

Founded in 1887 by G. STANLEY HALL

OFFPRINTED FROM

THE AMERICAN JOURNAL OF PSYCHOLOGY

EDITED BY

KARL M. DALLENBACH
UNIVERSITY OF TEXAS

AND

M. E. BITTERMAN
BRYN MAWR COLLEGE

E. B. NEWMAN
HARVARD UNIVERSITY

WITH THE COÖPERATION OF

E. G. BORING, Harvard University; W. K. ESTES, Indiana University; J. P. GUILFORD, University of Southern California; HARRY HELSON, University of Texas; E. R. HILGARD, Stanford University; FRANCIS W. IRWIN, University of Pennsylvania; G. L. KREEZER, Washington University; D. G. MARQUIS, University of Michigan; GEORGE A. MILLER, Harvard University; R. M. OGDEN, Cornell University; W. B. PILLSBURY, University of Michigan; LEO POSTMAN, University of California; W. C. H. PRENTICE, Swarthmore College; T. A. RYAN, Cornell University.

ESTIMATIONS OF LOUDNESS BY A GROUP OF UNTRAINED OBSERVERS

By JOSEPH C. STEVENS AND ENDEL TULVING
Harvard University

December, 1957, Vol. LXX, pp. 600-605

Published by The American Journal of Psychology, Department of
Psychology, University of Texas, Austin, Tex.

ESTIMATIONS OF LOUDNESS BY A GROUP OF UNTRAINED OBSERVERS

By JOSEPH C. STEVENS and ENDEL TULVING, Harvard University

In a recent issue of this JOURNAL S. S. Stevens reported several experiments, the results of which support the validity of the method of magnitude-estimation in establishing a scale of loudness.¹ He relied principally upon two methods of research. In one, *E* assigned a numerical value to a standard loudness and *O* estimated other variable loudnesses in terms of it. In the second, no standard was used and *O* estimated the loudnesses of various sounds on a scale of his own choosing.

Since the results reported in Stevens' paper were in general obtained from sophisticated *O*s (graduate students and members of the staff), the question may be asked whether less sophisticated *O*s would give comparable results. In assessing the validity of the loudness-scale derived from direct estimates of magnitude with no designated standard, it might be advisable to use naïve *O*s. That is to say, if *O*s have estimated loudness in experiments in which *E* designated a standard by means of a particular number, they may come to experiments in which no standard is designated with a predisposition to use numbers similar to those previously designated. If this is the case, it would be difficult to distinguish between the two procedures.

A second question to be considered is whether the method of 'magnitude-estimation' yields results when employed with groups of *O*s that are comparable to those obtained from *O*s serving alone. Except for an early study by Ham and Parkinson, most experiments on loudness have been done with individual *O*s under rather carefully controlled conditions.² From the point of view of acoustical engineering, it would be profitable to compare the loudness-scale based on the estimates of a large group of *O*s in a relatively diffuse sound-field with the scales obtained from individual *O*s wearing earphones.

It was with this problem in mind that we decided to test the two methods of estimating loudness described above.

Procedure. The experiments were performed on 70 Harvard and Radcliffe undergraduate students in introductory psychology during the course of a class session. As far as we know, only two members of the class had ever before participated in an experiment on loudness. Because of their number and their inexperience in judging loudness, it was thought that this group of students would provide a suitable population on which to test the generality of the sone scale.

Experiment I. The *O*s were informed that they were to take part in two experi-

* Received for publication May 29, 1956. This research was carried out under Contract N5ori-76 between Harvard University and the United States Office of Naval Research (Project NR142-201, Report PNR-184).

¹ S. S. Stevens. The direct estimation of sensory magnitudes—loudness, this JOURNAL, 69, 1956, 1-25.

² L. B. Ham and J. D. Parkinson, Loudness and intensity relations, *J. acoust. Soc. Amer.*, 3, 1932, 511-534.

ments in the scaling of psychological magnitudes. For Experiment I, the following instructions were read to the class:

We are going to present to you through the loudspeaker a series of noises which will last for a few seconds each, with a short time-interval between successive noises. When you hear the first noise put down on your paper, opposite No. 1, any number at all which you think appropriate to represent the loudness of that noise. When you hear the second noise assign a number to it, beside No. 2 on your paper. Try to make the ratios between the different noises correspond to the ratios between the loudnesses of the noises. In other words try to make the number proportional to the loudness as you hear it. Thus, if you give the number X to the first noise, and the second noise sounds six times as loud to you as the first, assign the number 6 times X to it. Or, if the second noise sounds 15 times as soft as the first, assign the number X divided by 15 to it. Be sure to avoid looking at your neighbor's paper, as independence of judgment is critical to the experiment.

