Psychophysics

Michael H. Birnbaum
California State University at Fullerton

Glossary

Absolute judgment Task in which subject is asked to judge stimuli using "absolute" standards. Important finding is that absolute judgments are relative: they depend on the context.

Birnbaum's subtractive theory Judgments of "differences" and "ratios" are both governed by subtraction, but "differences of differences" and "ratios of differences" are governed by two operations on the same scale of sensation.

Choice task (comparison) Subject is asked to compare two stimuli and judge a psychological relation between the two (e.g., choose which weight is heavier). Dependent variables can include proportion of choices and time to make the choice. Comparison task can also involve judgments of quantitative relations between stimuli.

Context The context includes all features of the environment that affect the judgment of a stimulus besides the value of the stimulus itself. The judgment of a given stimulus depends, for example, on the set of other stimuli also presented for judgment.

Cross-modality Task in which stimuli from different psychological modalities or dimensions are to be matched, compared, or combined. For example, a subject might be asked to adjust the loudness of a tone to match the heaviness of a lifted weight.

"Difference" task Subject is asked to compare two stimuli and estimate the psychological "difference" in psychological values. (Quotation marks are used to remind the reader that "difference" judgments may or may not be governed by subtraction.)

Fechner's law Subjective value is a logarithmic function of physical value, so equal physical ratios produce equal psychological differences.

Just noticeable difference Value of smallest difference that can be detected. Usually defined as the physical change in the comparison stimulus that would be required to change the proportion of judgments that the "comparison is greater than the standard" from 0.5 to 0.75.

Magnitude estimation Task in which subject is asked to assign a number to represent the strength of sensation using an unbounded scale.

Point of subjective equality (PSE) Value of comparison stimulus that is judged greater than the standard 50% of the time.

Rating Subject assigns a number to a stimulus to represent the strength of its sensation using a numerical scale that is bounded at both ends.

Recognition Subject reports that a stimulus is the same as a previously presented stimulus.

Scaling The process of assigning measurements to psychological values, developing a numerical scale.

Signal detection theory Stimuli are represented by distributions, as in Thurstone's law. The subject will report a "signal" if the momentary sensation produced by a stimulus exceeds a criterion that is established by decision processes.

Standard Stimulus against which a comparison or variable stimulus can be judged.

Stevens' law Subjective value is a power function of physical value, so equal physical ratios produce equal psychological ratios.

Encyclopedia of Human Behavior, Volume 3

Threshold Absolute threshold is the value of smallest stimulus that can be detected. Difference threshold is the smallest difference that can be detected. The exact definition of "threshold" changes in different theories of detection and discrimination.

Thurstone's law of comparative judgment Theory that each physical value produces a distribution of subjective values. Judge reports that the comparison exceeds the standard whenever its momentary sensation exceeds the sensation produced by the standard.

Weber's law Physical size of just noticeable difference is a constant proportion of the value of the standard for a given dimension.

PSYCHOPHYSICS is the study of the relationships between the physical and the psychological worlds. Corresponding to physical weight and sound energy, for example, it is assumed that people have sensations of heaviness of weights or loudness of sounds. Psychophysics includes the study of thresholds, discrimination, recognition, and scaling. It deals with such questions as "What is the weakest stimulus that can be detected?" "What is the smallest difference in stimuli that can be discriminated?" "How does a person recognize another presentation of the same stimulus?" "Can we assign numbers to represent the strengths of psychological sensations in a meaningful way?" Psychophysical methods, theories, and empirical findings have proven valuable in many fields of psychology in which stimuli, responses, or intervening psychological constructs are measured quantitatively. Psychophysics has had important influences on the study of perception, sensation, social psychology, memory, judgment, and decision making.

I. INTRODUCTION

Psychophysics is the oldest area of psychology and probably the most controversial. Regarded as "mad" and "moonshiney" by some and as the foundation of scientific psychology by others, the field has continued to provide new and interesting problems to each generation of psychologists. Psychophysics had its origin on October 22, 1850, the morning when Gustav Fechner lay in bed contemplating problems in the philosophy of mind and body. Fechner initially thought that psychophysics would resolve age-old controversies in philosophy. However, by 1860 (when he published *Elements of Psychophysics*), Fechner realized that psychophysics could be accepted as a new field of study, independent of its philosophical interpretations.

Fechner proposed that corresponding to the physical world there is a psychological world. In the physical world, material objects can be measured in physical units, and in the psychological world, sensations can be measured in psychological units. Based on ideas of D. Bernoulli and E. H. Weber, Fechner concluded that psychological sensations (Ψ) are a logarithmic function of physical values (Φ). Fechner's mathematical interpretation of Weber's results and his derivation of the logarithmic psychophysical law are presented in simple form below.

