

Single judgments of numerosity

LESTER E. KRUEGER

Ohio State University, Columbus, Ohio 43210

In previous studies, subjects generally underestimated the number of elements present in a display. To eliminate the range and intertrial effects that arise when several displays are judged in succession and that might have produced the underestimation, subjects in the present study judged only a single display. The single judgments fitted a power function having an exponent of .83, which is consistent with previous data. Single judgments of loudness, area, and duration, by contrast, have produced abnormally low exponents apparently because the built-in scale unit, or modulus, available on numerosity is lacking for other modalities. The tendency to underestimate numerosity was much stronger for female than for male subjects.

S. S. Stevens (1975) said that "given a large enough number of observers, it might prove possible to verify the power law with each observer assigning only one number to one stimulus" (p. 29). Numerosity seems an ideal dimension for just such a test. The scale unit, or modulus, which nominally is free to vary if the subject is not told what number to call the first stimulus (standard), may be fixed by the dimension itself in the case of numerosity, and thus there would be no need to transform numerosity judgments in order to compensate for differences in the way different observers use numbers. A given degree of perceived numerosity likely would evoke nearly the same number response from various observers, whereas a given level of perceived loudness, say, might evoke a large number from some observers, but a small number from others. Numerosity is seemingly unique in that it is scaled in nonarbitrary units. Ten elements perceived as such cannot be called 5 or 20. With line length, by contrast, a 3-ft extent might be called 36 (if the modulus was set for inches) or 1 (if the modulus was set for yards). Normally, if the modulus is not fixed, the subject must provide several judgments, so that information on the exponent (slope on a log-log plot) may be dissociated from that on the modulus (intercept on a log-log plot).

If the modulus varied in a regular manner across observers, however, and observers were consistent in matching the subjective value of loudness, say, to the subjective value of number, as Zwislocki and Goodman

(1980) claimed, then the average number assigned would provide an absolute scale of loudness. According to Zwislocki and Goodman, there is "an absolute coupling between numerals and psychological magnitudes" (p. 28), which originates from the concept of numerosity in which numbers have absolute meaning. Zwislocki and Goodman found a marked tendency for observers to match a particular level of number to a particular level of auditory intensity or line length, irrespective of context, task (magnitude estimation, magnitude production), or even age (first-graders, adults). However, they found that experience decreased the assigned numbers on magnitude estimation, especially at low sensation levels, and that this effect took place quite rapidly, resulting in a steepening of the power function.

Miller and Dawson (Note 1; Miller, Note 2) also found that the power function was flatter early in practice, and they pinpointed the flattening as occurring on the very first judgment of loudness, perceived duration, or perceived area. They concluded that their subjects did not map a loudness, say, onto the number having the same subjective value but, rather, judged the ratio of successive loudnesses, which the subjects could begin to do only on the second judgment. Thus, they regarded their findings as support for the relational theories advanced by Krantz (1972, 1974), Shepard (1978, Note 3), and R. Teghtsoonian (1973; Teghtsoonian & M. Teghtsoonian, 1978). However, they did not obtain an absolutely flat, nonincreasing function on the first judgments, so the coding was at least partly absolute rather than relative. The failure to obtain the regular power-function exponent with first judgments might be due simply to unsystematic variability in the modulus used by different observers. If so, then single judgments of numerosity, in which the modulus is assumed to be fixed and equated across observers, based on prior experience, ought yet to produce a normal power function.

This study was supported in part by National Institute of Mental Health Grant MH32295. The author is grateful to Matthew Blazek, Ralph Gott, Ernest Hallford, Larry Hettinger, Robert Hilkert, Steve Kelso, Cheryl Lawton, Nancy Morris, Cynthia Reagan, Ronald G. Shapiro, Steve Rogers, and Holly Simpson for assistance in collecting the data, and to William E. Dawson, Mark E. Miller, Ronald G. Shapiro, and John Volkmann for helpful comments on earlier versions of this report. Requests for reprints should be sent to Lester E. Krueger, Ohio State University, Human Performance Center, 404-B West 17th Avenue, Columbus, Ohio 43210.

Minimizing Judgmental Bias

Obtaining single judgments of numerosity ought, in fact, to provide a more accurate and valid scale of perceived numerosity, because it eliminates the context and sequential effects present when a person judges several different magnitudes in immediate succession (e.g., Cross, 1973; Poulton, 1968, 1975, 1979). Numerosity is so simple to judge that subjects require no prior training or instruction on magnitude estimation, which also might bias judgments (Poulton, 1968). By obtaining only one judgment per subject, the present study used a between-subjects rather than within-subjects experimental design. Poulton (1975) said that "the results of a within-subjects design should always be regarded with suspicion" (p. 3). Poulton (1979) noted that transfer effects might occur between blocks of trials or experimental sessions if graduate students in a particular laboratory were tested repeatedly. The present study avoided such effects by obtaining a single judgment from each of 800 undergraduate students, who were enrolled in an introductory psychology course and had had minimal exposure, if any, to the concepts and methods of psychophysics.