Special precautions were taken to avoid suggesting any particular number to designate the loudness of the first noise.

Eight intensive levels of white noise were presented for approximately 3 sec. each, with about 7 sec. intervening between presentations. The noises were delivered through a 12-in. loudspeaker in a bass-reflex cabinet placed at the front of the classroom. The bandwidth of the noise was limited by a low-pass filter having a cut-off at 10,000 \sim . The loudest sound presented had a sound-pressure level of approximately 105 db. (re 0.0002 dyne/cm.²) as measured on a General Radio sound-level meter located 3 ft. in front of the source. To produce and to measure sounds below this level we relied on an attenuator connected between the noise generator and an amplifier. The levels presented were 85, 100, 75, 55, 90, 105, 65, and 95 db., in that order. The sound-pressure levels of these noises were, of course, lower at points farther removed from the loudspeaker. A check with a sound-level meter showed that the levels were approximately 10 db. lower when measured in the middle of the classroom with the O s present. Since, however, the slope of the sone-scale is, as we believe, essentially invariant with sound-pressure level, the fact that our O s were exposed to different absolute levels should not be critical.

Each O recorded his numerical estimation of the loudness of each sound. The records were collected before Experiment II was begun. The results of the first experiment are discussed below.

Experiment II. For Experiment II, the following instructions were read:

In this experiment we are going to present to you pairs of noises in succession. The first of each pair is the standard, and it will stay the same throughout the experiment. The second of each pair is the variable and will differ from one pair to another. We are going to call the loudness of the standard '10.' Your task is to estimate the loudness of the variable. In other words, the question is: if the standard is called '10,' what would you call the variable? Use whatever numbers seem to you appropriate—fractions, decimals, or whole numbers. For example, if the variable sounds seven times as loud as the standard, write '70'; if the variable sounds 1/5 as loud as the standard, write '2'; if 1/20 as loud, write '1/2' or '0.5,' etc. Try not to worry about being consistent. Try to give the appropriate number to each noise regardless of what you may have called some previous one.

The sound-pressure level of the standard noise presented before each variable was 85 db., determined in the fashion described above. The levels of the variable noises were 90, 65, 105, 95, 55, 85, 100, and 75 db., in that order. The standard was in

each case presented for about 3 sec., followed by a variable of the same duration, with about 2 sec. of silence between them. Approximately 10 sec. were allowed for the Os to record their estimates before the next pair was presented.

Results. The estimates given in Experiment I were transformed by multiplying each O's judgment of the loudness of the first noise by a factor that made the resulting product equal 10, and then multiplying each of his successive judgments by the same factor. This operation puts all the judgments of loudness on a common scale. The medians of the transformed estimates of loudness for Experiment I are represented by circles in Fig. 1. In the same figure the median estimates of loudness obtained in Experiment II (in this case actual estimates) are designated by triangles. The interquartile ranges are indicated by the length of the vertical lines.

The two straight lines in Fig. 1 have the same slope as the sone-function derived by Stevens from the pooled results of about a dozen previous experiments.³ The slope of these lines is such that an increase of 10 db. in sound-pressure level corresponds to a twofold increase in estimated loudness. In view of the general conditions of our experiment, which were certainly less well controlled than those in the laboratory investigations, the close fit between our results and the line representing the sone-function is significant. The fit is apparently a little better for the data obtained in Experiment I where no standard was designated. Previous work would lead us to expect this result. Furthermore, in Experiment II the deviations from the predicted line are in the same direction as has been found in earlier experiments, i.e. when the standard is set near the middle of the range, very faint sounds appear to be underestimated and very loud ones overestimated.

This systematic effect has been interpreted by Stevens essentially as follows: In making estimates of loudness relative to a standard the O does two things: he judges the loudness of the variable with respect to the standard, but to a small extent he also tends to report something about the absolute loudness of the variable. It is as though a given estimate were influenced by two factors: the ratio of the variable to the standard and the absolute value of the variable. How this process will affect the slope of the loudness-function has been shown to depend on the level of the standard sound. In our experiment, the placing of the standard near the level of 'medium loudness' produced the expected increases in slope near the extreme ends of the function. For a more detailed discussion of the effects of various experimental procedures upon deviations of this kind the interested reader is referred to Stevens.⁴

As shown in Fig. 1 the variabilities of the judgments under these two experimental methods are roughly comparable. The overall variability is approximately the same as in the studies employing individual Os.

