To Honor Fechner and Repeal His Law

A power function, not a log function, describes the operating characteristic of a sensory system.

S. S. Stevens

One hundred years ago G. T. Fechner (*1*) published the fruits and findings of a ten-year labor—an event that we celebrate as the nascence of the discipline called psychophysics. In the century since the *Elemente der Psychophysik* first made its stir, the simple but controversial logarithmic law that goes by Fechner's name has invaded almost all the textbooks that mention human reactions to stimuli. It is fitting and proper, therefore, that we should gather this year in a symposium to mark the anniversary of these beginnings and to inquire how the issues stand in 1960.

Perhaps the most insistent question on this 100th anniversary of Fechner's monumental opus is how its author could have known so much and have made such a wrong guess. (He believed that, unlike errors in general, errors in perception are independent of the perceived magnitude.) Talent, erudition, originality—each of these gifts was his in generous measure, and he applied his skills with signal success to several different domains. Not only did he create psychophysics and pioneer in experimental esthetics, but he also laid the foundations for what von Mises (*2*) later transformed into a well-known theory of probability based on the concept of a "collective."

But it was Fechner's version of the psychophysical law that really made him famous. With it he founded psychophysics and sent it off on a curious tangent—a deflection that lasted for the better part of a century. If we regard a hundred years as a long time—and it certainly seems long in the fast-moving evolution of modern science—then Fechner was almost right in his defiant forecast of 1877 (*3*): "The Tower of Babel was never finished because the workers could not reach an understanding on how they should build it; my psychophysical edifice will stand because the workers will never agree on how to tear it down."

These words were published 17 years after the appearance of the *Elemente*. By that time Fechner had had full opportunity to correct his magnum error, for at least two different arguments had by then been made in favor of what has more recently appeared to be the correct relation between the apparent magnitude of a sensation and the stimulus that causes it.

1) Brentano (*4*) had suggested that Weber's law may hold at both levels: stimulus ϕ and sensation ψ. In other words, $\Delta\phi = k\phi$, and $\Delta\psi = k\psi$. This in itself is not truth, but a simple, if illegitimate, Fechnerian integration leads from these two equations directly to the correct general form of the psychophysical law. It was Fechner himself who argued that Brentano's suggestion had to be wrong, because it would entail a power-function relation between stimulus and sensation.

2) Plateau (*5*) had suggested that, when we vary the illumination on a scene made up of different shades of gray, it is not the subjective *differences* but the subjective *ratios* that remain constant. Plateau therefore conjectured that the psychophysical relation might be a power function—a conjecture that he later renounced for a wrong reason.

Because both these suggestions led to a power law, as opposed to a logarithmic law, Fechner was inspired to write long and bitterly in his denunciation of them. It was asking too much, perhaps, to expect a professor to change his mind after two decades of devotion to an ingenious theory. Moreover, by dint of his industry and his polemics, Fechner succeeded in making the logarithmic function the sole contender, so that little or nothing was heard of the power function for many decades. If a change is now setting in, it is because new techniques have made it plain that on some two dozen sensory continua the subjective magnitude grows as a power function of the stimulus magnitude (*6*).

It is understandable that Fechner should fight stubbornly throughout his later decades to salvage his intellectual investment in the thesis that a measure of the uncertainty or variability in a sensory discrimination can be used as a unit for the scaling of the psychological continuum. He had sensed the essence of this possibility as he lay abed on that famous morning of 22 October 1850, and he had put the idea promptly and tenaciously to work. But why should such an unlikely notion have persisted for so long in other circles, and why should it have blossomed out in such noted and provocative guises as those devised by Thurstone and his school? I have puzzled so often about the ability of this fancy to persist and grow famous that I have accumulated a list of possible reasons for it. I will run quickly through a few of them, not because they are the true causes, or are exhaustive of the possibilities, but only because listing them may inspire others to inquire further into this anomaly in the history of scientific thought.

Excuses for Unitizing Error

These, then, are some of the excuses we might offer for the popularity of methods that try to create measurement by unitizing the residual noise in a stimulus-response sequence of one kind or another.

No competition. There are (or were) no competing methods. Plateau, to be sure, had asked eight artists each to paint a gray that appeared to lie midway between black and white, but this method of bisection did not catch on in a way that could challenge the arsenal of procedures aimed by Fechner at the determination of the *jnd* (just noticeable difference). When a bad theory is punctured it does not shrink quietly away. It retires from the field of sci-

The author is director of the Psychological Laboratories, Harvard University, Cambridge, Mass. This article is adapted from the text of an address delivered at a centennial symposium honoring Fechner, sponsored by the American Psychological Association and the Psychometric Society, which was held in Chicago in September 1960.

entific contest only when pushed aside by a stronger theory. Plateau shot his perceptive shaft into the Fechnerian blimp by asserting that his bisection results entailed a power law rather than a logarithmic law, but he then turned the experimental attack over to his friend Delboeuf (*7*), who, for reasons not entirely clear, proceeded to obtain bisection data that approximated the Fechnerian logarithmic prediction. Plateau thereupon reversed his view and turned to other topics, leaving Fechner to patch up his defenses and carry the day. If Plateau, whom we remember well for his part in establishing the Talbot-Plateau law relating to the brightness of intermittent light, had fought back at this juncture instead of capitulating, the story might then have taken a happier turn. Plateau's method of bisection may not be the best of procedures, but it could have carried him a long way toward the establishment of the correct psychophysical law. If that had happened, psychophysics would have been spared a hundred years of futility.