However, even without previous experience in psychophysical scaling, a person's single judgment might yet be affected by the range and type of stimuli to which he or she had been exposed in everyday life (Greenwald, 1976; Poulton, 1975). People may constantly make implicit judgments concerning numerosity, loudness, brightness, etc. Therefore, the present experiment could not and did not exclude all contextual factors that might influence numerosity judgments. Some individuals, for example, may normally be exposed to an unusually wide range of numbers of elements and, as a result, might have a smaller exponent for perceived numerosity (Poulton, 1968, 1975, 1979). The present procedure ought to be more ecologically valid, though, in indicating how people, as conditioned or biased by the normal range and variety of external stimulation, judge numerosity in the real world.

The present study thus provided a further, seemingly less biased, test of whether people tend to overestimate or underestimate number. Which tendency predominates will depend on the scale constant or unit of measurement, k , and the exponent, β , in the formula $\psi = k\phi^\beta$, which depicts sensation magnitude, ψ , growing as a power function of the stimulus magnitude, ϕ (S. S. Stevens, 1975, p. 13). The scale constant, k , which determines the intercept on a log-log plot, reflects the size of the modulus, and the effect of modulus variability is sometimes called intercept variability (S. S. Stevens, 1975, p. 289). Even though the modulus is assumed to be built-in and fixed on the numerosity dimension, the scale constant, k , need not be fixed at 1.0.

Overestimation is favored to the extent that the

scale constant, k , and the exponent, β , exceed 1.0, and underestimation is favored to the extent that they fall below 1.0. If one value exceeds 1.0, but the other falls below 1.0, then overestimation may be found over part of the stimulus range and underestimation over the rest. In Abbey's (1962) case, k fell far below 1.0, so that, even with an exponent of 1.2, he consistently found underestimation for displays that ranged up to 200 dots in size. The more usual finding, however, is that k exceeds 1.0 but β does not, and that numerosity is judged accurately or overestimated until the number of elements reaches 10 to 20 (Kaufman, Lord, Reese, & Volkman, 1949; Minturn & Reese, 1951), beyond which underestimation occurs. S. S. Stevens (1957) relied on Taves's (1941) fractionation data in setting the exponent for perceived numerosity at 1.34, but on magnitude estimation, Taves's subjects underestimated the number of dots present and "on a log-log plot the slope of the function is less than unity" (Taves, 1941, p. 32). Bevan and Turner (1964) found a consistent tendency to underestimate numerosity, and the data for their control condition indicate an exponent of about .90. Krueger (1972) estimated the exponent to be .85, Indow and Ida (1977) estimated it to be .86, and Kaufman, Lord, Reese, and Volkman (1949) obtained data which indicate that the exponent is .87.

It is conceivable that the stimulus range used in the studies that found underestimation was such that it reduced the power-function exponent. The greater the stimulus range used, the lower the power-function exponent tends to be (Poulton, 1968, 1975, 1979; Teghtsoonian, 1973; Teghtsoonian & Teghtsoonian, 1978). On the other hand, the response range typically does not shrink or expand with the stimulus range when such familiar physical units as weight or numerosity are involved (Poulton, 1979). Range effects ought to be attenuated, too, insofar as subjects scale numerosity in an absolute fashion and are prone to assign a particular number to a particular sensory input regardless of context (Zwislocki & Goodman, 1980).

The power-function exponent may be lowered not only by increasing the stimulus range, but also by running a subject on a large number of trials (Baird, Green, & Luce, 1980) or by allowing concurrent variation on an extraneous dimension (Zwislocki & Goodman, 1980). The present study, by presenting only one stimulus to each subject, eliminated these factors as well as that of stimulus range. As long as the severe flattening found by Miller and Dawson on the functions for other modalities does not occur here, single judgments of numerosity ought to produce a higher exponent. The present findings thus may help to dispel the uncertainty concerning what the true exponent is for perceived numerosity, as well as to verify whether the general tendency to under-

estimate numerosity holds when range and intertrial effects are eliminated.

METHOD

Test Stimuli

A computer generated the printout sheets containing Xs. The lightly ruled (16 grey lines) sheet (37.94 cm wide \times 21.59 cm high) contained a 35 column \times 25 row matrix (26.19 cm wide \times 15.56 cm high) of possible X positions, which were randomly sampled to fill the number of X positions specified: 25, 35, 50, 75, 100, 150, 200, and 300. A separate display was generated for each subject, or 800 sheets in all. (For further details, see Krueger, 1972.)

