Journal of Experimental Psychology:
Human Perception and Performance
1994, Vol. 20, No. 2, 372-381

0096-1523/94/$3.00

A Dynamic Theory of Proportional Judgment: Context and Judgment of Length, Heaviness, and Roughness

Lawrence T. DeCarlo

Subjects judged the length of lines, the heaviness of weights, or the roughness of sandpaper in 2 conditions. In 1 condition, they were instructed to make all their judgments relative to a long-term reference point, which consisted of a reference response and sensation. In the other condition, they were told to use a short-term reference point, namely, the response and sensation of the previous trial. A dynamic model of proportional judgment (L. T. DeCarlo, 1989/1990) predicts that the autocorrelation of successive responses will be larger for the latter instructions. This prediction was confirmed for the 3 continua. In addition, fits of a recently proposed dynamic regression model show that there is little or no effect of the previous stimulus intensity on the current response, whereas the results for an earlier model suggest a large contrast effect. The theory and experiments provide insight into judgmental and sensory processes in magnitude scaling.

In the typical magnitude scaling experiment, different intensities of a stimulus are presented in a random (or pseudorandom) order, and the subject's task is to indicate how intense his or her resulting sensations are. It has long been recognized that, even though successive stimuli are by design uncorrelated, subjects' responses are not. A typical finding, for example, is that the response on the current trial is correlated with the response of the previous trial. The responses are said to be autocorrelated. The study of these "sequential effects" has been an active area of inquiry (e.g., DeCarlo, 1992; DeCarlo & Cross, 1990; Jesteadt, Luce, & Green, 1977; Morris & Rule, 1988; Schifferstein & Frijters, 1992; Ward, 1982, 1987).

The present article focuses on one source of sequential effects—the relativity of judgment. In particular, a basic implication of a dynamic model of proportional judgment (DeCarlo, 1989/1990) is that autocorrelation arises because of the influence of different frames of reference on judgment. Evidence in favor of the model has been provided by experiments where the instructions were varied so as to affect the relative influence of the different frames. For example, the magnitude of the observed autocorrelation has been shown to vary with the instructions for magnitude estimation (ME) of loudness and area, and cross-modality matching (CMM) of line-length to loudness (DeCarlo, 1989/1990; DeCarlo & Cross, 1990).

The purpose of the present research is to determine how different frames of reference affect judgment for three continua: the length of lines, the heaviness of lifted weights, and the roughness of sandpaper. In each case, the instructions are varied in a within-subjects design and the effect on autocorrelation is examined. It is also determined if there is an effect of the previous stimulus intensity on the current response. The article begins with a discussion of the judgmental model that motivated the experiments.

Judgment as a Process of Comparison

Subjects in magnitude scaling experiments are typically instructed to make proportional judgments of their sensation magnitudes. Although this seems fairly straightforward, it is important to recognize that there is more than one way to perform the task. One approach, for example, is to assign a reference response R_0 (a modulus) to a reference sensation ψ_0 (a standard). On each trial, subjects compare their current sensation ψ_t to their reference sensation ψ_0. They then choose a response R_t so that its relation to the reference response R_0 reflects the relation between ψ_t and ψ_0. This approach can be written as

$$\frac{R_t}{R_0} = \frac{\Psi_t}{\Psi_0}\upsilon_t,$$

where υ_t represents random judgmental error. The idea of the above is that judgment, and measurement in general, is a process of comparison. Subjects compare ψ_t and ψ_0, as well as R_t and R_0. The comparisons are modeled here using ratios, which is consistent with the instructions typically given in magnitude scaling experiments (although the theory can also be developed using differences). Note that it is not assumed that subjects estimate ratios when determining their responses; rather, they produce responses so that response relations reflect sensation relations. Rearranging terms and substituting $\alpha = R_0/\psi_0$ gives

$$R_t = \alpha\, \Psi_t \upsilon_t.$$

Thus, if the above approach is used, responses will be proportional to sensation magnitudes, as requested. An important aspect of the above is that it shows that, if the subject makes ratio comparisons, then the assignment of an

I thank Jennifer Simpson for conducting the heaviness and roughness experiments.

Correspondence concerning this article should be addressed to Lawrence T. DeCarlo, Department of Psychology, Fordham University, Bronx, New York 10458.

arbitrary response (R_0) to a reference sensation (ψ_0) establishes a psychological unit of measurement ($\alpha = R_0/\psi_0$). In the classical approach of Stevens (1986), the unit is assumed to remain constant over time. The model introduced below relaxes this assumption.

Another way to make proportional judgments is to use the response and sensation of the previous trial as reference points. In this case, subjects respond so that the relation between their current response and previous response matches the relation between their current sensation and previous sensation. This approach can be written as

$$\frac{R_t}{R_{t-1}} = \frac{\Psi_t}{\Psi_{t-1}} \upsilon_t,$$

which can be rearranged to give

$$R_t = \frac{R_{t-1}}{\Psi_{t-1}} \Psi_t \upsilon_t.$$

Repeated substitution in the above equation for R_{t-1}, R_{t-2}, and so on to R_0 (using lagged versions of the equation) shows that if the above approach is used, then responses will again be proportional to sensation magnitudes (see DeCarlo & Cross, 1990). The temporal structure of the responses, however, will differ. This can be seen by noting that, because the current response is compared to the previous response, the judgmental error that is a part of R_{t-1} also affects R_t. As a result, the errors will have effects that propagate over trials. Thus, although the mean responses will be unaffected by which of the two approaches outlined above is used, the temporal structure of the responses will be markedly affected.

The model introduced by DeCarlo (1989/1990) explicitly recognizes that there is more than one way to perform the task. The idea of the model is that both of the frames of reference discussed above can influence judgment. In particular, the long-term frame, R_0/ψ_0, and the short-term frame, R_{t-1}/ψ_{t-1}, are weighted in judgment as

$$R_t = (R_0/\Psi_0)^{1-\lambda}\, \Psi_t\, (R_{t-1}/\Psi_{t-1})^{\lambda}\, \upsilon_t, \qquad (1)$$

where the parameter λ measures the relative influence of the two frames. Note that the two approaches to making proportional judgments discussed above are special cases of Equation 1. For example, if $\lambda = 0$, then all judgments are made relative to the long-term frame of reference, R_0/ψ_0, and the model reduces to the first approach discussed above. If $\lambda = 1$, then all judgments are made relative to the short-term frame of reference, R_{t-1}/ψ_{t-1}, which is the second approach.

