Psychological Research (2008) 72:115–122
DOI 10.1007/s00426-006-0103-1

ORIGINAL ARTICLE

The influence of trial-to-trial recalibration on sequential effects in cross-modality matching

Martha Teghtsoonian · Robert Teghtsoonian ·
Lawrence T. DeCarlo

Received: 7 April 2006 / Accepted: 19 October 2006 / Published online: 15 November 2006

Abstract Sequential effects are examined in a cross-modality matching experiment where observers adjusted the loudness of a tone in response to presented lengths of a metal tape. In one condition, the initial level of the tone to be adjusted was the same as the final adjusted level of the previous trial, whereas in another condition, the tone to be adjusted was reset to a different level before each trial. A fit of the DeCarlo–Cross dynamic model shows that the primary effect of the manipulation was on a judgmental factor, with little or no effect on a perceptual factor. We suggest that starting a trial with the tone at the final adjusted level of the previous trial induced the observer to rely more heavily on the loudness-length pair of the previous trial as a frame of reference for relative judgment; we call this reliance trial-to-trial recalibration. In contrast, when the tone is set to a level independent of its value on the previous trial, there is virtually no effect of one trial on the next trial's performance, a result consistent with the observer maintaining a stable frame of reference. We argue that sequential effects are not unavoidable and that the technique described here adds to a growing list of methods for reducing or eliminating them.

Introduction

When a listener judges the loudness of a series of acoustic signals, the response may depend not only on the amplitude of the signal being judged, but also on events from previous trials. One focus of research on these *sequential effects* is on theoretical questions about underlying processes: For example, do features of the previous trial affect the perception of the stimulus on a given trial, the choice of a response, or both? A second focus is on determining the relevant features of the previous trial: Are the effects on the following trial governed by the stimulus, the response, or some combination of the two? However, beyond these specific issues, a more general question may be posed. Sequential effects belong to a broader set of factors—e.g., the composition of the set of stimuli being studied, the spacing among them, and the order in which they are presented—that can influence performance, independently of signal amplitude, on any given trial. Some (e.g., Lockhead, 1992; Schneider & Parker, 2002) have asserted that, taken together, these *context effects* make it impossible to find a simple general principle—like that proposed by Teghtsoonian (1971, 1973)—about the relation between a given sensory input and the response to it. Two of us have argued previously for a less pessimistic position and presented evidence that at least some context effects can be understood and controlled (Teghtsoonian & Teghtsoonian, 1983, 1978; Teghtsoonian, Teghtsoonian, & Baird, 1985). In this report, we show how sequential effects in a particular experimental paradigm can be made either to occur or to be nearly eliminated by a simple procedural adjustment. We also show that the effect of this adjustment can be described

---

The data on which this paper is based were presented as a poster by Martha Teghtsoonian at the 1992 meeting of the International Society for Psychophysics, Stockholm, Sweden.

M. Teghtsoonian (✉) · R. Teghtsoonian
Department of Psychology, Smith College,
Northampton, MA 01063, USA
e-mail: mtex@smith.edu

L. T. DeCarlo
Teachers College, Columbia University,
New York, USA

and understood by the application of a suitable mathematical model and an associated theory. Finally we argue that this outcome encourages further studies seeking to identify and control or eliminate other context effects. If such a program succeeds, the effort to develop and refine general laws relating subjective magnitudes to stimulus intensities will continue to be a feasible goal of psychological science.

Early in the development of direct scaling techniques (e.g., magnitude estimation), sequential effects were regarded as a potential source of error but one that might be controlled by randomizing or counterbalancing stimulus order (Stevens, 1986). The premise was that sequence effects were of no intrinsic interest. But others have focused on the phenomenon itself (e.g., Garner, 1953; Cross, 1973; Ward, 1973) and have worked to develop clear descriptions of the nature and degree of those effects. For example, both Cross and Ward demonstrated that, in magnitude estimation, the judged loudness of an auditory stimulus on the current trial is positively correlated with the stimulus intensity on the previous trial, an effect characterized by Cross as *assimilation*; Ward (1975) found a similar effect in cross-modality matching.

