Repeated magnitude estimations with a variable standard: Sequential effects and other properties*

LAWRENCE M. WARD†

Rutgers University, New Brunswick, New Jersey 08903

Twelve Ss made magnitude estimations of the loudness of each one of a sequence of pure tones according to the rule $R(N) = R(N - 1) \cdot [S(N)/S(N - 1)]$, where $R(N)$ is the response on Trial N, $R(N - 1)$ is the response on Trial N − 1, and $S(N)/S(N - 1)$ is the judged ratio of the "loudness" of the pure tone presented on Trial N to that of the pure tone presented on Trial N − 1. It was found that these magnitude estimations were assimilated toward the immediately preceding stimuli as far as five trials back in the sequence of stimuli. In addition, ratio judgments were consistently asymmetric and the data displayed a form of "time order error." In all cases, there are similar effects displayed in category judgment data. These and other data imply that at least some kinds of magnitude estimations may involve a judgment of the "difference" or "distance" between pairs of stimuli as a first step in the production of the response required by the judgment situation.

This paper presents data concerning various properties of magnitude estimations of the loudness of pure tones which have implications for at least two questions: (1) the similarity of magnitude estimation data to those of absolute category judgment, and (2) a theory of the processes underlying the performance of the magnitude estimation task.

Several investigators conceive of psychophysical judgment as a matching task. Attneave (1964), Ekman (1964), and Stevens (1966) have all suggested that Stevens's magnitude estimation task may be viewed as a cross-modality matching task in which stimulus sensation is matched to "number sensation" to produce the response; Ward and Lockhead (1971) have proposed that the classical absolute category judgment task is a task in which an internal representation of a stimulus is matched to a remembered response scale. Stevens (1966) believes *all* judgment to be a form of matching. In most psychophysical judgment, a particular stimulus range is matched by the S to a particular response domain (e.g., the numbers 1-10) according to a particular rule. In the method of magnitude estimation, the S is instructed to assign a number to a stimulus by considering the ratio between that stimulus and another, already numbered, stimulus. A great many Es have found that, for prothetic stimulus dimensions (Stevens & Galanter, 1957), responses produced by the method of magnitude estimation generally are a power function of the stimulus magnitude measured in physical units. Stevens prefers to express this relation as $R = a(S - S_0)^n$, where R is the response, S is the stimulus magnitude, S_0 is a stimulus scale translation to correct for distortions near absolute threshold [sometimes interpreted as the

"effective threshold" (Stevens & Stevens, 1963)], a is a constant representing the unit of the scale, and n is the exponent of the power function. On the other hand, in the method of category judgment, the S is instructed to assign stimuli to categories which represent intervals on the stimulus dimension. For prothetic continua, the category scale seems to be neither logarithmic with stimulus magnitude (Fechner's law) nor linear with the magnitude estimation scale (Stevens & Galanter, 1957). It is almost always concave downwards when plotted against the magnitude estimation scale, but is also well fit by a power function (Marks, 1968; Curtis, 1970; Ward, 1972).

There is also evidence that the two methods of judgment are subject to many similar kinds of judgmental bias. Poulton (1968) found six major variables that affect the exponent of a power function fitted to magnitude estimation data. They are: (1) stimulus range, (2) the distance of the stimuli from the "absolute" threshold, (3) the position of the standard in the stimulus ensemble, (4) the distance of the first variable from the standard, (5) whether the S has available a finite or an infinite set of numbers with which to respond, and (6) the size of the modulus (the number that is assigned to the standard in this type of experiment). Situational variables known to affect responses in the method of category judgment are: (1) the range and frequency distribution of the stimuli (Parducci, 1965), (2) interpolated anchor stimuli (Helson, 1959), (3) the position of the standard in the stimulus ensemble (Helson, 1959), and (4) the spacing of the stimuli (Stevens & Galanter, 1957). Also, Holland and Lockhead (1968), Ward and Lockhead (1970, 1971), and Ward (1972) have shown that absolute category judgments of various types are assimilated to the value of the immediately previous stimulus and response and are assimilated to or contrasted with those stimuli and responses further back in the sequence, depending on the presence or absence of information feedback.

The present study is designed to provide data which will allow stronger conclusions as to the similarity of the two forms of judgment and the judgmental processes underlying magnitude estimations. In this study, Ss judged loudness levels by a commonly used method of magnitude estimation. The results were analyzed for sequential effects and several other psychophysical properties, since these have implications for the judgmental processes underlying category judgments and presumably could perform the same function for magnitude estimations. To my knowledge, such an analysis has never before been performed on magnitude estimation data, nor have previous studies analyzed changes in response strategies over a great many magnitude estimation trials.