Equation 1 allows for an influence of both frames of reference on judgment. An important consequence of the model is that it shows that, if the short-term reference influences judgment, then the judgmental errors will be autocorrelated. This can be seen by substituting α for R_0/ψ_0 in Equation 1 and rearranging terms to get

$$R_t = \alpha\, \Psi_t\, (R_{t-1}/\alpha\, \Psi_{t-1})^{\lambda}\, \upsilon_t,$$

which can also be written as

$$R_t = \alpha\, \Psi_t \epsilon_t, \qquad (2)$$

where ϵ_t is equal to the terms on the right of ψ_t in the preceding equation. Equation 2 is a model of proportional judgment with nonrandom judgmental errors, ϵ_t, which is in contrast to the random judgmental errors υ_t of Equation 1 (it can be shown that ϵ_t is solely a function of υ_t and earlier random errors υ_{t-1}). Insight into the nature of the error process can be gained by noting that Equation 2 implies that $\epsilon_t = R_t/(\alpha\psi_t)$, from which it follows that $\epsilon_{t-1} = R_{t-1}/(\alpha\psi_{t-1})$, and substituting into the equation preceding Equation 2 gives

$$R_t = \alpha \Psi_t \epsilon_{t-1}^{\lambda}\, \upsilon_t.$$

The model can be transformed from a multiplicative model to an additive one by taking logarithms, which gives

$$\log R_t = \log \alpha + \log \Psi_t + \lambda e_{t-1} + u_t, \qquad (3)$$

where $e_{t-1} = \log \epsilon_{t-1}$ and $u_t = \log \upsilon_t$.

Equation 3 shows that, when $\lambda \neq 0$, the error process will be first-order autoregressive, or AR(1). Thus, Equation 1 provides a theoretical basis for autocorrelated errors (and responses) in magnitude scaling experiments. The theory shows that the autocorrelation parameter λ can be interpreted as a measure of the relative influence of the short- and long-term frames of reference on judgment. Note that each frame consists of a response–sensation pair, which together establish a psychological unit of measurement. Another (equivalent) view of the model is that it generalizes the simple model of proportional judgment by allowing the unit of measurement α to vary over time.

Equations 1 and 3 have several implications for the structure and interpretation of magnitude scaling data. For example, a basic implication is that it might be possible to manipulate the magnitude of the observed autocorrelation, as measured by λ, by varying the instructions. In particular, the autocorrelation should be larger or smaller depending on whether subjects are told to make all their judgments relative to short-term or long-term reference points, respectively (assuming of course that they are able to follow the instructions). As noted above, this has been confirmed for ME of loudness and area and CMM of line-length to loudness (DeCarlo, 1989/1990; DeCarlo & Cross, 1990). The experiments presented below examine the effect of varying the instructions within-subjects for three other continua: length of lines, heaviness of weights, and roughness of sandpaper.

A Dynamic Regression Model

Equation 1 is a model of the judgmental process. As shown in the preceding section, it leads directly to a simple psychophysical measurement model, namely, a model of proportional judgment with autocorrelated errors. The model shows how unobservable sensation magnitudes are related to observed responses in magnitude scaling experiments. To arrive at a regression model, which relates observed responses to measured stimulus intensities, the relation between sensation magnitudes and stimulus intensities, that is, the psychophysical function, must be specified. Stevens (1986) advocated a power psychophysical function

$$\Psi_t = S_t^{\beta}\delta_t,$$

where S_t is the measured value of the stimulus intensity on trial t, β is a characteristic of the sensory continuum (according to Stevens), and δ_t represents random noise in the sensation.

A number of researchers have suggested that the above equation be generalized to allow for a possible context effect of prior stimulation on the current perception (see DeCarlo & Cross, 1990). One possible generalization is

$$\Psi_t = S_t^{\beta}\, S_{t-1}^{\gamma}\, \delta_t,$$

where S_{t-1} is the stimulus intensity on the previous trial, and γ is a parameter that indicates the magnitude and direction of the effect of the previous stimulus intensity on the current perception. If $\gamma = 0$, then the current perception is not affected by the previous stimulus intensity. A positive value of γ, on the other hand, indicates an additive effect of S_{t-1}, whereas a negative value indicates a subtractive effect. It should be noted that nonzero values of γ can be interpreted as arising from a perceptual effect (e.g., an effect of prior stimulation on ψ_t) or (in cases where γ is positive) from a memory effect, such as a tendency to confuse the current and previous perceptions in memory (assimilation in memory; see DeCarlo & Cross, 1990).

Substituting the above into Equation 3 gives

$$\log R_t = \log \alpha + \beta \log S_t + \gamma \log S_{t-1} + \lambda e_{t-1} + u_t, \qquad (4)$$

where $u_t = \log \upsilon_t + \log \delta_t$. Equation 4 is the basic dynamic regression model considered in this article (see DeCarlo, 1992; DeCarlo & Cross, 1990). Note that the model attributes the observed autocorrelation of responses to two possible sources: a judgmental process (as shown by Equation 1) and a perceptual–memory process (as shown by the equation preceding Equation 4). The judgmental effect, which is the influence of short- and long-term frames of reference on judgment, is measured by the autocorrelation parameter λ. The perceptual and/or memory effect, which is the influence of prior stimulation on the current perception, is measured by the parameter γ. An important aspect of Equation 4 is that, according to the theory presented above, its coefficients provide direct estimates of the theoretical parameters λ and γ. The estimated parameters are examined below to determine (a) whether the judgmental effect can be manipulated for the three continua by varying the instructions, and (b) whether there is an effect of the previous stimulus intensity on the current response.

Method

Subjects

The subjects were 26 undergraduates enrolled in introductory psychology courses at Fordham University. They received course credit for participating in the experiment. Eight of the subjects participated in the line-length experiment, 10 in the heaviness experiment, and 8 in the roughness experiment.

Apparatus

The lines were presented on a videographics array color monitor (640 × 480). The lines were white against a black background; they were approximately 2 mm in thickness and approximately 2, 3, 6, 9, 16, 26, 43, 72, 121, and 202 mm in length. Each line was presented for 2 s. Subjects used a keyboard to enter numerical responses, which could contain decimals. Each trial began 2 s after the subject entered a response.

