

Consistency of individual exponents in cross-modal matching

MARTHA TEGHTSOONIAN and ROBERT TEGHTSOONIAN
Smith College, Northampton, Massachusetts

An important question about individual differences in the exponent of the psychophysical power law is how they should be interpreted. The differences may reflect permanent characteristics of individuals, and it has been argued that, if this is so, the range of these differences is so great as to identify the class of data as exceptional among the physical and biological sciences. Cited as evidence of such permanence has been the correlation between individual exponents obtained on two separate occasions. In a previous paper, we showed that increasing the time interval between occasions reduced the correlation to a nonsignificant level; we argued, therefore, that obtained individual differences in exponents, even though large, depended upon the operation of factors only incidentally associated with the particular observer. In a series of new studies of session-to-session correlation between individual exponents, we provide evidence that: (1) our original finding for magnitude estimates of visual size is repeatable, with the correlation dropping to nearly zero after 1 week; (2) when judged line length is matched to brightness, a delay of 1 week is sufficient to produce a nonsignificant correlation; (3) in contrast, magnitude estimates of loudness yield significant correlations after a week's delay; (4) but, when moduli are arbitrarily changed between sessions by the experimenter, these correlations for magnitude estimates of loudness drop to a nonsignificant level, even for a zero-delay condition. We conclude that, whereas in some scaling tasks the passage of time alone between sessions is sufficient to disrupt what appears to be the mnemonic basis for session-to-session correlation, in other (less familiar) tasks, more positive interference (in the form of a modulus change) is needed to achieve the same end. The evidence is consistent with the belief that enduring characteristics of the observer contribute only a small portion of the variability in individual power law exponents.

The method of cross-modal matching entails an observer's adjusting the stimulus magnitude on one perceptual continuum to achieve an intensive match with the perceived magnitude of a stimulus on a second continuum (see Marks, 1974, or Stevens, 1975, for a detailed account). Thus, an observer can adjust the luminance of a spot of light so that its brightness is equal in magnitude to the loudness of a tone of some designated intensity. Or, more conveniently for the experimenter, the observer may simply provide a number match to the target tone; in this case, the technique is referred to as the method of magnitude estimation (Stevens, 1975). Regardless of the identity of the continuum controlled by the observer, the results have, with remarkable consistency, taken the same form: The magnitude of the observer's setting on continuum A

(Φ_a) is a power function of the stimulus magnitude on continuum B (Φ_b), that is, $\Phi_a = \alpha \Phi_b^\beta$. This is the psychophysical relation sometimes referred to as Stevens' law.

At least as important as the form of the law itself is the finding that for any given response continuum, each target continuum can be characterized as having a distinctive value of β . Thus, for example, loudness has an exponent of about 0.6, and perceived intensity of electric current has an exponent of about 2.2, when each is scaled by the method of magnitude estimation; similarly, more than three dozen perceptual continua have been associated with characteristic exponents. Efforts have been made to provide interpretations for these values. Stevens (1958, 1960), for example, argued that the exponent indicated whether the relevant receptor system operated as a "compressor" or an "expander," and even that the behaviorally based exponent could be duplicated at the neurological level (Stevens, 1970). R. Teghtsoonian (1971) suggested that exponents encode information about the size of the dynamic range, and possibly about the resolving power as well, of the associated receptor systems.

But nearly all of the research literature on which

The research reported here was supported in part by Grant HD10857 and by Grant NS12597. We are indebted to Susan Crow, Barbara Czescik, Miriam Graham, and Ann Peterson, who assisted us in running subjects and in analyzing data. We are also indebted to John C. Baird, R. Duncan Luce, and Stanley Rule, who read a version of the manuscript and whose comments have helped to improve it. Reprint requests should be sent to R. Teghtsoonian, Clark Science Center, Smith College, Northampton, Massachusetts 01063.

such interpretive systems are based involves judgments averaged over groups of individual observers. Important questions have been raised about the degree to which such average values are representative of the behavior of individual observers, and it is those questions that are the subject of this paper.

When a group of subjects provides matches for the same set of stimulus values, it has been found that the individual exponents differ significantly from each other—that is, that the goodness of fit of the individual power functions is substantial and permits us to reject the null hypothesis that individual differences are within measurement error (e.g., Mitchell & Gregson, 1971; M. Teghtsoonian & R. Teghtsoonian, 1971; Wanschura & Dawson, 1974). But, if individual exponents do differ significantly from each other, and if these differences are large relative to the differences that distinguish continua, the interpretation of exponents based on pooled data is not easy to make and the idea that the group exponent is representative—that is, that it reflects characteristics of the individual transduction process—is put into doubt, as Curtis, Attneave, and Harrington (1968) and Mitchell and Gregson (1971) pointed out. In fact, Luce (1972) suggested that one fundamental difference between physical and psychophysical measures is that the latter do not exhibit a fixed relation to one another over individuals, that is, that exponents are not independent of the subject.

However, it does not necessarily follow from the fact that individual exponents differ from each other that these differences are attributable to relatively enduring characteristics, whether perceptual or response-based, of the individual subject. There is some evidence to support the conclusion that they are: A number of experiments (Ekman, Hosman, Lindman, Ljungberg, & Akesson, 1968; Jones & Woskow, 1962; Luce & Mo, 1965; Mitchell & Gregson, 1971) have shown that the performance characteristics of an individual on one occasion are good predictors of his performance on a subsequent occasion. All these studies, however, share at least one of two procedural aspects—many judgments of each stimulus, and judgments made close together in time. These conditions favor intraindividual consistency on a mnemonic basis, quite aside from perceptual factors. Repeated presentation of stimuli and repeated responding may permit the subject who strives for consistency to differentiate stimuli and responses and to remember and repeat the response already made to a stimulus. In an experiment previously reported (M. Teghtsoonian & R. Teghtsoonian, 1971), we attempted to separate mnemonic from perceptual factors by directly manipulating a variable that might be expected to influence memory for responses, the interval between

the first set of judgments of a stimulus series and the second set. If the consistency of individual exponents reflects characteristic individual differences, then the passage of time should exert little influence. In fact, number matches to the apparent visual area of circles in two sessions showed a high ($r = +0.8$) and reliable correlation between individual exponents in Sessions 1 and 2 with no time interval between the two sessions, but the correlation dropped to a nonsignificant value ($r = +0.4$) after 1 h and was at zero after 11 weeks. We concluded that, although there are real individual differences in exponents, they are produced by factors that are stable only for short time periods, and thus the differences do not represent enduring individual characteristics. That report has been criticized because of the small *N*s (9-12 per group) and because of the absence of data on variability between and within subjects. Subsequently, also, Englund and Dawson (1974) failed to confirm these results in a study in which subjects matched number to visual area of circles in two sessions, separated either by no delay or by 1 week; in their study, all session-to-session correlations (τ) were significant. In a further experiment, duration was matched to visual area, and a significant session-to-session correlation prevailed after a week's interval.