METHOD

Subjects

Twelve undergraduate volunteers, naive to psychophysical judgment, served as paid Ss for the study. All had no known hearing defects and all performed adequately on the experimental task.

Apparatus and Stimuli

The stimuli were 500-msec-duration 1,000-Hz sinusoids generated by an oscillator (Hewlett-Packard 200-CD). A tape reader selected one of 10 different attenuators on each trial, and the resulting amplitude was delivered dioticaIly through Superex ST-S headphones. The 10 amplitudes delivered to the phones were: 1.74, 2.76, 4.38, 6.94, 11.0, 17.4, 27.6, 43.8, 69.4, and 110.0 mV. These values were measured on a vacuum tube voltmeter, with an accuracy of 1% of the reading, placed in parallel to the headphones. The attenuators were reset each day to assure that the voltages were always the same. The stimuli were chosen to be equally spaced on a log scale, and adjacent pairs were 4 dB apart. The total stimulus range was thus 36 dB, with the softest stimulus about 40 dB re .0002 dynes/cm^2. The stimuli were gated by a timer-activated electronic switch to minimize transients.

Design

Each of the 12 Ss produced 500 magnitude estimations (50 per stimulus) on each of 2 days, for a total of 1,000 judgments per S and 12,000 magnitude estimations in all. There were five random orders of the 500 stimuli presented in a single day. Each S received the stimuli in one order for his first day and in a different order for his second day. All combinations of the five orders were used about equally often, and there were no differences in the data depending on which combination was received.

It was decided to have the Ss do a magnitude estimation task which was more analogous to the typical category judgment task than the task typically used. Typically, on each trial, the S is presented with two stimuli: a standard numbered stimulus and a variable unnumbered stimulus. The task of the S is to assign a number to the variable in such a way that the ratio of his response to the number (modulus) of the standard reflects what he believes to be the sensation ratio between the variable and the standard stimuli. This technique is analogous to the classical technique of comparative category judgment. However, since most category scales compared to magnitude scales have been of the absolute category judgment variety, with no standard stimulus present on each trial, it was thought better to use a

magnitude estimation technique more analogous to this task. Thus. in this study, the S was asked to judge the ratio between the present and previous stimulus in the stimulus series and to assign a number to the present stimulus so that the ratio of the present response to the previous response would reflect what he believed to be the ratio between the sensation values of the present and previous stimuli. For the very first stimulus, the E presented a particular stimulus and gave it a particular number (modulus). Thus. in essence, this was a task with a variable standard, where the standard on each trial was the stimulus that had been juded on the immediately previous trial. More specifically, the equation $R(N) = R(N - 1) \cdot [S(N)/S(N - 1)]$, where $R(N)$ is the response on Trial N. and $S(N)$ is the stimulus on that trial, describes the task of the S. Other Es have had success with this method (Curtis, 1970; Curtis, Attneave, & Harrington, 1968; Curtis & Fox, 1969), and it should also be particularly subject to sequential biases.

Before starting the loudness judgment task, the S judged line lengths in the same way that he would be judging loudnesses. This procedure has been recommended to minimize biases associated with the way Ss use numbers (Stevens, 1966). The range of line lengths was from 1 to 60 in. It was emphasized that there was no restriction on the numbers that the S could assign to stimuli except his own ratio judgments. Stimuli (line lengths) were presented that led the S to use very large numbers (in the hundreds) and very small numbers (close to zero). It was shown that there are a great many numbers between zero and one and that they were all usable if necessary (except for the zero). All Ss learned the task within a few trials.

After the E was satisfied that the S understood his task, he told the S that the first loudness stimulus would be presented by the E and that it was to be called "10." Ss were told not to take any excessively long (more than a few seconds) breaks during the session, as the accuracy of their ratio judgments depended on the accuracy of their memories of the previous stimulus. Ss worked at their own rates, about 5-7 sec per trial.

RESULTS

Psychophysical Functions

The psychophysical functions of each S were highly similar, and so only the function of the pooled data of the 12 Ss is plotted in log-log coordinates in Fig. 1. In this figure, the logarithms of the *geometric* mean responses are plotted, since the distributions of responses to the various stimuli were highly skewed, as is often the case for magnitude estimations. The psychophysical function is strikingly close to a straight line in the log-log plot. However, there is a sinusoidal bowing of the points about the straight line, similar to that found in category judgments (Ward, 1972).

The fact that the points lie on a straight line in Fig. 1 means that a power function would provide a good fit to the data, with the exponent of the power function being the slope of the best fitting straight line. Power functions of the form $R = aS^n + b$ (the form preferred by some investigators, e.g., McGill, 1960; Fagot & Stewart, 1968) and $R = a(S - S_0)^n$ (holding S_0 constant at 0) were fitted to the data for each S and to the pooled data by the method of least squares (using a modification of a computer program supplied by Dwight Curtis).

