Measurement invariance explains the universal law of
generalization for psychological perception

Steven A. Frank*1

*Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697-2525

Edited by Günter P. Wagner, Yale University, New Haven, CT, and approved August 15, 2018 (received for review June 7, 2018)

The universal law of generalization describes how animals
discriminate between alternative sensory stimuli. On an appropriate
perceptual scale, the probability that an organism perceives two
stimuli as similar typically declines exponentially with the differ-
ence on the perceptual scale. Exceptions often follow a Gaussian
probability pattern rather than an exponential pattern. Previous
explanations have been based on underlying theoretical frame-
works such as information theory, Kolmogorov complexity, or
empirical multidimensional scaling. This article shows that the
few inevitable invariances that must apply to any reasonable
perceptual scale provide a sufficient explanation for the univer-
sal exponential law of generalization. In particular, reasonable
measurement scales of perception must be invariant to shift by
a constant value, which by itself leads to the exponential form.
Similarly, reasonable measurement scales of perception must be
invariant to multiplication, or stretch, by a constant value, which
leads to the conservation of the slope of discrimination with per-
ceptual difference. In some cases, an additional assumption about
exchangeability or rotation of underlying perceptual dimensions
leads to a Gaussian pattern of discrimination, which can be under-
stood as a special case of the more general exponential form. The
description is invariant to shift, stretch, and rotation pro-
vide a sufficient explanation for the universally observed patterns
of perceptual generalization. All of the additional assumptions
and language associated with information, complexity, and empir-
ical scaling are superfluous with regard to the broad patterns of
perception.

I show that a few inevitable invariances of any reasonable
perceptual scale determine the exponential form for the univer-
sal law of generalization in perception. All of the other details
of information, complexity, and empirical scaling are superflu-
ous with respect to understanding why the universal law of
generalization has the exponential form.

I also show that, when the separation between stimuli depends
on various underlying perceptual dimensions, it sometimes
makes sense to assume that the perceptual scale will also
obey exchangeability or rotational invariance. When that addi-
tional invariance holds, the universal law takes on the Gaussian
form, which I show to be a special case of the general exponen-
tial form.

Basic Problem and Notation

Chater and Vitányi (ref. 2, p. 346) state the law as “the prob-
ability of perceiving similarity or analogy between two items, a
and b, is a negative exponential function of the distance d(a, b)
between them in an internal psychological space.”

Let the notation $P(R_b | S_a)$ describe the probability of a pos-
itive response, $R_b$, to the event $b$, given an initial stimulus, $S_a$,
by the event $a$. A positive response expresses the perceived sim-
ilarity of $b$ to $a$, which may also be thought of as expressing the
generalization that $b$ and $a$ belong to the same category.

The goal here is to understand how the perceived similarity of $b$ to $a$, observed as $R_b | S_a$, translates into a continuous
psychological measurement scale, $T_b | a$, so that

$$P(R_b | S_a) \equiv f(T_b | a)$$

[1]

for a suitably defined mapping $R_b | S_a \rightarrow T_b | a$ and probability distribution function, $f$. We seek the characteristics of the mapping and the associated function, $f$.

Significance

When an animal is presented with two stimuli, it may consider
them similar or different. Similarity often expresses a genera-
ized notion of a category, such as two circles with different
sizes, shadings, and colors both being circles. In many studies,
perception of similarity declines exponentially with the mea-
sure of separation, a pattern often called the universal law of
generalization. This article shows that the universal exponen-
tial law can be explained by simple properties any reasonable
perceptual scale must have. A shift of the scale by a con-
stant amount, or a stretch by a constant amount, should not
change the animal’s ability to perceive generalities or differ-
ences. Those invariant measurement properties by themselves
explain why perceived generalization follows an exponential
pattern.

Author contributions: S.A.F. designed research, performed research, and wrote the paper.
The author declares no conflict of interest.
This article is a PNAS Direct Submission.
Published under the PNAS license.

1Email: safrank@uci.edu.
Invariant Properties of Measurement

There are many different suggestions in the literature for how to define a perceptual scale, \( T_{b|a} \) (1, 2, 4). I focus on the minimal properties that any reasonable measurement scale must have, rather than on detailed assumptions motivated by external theories (5–7). I express the minimal properties as simple invariances. Before listing the invariances, consider two simple examples.

First, suppose we wish to analyze the perception of temperature for event \( b \), given that event \( a \) is at the freezing point for water. If we choose to measure the temperature on the Celsius scale, then \( T_{a|0} = 0 \) and \( T_{b|0} = C \). It would make sense to assume that perceptual generalization would be identical if we assigned numerical values on a Fahrenheit scale, \( T \), which we obtain by \( T_{b|a} = 32 + 1.8 T_{b|a} \).