Experiment 1 was undertaken as an extension of our original study: Groups of subjects made number matches to straight lines varying in extent, with session-to-session intervals ranging from a few minutes to 1 week. Experiment 2 was a replication of Experiment 1.

EXPERIMENTS 1 AND 2

Method

Stimulus materials. The stimuli were straight lines, drawn in black ink on white paper and ranging in length from 0.5 to 29.7 cm. They were presented in Multilithed booklets, one line to a 22 × 35 cm page, oriented vertically and centered. Each booklet contained the nine lengths, arranged in a different random order for each subject within a delay group and for Sessions 1 and 2. The random orders were repeated across groups. Instructions for Session 1 (or Session 2) constituted the cover page.

Subjects. The subjects were students in sections, varying somewhat in size, of an introductory psychology class at a women's college. We aimed for *N* = 25 in each delay group in Experiment 1; final *N*s were 29, 20, 24, and 23, for zero, 1-h, 1-day, and 1-week delays, respectively. We aimed for *N* = 50 in Experiment 2; final *N*s were 51, 33, 48, and 61, for zero, 40-min, 1-day, and 1-week delays, respectively.

Procedure. All subjects assigned numbers to the nine lengths on two occasions. There were four delay groups, with 3-5 min, 40 min (Experiment 2) or 1 h (Experiment 1), 1 day, or 1 week between sessions. The subjects were run in groups, during regularly scheduled class meetings. They were not forewarned of the second sessions.

For Session 1, the subjects read the cover page of instruc-

tions. These specified the task as assigning numbers to represent how long the lines looked, told the subject she could use any numbers she wanted (very large numbers as well as very small numbers, including fractions less than 1) except for zero, and cautioned her not to use a ruler or otherwise attempt to measure the lines, since we were interested in her subjective impressions. She was to write the number chosen at the bottom of each page. The subject worked at her own pace. For Session 2, the procedure was the same, except that the instructions said, "On a previous occasion, you have assigned numbers to represent how long some lines look to you. We would like you to continue the same task with some more lines." The subject was reminded that she could choose any number that seemed appropriate to reflect her impression of length.

Results and Discussion

For each subject for Session 1 and Session 2, a least squares fit of log judgments as a function of log length was obtained; the slope of this function is the exponent of an individual power function. The exponents from Experiments 1 and 2 were combined for correlational analysis.¹ The mean exponent for Session 1 was .98; standard deviations at different delays ranged from .12 to .15. The mean exponent for Session 2 was .98; standard deviations at different delays ranged from .11 to .14.

The exponents for Session 1 were correlated with those for Session 2. The resulting values of *r*, together with the number of subjects, at each delay are shown in Table 1. For delays of up to 1 day, correlations are low to moderate, positive, and significant. After 1 week, the correlation is no longer significant; at that delay, individuals are interchangeable with respect to exponents.

But, it could be argued that nonsignificant session-to-session correlations are simply artifacts of the error of measurement associated with individual exponents. If these values are large, that is, comparable in size to the variability in exponents across individuals, then estimated session-to-session correlations are bound to be low even if the "true" correlation were 1.0. To check this possibility, we calculated the standard error of each individual exponent and the means of these values for the various delay conditions at each session, and the corresponding standard deviations of individual exponents. The results are shown in Table 2. It is clear that the standard errors of the individual exponents are substantially smaller than the variability over subjects (on the order of 1/3) and therefore could not account for low, nonsignificant correlations.

Table 1
Magnitude Estimation of Length: Correlation Coefficients (*r*)
Between Session 1 and Session 2 Exponents

| | Delay | | | |
|----------|-------|-------|-------|--------|
| | Zero | 1 h | 1 Day | 1 Week |
| <i>r</i> | +.36* | +.55* | +.35* | +.09 |
| N | 79 | 52 | 70 | 83 |

**p* < .05.

Table 2
Mean Standard Error and Standard Deviation of Individual
Exponents for Two Sessions at Four Delays

| | Delay | | | | | | | |
|-----------|-------|-----|-----|-----|-------|-----|--------|-----|
| | Zero | | 1 h | | 1 Day | | 1 Week | |
| | SE | SD | SE | SD | SE | SD | SE | SD |
| Session 1 | .04 | .12 | .05 | .15 | .04 | .12 | .05 | .15 |
| Session 2 | .05 | .13 | .04 | .12 | .05 | .14 | .04 | .11 |

Note—The standard error of an individual exponent β_j is $[\beta_j \sqrt{(1/r_j^2) - 1}] / [1/\sqrt{(N - 2)}]$.

Another approach to the problem is to ask whether a nonsignificant test-retest reliability coefficient—that is, a session-to-session correlation—is produced by a low split-half reliability coefficient—that is, an unreliable estimate of the individual exponent. For Experiment 1 only, split-half reliability coefficients were determined by calculating for each subject two exponents, one based on stimulus values 1, 3, 6, and 8, and the other for values 2, 4, 7, and 9, and correlating them. These coefficients (corrected for number of stimulus values) ranged (over the two sessions and various delays) from +.46 to +.92, with a median of +.75. In comparison, the highest test-retest coefficient was +.56.

It seems clear that we have sufficiently reliable estimates of individual exponents, but that at any single session there are clear differences among individuals. The important result is that these differences do not entail a consistent ordering of subjects when the test-retest delay is relatively long. Whatever the factors may be that account for intersubject differences, those differences are not very durable over time, an attribute that seems incompatible with the usual view of individual differences.

The outcomes of Experiments 1 and 2 confirm the conclusion of the previous experiment (M. Teghtsoonian & R. Teghtsoonian, 1971), in which subjects assigned numbers to visual area, and gave us some confidence that the low session-to-session correlations in that experiment were not an artifact associated with small Ns and high intrasubject variability. But it seemed desirable to repeat the basic design with other target modalities, especially ones that are more purely "sensory" than are visual length and area. Experiment 3 used a repeated judging design in which delay between first and second sessions varied from zero to 1 week. The subjects were asked to match numbers to a 3000-Hz tone varying in intensity, and to match line length to luminosity.

EXPERIMENT 3

Method

Apparatus. The sound signal, a 3000-Hz tone produced by a Hewlett-Packard 241A oscillator, was attenuated to produce

variations from 68 to 104 dB SPL in 6-dB steps, and presented to the subject binaurally through PDR-8 earphones.

The light signal, a luminous disk 14.7 cm in diameter, was projected on a white wall 114 cm from the subject. Room lights were dimmed. A Carousel projector was used with neutral density filters to produce a variation in range of two log units of luminosity. Stimulus calibration with an SEI exposure photometer showed luminosity values ranging from 0.69 to 69 cd/m².

Directly in front of the subject during the matching of line length to luminosity was a band of white paper (adding machine tape, 5.7 cm wide) running the length of a 180-cm table. With a black Pentel, she drew a matching line on this paper, which was then unrolled to conceal the previous match and to expose a fresh surface for the next one.