It is of interest to examine the numerical values assigned by the Os to the first sound in Experiment I. Table I shows the distribution of these numbers. The number 10 was highly preferred by this group of Os. Other popular numbers were 20, 5, 50, and 100.

It appears that the Os' estimates of loudness in Experiment I depend to some

³S. S. Stevens, The measurement of loudness, *J. acoust. Soc. Amer.* 27, 1955, 815-829.

⁴S. S. Stevens, this JOURNAL, *op. cit.*, 6-22.

extent upon their choice of a low or high number to represent the loudness of the first sound. Evidence of this dependency is shown in Fig. 2 where curves are plotted for subgroups of the 70 *O*s divided as follows: (1) 20 *O*s who assigned

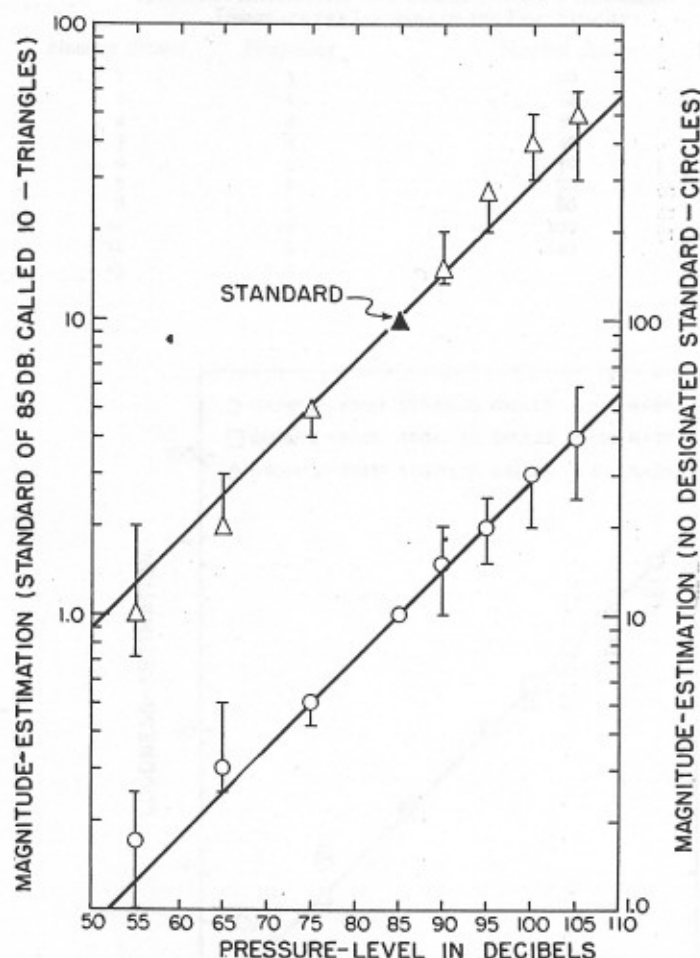


FIG. 1. MAGNITUDE-ESTIMATIONS OF THE LOUDNESS OF WHITE NOISE

Circles refer to Experiment I in which a standard was not used. Triangles refer to Experiment II in which the loudness of the standard (85 db.) was called 10. Each point is the median of the judgments of 70 *O*s. The vertical lines indicate the interquartile ranges. The straight lines show the slope of the sone-scale of loudness for a 1000~ tone.

a number smaller than 10 to the first stimulus; (2) 30 *O*s who used numbers from 10 to 30; and (3) 20 *O*s who used numbers 40 and above. The median estimates of Group 1 show the largest deviations from the predicted line at the low end of

TABLE I
FREQUENCY-DISTRIBUTION OF NUMBERS CHOSEN IN EXPERIMENT I TO
DESIGNATE THE LOUDNESS OF THE FIRST STIMULUS

Number chosen	Frequency	Number chosen	Frequency
1	3	20	7
2	1	24	1
4	4	30	1
5	5	40	3
6	3	50	7
7	2	70	2
8	2	80	1
10	18	100	6
12	2	120	1
17	1		

