

Stability of line-length estimates using the method of absolute magnitude estimation

RONALD T. VERRILLO
Syracuse University, Syracuse, New York

Absolute magnitude estimates of line length were made by two groups of nine subjects each. One group was tested and then retested at an intersession interval of 1 year and the other, at an intersession interval of 2 years. High temporal stability was observed in both groups in both the slope values, individual and group, of the power functions and in the numerical responses themselves.

The purpose of this study is to determine the temporal stability of responses to a line-length estimation task using the method of absolute magnitude estimation. The introduction of the method of magnitude estimation by S. S. Stevens (1955, 1965) was followed by an interest in the variability of individual power functions. Some studies reported considerable inter-individual variability of exponents in a number of sensory systems (Ekman & Åkesson, 1965; Jones & Marcus, 1961; Marks & J. C. Stevens, 1966; Schneider & Lane, 1963; J. C. Stevens & Guirao, 1964; J. C. Stevens & Mack, 1959). The stability of individual power functions over time, however, was not investigated systematically until M. Teghtsoonian and R. Teghtsoonian (1971) used magnitude estimation to scale length and area in five sessions separated by 24-h periods and again after an intersession interval of 1 year. They found individual exponents for both continua to be relatively stable for 24 h, but to lose stability beyond that interval.

Verrillo and Chamberlain (1971) compared the combined results of vibrotactile magnitude estimation and production from a single subject with those obtained 3 years later. The slopes of the two curves were identical, and the absolute positions were described as "remarkably close."

Engeland and Dawson (1974) found subjective area and loudness with a 1-week intersession interval to be stable (Kendall rank correlation, $p < .01$), as were cross-modality matches between these two continua. Logue (1976) reported individual exponents of loudness to be fairly stable ($r = +.7 - 50\%$ of variance) for up to 11 weeks, and Walsh and Browman (1978) reported stability over a 2-month period for a cross-modality matching procedure in vision and hearing. Hellman (1981) found individual exponent

differences between tones of different frequencies to be stable for periods up to 6 weeks.

Most previous investigators were concerned almost exclusively with the consistency of exponent values, often after normalization of the data, and virtually no attention was given to the numerical responses themselves. Experimental evidence based on the method developed by S. S. Stevens (1955, 1956) and modified by Hellman and Zwislocki (1961, 1963, 1964, 1968) suggest that subjects tend to use absolute, rather than ratio, scales. An absolute scale implies fixed units as well as absolute coupling between numerals and psychological magnitudes (Zwislocki & Goodman, 1980). The first evidence that subjects could couple numerals and sensations on an absolute scale was provided by Hellman & Zwislocki (1961). This observation led to a series of studies of absolute scaling of brightness (Barlow & Verrillo, 1976), loudness (Hellman, 1976; Hellman & Zwislocki, 1963, 1964, 1968), line length (Irvin & Verrillo, 1979; Verrillo & Irvin, 1979), surface area (Verrillo & Graeff, 1970), and vibrotaction (Verrillo, 1974; Verrillo & Chamberlain, 1971, 1972; Verrillo, Fraioli, & Smith 1969), all of which were consistent with the hypothesis that numbers and sensations could be matched by subjects on an absolute scale. Zwislocki and Goodman (1980) reviewed the literature and reported a series of experiments on loudness and line-length estimation that led them to conclude that numbers and sensation magnitudes can be paired on an absolute scale. The use of the method of absolute magnitude estimation permits a comparison not only of the exponent of the power curve, but also of the position of the curve with respect to the numerical responses in the investigation of the stability of subjects' responses. The method is described below.

METHOD

Two experiments were performed, one in which a 1-year interval and another in which a 2-year interval separated the two testing sessions. Nine subjects were tested in each experiment. No subject served in both experiments.

This research was supported by Grant NS-09940 from the National Institutes of Health, U.S. Department of Health and Human Services. The author's mailing address is: Institute for Sensory Research, Syracuse University, Syracuse, New York 13210.