At this point let me interrupt the story and try to justify my belief that a man of Plateau's great talent could have made an effective stand against Fechner by perfecting the method of bisection and using it to produce a sufficient body of experimental results. It is true that the method of bisection, like all methods that attempt to partition the distance between a pair of stimuli, contains systematic biases that prevent its generating a linear segment of the scale of sensation. Applied to sensory intensity—the so-called prothetic continua—these methods produce "partition scales" whose curvature is always downward when they are plotted against the ratio scale of subjective magnitude (*8*). Nevertheless, as Plateau had surmised, these procedures are capable of verifying the power law. The estimate they give of the exponent is generally too small, but often the discrepancy is only 10 to 15 percent.

A series of brightness bisections that R. J. Herrnstein and I obtained in 1953 are shown in Fig. 1. We were wondering whether brightness bisections would exhibit the hysteresis that manifests itself in other sense modalities when the order of presentation of the stimuli is reversed. The answer is yes (*9*), but more important to our immediate concern are three other features of the outcome. (i) The curvature of the functions in the semi-log plot of Fig. 1 demonstrates that the

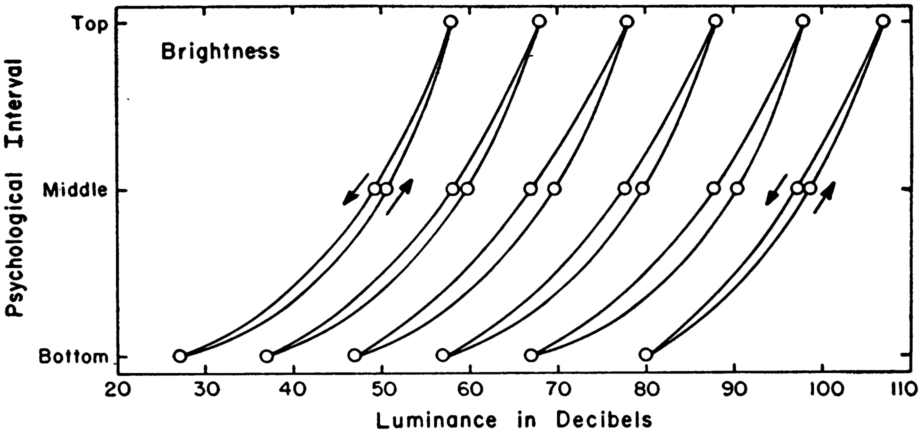

Fig. 1. Bisections of brightness intervals, showing the hysteresis effect produced by the order of stimulus presentation. The abscissa shows the luminances in decibels relative to 10^{-10} lambert. The apparent midpoint is at a higher luminance when the luminances are viewed in ascending order. Each bisection was made by from 14 to 19 observers, each of whom repeated the task three times, in both ascending and descending order. Thus, each point is based on from 45 to 57 bisections. The bisection values obtained by averaging the ascending and descending series are consistent with a power function having an exponent slightly under 0.3.

relation between sensation and stimulus is not logarithmic—that is, the bisection does not occur at the geometric mean. (ii) The similarity in the curvature obtained over a wide range of absolute stimulus levels is consistent with the power function. (iii) The bisection values obtained by averaging the ascending and descending series determine that the power function has an exponent of about 0.3. This compares with the slightly larger value, 0.33, typically found in experiments on the ratio scaling of the brightness of luminous targets seen in the dark.

Admittedly, modern apparatus makes the experiment easier now, but it is difficult to see how any ghost other than the *Zeitgeist* could have prevented the establishment of the power law back in the last century if Plateau had followed through. Merkel (*10*), incidentally, had the necessary apparatus all set up in 1888, and his limited results on bisections agree well with those in Fig. 1. But no one paid much attention to Merkel. He is the one who gave us the method of ratio production, "stimulus doubling," as he called it. With this new procedure he was fully equipped to rescue psychophysics, but neither he nor anyone else seemed to know it. Oh, to have the perceptual resolving power of hindsight!

The universality of variability. The unitizing of "noise" is attractive because confusion and variability, like death and taxes, are always with us. Normally a nuisance to science, dispersion among people's judgments becomes grist for the mill when Fechner makes dispersion into a "difference limen" and calls it the unit of his scale; and it keeps the mill wheel whirling when Thurstone enters dispersion into his "equation of comparative judgment" and computes scale values for the stimuli. While preserving the "logic," as he called it, of Fechner's procedure, Thurstone ventured to "extend the psychophysical methods to interesting stimuli" (*11*). Among other spirited questions, he asked, "which of two nationalities would you prefer to marry?" which to some people is a livelier issue than any possible question about lifted weights.

Here I must digress to modify the foregoing assertion regarding the universal presence of confusion and variability. Dispersion among data is always with us, to be sure, but sometimes it is not there in sufficient bulk for the smooth working of the Thurstonian transformations. Mischievous as it may seem, investigators who use these methods sometimes feel called upon to seek noisier data than those they happen to have gathered. The premium normally placed upon precision, repeatability, and the elimination of perturbations may turn into a liability when the perturbations themselves provide the alleged unit of measurement.

