Sequential effects in successive ratio estimation

LAWRENCE T. DeCARLO
Teachers College, Columbia University, New York, New York

Sequential effects are examined in four successive ratio estimation (RE) experiments. The procedure in successive RE is identical to that for magnitude estimation (ME), but the task in successive RE is to estimate the ratio of the current to the previous sensation magnitude, and not the separate magnitudes of the sensations. A positive stimulus context effect was found in successive RE for several continua, in agreement with results previously found for ME. The residual autocorrelation for successive RE was zero in many cases, but in some cases negative autocorrelation was found, which is in contrast to the positive autocorrelation that is typically found for ME and other magnitude scaling procedures. It is shown that, when the role of perceptual error is recognized, negative autocorrelation is predicted by a classical model of ratio estimation. Some aspects of response bias are also discussed.

Successive ratio estimation (RE; Lockhead & King, 1983) is a variation on the magnitude scaling procedures introduced by Stevens (see Gescheider, 1997; Stevens, 1986). In successive RE, stimuli are presented one at a time in a sequence, exactly as in magnitude estimation (ME), but the participant's task is to estimate the ratio of sensation magnitudes produced by successive stimuli, whereas in ME, the task is to estimate the magnitude of each sensation. Thus, successive RE offers an interesting comparison to ME, in that the physical procedure is exactly the same across RE and ME, but the nature of the task differs: estimating sensation ratios versus estimating sensation magnitudes. Any differences across the two procedures must therefore reflect differences in underlying cognitive and/or perceptual processes. It is well known that—in magnitude scaling procedures such as ME, cross-modality matching (CMM), or magnitude production (MP)—participants' responses are not independent over trials, which is usually discussed under the general title of sequential effects (see, e.g., Cross, 1973; DeCarlo, 1994, 2003; Holland & Lockhead, 1968; Ward, 1973, 1979; Ward & Lockhead, 1970). Given the interesting comparison that successive RE affords to ME, it is somewhat surprising that only one study, to my knowledge, has explicitly examined sequential effects in successive RE (Lockhead & King, 1983).

The present article has several goals. First, empirical evidence as to the nature of sequential effects in successive RE is presented. For example, the residuals from a basic regression model for successive RE are examined for autocorrelation; no prior studies, to my knowledge, have examined residual autocorrelation in successive RE. Also assessed is whether or not a stimulus context effect appears in successive RE, as in ME and CMM. The theoretical models underlying successive RE are also presented and discussed. Finally, some of the experiments

presented below include other conditions, such as ME and CMM; this allows one to examine effects of response bias across the various procedures (DeCarlo, 2005).

The Classical Model for Ratio Estimation

Procedures labeled as "ratio estimation" have a long history in psychophysics; here, it is simply noted that in the typical procedure, two stimuli are presented on each trial, with the participant's task being to estimate the ratio of the sensation magnitudes associated with the two stimuli (see, e.g., Stevens, 1986). In successive ratio estimation, on the other hand, only one stimulus is presented per trial, exactly as in ME, with the difference that the task is to estimate the ratio of the current sensation magnitude to the sensation magnitude of the previous trial. The classical model for ratio estimation can easily be applied to successive RE,

$$R_{t,t-1} = \frac{\Psi_t}{\Psi_{t-1}} v_t, \qquad (1)$$

where $R_{t,t-1}$ is the observed ratio estimate, Ψ_t is the sensation magnitude on trial t, Ψ_{t-1} is the sensation magnitude on trial $t-1$, and v_t is (random) judgmental error on trial t. Equation 1, without the judgmental error term, has previously been considered by Sjöberg (1971) and Fagot (1978). Equation 1 represents the simplest possible assumption, which is that participants' ratio estimates reflect the ratio of successive sensation magnitudes, as requested, with judgmental error v_t. Note that Equation 1 does not include a proportionality constant; the inclusion of a proportionality constant is one way to allow for response bias, as noted below.

The sensation magnitudes are of course not observed; so, to complete the model, the relation between sensation magnitudes and observed magnitudes of physical stimuli—that is, the *psychophysical function*—must be specified. One approach is to assume a power psychophysical function,

$$\Psi_t = S_t^\beta \delta_t, \qquad (2)$$

where β is the exponent for the psychophysical function, which Stevens claimed is a characteristic of the sensory

Correspondence should be addressed to L. T. DeCarlo, Department of Human Development, Box 118, Teachers College, Columbia University, 525 West 120th Street, New York, NY 10027 (e-mail: decarlo@tc.edu).

continuum (see Stevens, 1986), and δ_t is (random) perceptual error (i.e., δ_t allows for variation in the sensation magnitudes).

Substituting Equation 2 into Equation 1 and taking logarithms gives a basic regression model for successive RE,

$$\log R_{t,t-1} = \beta \log\left(S_t/S_{t-1}\right) + d_t - d_{t-1} + u_t, \quad (3)$$

where $d_t = \log(\delta_t)$ and $u_t = \log(v_t)$. The above shows that a simple regression of the log of the observed ratio estimates on the log of the presented stimulus intensity ratios should be a linear function with a slope of β and an intercept of zero; the experiments presented below provide evidence on this. An intercept can also be included in Equation 3, which gives what Fagot (1978) referred to as the *constant bias* model, in that it allows for constant underestimation or overestimation of the sensation ratios.

An important aspect of Equation 3 is that it shows that the error structure in the classical model is not random, which has been overlooked up to this point because error terms have typically not been included in the theoretical models (i.e., Equations 1 and 2). The next section examines the error structure of Equation 3 in more detail.

Autocorrelated Errors and the Classical Model

The above derivations show that if participants compare sensation magnitudes in successive RE as instructed, the result is a nonrandom error process. The nonrandomness arises because of the effect of perceptual variability (δ_t) combined with the ratio assumption of the judgmental model (Equation 1). As shown by Equation 3, the error process in logarithmic form is

$$e_t = d_t - d_{t-1} + u_t, \quad (4)$$

where e_t denotes nonrandom error. As shown in the Appendix, the first-order autocorrelation of the errors, $\text{corr}(e_t, e_{t-1})$, is therefore

$$\text{corr}\left(e_t, e_{t-1}\right) = -1/\left[2 + \left(\sigma_u^2/\sigma^2\right)\right], \quad (5)$$

where σ_u^2 is the variance of the judgmental error term u_t and σ^2 is the variance of the perceptual error term d_t (and d_{t-1}). It follows from Equation 5 that the correlation in successive RE should range from zero to -0.5; specifically, the correlation approaches zero (from the left) as σ_u^2/σ^2 gets large (e.g., for small perceptual error variance relative to judgmental error variance) and approaches -0.5 as σ_u^2/σ^2 gets small. Thus, the model makes a rather novel prediction, in that it suggests that residual autocorrelation in successive RE, if nonzero, should be negative, whereas the autocorrelation in ME, CMM, and MP is consistently positive (when nonzero), both for individual subjects and average data. Evidence about the error structure in successive RE is therefore of theoretical interest and is provided by the experiments presented below.