The weights for the lifted weight experiment were made from 8 commercially available (Pearl Paints) cylindrical plastic 8-oz containers that were 63.5 mm in diameter and 91 mm in height, with white screw-on covers. The translucent containers were lined with black paper so that the inside of the container was not visible. They were then filled with lead shot and tightly packed with cotton to eliminate movement of the lead. The final weights of the containers (to the nearest gram) were 50, 80, 120, 180, 270, 400, 600, and 900 g.

For the roughness experiment, aluminum oxide sandpaper of the following grit numbers was used: 36, 50, 60, 80, 100, 120, 180, 220. The sandpaper was cut into 50-mm squares.

For the line-length experiment, the presentation of the stimuli and collection of responses were controlled by Micro Experimental Laboratory software running on an IBM compatible computer. The 10 lines were randomly sampled with replacement for a total of 100 trials. For the heaviness and roughness experiments, the 8 stimuli were presented according to sequences of 60 trials generated by the function RANUNI of Statistical Analysis System (SAS; SAS Institute Inc., 1990).

Procedure

Line Length. Each subject participated in two conditions, which differed only with respect to the instructions (presented below). The order of conditions was counterbalanced across subjects. Subjects were first given 10 practice trials, which consisted of magnitude estimation of the length of the 10 lines. The instructions were the same as those used in the experiment. After the practice trials, subjects were given an opportunity to ask questions. They were then given a total of 100 trials of ME of the 10 line-lengths. Upon completion of the 100 trials, they were given approximately a 2-min break. The second condition was then begun. Again, subjects were given 10 practice trials before beginning the second 100 experimental trials.

The instructions for ME (referred to as fixed reference ME in DeCarlo & Cross, 1990) were as follows:

> You will be presented with a series of lines that vary in length. Your task is to indicate how long each line seems by assigning a positive number to its length. USE ONE OF THE LINES AS A REFERENCE POINT. Start by choosing a positive number for the reference length. Then try to make all your judgments RELATIVE TO THE REFERENCE LENGTH. That is, on each trial, COMPARE THE CURRENT LENGTH TO THE REFERENCE LENGTH and choose a number that has the same relation to the reference number. Your number on each trial should indicate how many times longer or less long the current line seems RELATIVE TO THE REFERENCE LINE. You may use any positive numbers you like, including decimals. If you have any questions, please ask them now.

For the practice ME session, subjects were told to use the first presented line (which was randomly selected) as the reference length (because they had not yet seen the lines used in the experiment). For the experiment, they were told that they could use any length they liked as the reference.

The instructions for ratio magnitude estimation (RME; prior reference instructions in DeCarlo & Cross, 1990) were as follows:

> You will be presented with a series of lines that vary in length. Your task is to indicate how long each line seems by assigning a positive number to its length. USE THE LINE OF THE PREVIOUS TRIAL AS A REFERENCE POINT. Start by choosing a positive number for the first length. Then try to make all your judgments RELATIVE TO THE LENGTH PRESENTED ON THE PREVIOUS TRIAL. That is, on each trial, COMPARE THE CURRENT LENGTH TO THE PREVIOUS LENGTH and choose a number that has the same relation to the previous number. Your number on each trial should indicate how many times longer or less long the current line seems RELATIVE TO THE LINE OF THE PREVIOUS TRIAL. You may use any positive numbers you like, including decimals. If you have any questions, please ask them now.

It was emphasized during the practice trials that subjects should compare each line-length to the reference length, for ME, or to the previous length, for RME.

Heaviness. The procedure was identical to that described above, with the following exceptions. To familiarize subjects with the task, they were first given 20 practice trials judging the length of lines, using the 10 lengths that appeared in the line-length experiment. The lines were presented in a random order with each length appearing twice. Subjects were then given 10 practice trials judging the heaviness of the weights.

Each condition of the experiment (ME or RME instructions) consisted of 60 trials where the heaviness of each weight was judged. The instructions were identical to those presented above, with the substitution of the word "heaviness" for "length" ("heavier" for longer, "heavy" for "long") and "weight" for "line." The order of instructions was counterbalanced across subjects. Because the responses were oral, subjects were told that, in addition to decimals, they could use fractions if they wished. For the ME condition, the 270-g weight was designated as the standard; subjects were allowed to choose its reference number (modulus). Subjects lifted the reference weight several times before the practice trials and again before the experimental trials.

Subjects were blindfolded and sat across from the experimenter at a table. Their arms (from elbow to hand) rested on the table, with the palm perpendicular to the tabletop. On each trial, the weight was placed against each subject's palm. The subject was instructed to lift each weight, being careful not to jiggle it, and to judge its heaviness. Subjects called out their responses, which were recorded by the experimenter.

Roughness. The procedure was identical to the above, with the following exceptions. Following 20 practice trials judging line-lengths, subjects were given 10 practice trials judging the roughness of the sandpapers. The experiment, which consisted of 60 trials, was then begun. The instructions were identical to those presented above, with the substitution of the word "roughness" for "length" and "sandpaper" for "line." The blindfolded subjects were told to stroke their first finger across the surface of the sandpaper twice on each trial. For the ME condition, the grit number 80 sandpaper was designated as the standard; subjects were allowed to choose its reference number (modulus). Subjects were asked to stroke the reference sandpaper several times before the practice trials and again before the ME experimental trials.

Results

Mean and Variability of Responses

The upper panels of Figure 1 present, separately for each instruction and continuum, the medians (across subjects) of the mean log responses to each log stimulus intensity. For the roughness estimation experiment, the log responses are plotted against the log of the inverse of the grit numbers (see Marks & Cain, 1972; Stevens & Harris, 1962). The plots for all three experiments are approximately linear.

The lower panels of Figure 1 show, for each instruction and continua, the medians and interquartile ranges (across subjects) of the standard deviations of the log responses for each log stimulus intensity. The standard deviations for the line-length experiment appear to be approximately constant throughout the range. The standard deviations for the heaviness and roughness experiments show a decrease for the higher stimulus intensities. This result is frequently found in magnitude scaling experiments (see Marley & Cook, 1986). The plots also show that the median standard deviations tend to be larger for RME than for ME (i.e., the response variability is larger). Greater response variability for RME is consistent with larger autocorrelation (see DeCarlo & Cross, 1990).