To describe the effect, Cross (1973) proposed a model in which, in addition to the stimulus presented on the current trial, S_t, the stimulus presented on the previous trial, S_{t-1}, was included. In the corresponding regression equation the coefficient of log S_{t-1} provides a measure of the degree of assimilation.[1] However, this model leaves open the question of whether assimilation is the result of a perceptual effect: It may be that the observer's experience with S_{t-1} affects the *perception* of S_t, but it may also be that it independently affects the *judgmental process* on trial t. Responding to this ambiguity, later investigators (e.g., Jesteadt, Luce, & Green, 1977; see Appendix A) have proposed more complex models, but the one we will consider here is based on the work of DeCarlo (1989/1990) and DeCarlo and Cross (1990); we refer to their work as the Relative Judgment Model. It has been described in detail and its implications have been explored elsewhere (DeCarlo, 1992, 1994, 2003). Here we provide only a summary of its major features.

The Relative Judgment Model can be written as a regression model as follows:

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

where R_t is the response on a given trial, S_t and S_{t-1} are the stimulus values on that trial and the preceding trial respectively, and u_t is a random error term for Trial t. The novelty (compared to earlier models) is the nonrandom error term e_{t-1}, and its meaning requires some explanation. Earlier models (notably that of Jesteadt et al., 1977) included a term for R_{t-1} as an attempt to evaluate the role of the response provided on the preceding trial. Although this is a plausible idea, there are some difficulties associated with it, some of which are noted in Appendix A.

Another approach focuses on the role of frames of reference used by the observer in order to make proportional responses, where the frame consists of the *pairing* of a response and sensation. In particular, it is important to recognize that there are two ways for observers to make proportional judgments (DeCarlo, 1989/1990). One option is to use one response–sensation pair as a fixed reference, and to make all comparisons to that reference. For example, one can assign a reference response, R_0, to a reference sensation, Ψ_0. This leads to a simple model of proportional judgment, as assumed by Stevens (1986). As noted earlier (e.g., DeCarlo, 2003), the response–sensation pair defines a unit of measurement, $\alpha_0 = R_0/\Psi_0$, which is assumed to remain constant over trials. Another option is to use the response–sensation pair from the previous trial as the reference. In this case, the unit of measurement defined by the observer, $\alpha_{t-1} = R_{t-1}/\Psi_{t-1}$, is not constant, but rather varies over trials. We think of this second strategy as constituting *trial-to-trial* recalibration: Each trial provides a new standard and modulus for use in making a judgment on the following trial. A consequence of this second strategy is that judgmental error will propagate over trials, and so will be correlated over trials.

In the Relative Judgment Model, both frames of reference are viewed as having an influence on judgment, which in turn leads to an autocorrelated error process, $\lambda\, e_{t-1} + u_t$, as shown in Eq. 1; the derivation of the autocorrelated error process from the weighted effects of the two frames of reference is shown in DeCarlo (1994, 2003) and is summarized in Appendix B. The parameter λ provides a measure of the relative influence of the two frames of reference: a value of zero indicates that a fixed response–sensation pair is used as the reference, whereas a value of unity indicates that only the response–sensation pair from the previous trial is used as the reference; values between zero and unity indicate how heavily each of these references is weighted in judgment. Thus, the magnitude of λ indicates the degree of trial-to-trial recalibration: When λ is high, judgments on a given trial are heavily influenced by the response–sensation pairing of the preceding trial, and when λ is low, judgments are heavily influenced by a stable response–sensation pairing (i.e., a stable unit).

---

[1] Cross included S_{t-1}/S_t as a term in his model, to allow for assimilation, with the result that S_{t-1} is included in the regression model (see DeCarlo, 2003, for details).

There is now considerable evidence showing the value of the DeCarlo–Cross model and the associated theory. For example, the Relative Judgment Theory suggests how to gain control over the autocorrelation: It has been shown in several studies (DeCarlo, 1989/1990, 1994; DeCarlo & Cross, 1990) that the magnitude of the autocorrelation is reduced or eliminated when participants are encouraged (by verbal instruction) to make their judgments relative to a fixed reference point (i.e., a fixed stimulus–response pair), whereas the magnitude of the autocorrelation is large when participants are instructed to make their judgments relative to the stimulus–response pair of the preceding trial. Second, the coefficient of S_{t-1} for the DeCarlo–Cross model is positive for loudness estimation, which is consistent with earlier findings of assimilation for loudness. Third, the coefficient of S_{t-1} decreases with an increase in the inter-trial interval (DeCarlo, 1992): The longer the time between trials, the smaller the value of the coefficient of S_{t-1}, as expected. In addition to these studies, which used the method of magnitude estimation (ME), a recent report has shown similar results for magnitude production (MP; DeCarlo, 2003). Thus, the value of the DeCarlo–Cross model and the associated theory of relative judgment have been demonstrated in a variety of experimental contexts.