Stimulus Context Effects and the Classical Model

Considerable research has suggested that there might be an effect of the previous stimulus magnitude on the cur-

rent perception (e.g., Cross, 1973; DeCarlo & Cross, 1990; Holland & Lockhead, 1968; Ward, 1973, 1979). The psychophysical function (Equation 2) can easily be generalized to allow for effects of prior stimulation on the current perception (a stimulus context effect). In particular, one approach to modeling this effect is to generalize Equation 2 as follows,

$$\Psi_t = S_t^\beta S_{t-1}^\gamma \delta_t, \quad (6)$$

where γ reflects an effect of the previous stimulus intensity on the current perception and its sign indicates the direction of the effect. The typical finding for ME and CMM of loudness and several other continua (length, heaviness, and roughness) is that γ is small and positive (see, e.g., DeCarlo, 1990, 1992, 1994; DeCarlo & Cross, 1990). Cross (1973) introduced a variation of the above in which the exponent of S_t, β, is replaced with $\beta - \gamma$, in which case a positive value of γ indicates an assimilative effect, in that the current perception is larger when S_{t-1} is greater than S_t and is smaller when S_{t-1} is less than S_t, whereas Equation 6 simply has an additive effect of S_{t-1} (see DeCarlo, 2003). It should be apparent that the only difference is whether or not the effect of S_{t-1} is viewed as lowering the exponent of S_t.

Substituting Equation 6 into Equation 1 and taking logarithms gives

$$\log R_{t,t-1} = \beta_0 + \beta \log\left(S_t/S_{t-1}\right) + \gamma \log\left(S_{t-1}/S_{t-2}\right)$$
$$+ d_t - d_{t-1} + u_t, \quad (7)$$

where an intercept β_0 has been added to allow for constant response bias, as noted above. Equation 7 shows that, in order to assess stimulus context effects (i.e., obtain an estimate of γ), one should include $\log(S_{t-1}/S_{t-2})$ as a regressor and not simply $\log S_{t-1}$; this shows why it is important to be explicit about the theory underlying the model. In fact, Equation 7 with a positive value of γ accounts for several results previously found for successive RE by Lockhead and King (1983), as shown in the Appendix.

Equation 7 is a theoretically based model that can be used to investigate the nature of sequential effects in successive RE. It is of interest to see, for example, if a stimulus context effect, as measured by γ, is found in successive RE and if it is similar in sign and/or magnitude to that found in ME and CMM. Second, a time series analysis of the residuals from a fit of the structural portion of Equation 7 (i.e., the part that doesn't include the error process) will be performed in order to obtain information about the nature of the error process in successive RE. Note that fitting the structural part of the model and ignoring the correlated error process still gives unbiased estimates of the model parameters.

The next section presents results from four experiments on successive RE: RE of loudness of a 1000-Hz tone, RE of area of circles, RE of heaviness of lifted weights, and RE of roughness of sandpaper. For the loudness experiment, the participants also participated in ME, CMM, and MP conditions; for the area experiment, the participants also participated in an ME condition.

METHOD

Participants

The participants were students enrolled in psychology courses; participation in the experiments was voluntary. There were different participants in each of the four experiments, with 8 participants in the loudness experiment, 12 in the area experiment, 16 in the heaviness experiment, and 20 in the roughness experiment.

Apparatus

For the loudness experiment, a General Radio Company oscillator was used to generate 1000-Hz tones. Seventeen tones, ranging from 40 to 88 dB (SPL), in 3-dB steps, were presented binaurally (1-sec duration) through Grason-Stadler headphones (TDH-39). The presentation of stimuli and recording of responses were controlled by a personal computer using BASIC programs written by the author. For the RE and ME conditions, the participants entered their responses by using a KAT, which is a pad that maps to the computer screen. A numerical keypad appeared on the terminal: The keypad consisted of the numbers 0 through 9, a period, a clear entry key, and an enter key. The participants were first given practice so that they could quickly enter their responses using the KAT. For the CMM condition, the participants produced line lengths, by brushing their finger across the KAT, that could range in length from 2 to 640 mm. For the MP condition, the participants produced intensities of a 1000-Hz tone by brushing their finger across the KAT, which incremented the tone in 0.5-dB steps from 10 to 90 dB (it took three to four strokes to adjust the stimulus across the full range). The presented numbers (2-sec duration) for MP were 2, 3, 4, 5, 7, 10, 15, 20, 25, 30, 40, 50, 75, 100, 125, 150, and 200. For both CMM and MP, the initial value of the stimulus to be adjusted (line length or 1000-Hz tone, respectively) was one of 50 values randomly selected from the range (2–640 mm for line length; 10–90 dB for loudness).

For the area experiment, 10 circles with areas (in cc) of 573, 962, 1,662, 2,923, 5,027, 6,648, 11,883, 20,613, 27,466, and 35,970 were presented; for the heaviness experiment, eight weights (8-oz cylindrical containers that weighed, in grams: 50, 80, 120, 180, 270, 400, 600, and 900) were presented; for the roughness experiment, eight squares (50 mm) of aluminum-oxide sandpaper (grit numbers: 36, 50, 60, 80, 100, 120, 180, and 220) were presented. In all of the experiments, the order of presentation of the stimuli was determined by sequences generated using the uniform probability generator of SAS; sequences were selected so that each stimulus was presented at least five times.

Procedure

In all of the experiments, the participants were first given practice sessions. For successive RE, the participants first gave ratio estimates of presented line lengths, for ME, the participants first gave magnitude estimates of presented line lengths, and for CMM and MP, the participants first produced line lengths in response to presented numbers. In a second practice session, the presented stimuli were those used in the experiments—that is, a 1000-Hz tone for RE, ME, and CMM of loudness (and a 1000-Hz tone for the response in MP), the circles used for RE and ME of area, and the weights and sandpapers used for RE of heaviness and roughness, respectively.

For the successive RE conditions, the participants were instructed to compare their current sensation to their previous sensation and to give an estimate of the ratio of sensation magnitudes. They were told that a response of "1" meant that the current sensation was equal in magnitude to the previous sensation, responses greater than "1" indicated that the current sensation magnitude was that many times greater than the previous sensation magnitude, and responses less than "1" indicated that the current sensation magnitude was that many times less than the previous sensation magnitude. The participants were encouraged to use decimals or fractions as needed.