Time Series Analysis of Residuals

The autocorrelation function (ACF), partial autocorrelation function (PACF), and cross-correlation function (CCF) for the residuals obtained from a fit of Stevens's power law in log–log form (i.e., a regression of log R_t on log S_t) were computed and plotted for each subject using the procedure PROC ARIMA of SAS (see SAS Institute Inc., 1988). The estimated ACF is the correlation of the residuals separated by lags of 1, 2, 3, and so on; the results for the first 8 lags are shown. The PACF is similar, except that correlations for intermediate lags are partialed out. The CCF is the cross-correlation of the residuals with the log stimulus intensities from previous trials (the lagged intensities). These functions are used in time series analysis to help identify the process (see Box & Jenkins, 1976). For an AR(1) process, the ACF is a geometrically decaying function of the lag, whereas the PACF shows a cutoff after the first lag.

Figures 2, 3, and 4 present, separately for each instruction and experiment, the group ACF, PACF, and CCF plots. The group plots present the medians and interquartile ranges (across subjects) of the individual correlations, and provide a summary of the results for each subject.

The ACF and PACF plots for the three experiments show that there is little or no autocorrelation for the ME instructions. In contrast, the plots for RME show relatively large autocorrelation. The ACF and PACF plots for RME suggest a first-order autoregressive error process: there is an approximate geometric decay in the ACF and a cutoff after the first lag in the PACF (cf. DeCarlo, 1992; DeCarlo & Cross, 1990). A comparison of the ACF and PACF plots across the two instructions show that the autocorrelation is clearly larger for RME, which is as predicted. The CCF plots show that the previous stimulus intensity has at most a small, positive effect on responses (the lag 1 correlation is small and positive), with perhaps a slight increase for RME.

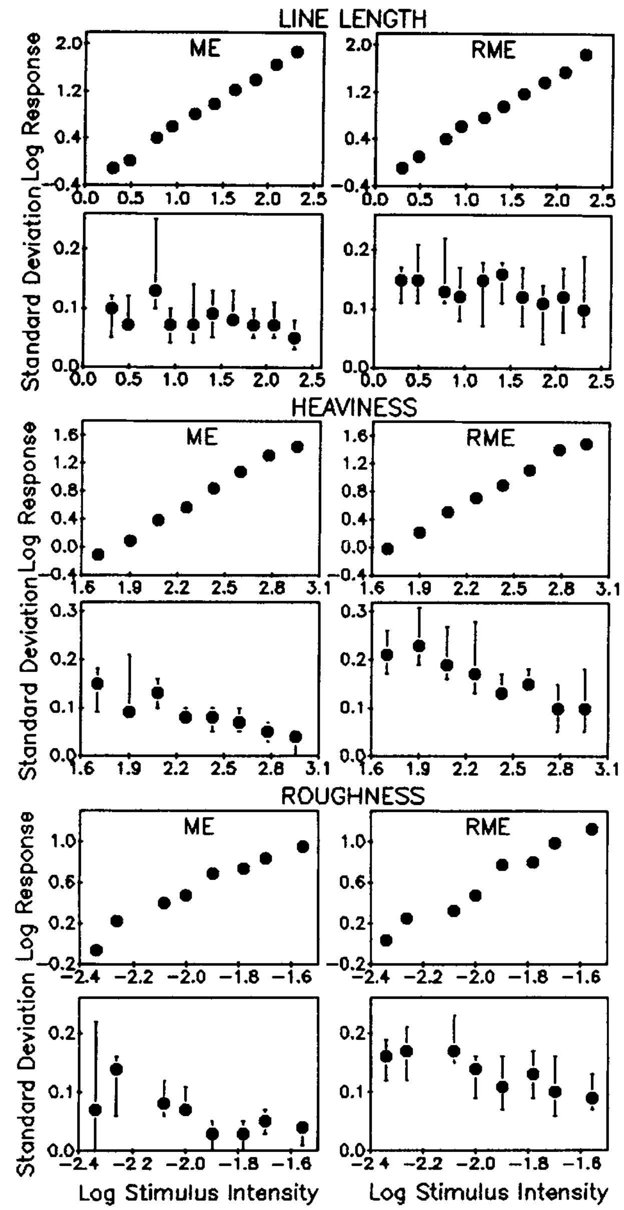

Figure 1. Magnitude estimation (ME) of length (top 4 panels), heaviness (middle 4 panels), and roughness (bottom 4 panels). The upper panels for each experiment present, separately for each instruction (ME or ratio magnitude estimation [RME]), the medians (across subjects) of the mean log responses to each log stimulus intensity (log length in millimeters for length, log weight in grams for heaviness, log inverse grit number for roughness). The lower panels for each experiment present the medians and interquartile ranges (across subjects) of the standard deviations of the log responses.

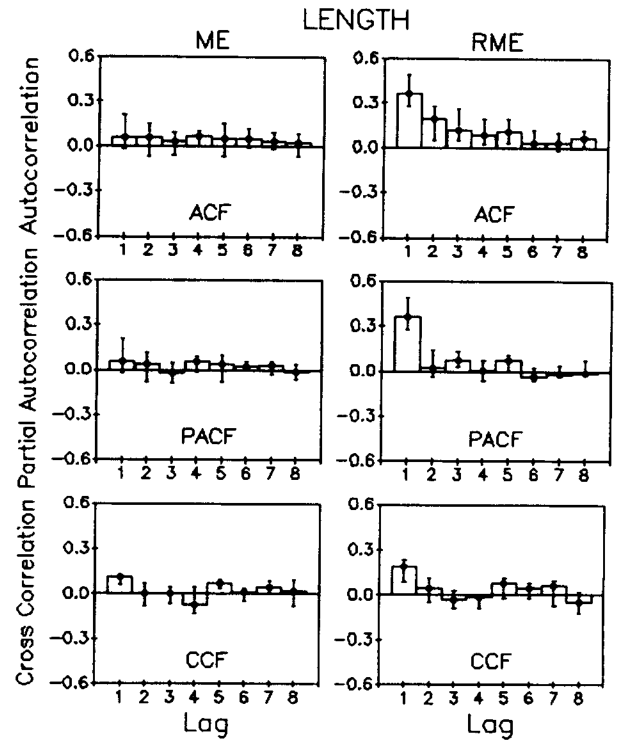

Figure 2. Magnitude estimation (ME) of length. The medians and interquartile ranges (across subjects) of the autocorrelation functions (ACFs), partial-autocorrelation functions (PACFs), and cross-correlation functions (CCFs) for the residuals from a regression of log R_t on log S_t. The functions are plotted separately for each instruction (ME and ratio magnitude estimation [RME]).