II. WEBER'S LAW

Weber's law states that the increase in a stimulus that is just noticeably different is a constant proportion of the stimulus. In a typical experiment, illustrated in Figure 1, the judge lifts a standard weight, followed by one of several comparison weights and judges whether the comparison seems "lighter" or "heavier" than the standard. The proportion of "heavier" judgments is plotted as a function of the comparison stimulus in Figure 1. [Quotation marks are used in this article to distinguish subject's judgments from actual or theoretical statements. In this case, Figure 1 shows that when the comparison stimulus is 190 g (actually lighter than the standard), the subject judges it "heavier" than the standard 20% of the time.] The point of subjective equality (PSE) is defined as the value of the comparison that is judged "greater" than the standard 50% of the time. The difference between the standard and the PSE is defined as the time order error. In this case, there is a negative time order error because the subject tends (more than half the time) to judge the 200-g weight to be "heavier" than the standard of the same value.

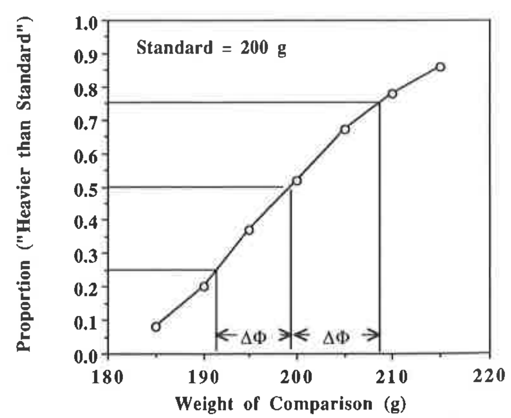

FIGURE 1 Typical results to determine a Weber fraction. The subject lifts the standard weight (200 g), followed by one of the comparison weights, and reports whether the sensation produced by the comparison exceeds that of the standard. In this case, it takes an increase ($\Delta\Phi$) of about 10 g to increase the proportion of judgments from 0.5 to 0.75, so the Weber fraction $k = 10/200 = 0.05$. If the standard were increased to 400 g, for example, Weber's law implies that the $\Delta\Phi$ would be 20 g to produce the same change in choice proportion.

The just noticeable difference (JND) is defined as the change in the stimulus required to increase the percentage of "greater" judgments from 50 to 75% (or some other arbitrary level). Weber found that this change in physical value ($\Delta\Phi$) is a constant fraction of the standard,

$$\Delta\Phi = k\Phi, \tag{1}$$

where k is the Weber constant for the continuum and the particular definition of a JND. For example, in this case the just noticeable difference at 200 g is a change of 10 g, so $k = 10/200 = 0.05$. According to Weber's law, it should take an increase of 20 g to make one JND from a 400-g standard, or an increase of 40 g to produce one JND from a standard of 800 g.

III. FECHNER'S LAW

Fechner theorized that if Eq. 1 held for any value of the standard, and if it would hold in the limits for any definition of a JND (for tiny changes in detection probability, $\Delta\Phi \rightarrow \delta\Phi$), and if all just noticeable differences are psychologically equal, then Weber's law implied:

$$k\,\delta\Psi = \delta\Phi/\Phi, \tag{2}$$

where $\delta\Phi$ is the physical differential value and $\delta\Psi$ the psychological differential. Integrating both sides,

$$\int k\,\delta\Psi = \int \delta\Phi/\Phi, \tag{3}$$

which reduces to Fechner's law:

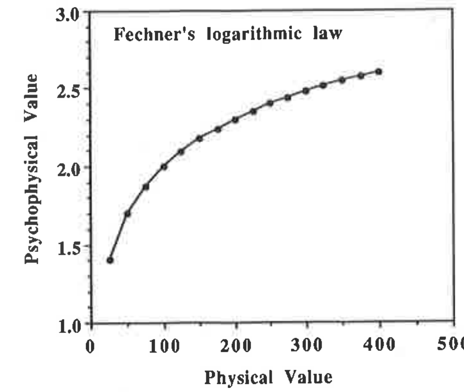

FIGURE 2 Fechner's law, illustrated for lifted weight.

$$\Psi = a \log \Phi + b, \tag{4}$$

where a and b are constants that reflect the value of the Weber fraction (k), the constant of integration, and the base of the logarithm.

A logarithmic relationship is illustrated in Figure 2, which shows Fechner's law applied to lifted weights. According to the logarithmic function, when stimuli bear the same physical ratio, the psychological differences will be equal, i.e., $\Phi_1/\Phi_2 = \Phi_3/\Phi_4$ implies $\Psi_1 - \Psi_2 = \Psi_3 - \Psi_4$. For example, $100/200 = 200/400$, so weights of 100, 200, and 400 g are equally spaced psychologically, according to Fechner's law.