Loudness. For the loudness experiment, 8 participants participated in four conditions, with each condition on a different day. Each participant was run separately. The four conditions were RE of loudness, ME of loudness, CMM of line length to loudness, and MP of loudness to presented numbers; each condition consisted of 200 trials. For each set of 4 participants, the conditions were presented in a digram-balanced Latin square order (i.e., each condition was preceded by every other condition in a balanced design, across participants; see Wagenaar, 1969). The instructions for RE were as described above. For the ME, CMM, and MP conditions, "prior reference" instructions were used—that is, the participants were instructed to compare their current sensation to their previous sensation and produce a response that had the same relation to their previous response; this type of instruction typically leads to large autocorrelation (see DeCarlo, 1994; DeCarlo & Cross, 1990).

For RE, ME, and CMM, each tone was presented for 1 sec. For MP, a number was presented for 2 sec in the middle of the terminal; immediately after the number was presented, a 1000-Hz tone was turned on at one of 50 random initial values ranging from 10 to 90 dB. The participant could then adjust the intensity of the tone in 0.5-dB steps by brushing a stylus or his or her finger across the surface of the KAT; movements to the right increased the intensity, whereas movements to the left decreased the intensity.

Area. For the area experiment, 12 participants participated in two conditions, RE and ME of area; each condition consisted of 120 trials. All 12 participants were run at the same time in a classroom, with RE first, followed by ME, after a 5-min rest period. The instructions for RE were as noted above. For ME, "fixed reference" instructions were used, in that the participants were told to compare each circle to a reference circle, which they were free to choose; this typically leads to small autocorrelation (see DeCarlo, 1994; DeCarlo & Cross, 1990). Each circle was presented for about 3 sec, and the participants wrote their responses on response sheets.

Heaviness and roughness. For the heaviness experiment, 16 participants participated in an RE experiment consisting of 80 trials, with each participant run separately. For the roughness experiment, 20 participants participated in an RE experiment consisting of 80 trials, with each participant run separately. Each weight or sandpaper was presented for about 3 sec, and the participants gave their responses orally, which were then recorded by the experimenter.

RESULTS

Figure 1 presents, for each RE experiment, the means (across participants) of the log responses to each log successive stimulus ratio. (Because there are many possible log stimulus ratios, they were rounded to the nearest 0.05 and the corresponding log responses were averaged.) The trends are clearly linear for the loudness, area, and heaviness experiments. The trend for the roughness experiment is also linear, but shows a slight tendency for stimulus ratios less than unity to be underestimated and stimulus ratios greater than unity to be overestimated.

Time Series Analysis of Residuals

Figure 2 presents, for all four conditions of the loudness estimation experiment, the group autocorrelation function (ACF) and partial autocorrelation function (PACF) for the residuals from fits of regression models. For RE, the regression model is Equation 7 (i.e., the structural part of the model). For ME and CMM, the model consists of a regression of the log response on the current and lagged log stimulus intensity (see, e.g., DeCarlo, 1994). For MP, the model consists of a regression of the log response on the current log stimulus intensity and the lagged log response (see DeCarlo, 2003). The group plots were determined
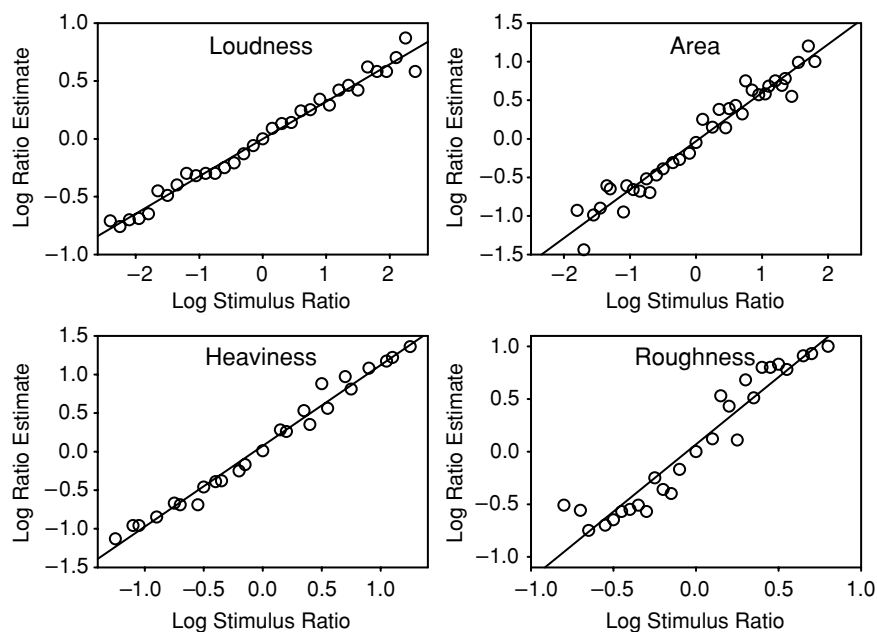
Figure 1. The mean log ratio estimates, across participants, as a function of the log stimulus ratios, for the four successive RE experiments.

in two steps. First, the individual ACFs and PACFs for the residuals were computed and plotted separately for each participant using PROC ARIMA of SAS. The medians and interquartile ranges of the correlations were then computed across participants for each lag and are shown in the figures. The group plots provide a summary of the individual plots. (Estimates of the first-order autocorrelation for each individual are shown in Table 1.)

The top panels of Figure 2 show that successive RE clearly differs from the other magnitude scaling procedures in that the residuals overall show little autocorrelation and the first-order autocorrelation (i.e., the autocorrelation at lag 1) tends to be negative. In contrast, the next two sets of panels show large positive autocorrelation for ME and CMM; the ACF plots show an approximate geometric decay and the PACF plots show a drop after the first lag, both of which are consistent with an AR(1) error process. This is the typical result found in ME and CMM experiments (see, e.g., DeCarlo, 1992, 1994; DeCarlo & Cross, 1990). The bottom set of panels shows that the autocorrelation is small and positive for MP; a comparison with the other panels shows that the magnitude of the autocorrelation is smaller for MP than for ME (with prior reference instructions), which replicates results found by DeCarlo (2003) and Green, Luce, and Duncan (1977).

The top four panels of Figure 3 show group ACF and PACF plots for the RE and ME conditions of the area estimation experiment. Both sets of plots show that the autocorrelation of the residuals is near zero, as has previously been found for ME of area with fixed reference instructions (DeCarlo & Cross, 1990). A comparison of the RE and ME plots shows that the median first-order autocorrelation, though clearly quite small, tends to be

negative for successive RE and positive for ME. The lower four panels of Figure 3 show ACF and PACF plots for the heaviness and roughness successive RE experiments. The plots show that, for both experiments, the autocorrelations are generally small; also note that the first-order autocorrelation, although small, tends to be negative, at least for the roughness experiment.