Regression Analysis

Table 1 presents the results for fits of Equation 4 for the ME instructions (left half of table) and the RME instructions (right half) for the length, heaviness, and roughness experiments, respectively. The estimated coefficients were obtained using PROC AUTOREG of SAS (using the Yule–Walker estimates; see SAS Institute Inc., 1988). The results of significance tests on the coefficient of log S_{t-1} (γ) and e_{t-1} (λ) are shown for the individual analyses. The test for nonzero γ is for the transformed regression model (see SAS Institute Inc., 1988). For the autocorrelation test (i.e., the test for nonzero λ), the Durbin–Watson test (Durbin & Watson, 1950, 1951) was performed on the residuals obtained from a regression of log R_t on log S_t and log S_{t-1}. The coefficient of determination, R^2, is also shown for each subject. This is the R^2 for the transformed model, which is a measure of the goodness of fit of the systematic part of the transformed model; it is referred to in the SAS output as the regression R^2 (see SAS Institute Inc., 1988). Note that the reported values of R^2 are for fits of the individual responses, not for mean responses. (Fits for mean responses are often reported, which leads to larger values of R^2.)

The mean estimates of the exponent β are close to those typically obtained for each continua (see Marks & Cain, 1972; Stevens, 1986; Stevens & Harris, 1962). The mean estimates of β are slightly less than unity for line-length estimation (.99 for ME, .92 for RME) and are greater than unity for heaviness (1.32 for ME, 1.32 for RME) and roughness estimation (1.36 for ME, 1.31 for RME). The values of R^2 are high for each subject, which reflects the fact that the linearity shown in Figure 1 generally holds for the individual data.

Table 1 also shows that the autocorrelation, as measured by λ, is considerably larger for the RME instructions than for the ME instructions for all three continua. The mean estimates of λ for line-length estimation are .08 for ME and .37 for RME, for heaviness they are .11 for ME and .48 for RME, and for roughness they are .11 for ME and .50 for RME. With respect to the results for each individual, an increase in autocorrelation for the RME instructions appears for 8 out of 8 subjects in the line-length estimation experiment, for 9 out of 10 subjects in the heaviness estimation experiment, and for 8 out of 8 subjects in the roughness estimation experiment. Table 1 also shows that the mean estimates of γ (the perceptual–memory parameter) are positive but close to zero. Tests for the individual subjects show that only a few of the estimates are significant.

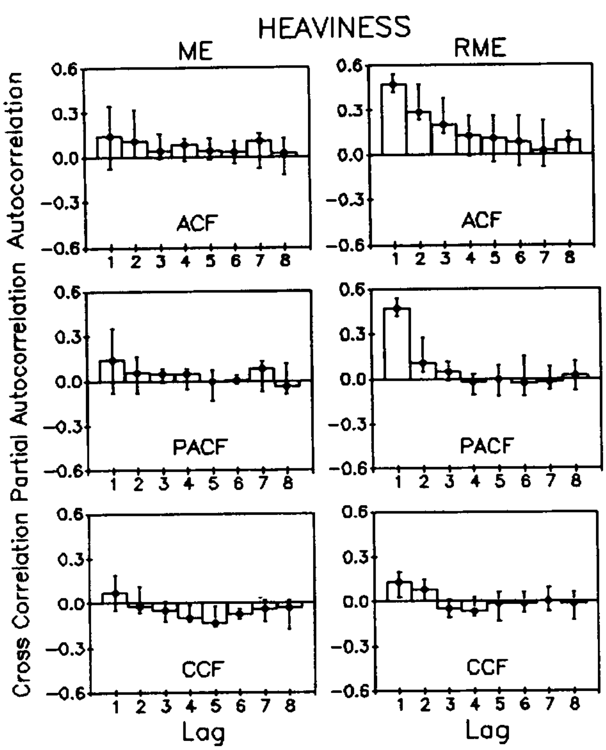

Figure 3. Magnitude estimation (ME) of heaviness. The medians and interquartile ranges (across subjects) of the autocorrelation functions (ACFs), partial-autocorrelation functions (PACFs), and cross-correlation functions (CCFs) for the residuals from a regression of log R_r on log S_r. The functions are plotted separately for each instruction (ME and ratio magnitude estimation [RME]).

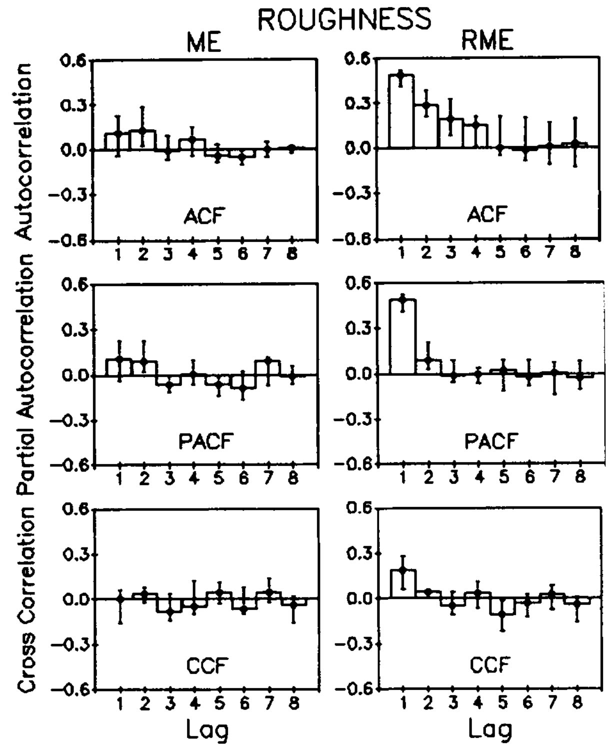

Figure 4. Magnitude estimation (ME) of roughness. The medians and interquartile ranges (across subjects) of the autocorrelation functions (ACFs), partial-autocorrelation functions (PACFs), and cross-correlation functions (CCFs) for the residuals from a regression of log R_r on log S_r. The functions are plotted separately for each instruction (ME and ratio magnitude estimation [RME]).