For 2 of the 12 Ss, the program was unable to reach a

satisfactory solution, indicating a poor fit to a power function. However, for the other 10 Ss, and for the pooled data, highly satisfactory solutions were achieved for both forms of power function. Since the functions were similar for the various Ss, only the functions fitted to the pooled data are reported here. The best fitting functions for the pooled data were $R = .58S^{.55}$ and $R = .27S^{.70} + .66$, with standard errors of estimate of .22 and .15, respectively. It may be seen that the power functions provide an excellent description of the data. However, the exponents of the two forms of the power function do differ somewhat. This is to be expected, since the two forms have different numbers of parameters, and is true also when the two forms are fitted to category judgment data (Ward, 1972). The average exponents for magnitude estimation are close to the typical exponent of .6 (for binaural stimuli measured in pressure units) reported by Stevens (1956). That exponent (.6) is the average of a great number of experiments in a standardized situation that differs from the present one in several ways. The value of .6 does lie about two standard errors (.02 and .05) away from the exponents estimated from the pooled magnitude estimation data.

Time Order Error

Any systematic deviation from a perfect match of the center of the stimulus range to the center of the response domain has classically been called "time order error." The data were analyzed for this property of classical psychophysical judgment. In the present experiment, Stimulus No. 5 was always the first stimulus presented, and it was always labeled "10" at that time. The number "10" is said to be the "modulus" in an experiment of this type. The response to this stimulus should continue to be "10" whenever it occurs (except for random error) if the modulus does not change over time. Es do not generally report whether the modulus changed in magnitude estimation studies, and most experimental situations are designed to minimize effects of this sort. One typical way of doing this is to present a standard, whose label is always the modulus, on every other trial so that the variable stimulus is always compared with it. This would assure that the modulus remained constant.

In the present experiment, however, no attempt of this sort was made. The theoretical "standard" stimulus (No. 5, 11.0 mV) occurred at random in the sequence of stimuli. If a tendency toward a time order error was operating in this situation, it should manifest itself by a systematic change over trials in the response to Stimulus No. 5. In 11 of the 12 Ss, it was found that the geometric mean response to Stimulus 5 was lower than 10, indicating a decrease of the modulus from the beginning value. For the pooled data, the geometric mean response to Stimulus 5 was 2.08, a drop of almost an order of magnitude in the modulus. This, of course,
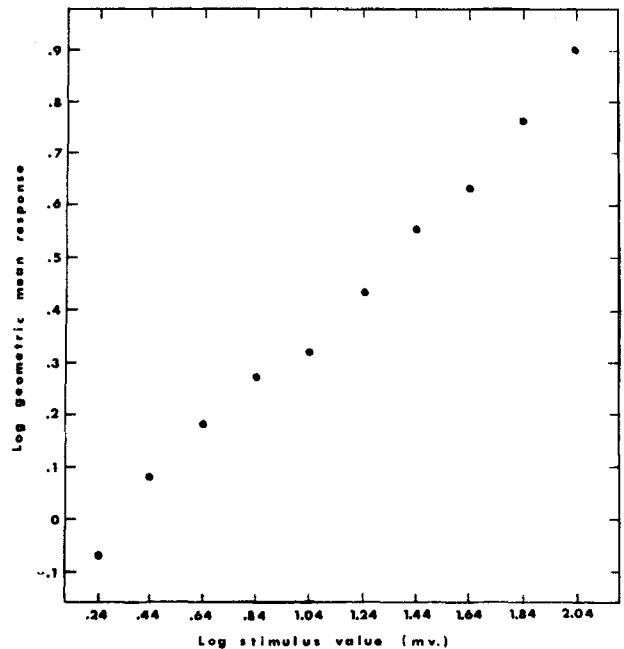

Fig. 1. Psychophysical function of the pooled data of 12 Ss for magnitude estimations of loudnesses. There are 1,200 observations per point.

represents only the average modulus over all 1,000 trials for each S; the actual modulus drop must have been greater. Ward (1971) analyzed this modulus drop in detail and found drops of several orders of magnitude over only 500 trials. He found that drops in the first few hundred trials were generally partially compensated for in the last few hundred trials, but the modulus of some Ss showed a continuous slow change in the negative direction over the entire 500 trials. For the present study, it can be noted that there seems to be something similar to a negative time order error present in these data.