Subjects. The subjects were volunteers from an introductory psychology course at a women's college; they received course credit for their participation. There were four groups. Originally, N=22 for each group. Any subject for whom a power function was a poor fit to the individual judgments ($r^2 < .70$) was replaced. Eventually, it was necessary to run 33, 42, 41, and 40 subjects, respectively, for the four groups to obtain a final N=30 for each group for the SPL part of the study, and a final N=25 for each group for the luminosity part.

Procedure. There were two tasks: In one (number matched to SPL), the subjects matched numbers to a 3000-Hz tone at seven sound pressure levels varying over a range of about 100 to 1 (36 dB); in the second (length matched to luminosity), they matched line lengths to a projected disk of light at seven luminosity levels varying over a range of 100 to 1. Task was a within-subject variable: Every subject performed both tasks twice, once at Session 1 and once at Session 2. There were four delays between sessions—zero, 35 min, 24 h, and 7 days; delay was a between-subjects variable. Table 3 shows the experimental design schematically. Thus Group 1 was a single group of subjects who provided data for zero delay for length-luminosity and for 35-min delay for number-SPL. Group 2 provided data for 35-min delay for length-luminosity and zero delay for number-SPL. Group 3 provided data for 1-day delay for both tasks; Group 4 provided data for 1-week delay for both.

Detailed instructions for both tasks were administered in Session 1. The instructions for length-luminosity are given below; those for the other task were the same, except that the subject was asked to match number to loudness:

In the first part of this experiment, I'm going to show you some circles of light, by projecting them on the wall in front of the table, and I want you to tell me how bright they look to you. You'll do this by drawing lines to match the brightness of the circles. For the first circle I show you, draw a line whose length seems appropriate to the brightness of the circle. For the next circle, draw a line to match *its* brightness—longer than the line for the first circle if it seems brighter, shorter if it seems dimmer. And however much brighter or dimmer the circle looks, make the line you draw that much longer or shorter. You may make the lines as long or as short

as you wish, so long as they are proportional to the brightness as it seems to you You may use as much space as you want for the line, because there is more paper in the box, and we will pull out more as you need it

At Session 2, the instructions were:

[Last week] [Yesterday] you looked at circles of light and drew lines whose lengths matched the brightnesses. [Today] I'll show you some more circles, and I'd like you to continue drawing lines so the lengths are proportional to the brightness of the circles

The material in square brackets was omitted for Groups 1 and 2. If the subject asked whether she should try to use the same lengths as last time, she was told she could make the lines any lengths she wished, as long as the lengths were proportional to the brightnesses.

For both tasks, in each session, each value on the target continuum was presented once, in a different random order for each subject within a group; across groups, the same set of random orders was used. Each value had an equal probability of occurring first, and thus of being treated by the subject as a "standard." A target value was presented once for 2 sec; after a 30-sec interval, the next value was presented. Each session contained a 10-min break, during which the subject left the experimental room. For Groups 3 and 4, she was reminded to return 1 day (or 1 week) later; if she asked what would happen on that occasion, she was told that she would perform similar tasks.

Results

Least squares fits to individual log judgments as a function of log stimulus intensity yielded exponents, scale factors, and r^2 s for individual power functions.

Number-SPL. The four delay groups had the same exponent relating number to SPL, for both sessions (one-way analyses of variance yield $F < 1$ in each case). The mean exponent, pooling over delay groups, is 0.38 for Session 1 and 0.44 for Session 2; the Session 1 exponent is smaller [$t(120) = 4.65, p < .01$] The value of 0.4 is one we have obtained previously (R. Teghtsoonian & M. Teghtsoonian, 1978) for naive subjects making magnitude estimations of a two-log-unit range of SPL. The mean r^2 is .89 for Session 1, .92 for Session 2. The standard deviation of the distribution of the exponents, pooled over delay groups, is 0.16 for both sessions; this should be compared with the means of the standard errors of individual exponents, pooled over delay groups—0.06 for both sessions.

Length-luminosity. The four delay groups had the same exponent relating length to luminosity, for both sessions (one-way analyses of variance yield $F < 1$ in each case). The mean exponent, pooling over delay groups, is 0.26 for Session 1 and 0.25 for Session 2; the small difference is not reliable [$t(98) = 1.90, p > .05$]. The mean r^2 is .84 for Session 1, .85 for Session 2. The standard deviation of the distribution of exponents, pooled over delay groups, is about 0.09 for both sessions; the means

Table 3
Design of Experiment 3

| Group | Session 1 | Intersession Interval | Session 2 |
|-------|-----------|-----------------------|-----------|
| 1 | SPL*Lum | None | Lum*SPL |
| 2 | Lum*SPL | None | SPL*Lum |
| 3 | Lum*SPL | 1 Day | SPL*Lum |
| 4 | Lum*SPL | 1 Week | SPL*Lum |

Note—SPL denotes number matched to sound pressure level; Lum denotes line length matched to luminance; * indicates a 10-min break.

of the standard errors of individual exponents, pooled over delay groups, are 0.04 and 0.05 for Sessions 1 and 2, respectively.

Tasks compared. A comparison of the two tasks shows (1) a higher exponent for number matched to SPL than for length matched to luminosity—the former is about 1.6 times the latter. On the assumption that length and number are related by an exponent of unity (supported by Stevens & Guirao, 1963, and M. Teghtsoonian, 1965), this difference is in the direction predicted from the values tabled by Stevens (1975); (2) a lower standard deviation of the distribution of exponents for length-luminosity than for number-SPL—the former is about half the latter; (3) lower values of r^2 for individual exponents for length-luminosity.²

The correlation between Session 1 exponents for number matched to SPL and line length matched to luminosity, +.18, is not significant ($df=74$, $p > .05$).

Session-to-session correlations. The session-to-session correlations r_{12} between individual exponents are shown in Figures 1a and 1b. For number matched to SPL (Figure 1a), the correlation is significant at all delays ($\alpha=0.05$); the median corre-

lation is +.65. For line length matched to luminosity (Figure 1b), the pattern is irregular, with r^2 rising from a nonsignificant value at no delay to significant levels at 35-min and 1-day delays and falling to a nonsignificant value at a 1-week delay ($\alpha=0.05$).

These correlations include all subjects for whom individual r^2 s were equal to or greater than .70 for both sessions. If we raise the criterion r^2 to .80, mean r^2 becomes .91 (Session 1) and .93 (Session 2) for SPL exponents, and .88 (Sessions 1 and 2) for luminosity exponents. For the resulting smaller groups, Figure 1a shows significant correlations ($N_s=28, 27, 29$, and 25 in Groups 1-4, respectively) between SPL exponents. Figure 1b shows the correlations for luminosity exponents, with significant values for delays up to 1 day and a nonsignificant value for the 1-week delay ($N_s=15, 13, 18$, and 11, for Groups 1-4, respectively). The standard deviations for these smaller distributions of exponents are around 0.16 for SPL and around 0.07 for luminosity. The mean standard errors of individual exponents are around 0.05 for SPL and around 0.04 for luminosity.

