(d) = 0$ (Eq. 8).

The exact form for the generalization function $g(d)$ depends on the particular probability density function $p(s)$. However, a sensitivity analysis suggests that this dependence is rather weak. The dotted curves in Fig. 3 are the functions $g(d)$ obtained by integration after substituting, for $p(s)$ in Eq. 6, the six quite different density functions shown in the shaded inserts, namely, functions $p(s)$ that are rectangular (A), triangular and decreasing (B), exponential (C), triangular and increasing (D), parabolic (E), and Erlangian (F). At least for these six shapes, $g(d)$ is not only monotonic decreasing and concave upward but reasonably close to a simple exponential decay function (the smooth curve). Evidently, the form of $g(d)$ is a relatively robust consequence of the probabilistic geometry of consequential regions.

The Erlang probability density function (the shaded inset in Fig.

3F), in particular, yields exactly the exponential decay function for $g(d)$. This choice for $p(s)$ has, moreover, a unique theoretical justification: In the absence of any information to the contrary, an individual might best assume that nature selects the consequential region and the first stimulus independently. In this case, the probability that that first stimulus would fall within the consequential region is proportional to its volumetric measure $m(s)$, which, here, is simply s. According to Bayes's rule (*26*), an individual who assumed a probability density function $q(s)$ before encountering the first stimulus, should revise that function, after finding that stimulus to be consequential, to a density function $p(s) = C \cdot m(s) \cdot q(s)$. Here, C is the normalizing constant determined by Eq. 3, and $q(s)$ is assumed to be subject to the constraints already stated for $p(s)$ in Eqs. 3 and 4.

In addition, if $q(s)$ is to represent a condition of minimum knowledge about the size of the consequential region, $q(s)$ should maximize the Shannon-Wiener entropy measure of uncertainty (*27*). The function $q(s)$ that both satisfies the stated constraints and maximizes this entropy measure is an exponential probability density function (*28*) of the form displayed in the shaded inset in Fig. 3C. Substituting such a function for $q(s)$ and solving for C, we obtain for $p(s)$, in the one-dimensional case, exactly the Erlang probability density function with shape parameter 2

$$p(s) = \left(\frac{2}{\mu}\right)^2 s \cdot \exp\left(-\frac{2}{\mu}s\right) \qquad (9)$$

This is just the density function that is displayed as the shaded inset in Fig. 3F and that yielded the exponential decay for the generalization function

$$g(d) = \exp\left(-2\frac{d}{\mu}\right) \qquad (10)$$

Derivation of the Two Metrics

In the multidimensional case, the consequential region is no longer merely an interval of a certain length s. However, just as the shape assumed for the density function $p(s)$ had little effect on the derived generalization function $g(d)$, the shape assumed for the