In summary, a comparison of residual plots for successive RE with those found for ME, CMM, and MP—both here and in earlier research—shows striking differences. First, the autocorrelation in successive RE appears to be generally small. Second, successive RE differs markedly from the other procedures in that the first-order autocorrelation ranges from zero to negative, which is in contrast to the positive autocorrelation that is typically found for ME, MP, and CMM. The next section further examines this result for individual data.

Regression Analysis

Table 1 presents, for the four successive RE experiments, the parameter estimates obtained from maximum likelihood fits of (the structural portion of) Equation 7 to each participant's data. (The individual results for ME, CMM, and MP are similar to those found in earlier research and so are not shown.) Table 2 presents the mean parameter estimates for all the experiments and conditions. With respect to the estimate of the intercept, β_0, Table 1 shows that it is in many cases close to zero and not significant; however, some significant values appear. Table 2 shows that the mean estimates of the intercept are close to zero and nonsignificant for both the loudness and area RE experiments, whereas the heaviness and roughness RE experiments show small and significant positive values.
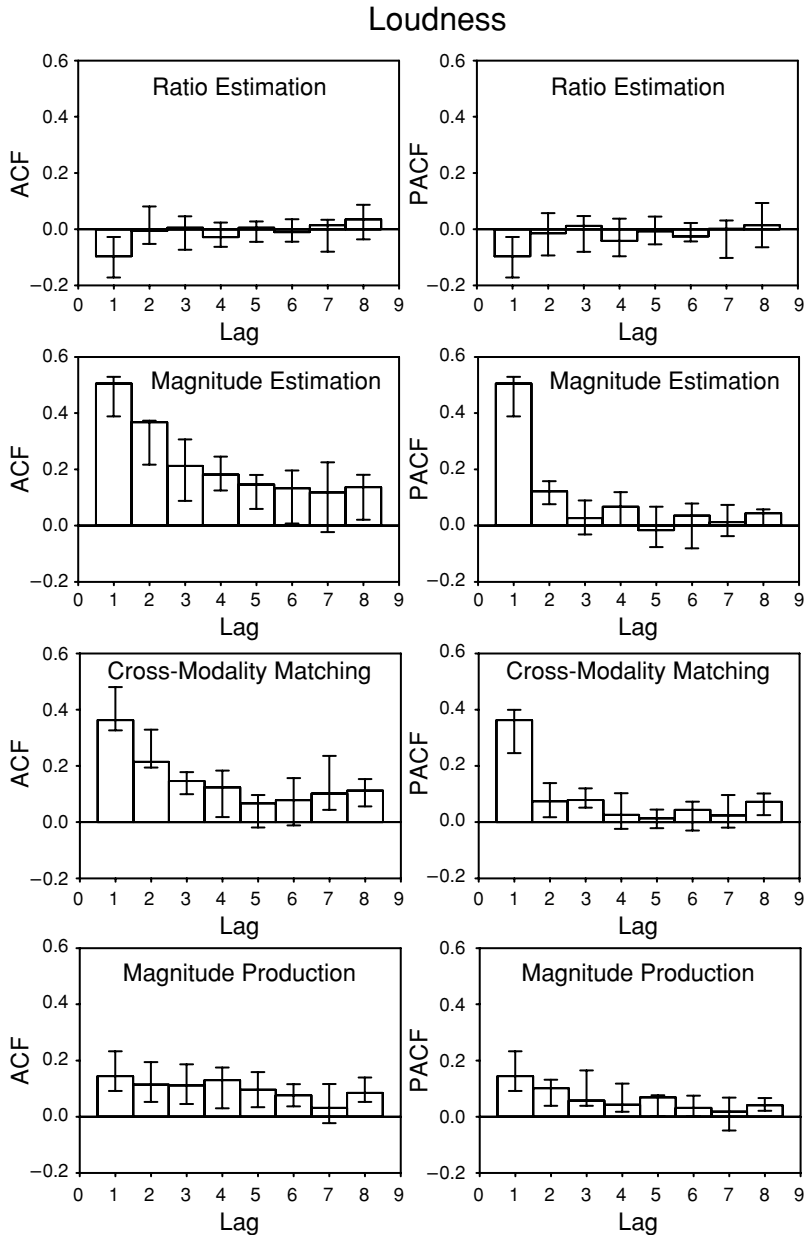
Figure 2. Group autocorrelation functions (ACF) and partial autocorrelation functions (PACF) for the residuals from fits of regression models to the data (as described in the text) from the four conditions of the loudness estimation experiment.

Thus, there appears to be a constant bias in some cases, but it generally appears to be rather small in magnitude.

Table 2 shows that, for all four conditions, the mean estimates of the exponents for loudness are less than unity, as is typically found. The mean exponents for area are less than unity but larger than those for loudness, as is also typically found (see DeCarlo & Cross, 1990; Stevens, 1986). The mean exponents for both heaviness and roughness are greater than unity, which is consistent with earlier research (e.g., DeCarlo, 1994; Stevens, 1986). The mean estimates of the exponent β will be discussed in more detail in the section on response bias, below.

Of main interest here are sequential effects in successive RE, which are reflected by the stimulus context effect parameter γ and the autocorrelation. With respect to γ, Table 1 shows that the individual estimates are small but generally positive, and significant in many cases. Table 2 shows that, for loudness estimation, the mean estimates of γ are significant and in the range of 0.07–0.10 across RE, ME, and CMM, which is similar to results found in other studies (DeCarlo, 1994; DeCarlo & Cross, 1990). For MP, the mean estimate of γ is close to zero and nonsignificant, which replicates results found by DeCarlo (2003) and Green et al. (1977). DeCarlo (2003) noted that a stimu-