Discussion

The design of this study required matching number to sound pressure level to follow matching length to luminosity for Groups 1, 3, and 4 but to precede it for Group 2. That all groups are equivalent in Session 1 number-SPL exponent and length-luminosity exponent indicates that the order of the tasks did not influence matching performance. The independence of SPL and luminosity exponents for Session 1 is also demonstrated by the nonsignificant correlation between them. Thus, the two parts of the design may be treated as separate studies, each varying delay for one of the matching tasks.

A major finding is the failure of session-to-session correlation to decrease at all with delay for matching number to sound intensity. If this result is considered together with that of Logue (1976), who also found no decrease with delay, that correlation is between +.6 and +.7 for delays up to 11 weeks, indicating that one-third to one-half of the variance in number-SPL exponents is attributable to individuals.³

The pattern of correlations for length-luminosity matching is not as clear. One reasonable interpretation of the data shown in Figure 1b is that there is a moderate correlation, around +.6, for the three shorter delays, but no significant correlation for the 1-week delay. Judgments of line length showed the same pattern in Experiments 1 and 2. The individual data exhibit more variability than those for number-SPL, and the failure of the zero-delay condition to exhibit a significant correlation is prob-

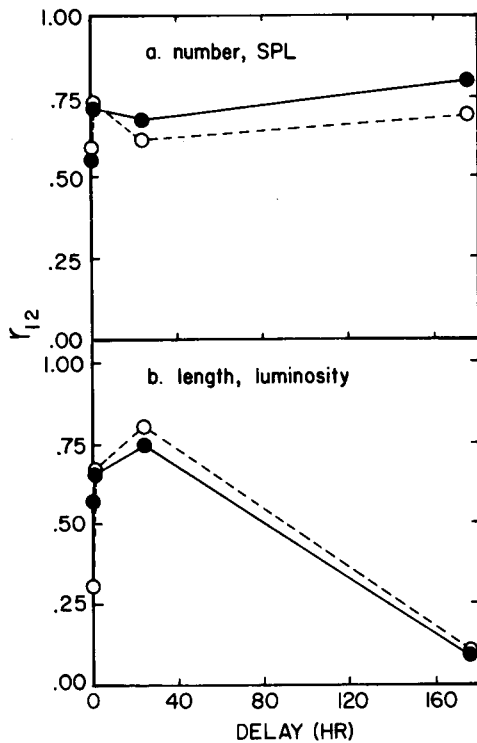


Figure 1. (a) Session 1-Session 2 correlation between exponents for number matched to SPL at four delays. Filled circles—subjects for whom $r^2 > .80$; open circles— $r^2 > .70$. (b) Session 1-Session 2 correlation between exponents for length matched to luminosity at four delays. Filled circles—subjects for whom $r^2 > .80$; open circles— $r^2 > .70$.

ably related to this higher variability, a possibility that receives support from the finding that the correlation is significant for a smaller group of subjects who met a more rigorous criterion for goodness of the power function fit.

It should be noted that whatever factors contribute to the reliable session-to-session correlations for a single task do not control them task-to-task. The correlation between number-SPL and length-luminosity exponents is not reliable. The hypothesis that individuals have characteristic ranges independent of the matching continuum is not supported by these data; the subjects did not consistently "constrict" or "expand" the ranges of number and length.

Some tentative conclusions can be drawn from the outcome of Experiment 3, taken together with the other work reported and reviewed here. In some procedures—number matched to visual area and to visual length, line length matched to luminosity—a 1-week delay between two judgment sessions is sufficient to reduce the correlation between exponents to a value close to zero. But in other procedures such a reduction does not occur: We and Logue have failed to find it for number matched to sound pressure level; Dawson and his colleagues have failed to find it for duration matched to area and for sound pressure level and duration matched to each other. The possibility arises that the durability of the individual exponent over time varies with the target continuum, with the matching continuum, or with the combination. The studies available at present have not varied the choice of continua systematically, and no conclusions can be drawn for the moment.

A question that should be raised is how much forgetting we can expect to occur as delays are increased up to 1 week. Underwood (1957) pointed out that, when retention of a deliberately learned list is tested after 24 h, rather little forgetting is measured *if* the subjects have not previously learned other lists. Underwood did not have data for forgetting as a function of time with partial learning or after minimal proactive interference, but he expected that the fall in the retention curve over time would be relatively small. We may, then, have been naive in expecting that delay alone would produce substantial forgetting in the repeated judging task unless there existed a source of substantial proactive inhibition. The average subject has little, if any, experience in matching numbers to sound pressure levels, whereas most have had considerable experience in matching numbers to lengths. We may speculate that proactive interference functions in the latter case to produce greater forgetting.

In the next studies, rather than relying on whatever effects might occur with the passage of time,

we sought to interfere actively with the subject's tendency (if she does have one) to use remembered numbers in the second series of judgments.

EXPERIMENTS 4 AND 5

Our tactic for preventing the subject from using the same responses in the two series of judgments was to assign a modulus for the first series and a different modulus for the second.⁴ Thus, whereas the subject could produce the same exponent in both series by repeating ratios of judgments, she could not do so by repeating the numbers themselves. If high session-to-session correlations depend on repeated judgments, then this procedure should reduce them to zero. If, on the other hand, such high correlations reflect enduring individual characteristics, this procedure should produce little change in the value $r = +.6$. Two experiments were carried out, one a replication of the other.

Method

Subjects. The subjects were recruited from an introductory psychology course and received credit for their participation. In Experiment 4, there were 43 subjects; of these, 11 were discarded (see Results for criterion), leaving a final $N = 32$. In Experiment 5, there were 47 subjects, of whom 12 were discarded, for a final $N = 35$.

Apparatus. The apparatus for producing a 3000-Hz tone of variable intensity was the same as that described for Experiment 3.

Procedure. The subjects were run singly. Each twice judged a series of eight intensities, covering a 32-dB range in 4-dB steps, with a different random order of presentation each time. Instructions before Session 1 were as given for number-SPL matching in Experiment 3, except that the following sentence was substituted for "for the first tone, choose a number that seems appropriate to its loudness": "The loudness of the first tone you hear will be called [assigned modulus 1]. That is, however, loud the first tone sounds to you, call that loudness [assigned modulus 1]." Instructions for Session 2, which came directly after Session 1 with an intersession interval of zero, were: "Now you'll hear some more tones and I'd like you to continue assigning numbers proportional to the loudness. This time, though, the loudness of the first tone you hear will be called [assigned modulus 2]; you should assign numbers to the other tones according to their loudness relative to the loudness of the tone called [assigned modulus 2]."

