Population norms of top sensory magnitudes and S. S. Stevens' exponents[1]

E. C. POULTON
APPLIED PSYCHOLOGY RESEARCH UNIT, CAMBRIDGE, ENGLAND

If S. S. Stevens' exponents indicate the rates at which sensations grow with increases in sensory intensity, they ought to correlate with the population norms of top sensory magnitudes. Using a comprehensive sample of eight sensory dimensions, the tau coefficient of rank correlation between Stevens' exponents and the medians of the top sensory magnitudes reported by 305 observers was found to be only +.15 (p >.05). With the geometric means tau fell to -.04. A split-half consistency check on the medians of the population norms suggested that they were not to blame for the low correlation. Direct comparisons of pairs of sensory dimensions on 146 additional observers produced results which confirmed the population norms. Since there is no way of comparing most of the top physical stimuli experienced in everyday life, it is not possible to make a joint prediction from exponents and top stimuli. S. S. Stevens' exponents thus appear to have little predictive value outside the experimental conditions under which they were measured.

In numerous experimental and theoretical papers, many of which can be found in the list of references, S. S. Stevens has described how the logarithms of estimated sensory magnitudes are related linearly to the logarithms of the physical intensities of the stimuli presented. The steepness of the slope for any one sensory dimension should indicate the rate at which the particular sensation grows with increasing physical intensity (S. S. Stevens, 1960). Since physical intensity is plotted in logarithmic units, the steepness of the slope does not change with changes in the size of the physical units of measurement (although it is sensitive to power transformations, for example, a change from sound pressure to energy flux density—see S. S. Stevens, 1956, p. 4). Thus, on the face of it, the slope should be a characteristic of sensation with considerable generality. Taken together with the top physical stimulus met in everyday life, the slope ought to predict the size of the most intense sensation experienced along its particular sensory dimension:

$$\text{Slope} = \frac{\text{Log (Top sensation)}}{\text{Log (Top stimulus)}} \qquad (1)$$

Unfortunately, it is not possible to compare the top stimulus values of most of the physical dimensions experienced in everyday life, owing to the arbitrary definitions of the units of measurement. Should lengths be measured in microns or light years? But even when the sizes of the top stimuli are neglected, there ought still to be some positive association (although not perfect) between the calculated slopes and the logarithms of the reported sizes of the most intense sensations experienced.

In his Klopsteg lecture, S. S. Stevens (1962) lists the exponents or slopes obtained in his laboratory for 27 sensory dimensions. The left side of Table 1 gives the exponents and stimulus ranges of the 21 out of 27 sensory dimensions for which adequate experimental data have been published. The dimensions are ordered according to the sizes of the exponents. The stimulus range is defined geometrically as the ratio of the largest variable to the smallest variable. This is clearly unsatisfactory from a theoretical point of view, since halving the smallest variable makes little difference to the arithmetic range; yet it has as great an effect on the geometric range as doubling the largest variable, which practically doubles the arithmetic range. But it is not possible to define the stimulus range arithmetically, since it then depends upon the sizes of the units used in each physical dimension, which are arbitrary. Where two or more stimulus ranges were used in different experiments, the range of the first experiment has been given, since it may have influenced both the E and some of the Os in subsequent experiments (see Poulton, 1967).

For the 21 dimensions listed in the table, the tau coefficient of rank correlation (Siegel, 1956, p. 213) between size of exponent and geometric stimulus range is -.60 (p< .001). This means that 36 percent of the variance of the tabulated exponents is accounted for simply by the geometric range of stimulus variables used in determining them. A sizeable correlation would be predicted from the results of experiments in which the range of variables has been varied within a single sensory dimension (see Poulton, 1967). The range of variables selected for an experiment is one of the more powerful determinants of the size of the exponent found.

The negative correlation with experimental stimulus range means that Stevens' exponents do not directly represent the theoretical slopes of expression (1). To obtain the theoretical slopes it is necessary to correct the exponents for the experimental stimulus ranges. How this can be done, and whether it can legitimately be done at all, is not yet clear. If Stevens' exponents are used instead of the theoretical slopes, the negative correlation with experimental stimulus range may well reduce the size of the expected relationship with the logarithms of the reported sizes