Procedure

The task involved modulus-free magnitude estimation. The subject was asked to look at the sheet for a few seconds, and to report how many Xs appeared to be present on it. The subject was told not to count the Xs, but just estimate how many there were. The sheet then was laid in front of the subject, who was permitted to view the sheet for no more than about 5 sec.

Subjects

All 800 subjects were volunteers serving in various other experiments in order to receive credit in an introductory psychology course. Each subject was shown a single sheet for the numerosity judgment, either before or after participating in the other experiment. There were 100 subjects at each of the eight stimulus levels. Gender was not recorded for 57 subjects; of the rest, 339 were males and 404 were females.

RESULTS AND DISCUSSION

Exponent of Power Function

As the filled circles in Figure 1 show, the (geometric) mean responses formed a straight line on a

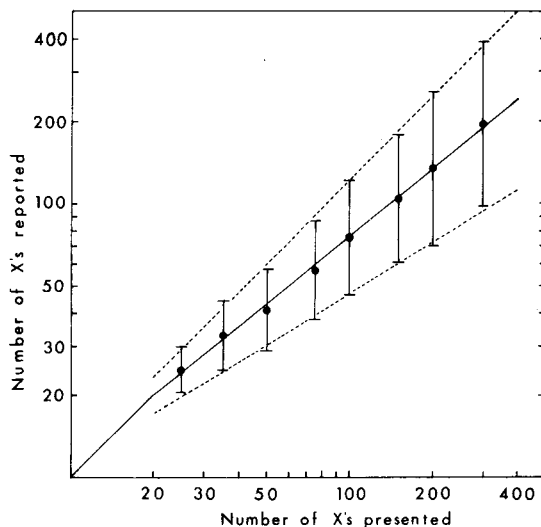


Figure 1. Geometric mean number of Xs reported (filled circles), by number of Xs presented, on log-log coordinates. The solid diagonal line has a slope of 1.0 (veridical function) until the number of Xs presented is 19.7, and then has a slope of .83 (least-squares fit to the geometric means). Each dotted diagonal line represents the least-squares fit to the standard deviation of the logarithmically transformed responses (solid vertical lines).

log-log plot, and thus verified Stevens's power law for single judgments. As the vertical bars in Figure 1 show, variability also produced a linear function on the log-log plot, and it will be discussed further below. Miller and Dawson (Note 1; Miller, Note 2) found that single judgments produced power functions having only 25% to 75% of the normal slope (exponent). In the present case, however, no flattening at all of the power function was evident. The present exponent of .83 was even slightly higher than the .76 magnitude-estimation exponent obtained by Krueger (1972). Thus, numerosity provides a surer means to obtain a reasonable facsimile of the normal power function with single judgments than does loudness, perceived duration, or perceived area, which Miller and Dawson used.

Why did numerosity produce different results? Probably because it has a fixed, built-in modulus. Or it may be because numerosity has been shaped by experience to a greater degree than the other dimensions, making it more natural to assign a particular number to a particular numerosity. The relational theory may hold to a greater extent for the other dimensions, as Miller and Dawson claimed. Relatively objective instructions were used in the present study (see Procedure section), which may have prompted some subjects to judge numerosity in an absolute rather than relative manner (Dawson, Note 4). On a set of regular, two-dimensional figures, M. Teghtsoonian (1965) found a larger exponent when subjects were asked to judge objective area (1.03) rather than apparent size (.76), a difference that, she argued, depended on the subject's estimating, then squaring, a linear dimension under the former instruction. In the present case, the Xs were scattered irregularly over a 35 \times 25 matrix, but the subjects still may have been able to estimate objective numerosity by multiplying the approximate number of rows by the approximate number of Xs per row. However, in spite of this and other factors, such as experience, which also ought to make numerosity a more objective dimension, the subjects underestimated the number of Xs present to a considerable extent. As Figure 1 shows, when 300 Xs were presented, the average (geometric mean) judgment was just under 200 Xs. The geometric mean attenuates the effect of the higher numbers assigned, but even when arithmetic means were used instead (Baird, Green, & Luce, 1980), the exponent was only .92.

That people can change their personal power-function exponent for numerosity is shown by the findings of Minturn and Reese (1951), who had subjects judge numerosity first without feedback as to the correct number of dots presented, then with feedback, and finally, 8 months later, without feedback. Providing feedback reduced the tendency to deviate from the correct function, either by overestimating

or underestimating numerosity. A general tendency favoring underestimation was virtually eliminated when feedback was provided, and it did not return when subjects were tested without feedback eight months later.