Table 1
Parameter Estimates for Each Participant in the Four RE Experiments

Participant	Loudness				Area				Heaviness				Roughness			
	β_0	β	γ	AR(1)	β_0	β	γ	AR(1)	β_0	β	γ	AR(1)	β_0	β	γ	AR(1)
1	-0.10**	0.35	0.05*	.05	-0.03	0.86	-0.11**	.02	0.03	1.35	0.16*	.02	0.02	1.63	0.29*	.02
2	0.06**	0.34	0.07**	.01	0.05**	0.69	0.03	-.08	0.04	0.97	0.02	-.03	0.01	0.92	0.05	-.14
3	-0.02	0.52	0.09**	-.22**	-0.06**	0.78	0.04*	-.14	0.07	0.73	0.11	.06	-0.02	1.61	0.24*	.10
4	0.03**	0.29	0.03*	-.28**	-0.02	0.79	0.05	-.21*	0.10*	0.63	0.07	-.08	0.02	1.18	-0.09	.01
5	-0.01	0.31	0.07**	-.10	0.04	0.95	0.01	-.04	0.02	1.29	0.11	-.17	0.06	1.15	0.15	.16
6	-0.02	0.28	0.09**	-.10	-0.04	0.68	0.08*	-.03	0.03	1.08	0.08*	-.22*	-0.04	1.04	0.13	-.12
7	0.00	0.42	0.10**	-.06	0.04	0.47	0.07*	-.03	0.14**	1.29	0.20**	-.11	-0.06	2.21	0.18	-.26**
8	0.09**	0.40	0.11**	-.13*	0.01	0.51	0.07*	.02	0.14**	1.65	0.12	.08	-0.00	1.28	0.11	-.24*
9					0.07**	0.40	0.14**	-.02	0.12**	0.88	0.16	.14	0.31**	2.01	-0.03	-.09
10					-0.19**	0.80	0.10	-.06	0.16**	1.01	0.12	.03	-0.02	2.41	0.24	-.36**
11					0.01	0.55	0.03	-.37**	0.11**	1.41	0.15	-.02	0.09*	2.12	0.25*	-.19*
12					-0.17**	0.64	0.25**	.07	-0.01	1.71	0.08	-.08	-0.02	1.16	0.23	.10
13									0.02	1.09	0.17**	.08	0.05	1.00	0.13	-.09
14									0.05	1.56	0.04	-.26**	0.05	1.51	0.24*	.03
15									0.03	0.73	-0.00	-.21*	0.01	1.92	0.01	-.35**
16									0.14**	1.12	0.28**	.14	0.04*	0.79	0.25**	-.04
17													0.15**	1.07	0.24*	.11
18													0.21**	1.38	0.23	.09
19													-0.02	1.11	0.20	-.13
20													0.09**	1.46	0.07	-.16

Note—AR(1) is an estimate of the first-order autocorrelation of the residuals, the significance of which was determined by using the Durbin–Watson test. Significance is not shown for β because all the values are greater than zero.   *$p < .05$.   **$p < .01$.
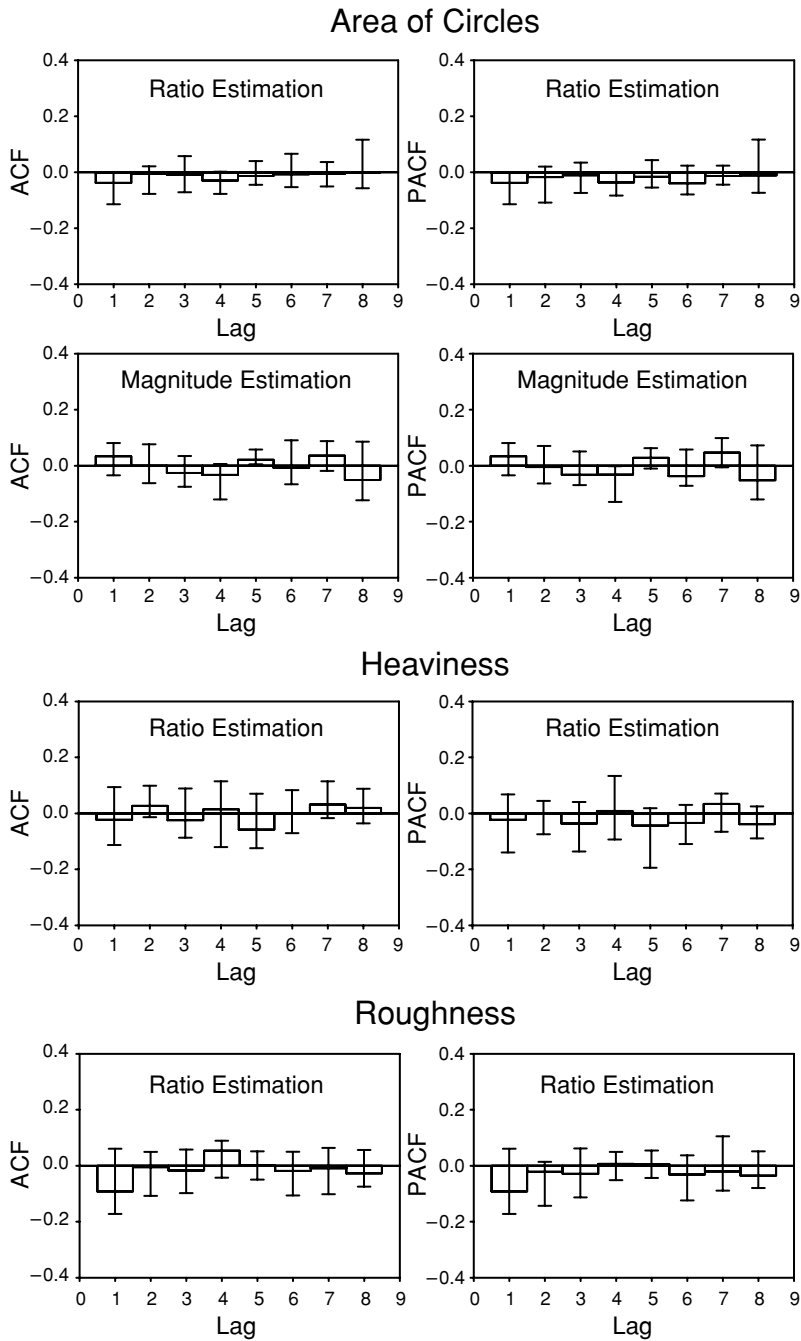
Figure 3. Group autocorrelation functions (ACF) and partial autocorrelation functions (PACF) for the residuals from fits of regression models to the data from the area, heaviness, and roughness experiments.

lus context effect might not be found in MP of loudness because the participant adjusts the loudness through different levels on each trial, which might eliminate any effect of the previously produced loudness on the currently produced loudness (note that this is why the effect is not included for the produced continuum in CMM). For area estimation, the mean estimates of γ are around 0.05 and

are significant for both RE and ME. Thus, the loudness and area experiments both show that the estimates of γ are consistent in sign and magnitude across RE and ME. For the heaviness and roughness RE experiments, the mean estimates of γ are positive and significant. Overall, the tables show that a stimulus context effect appears in successive RE of loudness, area, heaviness, and roughness;

Table 2
Mean Parameter Estimates for Each Experiment and Condition

Condition	β_0	β	γ	AR(1)
Loudness				
Ratio estimation	<0.01	0.36	0.08**	−.10*
Magnitude estimation	−1.73**	0.49	0.10**	.48**
Cross-modality matching	−0.96**	0.63	0.07**	.33**
Magnitude production	−0.93**[a]	0.74[b]	<−0.01[c]	.17**
Area				
Ratio estimation	−0.02	0.68	0.06*	−.07
Magnitude estimation	−3.37**	0.86	0.04*	.04
Heaviness				
Ratio estimation	0.07**	1.16	0.12**	−.04
Roughness				
Ratio estimation	0.05*	1.45	0.16**	−.08*

Note—Results of t tests on the parameter estimates are shown for β_0, γ, and AR(1); all the estimates of β are significant. [a]The estimate of the intercept in MP is obtained by multiplying the obtained intercept estimate (1.26) by the estimate of $-\beta$ (−0.74); see Equation 9 of DeCarlo (2003). [b]The estimate of β is the inverse of the obtained exponent estimate (1.36). [c]The estimate of γ in MP (−0.003) is obtained by multiplying the obtained exponent (0.004) by the estimate of $-\beta$ (see DeCarlo, 2003). *$p < .05$. **$p < .01$.

further, the effect is in the same direction (positive) as is found in ME and CMM and is similar in magnitude across RE and ME (for loudness and area).