Table 1. Population Stereotypes and S. S. Stevens' Sensory Dimensions

Reference		Sensory Dimension	Stevens' Exponent	Stevens' Range Bottom	Stevens' Range Top	Stevens' Range Units	Top ÷ Bottom	Experiment 1 Log10 Top Estimate Median	Experiment 1 Log10 Top Estimate Geometric Mean	SE	N
J. C. Stevens & Stevens,	1963	brightness 5 degree target	.33	44	104	decibels	10^6	2.54	2.90	.27	37
S. S. Stevens & Galanter,	1957	brightness point source	.50	80	110	decibels	10^3				
Reynolds & Stevens,	1960	loudness monaural	.54	40	100	decibels	10^6				
Reese & Stevens,	1960	coffee odor	.55	.0075	.60	relative concentration	80	1.62[a]	1.77	.18	38
S. S. Stevens,	1956	loudness binaural	.60	30	100	decibels	10^7	2.87	3.18	.34	44
S. S. Stevens & Guirao,	1963	visual area	.70	17	3200	sq cm	188				
Harper & Stevens,	1964	tactual hardness	.80	8.3	970	force/ identation	117	2.00	2.77	.25	42
S. S. Stevens,	1959	vibration intensity	.95	10	40	decibels	10^3				
J. C. Stevens & Stevens,	1960	cold	1.00	3.3	30.5	degrees below "neutral" temperature	9.2	2.00	2.42	.28	37
J. C. Stevens & Shickman,	1959	repetition rate	1.00	1	40	signals per sec	40				
S. S. Stevens & Galanter,	1957	visual length	1.00	4	111	cm	27.8				
S. S. Stevens & Galanter,	1957	duration	1.10	.25	4.0	sec	16				
J. C. Stevens & Mack,	1959	pressure on palm	1.10	.5	5.0	lb	10				
Lane, Catania & Stevens,	1961	vocal effort	1.10	2	27	decibels (relative)	320				
S. S. Stevens & Galanter,	1957	lightness of grays	1.20	65	80	decibels	32				
S. S. Stevens & Stone,	1959	finger span	1.30	2.3	63.7	mm.	27.7				
S. S. Stevens & Galanter,	1957	heaviness	1.45	19	193	g	10.1				
J. C. Stevens & Stevens,	1960	warmth	1.50	2.5	14.7	degrees above "neutral" temperature	5.9	2.30	2.61	.28	35
S. S. Stevens & Harris,	1962	tactual roughness	1.50	24	320	grit No. (proportional to reciprocal of grain diameter)	13.3	1.69[a]	2.14	.42	38
J. C. Stevens & Mack,	1959	force of handgrip	1.70	4	40	lb	10				
S. S. Stevens, Carton & Shickman,	1958	electric shock	3.50	.38	1.15	mA	3.0	3.03[b]	3.39	.23	34

a Coffee odor and tactual roughness both reliably smaller at the .001 level than any of the other sensory dimensions tested.
b Electric shock reliably greater at the .05 level or better than any other sensory dimension tested.

of the most intense sensations experienced. However, the negative correlation still leaves 64 percent of the variance, some of which ought to be related to the population norms of intense sensation.

The aim of Experiment I was to determine what the population norms were for the top sensory magnitudes of eight different sensory dimensions, so that the norms could be correlated with Stevens' exponents. The small but comprehensive sample of dimensions was selected by taking three of the dimensions with the largest exponents, three with the smallest, and two with exponents of intermediate size.

Experiment II was an additional check on the order of the median top sensory magnitudes obtained in Experiment I. Pairs of dimensions which were found to have reliably different distributions of top magnitudes were compared directly, to discover whether the results could be predicted from the top magnitude estimates of the separate dimensions. Pairs of dimensions with more similar distributions were also compared.

Procedure

The experiments were carried out on groups of adults seated at tables. Each person was given a pencil and a slip of paper. The instructions were read aloud by the E. In Experiment I the instructions for electric shock ran as follows: "It is possible to use numbers to give some indication of how intense a stimulus feels. For example, in the case of electric shock, the severer an electric shock, the higher the number which represents its severity. Please write down on your slip of paper the number which you think represents the severest electric shock you could experience." For other dimensions "the severer an electric shock" was replaced by "the louder a noise," "the brighter a light," "the stronger the smell of coffee," or "the hotter (colder, rougher, harder) a thing feels." The remainder of the instructions was modified appropriately.

The instructions deliberately avoided any reference to a particular range of stimuli or to a particular set of numbers. Nor was a standard stimulus mentioned, nor a numerical modulus. For it has been found (Poulton, 1967) that Os' magnitude estimates can be influenced reliably by specific information of this nature.

The instructions for Experiment II started with the same first sentence, and then continued: "The more intense the stimulus, the higher the number to represent its intensity." In comparing electric shock with brightness, the instructions then ran: "Please think of the severest electric shock you could experience and of the brightest light you could experience. I want you to write down on your slip of paper first, the one which would be the more intense, the severest electric shock or the brightest light; and second, a number representing how much more intense you think it would be than the other one." For half the Os electric shock always preceded light, as in the instructions quoted. For the other half of the Os the order was reversed.