We undertook the study reported here for two purposes. First, we wished to examine sequential effects in cross-modality matching (CMM) using the Cross paradigm (each of the six lengths preceded every other length, including itself, just once). Second, we wished to look at the effects on trial-to-trial recalibration of varying the accessibility of the response on Trial t–1 when S_t was judged. Observers adjusted the loudness of a tone to match the length of a metal tape. In one condition, each trial began with the tone at the level to which it had been adjusted on the previous trial; in another condition, the initial level of the tone was reset and was therefore independent of the intensity produced on the preceding trial. The DeCarlo–Cross model was fitted to the data, so that the roles of both the preceding stimulus and the preceding stimulus–response linkage could be evaluated in the two versions of the CMM procedure.

Method

Observers

The observers were 30 students, volunteers from an undergraduate psychology course at a women's college who received course credit for their participation.

Materials

Stimulus lengths were produced by extending a retractable metal tape to the designated length; the width of the tape was 1.2 cm. The experimenter and the observer sat on opposite sides of a table covered with a black cloth. The tape length was adjusted by the experimenter and was placed on the table in the observer's frontoparallel plane with the unmarked white side of the tape facing the observer. The midpoint of any given length was centered in the observer's field of view, at a viewing distance of approximately 75 cm. To make a matching response, observers produced sound intensities. A 1-kHz tone was provided by a tone generator (Philips GM 2306E) and passed through a sone potentiometer and a buffer amplifier unit to Madsen earphones worn by the observer. By turning the dial of the unit, the observer could vary the intensity of the tone from 40 to 100 dB SPL. An rms voltmeter (Ballantine 323-01) was used to measure sound intensity.

Design and procedure

Six lengths—1, 2.5, 6.3, 16, 40, and 100 cm—were presented as described above. The observer's task was to set the loudness of the tone so that it matched the length of the tape; she was to set the loudness to an appropriate level for the first length, then, on subsequent trials, to make the tone as much louder or softer as the tape appeared longer or shorter. After each length was presented, the observer was given as much time as she needed to make a setting. When she signaled satisfaction with her setting, the tone was turned off and the tape removed. For the next trial, a new length was presented and the tone was turned on at either a different arbitrarily-selected starting value or at the level of the previous response, depending on the condition.

There were two groups of observers, with 14 in the Changed group and 16 in the Unchanged group. In the Unchanged group, the initial level of the tone was the same as that selected by the observer on the previous trial; thus, when a trial began, the observer heard the tone at the level she had previously set, and then made her adjustments from that level. In the Changed group, the experimenter changed the intensity of the tone between trials, sometimes increasing and sometimes decreasing it by a variable amount; thus, when a trial began, the observer heard the tone at a different initial level, and then made her adjustments from that level.

The stimulus order was arranged so that each of the six lengths preceded every other length, including itself, just once (see Cross, 1973).[2] As a result, each observer made 37 judgments. A different stimulus order was used for each observer within a group; the same set of stimulus orders was used for each group.

Results

Figure 1 shows the mean produced sound pressure level (in dB) as a function of the presented tape length (on a log scale) for both conditions. The trends are clearly linear, with the data for the Unchanged condition showing a lower slope than the data for the Changed condition. The figure also shows lines that represent the fits of a power function (i.e., Stevens's model) to the data from each condition; the slope for the Changed condition is 0.94 with a 95% CI of (0.80, 1.09), and for the Unchanged condition the slope is 0.78 with a 95% CI of (0.67, 0.89); the overlap of the confidence intervals indicates no significant difference between the slopes at the .05 level.