Sequential Effects

Figure 2 shows the average response to the stimulus on Trial N as a function of the stimulus on Trial N − 1. The data are collapsed across Ss and pairs of adjacent stimuli. There are approximately 480 observations per point. The data of each S are identical in form to the average data in the figure, although of course more variable. It is apparent that for the most part, the higher the value of the stimulus on Trial N − 1, the higher the response to the stimulus on Trial N. These data imply that the response to the stimulus on Trial N is assimilated toward the value of the stimulus on Trial N − 1. This is identical in form to the findings of Holland and Lockhead (1968), Ward and Lockhead (1970, 1971), and Ward (1972) for absolute category judgments. The finding is now extended to the magnitude estimation situation.
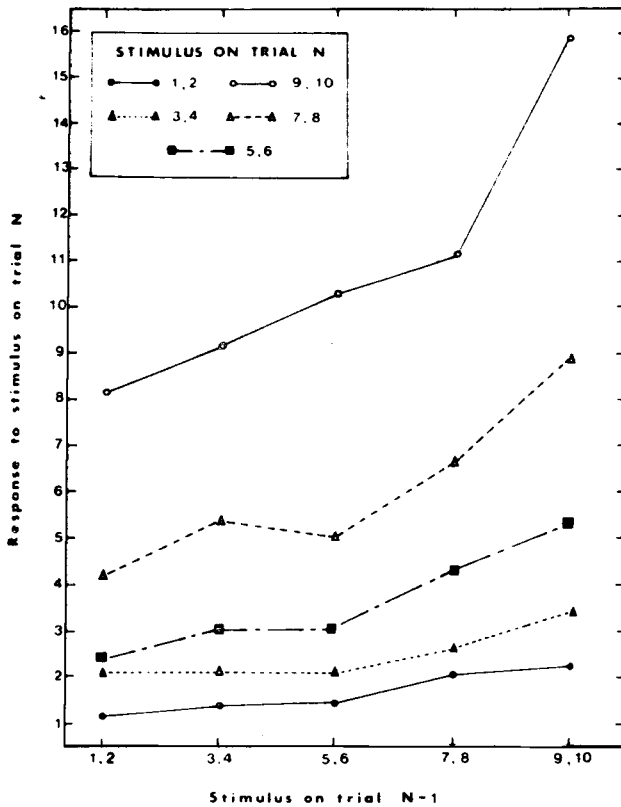
Fig. 2. Average response to the stimulus on Trial N as a function of the stimulus on Trial N − 1. The data are the geometric mean responses of 12 Ss. There are approximately 480 observations per point.

a particular value. There are about 1,200 responses per entry at each K level. Rank-order correlation coefficients of the stimulus value on Trial N − K with the average response over all stimuli on Trial N are shown to indicate the direction and reliability of the effects. The magnitude of the correlation coefficient only indicates how reliable the ordering of the stimuli and responses is, not the magnitude of the effect itself. The latter can be seen by inspecting the difference between the average response when the stimulus on Trial N − K was small, compared with the average response when that stimulus was large. A positive correlation indicates assimilation of the responses on Trial N to the stimulus on Trial N − K; a negative correlation indicates contrast. Inspection of the correlations in Table 1 reveals that responses on Trial N are clearly assimilated to stimulus values on Trials N − 1 through N − 5. There seem to be no reliable effects of stimuli 6 to 10 trials back in the sequence.

The overall sequential effects for the magnitude estimation situation are quite similar to those in a similar category judgment situation (Ward, 1972), with the exception that assimilation effects for magnitude estimations seem to extend at least as far back as five trials, and they do not seem to change to contrast at that point, but rather taper off to nothing. The analysis for the effects of previous responses was not done for the data since responses in the magnitude estimation situation are not limited and there are generally too few cases of each for such an analysis.

Table 1 shows responses on Trial N as a function of the stimulus on Trial N − 1 in a different way from Fig. 2, and extends the analysis of sequential effects to stimuli as far as 10 trials back in the stimulus sequence. Each entry is the average of all Ss' responses to all stimuli on Trial N when the stimulus on Trial N − K was

Symmetry Analysis

In this experiment, Ss were asked to produce responses according to the rule $R(N) = R(N − 1) \cdot [S(N)/S(N − 1)]$. We can get an estimate of the S's judgment of the ratio of the stimulus on Trial N [S(N)] to the stimulus on Trial N − 1 [S(N − 1)] by finding the ratio of the response on Trial N [R(N)] to the response on Trial N − 1 [R(N − 1)]. If the S was doing the task

Table 1
Geometric Mean Magnitude Estimations on Trial N as a Function of the Stimulus on Trial N − K (K = 1,10)