The standard stimulus was selected at random from the nine stimulus values, with the restriction that each value was the standard for Session 1 and the standard for Session 2 once for every nine subjects. The standard was presented twice, for 2 sec each time. In Experiment 4, the subject then judged the nine stimuli in random order; in Experiment 5, she judged the other eight stimuli in random order. Each stimulus was presented once for 2 sec.

The assigned moduli were selected from numbers actually given to these particular stimulus values by subjects in matching numbers to SPL (in an unpublished experiment) under free-modulus instructions. The selection followed four rules: (1) A different modulus was used in each session; (2) different modulus-standard pairings were used for each subject; (3) the modulus-standard pairing for Session 2 was done so that the subject would have to use absolute numbers different from those she had used in Session 1; (4) for half the subjects, the Session 1 modulus was larger than the Session 2 modulus.

Results and Discussion

Power functions were fitted to each subject's judgments for each session; the assigned modulus was treated as a judgment in Experiment 5. Any subject for whom r^2 was less than .80 was discarded and replaced. Table 4 shows means, standard deviations, and standard errors for individual exponents, and means of associated r^2 s. In comparison, these values from the no-delay group in Experiment 3, who matched number to sound pressure level with a free modulus, are exponents 0.36 and 0.41, SDs = 0.13 and 0.14, and r^2 s = .90 and .90 for Sessions 1 and 2, respectively. We find that the assigned modulus apparently does not disrupt matching behavior, since power functions fit individual judgments as well in Experiments 4 and 5 as in Experiment 3. When the modulus-change technique is used, the exponents are somewhat closer to the value of 0.67 for loudness suggested by Stevens (1972). Standard deviations are larger, an outcome discussed below.

Session-to-session correlation between exponents. For Experiment 4, the correlation r between Session 1 and Session 2 exponents is +.02, and for Experiment 5, it is +.27; neither value is significant ($\alpha = .05$). (For the experiments combined, $r = +.11$, $p > .05$.) When the subject is prevented from using the same numbers (though not from using the same or other appropriate ratios), the intersession correlation between number-SPL exponents is not reliably different from zero, even when there is no delay. This finding is consistent with the hypothesis that previously obtained correlations of +.6 after delays up to 11 weeks depend on subjects' using the same or similar numbers. Given that hypothesis, we speculate that the standard procedure exemplified in Experiment 3 does not completely interfere with retention when those numbers are matched to sound pressure levels, a unique pairing for most subjects.

An alternative interpretation. Although reduced correlation coefficients are usually interpreted as

evidence of a reduction in the degree of association, an alternative interpretation is possible. Since the correlation coefficient is stable as the ratio of covariance to the product of the two standard deviations, any experimental manipulation that increases these standard deviations (between-subjects variability of exponents in the present case) without a corresponding increase in covariance will lead to a reduced correlation. In short, it is possible that our modulus assignment procedure constitutes just such a manipulation, and that the degree of association between Session 1 and Session 2 exponents is unaffected (as indicated by their covariance), but that the reduced correlation is entirely attributable to the resulting increase in variability. (It should be noted in passing that for such an outcome to occur, the variability-enhancing operation must be unsystematic in its influence. If, for example, modulus assignment increased exponents by a multiplicative transformation, then the covariance as well as the univariate variances would increase, and the correlation coefficient would remain unchanged.)

Fortunately this alternative interpretation is subject to evaluation. First, a comparison was made between the variances in exponents for Experiment 3, on the one hand, and for Experiments 4 and 5, on the other. The data from Experiment 3 for subjects matching number to sound pressure level at all delays represent performance on a free-modulus, no-standard task, and can serve as a comparison for data from the assigned-modulus task. The variance in exponents in Experiment 3, Session 1, for the 107 subjects with r^2 greater than or equal to .80 is .028, whereas for the 67 comparable subjects in Experiments 4 and 5, it is .058; the difference is reliable [$F(66,106) = 2.04$, $p < .05$]. For Session 2, the comparable variances are .023 and .082; the difference is again reliable [$F(66,106) = 3.57$, $p < .05$]. However, we can also compare covariances: The covariance in Experiment 3 is .018, and in Experiments 4 and 5, .008, and the difference is reliable [$F(66,106) = 2.29$, $p < .05$]. Thus, the assigned-modulus technique both increases between-subject variance and decreases covariance.

We may, with some caution, even go a step further and ask what correlation coefficient would have resulted in Experiments 4 and 5, given the obtained covariance noted above, *if* the Session 1 and Session 2 variances remained unchanged from the results obtained in Experiment 3. To estimate this value, the covariance for Experiments 4 and 5 is divided by the product of the standard deviations for Sessions 1 and 2 of Experiment 3. The result, 0.32, evaluated for the sample size of 67 noted above for the pooled results of Experiments 4 and 5, is significant at the .01 level. While the results of such a post hoc analysis must be viewed with some caution, they do suggest that even when steps are taken

Table 4
Magnitude Estimation of Loudness With Modulus Change: Means and Standard Deviations of the Distributions of Individual Exponents, and Means of the Distributions of Standard Errors and r^2 s Associated With Those Exponents

| | Mean | SD | Mean SE | Mean r^2 |
|--------------|------|-----|---------|------------|
| Experiment 4 | | | | |
| Session 1 | .62 | .27 | .07 | .90 |
| Session 2 | .53 | .25 | .06 | .91 |
| Experiment 5 | | | | |
| Session 1 | .49 | .19 | .07 | .89 |
| Session 2 | .56 | .32 | .07 | .90 |

Note—The standard error of an individual exponent β_j is $[\beta_j \sqrt{(1/r_j^2) - 1}] / \sqrt{(N - 2)}$.

to insure that second-session judgments are truly independent of first-session judgments, there remains a reliable effect due to individual differences. What is notable, however, is that the magnitude of this effect is small, accounting for less than 10% of the total variance, and requires an unusually large sample size to achieve statistical significance. This very modest effect provides little support for the argument that psychophysical power law exponents, because of the unusually large individual differences that characterize them, stand apart from other physical and psychological measures. It seems clear that an estimate of persisting individual differences based only on the variability observed in a single session, or on session-to-session correlation obtained without vigorous efforts to assure session-to-session independence, will be too high.