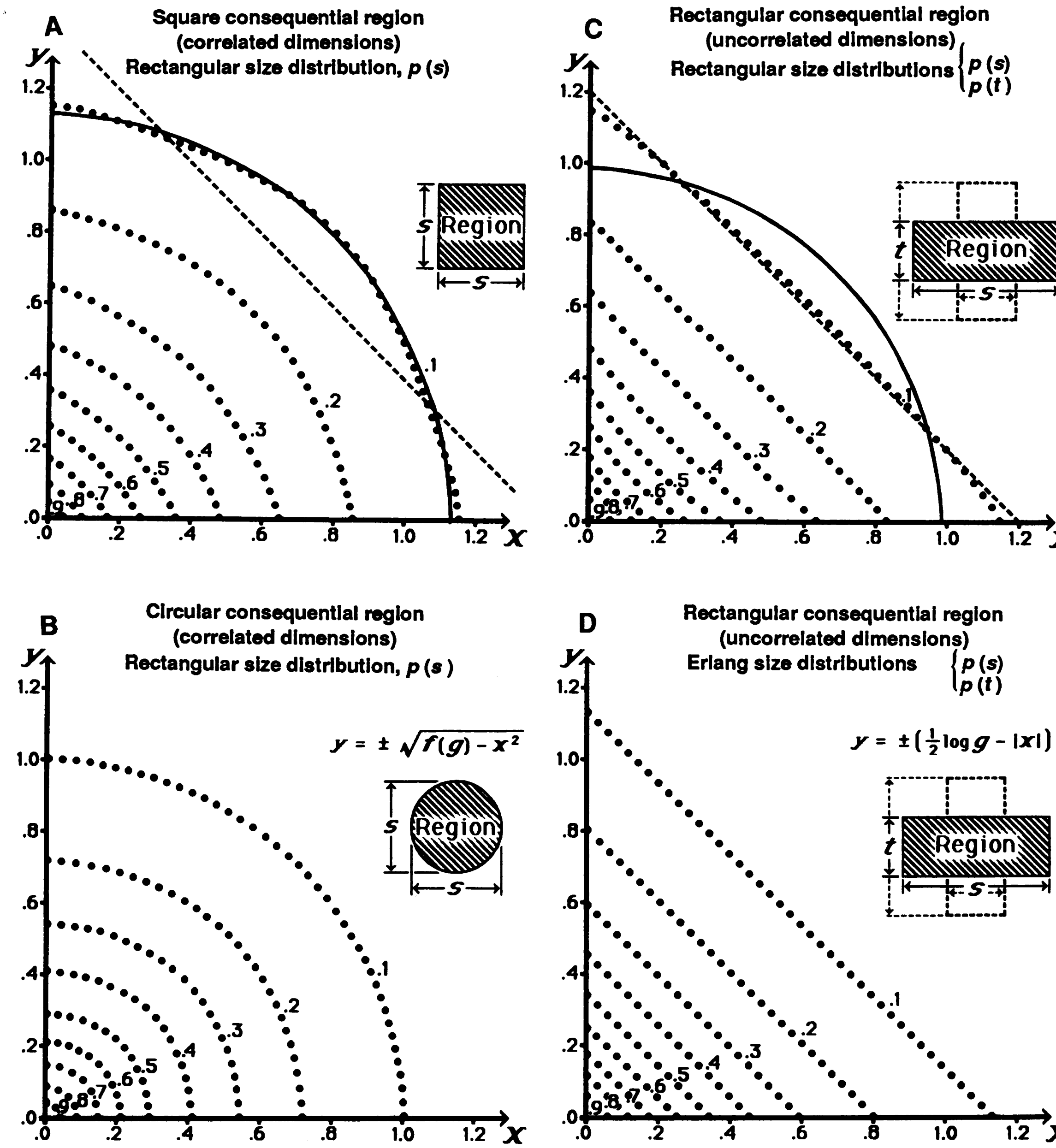

Fig. 4. Equal generalization contours plotted in one quadrant of two-dimensional psychological space. The contours on the left were derived on assumptions that the consequential region is either square (**A**) or circular (**B**) and, hence, that the extensions, s, of the consequential region in the two directions of the space are perfectly correlated. The contours on the right were derived on assumptions that the consequential region is rectangular and that its possible extensions, s and t, along the two directions are uncorrelated and have density functions, $p(s)$ and $p(t)$, that are either rectangular (**C**) or Erlangian (**D**). In all panels, equal generalization contours are dotted lines, with associated levels of generalization $g(x, y)$ indicated by adjacent numbers. In (A) and (C) L_2-norm is indicated by a solid curve; L_1-norm by a dashed line.

consequential region has (up to an affine normalization) little effect on the contours of equal generalization. This, too, is a consequence of a geometrical fact. The region can be quite irregular and even nonconvex but, as long as it is centrally symmetric, the locus of centers of such a region having a specified overlap with a given such region approximates the ellipse of the Euclidean metric (Fig. 2B).

Figure 4A shows, for one quadrant of a two-dimensional space, the contours of equal generalization around the stimulus (0, 0) that are obtained by carrying out the integration of Eq. 2 under the two assumptions (i) that the consequential region, though still of unknown size and location, has the shape of a square aligned with the coordinate axes, and (ii) that $p(s)$ is the rectangular distribution. Except for the (rhombic) contours very close to the original stimulus (0, 0), the resulting contours are more circular than square. The same is true when other density functions are substituted for $p(s)$.

For stimuli, like colors, that differ along dimensions that do not correspond to uniquely defined independent variables in the world, moreover, psychological space should have no preferred axes. The consequential region is then most reasonably assumed to be circular or, whatever other shapes may be assumed, to have all possible orientations in the space with equal probability. Symmetry then entails strictly circular contours of equal generalization (Fig. 4B) and, hence, the Euclidean metric (or L_2-norm).

