ON THE PROBLEM OF SCALES FOR THE MEASUREMENT OF PSYCHOLOGICAL MAGNITUDES

S. S. Stevens
Harvard University, Cambridge, Massachusetts 02138, USA

Abstract

[This paper has no abstract. Stevens presented it at the Fifth International Congress for the Unity of Science, which convened in Cambridge, Massachusetts, USA, on September 3-9, 1939. Stevens's paper, together with several other papers presented at the Fifth International Congress, was scheduled for publication in The Journal of Unified Science (Erkenntnis)*, volume 9. But that journal ended with volume 8, dated 1939/1940. The anticipated volume 9 was never published (see Marks, these* Proceedings*). Consequently, Stevens's paper is presented here – edited and reformatted for this volume – for the first time since he delivered it at the 1939 International Congress. The paper is printed by permission of Peter Stevens. – Lawrence E. Marks]*

We have it from one of their own distinguished number (Campbell, 1938) that "The most distinguished physicists, when they attempt logical analysis, are apt to gibber; and probably more nonsense is talked about measurement than about any other part of physics." If despite so sharp a warning, another sally into the domain of measurement is here attempted, two excuses can be offered. (1) Modern operational analysis (for a statement of the point of view adopted in this paper, see Stevens, 1939) has suggested new ways to lay hold on this ancient problem, and (2) something traditionally held to be unmeasurable has recently been measured – or so it seems to some of us. In the present essay I shall try first to review certain fundamentals of measurement and then to inquire what meaning can be given to the scales recently proposed for the measurement of sensation.

Numeral, Numerousness and Numerosity

An analysis of arithmetic from the point of view of syntactics and semantics reveals the following possible distinctions:

1. *Numerals* are certain signs we commonly make on paper. The names of the numerals are "one", "two", "three", etc.

2. *Numerousness* is a property or attribute which we are able to discriminate when we regard a collection of objects.

3. *Numerosity* is a property defined by certain operations performed upon groups of objects.

Numerals, as signs, are related to numerosity and to numerousness according to certain semantical rules (Note 1). The following is proposed as a possible way of looking at the syntactics and semantics of numerals.

Suppose that we have taken several handfuls of beans from a bag and have placed each handful in a separate region on a table. As we look at these piles of beans we are able to discriminate or pay attention to various aspects or attributes, such as color, pattern, etc. These are ways in which the piles of beans may vary. Among these discriminable aspects of the beans is one that we may call *numerousness*. And simply by looking at the groups of beans, we could order the groups into a series according to the degree to which they exhibited the attribute of numerousness. Perhaps repetition of the experiment would reveal a low reliability in the ranking, but reliability is a relative matter. The point is that such a ranking could be made on the basis of numerousness as directly perceived.

We know from experience, however, that greater reliability can be had if we rank-order the groups by pairing successively one bean from each group until one group is exhausted. Then if any beans remain in the other group, that group is said to have the greater *numerosity*. If the pairing exhausts both groups simultaneously, their numerosity is equal. By this operation (pairing object against object) which we have substituted for the operation of direct estimate, we can order piles of beans into a series in which each successive pile contains exactly one more bean than its predecessor.

Now, if we designate each of these piles by a separate sign, and if we decide to use the same sign to designate all groups showing the same numerosity, we find ourselves in possession of a series of numerals. The 'spatial' (topological) order in which we write the numerals depends on the degree of numerosity designated by each numeral. Then, in order to specify the numerosity of any group we have merely to pair successively each object in the group with a numeral from the numeral-series, beginning of course with the first numeral in the series. This operation we call *counting*. The numeral arrived at when the objects are exhausted is the [one] used to designate the numerosity of the group. We see that the numeral "5" like the word "red" can be used to designate a property of objects.

When we regard the numeral-series as originating in this fashion, we are not astonished to find that it exhibits certain important properties. The relations obtaining among groups, considered from the point of view of numerosity, are reflected in the relations obtaining among numerals – but with an important difference: degree of numerosity among groups corresponds to 'spatial' relations in the numeral-series. Likewise, the numerosity achieved by combining two groups (addition) corresponds to the numeral arrived at by stepping off two successive 'distances' along the numeral-series, and since the order for combining the groups and for stepping off the 'distances' is immaterial, we can demonstrate the associative, commutative and distributive laws both for groups of objects and for our series of numerals. Furthermore, to the extent that we can show transitivity, asymmetry, etc. among relations of numerosity, we can also show them among the topological 'spatial' relations in the numeral-series.