Elimination of Regression Effects

The .83 exponent obtained here is quite close to the .85 composite exponent on numerosity obtained by Krueger (1972), using magnitude estimation (.76) and magnitude production (.94). Magnitude production typically produces a smaller exponent than does magnitude estimation, presumably because an opposing bias is involved (S. S. Stevens, 1975). The present study did not use magnitude production, because it would have required exposing the subject to a range of stimuli even on the first judgment, but fortunately there was no bias or regression effect for it to eliminate, since obtaining only a single judgment from each subject increased the exponent on magnitude estimation to nearly the .85 level.

The disparity in exponent between the two tasks may reflect sequential effects (Cross, 1973). The judgment on one trial typically is assimilated towards that on the preceding trial (Lockhead, 1973; Ward, 1979), which tends to be located towards or beyond the center of the series, relative to the current stimulus. The judgment thus is pulled towards the center of the response range, which reduces the exponent for magnitude estimation but increases it for magnitude production. If the regression effect depends largely on sequential effects, however, then in the present study, which used stimulus materials very similar to those used by Krueger (1972) but eliminated sequential effects, the magnitude-estimation exponent ought to be about .85, which is just what was found. Thus, when the subject cannot reflect on how the present stimulus relates to other stimuli and thus any effects involving assimilation (Cross, 1973) or the avoidance of extreme response ratios (Teghtsoonian, 1973; Teghtsoonian & Teghtsoonian, 1978) are eliminated, the power-function exponent for magnitude estimation increases somewhat, but only enough to eliminate the regression effect.

If the regression effect is based on a contraction bias (i.e., a tendency to select a response that is closer to the middle of the response range; Poulton, 1979), rather than assimilation or the avoidance of extreme response ratios, then it will be present even when only a single judgment is obtained. Poulton (1979) cited cases in which the contraction bias was evident on the very first judgment. The contraction bias, if present, could not have been very large, however, because when the direction of bias was reversed, that is, by using magnitude production, the exponent obtained was no higher than .94 (Krueger, 1972). Furthermore, the increase in the magnitude-estimation exponent from .76 to .83 suggests that sequential effects

had a more important role and that little, if any, contraction bias was present.

However, other aspects of the present data are consistent with the presence of a contraction bias. As shown in Figure 1, the empirical function intersected the veridical function at about a numerosity of 20, but the (predicted) variability was still sizable at this point, so perhaps the function should be projected to indicate overestimation below 20. In this case, 20 would be the fulcrum or centering point towards which responses contract. Minturn and Reese (1951) found that numerosity tended to be underestimated when more than 20 elements were presented and overestimated when less than 20 elements (but more than 5 elements) were presented. Kaufman, Lord, Reese, and Volkman (1949) also found overestimation in the lower region, but the crossover point in their case was at 11, not 20.

Variability of Judgments

The vertical bars in Figure 1 indicate the standard deviation of the (log) number of Xs reported for each of the eight stimulus magnitudes. As Kaufman et al. (1949) and others have found, variability increased as the number of Xs presented increased. In fact, the variability itself formed a power function. The standard deviation of the (logarithmically transformed) responses increased as the .20 power of the number of Xs presented. This power function is shown by the two dotted lines in Figure 1, which, if extended to the left, would converge at a (linear) value of 8.85 Xs presented. This is close to the value of 6, which Kaufman et al. (1949) reported to be the largest number that can be subitized, that is, apprehended in a very rapid, confident, and nearly errorless manner.

The upper and lower vertical bars at each point in Figure 1 encompassed approximately the same proportion of judgments, because the (logarithmically transformed) responses at each stimulus level generally formed a symmetrical distribution. In fact, as Figure 2 shows, several of them closely approximated a normal distribution, as J. C. Stevens (1957) found. These log-normal distributions provided one justification for using geometric means rather than arithmetic means in plotting the power function shown in Figure 1 (cf. S. S. Stevens, 1971).

Some distributions in Figure 2 did not approximate the normal as well as others, though, owing to a more pronounced round-number tendency. Arrows in Figure 2 indicate the location on the abscissa of round-number responses given by 14 or more of the 100 subjects tested at the particular stimulus level. Only 11% (86) of the 800 responses were not round numbers, that is, did not end in 0 or 5. In fact, the four most popular responses (25, 30, 50, 100) constituted 32% (258) of the reports. Other investigators also have found that subjects prefer certain numbers (Baird & Noma, 1978). When z scores were combined