IV. CATEGORY RATINGS

In the method of category rating, the subject is instructed to assign integers to stimuli to represent categories of sensation. When subjects are asked to sort the heaviness of weights into categories, category ratings of heaviness appear to fit Fechner's logarithmic law fairly well.

Fechner's law seemed to connect discrimination and judgment, because category ratings (judgments) seemed to be a linear function of the logarithm of physical value. For example, the ancient astronomer Hipparchus (c. 150 B.C.) rated the brightness of stars on a six-category scale, from stars of the first magnitude (brightest) to stars of the sixth magnitude (faintest). When modern astronomers measured the intensity of light from the stars, the scale of Hipparchus was found to be approximately logarithmically related to intensity, consistent with Fechner's law. [*See* CATEGORIZATION.]

V. LIMITS OF SENSATION

Some stars are too dim to be seen by the naked eye. Sensory psychologists have been concerned with the thresholds of sensation, values below which judgments would be random guesses. The average adult can reliably detect a candle flame from 30 miles on a clear night, he or she could reliably hear the ticking of a watch in a quiet room from 20 feet, feel the wing of a bee falling on the cheek from 1/2 in., taste a teaspoon of sugar in 2 gallons of water, or smell a drop of perfume diffused in a classroom. Psychophysical studies of thresholds have led to useful procedures for studying sensory deficiencies.

At the other extreme, there also may be an upper threshold. In summary, as one increases the physical measure of sensation below the lower threshold, psychological value does not change; above the threshold, sensation increases with increases in the physical magnitude. Finally, as the stimulus exceeds the upper threshold, becoming painful, sensations on the original dimension are unaffected by further increases.

VI. THURSTONE'S LAW

L. L. Thurstone realized that the basic idea of discrimination could be used to scale stimuli without physical measures. If each stimulus produces a normal distribution of values on a subjective, discriminal continuum, and if the subject says "stimulus i exceeds j" whenever the momentary sensation of stimulus i exceeds the momentary sensation produced by stimulus j, then the probability of this judgment, $p(i, j)$, will be given by the following equation:

$$p(i, j) = \mathbf{N}[(\Psi_i - \Psi_j)/\sigma_{ij}], \quad (5)$$

where Ψ_i and Ψ_j are the means of the distributions of subjective values on the discriminal continuum, $\mathbf{N}$ is the cumulative normal distribution function, and σ_{ij} is the standard deviation of difference between the momentary sensations. Special cases of Eq. 5 provided testable theories that allow one to construct a scale of sensation that can be checked for internal consistency and independently tested against Fechner's law.

For example, if all of the σ_{ij} are equal, we can set the value of σ_{ij} to 1. Suppose there are three stimuli (A, B, and C); suppose $p(A,B) = 0.84$ and $p(B,C) = 0.84$, then we know that the interval from A to B and from B to C are both 1 [because $\mathbf{N}(1) = 0.84$]. Therefore, the interval from A to C is 2, so $p(A,C) = 0.98$ [because $\mathbf{N}(2) = 0.98$]. As Thurstone put it, if an experimenter obtained a matrix of choice proportions and one of the entries were erased from the matrix, the missing value could be reproduced by knowing the other values and the theory.

VII. THEORY OF SIGNAL DETECTION

Signal detection theory developed from a premise similar to that in Thurstone's law, that each stimulus produces a distribution of values. The classical problem studied in the theory of signal detection is how to interpret the subject's report that a "signal" has been detected. (Quotation marks distinguish the subject's response from the actual event.) For example, the subject might be asked to listen to auditory stimuli which consist of either a burst of white noise (Noise) or a 1000-Hz tone (Signal) embedded in white noise. The subject's task is to report "signal" or "noise" to identify the two types of trials. We can classify the outcomes in a two by two table as in Table I.

Figure 3 illustrates the basic tenants of signal detection theory. The distributions represent the variability of sensation produced by the two stimuli. When the sensation exceeds the limen, shown in the figure, the subject reports that a "Signal" occurred. The probabilities of hits and false alarms are the areas under the Signal and Noise curves to the right of the limen, shown as shaded regions in Figure 3.