For stimuli that differ along dimensions, such as size and orientation, that correspond to uniquely defined independent variables in the world, however, psychological space should possess, corresponding preferred axes. Whatever type of shape is then assumed for the consequential region, the degree to which that region is extended along one preferred axis should not be correlated with the degree to which it is extended along another such axis. Instead of assuming that the region is a square or circle, in the two-dimensional case, the individual might assume that it is a rectangle or an ellipse aligned with the preferred axes of the space. Integration must then be carried out over the two independently variable size dimensions of the consequential region, say s and t (as indicated on the right in Fig. 4), with corresponding probability density functions, $p(s)$ and $p(t)$.

As before, the curves of equal generalization depend very little on either the form chosen for these density functions or the shape chosen for the consequential region. However, in the absence of a correlation between the two size dimensions of the consequential region, the contours no longer approximate the circles associated with the L_2-norm. Instead, they approximate the rhombs associated with the L_1-norm. This is illustrated in Fig. 4C, for the assumptions that the consequential region is rectangular and that $p(s)$ and $p(t)$ are both the rectangular probability density function. Indeed, when the probability density functions $p(s)$ and $p(t)$ are taken to be the Erlang function (Eq. 9) derived from the assumption of maximum uncertainty about the two size dimensions of the consequential region, generalization falls away with distance in exact accordance both with the exponential decay function (Eq. 10) and with the metric of the L_1-norm (Fig. 4D).

Limitations and Proposed Extensions

The theory of generalization, as set forth here, strictly applies only to the highly idealized experiment in which generalization is tested immediately after a single learning trial with a novel stimulus. Existing evidence and theoretical considerations indicate that in the cases either of protracted discrimination training with highly similar stimuli (*29, 30*) or of delayed test stimuli (*8*), "noise" in the internal representation of the stimuli will manifest itself in two deviations from the functional relations derived here. The first is a deviation away from the simple exponential and toward an inflected, Gaussian function. In Fig. 1 such a deviation is evident in L, where the data (probabilities that similar stimuli were judged to be identical) do not represent generalization so much as failure of discrimination, and perhaps in E and H, where test stimuli continued to be presented long after the termination of reinforcement. The second is a deviation away from rhombic and toward elliptical curves of equal generalization, even for analyzable stimuli. To the extent that primitive organisms do not support the distinction between generalization and failure of discrimination, they too may manifest these deviations. Moreover, under the most natural extension of the present theory to multiple learning trials, differential reinforcement could shape the generalization function and contours around a particular stimulus into a wide variety of forms.

Here, space does not permit more than a brief mention of a few such directions in which I am currently extending the theory. (i) Phenomena of discrimination and classification learning, and possibly the asymmetries of generalization described by Tversky (*31*), require that over a series of trials, the probabilities that an individual associates with the alternative candidates for a consequential region are modified on the basis of the frequencies with which positive and negative stimuli fall inside or outside each such candidate region (*32*). In this connection, the assumption of sharply bounded consequential regions has the advantage that solely through this process of probability adjustment, an individual could come to discriminate stimuli that do from those that do not belong to such a sharply bounded region. (ii) Nevertheless, preliminary mathematical investigations indicate that the robust exponential function and two metrics are also derivable if the probability or magnitude of a consequence, instead of being assumed to drop off discontinuously at the boundary of a discrete consequential region, is assumed to decline gradually, in accordance with a continuous, unimodal distribution of, for example, Gaussian form but unknown location and dispersion. (iii) If the possible dispersions of the consequential region (or of the unimodal distribution) along preferred dimensions are assumed to be negatively correlated, the curves of equal generalization obtained by integration take on a concave, star-shaped form corresponding to a value $r < 1$ in Eq. 1. Such curves imply a violation of the triangle inequality for psychological distances, a violation for which Tversky and Gati have reported evidence with stimuli having highly separable dimensions (*30*). (iv) Finally, the idea of candidate regions furnishes a basis for explaining, also, a very prevalent chronometric finding, namely, that the time to discriminate between two stimuli is reciprocally (not exponentially) related to the distance between those stimuli in psychological space. We need merely suppose that a stimulus elicits internal representational events corresponding to candidate regions in accordance with probabilities, per unit time, proportional to the probabilities already defined, and that discrimination is precipitated by the first such event that corresponds to a region that includes one but not both of the two stimuli.