The subjects in the first experiment were six males and three females, ranging in age from 22 to 55 years. The subjects in the second experiment were four males and five females and ranged in age from 20 to 50 years. No attempt was made to balance the groups with regard to age, because it had been shown that this variable had little effect on performance in the estimation of line length (M. Teghtsoonian & Beckwith, 1976; Verrillo, 1979, 1981; Zwislocki & Goodman, 1980). All subjects had normal or corrected-to-normal vision. Four subjects in each group had had previous experience with magnitude-estimation procedures, some with line lengths and some with vibrotactile subjective magnitude.

The subjects were seated comfortably at a distance of 3.7 m from and facing a 1.53×1.53 m white projection screen in a large, well-illuminated room (fluorescent, 540 lx). The subjects were situated 5 deg from a perpendicular to the center of the screen to allow for front projection. Each subject made two subjective estimates for each of seven black lines projected one at a time in a horizontal orientation. The lines were presented in the middle of the screen, approximately 157 cm from the floor. The average viewing angle in the vertical direction was about 6 deg. The luminance of the lines was 86 cd/m^2 on a 170 cd/m^2 background. The physical lengths of the lines measured on the screen were 1.4, 2.8, 5.0, 13.2, 26.6, 53.3, and 132.5 cm. Lines were presented in a random sequence for each run, with the restrictions that the first stimulus was neither the longest nor the shortest line length and that two lines of the same physical length did not appear in succession.

The method of absolute magnitude estimation was used (Hellman & Zwislocki, 1961). The subjects read a set of instructions while the experimenter read it aloud (Appendix). The subjects were instructed to assign numbers, without a standard or modulus, such that the subjective magnitude they brought to mind matched the subjective magnitude of the line length viewed. Note that subjects in the absolute-scaling procedure are asked to match the psychological values of numbers to the psychological values of sensations, rather than to report numbers whose ratios correspond to the ratios of the sensations. The subjects were given as much time as they needed to respond, typically 1 sec, before the presentation of the next stimulus, which followed the response after about a 1-sec delay.

RESULTS AND DISCUSSION

An assessment of the temporal stability of absolute-magnitude-estimation responses must include changes in both the slope values and in the numerical responses themselves. Slope values were determined for individual subjects and for groups by fitting lines to the data, plotted on log-log coordinates, by the method of least squares. The slopes of these lines, which correspond to exponents of power functions, served as the scores when slopes were analyzed. The product-moment coefficient of correlation, sometimes referred to as the coefficient of stability, was used to estimate the stability of the individual slope values between trials. A simple *t* test (two-tailed) of the differences between the geometric mean values of all numerical responses between trials was performed to determine the significance of the change in the absolute position of the curve. A high correlation between slope values coupled with a low *t* value for the numerical responses would imply that the results produced using the method of absolute magnitude estimation in judging line length are stable over time.

Figure 1 shows the geometric means with standard-error bars (Alf & Grossberg, 1979) of the numerical responses of nine subjects plotted as a function of the physical length of lines for two trials separated by an intertrial interval of 1 year. Linear regression lines computed by the method of least squares yielded slope values of .96 for the first session (O, solid line) and .98 for the second session (●, broken line) a year later. The r^2 values indicating the goodness of fit of the two lines are both .99. The coefficient of correlation computed on the slopes of the individual curves in the two sessions yielded a value of .799 ($df=7$), which is significant beyond the .01 level of confidence. The mean percent change in slope amounted to approximately +1.4%.

Correlation coefficients of each subject's numerical response to each line length in the two sessions yielded *r* values within a narrow range from .98 to 1.00, well beyond the .01 level of confidence for $df=5$.

The geometric means of all responses made by each individual on Trial 1 and on Trial 2 were calculated, and a *t* test was used to assess the Trial 1 – Trial 2 difference. The difference was not significant [$t(8) = .04$, $p < .50$]. As a further test of the change between sessions, *t* tests were performed on the numerical estimates of each line length across all subjects. The values of *t* ranged from .53 to 1.47. None reached significance at the .10 level of confidence.