If each stimulus produces a normal distribution on a sensory continuum, and if there is a limen (boundary, or decision criterion) to decide whether each experience should be classified as a "Signal" or a "Noise" then we can represent the hit rate (conditional probability of saying "Signal" given a Signal was presented) and the false alarm rate (conditional probability of saying "Signal" given Noise was presented) as follows:

TABLE I
Classification of Events in Signal Detection

Actual stimulus	Response: "Noise"	Response: "Signal"
Noise	Correct rejection	False alarm
Signal	Miss	Hit

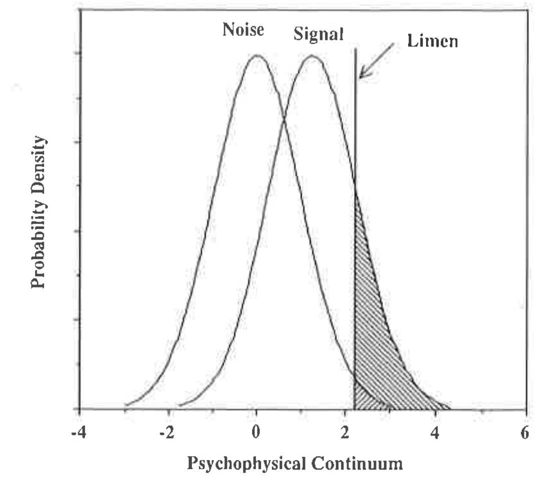

FIGURE 3 Illustration of a simple theory of signal detection. In this theory, each stimulus produces a normal distribution of subjective experiences. When the momentary experience exceeds a criterion, or limen, the subject reports the experience as a "signal." The probability of hits (saying "signal" given signal) and of false alarms (saying "signal" given the stimulus was actually noise) are shown as shaded areas to the right of the limen. Discriminability (d') is the difference between the means, divided by the standard deviation, which in this case is 1.25.

$$p(\text{"Signal"} \mid \text{Signal}) = \mathbf{N}[(\Psi_j - t_k)/\sigma_{jk}] \quad (6)$$

$$p(\text{"Signal"} \mid \text{Noise}) = \mathbf{N}[(\Psi_i - t_k)/\sigma_{ik}], \quad (7)$$

where $\mathbf{N}$ is the cumulative normal distribution, Ψ_j and Ψ_i are the mean strengths of sensation for the Signal and Noise, t_k is the limen (criterion to report "Signal" rather than "Noise"), and the standard deviations of the differences between the momentary values of the stimuli and the limen are σ_{jk} and σ_{ik}. In a simple special case, all of the standard deviations are assumed to be equal. In this case, the difference between the stimuli, $\Psi_j - \Psi_i$, in standard deviation units, is defined as d', the measure of discriminability. The criterion (or limen) is sometimes called the "bias" parameter. A scale of sensation can be constructed by cumulating successive differences between stimuli.

For example, suppose an experiment were performed in which the subject is paid \$1 for every hit and penalized \$20 for every false alarm. In this case, the judge would avoid saying "signal" because of the high cost of a false alarm. Suppose the data are as in Table II. The hit rates and false alarm rates are 0.198 and 0.018, respectively. In a table of cumulative normal probabilities, these values correspond to -2.1 and -0.85, yielding a difference, or d' of 1.25. These values correspond to the situation depicted in Figure 3. Considering the experiment and its results, the next logical question is to ask, "What would happen if the incentives for hits and false alarms were changed using the same stimuli?"

A large variety of rival detection, threshold, and choice theories have been developed to answer such questions. Some of these theories also have testable implications for decision times and the relations between decision times and response probabilities under different conditions of bias. An important graph for analyzing theories and data in signal detection experiments is a plot of hit rate against the false alarm rate obtained under various experimental conditions. This plot, called the ROC (receiver operating characteristic), could be used to compare different theories of signal detection and to examine the consequences of experimental manipulations.

Equations 6 and 7 imply that the ROC curve should be concave downward, if only the bias parameter (the limen) were affected by an experimental manipulation. Experiments that varied only the probability of a signal or the payoffs for hits and false alarms produce concave downward ROC curves that are consistent with the theory that the value of d' stays constant and that the motivational factors affected only bias. However, changes in the signal intensity affect d'. Thus, by increasing the signal intensity or by decreasing the level of noise, it is possible to increase the hit rate while holding false alarm rate constant. In contrast, changes in motivational factors such as payoffs produce correlated changes in both the probability of a hit and the probability of a false alarm.

Figure 4 illustrates two ROC curves. The higher curve illustrates the value of d' of 1.25 that is de-

TABLE II
Hypothetical Detection Results

Actual stimulus	Response: "Noise"	Response: "Signal"
Noise	0.982	0.018
Signal	0.802	0.198

Note: Each entry shows the proportion of each response, conditioned on the presentation of the stimulus.