Discussion

The focus of the present article is on the simple, dynamic judgmental model presented above as Equation 1. According to the model, subjects in magnitude scaling experiments choose responses, which can be nonnumerical, so that relations between their responses reflect relations between their sensations. This process of comparison forms a basis for proportional judgment, as shown previously. Equation 1 generalizes the basic model of proportional judgment by allowing for the influence of different frames of reference on judgment, namely, a long-term frame, which is a stable response–sensation relation, and a short-term frame, which is the response–sensation relation of the previous trial.

An important aspect of Equation 1 is that it shows that autocorrelated errors arise in magnitude scaling because of the influence of the different frames of reference. An understanding of this source of autocorrelation suggests how to gain a degree of experimental control over it. In particular, the experiments presented above, as well as previous research, show that there are sizeable changes in autocorrelation when the instructions are varied. A basic implication is that it is important to be explicit in the instructions about how the task should be performed. Simply asking for

Table 1
Results for Equation 4 for Length, Heaviness, and Roughness

Subject	Magnitude estimation				Ratio magnitude estimation			
	β	γ	λ	R^2	β	γ	λ	R^2
				Length				
1	1.20	.06	.06	.93	.99	.06*	.14	.93
2	1.03	.03	.15[a]	.93	1.05	.02	.22*	.98
3	.98	.02	.05	.98	.86	.05*	.32*	.94
4	.83	.04	−.05	.89	.77	−.02	.63*	.93
5	.95	.01	.27*	.98	.99	.01	.38*	.98
6	.98	.01	.02	.98	.91	.04	.31*	.94
7	.94	−.00	−.10	.97	.88	.09*	.38*	.94
8	1.04	.02	.24*	.97	.88	.06*	.55*	.94
M	.99	.02	.08		.92	.04	.37	
				Heaviness				
1	1.12	.10	−.09	.88	1.38	.05	.57*	.93
2	1.54	.11*	−.12	.97	1.58	.09	.54*	.88
3	1.43	−.04	−.19	.86	1.29	.01	.41*	.89
4	1.22	−.03	.26[a]	.93	1.15	.01	.45*	.90
5	1.42	.01	.31[a]	.96	1.51	.08	.65*	.96
6	1.53	.05	.39*	.95	1.74	.05	.14	.91
7	1.39	−.02	.14	.97	1.33	.08	.35*	.91
8	1.29	.00	.42*	.98	1.14	.04	.46*	.94
9	1.13	−.01	−.16	.95	1.00	.03	.51*	.95
10	1.15	.04	.16	.98	1.13	.00	.71*	.96
M	1.32	.02	.11		1.32	.04	.48	
				Roughness				
1	1.22	−.08	.39*	.91	1.70	.02	.51*	.91
2	1.37	−.01	−.09	.83	1.24	.10*	.73*	.93
3	1.13	.02	.20[a]	.94	1.26	.16*	.47*	.89
4	1.36	.01	−.01	.88	1.37	.11	.49*	.85
5	1.13	.13*	.26*	.86	1.20	.21*	.34*	.86
6	1.58	−.10	.05	.85	1.09	.01	.39*	.88
7	1.39	.02	.17	.91	1.46	.20*	.52*	.89
8	1.68	−.06	−.12	.97	1.13	−.05	.54*	.90
M	1.36	−.01	.11		1.31	.10	.50	

[a] Durbin–Watson test (Durbin & Watson, 1950, 1951) inconclusive.
* $p < .05$.

proportional judgments, for example, leaves it to the subject to determine how to perform the task.

Although autocorrelation is shown to be of theoretical interest in this article, it is often viewed in applied research as a nuisance. With respect to this point of view, the results show that in order to minimize autocorrelation, it should be made clear to subjects that they should make all of their judgments relative to a fixed reference point (a reference response and sensation), as in the ME instructions presented above. As shown in Table 1 and Figures 2, 3, and 4, this greatly reduces or eliminates the autocorrelation.

A variation on the method of ME is absolute magnitude estimation (AME; see chapters in Bolanowski & Gescheider, 1991; Zwislocki & Goodman, 1980). Subjects in AME are instructed to "match" their impression of number magnitude to their impression of sensation magnitude, where the magnitude impressions are assumed to lie on a common underlying continuum. The instructions presented here differ in that subjects are asked to match relations and not absolute impressions. The instructions follow from the focus on the *relative* nature of judgment and measurement; Equation 1 is in fact a dynamic model of relative judgment. Absolute judgment is considered here as being judgment with respect to a long-term frame of reference.

On an empirical level, little is known about how AME instructions affect sequential effects, as Gescheider (1988) has previously noted. Research on AME has focused on mean responses, because of interest in how AME affects "bias." Individual or mean autocorrelations have generally not been reported. The one exception (Ward, 1987) found relatively large autocorrelation for AME of loudness (an average of .39), which suggests that subjects' judgments were influenced by responses and sensations from previous trials. With respect to future research, the present article shows that Equation 4 offers a simple means of studying sequential effects in AME.

As discussed above, Equation 4 attributes sequential effects to a judgmental factor, as measured by λ, and a perceptual–memory factor, as measured by γ. The present research shows that the judgmental parameter λ behaves as expected when the instructions are varied: The mean estimates of λ were more than four times larger for the RME instructions. In addition, the experiments provide information about stimulus context effects for length, heaviness, and roughness. In particular, Table 1 shows that the mean estimates of γ tend to be small and positive for the three continua. These results agree with those found in previous research on line-length estimation (Morris & Rule, 1988) and heaviness estimation (Cross & Rotkin, 1975). The results for roughness estimation suggest that γ might be larger for RME (although its magnitude is still small relative to β), perhaps because subjects in the RME condition have a greater tendency to confuse their current and previous perceptions in memory. Further research is needed to explore this possibility.

There has been an emphasis in psychophysics on studying the psychophysical function, following the tradition established by Fechner (1860/1966). However, because all psychophysical techniques rely on observed responses to obtain information about unobservable sensation magnitudes, psychophysical judgmental models play an equally important role. In fact, it is difficult to draw conclusions about sensory processes from psychophysical studies without considering the judgmental process. For example, it has previously been shown that an apparent contrastive effect of the previous stimulus intensity on the current perception for loudness estimation might actually be an assimilative or additive effect (DeCarlo, 1992; DeCarlo & Cross, 1990). These seemingly contradictory conclusions arise in the context of an earlier regression model proposed by Jesteadt et al. (1977), which is

$$\log R_t = \alpha + \alpha_1 \log S_t + \alpha_2 \log S_{t-1} + \alpha_3 \log R_{t-1} + u_t. \quad (5)$$

Equation 5 differs from Equation 4 in that it includes a lagged dependent variable ($\log R_{t-1}$) in lieu of an AR(1) error process. Equations 4 and 5 represent two basic alternatives—whether to account for autocorrelation by introducing a lagged dependent variable (as in Equation 5) or a nonrandom error process (as in Equation 4). The discussion here focuses on how the choice of model affects the interpretation of the model parameters. The discussion is in terms of psychophysical models, but it has implications for dynamic modeling in general.