Second, suppose we wish to measure the perception of separation between two potentially dangerous prey items, such as noxious butterflies (4, 8). We begin by exposing a noxious butterfly, \( a \), to a predator. After the predator tastes butterfly \( a \), we then expose butterfly \( b \) to the same predator. For the exposure to \( b \), we measure the tendency for the predator to attack the potential prey item. Data may include the directions of movements relative to the butterfly, attacks per minute, or the probability of attack over repeated experiments. We now wish to find a scale, \( T_{b|a} \), that is a function of the data we have for the response to various butterflies, \( b \), relative to an initial stimulus butterfly, \( a \).

However we choose that scale, it makes sense to suppose that the information in \( T_{b|a} \) about the perceptual separation between \( b \) and \( a \) is the same as the information in \( \alpha + \beta T_{b|a} \) for some constants \( \alpha \) and \( \beta \). If that were not so, it would be equivalent to saying that the analogs of Celsius and Fahrenheit scalings would provide different information about the perceptual separation between the two butterflies.

For example, we may wish to set \( T_{a|0} = 0 \) to describe a zero separation between identical butterflies, or we may wish to let \( T_{a|0} \) express the amount of the baseline predator perception of the separation between identical stimuli. In either case, our scale \( T \) should contain the same information with respect to the probabilities of response given in Eq. 1. Here, similarity associates with the probability of avoidance response. We may also wish to express our scale standardized with respect to a unit response, \( T_{a|1} \), to \( b^* \), or with respect to a unit response, \( T_{b|1} \), to \( b^* \). The constant multiplications required to transform between units of measure should not alter the information in the perceptual scale, \( T \), about the probabilities of response.

Affine and Rotational Invariance

In other words, the way in which we measure perceptual distance between two stimuli should be independent of a shift and stretch of the scale by constant values. Formally, the scale should be shift invariant with respect to any constant, \( \alpha \), such that

\[
    f(T_{b|a}) = k_\alpha f(T_{b|a} + \alpha)
\]

for some constant of proportionality, \( k_\alpha \). The scale should also be stretch invariant to any constant, \( \beta \), such that

\[
    f(\lambda T_{b|a}) = f(\lambda \beta T_{b|a}),
\]

for which I show below that \( \lambda = \lambda_\beta / \beta \) is an invariant constant that is conserved in any particular application, set by the fact that \( 1/\lambda \) is the average value on the perceptual scale for positive responses to varying events \( b \) for a given stimulus \( a \).

Thus, the scale \( T_{b|a} \) has the property that the associated probability pattern is invariant to the affine transformation of shift and stretch, \( T_{b|a} \rightarrow \alpha + \beta T_{b|a} \). I will show that affine invariance by itself determines the exponential form for the universal law of generalization in perception.

In some cases, it makes sense to assume that the perceptual scale should also obey rotational invariance, such that the Pythagorean partition

\[
    T_{b|a} = y_1^2(\theta) + y_2^2(\theta)
\]

splits the measurement into components that add invariantly to \( T_{b|a} \) for any value of the parameter \( \theta \). The invariant quantity \( T_{b|a} \) defines a circle in the \((y_1, y_2)\) plane with a conserved radius \( R_{b|a} = \sqrt{T_{b|a}} \) that is invariant to \( \theta \), the angle of rotation around the circle, circumscribing a conserved area \( \pi R_{b|a}^2 = \pi T_{b|a} \).

Rotational invariance partitions a conserved quantity into additive components, for which the order may be exchanged without altering the invariant quantity. When rotational invariance holds, the universal law takes on a Gaussian form, which we will see to be a special case of the general exponential form.

The following sections develop the three invariances of shift, stretch, and rotation. I show that essentially all of the common properties of perceptual generalization follow from these invariances. The analysis here briefly summarizes the detailed development described in Frank (9). The novelty in this article concerns the simple understanding of widely observed psychological patterns.

Shift Invariance Implies the Exponential Form

To simplify notation, denote the perceptual scale by \( x \equiv T_{b|a} \) and the associated probability distribution by \( f(x) \equiv f(T_{b|a}) \). If we assume that the functional form for the probability distribution, \( f \), is invariant to a constant shift of the perceptual scale, \( x + \alpha \), then by the conservation of total probability

\[
    \int k_\alpha f(x) \, dx = \int k_\alpha f(x + \alpha) \, dx = 1
\]

holds for any magnitude of the shift, \( \alpha \), in which the proportionality constant, \( k_\alpha \), changes with the magnitude of the shift, \( \alpha \), independently of the value of \( x \), to satisfy the conservation of total probability.