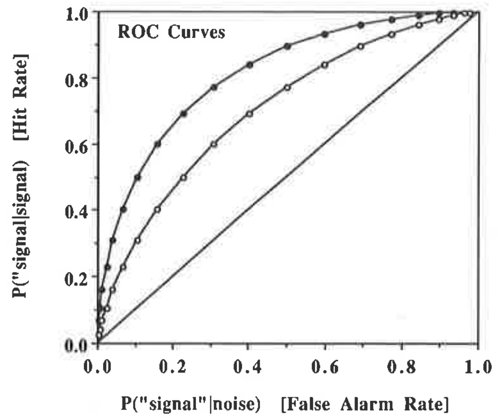

FIGURE 4 ROC curves. Filled circles show ROC curve for distributions in Figure 2 (d' is 1.25); different points represent hit rates and false alarm rates as the limen is moved across the scale. Open circles show ROC curve for a smaller value of d' (0.75).

picted in Figure 3. Points along this curve would be produced by changing the value of the category limen, t_k, in Eqs. 6 and 7, holding the value of d'(the separation between the distributions in Fig. 3) fixed. The lower curve illustrates the ROC curve for a smaller value of d'. Chance performance would fall on the identity line in Figure 4.

The study of signal detection, especially its distinction between discriminability and bias, has had important influences on basic and applied problems in psychology. The distinction is important to the analysis of recognition memory, social influences on memory, and the analysis of credibility of eyewitness testimony. Signal detection analysis has also proven very important in the study of such problems as medical diagnosis, weather forecasting, and the military problem of distinguishing "friendly" or "enemy" forces.

VIII. STEVENS' LAW

Stevens argued that the approaches of Fechner, Thurstone, and signal detection are "indirect" because they define ability to discriminate among stimuli as the basic unit of sensation. Stevens asked subjects to "directly" assign numbers to stimuli to represent the magnitudes of their sensations. In one variation of the magnitude estimation procedure, a standard stimulus is assigned a fixed number (called the modulus) and subjects are instructed to assign numbers to the comparison stimuli so that ratios of numbers would match psychological "ratios" of sensation [again, quotation marks distinguish judged "ratios" from actual or theoretical ratios].

Stevens and his associates found that magnitude estimations could be fit as power functions of physical value for a variety of continua. Stevens' power law of sensation can be written as follows:

$$ME = \alpha\Phi^{\beta}, \tag{8}$$

where ME is the average magnitude estimation, β is the exponent of the power function (which depends on the physical dimension), Φ is the physical value, and α is a constant that would depend on the modulus and standard. For example, for lifted weight, the exponent was reported to be 1.5. Thus, according to Fechner, the subjective difference in heaviness between 300 and 200 g is less than the subjective difference between 200 and 100 g, but according to Stevens, the differences in heaviness have the opposite order, getting larger as one moves up the scale. Stevens' power law is illustrated in Figure 5.

According to the power function, equal physical ratios produce equal sensation ratios, not equal sensation differences (i.e., if $\Phi_1/\Phi_2 = \Phi_3/\Phi_4$, then $\Psi_1/\Psi_2 = \Psi_3/\Psi_4$). Stevens introduced a new task, called cross-modality matching, in which the subject was instructed to report the sensation produced on one continuum by adjusting the value of a stimulus on another dimension whose sensation matched in intensity. For example, to express the heaviness of a lifted weight, the subject could adjust the loudness of a tone so that the loudness matched the heaviness. If subjective values are power functions of physical values on all continua, then cross-modality matches should be power functions of one another, with exponents predictable from the ratio of exponents determined by the method of magnitude estimation.

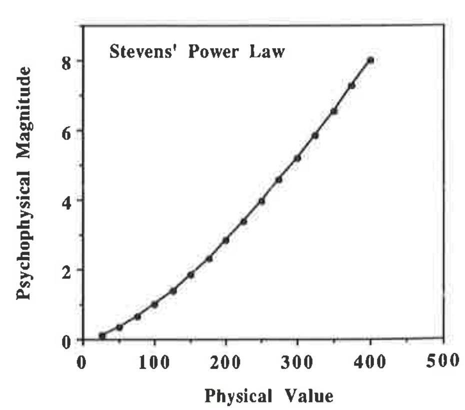

FIGURE 5 Stevens' power law, illustrated for lifted weight.

IX. CONFLICTS OF SCALES

Although cross-modality matches can be predicted from magnitude estimations, scales derived from different procedures do not agree. Compare Figures 2 and 5 for lifted weight. Because the two "laws" do not agree (Eq. 4 vs Eq. 8), Stevens said, "honor Fechner, but repeal his law." Magnitude estimations are not a linear function of category ratings, nor scales based on Thurstone's law, nor are they linearly related to scales based on the subtractive signal detection theory. Instead, magnitude estimations tend to be approximately an exponential function of category ratings. This conflict of scales caused much disagreement in psychophysical scaling.

The relationship between the two scales depends on the stimulus range, spacing, and frequency distribution and on other contextual features of the study. Some psychophysicists concluded that the conflicts of scales might be due to contextual effects, because the exact results for either category ratings or magnitude estimations could be altered by changing the procedures and stimuli used in the experiment.

