Perception, 1981, volume 10, pages 431–434

The impossible dream of Fechner and Stevens

David J Weiss
Department of Psychology, California State University, Los Angeles
Received 5 January 1981

Abstract. The idea that there is a single psychophysical function which describes how the human responds to stimulus intensity is rejected. The form of any empirical function depends upon the buried yet arbitrary assumption about how the stimuli are to be measured. Because psychophysical functions have this arbitrary basis, there can be no universal law, and further, no psychophysical function can reveal a general truth about the nervous system. The power law has been inappropriately reified; the descriptive usefulness of the power function has been incorrectly extended, perhaps because simplicity is appealing.

Much concern has been given to the question of whether a particular psychophysical method yields a valid scale. The debate waxes hot because there is no agreement on what is a sensible validity criterion. For example, the popular and simple technique of magnitude estimation almost always produces a power function. Critics have challenged the validity of the magnitude estimates, questioning the researcher's ability to determine whether the observer really does attend to ratios of sensations.

During the early 1960s, one of the central themes in this controversy was whether the subject applies an internal transformation during the attempt to execute the experimenter's instruction (Treisman 1964). It was noted that this hypothetical construct allowed the psychophysical function to have any (monotone) form (Phillips 1964), thus allowing Fechner's logarithmic candidate to stand (Ekman 1964); and by positing a different transformation for each response task one could resolve the discrepant sensory scales achieved through other scaling methods (Torgerson 1960). The idea of an internal transformation was fiercely opposed by the high priest of magnitude estimation, S S Stevens. A brilliant polemicist and tireless researcher, S S Stevens (1964) argued that there was no reason to resort to unobservables, that parsimony dictated that the data (buttressed by the validational technique of cross-modal matching) be accepted at face value.

Some twenty years later, S S Stevens has clearly won the day. An informal survey of current introductory-level texts on perception shows almost universal acceptance of the idea that sensation magnitude is a power function of stimulus intensity.

While I too (Weiss 1972) have rejected magnitude estimation, for present purposes I will accept the responses as exactly what S S Stevens (1975) says they are—direct estimates of sensory magnitude. Now I want to argue that, even if the judgments are valid, they cannot provide a general psychophysical function; they cannot solve Fechner's problem. Indeed, no data can.

The reason is that any empirical function depends not only on the responses, but also on how the stimuli are measured. A familiar example is the magnitude estimation scale of loudness. Most researchers measure the stimuli in pressure units, and they report the function to be Loudness = $aP^{0 \cdot 67}$. But if the stimulus measurements are made in energy units, the function is Loudness = $bE^{0 \cdot 33}$. The two functions are different in a nontrivial way, but the difference is not serious. Researchers know that the loudness function depends on how the stimuli are measured.

The logical consequence of this example seems to have escaped notice. Sound energy and sound pressure are related by squaring, and so the two reported functions are related by squaring. If a different measuring system for sound intensity were used, say one which was exponentially related to the usual system, then the loudness function would be exponentially related to the usually reported power function. The point is that any measuring system represents an arbitrary choice, and therefore the psychophysical law in which it is embedded has an arbitrary aspect as well.

My favorite illustration of the arbitrariness of measuring systems is length. Imagine a culture which measures length with what we call a slide-rule; the graduations on their rulers would be logarithmically related to those on ours. Operations on lengths in that culture would be orderly and internally consistent, but they would not be the familiar ones. For example, if I measure two sticks and lay them end to end, I find the length of the combination to be equal to the sum of the individual lengths. In the slide-rule culture, though, the length of the sum would be equal to the product of the two lengths. In a similar way, the 'universal' law of gravitational attraction would not be an inverse-square law in slide-rule land, since distance is a length term. Slide-rule length is in no way inferior to conventional length; our system is merely historically entrenched. Given identical sensations, observers in our laboratories and in those of slide-rule land will produce different psychophysical functions for subjective length.

Here we come to the crux of my argument. If a psychophysical function depends on how our culture has elected to scale the stimulus, then there can be no *universal* psychophysical law. The function for each continuum is inextricably linked to an arbitrary historical decision. A particular psychophysical function can have both practical utility (surely a loudness scale is valuable to a manufacturer of audio equipment) and theoretical significance, but its form cannot be compared to other such functions for other continua. It cannot tell us how the nervous system codes. Saying that doubling stimulus intensity leads to doubling the neuron's firing rate is meaningful only within the context of the particular stimulus measurement employed; the statement is descriptive only, and cannot be the basis for a general principle of neural transmission. The psychophysical function for a stimulus dimension is an important empirical fact, and it is meaningful within its context, but we cannot make the leap across continua to achieve 'THE LAW'.

Why then does the power law have such wide acceptance among psychologists? Surely this enthusiasm is based on the success of power functions in describing magnitude estimates of so many different stimulus sets in so many laboratories. I contend that this success is more important practically than theoretically, that it tells us more about power functions than about perception. This contention has two bases. The first is that in an approximate way, a power function can describe almost any subjective scale. The second is that in a detailed way the usual statistical evaluation of the power law is inadequate.

