Notes and Comment

Why unbiased numerical magnitude judgments of the loudness of noise are linear in decibels: A rejoinder to the Teghtsoonians

E. C. POULTON

Applied Psychology Unit, Cambridge, England

Teghtsoonian and Teghtsoonian (1986) have taken me to task for publishing data that are "consistent with the view that the subjectively equal stimulus spacing for the loudness of noise is linear in decibels, and that observers use numbers linearly in judging loudness" (Poulton, 1984, p. 338). The Teghtsoonians argue that the data can be accounted for equally well by power functions; they assume that the observers use a scale of numbers calibrated in ratios. Yet this is true of only some of the data, and there are other grounds for believing that the observers use a linear scale of numbers.

Two Ways of Using Numbers

In Poulton (1984), I described the linear relation between the loudness of noise and decibels without presenting the historical background, which will be found in greater detail elsewhere (Poulton, 1983). Torgerson (1960, p. 24) pointed out that observers have two ways of responding to the size of a difference between two stimuli. They can treat the difference as an interval, using for responses a linear scale of numbers similar to category ratings. Alternatively, they can treat the difference as a ratio, using a scale of numbers calibrated in ratios, as on a slide rule (Weiss, 1981).

As I have stated before (Poulton, 1981), there are two kinds of stimulus magnitudes. For stimulus magnitudes such as simple lengths, weights, and durations, observers have familiar physical units that they can use for their judgments. Once they have converted the stimulus magnitudes into these units, they can judge differences by subtracting the numerical values, and ratios by dividing them. By contrast, for stimulus magnitudes such as loudness, brightness, and lightness, ordinary observers do not have familiar physical units to use. They have only one way to perceive the difference between two magnitudes, although they can still respond using a numerical scale calibrated either linearly or in ratios.

The numerical scale that the observers use for their responses has to be inferred from their judgments. If the observers use a numerical scale that is calibrated linearly,

the numbers will be related linearly to the subjectively equal stimulus spacing. In the case of the loudness of noise, unbiased category ratings (Montgomery, 1975; see Poulton, 1979, Figure 5) indicate that the subjectively equal stimulus spacing is logarithmic, or linear in decibels. Thus the numbers will be related linearly to decibels. By contrast, if the observers use a numerical scale that is calibrated in ratios, the logarithms of the numbers will be related linearly to the subjectively equal stimulus spacing. For the loudness of noise, this gives the logarithmic relation favored by the Teghtsoonians between the numbers and decibels, which is a power function.

A more detailed discussion of Torgerson's (1960) distinction between the two ways of responding to the same pairs of stimuli was presented by Treisman (1963, 1964). The distinction has been demonstrated experimentally by Birnbaum (1980) and associates, who have used a number of different stimulus dimensions, including the loudness of a 1-kHz tone (Birnbaum & Elmasian, 1977). Birnbaum (1980, p. 304) showed that "judges compare two stimuli by subtraction," whether they respond using a numerical scale calibrated linearly or in ratios.

The Teghtsoonians (1986) acknowledge that observers can respond using a numerical scale calibrated either linearly or in ratios. But they do not point out that the two ways of responding to the same differences in loudness are responsible for the logarithmic and loglog, or power function, relations, respectively. Instead they merely express a preference for "subjects who both see the task as one of judging sensation ratios and have a sufficiently sophisticated understanding of the system of numbers to use it for that purpose" (pp. 74-75). In Teghtsoonian and Teghtsoonian (1978), in the section on magnitude estimation of loudness, they treated all responses as ratio responses. They fitted power functions to the median responses without regard for whether the observers had responded using a numerical scale calibrated linearly or in ratios. This inappropriate procedure is discussed in the last section of this note.