The results of this experiment show that over a 1-year period there was no significant change in slope value or in the absolute values of the numbers used in

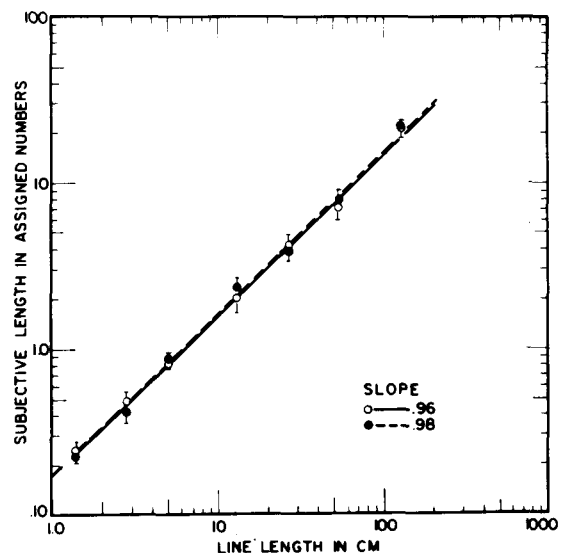


Figure 1. Geometric means and standard-error bars of the data of nine subjects tested and then retested after a period of 1 year. The data from the first test (O) are fitted by a solid line that has a slope of .99 ($r^2 = 1.0$). The data of the same subjects a year later (●) are fitted by a broken line that has a slope of .98 ($r^2 = .99$).

the scaling of line length by the method of absolute magnitude estimation.

The results of the experiment in which a 2-year interval separated the testing sessions (Figure 2) are a close approximation of the 1-year data. Mean slope values changed from .96 (O, solid line) to .97 (●, broken line) over the 2-year interval. The goodness of fit for the two functions was $r^2 = .99$ for both sets of data. The coefficient of correlation computed on the slopes of the two sets of individual data was .916 ($df = 7$), which is beyond the .01 level of confidence. The mean percent change in slope was +1.6%.

Correlation coefficients based on the raw numerical responses to each line length in the two sessions yielded r values between .96 and 1.00. A paired t test of the geometric mean over all absolute numerical estimates used in the two trials showed that the difference was not significant [$t(8) = 1.11, p < .10$].

The results of the retest after a 2-year interval are in accord with the data in which 1 year separated the two sessions; there were no appreciable changes either in slope or in the position of the power function for these nine subjects.

Table 1, in summarizing the results of the 1- and 2-year experiments, shows the geometric means of all responses made by each subject and the individual slope values for each trial.

Since neither slope nor the position of the curve changed over time, the two sessions for each group were collapsed over time to permit a comparison of the groups. The slope of the 1-year group is .97 and that of the 2-year group is .96; the fit of the data to the line was $r^2 = .99$ for both sets of data. Since the

individuals of the two groups could not be paired, a correlation coefficient based on the mean responses at each line length was calculated, yielding an r of .98, significant beyond the .01 level. A t test of the difference between the mean numerical estimates of the two groups collapsed over time yielded a t of 1.25 ($p < .10, df = 17$). The data show that there was a substantial agreement between the two groups in their responses.

A final analysis was performed to determine if there was a difference between the males and females of these groups in their estimation of line length. Verrillo (1979) reported that the perceived magnitude of a vibrotactile stimulus was greater for females than for males, but that no difference existed in the estimation of line length. Because time effects were not found in the present study, subjects were grouped according to sex from the 1- and 2-year groups. There were nine women and nine men. The percentage change from the first to the second test was 1.11% for women and 1.33% for men. The correlation of individual slope values between two trials was .928 ($p < .01, df = 7$) for women and .808 ($p < .01, df = 7$) for men. The absolute value of numerical estimates across line lengths was not significantly different over time for either the women [$t(8) = .75, p > .10$] or the men [$t(8) = .16, p > .50$]. Thus, men and women did not differ with respect to the stability of slope or absolute numerical estimation of line length. Neither the overall slope values nor the absolute numerical estimates across line lengths were different for the two sexes ($p > .10$ for the obtained value of t in each case).