Stimulus on Trial N − K	Geometric Mean Response on Trial N									
	K = 1	K = 2	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8	K = 9	K = 10
1	3.18	3.91	4.00	3.95	4.26	4.46	4.52	4.61	4.89	4.94
2	4.02	4.26	4.44	4.53	4.40	4.77	4.50	4.37	4.37	4.68
3	3.97	4.48	4.54	4.77	4.62	4.88	4.98	5.22	4.88	4.76
4	4.41	4.50	4.42	4.66	4.55	4.49	4.83	5.17	4.66	4.94
5	4.02	3.90	4.75	4.24	4.64	4.35	4.68	4.37	4.78	4.73
6	4.74	5.45	4.48	5.24	4.71	4.49	5.28	4.48	4.55	4.72
7	5.14	4.86	4.88	4.80	5.80	4.43	4.50	4.47	4.68	4.37
8	5.57	5.43	4.68	4.61	4.53	4.86	4.76	4.59	5.05	4.86
9	6.07	5.59	5.34	5.32	4.96	5.37	4.84	5.33	5.13	4.71
10	8.20	6.57	5.82	5.76	5.86	5.39	5.44	5.57	5.57	5.41
r′	.961	.885	.867	.740	.782	.458	.373	.391	.564	−.106

Note—Also shown are the rank-order correlation coefficients between the geometric mean responses on Trial N and the stimulus values on Trial N − K. There are 9 dfs for each r′. The two-tailed critical value for r′ at the .05 level of significance is .653 and at the .01 level is .860.

Table 2
Average Log R(N)/R(N − 1) for All Combinations of S(N) and S(N − 1) for Magnitude Estimation Data

Stimulus Value on Trial N − 1	Stimulus Value on Trial N									
	1	2	3	4	5	6	7	8	9	10
1	−.17	−.06	.04	.14	.29	.30	.42	.56	.70	.75
2	−.28	−.10	−.05	.05	.18	.23	.34	.47	.52	.72
3	−.35	−.19	−.10	.00	.10	.16	.29	.38	.45	.64
4	−.37	−.26	−.13	−.05	.02	.08	.21	.35	.49	.55
5	−.47	−.37	−.25	−.14	−.04	.04	.12	.33	.43	.51
6	−.50	−.38	−.30	−.14	−.05	.02	.07	.22	.33	.50
7	−.55	−.41	−.33	−.25	−.15	−.02	.08	.10	.25	.42
8	−.57	−.41	−.36	−.34	−.23	−.14	−.03	.06	.18	.36
9	−.66	−.53	−.42	−.37	−.22	−.19	−.10	−.03	.13	.31
10	−.70	−.57	−.47	−.39	−.30	−.26	−.13	−.06	−.04	.17

Note—Each entry is the average of about 120 observations.

consistently and well, then the ratio of responses for a particular pair of stimuli. N − 1 and N, in the stimulus series. say 1 and then 10, would be the reciprocal of the ratio of the responses when that pair of stimuli occurred in the reverse order, 10 and then 1. If this situation obtained, the geometric mean of the two fractions in question would be equal to one. For the purposes of the present paper, this situation will be defined as "symmetry." It is possible, in view of the variability of responses, that even when an average of a large number of occurrences of a particular pair of stimuli is taken, the geometric mean fraction will not be equal to one. This situation is considered to be "asymmetry" of the fractions. When the geometric mean fraction is less than one, this is defined as "negative asymmetry"; when it is greater than one, as "positive asymmetry." Asymmetry can be interpreted as judgment error; positive asymmetry implies overestimation of either or both of the stimuli involved, negative asymmetry implies underestimation of those stimuli.

Table 2 presents the results of the analysis of fractions for asymmetry. Each entry in the table is the average over all 12 Ss of the log geometric mean fraction estimated from adjacent responses for a particular ordered pair of stimuli. There are approximately 120 observations per entry. Since the averages of the *logarithms* of the fractions are entered in the table, the average of a complementary pair (e.g.. 1-10, 10-1) of the entries would have to be equal to *zero* for symmetry to obtain. If perfect symmetry obtained for the entire matrix. the main diagonal (stimulus repeats) would be all zeros and the rest of the matrix would be symmetric about the main diagonal. with the exception of a change of sign from positive in the upper right to negative in the lower left. Inspection of the matrix reveals that symmetry does not generally obtain.