Figure 2 presents the data (means across observers) in terms of the plots suggested by Cross (1973). The lines in the left panel (Changed condition) have a slope of 0.01 and those in the right panel (Unchanged condition) have a slope of 0.04, using the mean estimates of γ obtained by fitting Eq. 1 (see below). The finding of a small positive slope in the right panel is in agreement with Cross's finding of assimilation for loudness.

A more complete picture of the influence of previous events on the current response can be obtained by fitting the DeCarlo–Cross model described in the Introduction to the data from each condition. As noted in the Introduction, the Relative Judgment Model has parameters that serve as measures of a stimulus context effect and of a judgmental effect (see Eq. 1). Table 1 shows the mean parameter values for Changed and Unchanged conditions. For β, the exponent, the Changed group has a value of 0.94; the Unchanged, 0.77. For γ, a measure of the stimulus context effect, the Changed group has a value of 0.01 ($t(13) = 0.75$, $P = 0.46$, all tests are two-tailed); the Unchanged has a value of 0.04 ($t(15) = 2.74$, $P = 0.02$). For λ, a measure of the judgmental effect, the Changed group has a value of 0.12 ($t(13) = 1.95$, $P = 0.07$); the Unchanged has a value of 0.48 ($t(15) = 8.51$, $P < 0.01$).

Table 2 shows the parameter estimates and the coefficient of determination (R^2) for each observer in each condition. In all cases, the values of R^2 are generally above 0.80 and many are above 0.90. The estimates of γ are small, with two significant values in the Changed group (0.10 and –0.10) and four significant values in the Unchanged group (0.08, 0.16, 0.10, 0.10). In fact, for many subjects in the Changed condition, both sequential effects, as measured by γ and λ, are apparently zero, in which case the model simplifies to Stevens's power law (Stevens, 1986).

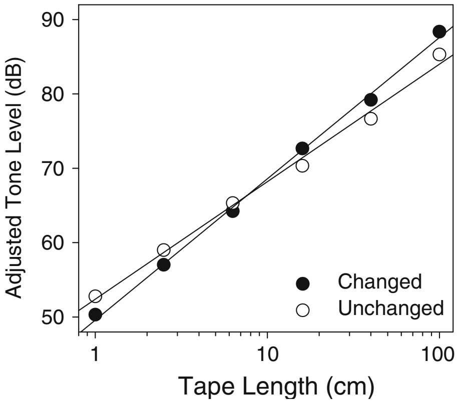

Fig. 1 Mean produced sound pressure levels (dB) plotted against the presented tape lengths (on a logarithmic scale) for the Changed (*filled circles*) and Unchanged (*unfilled circles*) conditions. Lines from a fit of a log-linearized version of Stevens's power law are also shown

Estimates of λ for individuals—which can be interpreted as a measure of the degree of trial-to-trial recalibration—differ markedly across the two conditions. For the Changed group, λ ranges from negative to positive, but most values are small and not significant: Their mean is 0.12. In contrast, for the Unchanged group, λ is always positive and most values are large and significant: Their mean is 0.48. Thus, the way the initial value of the tone was presented had a large effect on the autocorrelated error term. Note that the finding of small autocorrelation in the Changed condition is consistent with DeCarlo's (2003) results for MP; he also found relatively small autocorrelation when using a random initial value for the adjusted tone.

Discussion

The Relative Judgment Model and its associated regression analysis provide three values to be considered—β, the exponent; γ, the stimulus context effect; and λ, the autocorrelated error.

The exponent

The present CMM experiment yielded an exponent of 0.77 for the Unchanged group and 0.94 for the Changed group (which are not significantly different from

---

[2] Because of an error made during the experiment, the sequence was slightly unbalanced for Subject 13.