Wide applicability. The universality of noise imparts an equal universality to the Fechner-Thurstone scales. They can be erected with and without the knowledge of an underlying stimulus metric. Fechner scaled subjective weight by *jnd*'s and determined esthetic judgment by paired comparisons. Thurstone improved the machinery for handling the data of paired comparisons and made it applicable (given sufficient noise) to such elusive matters as attitudes, preferences, and the goodness of handwriting.

Wide applicability is a desirable trait, but it does not qualify as the decisive attribute in determining the merits of a scaling technique. For example, the operation of empirical addition, as practiced in the scaling of some of the "fundamental" magnitudes of physics, especially length and weight, has a very limited applicability but an overwhelming importance. One is led to suspect that universality may be less important than it seems at first sight.

It should be remarked that some of the modern psychophysical techniques—magnitude estimation, for example—can also be applied to stimuli for which there is no underlying metric. We have recently scaled the apparent roughness of sandpapers on both a ratio and a category scale and have demonstrated that roughness as felt by the finger is a prothetic continuum—all this despite our having no metric scale of the stimulus involved. In order to erect a psychological ratio scale, the experimenter needs only a nominal scale of his stimuli—he must be able to keep track of which is which.

The experiment on roughness was carried out by Judith Rich and Irma Silverman as a laboratory exercise. They presented 12 grades of sandpaper (nominally 24 to 320 grit) to the observer, who made two sweeps with his first and second fingers over each paper. The papers were presented twice each, in an irregular order, to each of 12 observers. For the magnitude estimations, a paper of medium roughness was presented first and called 10. The observer then assigned numbers proportional to the apparent roughness of each of the sandpapers as he felt it. For the category judgments, the smoothest paper was presented and called 1, and the roughest was presented and called 7. The observer then judged each paper twice, in irregular order, on a 7-point scale. Again there were 12 observers.

As shown in Fig. 2, the partition scale that results when the observers try to divide roughness into seven equally spaced categories is nonlinearly related to the scale erected by direct magnitude estimation. This outcome is the standard finding on prothetic continua. It is as though the observer, when he tries to partition a prothetic continuum into equal intervals, finds himself biased by the fact that a given difference at the low end of the scale is more noticeable and impressive than the same difference at the high end of the scale. This asymmetry is not present on metathetic continua (for example, pitch and apparent azimuth), and there the category scale is not systematically curved.

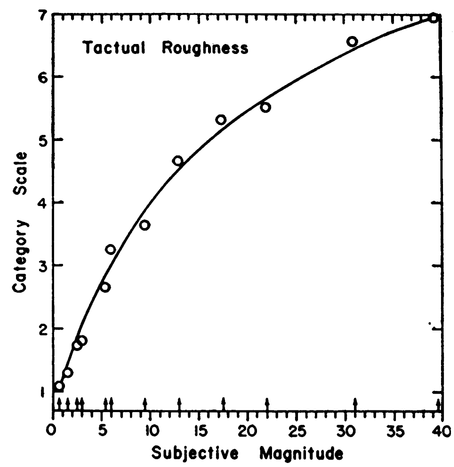

Fig. 2. Category and ratio judgments of the roughness of sandpaper. The arithmetic means of the category judgments (ordinate) are plotted against the scale derived from the geometric means of the magnitude estimations (abscissa). The marks along the abscissa show the locations of the sandpaper stimuli on the linear scale of subjective roughness.

The scaling of roughness should make it plain that the methods used in "the new psychophysics" are *not* restricted to psychological continua for which a stimulus metric is known at the level of an interval or a ratio scale —the scales that we commonly call quantitative (*12*). If further evidence is needed, look, for example, at Ekman's application of ratio scaling procedures to the esthetic value of handwriting (*13*). He cites evidence that the Thurstonian scale is nonlinearly related to the ratio scale in the way that one would expect for a prothetic continuum. Ekman clings to the view that there is still a use for the "indirect" methods that unitize confusion, because there may be psychological variables that cannot be directly observed. Whatever turn this particular argument may take, there is no denying that the challenge is out. Is there any substantive problem relating to the assessment of a subjective variable whose solution cannot be reached by direct ratio scaling procedures?

Easy on the observer. An argument for the variability methods is that they place minimal requirements on the person making the judgment. His task is easy; all he needs to be is variable. As a more specific instance of the argument for the use of a measure of dispersion as the unit of the subjective scale, we find Garner (*14*) saying that these methods are "more legitimate, valid, and meaningful for the scaling of loudness than are those methods which make use of various types of direct response on the part of the observer." Another specific merit claimed for these "discriminability" procedures is the stability of the results.

If it were true (and this we may well dispute) that *jnd* scales show more stability than ratio scales of subjective magnitude, we could still with justice ask that rude but pointed question: So what? It is conceivable that the noise ordinarily encountered when the observer responds to pairs of stimuli is much the same in magnitude from experiment to experiment, but so also may be the observer's temperature.

Relevance may be more crucial than precision.