In order to check for a pattern in the symmetry violations, the means of the two log fractions (one for each order of presentation) of each of the 45 off-diagonal complementary pairs was calculated. These numbers, along with the uncombined log fractions of the repeated pairs, are referred to hereafter as "symmetry

numbers." There were 55 symmetry numbers for analysis (45 from the off-diagonal complementary pairs and 10 from the repeats). The Spearman rank order correlation coefficient (r') between the symmetry numbers and several relevant factors was calculated. The symmetry numbers are correlated positively with a measure of the total strength of the two stimuli (the sum of the stimulus magnitudes in millivolts) involved with each symmetry number (r' = .947, df = 54, p ≅ 10⁻⁵). The symmetry numbers range from quite negative for stimulus pairs with a low total strength to quite positive for stimulus pairs with a high total strength. The greater the total strength of the stimulus pairs whose symmetry number is considered, the more positive the asymmetry, i.e., the more the ratios are overestimated. The symmetry numbers are also correlated with other factors (e.g., the difference in millivolts between the stimulus pair, the ratio between the pair, the stimulus value. etc.), but these in turn are correlated with total strength. Total strength handles more variance as a single factor than does any other variable. The correlation coefficients of the individual Ss were in all cases in the same direction as the coefficient for the average data, although generally slightly closer to zero because of the greater variability in the data of the individual Ss.

This analysis is not really relevant to category judgments. since in that task Ss are not instructed to pay attention to the ratios between stimuli. However. in view of the similarity of the data from the two methods so far. the analysis was performed on category judgments gathered by Ward (1972) in the same fashion as for the magnitude estimation data above. Table 3 presents the results for this analysis. It can be seen that the results take exactly the same form for category judgments as for magnitude estimations. The matrix is not symmetric. Once again. the calculated symmetry numbers correlate positively with the total strength of the relevant pairs of stimuli (r' = .815. df = 54. p ≅ 10⁻⁵).

DISCUSSION

The present study has investigated in detail some of

Table 3
Average Log R(N)/R(N − 1) for All Combinations of S(N) and S(N − 1) for the Category Judgment Data of Ward (1972)

Stimulus Value on Trial N − 1	Stimulus Value on Trial N									
	1	2	3	4	5	6	7	8	9	10
1	−.08	.02	.06	.18	.27	.35	.46	.57	.65	.82
2	−.13	−.06	−.01	.05	.12	.26	.34	.45	.57	.65
3	−.18	−.15	−.05	.02	.05	.17	.25	.55	.41	.54
4	−.28	−.18	−.11	−.01	.02	.11	.15	.26	.47	.52
5	−.34	−.22	−.21	−.11	−.06	.06	.14	.20	.47	.50
6	−.40	−.24	−.18	−.14	−.06	.01	.09	.10	.26	.40
7	−.47	−.31	−.24	−.20	−.13	−.08	.04	.07	.19	.29
8	−.58	−.36	−.33	−.24	−.20	−.21	−.06	−.01	.13	.23
9	−.55	−.52	−.38	−.37	−.31	−.23	−.16	−.07	.03	.10
10	−.78	−.65	−.50	−.41	−.42	−.30	−.22	−.16	−.05	.02

Note−Each entry is the average of about 120 observations.

the properties of magnitude estimations of loudnesses. It has been found that although instructions to Ss in the magnitude estimation situation are different from those in a typical category judgment situation, the basic psychophysical data are similar in many ways. The data of both methods are well fit by a power function. Both methods show the classical effects of "time order error." Responses in both situations are subject to the effects of the previous sequence of stimuli. And, in both situations, symmetry of ratios between stimuli does not generally obtain, and the asymmetry of ratios is correlated with the total strength of the relevant stimulus pairs. These results indicate a striking similarity of the judgment processes in magnitude estimation to those in category judgment. In what follows, this similarity is investigated in more detail and a tentative model of psychophysical judgment applicable to both magnitude estimations and category judgments is suggested.

Power Law

This study represents yet another confirmation of the power law, since the magnitude estimations were nicely fit by a power function. Such power functions are quite valuable as data descriptions whose parameters can be observed for invariances in the search for understanding of the processes underlying psychophysical judgment. An interesting example of this use of power functions is the observation that the exponents of power functions fitted to magnitude estimations are generally about twice as large as the corresponding exponents for category judgments (Ward, 1971). The present study confirms this finding for loudnesses; the exponents (.55 and .70) for the pooled magnitude estimation data are almost exactly twice those (.27 and .39) of similar data in a category judgment experiment (Ward, 1972). Since there are no differences in the form of the data for the two methods, the explanation for this invariance may lie in the actual numbers used for responses. Es typically limit the response domain in category judgments to 7-10 categories, and thus to Response Nos. 1-7 or 1-10, about