The syntactical rules governing the manipulation of numerals (as directed by such signs as +, x, =) are designed to make the numeral-series useful as a formal model for representing the numerosity of groups combined in various ways. Once in possession of these syntactical rules it was natural that mathematicians should extend their application to new kinds of numerals not designed to represent numerosity. And in the course of a long laborious history, mathematicians invented negative, imaginary and transcendental *numerals*. For our present purposes we might regard the series of all numerals as defining a 'topological space' and the syntactical rules of arithmetic as directions for moving about in that space.

The Relation of Numerousness to Numerosity

Obviously, the numeral-series provides us with an adequate scale for measuring numerosities by the operation of counting. Numerousness cannot be measured by the same operation of counting, but a scale of numerousness could be obtained by one of the following procedures:

(1) We could set up an *ordinal scale* of numerousness by arranging groups in order so that the numerousness of each would be either greater than or equal to that of the group preceding. (Reliability would here by limited by our powers of discrimination.) We could then assign numerals to the groups so ordinalized.

(2) We could set up an *intensive scale* by prescribing a semantical rule for determining the assignment of *adjacent* numerals: by assigning them to groups showing a just noticeable difference in numerousness.

(3) We could erect an *extensive scale* by determining when one group appeared half as numerous as another. If this judgment could be made, we could then assign to the smaller group the numeral lying in the numeral-series midway from the beginning of the series and the numeral assigned to the larger group. If we assigned numerals according to this procedure, the relations among groups exhibiting the property numerousness would be reflected in the 'spatial' relations of numerals within the numeral-series.

These scales of numerousness could be compared directly with the scale of numerosity.

A Psychological Scale of Loudness

Loudness is defined by the reaction of an observer to a sound-wave – he can discriminate various aspects of sounds just as he can discriminate various aspects of objects. And given an assortment of sounds, he can tell which exhibit greater or less loudness than a particular standard sound. Now, several procedures for the erection of scales are open to use. Since we are able to say when two sound-waves are equally loud, we could *define* the loudness of their *combination* as a loudness twice as great as that of either one alone. This is precisely the procedure we follow in setting up our fundamental scales of weight. The weight of a combination of two objects of equal weight is assigned the numeral twice as far along the numeral-series as the numeral assigned to each object alone. As in the case of weight we shuld have to prescribe for loudness eactly how the combining of equals is to be carried out, but this we could do; and the result would be a scale of loudness having all the properties of our scales of weight. But psychologists would not be happy to call such as scale a scale of loudness, for it would turn out to coincide with what we ordinarily call scales of *sound-intensity*. It would be the analogue of a scale of numerosity rather than a scale of numerousness.

As alternative procedures, then, first we could ordinalize any particular series of sounds and assign numerals in such a way as to achieve an *ordinal scale*. Secondly, we could employ as a rule for assigning adjacent numerals that two sounds should be just noticeably different in loudness. This was the procedure proposed by Fechner, and from it we obtain an *intensive scale*. (Fechner made the further assumption, which we shall presently examine, that the addition of just noticeable increments to the stimulus adds equal increments of loudness.) Thirdly, we could construct an *extensive scale* for loudness by having an observer tell us when one sound appears half as loud as another. To the louder sound we should assign the numeral standing in the numeral-series twice as far along as the numeral assigned to the weaker sound (for a consideration of sensory scales in the field of audition see Stevens & Davis, 1938).

Of what use are such psychological scales? Is there a pragmatic sanction for the effort involved in their construction? First of all it is instructive to compare the scales of loudness gotten by different operations. We find that the extensive scale constructed by the method of fractionation does not coincide with the scale of sound-intensity. And both of these scales differ in form from the intensive scale obtained by counting off just noticeable differences. Now, *provided* we accept the extensive scale as defining subjective loudness, we can decide upon the validity of Fechner's assumption regarding the equality of just noticeable increments. A comparison of the extensive scale with Fechner's scale shows that the addition of successive just noticeable increments to the stimulus adds larger and larger increments of loudness.