Model-maker's delight. Scaling by the unitizing of variability has the further advantage that it poses a challenge to our ingenuity and allows us to invoke elegant formal models in one or another aspect of the enterprise. In this respect it contrasts with those prosaic procedures of measurement that offer only the computation of a median or a geometric mean as the terminal ritual. As soon as it is decided that a measure of dispersion can be used for something more than the measurement of dispersion, new vistas open, and the model builders proceed to devise ingenious exercises in matters ranging from axiomatics to the programming of computers. All this ferment is interesting and good. It would be even better if it stood on some firmer base than disagreement among human judgments.

At the risk of giving aid and comfort to the Thurstonians, I have elsewhere (*15*) suggested that a closer approximation to the correct scale of subjective magnitude on a prothetic continuum would be achieved if it were assumed that discriminal (that is, subjective) dispersion is *proportional* to the psychological magnitude. The subjective dispersion is usually assumed to be constant, an assumption known as the Thurstonian "case V," which is essentially the same as the Fechnerian assumption that each *jnd* unit represents the same subjective difference on the psychological continuum. Fortunately for a sane approach to psychological measurement, the proportionality assumption, which I have suggested might be called "case VI" (*16*), turns out to have two defects. First, it is not quite true that variability in the subjective response is always proportional to subjective magnitude, and second, even if this were a good assumption, the resulting Thurstonian scale would turn out to be only a logarithmic interval scale. Psychological values separated by equal units of dispersion would not stand equidistant from one another; they would stand in a constant ratio to one another. But no one seems yet to have discovered an interesting use for a logarithmic interval scale—except to christen and describe it, and to point out that it remains invariant under the power group of transformations (*12*).

Prominent among the models contrived for the mirroring of human variability is the development known as detection theory (see *17*). Here there seems to be no thought of measuring sensations or other subjective magnitudes; the question at issue is the narrow problem of the rules that govern the ability of a person to detect a signal immersed in a noise. Since analogous problems have been addressed by engineers and mathematicians in other domains, an elaborate model incorporating certain aspects of statistical decision theory stands ready for application to the psychophysical case. Actually, the model is perhaps not quite so ready as I seem to imply, for arguments are in progress regarding the resemblance between the nature of the noise assumed by the model and the properties of the noise that might reasonably be assumed to reside in the observer (see *18*). In any case, we have here an instance in which the contentions regarding the model and its applicability serve to generate a sparkling interest in a problem that has little substantive flesh on its bones.

Pseudo differential equations. In some quarters the popularity of the methods that try to create measurement out of variability has been sustained by a misidentification. The standard deviations and quartile points of frequency distributions of judgments have become identified with differentials. The custom is to write ΔI when what we really mean is the scatter of some dial settings or the relative frequencies of some confusions. It is argued implicitly by some, explicitly by others, that ΔI is determined by the slope of the "operating characteristic" of the sensory system and therefore that it is as useful to psychology as differential equations are useful to physics. Here is what we read in a paper by Nutting (*19*). "As is well known, the visual *sensation* cannot be directly measured, but its derivative *sensibility* is readily measurable, and from this the sensation may be readily deduced just as the scale of an ammeter may be readily reconstructed if the sensibility is known for all currents."

Admittedly, that was an engineer talking, but I wish I could say that no modern psychophysicist holds similar views. Many recent arguments have turned on the question of whether it is possible for the psychophysical power function to vary in steepness (exponent) without a concomitant change in the resolving power. Some people claim to enjoy a compelling intuition to the effect that, when the psychophysical function gets steeper, the resolving power must get better. This, I suppose, is a tacit admission that the person in question has not made the proper distinction between a differential and a standard deviation.

Let us instruct ourselves on this issue by a brief sample of facts. The sensation of electric shock grows rapidly with intensity (exponent = 3.5), but the resolving power on this continuum seems to be little if at all better than in the vibration sense, where the exponent is several times smaller (*20*). The measured resolving power for differences in brightness changes by about 50-fold when the area of the stimulus changes from a smallish point to a wide field (see Hunt, *21*), but no comparable change occurs in the exponent of the brightness function (*9*). Actually, the brightness function is slightly steeper for point sources (exponent = 0.5) than for larger targets (exponent = 0.33), but the difference would be in the opposite direction if a smaller *jnd* went with a steeper function. In the auditory sense modality the *jnd* for intensity can be altered merely by changing the frequency of the tone. As Newman (*22*) pointed out, the difference in the resolving power at 80 and at 1000 cycles per second is such that the subjective size of the *jnd* cannot possibly be constant for all pure tones. These examples are just a few of the more dramatic instances in which variability, in the guise of the *jnd*, has failed to behave as a proper unit of subjective measurement.

Nevertheless, the number of man-hours devoted each year to the measurement of one or another aspect of sensitivity, sensibility, resolving power, detectability, or just plain *jnd* attests a persistent belief in the utility of these measures. By and large, they are remarkably tedious measurements; yet the cumulative curves in which they are displayed merely gauge the noise that happens to characterize the system under the circumstance chosen for the experiment. If the noise could somehow be reduced, the measured variability would grow smaller, but there is no reason to expect that the subjective magnitude, represented by some average value of the distribution, would be thereby altered. It is not required that the mean and the standard deviation vary together.