Logarithmic Response Bias

When Stevens (1975) instructed his observers to use a scale of numbers calibrated in ratios, he introduced what can be described as a *logarithmic response bias* (Poulton, 1979, p. 783) or *exponential response transformation* (Birnbaum & Elmasian, 1977). A logarithmic response bias can be produced also by responding with a range of numbers that contains a step change in the number of digits. After responding with numbers up to 10, observers tend to continue with 20, 30, 40, . . . rather than 11, 12, 13, . . . This can be described as a logarithmic response bias.

Financial support from the British Medical Research Council is gratefully acknowledged. Requests for reprints should be addressed to E. C. Poulton, M.R.C., Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England.

Figure 1. Very first multiple numerical judgments of the loudness of noise made against a standard of about 65 dB(A), called 1.0. Each dot represents the judgment of a separate observer. M = median, dB = decibels above the standard. (Poulton, 1984, Figure 3, after Poulton, 1982)

The very first judgments illustrated in Figure 1 (Poulton, 1984, Figure 3) clearly demonstrate the logarithmic response bias that is produced by a step change in the number of digits within the range of numbers used as responses. At 10 there is a complete discontinuity in the distribution of responses. Very first numerical judgments smaller than 10 almost invariably occur at digit intervals. Numerical judgments greater than 10 always occur at intervals greater than one digit. There are altogether 7 judgments of 9, but not a single judgment of 11. The difference is reliable ($p < .05$) on a binomial test. A similar reliable difference is found in the distributions of very first multiple numerical judgments of lightness. Out of a total of 350 responses, there are 13 responses of 9, but only 3 responses of 11 ($p < .05$) (Poulton, 1983, Figure 1.1).

In Poulton's (1984) Figures 1 and 2, the size of the stimulus range is small. Thus all the responses fall between 10 and 80. There is no step change within the range of responses to produce a logarithmic response bias. In addition, the instructions (p. 342) are designed to encourage the observers to judge in units comparable to units of length, using a demonstration line drawn on a blackboard. The instructions are not "deliberately ambiguous," as the Teghtsoonians state. And because in everyday life ordinary people generally compare magnitudes using a numerical scale calibrated linearly like a ruler, there is no reason to believe that the observers use a numerical scale calibrated in ratios.

Clear and Less Clear Distinctions Between Linear and Logarithmic Relations

In Poulton (1984), I claim only that the data presented are consistent with the view that there is a linear relation between loudness and decibels. Two of the results clearly favor a linear relation. The remaining results can be fitted fairly well by a logarithmic relation between loudness

and decibels, as the Teghtsoonians point out. I did not claim that none of the results can be fitted by a logarithmic relation, as the Teghtsoonians imply. If a logarithmic relation is to be ruled out, it has to be partly for the reasons that I have just given.

The magnitude judgments illustrated in Figures 1 and 2 of Poulton (1984) are all compatible with a linear relation between loudness and decibels. Figure 2A shows that Subject 1 had a median geometric response range of 7. This is large enough to show clearly that a linear relation between loudness and decibels is a better fit than a logarithmic relation. On a loglog plot (not shown) the best fitting function has considerable curvature.

The other clear evidence for a linear relation between loudness and decibels comes from the Poulton's (1984) Figure 3, which is reproduced here as Figure 1. Responses less than 10 are spaced at digit intervals, and give medians that are almost linear in decibels. The group of observers that judged the variable 5 dB above the standard gave a median of 2. Thus, adding 5 dB here adds one unit of loudness. The two groups of observers that judged the variable 20 dB above the standard gave a combined median of 6. This compares with a predicted median of 5 if each 5 dB adds one unit of loudness. The agreement is reasonably good. If the observers had used numbers as ratios, the median of 2 for the loudness of the variable 5 dB above the standard would predict a median of 16 for the variable 20 dB above the standard. This is far removed from the observed median of 6. Observers have also been shown to use numbers linearly in making very first fractional numerical judgments of lightness and darkness (Poulton & Simmonds, 1985).