$N=70$

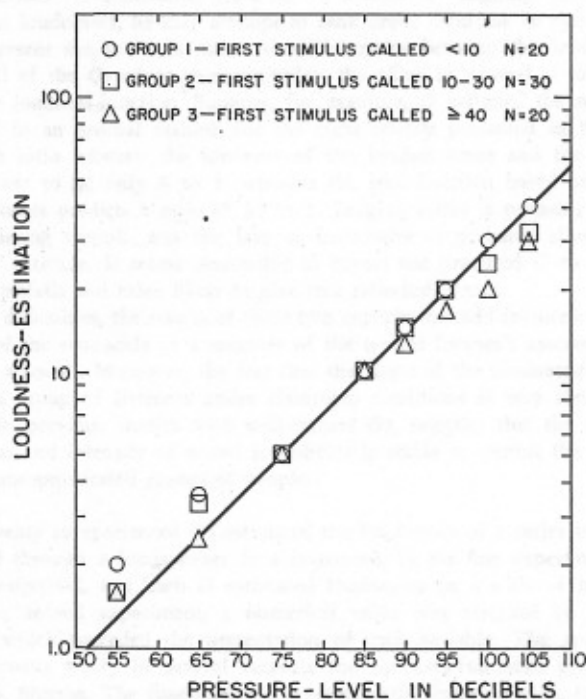


FIG. 2. EFFECT OF THE INITIALLY CHOSEN NUMBER UPON THE LOUDNESS-ESTIMATES IN EXPERIMENT I

The circles represent the median estimates of the 20 *O*s who selected a number smaller than 10 to represent the loudness of the first noise. The squares stand for the 30 *O*s whose first choice ranged from 10 to 30. The triangles stand for the 20 *O*s whose first choice was 40 and above. The straight line represents the sone-function.

the scale, whereas the medians of Group 3 deviate most markedly at the high end of the scale. Relative to the median judgments of the entire group, the *O*s who choose a small initial number tend to overestimate the fainter sounds, and the *O*s who choose a large initial number tend to underestimate the louder sounds. It would seem as though the *O*s felt some degree of constraint imposed by their relatively extreme initial choice of a number. The judgments of Group 2, who started with medium sized numbers, deviate from the line representing the sone-scale less markedly than the two other groups.

In Experiment I, in which there was no standard, the results from our unsophisticated *O*s followed a function whose slope is a little flatter than the function obtained from more experienced *O*s. It seems not unlikely that the judgments of naïve *O*s would tend in this direction, for the factors that produce errors in these experiments probably work more in one direction than in the other. One such factor may be the failure of the naïve *O* to make true ratio judgments. For example, he may not understand completely what is meant in the instructions: "make the numbers proportional" to something. As a result, instead of assigning numbers in proportion to the loudnesses, he may attempt to rank order them; or he may assign numbers to represent the intervals rather than the ratios between the loudnesses. When any or all of the *O*s adopt these attitudes, the effect in general is to flatten the slope of the loudness-function. Suppose, for example, *O* assigned the numbers 1 to 8 to stand in an ordinal fashion for the eight sounds presented in the first experiment. The ratio between the loudness of the loudest noise and the softest would then appear to be only 8 to 1, whereas the sone-function based on many previous experiments predicts a ratio of 32 to 1. Judging ratios is probably harder than merely ordering stimuli, and the lazy or inattentive *O* probably slips easily into an 'ordinal' attitude. It seems reasonable to expect the practiced *O* to be less subject to these pitfalls and more likely to give true ratio-judgments.

Despite these difficulties, the results of these two experiments add further evidence to the validity of the sone-scale as a measure of the typical listener's assessment of the loudness of a sound. Moreover, the fact that the slope of the loudness-function obtained from a group of listeners under classroom conditions is very similar to those obtained in previous studies with well-trained *O*s, suggests that the relation between loudness and intensity of sound is sufficiently stable to permit the scaling of loudness by unsophisticated groups of people.

Summary. Seventy inexperienced *O*s estimated the loudnesses of a series of white noises presented through a loudspeaker in a classroom. In the first experiment, no standard was designated, and each *O* estimated loudnesses on a scale of his own choosing. In the second experiment, a numerical value was assigned by *E* to a standard noise which preceded the presentation of each variable. The results of the two experiments verify in several respects the loudness-functions previously reported by S. S. Stevens. The functions are approximately linear on a log-log plot. In the experiment in which a standard was not designated, the slope was such that a twofold increase in loudness corresponds approximately to a 10-db. increase in intensity. Judgments were seen to depend somewhat upon *O*'s choice of a high or a low number to represent the loudness of the initial noise. The slope of the loudness-function was considerably steeper when a standard of medium loudness preceded the variable stimuli.