X. CONTEXTUAL EFFECTS

Although magnitude estimations appear to be a power function of physical value when data are obtained with certain methods, the results can be strongly affected by the stimulus range and distribution, and the response range and distribution. The power function exponent can be predicted from the assumption that subjects match a fixed response range to the stimulus range that the experimenter will likely present. The greater the stimulus range, the lower the exponent.

Even though the subject is free to use an unbounded range of responses, the range of example responses used to illustrate the task has a powerful effect on magnitude estimations. For example, subjects who are instructed to judge the "ratio" of two stimuli are willing to call a certain "ratio" 4 : 1 when the instructions mention a "ratio of 4" as the largest example. Other subjects receive the same stimuli and instructions, except a "ratio of 64" is mentioned; these subjects are willing to judge the same "ratio" 64 : 1.

H. Helson, E. C. Poulton, A. Parducci, and others demonstrated the effects of many contextual variables and proposed theories to explain them. Their research shows that absolute judgments, category ratings, magnitude estimations, and direct judgments obtained with other procedures all depend on the context.

The same stimulus can receive different judgments, depending on the distribution of other stimuli presented for judgment. Range-frequency theory, developed by Parducci, gives a good account of contextual effects in category ratings as a compromise between the subject's tendency to make responses linearly related to the subjective value of the stimuli relative to the end stimuli and the subject's tendency to use equal portions of the response range with equal frequency.

Some investigators saw contextual effects as a threat to the establishment of measurement, whereas others considered contextual effects to be lawful properties of quantitative judgments that could themselves be used to measure psychophysical magnitudes.

XI. COMPOSITION OF FUNCTIONS

In psychophysics, it is important to make clear distinctions between the physical measures of stimuli, the physiological events produced by the physical stimuli, the corresponding psychological sensations, perception of the object that produced the sensations, and the overt responses (judgments or decisions) made by the subject. For example, a subject lifts a lead weight with a mass of 400 g, nerves and muscles fire, events occur in the brain, and the subject says the weight is "heavy."

The problem for "direct" scaling from magnitude estimations or category ratings is that the data can be represented as a composition of a psychophysical function and a judgment function, as in the following equations:

$$\Psi = \mathbf{H}(\Phi) \tag{9}$$

and

$$R = \mathbf{J}[\Psi], \tag{10}$$

where R is the overt response to stimulus Φ, $\mathbf{H}$ is the psychophysical function, and $\mathbf{J}$ is the judgment function that assigns responses to sensations. Scales based on "direct" measurement implicitly assume that the **J** function is known. Within the context of "direct" measurement, it is not possible to disentangle the functions, because responses are a composition, $R = \mathbf{J}[\mathbf{H}(\Phi)]$. Thus, the disagreement between category ratings and magnitude estimations could not be pinpointed to **H** or **J**, nor could the locus of contextual effects be determined. Some argued for category ratings and others defended magnitude estimations as the "true" measures of sensation, but the disputes could not be resolved in the framework of so-called "direct" scaling. Nevertheless, they could be resolved, if subjects could judge two different operations on a common scale of sensation.

XII. "RATIOS" AND "DIFFERENCES"

In principle, judgments of "ratios" and "differences" allow one to derive a scale of sensation that would be unique to a ratio scale, and to analyze the **H** and **J** functions, if the following equations held:

$$R_{ij} = \mathbf{J}_R[\Psi_j/\Psi_i] \tag{11}$$

$$D_{ij} = \mathbf{J}_D[\Psi_j - \Psi_i], \tag{12}$$

where R_{ij} and D_{ij} are the judgments of "ratio" and "difference" between stimuli with sensations Ψ_i and Ψ_j; $\mathbf{J}_R$ and $\mathbf{J}_D$ are strictly monotonic judgment functions for "ratios" and "differences," respectively. This two-operation theory implies that judgments of "ratios" and "differences" will not be related by a function, but instead will show particular order relations. For example, $2/1 = 4/2$ but $2 - 1 < 4 - 2$; similarly, $3 - 2 = 2 - 1$, but $3/2 < 2/1$.

TABLE III
Predictions of Two-Operation Theory of "Ratios" and "Differences"

Row stimulus	"Differences": Minuend stimulus 1	2	3	4	"Ratios": Numerator stimulus 1	2	3	4
1	0	1	2	3	1	2	3	4
2	−1	0	1	2	0.5	1	1.5	2
3	−2	−1	0	1	0.33	0.67	1	1.33
4	−3	−2	−1	0	0.25	0.5	0.75	1

Note: Entries are calculated from Eqs. (11) and (12), using successive integers for scale values. These are actual ratios and differences and therefore are not related to each other by a monotonic function.