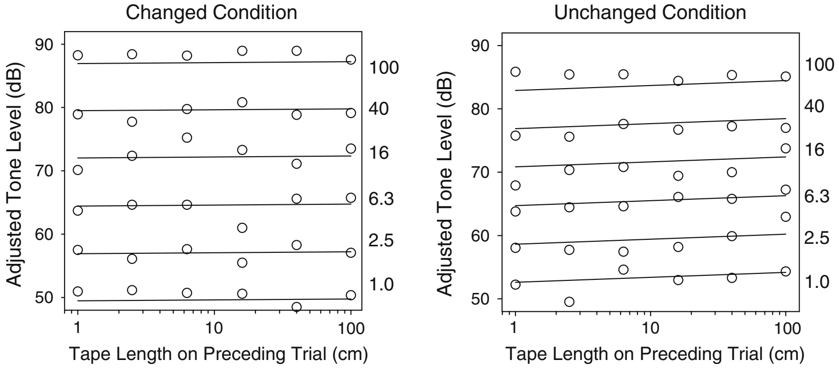

Fig. 2 Mean produced sound pressure level (dB) plotted against the tape length presented on the preceding trial (on a logarithmic scale), with the length of the current trial (in cm) listed to the right of each panel. The *left and right panels* show the data for the Changed and Unchanged conditions, respectively

Table 1 Mean parameter estimates for the DeCarlo–Cross Model (Eq. 1)

	Changed condition			Unchanged condition		
	β	γ	λ	β	γ	λ
Estimate	0.94	0.01	0.12	0.77	0.04*	0.48**

*$P < 0.05$. **$P < 0.01$

each other, as noted above). In the same kind of matching task with adult observers drawn from the same pool, Teghtsoonian (1980) found an exponent of 0.70, close to the range of values found here.

In Teghtsoonian's (1980) study, observers matched length to loudness as well as the reverse, yielding an estimate of the regression effect: For her adult observers, when length was matched to loudness, the inverse exponent was 1.75; when loudness was matched to length, the exponent was 0.70. The assimilation effect measured by γ in the present study cannot fully account for a regression effect this large.

The stimulus context effect

The finding that mean values of γ are essentially zero in the Changed group (0.01) and small but positive for the Unchanged group (0.04) suggests a small perceptual effect in the Unchanged group. The values of γ are consistent with Cross's finding that the stimulus context effect is generally not large, typically less than 0.10.

Autocorrelated error

Starting a tone on a new trial at the level to which it was adjusted on the previous trial resulted in much larger autocorrelation than starting the tone at an arbitrary level. This result has a very simple interpretation in terms of the DeCarlo–Cross model: The Unchanged procedure increases observers' reliance on the response–stimulus pair of the previous trial as a frame of reference, whereas the Changed procedure increases their reliance on a frame of reference that is stable from trial to trial. Indeed, only two subjects in the Changed group had values of λ significantly different from zero (see Table 2), whereas 12 subjects in

Table 2 Results for each observer for each condition

Observer	Parameter estimates			
	β	γ	λ	R^2
Changed condition				
2	0.84 (0.05)	0.05 (0.05)	0.31 (0.17)	0.88
4	0.79 (0.05)	0.05 (0.05)	0.29 (0.17)	0.88
6	0.84 (0.04)	0.10* (0.04)	0.49** (0.15)	0.93
8	1.08 (0.05)	0.04 (0.05)	0.04 (0.18)	0.94
10	1.16 (0.07)	0.04 (0.07)	–0.04 (0.19)	0.91
12	0.44 (0.04)	0.05 (0.04)	–0.33 (0.19)	0.81
14	0.94 (0.07)	0.01 (0.07)	0.14 (0.18)	0.84
16	0.98 (0.07)	0.09 (0.07)	–0.02 (0.18)	0.87
18	1.05 (0.06)	–0.12 (0.06)	0.34 (0.17)	0.90
20	1.31 (0.06)	0.00 (0.06)	–0.06 (0.18)	0.95
22	1.06 (0.07)	–0.03 (0.07)	0.41* (0.16)	0.90
24	0.55 (0.06)	0.01 (0.06)	–0.02 (0.18)	0.73
26	1.32 (0.05)	–0.02 (0.05)	–0.04 (0.18)	0.96
28	0.83 (0.05)	–0.10* (0.05)	0.16 (0.18)	0.90
Unchanged condition				
1	0.40 (0.03)	0.08** (0.03)	0.41* (0.16)	0.87
3	0.68 (0.05)	0.04 (0.05)	0.10 (0.18)	0.86
5	0.79 (0.05)	–0.01 (0.05)	0.51** (0.16)	0.90
7	0.94 (0.04)	0.16** (0.04)	0.35* (0.17)	0.95
9	0.99 (0.04)	0.10** (0.04)	0.24 (0.17)	0.96
11	0.73 (0.07)	0.05 (0.06)	0.23 (0.19)	0.80
13	0.97 (0.07)	–0.01 (0.07)	0.51** (0.15)	0.89
15	0.67 (0.07)	–0.01 (0.07)	0.54** (0.15)	0.78
17	0.97 (0.04)	0.06 (0.04)	0.53** (0.15)	0.94
19	0.96 (0.06)	–0.02 (0.06)	0.14 (0.18)	0.90
21	0.43 (0.03)	0.10** (0.03)	0.58** (0.15)	0.88
23	0.74 (0.04)	–0.03 (0.04)	0.70** (0.13)	0.93
25	0.95 (0.04)	0.02 (0.04)	0.52** (0.18)	0.94
27	0.77 (0.04)	0.06 (0.04)	0.62** (0.14)	0.90
29	0.37 (0.03)	0.00 (0.03)	0.95** (0.05)	0.90
30	0.92 (0.04)	–0.00 (0.04)	0.66** (0.13)	0.95