Conclusions

We generalize from one situation to another not because we cannot tell the difference between the two situations but because we judge that they are likely to belong to a set of situations having the same consequence. Generalization, which stems from uncertainty about the distribution of consequential stimuli in psychological space, is thus to be distinguished from failure of discrimination, which stems from uncertainty about the relative locations of individual stimuli in that space. Empirical results and theoretical derivations point toward two pervasive regularities of generalization. First, probability of generalization approximates an exponential decay function of distance in psychological space. Second, to the degree that the spreads of consequential stimuli along orthogonal dimensions of that space tend to be correlated or uncorrelated, psychological distances in that space approximate the Euclidean or non-Euclidean metrics associated, respectively, with the L_2- and L_1-norms for that space. I tentatively suggest that because these regularities reflect universal principles of natural kinds and of probabilistic geometry, natural selection may favor their increasingly close approximation in sentient organisms wherever they evolve.

Undoubtedly, psychological science had lagged behind physical science by at least 300 years. Undoubtedly, too, prediction of behavior can never attain the precision for animate that it has for celestial bodies. Yet, psychology may not be inherently limited merely to the descriptive characterization of the behaviors of particular terrestrial species. Possibly, behind the diverse behaviors of humans and animals, as behind the various motions of planets and stars, we may discern the operation of universal laws.