Table III illustrates two-operation theory using identity functions for $\mathbf{J}_R$ and $\mathbf{J}_D$, and using successive integers for the values of Ψ. The left half shows differences (column values minus the row values) and the right half of the table shows ratios (column values divided by row values).

W. Garner and W. Torgerson found results that they explained as inconsistent with the theory that subjects were using two operations on a common scale. Instead, they asked, what if subjects are doing the same thing despite the experimenter's instructions? Torgerson noted that if subjects only have one way of comparing two stimuli, it might be impossible to discover what that one operation is.

XIII. ONE-OPERATION THEORY

Judgments of "differences" and "ratios" are monotonically related in many studies for different continua such as heaviness of weights and loudness of tones. In these cases, it appears that subjects use the same operation to compare stimuli, despite instructions. These results can be represented by the following equations:

$$R_{ij} = \mathbf{J}_R[\Psi_i \otimes \Psi_j] \tag{13}$$

$$D_{ij} = \mathbf{J}_D[\Psi_i \otimes \Psi_j], \tag{14}$$

where $\otimes$ represents the (single) operation for both tasks.

For a variety of continua, it has been found that judgments of "ratios" and "differences" are indeed monotonically related, as predicted by Eqs. 13 and 14, rather than showing two orders as would be predicted by the theory that subjects use both operations. The data are consistent with the theory that if the common operation is subtraction, $\mathbf{J}_R$ is exponential and $\mathbf{J}_D$ is linear; on the other hand, if the common operation is a ratio, then $\mathbf{J}_R$ is a power function and $\mathbf{J}_D$ is logarithmic.

Table IV illustrates one-operation theory under the assumptions that the operation is subtraction and $\mathbf{J}_R$ is an exponential function. For simplicity, let $D_{ij} = \Psi_j - \Psi_i$ and $R_{ij} = 2^{[\Psi_j - \Psi_i]}$. Thus, judgments of "ratios" and "differences" are monotonically related.

XIV. BIRNBAUM'S SUBTRACTIVE THEORY

M. Birnbaum theorized that when the subjective scale is inherently an interval scale, subjects use the

TABLE IV
Predictions of One-Operation Theory of "Ratios" and "Differences"

Row stimulus	"Differences": Minuend stimulus 1	2	3	4	"Ratios": Numerator stimulus 1	2	3	4
1	0	1	2	3	1	2	4	8
2	−1	0	1	2	0.5	1	2	4
3	−2	−1	0	1	0.25	0.5	1	2
4	−3	−2	−1	0	0.125	0.25	0.5	1

Note: Theory of Eqs. (13) and (14). In this case, $R_{ij} = 2^{D_{ij}}$. Thus, "ratios" are an exponential function of differences, and do not show two rank orders characteristic of actual ratios and differences. See also Figure 6.

subtractive operation for both "ratios" and "differences" but that they indeed can use two operations as instructed when the ratio operation is meaningful on the subjective scale. Intervals of stimuli form a ratio scale even when the original scales are intervals; therefore, subjects should be able to use two operations to judge "ratios of differences" and "differences of differences." For example, it may not be meaningful to judge the "ratio" of the *easterliness* of Philadelphia to that of San Francisco because the stimuli may be represented as points on a mental map, rather than as magnitudes. However, even with such a map representation, it is meaningful to ask, "what is the ratio of the *distance* from San Francisco to Philadelphia relative to the *distance* from San Francisco to Denver?" because distances between stimuli have a well-defined zero, even when the stimuli themselves are defined on an interval scale.

Experimental tests of this theory have so far confirmed two orders for these more complex tasks, even when the simple "ratio" and "difference" tasks have yielded evidence of one order. Table V shows the predicted order for the more complex, four-stimulus task.

In Birnbaum's theory, magnitude estimations form an exponential function of subjective differences when the examples are geometrically spaced and the subject is asked to judge simple "ratios" of most continua. Figure 6 illustrates the theory of the $\mathbf{J}_R$ function for magnitude estimation. In this theory, nominal "ratios" form a category scale on which a "ratio" and its nominal reciprocal are equidistant from zero. In addition, if the examples are geometrically spaced, the subject treats them as equally spaced on a category scale of differences.

Indeed, manipulation of examples appears capable of changing "ratio" judgments drastically. For example, in an experiment on prestige of occupations, the median judged "ratio" of the prestige of a *doctor* to that of a *trash collector* was either "4" or "64" depending on the range of examples in the instructions. Similar results have been obtained with lifted weight and other continua.

In summary, experiments favor the subtractive theory of psychophysical comparison, which resolves the conflict between scales by theorizing that different procedures have different **J** functions. The **J** function for category ratings is nearly linear, but will take on different forms in different contexts, as predicted by range-frequency theory, as the frequency of presentation of the stimuli and spacing of the stimuli are varied. The **J** function for magnitude estimations can take on a roughly exponential form, and this function appears to be malleable, depending on the examples given to illustrate the task as well as the range and distribution of the stimuli.