If so much trouble has resulted from the identification of ΔI with a measure of the slope of the magnitude function, a simple remedy suggests itself. Perhaps, if we stop writing ΔI when we mean a measure of dispersion on a frequency distribution, the confusion may get itself cleared up in a generation or two. The most common measure of resolving power in sensory psychophysics is the median deviation—the 75-percent point on the cumulative "psychometric function" relating proportion of correct judgments to stimulus difference. Since this value is sometimes called the quartile point, why not replace the abbreviation ΔI by the abbreviation Q? We could then write the general linear form of Weber's law as $Q = k(I + I_0)$, and in the process we would discourage the view that Q's are something that can profitably be added up, or that can be regarded as indicative of the slope of the function relating sensation to stimulus.

The Question of the Neural Quantum

The ubiquitous variability of the human response has not only provided a tempting basis on which to build a deceptive theory of psychological measurement, it has also obscured the inner workings of the discriminatory mechanism. Under most procedures used to measure the *jnd*, some variety of noise sets bounds on the observer's resolving

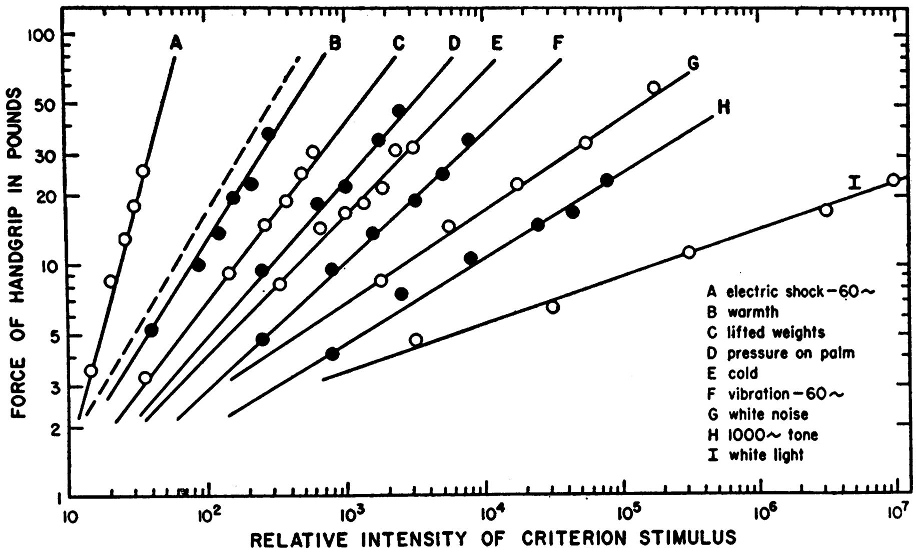

Fig. 3. Nine equal-sensation functions, obtained by matching force of handgrip to various criterion stimuli. The relative positioning of the functions along the abscissa is arbitrary. The dashed line shows a slope of 1.0 in these coordinates. Each point stands for the median force exerted to match a criterion stimulus. Ten or more observers participated in each of the nine experiments.

power and thereby determines the measured size of the resulting error distribution—the *Q*. What would happen if, by some contrivance, we could rid the experiment of noise, or at least of much of it? No doubt, if sufficient "quiet" could be achieved, the "grain" and discontinuities in the action of the sensory system would begin to manifest themselves, for it is highly unlikely that the neural processes that mediate discrimination are devoid of some all-or-none property. As a matter of fact, direct evidence for an all-or-none step function in the action of sensory discrimination has been observed in several experiments that have been undertaken to perfect the experimental arrangement to the point of a sufficient suppression of variability (for a recent review, see *23*). The success of some of the experiments that have sought evidence for a "neural quantum" (NQ) suggests that valuable findings may attend the suppression of noise and variability. It may prove better to struggle for the reduction of uncertainty than to try to enshrine it as the measure of psychological magnitudes.

The Psychophysical Law

All the evidence that was heaped together in previous decades for the purpose of refuting the Fechnerian dogma made no great hole in the "psychophysical edifice." Even the eloquent ridicule by William James, who went so far as to mix his metaphor in "striking Fechner's theories hip and thigh and leaving not a stick of them standing," could do little to change the content of the textbook discussions of "Fechner's Law." The task of clearing

Table 1. Representative exponents of the power functions relating psychological magnitude to stimulus magnitude on prothetic continua.

Continuum	Exponent	Conditions
Loudness	0.6	Binaural
Loudness	0.54	Monaural
Brightness	0.33	5° Target, dark-adapted eye
Brightness	0.5	Point source, dark-adapted eye
Lightness	1.2	Reflectance of gray papers
Smell	0.55	Coffee odor
Smell	0.6	Heptane
Taste	0.8	Saccharine
Taste	1.3	Sucrose
Taste	1.3	Salt
Temperature	1.0	Cold, on arm
Temperature	1.6	Warm, on arm
Vibration	0.95	60 cy/sec, on finger
Vibration	0.6	250 cy/sec, on finger
Duration	1.1	White-noise stimulus
Repetition rate	1.0	Light, sound, touch, shocks
Finger span	1.3	Thickness of wood blocks
Pressure on palm	1.1	Static force on skin
Heaviness	1.45	Lifted weights
Force of handgrip	1.7	Precision hand dynamometer
Autophonic response	1.1	Sound pressure of vocalization
Electric shock	3.5	60 cy/sec through fingers

the scientific bench top of the century-long preoccupation with the *jnd*, and the consequent belief in logarithmic functions, demands the cleansing power of a superior replacement. My optimism on this score has been recorded in other places, but I would like here to suggest that, if I seem to feel a measure of enthusiasm for the power law relating sensation magnitude to stimulus intensity, it is only because that law seems to me to exhibit some highly desirable features. Not the least of these desiderata is its apparent generality. On more than a score of sensory continua, the subjective magnitude ψ has been shown to grow as the stimulus magnitude ϕ raised to a power n. More specifically,

$$\psi = k(\phi - \phi_0)^n$$

where ϕ_0 is the effective threshold. As yet we have encountered no exception to this rule. Some of the exponents, a few of them not yet very firmly determined, are listed in Table 1. If the psychophysicists complete the sweep of the sensory domain and recover power functions at every turn, we may anticipate that little room will remain for Fechner's logarithmic relation. Perhaps Luce's (*24*) penetrating analysis has left no room for it anyhow.