Experimental subjects

The 305 Os used in Experiment I, and the 146 of Experiment II, were members of a panel maintained at the Applied Psychology Research Unit at Cambridge. Their ages ranged from 21 to 68 years. About 85 percent were women, mainly housewives. None had ever previously judged sensory magnitudes.

EXPERIMENT I: TOP MAGNITUDE ESTIMATES

The right side of Table 1 gives the logarithms of the medians of the top magnitude estimates obtained in Experiment I, the log geometric means with their standard errors, and the number of Os in each group. In computing the geometric means it was necessary to exclude one O from the group estimating cold who gave infinity. An O in the group estimating brightness who gave 10^{100} was also excluded, leaving the highest estimate 10^{17} coming from the group judging roughness. The next highest estimate was 10^{12} for loudness, and there was an increasing number of estimates below 10^{10}.

The geometric means are all larger than their corresponding medians, but the rank order does not change very much. The three sensory dimensions with the largest average estimates, electric shock, loudness, and brightness, are in the same rank order. So are the two dimensions with the smallest averages, coffee odor and roughness. The log median is probably a better measure than the log geometric mean, since it is less affected by the odd O who gives a very large estimate. Except in the case of infinity, it is difficult to decide whether or not to discard very large estimates in computing the geometric mean.

Comparing the log medians with Stevens' exponents, the median top estimate for shock (3.03) is the largest, just as S. S. Stevens' exponent for electric shock (3.50) is the largest. But the two next largest median top estimates, for loudness (2.87) and brightness (2.54), are associated with two of the smallest exponents (.60 and .33, respectively). Also of the two smallest median top estimates, the one for coffee odor (1.62) is associated with one of the smallest exponents (.55); while the one for tactual roughness (1.69) is associated with one of the largest exponents (1.50). The tau correlation between median top estimates and exponents is only +.15, which is not reliable ($p > .05$). The tau correlation between median top estimates and the stimulus ranges listed in the table is smaller still, only +.04. When geometric means are used instead of medians, the tau correlations fall to -.04 and -.15, respectively.

Stability of the population norms

There are several possible reasons why the order of the exponents in Table 1 does not predict the order of the average top magnitude estimates. One reason is possible unreliability of the top estimates. The distributions of the top estimates are given in Fig. 1, grouped in logarithmic orders of magnitude to the base 10. There is a certain amount of skewing towards the higher values, but the skewness is much less marked than when the data are grouped arithmetically. This suggests that in estimating top sensory magnitudes people may use numbers logarithmically. The data are also consistent with Attneave's (1962, pp. 626-627) alternative suggestion that people use numbers in much the same way as they judge sensory magnitudes, numerical estimates being a power function of experienced magnitude, for which Attneave proposes an exponent of about .4.

The distributions in Fig. 1 fall into four groups. At the top end the distribution for electric shock has its mode in the fourth column from the left. This is one column further to the right than the mode of any of the other distributions. On two-tailed Mann-Whitney U tests (Siegel, 1956, p. 116) the distribution was

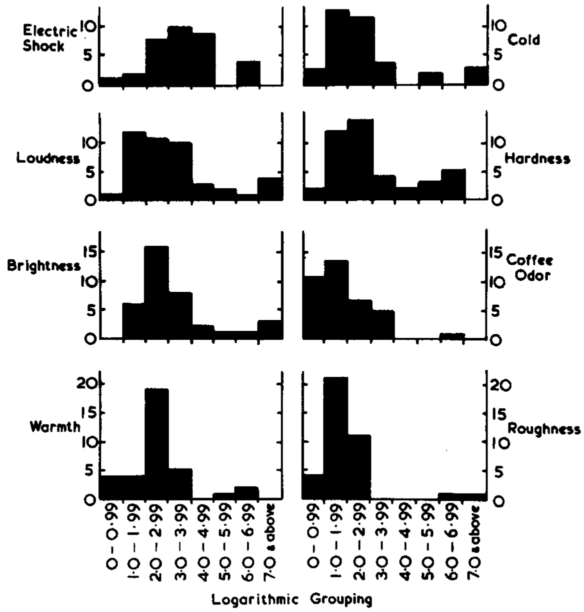

Fig. 1. Distributions of top sensory magnitudes for eight sensory dimensions obtained from separate groups of observers. The estimates have been grouped in logarithmic units to the base 10.

reliably different from all others (p< .05 or better). At the bottom end the distributions for coffee odor and tactual roughness have modes in the second column from the left. Both are reliably different from all the other six distributions (p< .001). The remaining distributions all have modes in the third column from the left, except for loudness where the third column is the middle of three tall columns. None of these distributions differ reliably from each other.