1 log unit. In magnitude estimation, the S is free to use any positive real number as a response. Given a modulus of 10, he generally uses numbers between .5 and 50 or so, about two log units as compared to the approximately one log unit available for use in category judgment situations. Teghtsoonian (1971) points out. that, from the simple form of the power law, it follows that $n = \log R\psi/\log R\varphi$, where n is the exponent of the power function, $R\varphi$ is the ratio of the largest to the smallest stimulus magnitude, and $R\psi$ is the ratio of corresponding sensory magnitudes. If we assume that $\log R\varphi$ is relatively constant [Ward (1971) used only studies in which it was the same, or average exponents from a great many studies in which it varied over a fixed range], then we get the following when we insert the respective ranges of response estimated above: $n_{cj} = \log R\psi/k = 1/k$, and $n_{me} = \log R\psi/k = [2/k] = 2[1/k]$. Therefore, the exponent of the magnitude estimation power function is twice as large as the exponent of the category judgment power function, no matter what the continuum of interest, simply because of the relationship of the range of responses allowed. These assertions can, of course, be tested by varying the range of responses in a category judgment experiment to see whether the exponent can be brought more into line with the magnitude estimation exponent. Stevens and Galanter (1957) and Gibson and Tomko (1972) have provided evidence that the category scale approaches linearity with the magnitude scale as the number and range of responses allowed in the category judgment situation approaches that of the magnitude estimation situation.

Sequential Effects

It was found in this study that the response to a stimulus in a magnitude estimation situation is assimilated to the values of the stimuli as far as five trials back in the sequence of stimuli presented for judgment. From this point on, there seems to be little effect of any kind of the previous stimuli on responses. These sequential effects are very similar in form to those found

for category judgments in the absence of an E-induced identification function (Ward, 1972), with the exception that the contrast effects found for category judgments after the assimilation effects had dwindled were not found in the present study for magnitude estimations. Ward and Lockhead (1971) proposed that the sequential effects found in category judgments were due to two response system processes employed in producing a response in the presence of uncertainty of the "correct" response. The first is a process of guessing more small and fewer large differences between succeeding stimuli than in fact occur in a random sequence of equiprobable stimuli. The S seems to follow a guessing strategy between the extremes of probability matching and maximizing with respect to the expected distribution of differences in category steps between stimuli adjacent in the sequence. Ward and Lockhead (1971) show how such a guessing strategy would necessarily result in assimilation of responses to the previous stimulus and response; the effect is compounded to other stimuli and responses in the sequence if the previous response is used as an estimate of the value of the previous stimulus. The second is a process of trying to use the available responses equally often, which would result in contrast with the previous stimuli and responses. Presumably, the assimilation and contrast observed in category judgments made in the absence of an E-induced identification function (Ward, 1972) arise from these sources as well.

It is quite possible that the tendency to use responses that are closer together than demanded by the stimuli could be operating to produce the assimilation in the magnitude estimation situation. This tendency would be expected if the S was paying attention to the differences between pairs of stimuli and had learned the expected distribution of differences between stimuli in a random sequence. No contrast effects were observed in the magnitude estimation data. This is presumably because, in a situation where the potential response set is practically unlimited, there is no tendency to use all responses equally often, and thus no contrast. And, in fact, the data of this study show that responses in the magnitude estimation situation definitely were not used with equal frequency. The study by Baird, Lewis, and Romer (1970) confirms this assertion for magnitude estimation.

The fact that both judgment situations give rise to assimilative effects implies that the locus of the bias lies in a fundamental judgment process common to the two situations. Ss are judging pairs of stimuli as more similar than they would be judged in the absence of this bias. This could mean that the memories or internal representations of the previous and present stimulus are assimilated toward each other. Since Ward and Lockhead (1971) found the same biases operating in a guessing situation, however, it is more likely that the underestimation results from aberrations in the Ss' construction of some underlying judgment scale (see next section). The source of the aberrations may lie in the Ss' attempts to maximize with respect to the expected distribution of differences between previous and present stimuli.

Symmetry Analysis

It was shown that the geometric means of ratios of complementary pairs of stimuli, the symmetry numbers, were not generally equal to one, i.e., the ratio judgments were not generally symmetric for the data of either experiment. The pattern of symmetry violations was shown to be identical for magnitude estimations and category judgments; asymmetry changed from quite negative through zero to quite positive as the total strength of the stimulus pair, responses to which were used to calculate the ratios, increased. Why this happened will be the subject of another paper. For the present, it is observed that the result implies similar underlying processes of judgment in the two methods. Since category judgment presumably forced Ss to ignore ratios (Stevens, 1966), it seems likely that the judgment process common to both methods is a process of category or interval judgment.