William James (*25*) once addressed himself to Fechner's *Massformel*, $S = c \log R$, by pulling himself up to a towering indignation and letting go with: "No human being, in any investigation into which sensations entered, has ever used the numbers computed in this or any other way in order to test a theory or reach a new result."

Whether James was precisely correct in this stricture is beside the point. What he was invoking was the pragmatic test by which all scientific principles must be judged—including my own candidate, the power law. That brings us to the pay-off question: In what ways has the new psychophysical law done better than the old one in serving a scientific purpose?

Applications and Validations

As positive answers to this question, four examples can be marshaled, each concerned with a different asset of the new approach to psychophysics.

1) We have already noted that the power function governing the growth of subjective brightness (exponent = 0.33) predicts, with only a minor systematic error, the behavior of an ob-

server who undertakes to bisect the interval between two levels of luminance (see Fig. 1). A similar story can be told for loudness (*26*). One success of the new psychophysics has been the pulling together of many loose ends in a way that discloses consistencies and uniformities where none were apparent before. The nature of the biases in partition scales and the relation of these scales to ratio scales are fast becoming clear. Not only that, but the relation between subjective magnitude and the scale that is generated by Thurstone's method of successive intervals, which "unitizes" the confusions among category judgments, has yielded to orderly analysis by Galanter and Messick (*27*). Since the noise and confusion in judgments of loudness tend to grow in direct proportion to the subjective magnitude, it is not surprising that the confusion scale generated by discarding the mean and processing the variability turns out to resemble a logarithmic transform of the ratio scale of loudness.

2) In some ways the most dramatic validation of the scales generated by asking observers to make numerical estimations of sensory intensity is the demonstration that these same scales can be generated even if no appeal is made to "number behavior" at all. By means of cross-modality comparisons, each subjective continuum can be related to each other continuum, and, for the critic who thinks he will feel better if all reference to numerical judgments is avoided, the family of power functions governing the various sensory continua can all be assigned their appropriate exponents relative to that of some "base continuum," such as apparent length of lines. In practice, of course, we have been content to go along with results of the several procedures involving numerical methods, because these findings have stood the test of cross-modality validation. The argument runs as follows.

If, given an appropriate choice of units, two modalities are governed by the equations

$$\psi_1 = \phi_1^m$$

$$\psi_2 = \phi_2^n$$

and if the subjective values ψ_1 and ψ_2 are equated by asking the observer to make the one sensation seem as strong as the other at various levels, then the resulting equal-sensation function will be given by

$$\phi_1^m = \phi_2^n$$

In terms of logarithms

$$\log \phi_1 = \frac{n}{m}(\log \phi_2)$$

In log-log coordinates, therefore, the equal sensation function should be a straight line with a slope equal to the ratio of the two exponents.

This prediction was nicely borne out by a series of cross-modality matches between all possible pairs of the three continua, loudness, vibration on the finger tip, and electric shock to the fingers (*20*). From this encouraging beginning, the procedure of cross-modality matching has been extended to numerous other pairs, with special emphasis on what might be called scaling by squeezing.

Using a precision dynamometer, J. C. Stevens and Mack (*28*) worked out the subjective scale relating the apparent force of handgrip to the physical force exerted by the subject. This relation turned out to be a power function with an exponent of 1.7. Equipped with this scale, we then proceeded to take the measure of other sensory continua by asking observers to squeeze the dynamometer until the sensation of strain matched the apparent intensity of a criterion sensation in some other modality (*29*). A sample of the results is shown in Fig. 3, where two important facts stand out. All the data approximate straight lines in the log-log plot, and the slopes stand in the same order as the respective exponents listed in Table 1. Less obvious but even more crucial are the exact values of the slopes in Fig. 3. If these values are multiplied by the factor 1.7, the products agree reasonably closely with the values of the exponents listed in Table 1 (*6*).