A method of assessing the stability of the order of the median top magnitude estimates in Table 1 is to split each experimental group into two, and to compare the rank order of the medians of the first halves with the rank order of the second halves. Four of the groups each comprised two subgroups tested at different times, which were taken to be the two halves. The other four groups each had three or four subgroups. Here the first two subgroups were taken as the first half, the remaining subgroup(s) as the second half. The largest discrepancy in size between two halves was 10 for the group comprising 44 Os. The tau correlation between the rank orders of the medians of the two halves was +.55. When the rank order of the medians of each half was correlated separately against the rank order of S. S. Stevens' exponents, the taus were zero and +.19. Dividing the average of these two values by +.55 gives a tau corrected for attenuation (Guilford, 1936, p. 366) of +.17, which is little different from the uncorrected tau of +.15. Thus the small size of the rank correlation between Stevens' exponents and the median top magnitude estimates cannot be blamed upon the unreliability of the top magnitude estimates.

EXPERIMENT II: COMPARISONS BETWEEN DIMENSIONS

As a further check on the reproducibility of the top magnitude estimates, pairs of sensory dimensions were compared in Experiment II. In addition to stating which extreme of two sensory dimensions was the more intense, the Os were required to estimate how much more intense it was. When the estimates were grouped in logarithmic units to the base 10, and the group of smallest estimates in favor of one of the two dimensions was placed alongside the group of smallest estimates in favor of the other dimension, the combined distribution was always unimodal, as shown in Fig. 2.

In comparing electric shock with brightness, 24 Os thought the severest electric shock would be more intense than the brightest light; only four thought the reverse. The difference was reliable at the .001 level on a Binomial test (Siegel, 1965, p. 36). The median estimate put electric shock 15 times more intense, the geometric mean 7.3 times. In comparing electric shock with warmth, the voting was 21 to 5 (p< .005). The median estimate placed electric shock

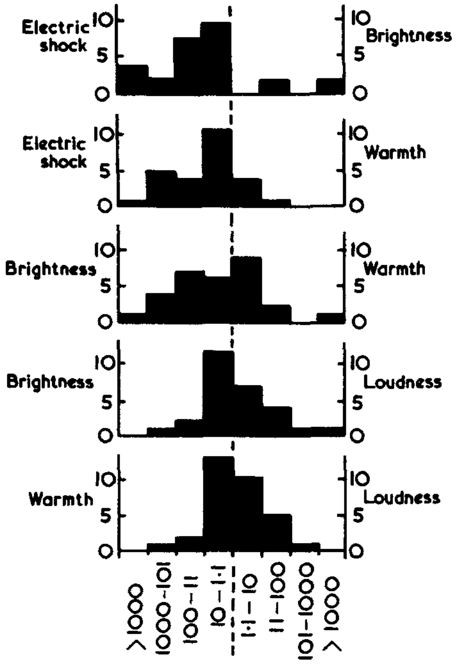

Fig. 2. Numbers of observers stating which top magnitude of two sensory dimensions was the more intense, and how much more intense it was. The ratio estimates have been grouped in logarithmic units to the base 10.

six times more intense, the geometric mean 13.1 times. Thus, as a population norm, the severest electric shock is reliably more intense than either the brightest light or the hottest feeling as was found in Experiment I, although the sizes of the differences cannot be specified precisely from the present data.

No reliable difference was found in paired comparisons of brightness with warmth, the voting being 18 to 12. The median estimate placed brightness three times more intense, the geometric mean 3.2 times. No difference at all was found in comparisons of brightness with loudness, where the voting was 15 to 15, nor of warmth with loudness, where the voting was 16 to 16. These results again correspond to those of Experiment I, where the distributions of top magnitude estimates on the three sensory dimensions were not found to differ reliably. Thus, the population norms appear to have some consistency within the population of Os, although the data do not permit exact numerical predictions between the two experiments.

S. S. Stevens' exponents predict the highly reliable difference between electric shock (exponent 3.50) and brightness (exponent .33). They also predict the reliable difference between electric shock (exponent 3.50) and warmth (exponent 1.50). The similarity between brightness (exponent .33) and loudness (exponent .60) is a little harder to reconcile, unless brightness is taken to be the brightness of a point source (exponent

.50) and loudness is monaural (exponent .54). And the similarities between warmth (exponent 1.50) on the one hand, and brightness (exponent .33) and loudness (exponent .60) on the other, are not predicted.