Finally, I should like to point out the possible importance of these scales in the field of psychophysiology. As already indicated, the scale of loudness obtained by the direct judgment of fractional relations does not correspond to that obtained by counting off just noticeable differences. It has been found, however, that in the case of pitch these two types of scales coincide rather precisely (Stevens, Volkmann, & Newman, 1937). Now, we believe that a tone appears louder to an observer when the *amount* of physiological excitation is increased, but we believe that a tone appears higher in pitch when the *distribution* of stimulation on the basilar membrane is shifted, i.e., when stimulation of one sort is substituted for that of another. Discrimination, then, can be mediated by these two basic mechanisms: (1) the addition of excitation to excitation and (2) the substitution of one excitation for another (change in the distribution of excitation). The contrasting relations among the pitch-scales as opposed to those among the loudness-scales suggest that these basic physiological mechanisms are reflected in the scales we achieve when we assign numerals to sounds. Furthermore, the behavior of pitch and loudness, as regards the relation between their extensive and intensive scales, suggests the hypothesis that perhaps all attributes based upon a physiological mechanism of the first type (addition of excitation) will show a divergence betwen the sensory scales based upon fractionation and the scales obtained by counting up just noticeable differences, whereas the scales of all attributes mediated by the second type (substitution or change in distribution of excitation) will not show this divergence (Note 2). In other words, a study of the scale-forms for the various attributes of sensation may lead eventually to the demonstration of psychophysiological laws.

Acknowledgements

[The efforts of Sergio Cesare Masin and Richard E. Kaufman are greatly appreciated and hereby acknowledged. Masin's initial search for Stevens's paper instigated the events that eventually led to its discovery. Kaufman is the Librarian for the Department of Psychology, Harvard University, and it was Kaufman who eventually located a copy of this never-published paper. Copies of several papers presented at the Fifth International Congress for the Unity of Science, including Stevens's, were printed for the members of the Congress, in advance of the meeting. At the top (upper right) of the first page of Stevens's paper, as printed for the Congress, there appears the following:

Reprinted for
the members of the fifth International Congress for the Unity of Science,
Cambridge, Mass. U.S.A. 1939, from:
The Journal of Unified Science (Erkenntnis) Volume IX
Editors: Rudolf Carnap and Hans Reichenbach.
Associate Editors: Philipp Frank, Joergen Joergensen, Charles W. Morris,
Otto Neurath, Louis Rougier, L. Susan Stebbing.
Publisher W. P. van Stockum & Zoon, 36 Buitenhof, The Hague, (Holland)
American Agents: University of Chicago Press, 5750 Ellis Avenue, Chicago, Ill.

The pages of Stevens's paper that were printed for the Congress are numbered 94 through 99. These presumably would have been the page numbers of the article in The Journal of Unified Science (Erkenntnis)*, had volume 9 been published. The copy of the article found by Kaufman is bound in a volume, located in the Harvard Psychology Department Library, containing a collection of articles by Stevens. That copy of the article itself originally belonged to Stevens, as indicated by the note, in his distinctive hand, in the upper left of the first page, "Please return to SS Stevens". – Lawrence E. Marks]*

References

Campbell, N.R. (1938). Symposium: Measurement and Its Importance for Philosophy. *Aristotelian Society*, 17 (Supplement). London: Harrison and Sons.

Judd, D.B. (1933). Saturation scale for yellow colors. *Journal of the Optical Society of America*, *23*, 35-40.

Merkel, J. (1888). Die Abhängigkeit zwischen Reiz und Empfindung. [The relation between stimulus and sensation.] *Philosophische Studien*, *4*, 541-594.

Stevens, S.S. (1939). Psychology and the science of science. *Psychological Bulletin*, *36*, 221-263.

Stevens, S.S., & Davis. H. (1938). *Hearing: Its Psychology and Physiology*. New York: Wiley.

Stevens, S.S., Volkmann, J., & Newman, E.B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, *8*, 185-190.