From this equality for total probability, which holds for any shift \( \alpha \) by adjustment of the constant, \( k_\alpha \), the condition for \( x \equiv T_{b|a} \) to be a shift-invariant scale is equivalent to

\[
    f(x + \alpha) = \kappa_\alpha f(x),
\]

in which \( \kappa_\alpha \) depends only on \( \alpha \) and is independent of \( x \). Because the invariance holds for any shift, \( \alpha \), it must hold for an infinitesimal shift, \( \alpha = \epsilon \). We can write the Taylor series expansion for an infinitesimal shift as

\[
    f(x + \epsilon) = f(x) + \epsilon f'(x) = \kappa_\epsilon f(x),
\]

with \( \kappa_\epsilon = 1 - \lambda \epsilon \), because \( \epsilon \) is small and independent of \( x \), and \( \kappa_0 = 1 \). Thus,

\[
    f'(x) = -\lambda f(x)
\]

is a differential equation with solution

\[
    f(x) = ke^{\lambda x},
\]

in which \( k \) is determined by the conservation of total probability. When the perceptual scale ranges over positive values, \( x > 0 \), then \( k = \lambda \).

The assumption that a perceptual scale must be shift invariant is, by itself, sufficient to explain the exponential form of the universal law of generalization.
The Exponential Form Implies Shift Invariance

The previous section showed that if the perceptual scale, $x$, is shift invariant, then the exponential form of the universal law of generalization follows. This section shows that if the universal law of generalization takes on the exponential form, then the underlying perceptual scale must be shift invariant. Thus, shift invariance is necessary and sufficient for the exponential form. Any assumptions about the perceptual scale beyond shift invariance must be superfluous with respect to the exponential form.

Begin with the assumption of the exponential form in Eq. 7 and write the consequence of a shift of the scale $x$ by $\alpha$ as

$$f(x + \alpha) = k_\alpha e^{-\lambda(x + \alpha)} = k_\alpha e^{-\lambda \alpha} e^{-\lambda x} = ke^{-\lambda x}$$

in which $k_\alpha = ke^{-\lambda \alpha}$ because the constant multiplier of the exponential must be chosen to satisfy the conservation of total probability, in other words, to normalize the total probability to be one. Thus, the exponential form implies shift invariance of the perceptual scale, $x$.

Stretch Invariance and Rate of Perceptual Change

If we assume that the perceptual scale is defined for positive values, $x > 0$, then the average value of $\lambda x$ is always one, because

$$\int_{0}^{\infty} \lambda xf(x)dx = \lambda \int_{0}^{\infty} \lambda xe^{-\lambda x}dx = 1.$$ 

Thus, for average value, $\bar{x}$, the value of $\lambda$ is $1/\bar{x}$. We can think of $\bar{x}$ as the average discrimination of various events, $b$, relative to an initial stimulus, $a$, in which the set of events $b$ corresponds to a uniform continuum along the perceptual scale, $x$.

It makes sense to assume that the average discrimination would not change if we arbitrarily multiplied our numerical scale for perception, $x$, by a constant, $\beta$. The conservation of average value and stretch invariance are equivalent, because

$$\lambda \int_{0}^{\infty} \lambda \beta xe^{-\lambda \beta x}dx = 1$$

when we allow $\lambda _\beta$ to adjust to satisfy the conservation of average value so that $\lambda = \lambda _{\beta} \beta$ or, equivalently, we assume stretch invariance of the scale, $x = T_{\beta}a$.

The constant $\lambda = 1/\bar{x}$ can be thought of as the slope or rate of change in the logarithm of discrimination, because

$$\log f(x) = -\lambda x.$$ 

Stretch invariance, or the conservation of average value, is sufficient to set the rate of change in the logarithm of discrimination. The average value of $-\log f(x)$ is a common definition of information or entropy and is related to many interpretations in terms of information theory (4, 10).

Rotational Invariance and Gaussian Patterns

The scale, $x$, measures the perceptual difference between two entities or events. In some cases, the total difference, $x$, depends on the perceived differences along several distinct underlying dimensions. With two underlying dimensions, we may write

$$x = z_1(\theta) + z_2(\theta).$$

For a particular value of $x$, the parameter $\theta$ describes all of the combinations of the two underlying dimensions that add invariably to $x$. If we let $x = r^2$ and let the dependence of $z$ on $\theta$ be implicit, we can write the prior expression equivalently as

$$x = r^2 = \sqrt{z_1^2 + z_2^2},$$

which defines a circle with coordinates along the positive and negative values of $\sqrt{z_1^2 + z_2^2}$, with a constant radius $r$ that is rotationally invariant with respect to the parametric angle, $\theta$. Traditionally, one uses $y_i = \sqrt{z_i}$, so that the radius, $r$, of a sphere has the familiar definition of a Euclidean distance

$$r^2 = \sum y_i^2.$$ 

For each radial value, $r = \sqrt{x}$, we can write $\{y_i(\theta)\}$ as the sets indexed by the parameter $\theta$ for which the individual dimensional measures combine to the same invariant radius. If the angles of rotation with equivalent radius occur with equal probability or without prior bias, then radial values are rotationally invariant with respect to probability or prior likelihood.