Unusual matching behavior. The matching behavior of some of our subjects deserves scrutiny. Several subjects in Experiment 5 produced an exponent near the mean value for one session and an extremely low exponent (0.05-0.10) for the other. Figure 2 shows data from one such subject. The source of the discrepancy is obvious. It is as if the subject instructed herself to assign a different number to each loudness, but to have numbers for adjacent loudness separated by the same interval. If the modulus is small, and the interval appropriately selected relative to the stimulus values (in Session 1, the modulus was 8.5, and, apparently, the interval was 1), the exponent will be appropriate. If, however, the modulus is large, and the same

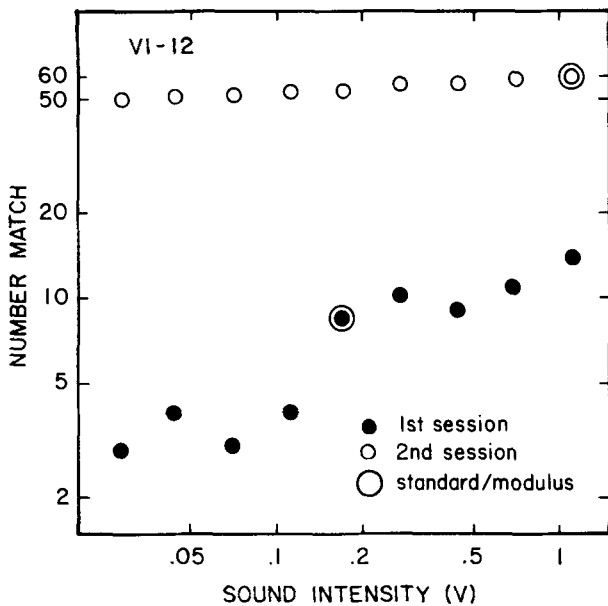


Figure 2. Number matches made by Subject VI-12 to a 3000-Hz tone varying in sound intensity, in two series of judgments with different standards and assigned moduli.

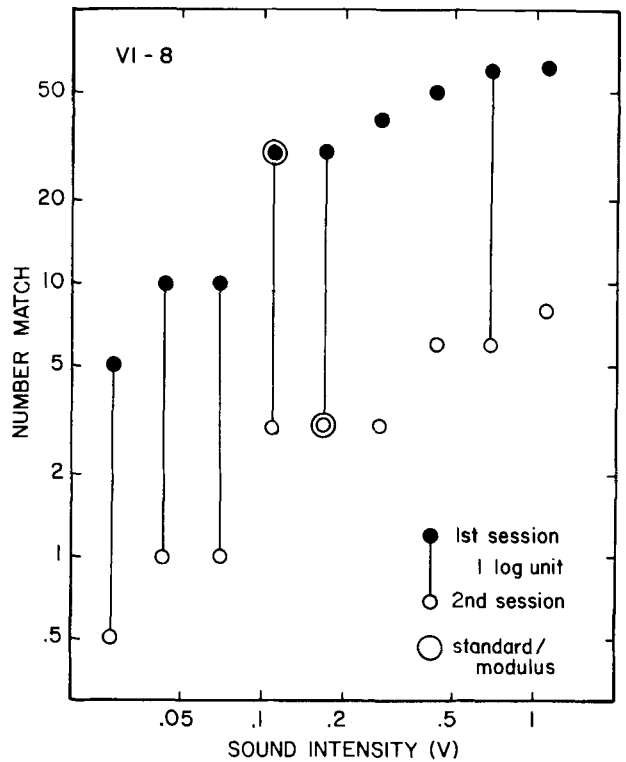


Figure 3. Number matches made by Subject VI-8 to a 3000-Hz tone varying in sound intensity, in two series of judgments with different standards and assigned moduli.

interval is maintained, the exponent will be very small, as in Session 2. As long as the subject can discriminate the stimulus levels, her power function will have a high r^2 , but as the modulus increases in size, the exponent will become smaller and smaller. Altogether, five subjects with similar matching behavior were identified.

Many experimenters who use number matching suspect that some subjects engage in category judging, in spite of instructions to make the ratio of numbers correspond to the ratio of sensations. The modulus-change procedure may be a way to discover such subjects, and provides an objective criterion for discarding their data. In the present study, the mean exponents increase somewhat, to 0.51 and 0.63, for Sessions 1 and 2, respectively, without the five subjects described above.

Another group of subjects can be identified who used particular numbers a_1, \dots, a_n for stimulus values 1 to n in Session 1, and used numbers $10a_1, \dots, 10a_n$ (or $100a_1, \dots, 100a_n$) in Session 2. Figure 3 shows the most striking example of this type of performance: The open circles show the numbers assigned in Session 1; the closed circles show the numbers assigned in Session 2, which are exactly 10 times greater for six of nine stimulus values. Perhaps six subjects performed in this way. The moral is that

subjects prefer to use the same numbers again, and may find a way to do so even when the experimenter attempts to forestall it.

The data from these two groups of subjects whose number matching is in some respects unlike that of the majority do not, however, change the conclusions about interindividual consistency in exponent: If all 11 subjects identified in the preceding paragraphs are excluded, the correlation between exponents for the two sessions remains no different from zero [$r(22) = +.11, p > .05$].

The results of these two experiments show that, by the fairly simple tactic of preventing the subject from repeating responses to a re-presented stimulus—while not preventing the subject from repeating a *response ratio* to a re-presented *pair of stimuli*—it is possible to reduce the correlation between individual exponents on two occasions of judging from a moderate value of $+.6$ (which may be maintained even when the intersession interval is as long as 11 weeks) to a value not different from zero even with no delay.

CONCLUSIONS

The observation in which this series of studies had its inception is that the exponent of the psychophysical power function shows wide variation across continua and, for a given continuum, among individuals. If the mean exponent characterizing a given continuum is taken as an indicator of the transducing properties of the associated sensory system, then the individual exponents, whose distribution provides an estimate of that mean, must reflect the transduction that occurs within each individual. But the variation in exponent among individuals is much greater than would be expected from what is known about individual differences in the structure of sensory systems. An alternative hypothesis to account for variation in exponent is that it is attributable to individual differences in ways of responding. Both these alternatives are based on an assumption that differences in exponent depend on enduring characteristics (whether sensory or response-based) of the individual. The assumption is based on the argument that a high test-retest reliability for scores (exponents) demonstrates that a high proportion of variance is attributable to enduring characteristics of the individual. It is the conclusion of this paper that the argument, and hence the assumption it supports, is incorrect, and that high test-retest correlations can be obtained in the absence of characteristic sensory or response features, simply because subjects remember what they have done before. A number of situational factors can determine differences in exponent on a first set of judgments, but on the

second set, in addition, the subject is influenced by how he responded on the first set.

We have provided one test of this nonindependence hypothesis by manipulating two variables—test-retest delay and modulus assignment—that can plausibly be imagined to increase the independence of one set of judgments from another, and have observed the resulting decrease in test-retest correlations. Nonetheless, it would be valuable if alternative tests could be developed. For example, we would like to measure the effect of test-retest delay on the similarity of the two sets of judgments. But it is not obvious how best to evaluate similarity. Whether repetition of single judgments, duplication of range or central tendency, or some combination of these should be used remains unclear and further investigation on this problem is in progress.

The studies reported here show that on some continua, the passage of time (and, presumably, the occurrence of forgetting) is sufficient to lessen the constraint exerted by the first set of responses on the second; on other continua, more active interference that makes the first set of responses unavailable to the subject is required to achieve that effect. For three continua, at least—sound intensity, luminance, and visual size—it is possible to obtain test-retest correlations not significantly different from zero. That fact makes it necessary to scrutinize the design of any experiment claiming a significant reliability coefficient to ascertain that memory effects could not have inflated its value.