Notes

Note 1. In choosing the two words numerousness and numerosity (barbarous as they may sound) I am deliberately avoiding the term number. Number usually refers to numerosity and is therefore related to numerals by means of semantical rules, but all too often the distinction between number and numeral is left ambiguous. Both Campbell (1938) and R. R. Ravven, in a thesis submitted to Harvard University, have emphasized the distinction between number and numeral and have pointed out that numerals, and not numbers, are what we assign to objects when we measure them. In order to emphasize this distinction I propose to use the term numerosity instead of number.

Note 2. The experiments necessary to test the hypothesis, that the subjective size of just noticeable differences is constant for attributes based upon a substitutive process and is a function of the magnitude of the stimulus for attributes based upon the additive process, have not been performed for all senses-departments. Nevertheless, what little evidence is available appears to support the notion. In the field of visual brightness the somewhat meager data indicate the expected divergence between the intensive and the extensive scales. I am indebted to Alfred L. Baldwin for a re-examination of data secured by Merkel (1888) who appears to have been the first to employ the method of fractionation. Merkel was not interested in our present problem, but his data show clearly that the subjective size of just noticeable differences increases with the intensity of the visual stimulus. On the other hand the size of the just noticeable difference for visual saturation appears to remain constant (Judd, 1933). This fact suggests that the discrimination of saturation is based upon a physiological process which is substitutive in nature.

S. S. STEVENS – A BRIEF SCIENTIFIC BIOGRAPHY

Lawrence E. Marks
John B. Pierce Laboratory and Yale University, New Haven, CT 06519, USA
(marks@jbpierce.org)

Abstract

This brief biography recounts the highlights of the scientific career of S.S. Stevens, 1906-1973.

Stevens was born in Ogden, Utah, on November 4, 1906. After two years each at the University of Utah and Stanford University, where he received his undergraduate degree, Stevens moved in 1931 from Palo Alto to Cambridge, entering Harvard as a post-graduate student. There he would remain – as a student, post-doctoral researcher, and member of the faculty – until his death, in Vail, Colorado, on January 18, 1973.

Although intending initially to study medicine at Harvard, Stevens found himself drawn to experimental psychology – and especially to psychophysics, working under the aegis of E.G. Boring, at the time Harvard's only Professor of Psychology (psychology still being a division within the Department of Philosophy). Three of Stevens's earliest publications dealt with the attributes of tones, and are noteworthy in recognizing the critical role of *invariance*, a notion that would be central to Stevens's classification of scales of measurement. Awarded his Ph.D. in 1933, Stevens conducted post-doctoral work for three years, first as an assistant to Boring, later as a fellow with Hallowell Davis, and finally as research fellow in physics. The collaboration with Davis led to several important articles in psychoacoustics and, notably, to their classic book on *Hearing*, published in 1938.

In 1936, Stevens became a junior faculty member in Harvard's newly independent Department of Psychology, eventually rising to the rank of Professor of Psychology and, finally, to Professor of Psychophysics. The early 1940s and the Second World War brought Stevens to the basement of Memorial Hall, where he, along with students and colleagues in the Psycho-Acoustic Laboratory, examined communication under conditions of intense background noise, like those encountered by military pilots, while the late 1940s brought him to the ski slopes. During this period, Stevens managed also to edit the mammoth and seminal *Handbook of Experimental Psychology*, published in 1951 – a work of 36 chapters, whose authors included 17 eventual members of the National Academy of Sciences, that would serve as a secular bible for at least a generation of graduate students.

Stevens wrote or edited a total of nine books, the three most important of which were spaced roughly equally over his four decades at Harvard: *Hearing* near the start, the *Handbook* near the middle, and *Psychophysics: An Introduction to Its Perceptual, Neural, and Social Prospects* at the end. In addition to the books were 137 major articles and chapters; among the most significant were his 1946 article proposing the classification of measurement scales, his 1951 chapter relating measurement, mathematics, and psychophysics, and his 1957 article formally proposing the power function as a general law governing intensity perception in sensory systems. It was wholly fitting and proper that over his eminent career Stevens garnered many awards from his scientific colleagues, including the Rayleigh Gold Medal, the Warren Medal of the Society of Experimental Psychologists, and election to the National Academy of Sciences.