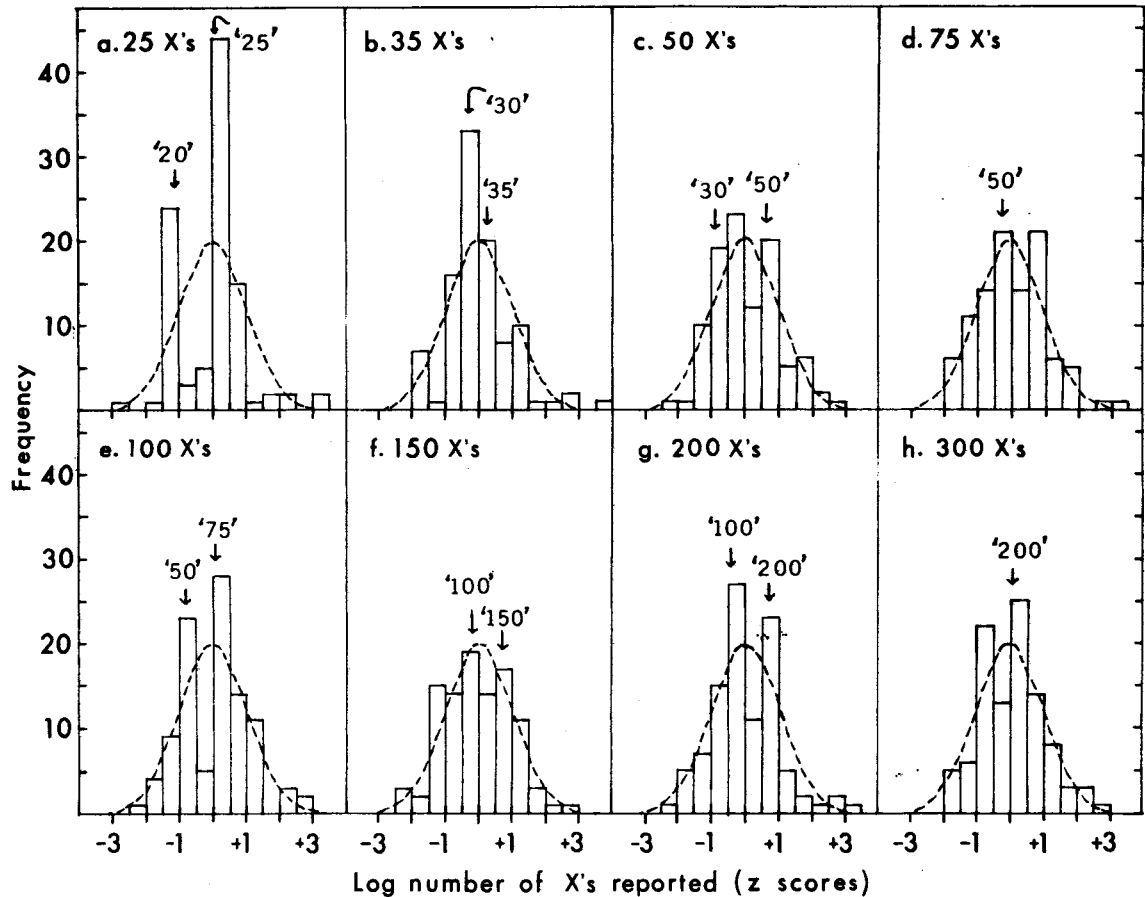


Figure 2. Distributions of z scores of logarithmically transformed responses at each stimulus level. Arrows indicate locations on the abscissa of responses made by 14 or more of the 100 subjects at that stimulus level. The dotted lines depict normal distributions of z scores.

across stimulus level, however, the resulting distribution, shown in Figure 3, was little perturbed by the round-number tendency. A Kolmogorov-Smirnov one-sample test (Siegel, 1956) indicated that the 800 z scores plotted in Figure 3 did not deviate significantly from the normal distribution (shown as a dotted line) ($D = .025, p > .20$).

The log-normal distribution of responses shown in Figure 3 may reflect random variation at the stimulus or response stage. Each subject received the Xs in a different random arrangement, and some arrangements might have favored a larger numerical response than others. It is not clear, though, how this variation could produce a log-normal distribution of responses. There also may have been variation in the way numbers were assigned to particular sensory impressions. The absolute coupling of number to sensory impression, as postulated by Zwislocki and Goodman (1980), may be merely a central tendency. A log-normal distribution would be produced if variation in the coupling occurred on a logarithmically transformed number scale. Poulton (1979) and others

have referred to such a logarithmic bias in the way in which numbers are used.

Variation in the Exponent

A third way to account for the log-normal distribution is to posit variation in the size of the exponent. If the exponents for the 800 subjects were distributed normally, with a mean value of .83 and a standard deviation of .20, and if the variability in judgment reflected solely variability in exponent, then the data predicted would match those observed. Thus, it might be more appropriate to report the exponent obtained here as $.83 \pm .20$.