Standard errors are in parentheses; significance tests are not shown for the estimates of β because all the estimates are significant (i.e., several times larger than their standard errors)

*$P < 0.05$, **$P < 0.01$

the Unchanged group had significant values. Thus, it appears that making the previous response available to the observer at the beginning of each trial induces a greater degree of trial-to-trial recalibration, whereas starting the trial with a different value induces the use of a stable frame of reference.

Controlling trial-to-trial recalibration

The present study adds to the evidence that procedures that increase reliance on the previous trial's stimulus–response pairing—the instructions (''make your judgment on this trial with respect to the response you made on the last trial''), or the physical availability of the previous response (providing it at the beginning of the next trial)—will increase the magnitude of the sequential effects that are indexed by the autocorrelation parameter. Conversely, procedures that encourage reliance on a stable frame of reference—the instructions (''make your response on this trial in terms of a fixed reference point''), or presenting a different intensity in CMM at the start of a new trial—will reduce or eliminate their magnitude.

To summarize, we conclude that our findings are of both theoretical and practical significance. First, we have shown that Relative Judgment Theory and the DeCarlo–Cross model of sequential effects are applicable to a CMM paradigm. Second, we have added to the evidence that the theory offers an original and useful conceptualization of the processes underlying sequential effects. Whereas earlier models focused attention on the roles of S_{t-1} and R_{t-1}, the DeCarlo–Cross model points to the role that can be played on every trial by the connection established by the observer between the response and sensation of the preceding trial. This influence is what we have termed trial-to-trial recalibration, and we have described ways in which its influence may be minimized: the results clearly show that starting each trial in CMM with a different intensity to be adjusted greatly reduces or eliminates sequential effects.

Once specified, the possibility of trial-to-trial recalibration seems obvious: There is nothing to prevent an observer on any given trial from relying on the scale implied by his previous judgment. The earliest version of ME (Stevens, 1986) employed a designated standard and an assigned modulus which were identified to the observer before the first trial—e.g., ''Here is an example of the tones you will hear; you should represent its loudness by the number 100.'' It soon became apparent that this preliminary step was not required, and indeed that it could exert a biasing effect on the results, so it was dropped in favor of allowing the observer complete freedom in selecting a judgment for the first stimulus. Oddly, no consideration was given to the possibility that the scale implied by that first judgment might not serve as a fixed reference for all observers for all subsequent judgments. We now know that some observers rely on the scale implied by their response to the previous stimulus, and that the scale may drift substantially over trials because of this influence. It is this process that the DeCarlo–Cross model identifies and measures.

In conclusion, we suggest that there are two rather different ways in which the results reported here can be interpreted. One could note that we have added further evidence showing how a response may depend on factors other than the current stimulus. But we have also shown how this particular effect can be both measured and experimentally controlled. Perhaps most importantly, we have shown that our findings can be described by a general model that includes both a systematic relation between stimulus levels and reports of their subjective magnitudes, and, in addition, a contextual effect that can influence any given judgment. Both features are present in the model, and, in our view, each reflects an underlying psychological process.