The last columns of Tables 1 and 2 show estimates of the first-order autocorrelation [labeled as AR(1)] of the residuals from a fit of (the structural part of) Equation 7. Table 1 shows that estimates of the residual correlation for each individual in successive RE are small and nonsignificant in many cases. In some cases, however, the estimates differ from zero and are significant. Note that in all cases in which the autocorrelation is significant, it is negative. Successive RE is the only magnitude scaling procedure to date in which significant negative autocorrelation has been found, to my knowledge. For example, examination of individual results for 84 participants in ME, MP, and CMM experiments presented in DeCarlo (1992, 1994, 2003) and DeCarlo and Cross (1990) shows that there were no cases with significant negative autocorrelation. With respect to average data, the mean autocorrelation coefficients in these and other studies were also all positive, as was found here for ME, CMM, and MP, as shown in Table 2. In contrast, Table 2 shows that, for all four successive RE experiments, the mean residual autocorrelation was small and negative, and was significant for the loudness and roughness experiments.

DISCUSSION

Autocorrelated Errors in Successive RE

The participants' task in successive RE is to estimate the ratio of their current sensation magnitude to their previous sensation magnitude. If they in fact do this, the error process will not be random over time, because of the consequences of comparing sensation magnitudes and perceptual error, as shown by Equations 1–3. In particular,

it follows from Equation 3 that the errors in successive RE should show autocorrelation ranging from zero to −0.5, as shown by Equation 5; Tables 1 and 2 show that this was indeed the case. Most participants did not show significant autocorrelation, but when the autocorrelation was significant, it was negative. This is important evidence that is consistent with the idea that participants compare sensation magnitudes as in Equation 1.

Equation 5 shows that magnitude of the (negative) residual autocorrelation for Equation 7 depends on the ratio of the variability of the judgmental error to perceptual error. This reveals an interesting aspect of the successive RE procedure, in that it suggests that the residual autocorrelation provides information about the ratio of judgmental to perceptual error variance (on a log scale). As shown in Table 2, the mean autocorrelation for the loudness study was −0.1, which implies a ratio of judgmental to perceptual error variance of 8 (from Equation 5). This suggests that the sensation variability was quite small compared with variability due to judgment, which is quite plausible in my view. Note that Equation 5 also suggests some interesting experiments. If one could manipulate the variability of the sensation, for example (perhaps by increasing the intertrial interval), one could obtain evidence as to the validity of the model: Greater variability in the sensation (relative to judgment), for example, implies larger negative autocorrelation. Predictions of this sort remain to be explored.

Stimulus Context Effects in Successive RE

A simple model of stimulus context effects (Equation 2), when incorporated into the classical model for ratio estimation (Equation 1), shows how to obtain an estimate of the context effect, as measured by γ. In particular, Equation 7 shows that in order to obtain a direct estimate of γ, $\log(S_{t-1}/S_{t-2})$ should be included as a regressor and not simply $\log S_{t-1}$ alone. This shows why it is important to explicitly present the mathematical formulation of the theories. With this clarification in hand, the above experiments show that a positive stimulus context effect was present for successive RE of loudness, area, heaviness, and roughness, which agrees with earlier research on ME and CMM with these continua. The effect also appears to be similar in magnitude to that found in ME and CMM. The finding of a positive value of γ is also consistent with several results for successive RE (of loudness) presented by Lockhead and King (1983), as shown in the Appendix. Thus, the present results as well as earlier research suggest that a positive stimulus context effect appears in successive RE for several continua, as is also found in ME and CMM.

On Some Extensions and Alternatives

There are many possible ways to extend the models given above; here, I briefly comment on some possible extensions. The focus here has been on effects extending back to the previous trial, although one can consider effects extending back to more than one trial, either for the error process (i.e., autocorrelation) or for the stimulus context model. For example, for successive RE, the present results show that the first-order autocorrelation is zero

or at most is small, and there is no apparent evidence of any higher-order effects, as shown by the PACFs (which are used in time series analysis to determine the depth of effects in autoregressive processes; see Box & Jenkins, 1976; DeCarlo, 1994). For ME, CMM, and MP, prior research has shown that considering effects from one trial back is often adequate, in that the PACFs tend to have a value near zero at lag 2, or at most a small positive value (see, e.g., DeCarlo, 1994; 2003), and this was also found here for ME, CMM, and MP (Figures 2 and 3). Thus, a first-order error structure seems to be adequate at this point, particularly for successive RE.

With respect to stimulus context effects, Equation 6 can be extended to allow for effects from earlier trials. For example, if S_{t-2} is included in Equation 6, then it follows from the classical model of Equation 1 that $\log(S_{t-2}/S_{t-3})$ should be included as a regressor in the model. Doing this for the present RE experiments gave average (across participants) estimates of the coefficient of $\log(S_{t-2}/S_{t-3})$ of 0.03, 0.06, 0.09, and 0.09 for the loudness, area, heaviness, and roughness RE experiments, respectively (all the estimates are significant at the .05 level). For the ME, CMM, and MP conditions, including a term for effects from two trials back gives a value of 0.03 ($p < .01$) for ME of loudness, but small (0.01 or less), nonsignificant values for CMM and MP of loudness and for ME of area. Thus, for the present experiments, there appears to be a small positive stimulus context effect that extends two trials back for successive RE, but there is little or inconsistent evidence of an effect for ME, CMM, and MP. Perhaps the effect of earlier stimuli is larger in successive RE than in ME because the RE task requires participants to focus on the previous trial; further research on this is needed.

As shown above, the classical model for RE combined with the psychophysical function (with an error term) together account for the negative autocorrelation found in some cases for successive RE. Although this is probably the simplest possible account of this result, it is not the only one. Here, I briefly note that another approach is to assume that ratio estimates are derived from unobserved (latent) magnitude estimates. If one then uses the dynamic model of judgment (see, e.g., DeCarlo, 1994), which was developed for ME, as the model for the latent magnitude estimates, it can be shown that this predicts zero to negative autocorrelation, but with a different interpretation of the autocorrelation. This approach offers the benefit of unifying models for ME, MP, and CMM with the model for successive RE. Although I detailed this approach in an earlier version of this article, the reviewers pointed out that validity studies need to be done first, to demonstrate any advantages of a more complex approach. In short, at this point, Equations 1 and 2 probably provide the simplest possible account of negative autocorrelation in successive RE.