The data to which these results are most directly comparable are those of M. Teghtsoonian and R. Teghtsoonian (1971), whose subjects estimated line length in sessions separated in time from 24 h to 1 year. Their data did not show the stability of the data reported here, but the experiments differed in several respects. The first difference was in the instructions to the subjects. Although Teghtsoonian and Teghtsoonian did not use a standard or modulus, they instructed subjects to "assign numbers according to the apparent length" of the stimulus. Our subjects were asked to assign numbers whose subjective magnitude matched the subjective magnitude of the line length. Perhaps this is a subtle difference, but slight differences in instructions can often account for large differences in results.

A second, and perhaps more significant, difference in the two studies was in the stimulus presentation. The subjects of M. Teghtsoonian and R. Teghtsoonian (1971) viewed two black markers moved along an aluminum bar to delineate line segments. The subjects in the present experiment viewed horizontal black lines projected on a white screen. Contrast, illumination conditions in the room, viewing angle, viewing distance, and height of the projected lines

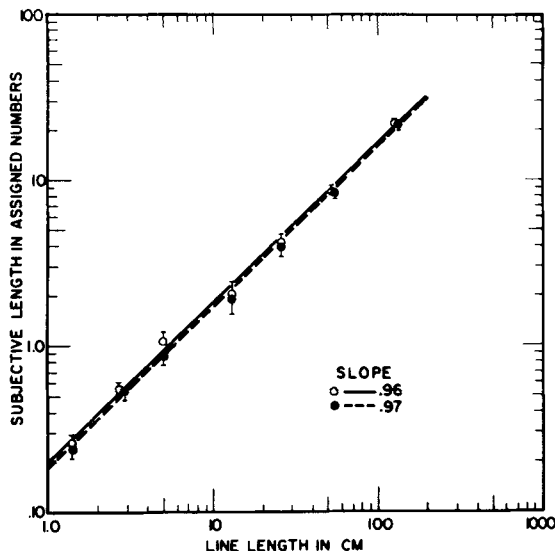


Figure 2. Same as Figure 1, but for nine subjects (not those in Figure 1) retested after a 2-year interval. The data of the first test (O) are fitted by a solid line that has a slope of .96 ($r^2 = .99$). The data of the same subjects 2 years later (●) are fitted by a broken line that has a slope of .97 ($r^2 = .99$).

Table 1
Geometric Means of Individual Numerical Responses and Individual Slope Values

Subject	1-Year Interval				Subject	2-Year Interval			
	Geometric Mean		Slope			Geometric Mean		Slope	
	Trial 1	Trial 2	Trial 1	Trial 2		Trial 1	Trial 2	Trial 1	Trial 2
D.G.	1.59	1.61	1.06	1.01	E.V.	1.65	1.64	.95	.93
B.R.	2.05	2.29	.95	.99	G.R.	2.06	1.98	.94	.94
S.C.	2.11	2.13	.94	.98	S.D.	3.68	3.73	.94	.93
R.B.	2.30	2.26	.92	.92	V.V.	2.66	2.52	.82	.87
L.G.	2.67	2.69	.93	.90	B.D.	1.58	1.37	1.12	1.18
J.L.	1.14	1.23	1.10	1.10	S.G.	2.01	2.16	.96	.95
D.M.	2.13	2.26	.98	.98	L.E.	2.73	2.13	.91	.95
S.L.	1.91	1.98	.95	1.03	L.D.	1.87	2.11	.93	.95
R.G.	2.56	1.87	.95	.94	A.D.	1.68	1.75	1.00	1.01
Mean*	2.00†	2.1†	.96	.98	Mean*	2.2†	2.1†	.96	.97
SE	.38	.41			SE	.41	.40		
r ²			.99	.99	r ²			.99	.99

*Grand geometric mean. †Based on all individual responses.

were carefully controlled. Teghtsoonian and Teghtsoonian controlled for the height of the bar and viewing distance, but did not mention the other experimental conditions in their report. Except for viewing distance (Verrillo & Irvin, 1979), it is not known what effects these experimental conditions can have on the outcome of experiments, or if they can account for the differences found between experiments. They are mentioned as possible reasons for the differences observed.