VALIDITY OF S. S. STEVENS' EXPONENTS

A second possible reason for the failure of the order of the exponents in the table to predict the order of the median top magnitude estimates is that the sizes of the exponents are merely a function of the experimental conditions under which they were determined. As already indicated, 36 percent of their variance is attributable to the geometric ranges of stimuli used in their determination. The influences of other independent variables, such as the choice of the first stimulus or standard, of the distance from this of the second stimulus or first variable, of the modulus, and of the set of numbers for the O to use, have been discussed by Poulton (1967). Until the sizes of these effects, both singly and combined, have been measured for the different sensory dimensions, it will not be possible to decide whether or not S. S. Stevens' exponents are determined exclusively by the choices of the independent variables used in his experiments.

A third and more charitable reason for the low correlation between exponents and median top estimates is suggested by expression (1) in the introduction. The top estimates are a joint function of the exponents or slopes and of the sizes of the top physical stimuli experienced in everyday life. It is conceivable that the top physical stimuli account for most of the variance, leaving only a small part attributable to the slopes. This hypothesis cannot be tested, since there is no sensible way of comparing the sizes of top physical stimuli. But if it is the case, the relative sizes of the individual exponents do not have as great a predictive value outside the experimental conditions under which they were determined as might be supposed from their face value.

References

Attneave, F. Perception and related areas. In S. Koch (Ed.), *Psychology: a study of science*. Vol. 4. New York: McGraw-Hill, 1962. Pp. 619-659.

Guilford, J. P. *Psychometric methods*. New York: McGraw-Hill, 1936.

Harper, R., & Stevens, S. S. Subjective hardness of compliant materials. *Quart. J. exp. Psychol.*, 1964, 16, 204-215.

Lane, H. L., Catania, A. C., & Stevens, S. S. Voice level: autophonic scale, perceived loudness, and effects of sidetone. *J. Acoust. Soc. Amer.*, 1961, 33, 160-167.

Poulton, E. C. The new psychophysics: six models for magnitude estimation. 1967, in preparation.

Reese, T. S., & Stevens, S. S. Subjective intensity of coffee odor. *Amer. J. Psychol.*, 1960, 73, 424-428.

Reynolds, G. S., & Stevens, S. S. Binaural summation of loudness. *J. Acoust. Soc. Amer.*, 1960, 32, 1337-1344.

Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Stevens, J. C., & Mack, J. D. Scales of apparent force. *J. exp. Psychol.*, 1959, 58, 405-413.

Stevens, J. C., & Shickman, G. M. The perception of repetition rate. *J. exp. Psychol.*, 1959, 58, 433-440.

Stevens, J. C., & Stevens, S. S. Warmth and cold: dynamics of sensory intensity. *J. exp. Psychol.*, 1960, 60, 183-192.

Stevens, J. C., & Stevens, S. S. Brightness function: effects of adaptation. *J. Opt. Soc. Amer.*, 1963, 53, 375-385.

Stevens, S. S. The direct estimation of sensory magnitudes-loudness. *Amer. J. Psychol.*, 1956, 69, 1-25.

Stevens, S. S. Tactile vibration: dynamics of sensory intensity. *J. exp. Psychol.*, 1959, 57, 210-218.

Stevens, S. S. The psychophysics of sensory function. *Amer. Sci.*, 1960, 48, 226-253.

Stevens, S. S. In pursuit of the sensory law. Second Klopsteg lecture, November, 1962, Technological Institute, Northwestern University.

Stevens, S. S., Carton, A. S., & Shickman, G. M. A scale of apparent intensity of electric shock. *J. exp. Psychol.*, 1958, 56, 328-334.

Stevens, S. S., & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. *J. exp. Psychol.*, 1957, 54, 377-411.

Stevens, S. S., & Guirao, M. Subjective scaling of length and area and the matching of length to loudness and brightness. *J. exp. Psychol.*, 1963, 66, 177-186.

Stevens, S. S., & Harris, J. R. The scaling of subjective roughness and smoothness. *J. exp. Psychol.*, 1962, 64, 489-494.

Stevens, S. S., & Stone, G. Finger span: ratio scale, category scale, and JND scale. *J. exp. Psychol.*, 1959, 57, 91-95.

Note

1. W. R. Garner kindly commented on a draft manuscript, P. M. E. Altham advised on the statistics, and C. H. Brock helped to calculate the geometric means. Financial support from the British Medical Research Council is also gratefully acknowledged.

(Accepted for publication March 29, 1967.)