Some Notes on Response Bias in Successive RE

One type of response bias was considered above—namely, constant response bias, which was found for some participants. Constant response bias can be viewed as a ten-dency to consistently overestimate or underestimate sensation ratios. Another type of response bias, recently discussed for ME, MP, and CMM (DeCarlo, 2005), allows for deviations of judgments from proportionality. Here, I note some implications of this approach for successive RE.

To allow for deviations from ratio estimation in successive RE, the classical model of Equation 1 can be generalized as follows:

$$R_{t,t-1} = \left(\frac{\Psi_t}{\Psi_{t-1}}\right)^b \upsilon_t, \qquad (8)$$

where b is a bias parameter. This type of generalization has been shown to be quite useful for ME, MP, and CMM (see DeCarlo, 2005). Equation 8 allows for response bias in successive RE—that is, systematic deviations of the ratio estimates from the sensation ratios.

The implications of Equation 8 for the models presented here are straightforward. The only changes in Equation 7 are that the exponent of S_t/S_{t-1} is now $b\beta$ instead of simply β and the exponent of S_{t-1}/S_{t-2} is $b\gamma$ instead of γ; this is also the case for ME and CMM. Note that the inclusion of a bias parameter only affects the parameters of the structural portion of the model; it has no effect on the error process of Equation 7 (and similarly for ME, CMM, and MP). For MP, the exponents of the model are estimates of b/β and $-b\gamma/\beta$, respectively (see DeCarlo, 2003, 2005). Note that if it is assumed that the bias is the same across ME and MP, then taking the geometric mean of the ME exponent, $b\beta$, and the inverse of the MP exponent, β/b, gives an estimate of β corrected for bias, as earlier conjectured by Stevens, whereas taking the square root of their ratio gives an estimate of the bias b (see DeCarlo, 2005).

The loudness experiment presented here included both ME and MP conditions, so one can use the approach noted above to obtain estimates of the corrected exponent and the bias. Using the mean estimates shown in Table 2, this gives an estimate of β of 0.59, which is consistent with values found in other loudness estimation experiments; these tend to be in the range of 0.6–0.7, as shown in Table 1 of DeCarlo (2005). The estimate of the bias b is 0.81, which is also consistent with values obtained in other studies (see Table 1 of DeCarlo, 2005). Thus, the results for the ME and MP conditions of the present loudness experiment are in agreement with those found in several other ME and MP loudness experiments, with respect to both corrected exponents and bias.

The loudness experiment also included CMM and successive RE conditions. If it is assumed that the corrected exponent obtained via ME and MP (about 0.6) also applies to these other conditions, then one can obtain an estimate of the bias in CMM and successive RE by dividing the obtained exponents (which are estimates of $b\beta$) by the corrected exponent of 0.6 (in which case β cancels out and one is left with an estimate of the bias, that is, $b\beta/\beta = b$). Note that this approach only assumes that the bias is equal across ME and MP, whereas no assumptions about the bias in CMM are made (see DeCarlo, 2005). For the present loudness experiments, this gives an estimate of bias for

the CMM experiment of 1.05, which suggests that there was no bias for CMM. For the successive RE experiment, dividing the obtained exponent by 0.6 gives an estimate of bias of 0.6, which is fairly large (values further from 1.0 indicate larger bias). Thus, the results suggest that response bias in successive RE of loudness (0.6) was larger than that for ME (0.8). This is summarized by the ratio of estimated biases (RE/ME), which is less than unity, $0.6/0.81 = 0.74$.

The area estimation experiment only included successive RE and ME conditions, so one cannot determine the separate biases for RE and ME, only their ratio (i.e., the exponent is $b'\beta$ for RE, where b' is the bias for RE, and the exponent is $b\beta$ for ME, so taking the ratio of obtained exponents gives $b'\beta/b\beta = b'/b$, which is the ratio of RE to ME biases). An estimate obtained in this way, using the mean exponents shown in Table 2, gives a value of 0.79. This is important, because it shows that the results for area estimation are consistent with those for loudness estimation, in that they again suggest that response bias for successive RE was larger than that for ME (because the ratio of RE/ME biases is again less than unity). Also note that the estimate of the ratio of the biases for area estimation, 0.79, is consistent in magnitude with that found above for loudness estimation, 0.74.

In summary, a simple model of response bias applied to successive RE reveals some interesting results. In particular, the results for the present experiments suggest that there was a consistent response bias that was less than unity for both successive RE and ME, with the bias being larger for RE than for ME. Simply put, this suggests that participants tended to constrict their range of responses more in RE than in ME, perhaps reflecting an aspect of using fractions and ratios. Note that the bias also affects the estimates of the stimulus context parameter γ, but the corrections in this case are rather small and make little difference (e.g., the estimate of γ for loudness changes from 0.08 to 0.10).

CONCLUSIONS

Successive RE, though not widely investigated, is an interesting procedure, in that the physical procedure is exactly the same as in ME, but the nature of the task is different. Thus, any differences in sequential effects across successive RE and ME must be due to the different nature of the task. The present experiments show that successive RE gives results that are both similar to and different from those found for ME. First, it was shown that a stimulus context effect appears in successive RE, as is also found in ME and CMM. Further, the effect is in the same direction (positive) and is of about the same magnitude across successive RE and ME (but is perhaps slightly larger for RE). A positive stimulus context effect is also consistent with several results presented earlier by Lockhead and King (1983). Second, it is shown that the residuals from a fit of the classical RE model generally show only small autocorrelation, if any. Successive RE is somewhat unique, however, in that if the regression residuals are autocorrelated, the autocorrelation tends to be negative, whereas

positive autocorrelation is typically found for ME, CMM, and MP. It is shown that the negative autocorrelation follows from the classical model of ratio estimation when the effects of perceptual variability are recognized. The finding of negative autocorrelation is evidence in favor of the idea that participants compare sensation ratios, as instructed. Finally, response bias with values of $b < 1$, as is typically found for ME, CMM, and MP (DeCarlo, 2005), also appears in successive RE. In addition, both the loudness and area experiments suggest that the bias was larger for successive RE than for ME. Overall, the present experiments show that there are invariant aspects of sequential effects and response bias across a variety of magnitude scaling procedures. Explicit development of the mathematical theory associated with these procedures helps to reveal their similarities and differences, and sheds light on underlying cognitive and perceptual processes.

REFERENCES

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

Cross, D. V. (1973). Sequential dependencies and regression in psychophysical judgments. *Perception & Psychophysics*, **14**, 547-552.