I now show that rotational invariance leads to the Gaussian pattern as a special case of the general exponential form. In the exponential form derived in earlier sections, $\lambda x$ described the stretch-invariant perceptual scale. To express that scale in terms of a rotationally invariant radial measure, $r$, we note that $x = r^2$ and we let $\lambda = \pi v^2$. Thus, we can write the stretch-invariant incremental perceptual measure as

$$\lambda dx = \pi v^2 dr = 2\pi v^2 dr.$$ 

The general exponential form is

$$f(x) dx = \lambda e^{-\lambda x} dx = 2\pi v^2 re^{-\pi r^2} dr.$$ 

At a given radius, $vr$, if, by rotational invariance, all combinations of values for the underlying measurement dimensions occur without bias or prior information, then the total probability in a radial increment, $v dr$, is spread uniformly over the circumferential path with length $2\pi vr$.

A radial vector intersects a fraction of the total probability density in the circumferential path in proportion to $1/2\pi vr$. Thus, the probability along an increment $v dr$ of the radial vector is

$$\left(1/2\pi vr\right)f(x) dx = ve^{-\pi r^2} dr = ve^{-\lambda r^2} dr,$$

invariantly with respect to the angle of orientation of the radial vector. This expression is the Gaussian distribution, with $r^2$ as the squared deviation from the mean or central location and with parameters commonly written as $\lambda = 1/2\pi v^2$ and $v = 1/\sqrt{2\pi \sigma^2}$ for variance $\sigma ^2$. The variance is simply the average value of the squared radial deviations, $r^2 = \bar{x}$.

We can also write the Gaussian in terms of the standard perceptual scale, $x$, as

$$g(x) d\sqrt{x} = ve^{-\lambda x} d\sqrt{x}.$$ 

When we consider the standard perceptual scale, $x$, with respect to the incremental square-root scale, $d\sqrt{x}$, we obtain a Gaussian. The incremental square-root scale makes sense when we consider $x$ as an aggregate measure of the sum of underlying perceptual dimensions. Each dimension naturally takes on a square-root scaling relative to the invariant total distance, because of the Euclidean measure of squared distance as the sum of squares along each dimension.

Discussion

Any reasonable perceptual scale must satisfy the simple affine invariances of shift and stretch. I have shown that those invariances are sufficient to explain the exponential form of the universal law of generalization. I have also shown that an additional
common invariance of rotation explains why some observed patterns of generalization follow a Gaussian rather than an exponential pattern. The Gaussian pattern is, in fact, a special case of exponential scaling, when the scale is a squared Euclidean distance over several underlying dimensions.

Previous explanations also generate the exponential pattern of the universal law (1, 2, 4). The reason those explanations succeed is that they include assumptions about shift invariance, which by itself generates an exponential pattern. All of the other assumptions and language associated with those prior explanations are superfluous with respect to the exponential form. Conclusions about rate of change in discrimination typically associate with an assumption about stretch invariance or, equivalently, conservation of average value.

It is certainly true that additional assumptions will lead to more precise predictions, which may then be tested to rule out particular mechanisms. But those additional assumptions and tests do not directly bear on the general exponential form itself.

I do not know of explicit prior explanations that unify the Gaussian pattern with the universal exponential law. Such explanations, if they exist, will generally reduce to the assumption of rotational invariance. Again, additional assumptions or arguments about particular underlying mechanisms are superfluous with regard to the general pattern.

It is, of course, interesting to consider what underlying perceptual mechanisms lead to the universal law. However, almost certainly, there is no single mechanism that could explain such a widely observed pattern. General patterns require general explanations that apply broadly. The simple invariances of meaningful measurement scales provide that general explanation for the observed patterns of perceptual scaling.

ACKNOWLEDGMENTS. National Science Foundation Grant DEB–1251035 and the Donald Bren Foundation support my research. I completed this work while on sabbatical in the Theoretical Biology group of the Institute for Integrative Biology at Eidgenössische Technische Hochschule (ETH) Zürich.