Zwislocki and Goodman (1980) and others have provided evidence that subjects tend to scale sensation magnitudes on absolute, rather than ratio, scales and that their responses are independent of the first stimulus scaled, the location of the stimulus range within the available range of stimuli, the extent of the range, and the range of available numbers. The present study adds a dimension of stability to the method. It is unlikely that the observed stability derives from a memory of numerals assigned to line lengths over periods of 1 and 2 years. After the retest session, the subjects were asked if they had remembered their previous responses. One (R.B., 1-year group) answered in the affirmative, but he did not produce a perfect replication of his earlier performance. Zwislocki and Goodman take the position that both sensations and numbers acquire absolute psychological magnitudes and that subjects can perform an absolute coupling of numbers with sensation magnitudes. This coupling has been shown to be stable across different age levels (Verrillo, 1981; Zwislocki & Goodman, 1980) and is here shown to be stable over time within individuals.

In summary, it is concluded that, when the method of absolute magnitude estimation of line length is used under carefully controlled conditions and with the instructions used in these experiments, the resulting power functions are stable with respect both to the slope of the power function and to the absolute

numerical responses of the subjects. The numbers used by the subjects to represent line length do not change significantly, and the rate of growth of subjective magnitude does not change for periods up to 2 years, nor do they change significantly between different groups of subjects.

REFERENCES

- ALF, E. F., & GROSSBERG, J. M. The geometric mean: Confidence limits and significance tests. *Perception & Psychophysics*, 1979, 26, 419-421.
- BARLOW, R. B., & VERRILLO, R. T. Brightness sensation in a ganzfeld. *Vision Research*, 1976, 16, 1291-1297.
- EKMAN, G., & ÅKESSON, C. Saltiness, sweetness, and preference: A study of quantitative relations in individual subjects. *Scandinavian Journal of Psychology*, 1965, 6, 241-253.
- ENGELAND, W., & DAWSON, W. Individual differences in power functions for a 1-week intersession interval. *Perception & Psychophysics*, 1974, 15, 349-352.
- HELLMAN, R. P. Growth of loudness at 1000 and 3000 Hz. *Journal of the Acoustical Society of America*, 1976, 60, 672-679.
- HELLMAN, R. P. Stability of individual loudness functions obtained by magnitude estimation and production. *Perception & Psychophysics*, 1981, 29, 63-70.
- HELLMAN, R. P., & ZWISLOCKI, J. Some factors affecting the estimation of loudness. *Journal of the Acoustical Society of America*, 1961, 33, 687-694.
- HELLMAN, R. P., & ZWISLOCKI, J. Monaural loudness function of 1000 cps and interaural summation. *Journal of the Acoustical Society of America*, 1963, 35, 856-865.
- HELLMAN, R. P., & ZWISLOCKI, J. Loudness function of a 1000 cps tone in the presence of a masking noise. *Journal of the Acoustical Society of America*, 1964, 36, 1618-1627.
- HELLMAN, R. P., & ZWISLOCKI, J. J. Loudness determination at low sound frequencies. *Journal of the Acoustical Society of America*, 1968, 43, 60-64.
- IRVIN, G., & VERRILLO, R. T. Absolute estimation of line length as a function of orientation and contrast polarity. *Sensory Processes*, 1979, 3, 261-274.
- JONES, F. N., & MARCUS, M. J. The subject effect in judgments of subjective magnitude. *Journal of Experimental Psychology*, 1961, 6, 40-44.
- LOGUE, A. W. Individual differences in magnitude estimation of loudness. *Perception & Psychophysics*, 1976, 19, 279-280.