Equation 5, which has been used in earlier research, gives results that differ from those for Equation 4 in that the coefficient of $\log S_{t-1}$ is typically negative for loudness estimation, which is what led to the conclusion of contrast (see Ward, 1982). This conclusion, however, depends on the judgmental model that is explicitly or implicitly assumed. For example, it has previously been shown that a negative value of α_2 can be interpreted as indicating contrast if the autocorrelation arises from the use of a response heuristic, such as a tendency to choose a response close to the previous response (see Equations 14 and 15 of DeCarlo & Cross, 1990). This is not the case, however, if the process is one of relative judgment as in Equation 1. In particular, it follows from Equation 1 that the coefficient of $\log S_{t-1}$ in Equation 5 is not simply an estimate of the perceptual–memory parameter γ, but rather is an estimate of $\gamma - \lambda\beta$ (see DeCarlo, 1992; DeCarlo & Cross, 1990). This can be seen by substituting the perceptual context model presented above directly into Equation 1 and rearranging terms. A basic consequence of the theory presented here, therefore, is that the coefficient α_2 in Equation 5 does not provide a direct estimate of γ (the perceptual–memory parameter), but also reflects effects of λ (the judgmental parameter) and β (the psychophysical exponent).

Two lines of evidence suggest that the coefficient α_2 in Jesteadt et al.'s (1977) model does indeed confound perceptual and judgmental effects. First, it has been shown that the coefficient of $\log S_{t-1}$ in Equation 5 systematically varies in magnitude with the autocorrelation (see DeCarlo & Cross, 1990). In particular, an increase in autocorrelation is typically accompanied by a sizeable increase in negative α_2 (in situations where β and γ remain relatively constant). It is important to recognize that this result is consistent with the parameter constraint implied by Equation 1. To see why, note that previous research has shown that γ is small and positive for loudness estimation, and that the product $\lambda\beta$ is greater than γ. It then follows from the relation $\alpha_2 = \gamma - \lambda\beta$ that α_2 will be negative and that an increase in λ will lead to an increase in the magnitude of negative α_2, which is exactly what was found.

A second line of evidence comes from recent research, which has shown, that, for loudness estimation, an increase in the intertrial interval (ITI) leads to a decrease in the magnitude of γ for Equation 4, which is as expected, whereas α_2 in Equation 5 increases in magnitude (DeCarlo, 1992). The decrease in γ for Equation 4 with increased ITI is consistent with the view that γ reflects the effect of the previous stimulus intensity on the current perception. It is expected that this effect will decrease in magnitude when the previous stimulus intensity is further away in time. The increase in α_2 for Equation 5 with increased ITI, on the other hand, is not consistent with the view that α_2 reflects a perceptual contrast effect, because the contrast should decrease with increased ITI, and not increase. The parameter constraint implied by Equation 1, however, sheds light on this result. In particular, the increase in negative α_2 is consistent with the relation $\alpha_2 = \gamma - \lambda\beta$: if γ is small, positive, and less than the product $\lambda\beta$ for loudness estimation, then a decrease in its magnitude with increased ITI will lead to an increase in the magnitude of negative α_2.

The present experiments allow a further comparison of Equations 4 and 5. Table 2 presents the mean (across subjects) coefficients obtained for fits of Equations 4 (left side of table) and 5 (right side of table) to the data from all three experiments. The results for each instruction (ME and RME) are shown separately.

Table 2
Mean Parameter Estimates for Equations 4 and 5

Condition	Equation 4 β	Equation 4 γ	Equation 4 λ	Equation 5 α_1	Equation 5 α_2	Equation 5 α_3
Length						
ME	.99	.02	.08	.99	−.07	.08
RME	.92	.04	.34	.92	−.30	.39
Heaviness						
ME	1.32	.02	.11	1.32	−.16	.13
RME	1.32	.04	.48	1.33	−.58	.49
Roughness						
ME	1.36	−.01	.11	1.36	−.13	.11
RME	1.31	.10	.50	1.31	−.57	.52

Note. ME = magnitude estimation; RME = ratio magnitude estimation.

The left side of the table presents a summary of the results discussed previously: The autocorrelation, as measured by λ, is clearly larger in magnitude for the RME instructions, and the mean estimates of γ are positive, but close to zero. The right side of the table shows that the results for Equation 5 differ in that the mean estimates of α_2 are large and negative. This is the typical finding of "contrast." Thus, Equation 5 suggests that the previous stimulus intensity exerts a contrastive influence on the current response, whereas Equation 4 shows that this is not necessarily the case.

The results shown in Table 2 provide further evidence in favor of the parameter constraint discussed above. As noted above, it follows from the relation $\alpha_2 = \gamma - \lambda\beta$ that α_2 should covary with the autocorrelation (for constant β and γ) and this result is clearly evident in the table. For length estimation, α_2 increases from −.07 for ME to −.30 for RME, for heaviness estimation from −.16 for ME to −.58 for RME, and for roughness estimation from −.13 to −.57. Thus, the present experiments provide further evidence of the confounding of effects by α_2 in Equation 5. In fact, the negative values of α_2 shown on the right side of Table 2 are closely predicted by the parameter constraint. For example, for ME of line-length, the predicted value of α_2 is $.02 - .08 \times .99 = -.06$ (from the results shown on the left side of the table), which is close to the mean value of −.07 obtained for Equation 5. The predictions for the rest of the table are similar.

Table 2 also sheds light on the observation of Morris and Rule (1988, p. 71) that "the lack of a contrast effect for length and numerousness in the present study differs from previous findings of such an effect for loudness and brightness." As discussed above, the contrast effect previously found for loudness and brightness most likely arose from the use of Equation 5 to study sequential effects. In fact, the results for length estimation shown on the right side of Table 2 suggest that, if Morris and Rule (1988) had fit Equation 5, they would have found "contrast" for length estimation. The results for Equation 4, on the other hand, show the "lack of contrast" noted by Morris and Rule. Another interesting aspect of Morris and Rule's results is that the autocorrelation was slightly smaller when an explicit modulus (reference number) and standard (reference stimulus) were designated.