This could occur not only when target and judgmental continua are the same in repeated testing, but also when the target continuum changes but the judgmental continuum does not (e.g., Jones & Marcus, 1961; Rule & Markley, 1971). Rule and Markley, for example, had subjects assign number and force of handgrip to circle area and to numerosness of dots. The test-retest reliability of coefficients was around $+.8$ when target continuum alone changed, but around $-.1$ to $+.3$ when the judgmental continuum alone changed. Rule and Markley concluded that interindividual variability in exponents stems primarily from idiosyncratic response biases. Certainly their results contradict the hypothesis that differences in sensory operating characteristics account for most of the variability. Their results do not, however, rule out the possibility that it is a tendency to repeat responses already made (whether numbers or handgrips), rather than characteristic response biases, that accounts for the correlations when the judgmental continuum was unchanged.

It is interesting to note that, when the judgmental continuum *is* changed, the test-retest correlation goes toward zero in the study by Rule and Markley

(1971). It is also true that when both target and judgmental continua are changed, a situation in which positive correlations should emerge if subjects have typical response biases (e.g., to constrict the range on the judgmental continuum), no such correlations are found. Thus, in Experiment 3, the correlation between exponents for number matched to sound intensity and line length matched to luminosity, with an $N=74$, was a nonsignificant $+ .2$. In a study by Wanschura and Dawson (1974), subjects matched sound intensity to duration, and duration to sound intensity; the correlation expected on the response-bias hypothesis was not found.

In short, our contention is that there is no compelling evidence that the reported variability in individual exponents in the typical cross-modal matching study is associated with comparable variability in enduring individual characteristics, whether of sensory systems or of response mechanisms, and that the apparent association depends upon the operation of memory for previously made responses.

We also issue a caveat about the use of magnitude estimation procedures (such as those reported by Luce & Mo, 1965, and more recently by Green & Luce, 1974, and Green, Luce, & Duncan, 1977) that entail many (30 to 60) repetitions of each stimulus value. The purpose of such a procedure is clear and reasonable—to improve reliability of estimated values, and to provide estimates of variability for each subject's judgments. However, from the perspective of our research, that procedure resembles our zero-delay condition, with many repetitions of the whole ensemble of stimuli rather than just one. Indeed, it is typical of the studies cited to have provided subjects with extensive prior practice, offering still more opportunity for the subject to become familiar with the stimulus set and his own set of responses. Despite the many differences between the two- and extended-repetition procedures, our results with the former suggest that caution is in order in accepting the assumption, implicit in the latter, that successive groups of judgments are independent of one another. It should be emphasized that this is quite a different matter from the question of whether the judgment of a single stimulus is independent of the immediately preceding responses or stimuli. The finding of Jesteadt, Luce, and Green (1977), for example, that "sequential effects do not extend over more than one trial" does not address the question of whether successive judgments of the same stimulus are independent. If, as our results lead us to suspect, they are not, then estimates of both the means and variances of such groups of judgments may be biased, and as a result compromise any analysis based upon them.

What, then, are the sources of variability in in-

dividual exponents? Among factors on the stimulus side are variability in target values (Cain, 1977, demonstrated that, on the continuum of odorant concentration, large differences from the nominal level of the target may occur and are not ordinarily recognized by the experimenter), and differences in stimulus context in random presentation orders (Cross, 1973, and R. Teghtsoonian & M. Teghtsoonian, 1978, have shown that the immediate context in which a stimulus occurs influences the judgment made); on the response side, a subject's choice of modulus (Engen & Ross, 1966) can influence the exponents, as may features of the response apparatus when the subject selects response values on some physical continuum. Walsh and Browman (1978) showed that, when care is taken to eliminate as many nonsensory sources of variability as possible, there is very little *intraindividual* exponent variability; unfortunately, their N of 3 does not permit us to estimate the *interindividual* variability under those circumstances. It remains a tenable hypothesis that very little of the variability in individual exponents is a reflection of enduring idiosyncratic characteristics, and that small individual differences in sensory systems may adequately account for that little.

However, some caution should be observed in evaluating our findings or other results based on similar techniques. First, the kind of correlational analysis we have used may not always provide a definitive resolution of the individual differences problem. If the error associated with estimating individual exponents is high, then nonsignificant session-to-session correlations could result even though there were real individual differences. For the experiments reported in this paper, we have provided evidence against the applicability of this objection: Error in individual exponent estimation is low relative to variability over subjects, and therefore is an unlikely explanation of the absence of session-to-session correlation. Second, we have not investigated the problem using the technique employed by Green, Luce, and their associates, and there is no direct evidence that the large and striking individual differences that they find are not due to permanent characteristics of individuals. An attempt to apply the modulus-shift technique to their version of the magnitude estimation task is certainly warranted. However, there is evidence (King & Lockhead, 1982) that, even when many closely spaced signals are employed (as in the Green-Luce technique), subjects can learn, when given feedback, to provide for each of those signals an average judgment in close agreement with an experimenter-assigned value, even one sharply different from the one the subject could be expected to generate spontaneously. This finding, that sub-

jects can learn and remember assigned responses for as many as 30 signals spaced at 1-dB intervals, suggests that the Green-Luce technique is not immune to the influence of mnemonic processes. Thus, there is no more reason for regarding the large differences among individuals obtained by Green and Luce as evidence of permanent individual differences than there is in the case of the more traditional magnitude estimation procedures that we have employed. Third, it must be emphasized that we are *not* asserting that there are no individual differences in power law exponents. It would be astonishing if this measure were unique in being free of variability over individuals. What is at issue is whether the very large variability that has been frequently observed and reported is entirely due to permanent qualities of the individual subjects or whether a major part of that variability simply reflects a variety of uncontrolled procedural artifacts. Our correlational analysis suggests that the latter may be true. But only when we have learned to identify and control all of those artifacts will it be possible to generate the strongest evidence for the low-individual differences hypothesis—low variability in individual exponents within a single session. In the meantime, we feel that caution needs to be exercised from the other side of the issue as well: It is a mistake to assume, as has been commonly done in the past, that large intrasession differences in exponents prove that there are enduring and distinctive characteristics of the individual subjects' sensory or response mechanisms. They do not.

Finally, something needs to be said about the theoretical significance of this issue. The question of individual differences has often been raised in contexts in which little more is implied than the fact that individuals differ in some trait or quality. But in the case of psychophysical exponents, Luce (1972) has made clear the argument that the fundamental nature of psychophysical measurement may hinge on the question of variability over subjects. Summarizing his argument, Luce (1972) says that psychophysical measures are unlike physical measures because they "do not exhibit any fixed relation to physical measures . . . when examined over individuals" (p. 96). In our view, the evidence for this argument is not strong. What has been widely accepted as evidence for very large individual differences may be only the cumulative effects of such arbitrary experimenter variables as stimulus order and such incidental subject variables as modulus selection and memory. No matter how plausible Luce's suggested alternative view of psychophysical measurement as the outcome of observers' functioning as transducers calibrated differently from one another, the available evidence on individual differences in exponents does not require us to accept it.