The descriptive capability of the power function when the data show a monotone relation has long been recognized by statisticians (eg Mandel 1964; Tukey 1977). In a similar way, power transformations have been recommended to achieve simple representations (Box and Cox 1964). Since any reasonable scaling data exhibit a monotone relation between stimulus and response, it is not surprising that a power function comes close to the data points. Perhaps psychologists should have been alerted when the ubiquitous power function appeared where there was no theoretical rationale (J C Stevens and Savin 1962), or when it was not supposed to because the responses were deemed invalid (Marks 1968).

But the eyeball test and its statistical counterpart, assessing the correlation between logarithmically transformed stimulus and response, are insensitive to departures from the hypothesized power relation. Use of the correlation coefficient as an index of goodness-of-fit has been criticized in a somewhat different context (Birnbaum 1973; Anderson and Shanteau 1977), but here the difficulty is simply that any monotone function will have a large linear component. Indeed, Good (1972) has shown that for even very nonlinear monotone functions the correlation coefficient is quite close to unity. From a logical standpoint the major problem is that the correlation coefficient can be increased at the experimenter's will merely by spacing the stimuli farther apart. Since difficulties with the correlation coefficient are well known, it is surprising that the much more appropriate analysis of variance technique introduced in this context by Bruvold and Gaffey (1965) (see also Pradhan and Hoffman 1963), in which the significance of the departures from linearity is assessed, is not more widely used (cf. Sclove 1972).

This is not to say that describing a psychophysical function simply by giving its exponent is misleading. It can be quite useful to convey concisely (even if approximately) the idea of how reported pain varies with current. But our familiarity with conventional physical measures should not lead us to think that the nervous system is calibrated according to them, and at the theoretical level this presumption cannot be justified. Any nonlinear transformation on the physical scale gives rise to a different psychophysical function. The resulting function need not preserve the power-law character of the original one.

It would be nice to have simple psychophysical laws. The notion is so attractive that Luce (1959) advocated principles of theory construction that severely limited the possible laws. The present argument suggests an even more drastic option. If we desire to find commonality in the laws for various perceptual continua, we can define the stimuli so that the same law obtains for all the continua; but we must keep in mind that the form of the laws is definitional rather than empirical. Transformations on the independent variable (Box and Tidwell 1962; Ramsay 1977) would be useful in accomplishing 'powerization'. Validation would no longer be a concern, since the scales would be valid by definition.

This facetious scheme is not as devoid of empirical content as it might seem at first glance. Obviously it will not work if scaling data are nonmonotone. Also, we would want to require that physical continua which are logically linked would maintain that linkage in their transformed versions—for example, area should be the square of length—and it might not be possible to achieve this in the rescalings.

References

Anderson N H, Shanteau J, 1977 "Weak inference with linear models" *Psychological Bulletin* **84** 1155-1170

Birnbaum M H, 1973, "The devil rides again: Correlation as an index of fit" *Psychological Bulletin* **81** 854-859

Box G E P, Cox D R, 1964, "An analysis of transformations" *Journal of the Royal Statistical Society, Series B* **26** 211-243

Box G E P, Tidwell P W, 1962, "Transformation of the independent variables" *Technometrics* **4** 531-550

Bruvold W H, Gaffey W R, 1965 "Subjective intensity of mineral taste in water" *Journal of Experimental Psychology* **69** 369-374

Ekman G, 1964 "Is the power law a special case of Fechner's law?" *Perceptual and Motor Skills* **19** 730

Good I J, 1972 "Correlation for power functions" *Biometrics* **28** 1127-1129

Luce R D, 1959 "On the possible psychophysical laws" *Psychological Review* **66** 81-95

Mandel J, 1964 *The Statistical Analysis of Experimental Data* (New York: John Wiley)

Marks L E, 1968 "Stimulus-range, number of categories, and form of the category scale" *American Journal of Psychology* **81** 467-479

Phillips J P N, 1964 "A note on Treisman's model" *Quarterly Journal of Experimental Psychology* **16** 386

Pradhan P L, Hoffman P J, 1963 "Effect of spacing and range of stimuli on magnitude estimation judgments" *Journal of Experimental Psychology* **66** 533-541
Ramsay J O, 1977 "Monotonic weighted power transformations to additivity" *Psychometrika* **42** 83-109
Sclove S L, 1972 "(*Y* vs. *X*) or (log *Y* vs. *X*)?" *Technometrics* **14** 391-403
Stevens J C, Savin H B, 1962 "On the form of learning curves" *Journal of the Experimental Analysis of Behavior* **5** 15-18
Stevens S S, 1964 "Concerning the psychophysical law" *Quarterly Journal of Experimental Psychology* **16** 383-385
Stevens S S, 1975 *Psychophysics* (New York: John Wiley)
Torgerson W S, 1960 "Quantitative judgment scales" in *Psychological Scaling: Theory and Applications* Eds H Gulliksen, S Messick (New York: John Wiley)
Treisman M, 1964 "Sensory scaling and the psychophysical law" *Quarterly Journal of Experimental Psychology* **16** 11-22
Tukey J W, 1977 *Exploratory Data Analysis* (Reading, MA: Addison Wesley)
Weiss D J, 1972 "Averaging: an empirical validity criterion for magnitude estimation" *Perception & Psychophysics* **12** 385-388