Baird and Noma (1978, Figure 6.1) presented distributions of exponents that were symmetrical and, at least for line length, approximately normal. To determine whether the distribution would be normal on numerosity, the exponents obtained for each of Krueger's (1972) 94 magnitude-estimation subjects were plotted, as shown in Figure 4. On a Kolmogorov-Smirnov one-sample test (Siegel, 1956), the deviation of the observed distribution of exponents from the

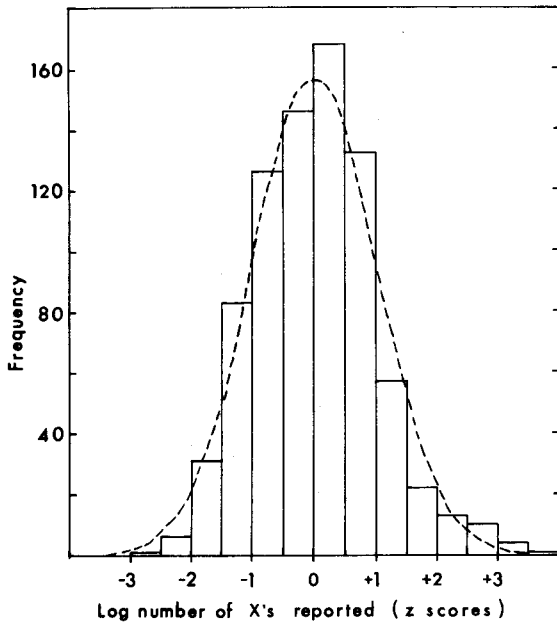


Figure 3. Distribution of z scores of logarithmically transformed responses, summed across the eight stimulus levels as shown in Figure 2. The dotted line depicts a normal distribution of z scores.

theoretical (normal) distribution (dotted line in Figure 4) was not significant ($D = .0319$, $p > .20$). The exponents had a mean value of .75 and a standard deviation of .20. Thus, the variability was the same as in the present study, though the mean value was lower.

Variation in the size of the exponent could be due in part to individual differences. Some people might tend to underestimate numerosity more than others. Such consistency within subjects was evident in a reanalysis of Krueger's (1972) data. In Experiment 1, the correlations of the exponent between pairs of the four sets of trials, representing different density con-

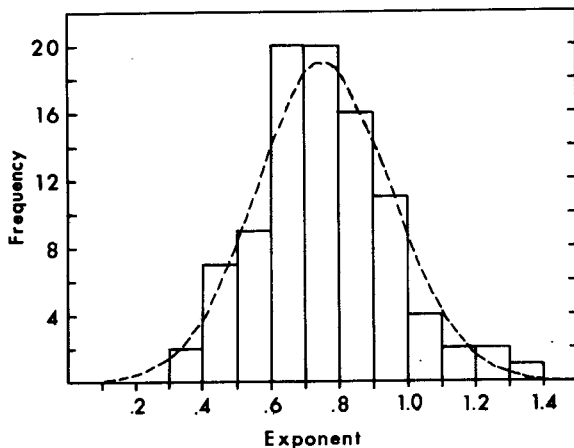


Figure 4. Distribution of exponents for Krueger's (1972) 94 magnitude-estimation subjects. The dotted line depicts a normal distribution.

ditions, ranged from $r(28) = +.58$ ($p < .005$) to $r(28) = +.68$ ($p < .005$). In Experiment 2, the correlations between pairs of the three sets of trials, representing different density conditions, ranged from $r(30) = +.33$ ($p < .05$) to $r(30) = +.67$ ($p < .005$). Minturn and Reese (1951) likewise found that 4 of their 10 subjects tended to overestimate numerosity and 6 tended to underestimate it. Rule (1966) found a correlation coefficient of .75 between the numerosity exponent obtained in his first two sets of trials and the last two sets of trials. He studied circle size and line length, as well as numerosity, and the correlations between continua on the size of the exponent ranged from .45 to .63.

The Gaussian distribution of the exponent (Figure 4) indicates that its value for a particular person and/or occasion depends on many random, independent events. However, there also may be major systematic influences, reflecting the effects of sex differences (discussed below), training, feedback (Minturn & Reese, 1951), and motivational factors. Children in Nelson and Lechelt's (1970) less-favored socio-economic group had a higher exponent on perceived numerosity when coins were shown (1.11) than when slugs were shown (.92).

An alternative possibility is that a lower exponent for a particular person reflects not an increased tendency to underestimate numerosity, but an increased contraction bias. There may be individual differences with respect to the regression effect (Stevens, 1971), that is, the tendency to center judgments around the average value of the response and to avoid extreme responses. A large centering tendency would lead to a relatively low exponent on magnitude estimation but a high exponent on magnitude production; a small centering tendency would produce the reverse pattern. Thus, there would be a negative correlation between the magnitude-estimation exponent and the magnitude-production exponent, and balancing for regression effects would eliminate individual differences in the size of the exponent. On the other hand, if variation in the size of the exponent reflects variation in the underestimation tendency, then the correlation ought to be positive rather than negative.