By contrast, Figure 1 shows that responses greater than 10 always occur at intervals greater than one digit. This logarithmic response bias puts the median of 20 for the variable 35 dB above the standard in a different category from the remaining medians. Yet the Teghtsoonians

(1986, p. 74) fit a power function to all the medians in Figure 1 together, disregarding the visible discontinuity at 10 between responses greater than and less than 10.

Logarithmic Response Bias Counteracts the Stimulus Range Equalizing Bias

Figure 2 provides the background for the Teghtsoonians' (1986) Appendix. The solid curved function illustrates R. Teghtsoonian's (1971) reciprocal relation across stimulus dimensions between some of Stevens's (1962) exponents and the log geometric stimulus ranges that Stevens and his students used to determine the exponents. The label of 30 attached to the solid function represents an average geometric response range of 30 in the equation

$$\text{Exponent} = \frac{\text{Log geometric response range}}{\text{Log geometric stimulus range}} \quad (1)$$

The dotted functions labeled 13, 5, and 2.5 represent the reciprocal relation with smaller geometric response ranges. In his original article, R. Teghtsoonian (1971) quoted an average geometric response range of 34 for Stevens's exponents. The exact value depends on the procedure used to calculate it (Poulton, 1983).

The reciprocal relation is produced by what can be called a *stimulus range equalizing bias* (Poulton, 1979, p. 792). The average observer uses as responses much the same geometric range of numbers, whatever the size of the geometric stimulus range.

The points connected by straight lines in Figure 2 show R. Teghtsoonian's (1973) inverse, but not fully recipro-

Figure 2. Stimulus range equalizing biases across and within stimulus dimensions. The curved functions represent constant geometric response ranges of the sizes printed against the functions. The solid curved function illustrates the relation across stimulus dimensions between many of Stevens's (1962) exponents and the corresponding geometric stimulus ranges. The points connected by straight lines illustrate the relation reported by other investigators within stimulus dimensions between exponents and the corresponding geometric stimulus ranges (Engen & Levy, 1958; Frederiksen, 1975; Montgomery, 1975; R. Teghtsoonian, 1973). The figure is an elaboration of R. Teghtsoonian's (1973) Figure 4.

cal, relation within a single stimulus dimension that has been reported by investigators other than Stevens. Each point represents a separate group of observers judging stimuli in a single geometric stimulus range. As the Teghtsoonians (R. Teghtsoonian, 1973, Figure 4; Teghtsoonian & Teghtsoonian, 1986) point out, the slopes of the straight lines are less steep than are the slopes of the nearest parts of the functions that represent the fully reciprocal relation across stimulus dimensions.

The Teghtsoonians' (R. Teghtsoonian, 1973; Teghtsoonian & Teghtsoonian, 1978) own median results, within the stimulus dimension of a 3-kHz tone, are represented by the unfilled squares and the dashed function. The dashed function differs from the remaining straight-line functions in two respects. First, the three left-hand unfilled squares all lie well to the left of or below the remaining points, and so indicate smaller geometric response ranges. Doubling the geometric stimulus ranges from 5 dB through 10 dB to 20 dB increases the median geometric response ranges only from 1.5 through 1.8 to 2.6. Following Equation 1, this gives exponents of .69, .51, and .41, respectively. But as the Teghtsoonians (1986, p. 73) point out, with geometric response ranges as small as this, it is not easy to tell whether the median observers respond linearly or with ratios. They may well respond linearly.

There is a second difference between the Teghtsoonians' results and the remaining straight-line functions. The dashed line connecting the two right-hand unfilled squares is almost horizontal. Here the geometric stimulus ranges are 20 dB and 40 dB and the corresponding geometric response ranges are 2.6 and 6.3. Following Equation 1, this gives exponents of .41 and .40, respectively, which show practically no stimulus range equalizing bias.