I propose that a process of category judgment of differences (or distances) between pairs of stimuli underlies the two types of judgment. Distances and ratios between pairs of stimuli in a stimulus set are monotonically related; in fact, ratios are differences in the logarithmic transform domain. It is possible that, in the magnitude estimation situation, Ss first make category judgments of the distance or difference between the immediately previous stimulus and the present stimulus. They would then have to transform this distance or difference judgment into another scale representing ratios between stimuli, a process which could be represented mathematically by the operation of exponentiation. The transform operation proposed here is a crude one, in which a distance judged "large," for example, would be transformed to a "large" ratio, represented by a number that the S considered to be "large." The number so obtained would then be used in the production of a response according to the multiplicative rule of judgment of that situation. In the category judgment situation, judgment of the difference between the previous and present stimulus (represented by an appropriate number of category units) along with knowledge of the previous stimulus value (or an estimate based on the previous response) could be used to obtain a response to the present stimulus by algebraic addition. Thus, the route from the basic difference judgment to the actual responses would differ in the two methods according to the rule of judgment employed, and the rule of judgment defines the range and number of responses allowed. This is consistent with the explanation above of the source of the difference in exponents between the two methods for the same continuum. In both types of judgment, underestimation of these differences due to guessing strategies, etc.,

would lead to the sequential effects discussed above.

The above proposal is very similar to one put forth by Torgerson (1961), who argued that Ss can only perceive a single fundamental relation between a pair of stimuli and that they interpret it according to the rules of the judgment situation. I am proposing that the relation perceived is a difference or distance, that it is always judged in terms of crude categories constructed for the particular stimulus situation, and that the difference in interpretation of the basic sensory event comes only after the fundamental judgment process has been completed. Thus, the difference between the two methods is argued to be at a higher level of information processing than that suggested by Torgerson (1961).

REFERENCES

Attneave, F. Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 4. New York: McGraw-Hill, 1964. Pp. 619-659.

Baird, J. D., Lewis, C., & Romer, D. Relative frequencies of numerical responses in ratio estimation. Perception & Psychophysics, 1970, 8, 358-362.

Curtis, D. W. Magnitude estimations and category judgments of brightness and brightness intervals: A two-stage interpretation. Journal of Experimental Psychology, 1970, 83, 201-208.

Curtis, D. W., Attneave, F., & Harrington, T. L. A test of a two-stage model of magnitude judgment. Perception & Psychophysics, 1968, 3, 25-31.

Curtis, D. W., & Fox, B. E. Direct quantitative judgments of sums and a two-stage model for psychophysical judgment. Perception & Psychophysics, 1969, 5, 89-93.

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

Fagot, R. F., & Stewart, M. R. An experimental comparison of stimulus and response translated power functions for brightness. Perception & Psychophysics, 1968, 3, 297-305.

Gibson, R. H., & Tomko, D. L. The relation between category and magnitude estimates of tactile intensity. Perception & Psychophysics, 1972, 12, 135-138.

Helson, H. Adaptation-level theory. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 1. New York: McGraw-Hill, 1959.

Holland, M. K., & Lockhead, G. R. Sequential effects in absolute judgments of loudness. Perception & Psychophysics, 1968, 3, 409-414.

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

McGill, W. The slope of the loudness function. In H. Gulicksen and S. Messick (Eds.), *Psychological scaling: Theory and applications.* New York: Wiley, 1960.

Parducci, A. Category judgment: A range-frequency model. Psychological Review, 1965, 72, 407-418.

Poulton, E. C. The new psychophysics: Six models for magnitude estimation. Psychological Bulletin, 1968, 69, 1-19.

Stevens, S. S. The direct estimation of sensory magnitudes–loudness. American Journal of Psychology, 1956, 69, 1-25.

Stevens, S. S. On the operation known as judgment. American Scientist, 1966, 54, 385-401.

Stevens, S. S., & Galanter, E. Ratio scales and category scales for a dozen perceptual continua. Journal of Experimental Psychology, 1957, 54, 377-411.

Stevens, J. C., & Stevens, S. S. Brightness functions: Effects of adaptation. Journal of the Optical Society of America, 1963, 53, 375-385.

Teghtsoonian, R. On the exponents in Stevens' law and the constant in Ekman's law. Psychological Review, 1971, 78, 71-80.

Torgerson, W. S. Distances and ratios in psychophysical scaling. Acta Psychologica, 1961, 19, 201-205.

Ward, L. M. Some psychophysical properties of category judgments and magnitude estimations. Doctoral dissertation, Duke University, 1971.

Ward, L. M. Category judgments of loudnesses in the absence of an experimenter-induced identification function: Sequential effects and power function fit. Journal of Experimental Psychology, 1972, 94, 179-184.

Ward, L. M., & Lockhead, G. R. Sequential effects and memory in category judgments. Journal of Experimental Psychology, 1970, 84, 27-34.

Ward, L. M., & Lockhead, G. R. Response system processes in absolute judgment. Perception & Psychophysics, 1971, 9, 73-78.