Wanschura and Dawson (1974) found that consistent individual differences did not disappear when they balanced for regression effects. They found an overall positive correlation ($r = +.30$) between exponents based on matching loudness to apparent duration and those based on matching apparent duration to loudness. The correlation was always positive, and was significant in one session, though not for the combined data. A reanalysis of the data from Krueger's (1972) Experiments 3 and 4 likewise revealed a positive correlation of $r(30) = +.30$ ($p < .05$) between magnitude estimation and magnitude production. The correlation within the magnitude-production task, that is, between trials in which numerosity was

adjusted upward and trials in which it was adjusted downward, was only slightly larger [$r(30) = +.34$, $p < .05$]. Thus, the downward dip in the magnitude-estimation exponent for some subjects in the present study more likely reflects an increased tendency to underestimate numerosity than an increased contraction bias.

Sex Differences

The tendency to underestimate numerosity was much stronger for female than for male subjects. At seven of the eight stimulus levels, the geometric-mean response was lower for females than for males, as Figure 5 shows. To equalize cell size for an analysis of variance, the subject population for each display size was pared to 39 males and 39 females by randomly removing subjects. In this test, males had significantly larger reports than did females [$F(1,608) = 6.51$, $p < .025$]. The sex effect was significantly greater [$F(1,608) = 5.21$, $p < .025$] on larger displays (100, 150, 200, 300 Xs) than on smaller displays (25, 35, 50, 75 Xs). The sex difference was mainly in the exponent (slope effect), which was .87 for males and .78 for females; the (common log) intercept was .16 for males and .29 for females. With the scores logarithmically transformed, the mean (common log) variance across the eight stimulus levels was virtually identical for males (.04254) and females (.04277). Thus, the variability of the exponent was the same for males and females.

Many factors might account for the sex differences shown in Figure 5. Males generally appear to

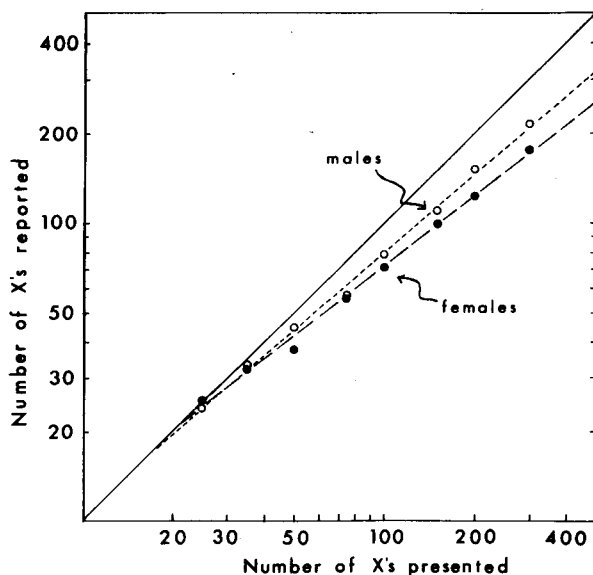


Figure 5. Geometric mean number of Xs reported by 339 male subjects (open circles) and 404 female subjects (filled circles), by number of Xs presented, on log-log coordinates. The solid diagonal line (veridical function) has a slope of 1.0.

have much better visual acuity under photopic conditions (Burg, 1966; Roberts, 1964), and they may have given larger numerosity responses simply because they saw the Xs more clearly and distinctly. Males may have been more accurate because they generally do better on visual-spatial tasks, such as the embedded or hidden figures test (Maccoby & Jacklin, 1974). A smaller exponent would be predicted for females if they were normally exposed to a wider range of number of elements than were males (cf. Poulton, 1968, 1975, 1979). The present effect may be task-specific, reflecting sex differences in the use of numbers or in susceptibility to regression effects and the like. Given the male advantage on arithmetical-reasoning and quantitative tasks (Maccoby & Jacklin, 1974), males and females might differ in general on how they do magnitude estimation and magnitude production.

REFERENCE NOTES

1. Miller, M. E., & Dawson, W. E. *Repeated-measures designs, the range effect, and relativity in psychophysical scaling*. Paper presented at the meeting of the Midwestern Psychological Association, St. Louis, May 1980.
2. Miller, M. E. *Relational judgment vs. stimulus mapping in direct ratio scaling*. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1980.
3. Shepard, R. N. *What does a psychophysicist measure?* Unpublished paper, Harvard University, 1966.
4. Dawson, W. E. Personal communication, November 1980.