DeCarlo, L. T. (1990). Sequential effects in magnitude scaling (Doctoral dissertation, SUNY, 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 & Performance*, **18**, 1080-1088.

DeCarlo, L. T. (1994). A dynamic theory of proportional judgment: Context and judgment of length, heaviness, and roughness. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 372-381.

DeCarlo, L. T. (2003). An application of a dynamic model of judgment to magnitude production. *Perception & Psychophysics*, **65**, 152-162.

DeCarlo, L. T. (2005). On bias in magnitude scaling and some conjectures of Stevens. *Perception & Psychophysics*, **67**, 886-896.

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

Fagot, R. F. (1978). A theory of relative judgment. *Perception & Psychophysics*, **24**, 243-252.

Gescheider, G. A. (1997). *Psychophysics: The fundamentals* (3rd ed.). Mahwah, NJ: Erlbaum.

Green, D. M., Luce, R. D., & Duncan, J. E. (1977). Variability and sequential effects in magnitude production and estimation of auditory intensity. *Perception & Psychophysics*, **22**, 450-456.

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

Lockhead, G. R., & King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 461-473.

Sjöberg, L. (1971). Three models for the analysis of subjective ratios. *Scandinavian Journal of Psychology*, **12**, 217-240.

Stevens, S. S. (1986). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New Brunswick, NJ: Transaction.

Wagenaar, W. A. (1969). Note on the construction of digram-balanced Latin squares. *Psychological Bulletin*, **72**, 384-386.

Ward, L. M. (1973). Repeated magnitude estimations with a variable standard: Sequential effects and other properties. *Perception & Psychophysics*, **14**, 193-200.

Ward, L. M. (1979). Stimulus information and sequential dependencies in magnitude estimation and cross-modality matching. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 444-459.

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

APPENDIX

Autocorrelated Errors and the Classical Model

This section derives the first-order autocorrelation for the errors of Equation 7. The error process is $e_t = d_t - d_{t-1} + u_t$. The perceptual and judgmental error terms, d_t and u_t, respectively, are assumed to be random with standard assumptions: $E(d_t) = E(d_{t-1}) = E(u_t) = 0$, $\text{Var}(d_t) = \text{Var}(d_{t-1}) = \sigma^2$, $\text{Var}(u_t) = \sigma_u^2$, $\text{Cov}(d_t, d_{t-1}) = 0$, $\text{Cov}(d_t, u_t) = 0$, and $\text{Cov}(d_{t-1}, u_t) = 0$, where E is the expectation operator, Var is the variance operator, and Cov is the covariance operator. It follows that

$$\text{Var}(e_t) = \text{Var}(d_t) + \text{Var}(d_{t-1}) + \text{Var}(u_t) - 2\text{Cov}(d_t, d_{t-1}) - 2\text{Cov}(d_{t-1}, u_t) + 2\text{Cov}(d_t, u_t)$$
$$= 2\sigma^2 + \sigma_u^2, \tag{A1}$$

which also holds for $\text{Var}(e_{t-1})$. Next, note that

$$\text{Cov}(e_t, e_{t-1}) = E(e_t \times e_{t-1}) = E\big[(d_t - d_{t-1} + u_t) \times (d_{t-1} - d_{t-2} + u_{t-1})\big] = -\sigma^2, \tag{A2}$$

(because the expectations of all of the cross-product terms are equal to zero). It then follows from A1 and A2 that

$$\text{Corr}(e_t, e_{t-1}) = \text{Cov}(e_t, e_{t-1}) / \big[\text{Var}(e_t) \times \text{Var}(e_{t-1})\big]^{1/2} = -\sigma^2 / \big(2\sigma^2 + \sigma_u^2\big) = -1 / \big[2 + \big(\sigma_u^2 / \sigma^2\big)\big], \tag{A3}$$

as noted in the text. Note that the bias generalization given by Equation 8 in the text only affects the mean responses and has no effect on the error structure, so Equation A3 still holds.

Previous Evidence on the Sign of γ

Equation 7 clarifies several results found by Lockhead and King (1983). Lockhead and King examined successive RE in a loudness estimation experiment (1000-Hz tone); their experiment differed slightly from those presented here in that participants received "feedback" on each trial. Nevertheless, the models presented here account for several of their results. For example, Figure 2 of Lockhead and King (1983) shows that the ratio estimates were larger when S_{t-1} was smaller (actually, categories of S_{t-1} were used) and were smaller when S_{t-1} was larger. Thus, the previous stimulus intensity S_{t-1} appears to exert a negative or "contrast" effect on the current ratio estimate.[A1] This result follows immediately from Equation 7, which can be seen by rewriting it as follows,

$$\log R_{t,t-1} = \beta_0 + \beta \log S_t + (\gamma - \beta) \log S_{t-1} - \gamma \log S_{t-2} + u_t, \tag{A4}$$

(note that the error structure doesn't affect the results presented here and so a simple error term u_t is used). The above shows that if γ is positive and less than β, which is the usual case, then $\gamma - \beta < 0$ and so $\log S_{t-1}$ will appear to exert a "contrast" effect (i.e., a negative effect) on the ratio estimates, even though the actual effect (γ) is positive. Thus, Equation 7 accounts for the contrast effect shown in Figure 2 of Lockhead and King (1983).

Lockhead and King (1983) also considered the situation in which the current and previous stimulus intensities were equal, $S_t = S_{t-1}$. The left panel of Figure 3 of Lockhead and King shows that when successive stimulus intensities were equal, the ratio estimates tended to be smaller when S_{t-1} was small and larger when S_{t-1} was large (categories of S_{t-1} were again used). It is important to recognize that this result follows immediately from Equation 7 with a positive value of γ. To see this, note that when $S_t = S_{t-1}$, Equation 7 reduces to the following,

$$\log R_{t,t-1} = \beta_0 + \gamma \log S_{t-1} - \gamma \log S_{t-2} + u_t. \tag{A5}$$

Equation A5 shows that, if γ is positive, then the ratio estimates will be smaller when S_{t-1} is small and larger when S_{t-1} is large, exactly as found by Lockhead and King. Equation A4 also accounts for a result shown in Figure 5 of Lockhead and King. This figure shows that, when successive stimuli were equal, the ratio estimates tended to be smaller when S_{t-2} was large and larger when S_{t-2} was small, and so S_{t-2} appears to exert a "contrast" effect on the ratio estimates. Note that this result also follows immediately from Equation A5, because if γ is positive, then the coefficient of S_{t-2} is negative, which is consistent with the apparent contrast shown in Figure 5 of Lockhead and King.

NOTE

A1. The plots of Lockhead and King (1983) were for response and stimulus intensities, and not their logarithms, which might account for why some curvature is apparent in their plots. In any case, the predictions discussed in the Appendix remain the same.