Appendix A: Appraisal of an alternative model

Jesteadt et al. (1977) extended the regression model suggested by Cross (1973) by including, in addition to S_t and S_{t-1}, the previous response, R_{t-1}, as a regressor,

$$\log R_t = \alpha_0 + \alpha_1 \log S_t + \alpha_2 \log S_{t-1} + \alpha_3 \log R_{t-1} + u_t. \tag{2}$$

A theoretical basis for this model has been presented by DeCarlo (1989/1990; also see DeCarlo & Cross, 1990). The basic idea is that the coefficient of R_{t-1} reflects a judgmental effect, such as choosing a response close to the response given on the previous trial. However, despite its plausibility, a number of difficulties emerge when the model of Jesteadt et al. is applied; DeCarlo and Cross (1990) and DeCarlo (1992, 1994) have detailed these at length, so only two examples will be noted here. For one, the coefficient of S_{t-1} for a fit of Jesteadt et al.'s model is negative for loudness estimation, which indicates contrast, whereas Cross found assimilation for loudness (i.e., a positive coefficient of S_{t-1}), as have a number of subsequent studies. Second, DeCarlo (1992) found that increasing the time between trials led to an increase in the magnitude of the coefficient of S_{t-1} for the Jesteadt et al.

Table 3 Mean parameter estimates for the model of Jesteadt et al. (1977)

	Changed condition			Unchanged condition		
	α_1	α_2	α_3	α_1	α_2	α_3
Estimate	0.94	–0.11	0.12	0.79	–0.29**	0.45**

*$P < 0.05$, **$P < 0.01$

model, a finding that is both counter-intuitive and conflicts with what can be shown with a simple Cross analysis.

When the JLG model is applied to the present data set, the parameter estimates obtained for fits of Eq. 2 to the individual data are shown in Table 3. The mean estimate of α_1 is the same as or very close to the mean estimate of β obtained for Eq. 1 and the mean estimate of α_3 is very close to the mean estimate of λ obtained for Eq. 1, which is a typical result reported in previous studies. The main difference found with Eq. 2 is that the coefficient of the lagged log stimulus (α_2) is considerably different from that found for Eq. 1 (i.e.,γ), in that the coefficient is large and negative in both conditions, and is considerably larger in the Unchanged condition (–0.29) than in the Changed condition (–0.11). This suggests a large stimulus contrast effect for length (at least in the Changed condition), which in our view does not seem likely. It has been shown, however, that in terms of the Relative Judgment Theory, the effect occurs because the parameter α_2 in Jesteadt et al.'s model confounds perceptual and judgmental effects (e.g., see DeCarlo, 1992, 1994, 2003). The results for the present study provide further support for this view. Also note that Table 1 shows that γ of Eq. 1 is at or close to zero in many cases, and so Eq. 1 offers a more parsimonious model than Eq. 2. We believe that the results obtained for the current data set with the JLG model add to the now considerable evidence that, despite its seeming plausibility, this model does not provide an acceptable description of sequential effects.

Appendix B: Relation of Eq. 1 to the Relative Judgment Model

The Relative Judgment Model weights the two frames of reference as follows:

$$R_t = (R_0/\Psi_0)^{1-\lambda}\ \Psi_t\ (R_{t-1}/\Psi_{t-1})^{\lambda}\ \nu_t, \tag{3}$$

(e.g., Eq. 1, DeCarlo, 1994; Eq. 3, DeCarlo, 2003). Thus, when $\lambda = 0$, all judgments are made with respect to a stable response–sensation pair (R_0/Ψ_0) and the model reduces to a simple model of proportional judgment; when $\lambda = 1$, all judgments are made with respect to the previous response–sensation pair (R_{t-1}/Ψ_{t-1}) and the model reduces to the response ratio hypothesis of Luce and Green (1974). Values of λ between zero and unity indicate the relative influence of the two frames of reference on judgment.