XV. PSYCHOPHYSICAL PERSPECTIVES

Although its core ideas are still under debate, psychophysics has developed as a cumulative science.

TABLE V
Predictions of Subtractive Theory of "Ratios of Differences" and "Differences of Differences"

Row difference	"Differences of differences": Minuend difference (3,2)	(3,1)	(5,2)	(5,1)	"Ratios of differences": Numerator difference (3,2)	(3,1)	(5,2)	(5,1)
(3,2)	0	1	2	3	1	2	3	4
(3,1)	−1	0	1	2	0.5	1	1.5	2
(5,2)	−2	−1	0	1	0.33	0.67	1	1.33
(5,1)	−3	−2	−1	0	0.25	0.5	0.75	1

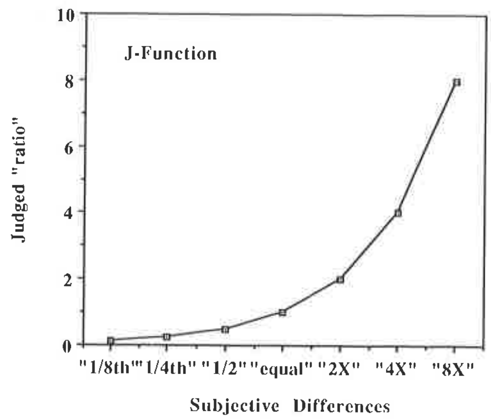

FIGURE 6 Theoretical judgment function, J_R, for magnitude estimation, including "ratio" judgment. In this theory, subjective differences are mapped into judged "ratios" by the examples, which form a category scale. If the examples are geometrically spaced, the function is an exponential function.

New data have forced old theories to be abandoned in favor of new ones.

Psychophysics seems to provide an ideal, simple psychological paradigm for investigating complex questions such as, "how can we measure a psychological value?" and "What produces a happy life?"

The techniques and theories of psychophysics are inherent in many areas of study in psychology, including those fields in which the investigator uses numbers to represent the psychological values of stimuli or uses numerical dependent variables to measure psychological reactions. People who use the results of psychophysical studies include people studying sensation, physiological mechanisms, individual differences, perception, memory, and social psychology. Ideas from signal detection have important interconnections to studies of judgment and decision making, in memory research, and in analyses of the reports of courtroom witnesses. [*See* DECISION MAKING, INDIVIDUALS; EYEWITNESS TESTIMONY; MEMORY.]

Many areas of psychology have debated whether it makes sense to discuss intervening variables that refer to internal psychological events. Modern psychophysics provides the best affirmative answer to this basic question of any field of psychology, but it will continue to remain controversial until a more complete, coherent system of measurement has been developed.

Bibliography

Birnbaum, M. H., Anderson, C. J., and Hynan, L. G. (1989). Two operations for "ratios" and "differences" of distances on the mental map. *J. Exp. Psychol. Hum. Percep. Perform.* **15,** 785–796.

Bolanowski, S. J., Jr., and Gescheider, G. A. (Eds.) (1991). "Ratio Scaling of Psychological Magnitude: In Honor of the Memory of S. S. Stevens." Erlbaum, Hillsdale, NJ.

Geissler, H.-G., Mueller, M. H., and Prinz, W. (Eds.) (1990). "Psychophysical Explorations of Mental Structures." Hogrefe & Huber, Toronto.

Gescheider, G. A. (1988). Psychophysical scaling. *Annu. Rev. Psychol.* **39,** 169–200.

Green, D. M., and Swets, J. A. (1966). "Signal Detection Theory and Psychophysics." Wiley, New York.

Guilford, J. P. (1954). "Psychometric Methods." McGraw-Hill, New York.

Link, S. W. (1992). "The Wave Theory of Difference and Similarity." Erlbaum, Hillsdale, NJ.

Luce, R. D., and Krumhansl, C. L. (1988). Measurement, scaling, and psychophysics. In "Stevens' Handbook of Experimental Psychology" (R. C. Atkinson, G. Lindzey, and R. D. Luce, Eds.), Vol. I, pp. 3–74. Wiley, New York.

Poulton, E. C. (1989). "Bias in Quantifying Judgments." Erlbaum, Hillsdale, NJ.

Stevens, S. S. (1986). "Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects." Transaction, New Brunswick, NJ.

Torgerson, W. S. (1958). "Theory and Methods of Scaling." Wiley, New York.

Wegener, B. (Ed.) (1982). "Social Attitudes and Psychophysical Measurement." Erlbaum, Hillsdale, NJ.