REFERENCES

- CAIN, W. S. Differential sensitivity for smell: "Noise" at the nose. *Science*, 1977, **195**, 796-798.
- CROSS, D. V. Sequential dependencies and regression in psychophysical judgments. *Perception & Psychophysics*, 1973, **14**, 547-552.
- CURTIS, D. W., ATTNEAVE, F., & HARRINGTON, T. L. A test of a two-stage model of magnitude judgment. *Perception & Psychophysics*, 1968, **3**, 25-31.
- EKMAN, G., HOSMAN, B., LINDMAN, R., LJUNGBERG, L., & ÅKESSON, C. Interindividual differences in scaling performance. *Perceptual and Motor Skills*, 1968, **26**, 815-823.
- ENGELUND, W., & DAWSON, W. E. Individual differences in power functions for a 1-week intersession interval. *Perception & Psychophysics*, 1974, **15**, 349-352.
- ENGEN, T., & ROSS, B. M. Effect of reference number on magnitude estimation. *Perception & Psychophysics*, 1966, **1**, 74-76.
- GREEN, D. M., & LUCE, R. D. Variability of magnitude estimates: A timing theory analysis. *Perception & Psychophysics*, 1974, **15**, 291-300.
- GREEN, D. M., LUCE, R. D., & DUNCAN, J. E. Variability and sequential effects in magnitude production and estimation of auditory intensity. *Perception & Psychophysics*, 1977, **22**, 450-456.
- HELLMAN, R. P. Stability of individual loudness functions obtained by magnitude estimation and production. *Perception & Psychophysics*, 1981, **29**, 63-70.
- JESTEADT, W., LUCE, R. D., & GREEN, D. M. Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, **3**, 92-104.
- JONES, F. N., & MARCUS, M. J. The subject effect in judgments of subjective magnitude. *Journal of Experimental Psychology*, 1961, **61**, 40-44.
- JONES, F. N., & WOSKOW, M. On the relationship between estimates of magnitude of loudness and pitch. *American Journal of Psychology*, 1962, **75**, 669-671.
- KING, M. C., & LOCKHEAD, G. R. Response scales and sequential effects in judgment. *Perception & Psychophysics*, 1981, **30**, 599-603.
- LOGUE, A. W. Individual differences in magnitude estimation of loudness. *Perception & Psychophysics*, 1976, **19**, 279-280.
- LUCE, R. D. What sort of measurement is psychophysical measurement? *American Psychologist*, 1972, **27**, 96-106.
- LUCE, R. D., & MO, S. S. Magnitude estimation of heaviness and loudness by individual subjects. *British Journal of Mathematical Statistical Psychology*, 1965, **18**, 159-174.
- MARKS, L. W. *Sensory processes: The new psychophysics*. New York: Academic Press, 1974.
- MITCHELL, J. J., & GREGSON, R. A. M. Between-subject variation and within-subject consistency of olfactory intensity scaling. *Journal of Experimental Psychology*, 1971, **89**, 314-318.
- RULE, S. J., & MARKLEY, R. P. Subject differences in cross-modality matching. *Perception & Psychophysics*, 1971, **9**, 115-117.
- STEVENS, S. S. Measurement and man. *Science*, 1958, **127**, 384-389.
- STEVENS, S. S. The psychophysics of sensory function. *American Scientist*, 1960, **48**, 226-253.
- STEVENS, S. S. Neural events and the psychophysical law. *Science*, 1970, **170**, 1043-1050.
- STEVENS, S. S. Perceived level of noise by Mark VII and decibels (E). *Journal of the Acoustical Society of America*, 1972, **51**, 575-601.
- STEVENS, S. S. *Psychophysics: Introduction to its perceptual, neural, and social prospects* (G. Stevens, Ed.). New York: Wiley, 1975.
- STEVENS, S. S., & GUIRAO, M. Subjective scaling of length and area and the matching of length to loudness and brightness. *Journal of Experimental Psychology*, 1963, **66**, 177-186.
- TEGHTSOONIAN, M. The judgment of size. *American Journal of Psychology*, 1965, **78**, 392-402.

- TEGHTSOONIAN, M., & TEGHTSOONIAN, R. How repeatable are Stevens' power law exponents for individual subjects? *Perception & Psychophysics*, 1971, **10**, 147-149.
- TEGHTSOONIAN, R. On the exponents in Stevens' law and the constant in Ekman's law. *Psychological Review*, 1971, **78**, 71-80.
- TEGHTSOONIAN, R., & TEGHTSOONIAN, M. Range and regression effects in magnitude scaling. *Perception & Psychophysics*, 1978, **24**, 305-314.
- UNDERWOOD, B. J. Interference and forgetting. *Psychological Review*, 1957, **64**, 49-60.
- WALSH, J. K., & BROWMAN, C. P. Intraindividual consistency on a cross-modality matching task. *Perception & Psychophysics*, 1978, **23**, 210-214.
- WANSCHURA, R. G., & DAWSON, W. E. Regression effect and individual power functions over sessions. *Journal of Experimental Psychology*, 1974, **102**, 806-812.

NOTES

1. Data from the two experiments were combined to provide more stable estimates of the value of r . Analyses carried out on the data from each experiment separately lead to the same conclusion about the effect of delay.
2. These means of exponents and associated standard deviations,

and mean r 's and standard errors, are based on sets of subjects that do not overlap completely for number-SPL and length-luminosity exponents. However, if only scores from subjects who contributed both exponents ($N=79$) are considered, the values given here are virtually unchanged.

3. A pilot study, in which subjects matched numbers to sound pressure level at delays of up to 1 day, obtained similar correlations: For zero delay, $r(23)=+.66$ ($p < .05$); for 1-h delay, $r(26)=+.80$ ($p < .05$); for 1-day delay, $r(25)=+.70$ ($p < .05$). Hellman (1981) also reported a correlation of $+.95$ for seven subjects matching number to sound pressure level at 4-6-week delays (although a correlation at a 6-month delay was not significant).

4. We preferred this tactic to the alternative of changing the signal intensities from session to session. Unless the shift in signal intensity were large, we feared that the new set of signals might not be discriminated as being different from the first set and would be judged with the same, remembered set of responses. If the signal-intensity shift were large, any decorrelation of exponents might be attributed to the change in the region of the intensity scale being explored. In principle, the intensity-shift tactic is an attractive one, but seems to us to be burdened by more potential problems than does the one we chose.

(Manuscript received May 17, 1982;
revision accepted for publication November 30, 1982.)