To simplify notation, let $\alpha_0 = R_0/\Psi_0$ so that Eq. 3 can be written as

$$R_t = \alpha_0^{1-\lambda}\ \Psi_t\ (R_{t-1}/\Psi_{t-1})^{\lambda}\ \nu_t. \tag{4}$$

Next, note that

$$\alpha_0^{1-\lambda} = \alpha_0(1/\alpha_0)^{\lambda}.$$

It follows that Eq. 4 can be written as

$$R_t = \alpha_0\Psi_t[R_{t-1}/(\alpha_0\Psi_{t-1})]^{\lambda}\nu_t. \tag{5}$$

Let ε_{t-1} equal the term in square brackets, that is $\varepsilon_{t-1} = R_{t-1}/(\alpha_0\Psi_{t-1})$. Eq. 5 can then be written as

$$R_t = \alpha_0\ \Psi_t\ \varepsilon_{t-1}^{\lambda}\nu_t, \tag{6}$$

which shows that the Relative Judgment Model can be written as a simple model of proportional judgment (with a nonrandom error process). Taking logarithms of Eq. 6 and using a version of the psychophysical function that allows for effects of the previous stimulus intensity, such as $\Psi_t = S_t^{\beta}\ S^{\gamma}_{t-1}$, gives Eq. 1 as given in the text, with $\alpha = \log \alpha_0$, $e_{t-1} = \log \varepsilon_{t-1}$, and $u_t = \log \upsilon_t$. This shows the relation of Eq. 1 to the Relative Judgment Model.

It might be asked why one can substitute $\varepsilon_{t-1} = R_{t-1}/(\alpha_0\Psi_{t-1})$ in Eq. 5 to arrive at Eq. 6. Note that a simple model of proportional judgment with non-random error is

$$R_t = \alpha_0\Psi_t\varepsilon_t,$$

and so

$$\varepsilon_t = R_t/(\alpha_0\Psi_t),$$

from which it follows that

$$\varepsilon_{t-1} = R_{t-1}/(\alpha_0\Psi_{t-1}),$$

which is exactly as used in Eq. 5 to arrive at Eq. 6. Thus, the Relative Judgment Model identifies a systematic component to the error term, that is, the effects are due to the relative influence of a prior frame of reference and a long-term frame. Further details, as well as notes about estimation, can be found in DeCarlo (1994, 2003).

References

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 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. (1994). A dynamic theory of proportional judgment: Context and judgment of length, heaviness, and roughness. *Journal of Experimental Psychology: Human Perception and 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., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General, 119*, 375–396.

Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology, 46*, 373–380.

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

Lockhead, G. R. (1992). Psychophysical scaling: Judgments of attributes or objects? *Behavioral and Brain Sciences, 15*, 543–601.

Luce, R. D., & Green, D. M. (1974). The response ratio hypothesis for magnitude estimation. *Journal of Mathematical Psychology, 11*, 1–14.

Schneider, B., & Parker, S. (2002). Psychophysics in the 21st century: Abandoning the quest for sensory scales. Fechner Day 2002. In: J. A. Da Silva, E. H. Matsushima, & N. P. Ribeiro Filho (Eds.), Eighteenth Annual Meeting of the International Society for Psychophysics, (pp. 33–38). Rio de Janeiro, RJ, Brazil: The International Society for Psychophysics.

Stevens, S. S. (1986). Psychophysics: Introduction to its perceptual, neural, and social prospect. New Jersey: Transaction.

Teghtsoonian, M. (1980). Children's scales of length and loudness: A developmental application of cross-modal matching. *Journal of Experimental Child Psychology, 30*, 290–307.

Teghtsoonian, M., & Teghtsoonian, R. (1983). Consistency of individual exponents in cross-modal matching. *Perception & Psychophysics, 33*, 203–214.

Teghtsoonian, R. (1971). On the exponents in Stevens's law and the constants in Ekman's law. *Psychological Review, 78*, 71–80.

Teghtsoonian, R. (1973). Range effects in psychophysical scaling and a revision of Stevens's law. *American Journal of Psychology, 86*, 3–27.

Teghtsoonian, R., & Teghtsoonian, M. (1978). Range and regression effects in magnitude scaling. *Perception & Psychophysics, 24*, 305–314.

Teghtsoonian, R., Teghtsoonian, M., & Baird, J. C. (1985). On the nature and meaning of sinuosity in magnitude estimation functions. *Psychological Research, 57*, 63–69.

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

Ward, L. M. (1975). Sequential dependencies and response range in cross-modality matches of duration to loudness. *Perception & Psychophysics, 18*, 217–223.