REFERENCES

- ABBEY, D. S. Cross-modality matching of numerosity and pitch. *Canadian Journal of Psychology*, 1962, 16, 283-290.
- BAIRD, J. C., GREEN, D. M., & LUCE, R. D. Variability and sequential effects in cross-modality matching of area and loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 1980, 6, 277-289.
- BAIRD, J. C., & NOMA, E. *Fundamentals of scaling and psychophysics*. New York: Wiley, 1978.
- BEVAN, W., & TURNER, E. D. Assimilation and contrast in the estimation of number. *Journal of Experimental Psychology*, 1964, 67, 458-462.
- BURG, A. Visual acuity as measured by dynamic and static tests: A comparative evaluation. *Journal of Applied Psychology*, 1966, 50, 460-466.
- CROSS, D. V. Sequential dependencies and regression in psychophysical judgments. *Perception & Psychophysics*, 1973, 14, 547-552.
- GREENWALD, A. G. Within-subjects designs: To use or not to use? *Psychological Bulletin*, 1976, 83, 314-320.
- INDOW, T., & IDA, M. Scaling of dot numerosity. *Perception & Psychophysics*, 1977, 22, 265-276.
- KAUFMAN, E. L., LORD, M. W., REESE, T. W., & VOLKMAN, J. The discrimination of visual number. *American Journal of Psychology*, 1949, 62, 498-525.
- KRANTZ, D. H. A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology*, 1972, 9, 168-199.
- KRANTZ, D. H. Measurement theory and quantitative laws in psychophysics. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical*

- psychology (Vol. 2): *Measurement, psychophysics, and neural information processing*. San Francisco: Freeman, 1974.
- KRUEGER, L. E. Perceived numerosity. *Perception & Psychophysics*, 1972, 11, 5-9.
- LOCKHEAD, G. R. Choosing a response. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.
- MCCOBY, E. E., & JACKLIN, C. N. *The psychology of sex differences*. Stanford, Calif: Stanford University Press, 1974.
- MINTURN, A. L., & REESE, T. W. The effect of differential reinforcement on the discrimination of visual number. *Journal of Psychology*, 1951, 31, 201-231.
- NELSON, T. M., & LECHHELT, E. C. Socioeconomic status, value, and response to number. *Perception & Psychophysics*, 1970, 8, 76-80.
- POULTON, E. C. The new psychophysics: Six models for magnitude estimation. *Psychological Bulletin*, 1968, 69, 1-19.
- POULTON, E. C. Range effects in experiments on people. *American Journal of Psychology*, 1975, 88, 3-32.
- POULTON, E. C. Models for biases in judging sensory magnitudes. *Psychological Bulletin*, 1979, 86, 777-803.
- ROBERTS, J. *Binocular visual acuity of adults*. Washington, D.C: U.S. Department of Health, Education, and Welfare, 1964.
- RULE, S. J. Subject differences in exponents of psychophysical power functions. *Perceptual and Motor Skills*, 1966, 23, 1125-1126.
- SHEPARD, R. N. On the status of "direct" psychophysical measurement. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science* (Vol. 9): *Perception and cognition issues in the foundations of psychology*. Minneapolis: University of Minnesota Press, 1978.
- SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- STEVENS, J. C. *A comparison of ratio scales for the loudness of white noise and the brightness of white light*. Doctoral dissertation, Harvard University, 1957.
- STEVENS, S. S. On the psychophysical law. *Psychological Review*, 1957, 64, 153-181.
- STEVENS, S. S. Issues in psychophysical measurement. *Psychological Review*, 1971, 78, 426-450.
- STEVENS, S. S. (G. Stevens, Ed.). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: Wiley, 1975.
- TAVES, E. H. Two mechanisms for the perception of visual numerosity. In *Archives of Psychology* (No. 265). New York: 1941.
- TEGHTSOONIAN, M. The judgment of size. *American Journal of Psychology*, 1965, 78, 392-402.
- TEGHTSOONIAN, R. Range effects in psychophysical scaling and a revision of Stevens' law. *American Journal of Psychology*, 1973, 86, 3-27.
- TEGHTSOONIAN, R., & TEGHTSOONIAN, M. Range and regression effects in magnitude scaling. *Perception & Psychophysics*, 1978, 24, 305-314.
- WANSCHURA, R. G., & DAWSON, W. E. Regression effect and individual power functions over sessions. *Journal of Experimental Psychology*, 1974, 102, 806-812.
- WARD, L. M. Stimulus information and sequential dependencies in magnitude estimation and cross-modality matching. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, 5, 444-459.
- ZWISLOCKI, J. J., & GOODMAN, D. A. Absolute scaling of sensory magnitudes: A validation. *Perception & Psychophysics*, 1980, 28, 28-38.

(Manuscript received July 31, 1981;
revision accepted for publication October 21, 1981.)