- MARKS, L., & STEVENS, J. C. Individual brightness functions. *Perception & Psychophysics*, 1966, 1, 17-24.
- SCHNEIDER, B., & LANE, H. Ratio scales, category scales, and variability in the production of loudness and softness. *Journal of the Acoustical Society of America*, 1963, 35, 1953-1961.
- STEVENS, J. C., & GUIRAO, M. Individual loudness functions. *Journal of the Acoustical Society of America*, 1964, 36, 210-213.
- STEVENS, J. C., & MACK, J. D. Scales of apparent force. *Journal of Experimental Psychology*, 1959, 58, 405-413.
- STEVENS, S. S. The measurement of loudness. *Journal of the Acoustical Society of America*, 1955, 27, 815-820.
- STEVENS, S. S. The direct estimate of sensory magnitudes—loudness. *American Journal of Psychology*, 1956, 69, 1-25.
- TEGHTSOONIAN, M., & BECKWITH, J. B. Children's size judgments when size and distance vary. Is there a developmental trend to overconsistency? *Journal of Experimental Child Psychology*, 1976, 22, 23-39.
- TEGHTSOONIAN, M., & TEGHTSOONIAN, R. How repeatable are Stevens's power law exponents for individual subjects? *Perception & Psychophysics*, 1971, 10, 147-149.
- VERRILLO, R. T. Vibrotactile intensity scaling at several body sites. In F. A. Geldard (Ed.), *Cutaneous communication systems and devices*. Austin, Tex: Psychonomic Society, 1974.
- VERRILLO, R. T. Comparison of vibrotactile threshold and supra-threshold responses in men and women. *Perception & Psychophysics*, 1979, 26, 20-24.
- VERRILLO, R. T. Absolute estimation of line length in three age groups. *Journal of Gerontology*, 1981, 36, 625-627.
- VERRILLO, R. T., & CHAMBERLAIN, S. C. Direct scaling of vibrotaction: An individual replication. *Psychonomic Science*, 1971, 22, 225-226.
- VERRILLO, R. T., & CHAMBERLAIN, S. C. The effect of neural density and contactor surround on vibrotactile sensation magnitude. *Perception & Psychophysics*, 1972, 11, 117-120.
- VERRILLO, R. T., FRAIOLI, A. J., & SMITH, R. L. Sensation magnitude of vibrotactile stimuli. *Perception & Psychophysics*, 1969, 6, 366-372.
- VERRILLO, R. T., & GRAEFF, C. K. The influence of surface complexity on judgments of area. *Perception & Psychophysics*, 1970, 7, 289-290.
- VERRILLO, R. T., & IRVIN, G. Absolute estimation of line length and afterimage as a function of viewing distance. *Sensory Processes*, 1979, 3, 275-285.
- WALSH, J. K., & BROWMAN, C. P. Intraindividual consistency on a cross-modality matching task. *Perception & Psychophysics*, 1978, 23, 210-214.
- ZWISLOCKI, J. J., & GOODMAN, D. A. Absolute scaling of sensory magnitude. *Perception & Psychophysics*, 1980, 28, 28-30.

APPENDIX

In this experiment we are trying to find out how people match the subjective magnitudes of numbers to the subjective magnitude lengths of lines.

By subjective magnitude we mean the impression that a person may have in their mind of how large a number is or how long a line is. We all have ideas about how big or small things are even without measuring them. It is this impression that we are attempting to measure.

I will show you one at a time on the screen a series of lines having different lengths. I would like you to assign a number to each line so the impression, or subjective magnitude, that the number suggests to you matches your subjective magnitude of the line. Don't think in terms of physical units of measurement, such as inches or miles. Just choose numbers so that the subjective magnitudes they bring to your mind match the subjective magnitudes of the line lengths. You may use any numbers that seem appropriate—whole numbers, fractions, or decimals.

Treat each presentation individually. Do not try to remember what numbers you assigned to preceding lines.

Do you have any questions?

(Manuscript received May 3, 1982;
revision accepted for publication December 13, 1982.)