In another investigation, the cross-modality comparison of loudness and

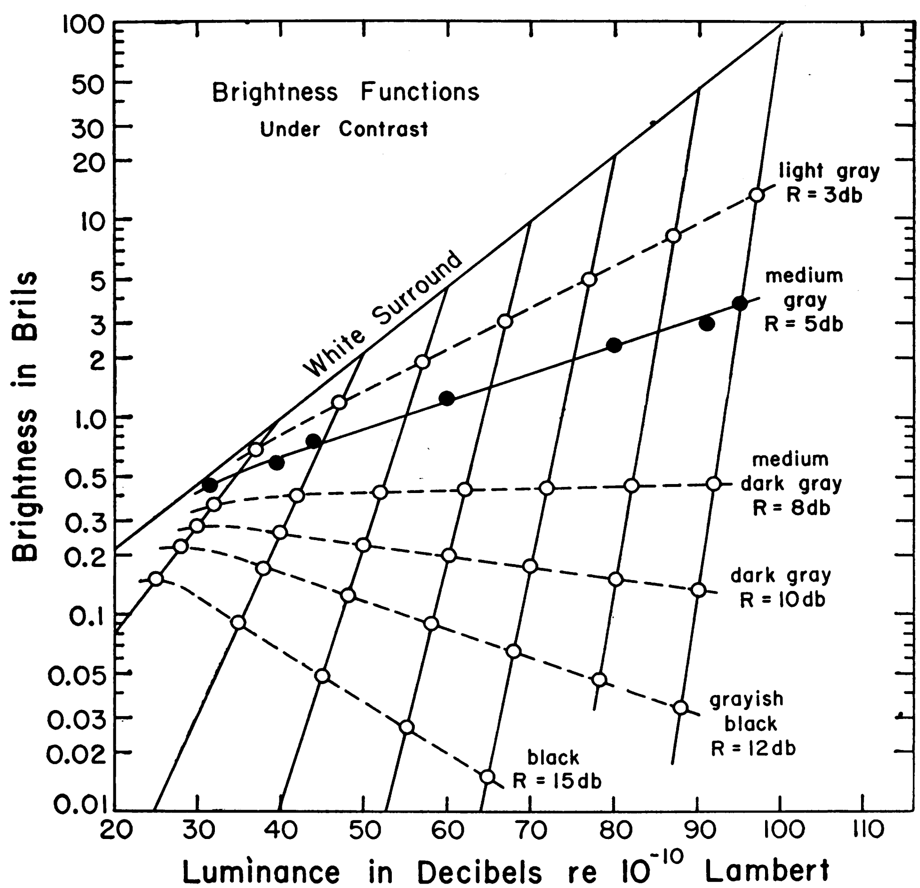

Fig. 4. A generalized set of functions showing the brightness of a target disk seen in the presence of a surround. The brightness function for the white surround is shown by the top line (slope = 0.33). Depending on the luminance of the surround, the brightness function for the target follows one or another of the steeper straight lines. The dashed curves pass through the points (circles) representing given fixed ratios, *R*, between the luminances of disk and surround. These ratios are stated in decibels. Thus, the top dashed curve is for a target disk 3 decibels below the surround in brightness. If the surround is a white paper, a darker paper 3 decibels down would be called light gray, as indicated. So-called brightness constancy is a manifestation of the relative flatness of the dashed curves. The filled circles represent data obtained by Leibowitz *et al.* (*33*), who matched a luminance seen in a black surround to a medium gray seen in a white surround under a wide range of illuminations.

vibration was one of the several procedures used to verify the mysterious fact that listening with two ears is different from listening with one. For some strange reason, the binaural exponent of the loudness function is about 10 percent larger than the monaural exponent (Table 1), and this small difference showed up when the two loudness functions were determined by comparisons with vibration applied to the tip of the finger, just as it manifested itself when the observers "matched" numbers to binaural and monaural loudnesses under the methods of magnitude estimation and magnitude production (*30*).

These procedures established the fact that for a sound pressure level of about 90 decibels the loudness of a sound as heard in two ears is precisely twice as great as the loudness heard when the same sound is delivered to only one ear. This 2-to-1 relation obtains only at one level, however, because of the difference in the size of the exponents for binaural and monaural listening. At lower levels the binaural-monaural ratio is smaller, whereas at levels greater than 90 decibels the binaural sound seems more than two times louder than the monaural sound.

3) In the domain of practical applications, an area that is not always without interest to the academic mind, the sone scale of loudness, the first and the most carefully documented of the modern ratio scales of sensation, has long since proved its utility to the acoustical engineer (the sone is the subjective unit of loudness). This scale recently performed its bit as an essential link in the development of a method for computing the total binaural loudness of a complex sound spectrum, given an analysis of the sound in terms of octave or third-octave bands (*31*). The loudness in sones of each band is determined from a set of equal-loudness contours, and the loudness values are added up according to a simple weighting function. To the loudness of the loudest octave band is added 0.3 times the sum of the loudness in the remaining bands. A version of this procedure is in fairly widespread use and is being readied as a secretariat proposal for general adoption by the International Standards Organization (*32*). The relevance of all this to our present concern is merely to show that ratio scales of sensation have their utility in the world of practical decisions.

4) The extension of brightness scaling to those circumstances under which a target is surrounded by a brighter background has led to some predictions that to me are rather startling. The full story is told elsewhere (*9*), but briefly it is this. We know that a bright surround depresses or inhibits the subjective brightness of a target in a most striking manner. This is the same inhibitory effect that we all experience when the glare of the oncoming headlights renders objects beside the road either dim or invisible. When a target luminance is subjected to the inhibitory contrast imposed by a brighter surround, the exponent governing the apparent brightness of the target jumps to a larger value. At high over-all levels the exponent grows by a factor as great as 10, but at lower levels the value of the exponent is smaller. These exponents are depicted by the slopes of the steep lines in Fig. 4, where we see that the steepness is less at the lower levels. These functions were determined in a long series of experiments aimed directly at determining the slopes of the brightness functions for targets seen under contrast. The brightness values are expressed in subjective units called brils.