REFERENCES AND NOTES

1. I. Newton, *Philosophiae Naturalis Principia Mathematica* (Royal Society, London, 1687).
2. E. N. Sokolov, *Annu. Rev. Physiol.* **25**, 545 (1963).
3. I. P. Pavlov, *Conditioned Reflexes*, G. V. Anrep, Transl. and Ed. (Oxford Univ. Press, London, 1927).
4. N. Guttman and H. I. Kalish, *J. Exp. Psychol.* **51**, 79 (1956).
5. B. F. Skinner, *The Behavior of Organisms* (Century Psychology Series, Appleton-Century, New York, 1938).
6. Incidentally, this may itself be a manifestation of generalization: As several behavior theorists have noted, 0% reinforcement is more similar to occasional reinforcement than it is to 100% reinforcement.
7. R. N. Shepard, thesis, Yale University (1955).
8. ——, *Psychol. Rev.* **65**, 242 (1958).
9. ——, *Science* **210**, 390 (1980).
10. ——, *Psychol. Rev.* **89**, 305 (1982); also for early indications of octave generalization in rats, see H. R. Blackwell and H. Schlosberg [*J. Exp. Psychol.* **33**, 407 (1943)].
11. R. N. Shepard, in *Stimulus Generalization*, D. I. Mostofsky, Ed. (Stanford Univ. Press, Stanford, CA, 1965), pp. 94–110.
12. —— and J. E. Farrell, *Acta Psychol.* **59**, 103 (1985).
13. K. S. Lashley and M. Wade, *Psychol. Rev.* **53**, 72 (1946); R. R. Bush and F. Mosteller, *ibid.* **58**, 413 (1951); G. Razran, *Psychol. Bull.* **46**, 337 (1949). An exponential form had previously been postulated for the gradient of generalization by C. L. Hull, *Principles of Behavior* (Century Psychology Series, Appleton-Century, New York, 1943). However, Hull based his postulate on data, from C. I. Hovland [*J. Genet. Psychol.* **17**, 125 (1937)] that were later shown to be insufficiently reliable for a determination of shape—see K. W. Spence [*Psychol. Rev.* **44**, 430 (1937)]. More fundamentally, I claim that because Hull started with a physical scale of stimulus difference, even though he proposed to adjust it so as to equalize just-noticeable differences (and thus discriminability) along the scale, Hull forfeited the possibilities of monotonicity and of invariance.
14. A. Einstein, *Ann. Phys. (Leipzig)* **49**, 769 (1916).
15. R. N. Shepard, *Psychometrika* **22**, 325 (1957).
16. L. M. Blumenthal, *Univ. Mo. Stud.* **13** (no. 2) (1938); G. Young and A. S. Householder, *Psychometrika* **3**, 19 (1938).
17. R. N. Shepard, *J. Exp. Psychol.* **55**, 509 (1958).
18. ——, *Psychometrika* **27**, 125 (1962); *ibid.*, p. 219.
19. ——, *J. Math. Psychol.* **3**, 287 (1966).
20. J. B. Kruskal, *Psychometrika* **29**, 1 (1964); *ibid.*, p. 28.
21. I. Newton, *Opticks* (Smith & Walford, London, 1704), book 1, part 2, prop. 6.
22. W. S. Torgerson, *Theory and Methods of Scaling* (Wiley, New York, 1958); R. N. Shepard, *Psychometrika* **39**, 373 (1974).
23. R. N. Shepard, *J. Math. Psychol.* **1**, 54 (1964).
24. For a review, see W. R. Garner, *The Processing of Information and Structure* (Experimental Psychology Series, Halsted, New York, 1974).
25. F. Attneave, *Am. J. Psychol.* **63**, 516 (1950).
26. W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley Mathematical Statistics Series, Wiley, New York, 1950).
27. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948); *ibid.*, p. 623; N. Wiener, *Cybernetics* (Principles of Electrical Engineering Series, MIT Press, Cambridge, MA, 1948).
28. A. M. Kagan, Yu. V. Linnik, C. R. Rao, *Characterization Problems in Mathematical Statistics*, B. Ramachandran, Transl. (Wiley-Interscience, New York, 1973).
29. R. M. Nosofsky, *J. Exp. Psychol. Gen.* **115**, 39 (1985); R. N. Shepard, *ibid.*, p. 58.
30. A. Tversky and I. Gati, *Psychol. Rev.* **89**, 123 (1982).
31. A. Tversky, *ibid.* **84**, 327 (1977).
32. The modification of probabilities might proceed as proposed in early stimulus sampling theories: W. K. Estes, *Psychol. Rev.* **57**, 94 (1950); E. D. Neimark and W. K. Estes, Eds., *Stimulus Sampling Theory* (Holden-Day, San Francisco, 1967); or in recent connectionist theories: D. Rumelhart and J. McClelland, Eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, MA, 1986), two volumes. However, my assumption of an evolutionarily shaped psychological space confers on candidate consequential regions a structure beyond what has so far been attributed to stimulus elements or to processing nodes.
33. W. J. McGuire, *J. Exp. Psychol.* **62**, 335 (1961).
34. G. A. Miller and P. E. Nicely, *J. Acoust. Soc. Am.* **27**, 338 (1955).
35. R. N. Shepard, in *Human Communication: A Unified View*, E. E. David and P. B. Denes, Eds. (McGraw-Hill, New York, 1972), pp. 67–113.
36. D. S. Blough, *J. Exp. Anal. Behav.* **4**, 31 (1961).
37. G. E. Peterson and H. L. Barney, *J. Acoust. Soc. Am.* **24**, 175 (1952).
38. R. N. Shepard and G. W. Cermak, *Cognit. Psychol.* **4**, 351 (1973).
39. G. Ekman, *J. Psychol.* **38**, 467 (1954). (Strictly, in Fig. 1, J and K, the ordinates are not the probability that a response learned to one stimulus generalizes to a second but the probability that the second stimulus is judged to attain a certain level of similarity to the first. I suggest, however, that both probabilities arise from the same basic process.)
40. E. Z. Rothkopf, *J. Exp. Psychol.* **53**, 94 (1957).
41. J. P. Cunningham and R. N. Shepard, *J. Math. Psychol.* **11**, 335 (1974). See also R. N. Shepard, *Hum. Factors* **5**, 33 (1963).

42. Supported by National Science Foundation grants BNS 80-05517 and BNS 85-11685. I first presented the empirical results, as summarized in Fig. 1, in a presidential address to the Experimental Division of the American Psychological Association, Los Angeles, August 1981. I first presented the mathematical theory of generalization at the annual meeting of the Psychonomic Society, San Antonio, TX, November 1984. This article was drafted while I was Honorary Research Fellow in the Department of Psychology at University College London, and revised while I was the first Fowler Hamilton Visiting Research Fellow at Christ Church, Oxford. For helpful comments, I thank, particularly, D. Blough, T. Cover, J. Freyd, R. D. Luce, L. Maloney, R. Nosofsky, P. Suppes, and E. Thomas.