Further insight into the results shown in Table 2 can be gained by noting that it follows from the relation $\alpha_2 = \gamma - \lambda\beta$ that the contrast found for Equation 5 depends in part on the magnitude of β. In particular, the relation implies that the apparent contrast will be large when β is large (for $\lambda > 0$). The results for heaviness and roughness are of interest with respect to this prediction, because β for these continua is typically "large" (i.e., greater than unity; see Stevens, 1986). It follows that the "contrast" found for these continua should be large, particularly when λ is large as in RME. The right side of Table 2 shows that this result was indeed found: The mean estimates of α_2 for the RME instructions are relatively large for both heaviness (−.58) and roughness (−.57). The relation between α_2 and the parameters of Equation 4, $\alpha_2 = \gamma - \lambda\beta$, suggests that this result was found not because there is a large contrast effect for heaviness and roughness, but because β is large for these continua.

Overall, Table 2 shows that in order to evaluate stimulus context effects, it is important to separate the effects of autocorrelation from the effects of the previous stimulus intensity. This is true irrespective of whether γ is positive, as found above, or negative. An example of possible confounding in a situation where γ might be negative has recently been provided by Schifferstein and Frijters (1992). Subjects were asked to estimate the sweetness of demineralized water containing different concentrations of sucrose. Subjects sipped and spit the solutions and then rinsed with demineralized water. Trials were separated by 50 s. The researchers noted that previous research has suggested that there might be a contrast effect for sweetness. However, considering that a relatively long ITI was used (50 s) and that subjects rinsed after each trial, it seems that the magnitude of the effect should be small. The analysis performed by Schifferstein and Frijters differed somewhat from that presented here, but they nevertheless showed that the mean coefficient for the previous stimulus intensity for a variant of Equation 5 was −.43 (p. 252), whereas the estimate of β was close to unity, and the estimate of λ was .36. From the relation discussed above, the predicted coefficient of log S_{t-1} for Equation 4 is therefore $-.43 + .36 \times 1 = -.07$. Thus, the estimate of γ for Equation 4 (−.07) is considerably smaller than the estimate of α_2 for Equation 5 (−.43). The results suggest that the magnitude of the apparent contrast was inflated by Equation 5 (from −.07 to −.43), because of the confounding of judgmental effects, as discussed above.

In summary, Equation 1 presents a simple model of dynamic judgment in magnitude scaling experiments. Combined with a generalization of the psychophysical function that allows for stimulus context effects, it leads directly to Equation 4. According to the theory, the coefficients of Equation 4 provide direct estimates of the theoretical parameters λ (a judgmental parameter) and γ (a perceptual–memory parameter). The results for estimation of length, heaviness, and roughness show that the magnitude of the judgmental effect, as measured by λ, can be manipulated by varying the instructions. In addition, the results show that the perceptual–memory effect, as measured by γ, tends to be small and positive for the three continua. With respect to Equation 5, the results show that magnitude of the coefficient of log S_{t-1} (α_2) varies with the autocorrelation in a manner consistent with the parameter constraint implied by Equation 1. The implication is that α_2 in Equation 5 reflects both perceptual and judgmental effects. The confounding appears to be responsible for the large, negative "contrast" typically found for Equation 5. Equations 1 and 4 provide insight into the nature of judgmental and sensory processes in magnitude scaling experiments.

References

Bolanowski, S. J., Jr., & Gescheider, G. A. (Eds.). (1991). *Ratio scaling of psychological magnitude: In honor of the memory of S. S. Stevens.* Hillsdale, NJ: Erlbaum.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (rev. ed.). San Francisco: Holden-Day.

Cross, D. V., & Rotkin, L. (1975). The relation between size and apparent heaviness. *Perception & Psychophysics, 18,* 79–87.

DeCarlo, L. T. (1990). Sequential effects in magnitude scaling (Doctoral dissertation, State University of New York at Stony Brook, 1989). *Dissertation Abstracts International, 50,* 3730B.

DeCarlo, L. T. (1992). Intertrial interval and sequential effects in magnitude scaling. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1080–1088.

DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General, 119,* 375–396.

Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression, I. *Biometrika, 37,* 409–427.

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression, II. *Biometrika, 38,* 159–178.

Fechner, G. T. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). NY: Holt, Rinehart & Winston. (Original work published 1860)

Gescheider, G. A. (1988). Psychophysical scaling. *Annual Review of Psychology, 39,* 169–200.

Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 92–104.

Marks, L. E., & Cain, W. S. (1972). Perception of intervals and magnitudes for three prothetic continua. *Journal of Experimental Psychology, 94,* 6–17.

Marley, A. A. J., & Cook, V. T. (1986). A limited capacity rehearsal model for psychophysical judgments applied to magnitude estimation. *Journal of Mathematical Psychology, 30,* 339–390.

Morris, R. B., & Rule, S. J. (1988). Sequential judgement effects in magnitude estimation. *Canadian Journal of Psychology, 42,* 69–77.

SAS Institute, Inc. (1988). *SAS/ETS user's guide* (Version 6, 1st ed.). Cary, NC: Author.

SAS Institute, Inc. (1990). *SAS language: Reference* (Version 6, 1st ed.). Cary, NC: Author.

Schifferstein, H. N. J., & Frijters, J. E. R. (1992). Contextual and sequential effects on judgments of sweetness intensity. *Perception & Psychophysics, 52,* 243–255.

Stevens, S. S. (1986). *Psychophysics.* NJ: Transaction Books.

Stevens, S. S., & Harris, J. R. (1962). The scaling of subjective roughness and smoothness. *Journal of Experimental Psychology, 64,* 489–494.

Ward, L. M. (1982). Mixed-modality psychophysical scaling: Sequential dependencies and other properties. *Perception & Psychophysics, 31,* 53–62.

Ward, L. M. (1987). Remembrance of sounds past: Memory and psychophysical scaling. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 216–227.

Zwislocki, J. J., & Goodman, D. A. (1980). Absolute scaling of sensory magnitudes: A validation. *Perception & Psychophysics, 28,* 28–38.

Received January 20, 1993
Revision received June 30, 1993
Accepted July 21, 1993 ■