If we want to use the functions in Fig. 4 to predict the appearance of a gray paper of a given reflectance viewed against a white background under various levels of illumination, we introduce the dashed lines, each of which shows the locus of the target luminances that bear a fixed ratio to the luminances of the surround. Put more simply, each shade of gray has its own dashed line in Fig. 4, and the line for a given gray shows how the brightness of the gray behaves when we change the illumination falling on the scene, including both the gray and its white surround.

What seems startling in Fig. 4 is that for some shades of gray the apparent brightness is supposed to decrease when the illumination is increased. Turn on more light and the target looks darker! That is the verdict of the dashed lines that slope down toward the right. This prediction has been checked with six observers who, having spent about 10 minutes in adapting to darkness, viewed a dark gray on a white background. The darkness of the target seen in dim light grew suddenly deeper when the illumination was suddenly increased by 10 or 20 decibels. The observers found it especially interesting to watch the target turn gradually blacker as the illumination was gradually increased.

Many other interesting deductions can be made from the functions in Fig. 4. But what we have considered is enough to show that it is indeed possible to use the new psychophysical law and the procedures by which it was established in order, as James put it, "to test a theory or to reach a new result."

References

1. G. T. Fechner, *Elemente der Psychophysik* (1860).
2. R. von Mises, *Probability, Statistics and Truth* (Macmillan, New York, 1939).
3. G. T. Fechner, *In Sachen der Psychophysik* (Leipzig, 1877), p. 215.
4. F. Brentano, *Psychologie vom empirischen Standpunkt* (Dunker and Hunblot, Leipzig, 1874), vol. 1.
5. J. Plateau, *Bull. Acad. roy. méd. Belg.* **33**, 376 (1872).
6. S. S. Stevens, *Am. Scientist* **48**, 226 (1960).
7. J. Delboeuf, *Etude psychologique. Recherches théoriques et experimentales sur la mesure des sensations et spécialement des sensations de lumière et de fatigue* (1873).
8. S. S. Stevens and E. H. Galanter, *J. Exptl. Psychol.* **54**, 377 (1957).
9. S. S. Stevens and J. C. Stevens, "The Dynamics of Visual Brightness," *Psychophysical Project Report, Harvard University, No. PPR-246* (August 1960); *J. Opt. Soc. Am.* **50**, 1139 (1960).
10. J. Merkel, *Phil. Stud.* **4**, 541 (1888).
11. L. L. Thurstone, *The Measurement of Value* (Univ. of Chicago Press, Chicago, 1959), chap. 24.
12. S. S. Stevens, in *Measurement: Definitions and Theories*, C. W. Churchman and P. Ratoosh, Eds. (Wiley, New York, 1959).
13. G. Ekman, "Some aspects of psychophysical research," in *Sensory Communication*, W. A. Rosenblith, Ed. (Technology Press and Wiley, New York, in press).
14. W. R. Garner, *J. Acoust. Soc. Am.* **30**, 1005 (1958).
15. S. S. Stevens, *Psychol. Rev.* **64**, 163 (1957).
16. ———, *Contemp. Psychol.* **4**, 388 (1959).
17. J. C. R. Licklider, "Three auditory theories," in *Psychology: A Study of Science*, S. Koch, Ed. (McGraw-Hill, New York, 1959), vol. 1.
18. M. V. Mathews, *J. Acoust. Soc. Am.* **32**, 931 (1960), abstract.
19. P. G. Nutting, *Trans. Illum. Eng. Soc. N.Y.* **11**, 939 (1916).
20. S. S. Stevens, *J. Exptl. Psychol.* **57**, 201 (1959).
21. R. W. G. Hunt, *J. Phot. Sci.* **1**, 149 (1953).
22. E. B. Newman, *Trans. Kansas Acad. Sci.* **36**, 172 (1933).
23. S. S. Stevens, "Is there a quantal threshold?" in *Sensory Communication*, W. A. Rosenblith, Ed. (Technology Press and Wiley, New York, in press).
24. R. D. Luce, *Psychol. Rev.* **66**, 81 (1959).
25. W. James, *The Principles of Psychology* (Holt, New York, 1890), vol. 1, p. 539.
26. S. S. Stevens, *J. Acoust. Soc. Am.* **27**, 815 (1955).
27. E. H. Galanter and S. Messick, "The relation between category and magnitude scales of loudness," paper presented at the 1st annual meeting of the Psychonomic Society, Chicago, Ill., 1960.
28. J. C. Stevens and J. D. Mack, *J. Exptl. Psychol.* **58**, 405 (1959).
29. J. C. Stevens, J. D. Mack, S. S. Stevens, *ibid.* **59**, 60 (1960); J. C. Stevens and S. S. Stevens, *ibid.* **60**, 183 (1960).
30. G. S. Reynolds and S. S. Stevens. *J. Acoust. Soc. Am.* **32**, 1337 (1960).
31. S. S. Stevens, *ibid.* **28**, 807 (1956).
32. ———, *Noise Control* **3**, No. 5, 11 (1957).
33. H. Leibowitz, N. A. Myers, P. Chinetti, *J. Exptl. Psychol.* **50**, 15 (1955).