As I have suggested (Poulton, 1984, p. 341), the stimulus range equalizing bias would disappear almost completely if the median undergraduates were to follow the median observers of Figure 1 and use a linear scale of numbers for stimulus ranges between 5 dB and 20 dB, but a scale of numbers calibrated in ratios for the largest stimulus range of 40 dB. The logarithmic response bias would then be responsible for much of the increase in the median geometric response range from 2.6 to 6.3, which leaves the exponent almost unchanged.

Without the logarithmic response bias, it can be predicted by extrapolation from the three smaller geometric response ranges that the stimulus range equalizing bias would produce a median geometric response range of about 4.0. Following Equation 1, this would give an exponent of about .30, continuing the downward trend of the other three exponents, .69, .51, and .41. It is difficult to see what artifact other than a logarithmic response bias could be responsible for almost halting the downward trend in the exponents.

REFERENCES

- BIRNBAUM, M. H. (1980). Comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General*, 109, 304-319.

- BIRNBAUM, M. H., & ELMASIAN, R. (1977). Loudness "ratios" and "differences" involve the same psychophysical operation. *Perception & Psychophysics*, **22**, 383-391.
- ENGEL, T., & LEVY, N. (1958). The influence of context on constant-sum loudness-judgments. *American Journal of Psychology*, **71**, 731-736.
- FREDERIKSEN, J. R. (1975). Two models for psychophysical judgment: Scale invariance with changes in stimulus range. *Perception & Psychophysics*, **17**, 147-157.
- MONTGOMERY, H. (1975). Direct estimation: Effect of methodological factors on scale type. *Scandinavian Journal of Psychology*, **16**, 19-29.
- POULTON, E. C. (1979). Models for the biases in judging sensory magnitude. *Psychological Bulletin*, **86**, 777-803.
- POULTON, E. C. (1981). Schooling and the new psychophysics. *Behavioral & Brain Sciences*, **4**, 201-203.
- POULTON, E. C. (1982). Biases in quantitative judgments. *Applied Ergonomics*, **13**, 31-42.
- POULTON, E. C. (1983). *Bias in judgment*. Unpublished manuscript.
- POULTON, E. C. (1984). A linear relation between loudness and decibels. *Perception & Psychophysics*, **36**, 338-342.
- POULTON, E. C., & SIMMONDS, D. C. V. (1985). Subjective zeros, subjectively equal stimulus spacing, and contraction biases in very first judgments of lightness. *Perception & Psychophysics*, **37**, 420-428.
- STEVENS, S. S. (1962, November). *In pursuit of the sensory law* (2nd Klopsteg lecture). Evanston, IL: Northwestern University, Technological Institute.
- STEVENS, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: Wiley.
- TEGHTSOONIAN, R. (1971). On the exponents in Stevens' law and the constant in Ekman's law. *Psychological Review*, **78**, 71-80.
- TEGHTSOONIAN, R. (1973). Range effects in psychophysical scaling and a revision of Stevens' law. *American Journal of Psychology*, **86**, 3-27.
- TEGHTSOONIAN, R., & TEGHTSOONIAN, M. (1978). Range and regression effects in magnitude scaling. *Perception & Psychophysics*, **24**, 305-314.
- TEGHTSOONIAN, R., & TEGHTSOONIAN, M. (1986). Scaling loudness over short ranges: A reply to Poulton. *Perception & Psychophysics*, **39**, 73-75.
- TORGERSON, W. S. (1960). Quantitative judgment scales. In H. Gulliksen & S. Messick (Eds.), *Psychological scaling: Theory and applications*. New York: Wiley.
- TREISMAN, M. (1963). Laws of sensory magnitude. *Nature*, **198**, 914-915.
- TREISMAN, M. (1964). Sensory scaling and the psychophysical law. *Quarterly Journal of Experimental Psychology*, **16**, 11-22.
- WEISS, D. J. (1981). The impossible dream of Fechner and Stevens. *Perception*, **10**, 431-434.

(Manuscript received February 21, 1986;
revision accepted for publication June 1, 1986.)